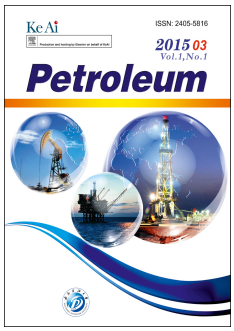
Accepted Manuscript



Development of machine learning methods for estimating permeability and porosity of oil reservoirs via petro-physical logs

Mohammad Ali Ahmadi, Zhangxin Chen

PII: S2405-6561(17)30163-3

DOI: 10.1016/j.petlm.2018.06.002

Reference: PETLM 218

To appear in: *Petroleum*

Received Date: 12 August 2017

Revised Date: 2 February 2018

Accepted Date: 8 June 2018

Please cite this article as: M.A. Ahmadi, Z. Chen, Development of machine learning methods for estimating permeability and porosity of oil reservoirs via petro-physical logs, *Petroleum* (2018), doi: 10.1016/j.petlm.2018.06.002.

Development of Machine Learning Methods for Estimating Permeability and Porosity of Oil Reservoirs via Petro-physical Logs

Mohammad Ali Ahmadi[1*], Zhangxin Chen[1]

[1)] Department of Chemical and Petroleum Engineering, University of Calgary, Calgary, Alberta, Canada

[*)] Address to Corresponding Author: Department of Chemical and Petroleum Engineering, University of Calgary, Calgary, Alberta, Canada. Email: mohammadali.ahmadi@ucalgary.ca, TEL: 001-587-2889700

Abstract

This paper deals with a new model for predicting porosity and permeability of oil reservoirs by coupling a machine learning concept and petrophysical logs. A rigorous machine learning technique called Least Squares Support Vector Machine (LSSVM) was employed. To improve the ability of the LSSVM model a Genetic Algorithm (GA) was employed. The machine learning approach was constructed and tested via data samples recorded from northern Persian Gulf oil reservoirs. Furthermore, other intelligent methods including fuzzy logic, artificial neural network, and hybridized methods were carried out and compared to the proposed GA-LSSVM model. The results gained from the machine learning model proposed here were compared to the relevant real petrophysical data and the outputs achieved from other methods employed in this study. The average relative absolute deviation between the approach estimations and the relevant actual data is found to be less than 1% for the least squares SVM approach. The results reported in this

paper indicate that implication of least squares SVM in porosity and permeability estimations can lead to construction of more reliable static reservoir models in simulation plans.

Keywords: Machine Learning; Neural Network; Support Vector Machine; Porosity, Permeability, Well Logs, Petro-physic

1. Introduction

Transport and storage characterizations including porosity and permeability are crucial in any reservoir evaluation plans. They regulate the directions of the reservoir fluids and flow through porous media [1-10]. Accurate estimation of these properties, particularly permeability, is vital for improvement of oil/gas recovery, CO_2 sequestration, selection of cost-effective production schemes, optimization of well placement, developing geothermal energy schemes, a secure design of noxious, management of water supplies, and radioactive waste storerooms [9-12]. It is worth to mention that a well test analysis as one of the main tools to characterize formation permeability is more accurate than well logs; however, in fractured reservoirs or layered reservoirs, an analysis of well testing is more complicated than in other types of reservoirs. A suitable magnitude for permeability can be gained experimentally from well logs or core samples. Though the coring process is very costly and time intensive, it can influence the outcome in not having enough cores to evaluate [8,10-11]. This may cause a high uncertainty accompanied with extrapolated or predicted geological data. Thus, a cost-effective, quick, and proficient model is essential in reservoir evaluation and characterization to specify permeability from well logs and the data from the existing cores to certify the outcomes [8-12].

Various formulas have been suggested in the literature for permeability and porosity prediction. However, most of these correlations are only appropriate for permeability estimation in sandstones, unconsolidated sands, and comparatively homogeneous porous media [8-12]. It is

very challenging to develop a widespread porosity/permeability correlation for very heterogeneous reservoirs including triple or dual porosity fractured carbonate rocks. Most of the proposed formulas include static rock character data while permeability is a dynamic variable and its estimation of higher confidence needs an instant production history [9,11]. However, having dynamic data can require more time and also higher costs [8,10-12].

An accurate prediction of the reservoir permeability and porosity is vital in any E & P systems in petroleum engineering. Porosity can be specified quite accurately from well logs; however, permeability should be gained through core analysis. Coring is a time and cost-intensive method contrasted with the electronic survey methods. Theoretically, specifying permeability through employing electrical measurement methods comprises solving equations in rock physics. Practically, several problems make a straight practical correlation impossible or complicated to determine. This turns into more complexes when one presumes description of typically heterogeneous reservoirs including naturally fractured carbonate rocks with a complicated triple or dual porosity system. It is not straightforward even in the lab to develop a correlation between a physical variable and a log response. Various upsetting parameters including reservoir fluids, mineralogy, damages of drilling to core plugs, and drilling fluid invasion can influence the measurement of permeability unfavorably [8-12].

Permeability calculation via employing well logs is a conventional job where adequate cores are not accessible. Throughout this circumstance, there are two separate approaches. The first approach is evolved according to the hypothetical theories related to fluid flow in capillary tubes [9-11]. This method comprises math based methodologies which imitate porous media via a bundle of capillary tubes. The accuracy of such theoretical approaches is low when porous media undertake geological events such as compaction, cementation, and dissolution [9-13]. The

second method is by empirical formulas which connect well logs and permeability. For example, permeability was formulated via porosity predicted from well logs. If more parameters like a sorting and grain size, connate water saturation, and a specific surface area are added to these formulas, then the precision of these formulas will increase remarkably [9-12]. It is worth to mention that prediction of permeability from well logs is also dealing with the scale-up problem meanwhile well logs typically have a vertical resolution of 2 feet in comparison with 2 inches in core samples. Furthermore, both the core and log approaches typically confront several measurement deviations. There remains the main question in construction and design of sensors to specify heterogeneous formations including geologically difficult reservoirs and carbonates [8-10].

Various permeability scatterings for heterogeneous reservoirs such as fractured carbonate media are demonstrated, for instance, normal, power-law, and random [14-16]. Bogdanov et al. (2007) [14] carried out a study on petrophysical characters of fractured porous media with a fracture size which had a power-law distribution. They studied the permeability of geological formations numerically by gaining the answers from combined Darcy's formulas throughout fractures and in the neighboring porous media. Two reduced parameters were proposed by these researchers to facilitate a collective explanation over a broad range of the fracture characters such as fracture shape, aperture, size distribution, and density. Furthermore, they suggested two common approaches for intense and loose fracture networks [14].

Permeability scattering is frequently very high in carbonate fractured rocks, whereas the porosity in the real major circumstances is in the range of 0.1-0.3 [16-17]. According to this fact, it is challenging in practice to determine a precise K-ϕ relationship for complex and heterogeneous porous media such as naturally fractured carbonate reservoirs [2,6,14-16]. Various empirical,

semi-theoretical, and theoretical relationships regarding a grain size and porosity are demonstrated in the literature in order to estimate permeability of a reservoir rock. Among them, familiar formulas are illustrated in Table 1 [2,6].

Insert Table 1

The theoretical formulas for permeability estimation can result in a high degree of precision if a pore-throat size distribution and three-dimensional networks are modeled appropriately. However, accomplishing such network modeling rises calculation costs and efforts enormously. Empirical approaches according to the statistical relationships between effective variables and permeability are also simple to use and quick. However, values gained from such approaches are useful just for wells from which the data is collected.

This paper deals with a new model for predicting porosity and permeability of oil reservoirs by coupling a machine learning concept and petrophysical logs. The machine learning approach was constructed and tested via data samples recorded accurately from northern Persian Gulf oil reservoirs. Furthermore, other intelligent methods including fuzzy logic, artificial neural network, and hybridized methods were carried out and compared to the proposed GA-LSSVM model.

2. Data Preparation

A total of 1000 real data points were employed to construct and test the accuracy and reliability of the machine learning model proposed in this paper. In this regard, data points were gathered from northern Persian Gulf oil reservoirs. The data samples employed in this paper include real values for Sonic Transit Time (DT), Density Tool Reading (NPHI), Bulk Density (RHOB), PHIT (total porosity), and the matched permeability and porosity in a core laboratory. It should be noted that real permeability/porosity in this paper means the value of permeability/porosity

measured in a core laboratory. This value of permeability/porosity was reported in routine core analysis (RCAL) reports. It is worth to highlight the statistical indexes of the data points employed in this paper. In this regard, Table 2 demonstrates the statistical indexes of the input and outputs parameters of the machine learning model.

Insert Table 2

3. Methodology

3.1. LS-SVM Method

Suykens and Vandewalle in 1999 [18] proposed the fundamental LSSVM based on the first type of SVM. One of the central schemes of the LSSVM is defeating the over-fitting issue which can happen in original SVM and ANNs. Before firing on details of the LSSVM approach, a graphical illustration of the LSSVM regression algorithm is illustrated in Figure 1. Suppose given inputs *X*i (PHIT, Sonic Transit Time (DT), Bulk Density (RHOB), Density Tool Reading (NPHI)) and output *Y*i (Permeability or Porosity) time series. Consider the training assortment $\{x_k, y_k\}$, $k = 1, 2, \ldots, N$, in which $x_k \in \mathbb{R}^n$ represents the kth input data in the input space and $y_k \in \mathbb{R}$ denotes the output value for a given value of the input variable (i.e., x_k) and N stands for the number of the training points. Utilizing the nonlinear function $\varphi(\cdot)$ that relates the training series in the input space to the multiple dimensional space, the following regression relationship is obtained [18-35]:

$$y = w^T.\varphi(x) + b \quad with \quad w \in \mathbb{R}^n, \quad b \in \mathbb{R}, \quad \varphi(\cdot) \in \mathbb{R}^n \to \mathbb{R}^{n_h}, \; n_h \to \infty \tag{1}$$

where $w$denotes the weight vector and b stands for the bias term. It is worth to note that the superscript "n" stands for the dimension of the data space, and n_h denotes the dimension of the passive typical space [18-21]. Optimization of the LSSVM method is as follows [31-35]:

$$\underset{w,b,e}{min} \quad \mathcal{J}(w, e) = \frac{1}{2} w^T w + \frac{1}{2} \gamma \sum_{k=1}^{N} e_k^2 \tag{2}$$

In order to minimize the equation mentioned above, the following equality limitation should be considered [18,20-21]:

$$y_k = w^T \phi(x_k) + b + e_k \qquad k = 1, 2, \ldots, N \tag{3}$$

where γ represents the factor of regularization, including the effects of the training error and the model's complication [18,20], and e_k stands for the regression error. In order to determine the location of the solution of the restricted optimization problem as shown below, the Lagrangian was employed [18-21]:

$$\mathcal{L}(w, b, e, \alpha) = \mathcal{J}(w, e) - \sum_{k=1}^{N} \alpha_k \{w^T \phi(x_k) + b + e_k - y_k\} \tag{4}$$

where α_k stands for the Lagrange multipliers or support values. The above equation was differentiated regarding $w, b, e_k,$ and α_k in order to determine the answer which results in the next equations [18-21]:

$$\frac{\partial \mathcal{L}(w,b,e,\alpha)}{\partial w} = 0 \rightarrow w = \sum_{k=1}^{N} \alpha_k \varphi(x_k) \tag{5}$$

$$\frac{\partial \mathcal{L}(w,b,e,\alpha)}{\partial b} = 0 \rightarrow \sum_{k=1}^{N} \alpha_k = 0 \tag{6}$$

$$\frac{\partial \mathcal{L}(w,b,e,\alpha)}{\partial e_k} = 0 \rightarrow \alpha_k = \gamma e_k, \quad k = 1, \ldots, N \tag{7}$$

$$\frac{\partial \mathcal{L}(w,b,e,\alpha)}{\partial \alpha_k} = 0 \rightarrow y_k = \varphi(x_k).w^T + b + e_k, \quad k = 1, \ldots, N \tag{8}$$

Excluding the variables w and e through the re-ordering of the aforementioned formulas, the Karush-Kuhn-Trucker system was achieved as follows [18-21,31-35]:

$$\begin{bmatrix} 0 & 1_v^T \\ 1_v & \Omega + \gamma^{-1} I \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix} \tag{9}$$

where $y = [y_1 \dots y_N]^T$, $1_N = [1 \dots 1]^T$, and $\alpha = [\alpha_1 \dots \alpha_N]^T$. In Equation (9), I introduces an identity matrix, $\Omega_{kl} = \varphi(x_k)^T . \varphi(x_l) = K(x_k, x_l), \forall\, k, l = 1, \dots, N,$ and $K(x_k, x_l)$ represents the kernel function which is required to encounter Mercer's situation [18-21]. Three typical forms for the kernel function are provided as follows:

- $K(x, x_k) = x_k^T x\ (Linear\ kernel)$
- $K(x, x_k) = (\tau + x_k^T x)^d\ (Polynomial\ kernel\ of\ degree\ d)$
- $K(x, x_k) = exp(-\|x - x_k\|^2/\sigma^2)\ (Radial\ basis\ function\ (RBF)\ kernel)$

The final expression of the LSSVM method for function estimation can be formulated as follows:

$$y(x) = \sum_{k=1}^{N} \alpha_k K(x, x_k) + b \tag{10}$$

where (b, α) stands for the answer of the linear system expressed through Equation (9).

Throughout regression faults, the RBF kernel was normally employed while it can be formulated below [18-21,34-35]:

$$K(x, x_k) = \exp(-\|x_k - x\|^2 / \sigma^2) \tag{11}$$

where σ^2 stands for the squared bandwidth that can be optimized by a robust and effective optimization approach like GA or PSO (Particle Swarm Optimization) during the calculation process [18-21,33-35].

To indicate the mean squared error (MSE) of the gained outcomes of the addressed vector machine approach the formula below was executed [18-21,31-35]:

$$MSE = \frac{\sum_{i=1}^{n} (P / K_{est_i} - P / K_{act_i})}{ns} \tag{12}$$

where P/K, subscripts *est* and *act* stand for the estimated and real porosity/permeability values, correspondingly, and *ns* denotes the number of data points from the initial population. In this paper, the suggested LS-SVM method by Suykens and Vandewalle was executed [18,21,32-35].

3.2.Fuzzy Decision Tree (FDT)

A FDT is taken as one of the advanced, well known and distinguished methods of automatic rule producers. In other words, carrying the FDT out can result in the automatic generation of fuzzy rules [36-40]. Its advantages are briefly summarized in quantifying cognitive uncertainties and lessening classification obscurity when fuzzy records are accessible. In short, a membership function (MF) $\mu(x)$ of a fuzzy variable Y can be assumed as a possibility distribution of Y on X, which is $\pi(x) = \mu(x)$ for all $x \in X$. The possibility measure $E_{\propto}(Y)$ of obscurity is formulated as follows [36-41]:

$$E_{\alpha}(Y) = g(\pi) = \sum_{i=1}^{n}(\pi_i^* - \pi_{i+1}^*)\ \ln i\ , \tag{13}$$

where $\pi^* = \{\pi_1^*, \pi_2^*, \dots, \pi_n^*\}$ stands for the permutation of the possibility dispensation $\pi = \{\pi(x_1), \pi(x_2), \dots, \pi(x_n)\}$ ranked so that $\pi_i^* \geq \pi_{i+1}^*$ for $i = 1, \dots, n$, and $\pi_{n+1}^* = 0$. The vagueness of characteristic A is then computed as [36-41]:

$$E_{\alpha}(A) = \frac{1}{m}\sum_{i=1}^{m} E_{\alpha}\big(A(u_i)\big) \tag{14}$$

in which $E_{\alpha}\big(A(u_i)\big) = g\left(\mu T_s(u_i)/max_{1\leq j\leq s}\left(\mu T_j(\mu_i)\right)\right)$, with T_j the linguistic scale utilized within a characteristic for m points. Overlapping exists when linguistic terms cover each other. The level in which A denotes a subset of B stands for recorded through fuzzy subset hood $S(A, B)$ and is determined by [36-41]:

$$S(A, B) = \frac{\sum_{u\epsilon U} min(\mu_A(u), \mu_B(u))}{\sum_{u\epsilon U} \mu_A(u)}, \tag{15}$$

where all characteristics are the same assortment of objects. The possibility of classifying an object to class C_i on the basis of the determined fuzzy record E can be formulated as follows [36-41]:

$$\pi = (C_i|E) = \frac{S(E,C_i)}{max_j\, S(E\,,C_j)}, \tag{16}$$

in which $S(E, C_i)$ determines the level of truth for the classification rule (that is $E \Rightarrow C_i$). The classification obscurity owing to a single piece of record (a fuzzy value for a characteristic) can be determined as follows [36-41]:

$$G(E) = g\big(\pi(C|E)\big). \tag{17}$$

$G(P|F)$ represents the classification obscurity with fuzzy partitioning $P = \{E_1, \dots, E_k\}$ on the fuzzy record F which is the weighted average of classification obscurity with each subset of partition [36-41]:

$$G(P|F) = \sum_{i=1}^{k} w(E_i|F)G(E_i \cap F), \tag{18}$$

in which $G(E_i \cap F)$ represents the classification obscurity with fuzzy record$E_i \cap F$, where $w(E_i|F)$ denotes the weight $\big(w(\cdot)\big)$ that applies the relative size of assortment $E_i \cap F$ in F [36-41]:

$$w(E_i|F) = \frac{\sum_{u\in U} min\Big(\mu_{E_i}(u),\mu_F(u)\Big)}{\sum_{j=1}^{k}\Big(\sum_{u\in U} min\Big(\mu_{E_j}(u),\mu_F(u)\Big)\Big)}. \tag{19}$$

Finally, characteristics are assigned to nodes based on the lowermost degree of ambiguity. It is stressed here that if the degree of subset hood is higher than a truth value β then a node became a leaf node. In other words, a truth value determines each node terminate or not [36-41].

3.3.Artificial Neural Network (ANN)

Briefly, ANNs are designed based on simulation of the human brain with the purpose of determining the relationship between outputs and inputs of a system. An ANN is trained with the available experimental data throughout the training step and is employed for estimating the unknown data [42-43]. Neural networks include simple synchronous processing components that are known as nodes or neurons located throughout layers. Usually, an artificial neural network (ANN) has three layers: an output layer, a hidden layer, and an input layer. The most common type of artificial neural networks (ANN) in petroleum engineering application is multi-layer perceptron (MLP) which is trained with aim of a back-propagation (BP) approach [44-49]. Multi-layer perceptron (MLP) networks score over conventional methods with respect to their much declined improvement time and their capability to make implementation of associated information [44-45,47-53].

In general, the measured permeability/porosity data was split into two assortments namely testing and training sets. The training process of a network is summarized to bias and weight optimization which selects each node interconnection until the target values at the output layer are as close as possible to the corresponding experimental data. The MSE of the network is expressed as in Eq. 12 [44-49].

3.4.Evolutionary Algorithms

3.4.1. Genetic Algorithm (GA)

Genetic Algorithms (GA), to gain a robust search engine and optimization approach, implement the origin of natural evolution with the genetic proliferation of characteristics which is called the principle of “survival of the fittest”. The important characteristic of a GA is that it specifies simultaneously many feasible answers and explores various areas in the desired space chosen by

the researcher [44,46,54-55]. GA employs a straightforward answer to genetics in biological systems, and the Darwinian natural selection is a robust further to the classical methods. According to the Darwinian theory of 'survival of the fittest', a GA can gain the best answer after a series of loop calculations. Artificial selection, crossover, and mutation operators constitute the search procedure.

In the first stage of a GA, an initial population, demonstrating deputies of the possible route, is generated to commence the search procedure. The components of the population are coded into chromosomes that are bit-strings. With the aim of another index that is named a fitness value, the robustness of the strings is assessed that illustrates the restrictions of the issue. Chromosomes are chosen for further genetic manipulation by their fitness value. It is worth mentioning that the process of selection is chiefly dependable on guaranteeing the survival of the best-fit chromosomes. Further choosing of the chromosomes strings in the genetic treatment procedure including two stages is implemented. Throughout the prior stage, the crossover operation which recombines the genes of each two chosen chromosomes is utilized. The next stage in the genetic utilization procedure is called a mutation process, where the bits of the chromosomes are changed at one or more randomly chosen positions. The operation of mutation aids to defeat the issue of trapping at a local optimum. The off-springs generated by the genetic utilization procedure are the subsequent population to be assessed [44,46,54-55]. The graphical demonstration of the GA-based training algorithm for the ANN is illustrated in Figure 2.

Insert Figure 2

3.4.2. Particle Swarm Optimization (PSO)

The PSO approach is evolved by the behavior of social organisms such as a flock of birds. Like other population-based methods, for instance, GA, the PSO algorithm is initiated with a population of chance solutions, named particles. These particles travel over the search area with a flexible velocity and keep the best position it has discovered in the search space. Each particle can revise its velocity vector to explore the best position owing to its flying expertise and the flying expertise of the other particles in the search space [48-49]. The graphical demonstration of the PSO-based training algorithm for the ANN is demonstrated in Figure 3.

Insert Figure 3

3.4.3. Imperialist Competitive Algorithm (ICA)

An ICA is a novel evolutionary approach in the evolutionary computation theme established based on the social and political evolution in societies [56]. Similar to other evolutionary approaches the ICA begins with primary individuals named countries. There are two types of countries: imperialist and colony that are composed to create an empire. Throughout the imperialistic competition procedure, imperialists attempt to gain additional colonies. Therefore, throughout the competition, the weak imperialists will lose power and powerful ones will achieve additional power. If an empire wastes all of its colonies, then it is presumed to be ended. In the end, the supreme potent imperialist will endure inside the world, and all the countries are colonies of this sole empire. In this stage, colonies and imperialist have identical power and position [45,47,56].

3.4.3.1.Generating initial empire

The objective of optimization is to obtain an optimum answer regarding parameters of the issue. They create an array of variable magnitudes to be optimized. In GA terms, this array is named “chromosome”, but here the terminology “country” is employed for this array, alternatively. In a

N_{var} –dimensional optimization issue, a country is a $1 \times N_{var}$ array. This array is given as follows [45,47,56]:

$$Country = [P_1,P_2,,P_3,\ldots,P_{Nvar}] \quad (20)$$

By assessing the cost function *f*, the cost of a country is specified [45,47,56]*:*

$$Cost=f\,(country) = f\,([P_1,P_2,,P_3,\ldots,P_{Nvar}]) \quad (21)$$

The approach commences with the number of imperialists (N_{imp}), the number of initial countries ($N_{country}$), and the number of the rest countries as colonies each of which belongs to an empire (N_{col}). The primary number of colonies of an empire is in closeness with their power. To split the colonies between imperialists proportionately, the normalized cost of an imperialist is given as follows [45,47,56]:

$$C_n=c_n - max_i\{c_i\} \quad (22)$$

in which c_n stands for the cost of the *n*-th imperialist and C_n denotes its normalized cost. Via the normalized costs of all imperialists, the power of every imperialist can be calculated as follows [45,47,56]:

$$P_n = \left|\frac{C_n}{\sum_{i=1}^{N_{imp}} c_i}\right| \quad (23)$$

Alternatively, the normalized power of an imperialist is specified via its colonies. Subsequently, the primary number of an imperialist can be given below [45,47,56]:

$$N.C_n = round\,\{P_n.N_{col}\} \quad (24)$$

where $N.C_n$ represents the initial number of colonies of the *n*-th empire and N_{col} denotes the number of all colonies. To split the colonies between imperialists, $N.C_n$ of the colonies are

chosen arbitrarily and allocated to every imperialist. The colonies associated with the imperialist create the *n*-th empire [45,47,56].

3.4.3.2. Moving colonies of an empire toward the imperialist

The imperialist countries attempt to expand their colonies and make them a part of themselves. This circumstance is demonstrated via repositioning all colonies in the direction of their corresponding imperialist. A colony shifts toward an imperialist by x units (x is a random parameter with uniform scattering) [45,47,56]:

$$x \sim U\ (0, \beta \times d) \tag{25}$$

where β stands for a number greater than 1 and d denotes the space between an imperialist and a colony. Throughout the shifting process, a colony can attain a position with a lower cost than that of its imperialist. Throughout this circumstance, the colony and the imperialist exchange their positions. So the approach will endure by the imperialist throughout the fresh position, and then colonies commence shifting in the direction of this position [45,47,56].

3.4.3.3. The total power of an empire

The total power of an empire depends on both the power of imperialist's colonies and the power of the imperialist country. This circumstance is demonstrated via identifying the total cost as given below [45,47,56]:

$$To\ .\ C_n = Cost\ (Imperialist_n\ ) + \xi\ mean\ \{Cost\ (Colonies\ of\ impire_n)\ \} \tag{26}$$

where $To\ .\ C_n$ stands for the total cost of the *n*-th empire and ζ denotes a positive number that is presumed to be lower than 1. An infinitesimal value for ζ indicates that the total power of an empire to be specified via just the imperialist and rising it will raise the responsibility of the

colonies in specifying the total power of an empire. The magnitude of 0.1 for ζ is an appropriate magnitude for most of the cases [45,47,56].

3.4.3.4. Imperialist competition

Each empire attempts to take the ownership of colonies of other empires and manage them. The imperialistic contest progressively carries a reduction in the power of weaker empires and an improvement in the power of more prevailing ones. This contest is demonstrated via just selecting some, normally one, of the fragile colonies of the fragile empires and creating a challenge between all empires to have these colonies. To commence a challenge, first, the possession probability of each empire should be specified according to their total power. The normalized total cost is given as follows [45,47,56]:

$$N.To.C_n = To.C_n - max_i\{To.C_i\} \quad (27)$$

where $N.To.C_n$ and $To.C_n$ stand for the normalized total cost and the total cost of the *n*-th empire, correspondingly. Via the normalized total cost, the ownership possibility of every empire is defined as follows [45,47,56]:

$$P_{P_n} = \left| \frac{N.To.C_{.n}}{\sum_{i=1}^{N_{imp}} N.To.C_{.i}} \right| \quad (28)$$

To split the colonies between empires, vector *P* is created as follows [45,47,56]:

$$P = [P_{P_1}, P_{P_2}, P_{P_3}, \ldots, P_{P_{N_{imp}}}] \quad (29)$$

So the vector *R* with an identical size as *P* whose components are homogeneously scattered arbitrary numbers is generated as follows [45,47,56]:

$$R = [r_1, r_2, r_3, \ldots, r_{N_{imp}}] \quad (30)$$

Therefore, vector *D* is created via subtracting vector *R* from vector *P* [45,47,56]:

$$D = P - R = [D_1, D_2, D_3, \ldots, D_{N_{imp}}] \quad (31)$$

Addressing vector *D*, the colonies are passed to an empire whose corresponding index throughout *D* is maximized. A weak empire will fall throughout the imperialistic challenge, and their colonies will be split between other empires. Finally, all the empires excluding the greatest prevailing one will fall, and all the colonies will be under the management of this single empire. Throughout this step, colonies and imperialists have the identical power and position. The utilization process of the suggested optimization strategy by ICA is depicted in Figure 4 [45,47,56].

Insert Figure 4

In this section, the ICA was utilized as an ANN optimization approach, and the *MSE* was considered as a cost function of the ICA approach. The objective here is to minimize the cost function through the proposed algorithm.

3.4.4. Hybrid Particle Swarm Optimization and Genetic Algorithm (HGAPSO)

Albeit the fact that the implementation of GA has accomplished effectively within a broad variety of engineering and design issues [57-58], it is a time-consuming approach if carried out for large-scale optimizations which need numerous function assessments for convergence. Thus, to unravel this addressed issue (for instance, high computation costs and efforts), it is highly appreciated to merge particle swarm optimization and a genetic algorithm in a single approach with the aim of advantage from the suitable features and searching capabilities of both methods in predicting required factors. The hybrid of particle swarm optimization (PSO) and the genetic

algorithm (GA) named HGAPSO and first suggested by Juang (2004) [59], was employed to predict the porosity/permeability of reservoirs. Figure 3 demonstrates the procedure of the introduced hybrid GA and PSO approaches [44,46,59].

The floating-point representation of variables in a GA and a search operator which recognizes neighboring areas in the search region is capable of removing the two obstacles associated with single-point crossover and binary coding. Therefore, a floating point coding pattern is approved here for the entire genetic algorithm (GA), particle swarm optimization (PSO) and hybrid PSO and GA (HGAPSO) [44,46,59].

Insert Figure 5

4. Results and Discussion

4.1.Permeability Model

Distributions of the addressed output (permeability) versus independent variables including Sonic Transit Time (DT), Bulk Density (RHOB), Density Tool Reading (NPHI), and PHIT logs are demonstrated in Figures 6 to 9. Figure 6 depicts the variation of permeability versus the corresponding neutron log data samples. According to Figure 6, permeability varies from 0.001 to around 55 mD for Density Tool Reading (NPHI) from -0.1 to 0.35. Figures 7 to 9 illustrate the correlations between permeability and Sonic Transit Time (DT), Bulk Density (RHOB), Density Tool Reading (NPHI), and PHIT logs data samples, correspondingly.

Insert Figure 6

Insert Figure 7

Insert Figure 8

Insert Figure 9

As noted previously, the most important parameters of the LSSVM approach are summarized to γ and σ^2. To figure out optimum values of theses addressed crucial parameters, a high-performance optimization algorithm should be utilized. The objective function of the optimization algorithm was a mean squared error (MSE). Via Random division the data samples were divided into two main groups comprising the "Testing" set and the "Training" set. About 25 % and 75 % of the main data samples were chosen for the "Test" set (250 data samples) and the "Training" set (750 data samples), correspondingly. Numerous distribution cuts were put forth to prevent the local build-ups of the data in the workable area of the addressed issue. Consequently, the acceptable scattering is the one with homogeneous build-ups of the data on the branch of the two referred divided data banks [33-35]. The GA was employed to determine the optimum values of the LSSVM parameters accurately [35,60-61]. The optimization toolbox in MATLAB® software was employed to figure out the optimum values of the LSSVM parameters [22-24]. In the GA the number of individuals was set 1500, and enormous efforts have been made to avoid any pre-convergence [33-35].

The optimum values of the LSSVM parameters were γ= 0.685262 and σ^2 = 0.07962. The number of digits for the addressed LSSVM approach has effects on the efficiency and accuracy of the LSSVM approach. Sensitivity analysis was performed to figure out optimum digits of the variables [32-35].

A contrast between the monitored permeability values and the actual values is demonstrated in Figure 10. Figure 10 demonstrates the scatter diagram that compares real permeability against the results achieved from a machine learning model. A tight cloud of points about the diagonal line (Y=X) for testing and training data assortments present the high accuracy of the LSSVM model.

The addressed figure depicts that superior agreement exists between the outputs achieved from the machine learning approach and the permeability measured in the laboratory.

Figure 11 depicts the distribution for the deviation of the gained outputs by performing the LSSVM model versus NPHI values of permeability for the whole data sets that were implemented for evolving the LSSVM approach. Also, Figures 12 to 14 depict the deviation distribution against the corresponding PHIT, RHOB, and DT, correspondingly. These plots illustrate the level of agreement between the actual recorded data and the estimated values. Finally, the relative error distribution of the machine learning method against the corresponding measured permeability is depicted in Figure 15. As depicted in Figure 15, the maximum error of the LSSVM results was referred to the lower values of reservoir permeability. This demonstrates that the permeability prediction via the machine learning model proposed here has a reasonable harmony with the measured reservoir permeability.

Finally, to compare between the vector model proposed in this paper and previously applied intelligent models such as hybridized methods, a fuzzy decision tree (FDT) and artificial neural network, we applied all of the methods on the used data banks in this paper. Figure 16 represents the scatter plot of the ANN approach against the relevant actual permeability of the reservoir. As demonstrated in Figure 16, the correlation coefficient between the estimated and actual data points is 0.4335 which reveals that the ANN model was unable to estimate reservoir permeability accurately. Figure 17 demonstrates the outputs gained from a fuzzy decision tree (FDT) versus the relevant real values of reservoir permeability. As depicted in Figure 17, the same as for the ANN model, the FDT model is unable to estimate reservoir permeability correctly.

Figure 18 illustrates the outputs gained from the ICA-ANN model versus relevant real reservoir permeability. As illustrated in Figure 18, the ICA-ANN model has a better estimation of reservoir permeability with a higher degree of accuracy that the ANN and FDT models.

The minimalizing procedure of the MSE magnitude is the optimizing and adjusting process of thresholds and weights of the neural network. Thus, the HGAPSO is employed to optimize the thresholds and weights of the neural network model. The optimization of ANN weights is employed in this paper. HGAPSO is utilized here as the ANN optimization approach, and the MSE is considered as a cost function throughout the HGAPSO method. The main objective of the proposed algorithm is to minimalize the MSE. Figure 19 depicts reservoir permeabilities gained from the HGAPSO-ANN model versus field reservoir permeability. As depicted in Figure 19, HGAPSO-ANN has a good agreement with the field reservoir permeability, but the performance of the model is higher than for other intelligent approaches but lower than the vector model that was evolved previously in this paper. Statistical parameters of the outcomes gained from the approaches mentioned above are reported in Table 3.

Insert Figure 10

Insert Figure 11

Insert Figure 12

Insert Figure 13

Insert Figure 14

Insert Figure 15

Insert Figure 16

Insert Figure 17

Insert Figure 18

Insert Figure 19

Insert Table 3

4.2.Porosity Model

Distributions of the porosity versus independent variables including Sonic Transit Time (DT), Bulk Density (RHOB), PHIT, and Density Tool Reading (NPHI) logs are demonstrated in Figures 20 to 23. Figure 20 depicts the variation of porosity versus corresponding neutron log data samples. Owing to Figure 20, porosity varies from 0.033% to about 23.6% for Density Tool Reading (NPHI) from -0.1 to 0.35. Figures 21 to 23 illustrate the correlations between porosity and Sonic Transit Time (DT), Bulk Density (RHOB), PHIT, and Density Tool Reading (NPHI) logs data samples, correspondingly. The same procedure as for the permeability prediction was followed throughout this section. The optimum values of the LSSVM parameters were γ= 0.885313 and σ^2= 0.007753. The number of digits for the addressed LSSVM approach has effects on the efficiency and accuracy of the LSSVM approach. Sensitivity analysis was performed to figure out optimum digits of the variables.

The same as in the previous section, a contrast between the porosity values estimated by machine learning and the porosity values measured in the lab are depicted in Figure 24. As depicted in Figure 24, the porosity values estimated by the machine learning model followed the diagonal line (Y=X) for both the training and testing data sets depicting the high accuracy of the machine learning model. The addressed figure depicts that superior agreement exists between the outputs achieved from the machine learning approach and the porosity measured in the laboratory.

The relative error distribution of the machine learning model against the corresponding porosity measured in the laboratory is demonstrated in Figure 25. As depicted in Figure 25, the maximum error of the LSSVM results was referred to the lower values of reservoir porosity. This demonstrates that the porosity prediction via the machine learning model proposed here has a reasonable harmony with the measured reservoir porosity. Figure 26 demonstrates the

distribution of the deviation of the gained outputs by applying the LSSVM model versus DT values of porosity for all of the data sets that were implemented for developing the LSSVM approach. Also, Figures 27 to 29 depict the deviation distribution against the corresponding PHIT, RHOB, and NPHI, correspondingly. These plots demonstrate the level of agreement between the actual recorded data and the estimated values.

Finally, to compare between the vector model proposed in this paper and previously applied intelligent models such as hybridized methods, a fuzzy decision tree (FDT) and artificial neural network, we applied all of the methods to the used data banks in this paper. Figure 30 represents the scatter plot of the ANN approach versus the relevant actual porosity of the reservoir. As demonstrated in Figure 30, the correlation coefficient between the estimated and actual data points is 0.6287 which reveals that the ANN model is unable to estimate reservoir porosity accurately. Figure 31 demonstrates the outputs gained from the fuzzy decision tree (FDT) versus the relevant real values of reservoir porosity. As depicted in Figure 31, the same as for ANN model, the FDT model is unable to estimate reservoir porosity correctly. Figure 32 illustrates the outputs gained from the ICA-ANN model versus relevant real porosity. As illustrated in Figure 32, the ICA-ANN model has a better estimation of reservoir porosity with a higher degree of accuracy that the ANN and FDT models. Figure 33 depicts reservoir porosities gained from the HGAPSO-ANN model versus field reservoir porosity. As depicted in Figure 33, HGAPSO-ANN has a good agreement with the field reservoir porosity, but the performance of the HGAPSO model is higher than other intelligent approaches but lower than the vector model that was evolved previously in this paper. Statistical parameters of the outcomes gained from the approaches mentioned above are reported in Table 4.

Insert Figure 20

Insert Figure 21

Insert Figure 22

Insert Figure 23

Insert Figure 24

Insert Figure 25

Insert Figure 26

Insert Figure 27

Insert Figure 28

Insert Figure 29

Insert Figure 30

Insert Figure 31

Insert Figure 32

Insert Figure 33

Insert Table 4

5. Conclusions

Enormous attempts have been made to provide an accurate and robust approach for determining values of permeability/porosity of petroleum reservoirs. Least squares support vector machine (LSSVM), hybridized methods, a fuzzy decision tree (FDT) and ANN as robust and effective intelligent approaches were implemented to develop a predictive model. Also, precise actual porosity and permeability data from one of the northern Persian Gulf oil fields were employed to figure out effectiveness and accuracy of the proposed predictive tool. Based on the gained outcomes from this work the following key deductions can be extracted:

1. A least squares support vector machine (LSSVM) model has the capability of avoiding being trapped in local optima for reservoir permeability/porosity prediction since this new approach proposed here has both global and local searching competences.
2. When employing the LSSVM method for reservoir permeability/porosity prediction, finding the optimum values of the ɣ and δ^2 present a big challenge which requires more investigation.

References

[1]. Nelson, P., Permeability-porosity relationships in sedimentary rocks. The log analyst , (1994), 1, 38-62.

[2]. Ezekwe, N. (2010). Petroleum Reservoir Engineering Practice. Prentice Hall.

[3]. Edlmann, K., M.Somerville, J., Smart, B. G., Hamilton, S., B.R.Crawford. (1996). Predicting Rock Mechanical Properties from Wireline Porosities. SPE (pp. 169-175). Trondheim,Norway: SPE 47344.

[4]. Zhang, L. (2013). Aspects of rock permeability. Tucson, Arizona, USA, Springer.

[5]. Zhang, Y., Lollback, P., Rojahn, J., Salisch, H., & Stuart, W., A Methodology for Estimating Permeability from Well Logs in a Formation of Complex Lithology. SPE (pp. 561-566). Adelaide, Australia, 28-31 October 1996.

[6]. Johnson, A. I. (1963). Application of laboratory permeability data. Denver, Colorado: Water resources division.

[7]. Ahmed, T. (2006). Reservoir Engineering handbook. Second Edition, Oxford, UK, Elsevier .

[8]. Newman, G.H; Martin, J.C. Equipment and experimental methods for obtaining laboratory compression characteristics of reservoir rocks under various stress and pressure conditions. SPE paper no. 6855, 52nd Annual Technical Conference of SPE, Denver, Colorado, October 9-12, 1977.

[9]. Dullien, F. A. L., (1992), Porous Media - Fluid Transport and Pore Structure. 2nd edition, Academic, San Diego, Calif., United States

[10]. Byrnes, A.P., Measurements of independent and dependent variables for petrophysical properties prediction. Reservoir Quality Assessment and Prediction in Clastic Rocks: M.D., Wilson edition, SEPM short course #30, 1994, p. 231-247 & p. 293-31

[11]. Wu, T., Ph.D. Dissertation, Department of Geology and Geophysics, Texas A&M University, College Station, Texas, United States, December 2004.

[12]. Thomas, D.C., Pugh, V.J., A statistical analysis of accuracy and reproducibility of standard core analysis. Proceeding of Society of Core Analysis, Paper no. 8701, 1987.

[13]. Rezaee, R., Saeedi, A., Clennell, B., Tight gas sands permeability estimation from mercury injection capillary pressure and nuclear magnetic resonance data. Journal of petroleum science and engineering , (2012), 92-99.

[14]. Bogdanov, I. I.; Mourzenko, V. V.; Thovert, J.F. Effective permeability of fractured porous media with power-law distribution of fracture sizes. Physics Review E , (2007), 76 (3), p.15.

[15]. Saidi, A. M., (1987), Reservoir Engineering of Fractured Reservoirs. Total Edition Press: Paris, France.

[16]. Van Golf-Racht, T.D., Fundamentals of Fractured Reservoir Engineering. Netherland, Elsevier, 1982.

[17]. Lamur, A., Kendrick, J., Eggertsson, G., Wall, R., Ashworth, J., Lavallée, Y., 2017. The permeability of fractured rocks in pressurised volcanic and geothermal systems. Scientific Reports 7, 6173.

[18]. Suykens, J. A. K.; Vandewalle, J. Least squares support vector machine classifiers. Neural Process. Lett. (1999), 9, 293-300.

[19]. Cortes, C., Vapnik, V. Support-Vector Networks. Machine Learning, (1995), 20, 179-188.

[20]. Suykens, J. A. K., Van Gestel, T., De Brabanter, J., De Moor, B., Vandewalle, J., (2002), Least Squares Support Vector Machines, World Scientific, Singapore, 2002.

[21]. Pelckmans, K.; Suykens, J. A. K.; Van Gestel, T.; De Brabanter, D.; Lukas, L.; Hamers, B.; De Moor, B.; Vandewalle, J. LS-SVMlab: a Matlab/C Toolbox for Least Squares Support Vector Machines; Internal Report 02-44, ESATSISTA; K. U. Leuven: Leuven, Belgium, 2002.

[22]. Ahmadi, M.A., Ebadi, M., Hosseini, S.M., Prediction Breakthrough Time of Water Coning in the Fractured Reservoirs by Implementing Low Parameter Support Vector Machine Approach, Fuel, (2014a), 117 , 579–589.

[23]. Ahmadi, M.A., Ebadi, M., Soleimani Marghmaleki, P., Mahboubi Fouladi, M., Evolving Predictive Model to Determine Condensate-to-Gas Ratio in Retrograded Condensate Gas Reservoirs, Fuel, (2014b), 124C, 241-257.

[24]. Ahmadi, M.A., Masoumi, M., Askarinezhad, R., Evolving Connectionist Model to Monitor Efficiency of the In-situ Combustion Process: Application to Heavy Oil Recovery, Journal of Energy Technology, (2014c), 2(9-10), 811–818

[25]. Ahmadi, M.A., Kashiwao, T., Bahadori, A., Prediction of Oil Production Rate Using Vapor-extraction Technique in Heavy Oil Recovery Operations, Petroleum Science and Technology, (2015a), 33:20, 1764-1769

[26]. Ahmadi, M.A., Galedarzadeh, M., Shadizadeh, S.R., Low Parameter Model to Monitor Bottom Hole Pressure in Vertical Multiphase Flow in Oil Production Wells, Petroleum, (2015b), In Press, doi:10.1016/j.petlm.2015.08.001

[27]. Ahmadi, M.A., Zahedzadeh, M., Shadizadeh, S.R., Abassi, R., Connectionist Model for Predicting Minimum Gas Miscibility Pressure: Application to Gas Injection Process, Fuel, (2015c), 148, 202–211

[28]. Ahmadi, M.A., Lee, M., Bahadori, A., Prediction of a solid desiccant dehydrator performance using least squares support vector machines algorithm, Journal of the Taiwan Institute of Chemical Engineers, (2015d), 50,115-122

[29]. Ahmadi, M.H., Ahmadi, M.A., Sadatsakkak, A., Connectionist Intelligent Model Estimates Output Power and Torque of Stirling Engine, Renewable and Sustainable Energy Reviews, (2015e), 50, 871–883

[30]. Ahmadi, M.A., Masoumi, M., Askarinezhad, R., Evolving Smart Model to Predict Combustion Front Velocity throughout In-Situ Combustion Process Employment, Energy Technol. (2015f), 3, 128 – 135

[31]. Ahmadi, M.A., Ebadi, M., Evolving Smart Approach for Determination Dew Point Pressure of Condensate Gas Reservoirs, Fuel, (2014), 117(B), 1074-1084

[32]. Ahmadi, M.A., Pournik, M., A Predictive Model of Chemical Flooding for Enhanced Oil Recovery Purposes: Application of Least Square Support Vector Machine , Petroleum, (2015), In Press, DOI:10.1016/j.petlm.2015.10.002

[33]. Ahmadi, M.A., Connectionist Approach Estimates Gas-Oil Relative Permeability in Petroleum Reservoirs: Application to Reservoir Simulation, Fuel, (2015a), 140C, 429-439

[34]. Ahmadi, M.A., Mahmoudi, B., Yazdanpanah, A., Development of robust model to estimate gas–oil interfacial tension using least square support vector machine: Experimental and modeling study, The Journal of Supercritical Fluids, (2016), 107,122–128

[35]. Ahmadi, M.A, Bahadori, A., A LSSVM Approach for Determining Well Placement and Conning Phenomena in Horizontal Wells, Fuel, (2015), 153, 276-283

[36]. Adamo, J. M., Fuzzy decision trees, Fuzzy Sets and Systems, (1980), 4(3), 207-219.

[37]. Yuan, Y. and M. J. Shaw., Induction of fuzzy decision trees, Fuzzy Sets and Systems, (1995), 69(2): 125-139.

[38]. Wang, X., B. Chen, et al., On the optimization of fuzzy decision trees, Fuzzy Sets and Systems, (2000), 112(1), 117-125.

[39]. Olaru, C. and L. Wehenkel, A complete fuzzy decision tree technique, Fuzzy Sets and Systems, (2003), 138(2): 221-254.

[40]. Beynon, M. J., M. J. Peel, et al., The application of fuzzy decision tree analysis in an exposition of the antecedents of audit fees, Omega, (2004), 32(3): 231-244.

[41]. Ebadi, M., Ahmadi, M.A., Farhadi Hikoei, K., Application of Fuzzy Decision Tree Analysis for Prediction Asphaltene Precipitation Due Natural Depletion; Case Study, Australian Journal of Basic and Applied Sciences, (2012), 6, 190-197.

[42]. Bain, A. Mind and Body: The Theories of Their Relation. New York: D. Appleton and Company, 1873.

[43]. James, W., (1980), The Principles of Psychology. New York: H. Holt and Company.

[44]. Ahmadi, M.A.; Golshadi, M., Neural Network Based Swarm Concept for Prediction Asphaltene Precipitation due Natural Depletion, Journal of Petroleum Science Engineering. (2012), 98–99, 40–49.

[45]. Ahmadi, MA, Ebadi, M, Shokrollahi, A, Majidi, SMJ: Evolving artificial neural network and imperialist competitive algorithm for prediction oil flow rate of the reservoir. Applied Soft Computing, (2013a), 13(2):1085-1098.

[46]. Ahmadi, M.A., Zendehboudi, S., Lohi, A., Elkamel, A., Chatzis, I., Reservoir permeability prediction by neural networks combined with hybrid genetic algorithm and particle swarm optimization ,Geophysical Prospecting, (2013b), 61, 582–598

[47]. Ahmadi, M.A., Prediction of asphaltene precipitation using artificial neural network optimized by imperialist competitive algorithm, Journal of Petroleum Exploration and Production Technology, (2011), 1, 99-106.

[48]. Ahmadi, M.A., Neural Network Based Unified Particle Swarm Optimization for Prediction of Asphaltene Precipitation, Fluid Phase Equilibria, (2012), 314, 46-51.

[49]. Ahmadi MA, Shadizadeh SR: New approach for prediction of asphaltene precipitation due to natural depletion by using evolutionary algorithm concept. Fuel, (2012), 102(0),716-723.

[50]. Ahmadi, M.A., Developing a Robust Surrogate Model of Chemical Flooding Based on the Artificial Neural Network for Enhanced Oil Recovery Implications, Mathematical Problems in Engineering, (2015b), Volume 2015, Article ID 706897, 1-9

[51]. Hornick, K., Stinchcombe, M., White, H., Multilayer feed forward networks are universal approximators. Neural Networks, (1989), 2, 359–366.

[52]. Hornick, K., Stinchcombe, M., White, H., Universal approximation of an unknown mapping and its derivatives using multilayer feed forward networks. Neural Networks, (1990), 3 (5), 551–600.

[53]. Garcia-Pedrajas, N.; Hervas-Martinez, C.; Munoz-Perez, J. COVNET: A cooperative co evolutionary model for evolving artificial neural networks. IEEE Transaction on Neural Networks, (2003), 14, 575–596.

[54]. Holland J. H., (1975), Adaptation in Natural and Artificial Systems, 2ed edn. MIT Press, Cambridge, MA, United States.

[55]. Hassan R., Cohanim B. and Weck O. 2005. A comparison of particle swarm optimization and genetic algorithm. Proceedings of 46th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics & Materials Conference, Austin, Texas, 18-21.

[56]. Atashpaz-Gargari, E., Lucas, C., Imperialist Competitive Algorithm: An Algorithm for Optimization Inspired by Imperialistic Competition. IEEE Congress on Evolutionary Computation, 2007, IEEE: New York, 2007; 4661−4667.

[57]. Qul, X., Feng, J., Sun, W., Parallel genetic algorithm model based on AHP and neural networks for enterprise comprehensive business. Intelligent Information Hiding and Multimedia Signal Processing. Harbin, 15–17 August 2008.

[58]. Balan, B., Mohaghegh, S.,Ameri, S., State of the art in permeability determination fromwelllogdata:Part1—a comparative study, model development. SPE Eastern Regional Conference and Exhibition, Morgantown, SPE 30978. West Virginia, 17–21September1995.

[59]. Juang, C.F., A hybrid of genetic algorithm and particle swarm optimization for recurrent network design. IEEE Trans. Syst. Man Cybern. B Cybern, (2004), 34 (2), 997–1006.

[60]. Schwefel, H.P., 1981. Numerical Optimization of Computer Models. John Wiley & Sons, New York.

[61]. Davis, L., (1991), Handbook of Genetic Algorithms. Van Nostrand Reinhold, New York.

[62]. Timur, A., An Investigation Of Permeability, Porosity, & Residual Water Saturation Relationships For Sandstone Reservoirs. The Log Analyst , (1968), IX(4) ,8-17.

[63]. Coates, G.R. and J.L. Dumanoir, 1981. A new approach to improved log derived permeability. The Log Analyst, pp: 17.

[64]. Tixier, M.P., 1949. Evaluation of permeability from electric log resistivity gradient. Earth Sci. J., 2: 113-113.

[65]. Aigbedion, I., A case study of permeability modeling and reservoir performance in the absence of core data in the Niger Delta, Nigeria. Journal of Applied Sciences , (2007), 772-776.

[66]. Kozeny, J. "Ueber kapillare Leitung des Wassers im Boden." Sitzungsber Akad. Wiss., Wien, 136(2a): 271-306, 1927

[67]. Pape, H., Clauser, C., Iffland, J., Permeability prediction based on fractal pore space geometry. Geophysics, (1999), 64, 1447-1460.

[68]. Jorgensen DG. Estimating Geohydrologic Properties from Borehole-Geophysical Logs. Groundwater Monitoring & Remediation. 1991 Aug 1;11(3):123-9.

[69]. Ahmadi, M. A., Ahmadi, M. R., Hosseini, S. M., Ebadi, M. Connectionist model predicts the porosity and permeability of petroleum reservoirs by means of petro-physical logs: Application of artificial intelligence. Journal of Petroleum Science and Engineering, (2014), 123, 183-200.

Tables

Table 1: Selected model for permeability prediction

Formula	Type	Reference
$k = 0.136\frac{\Phi^{4.4}}{S_{wr}{}^{2}}$	Empirical	Timur model (1968) [62]
$k^{0.5} = 100\frac{\Phi^{2}(1 - Sw_{ir})}{Sw_{ir}}$	Empirical	Coates and Dumanoir (1981) [63]
$k^{0.5} = 250\frac{\Phi^{3}}{Sw_{ir}}$	Empirical	Tixier (1949) [64]
$Log\ K = -0.83565 + 13.069 \times \emptyset$	Empirical	Aigbedion (2007) [65]
$K = 18044\Phi^{3.104}$	Empirical	Aigbedion (2007) [65]
$k = \frac{1013000\emptyset^{3}D_{p}^{2}}{180(1 - \emptyset)^{2}}$	Theoretical	Kozeny-Carman (1927-1956) [66]
Sandstone: $k = 0.031\emptyset + 7.463\ \emptyset^{2} + 0.191\ (10\emptyset)^{10}$ Shaly sandstone: $k = 0.0062\emptyset + 1.493\emptyset^{2} + 0.058\ (10\emptyset)^{10}$	Semi-Theoretical	Pape *et al.* (1999) [67]
$k = 84105\ \frac{\emptyset^{m+2}}{(1 - \emptyset)^{2}}$	Empirical	Jorgensen (1987) [68]

Table 2: Statistical indexes of the petrophysical data points employed in this paper

Parameter	DT	RHOB	PHIT	NPHI	Porosity	Permeability
Min	48.3975	2.2789	0.0002	-0.10513	0.033	0.0139
Max	81.7146	2.7474	22.3336	0.350546	23.604	48.397
Average	57.87668	2.569957	7.187092	0.075596	7.75259	2.860056

Table 3: Statistical parameters of the used models for estimation reservoir permeability

Parameter	FDT	ANN	ICA-ANN	HGAPSO-ANN	LSSVM
R^2	0.4248	0.4335	0.9617	0.9747	0.994
Mean Squared Error (MSE)	21.959	20.841	1.4152	0.9606	0.4176
AARD[a]%	745.308	600.281	40.515	36.3844	3.5093

a) AARD%=Average Absolute Relative Deviation

Table 4: Statistical parameters of the models used for estimation reservoir porosity

Parameter	FDT	ANN	ICA-ANN	HGAPSO-ANN	LSSVM
R^2	0.4034	0.6287	0.8216	0.8505	0.9716
Mean Squared Error (MSE)	29.132	18.275	8.494	5.334	1.084
AARD%	178.612	148.376	80.863	60.321	7.497

Figures

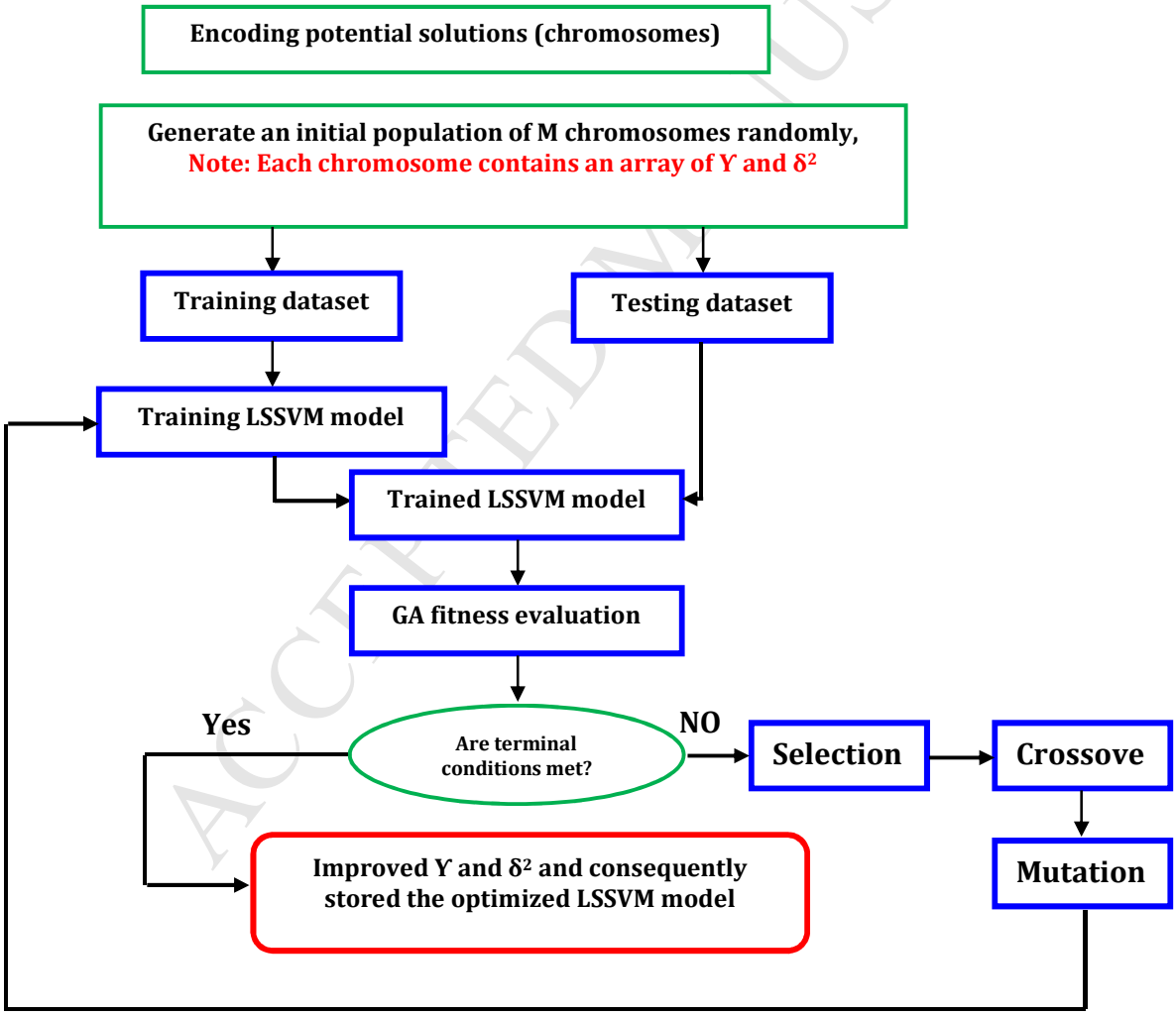


Figure 1: Flowchart of the LSSVM approach [25-35]

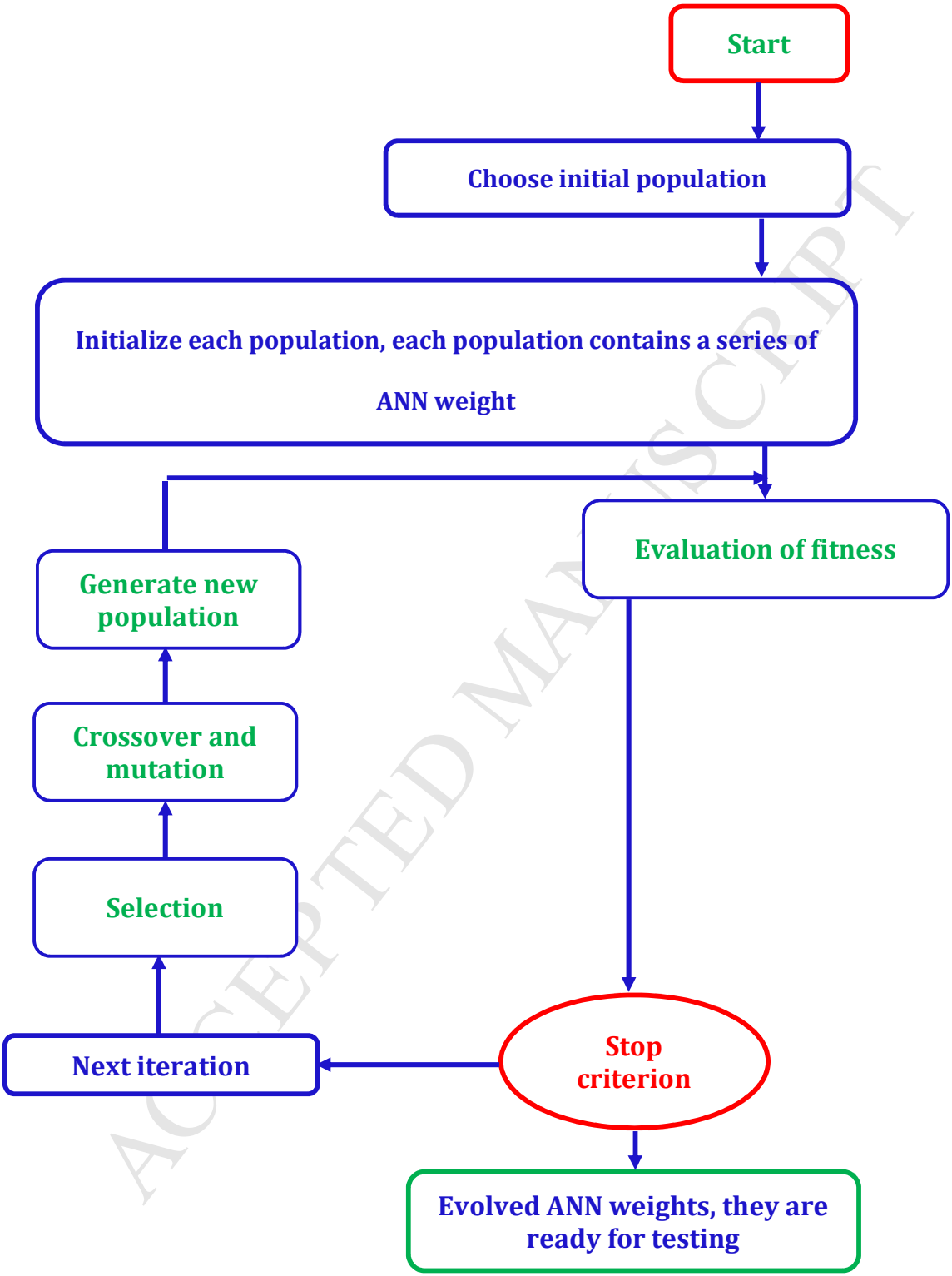


Figure 2: GA-based algorithm box chart in optimization of the weights of ANN

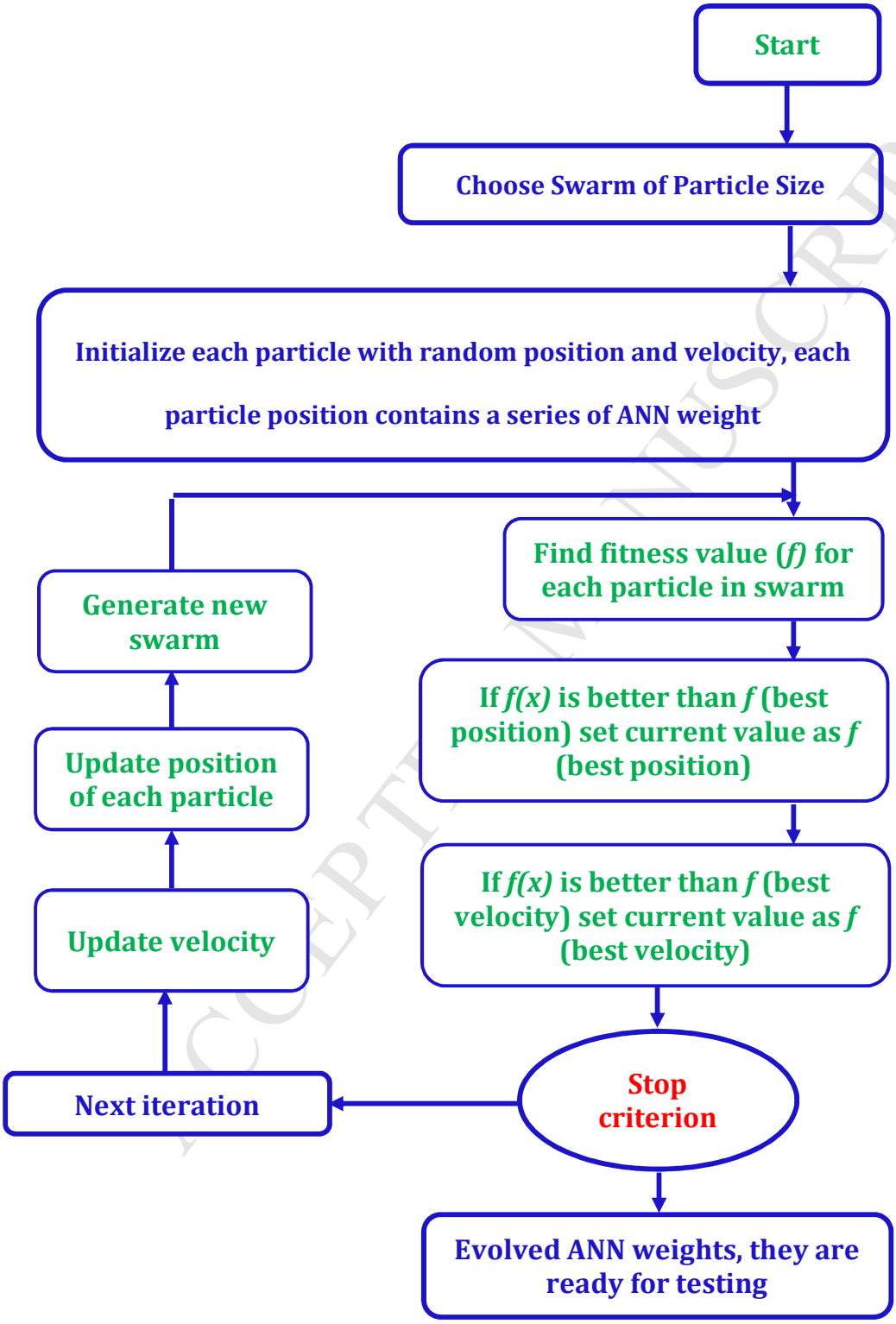


Figure 3: PSO-based algorithm box chart in optimization of the weights of ANN

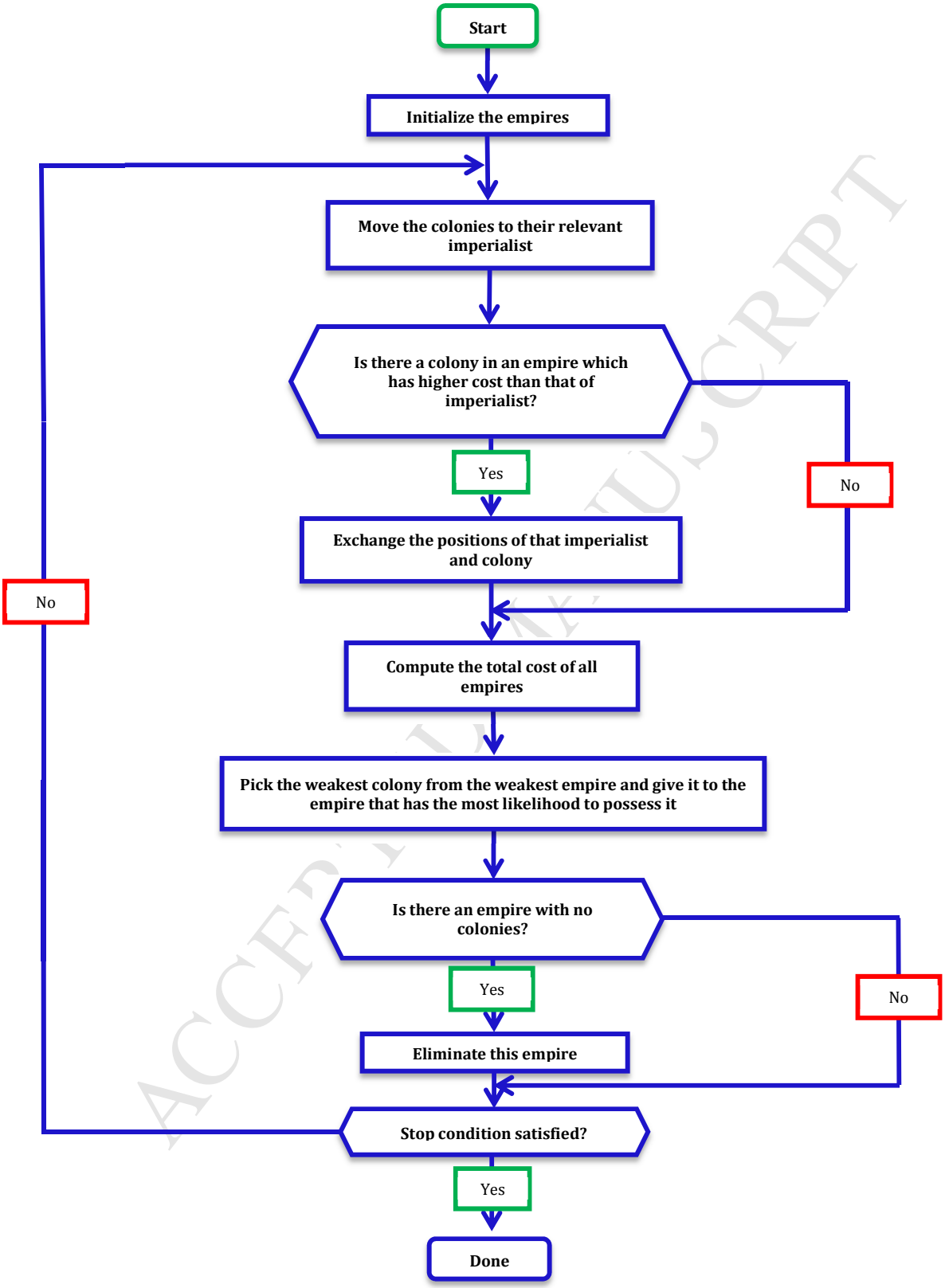


Figure 4: Box chart of imperialist competitive algorithm process [45,47,56]

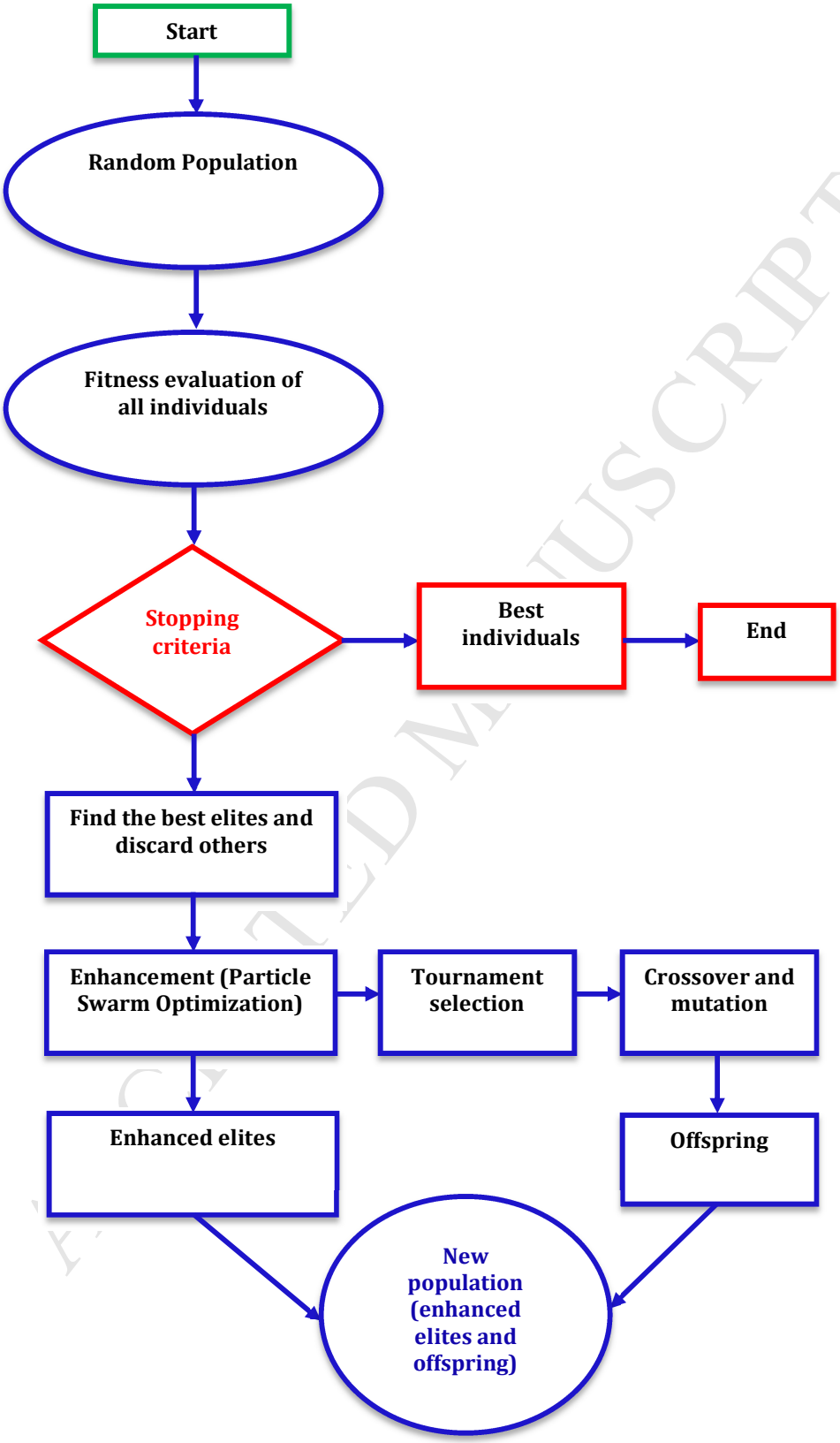


Figure 5: Box chart of HGAPSO approach [44,46,59]

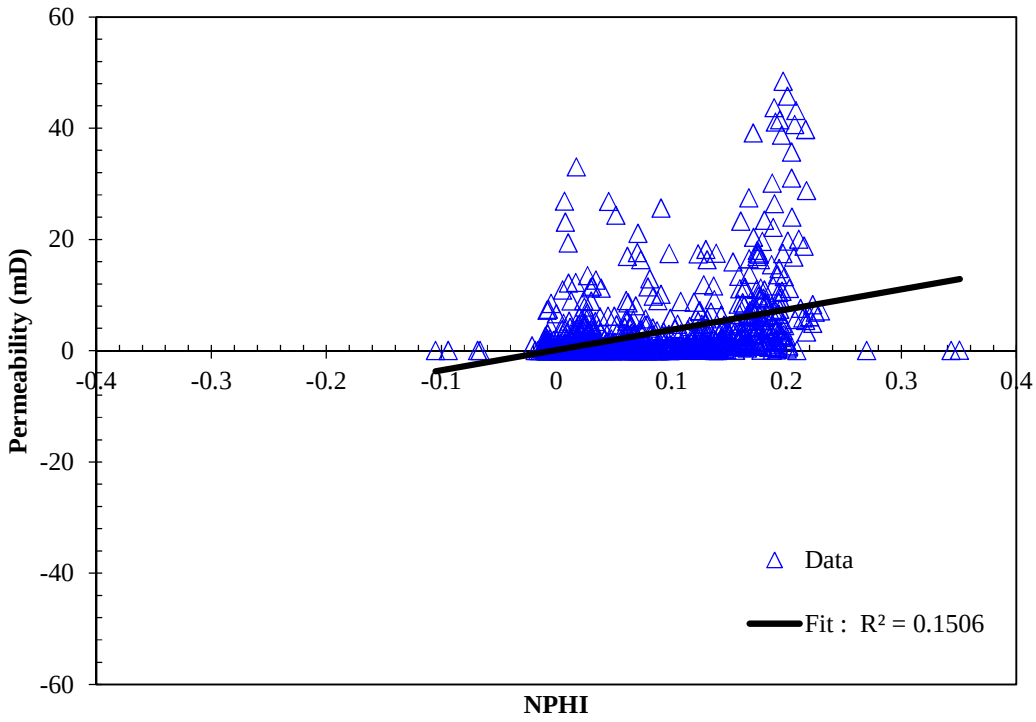


Figure 6: Dependency of permeability on neutron log (NPHI)

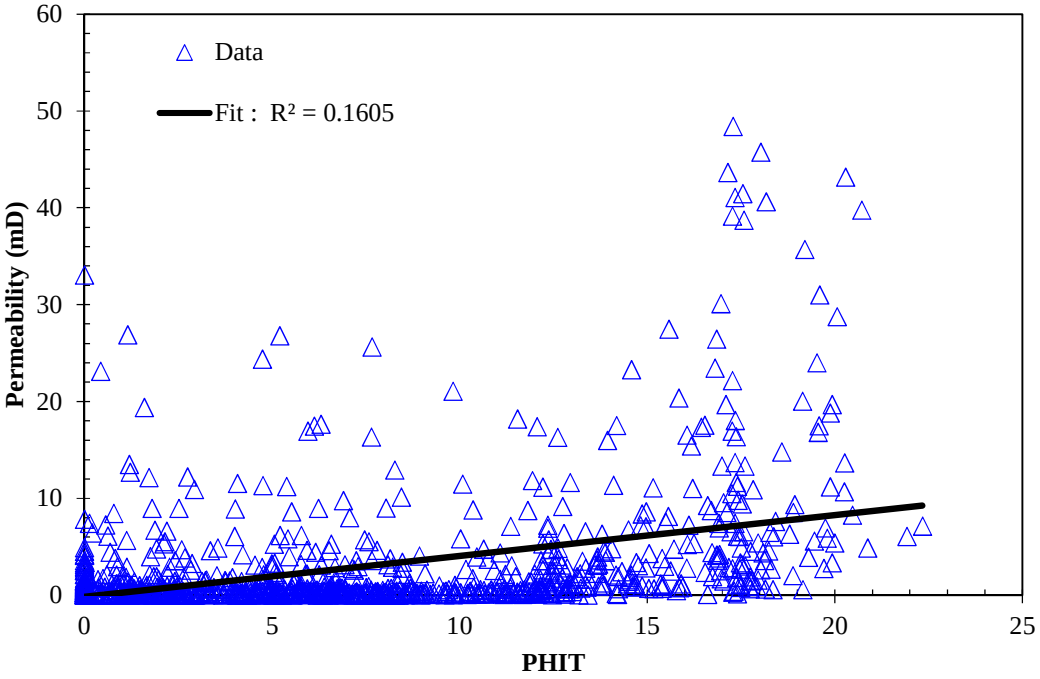


Figure 7: Dependency of permeability on PHIT log

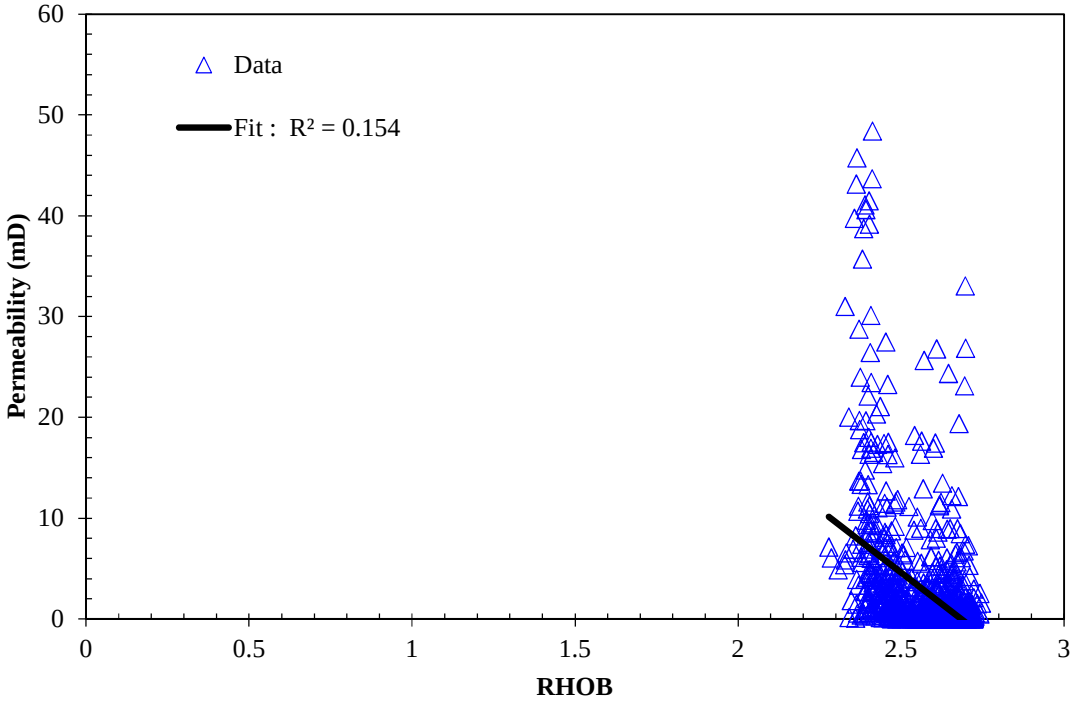


Figure 8: Dependency of permeability on RHOB log

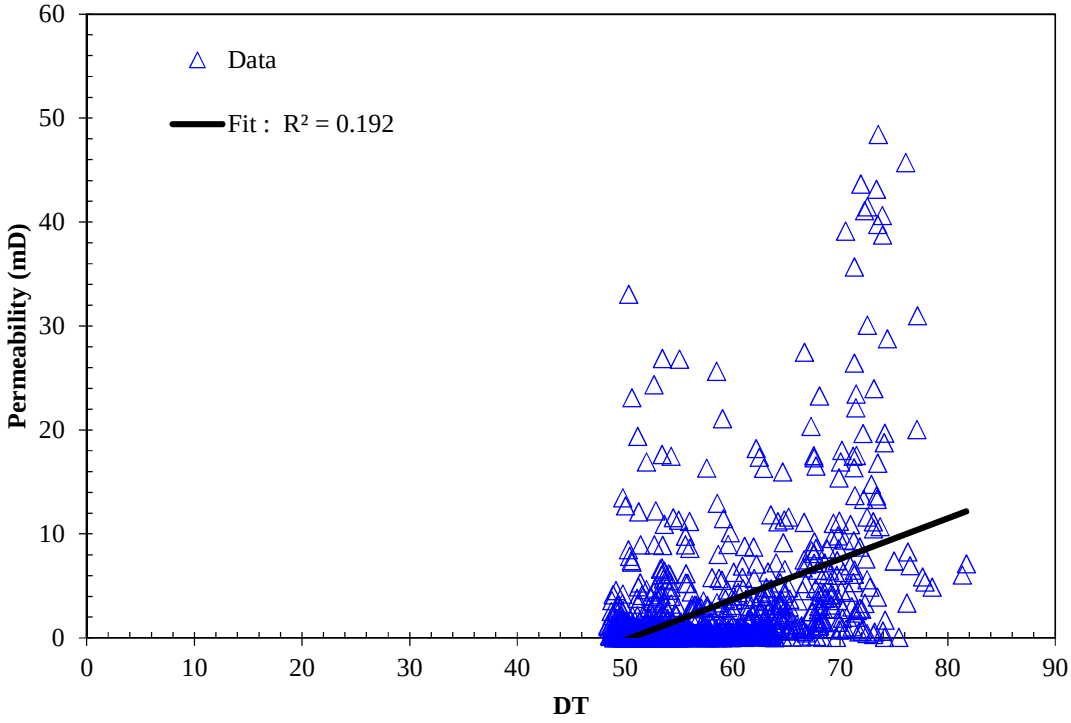


Figure 9: Dependency of permeability on DT log

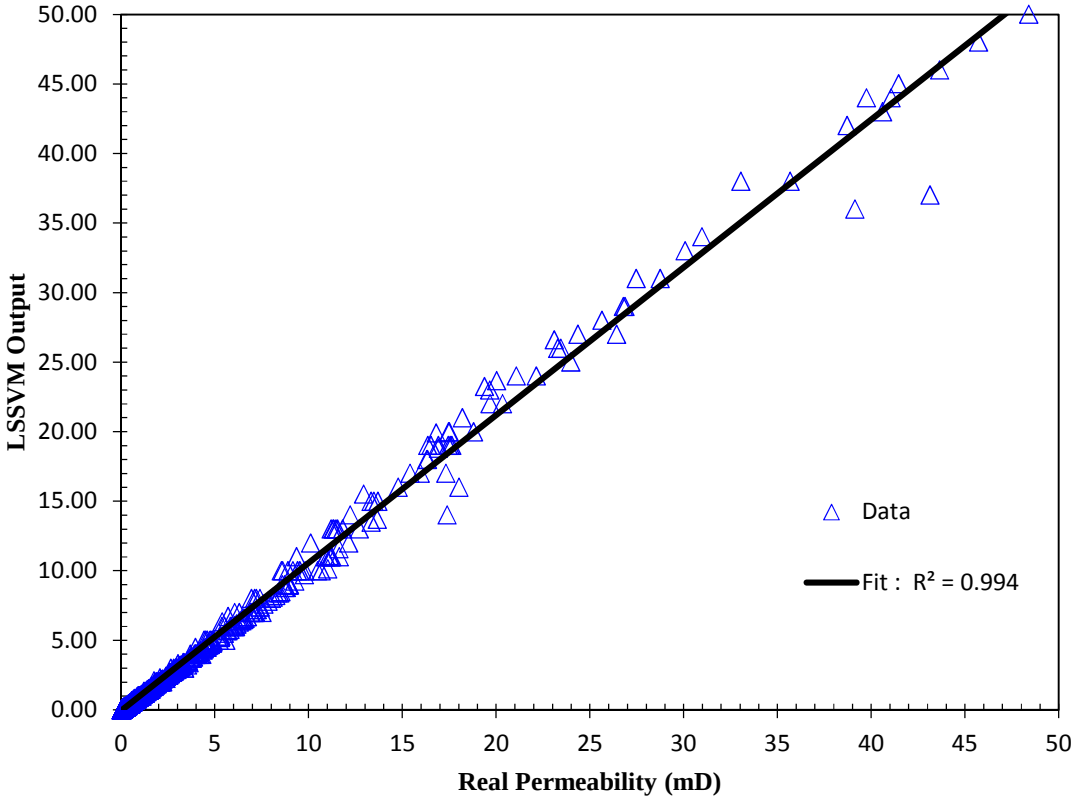


Figure 10: Regression plot of LSSVM output versus relevant real permeability data samples

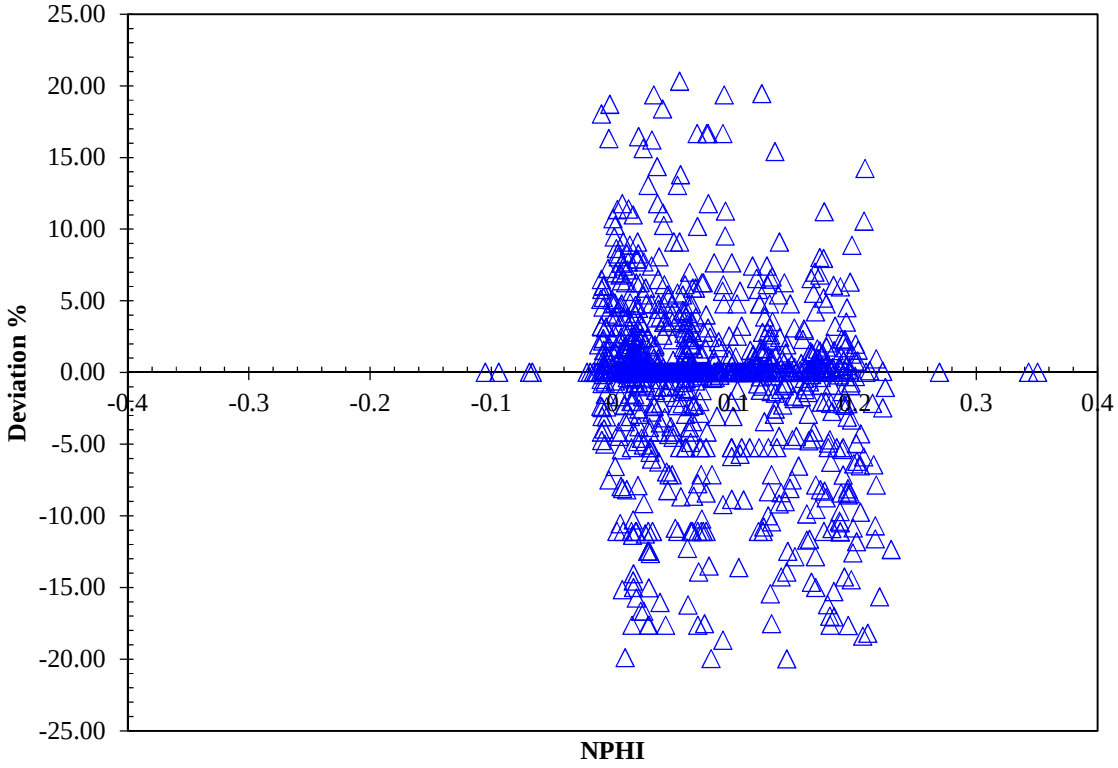


Figure 11: Relative deviation of developed approach versus relevant neutron log (NPHI) data samples

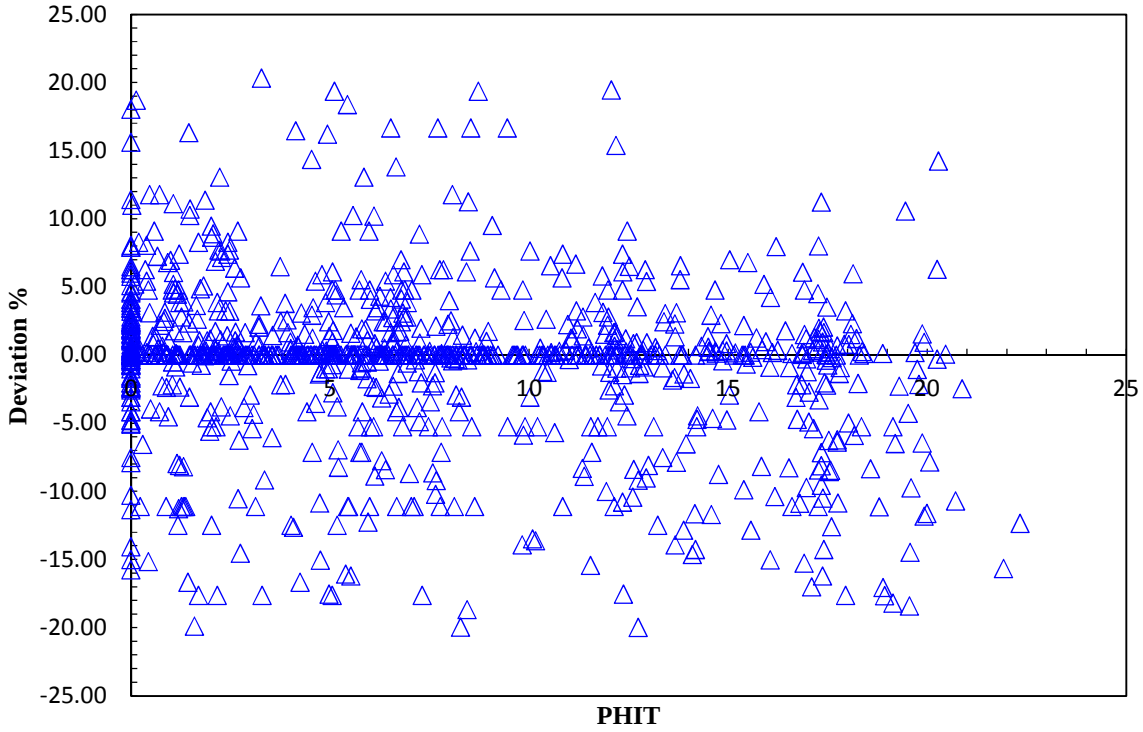


Figure 12: Relative deviation of developed approach versus relevant PHIT log data samples

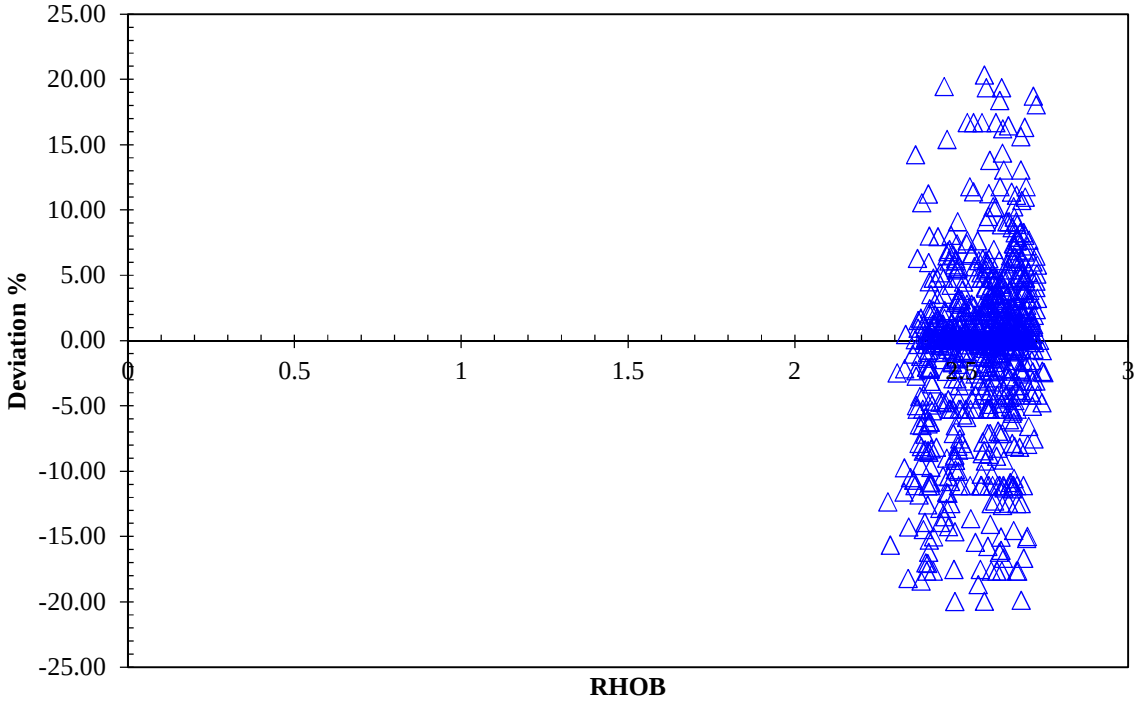


Figure 13: Relative deviation of developed approach versus relevant RHOB log data samples

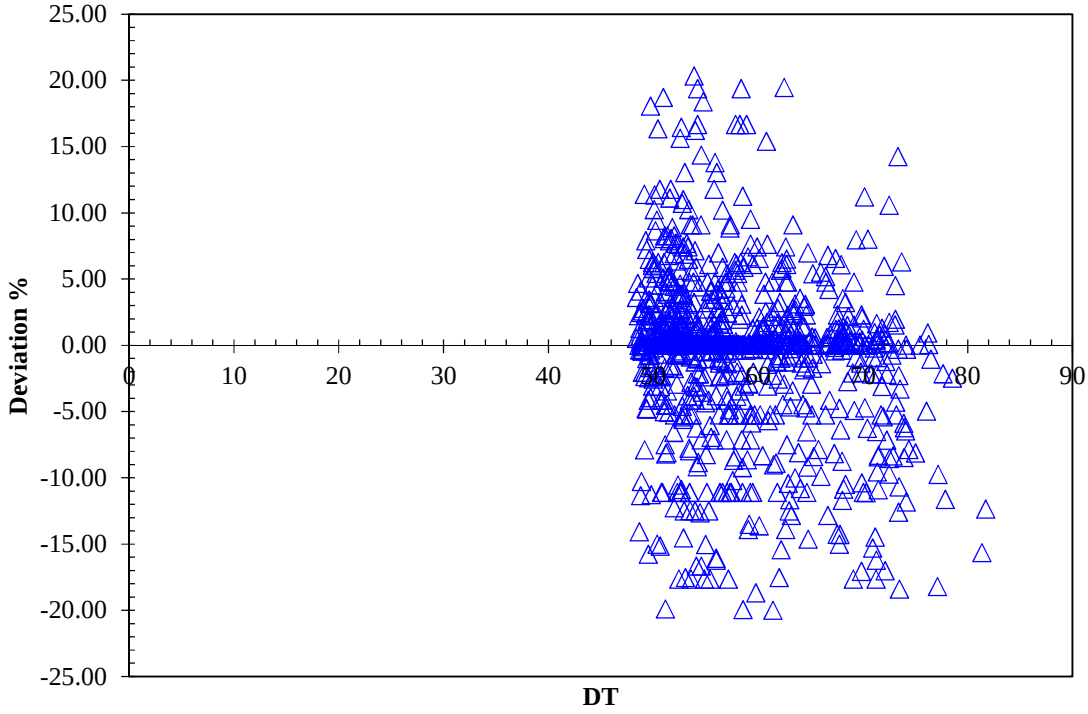


Figure 14: Relative deviation of developed approach versus relevant DT log data samples

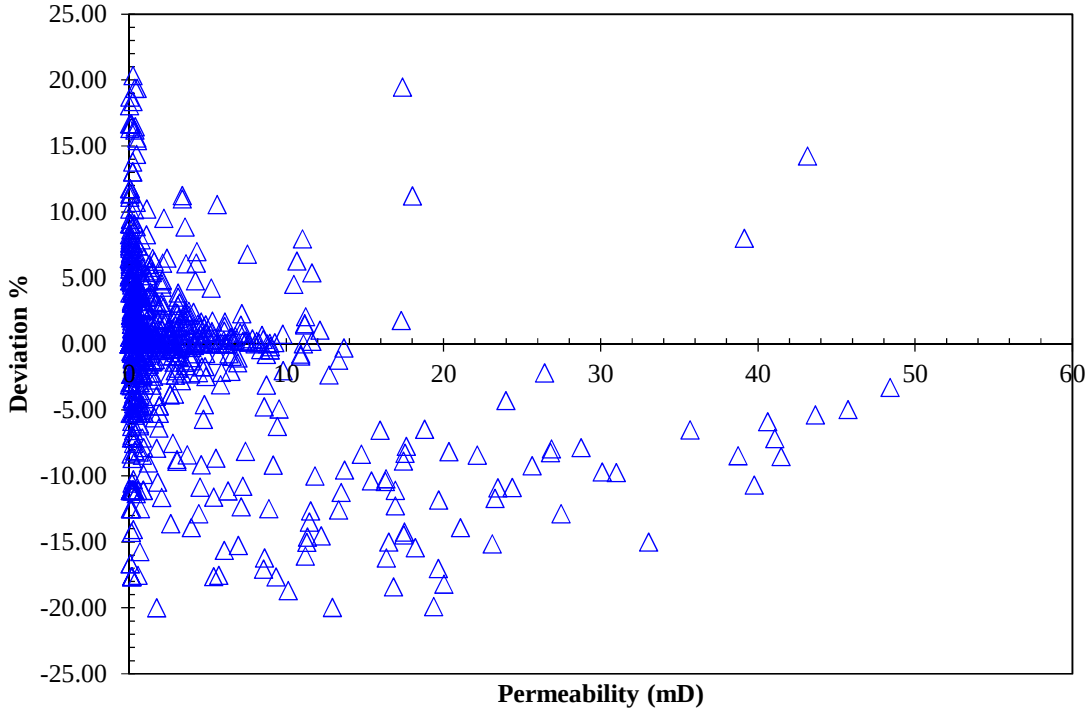


Figure 15: Relative deviation of developed approach versus relevant DT log data samples

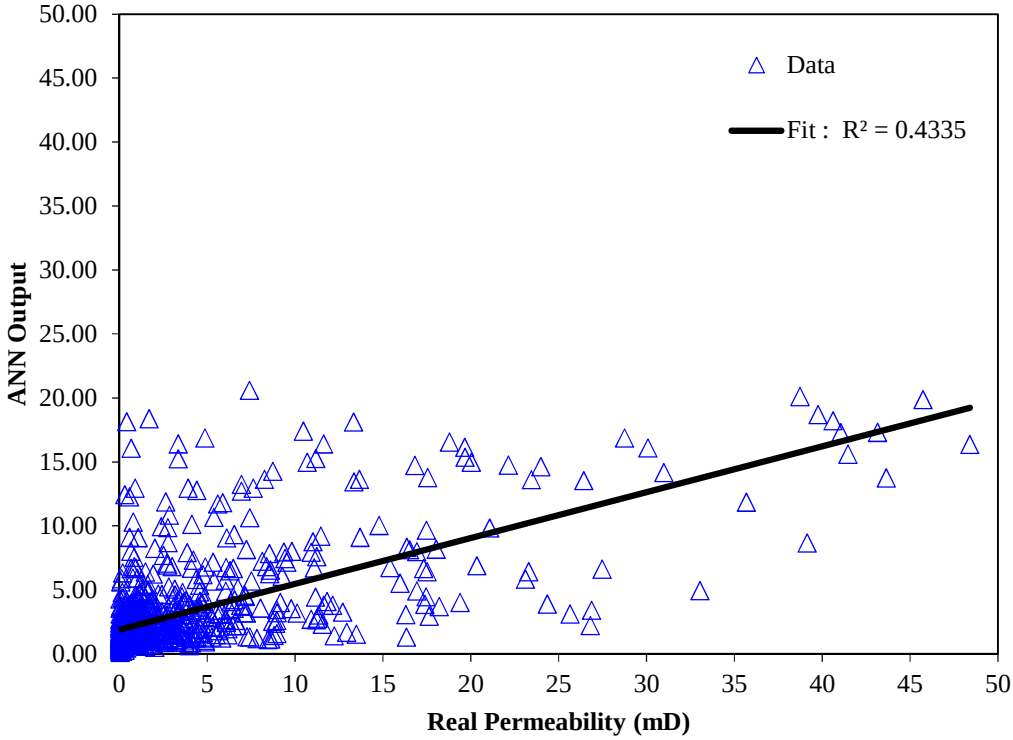


Figure 16: Regression plot of ANN output versus relevant real permeability data samples

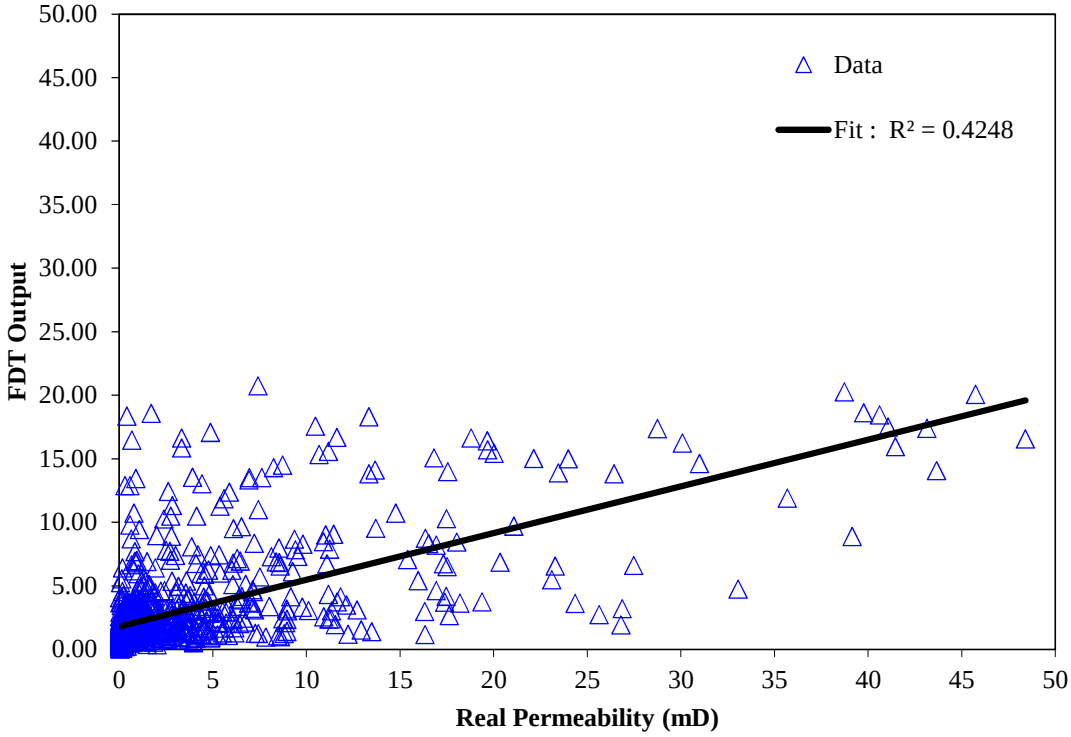


Figure 17: Regression plot of FDT output versus relevant real permeability data samples

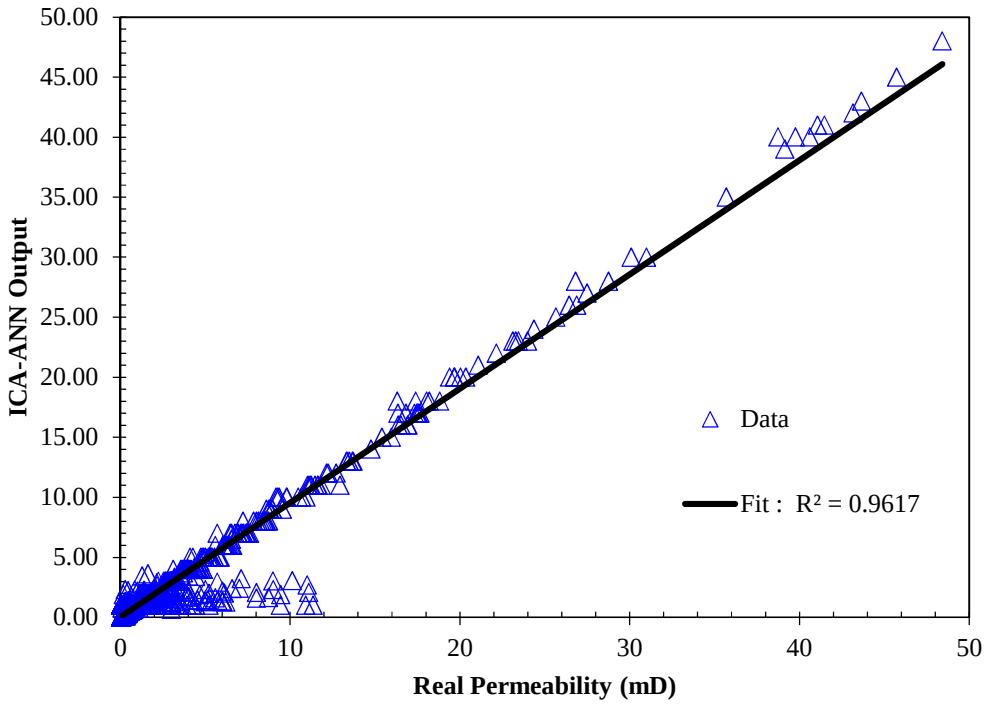


Figure 18: Regression plot of ICA-ANN output versus relevant real permeability data samples

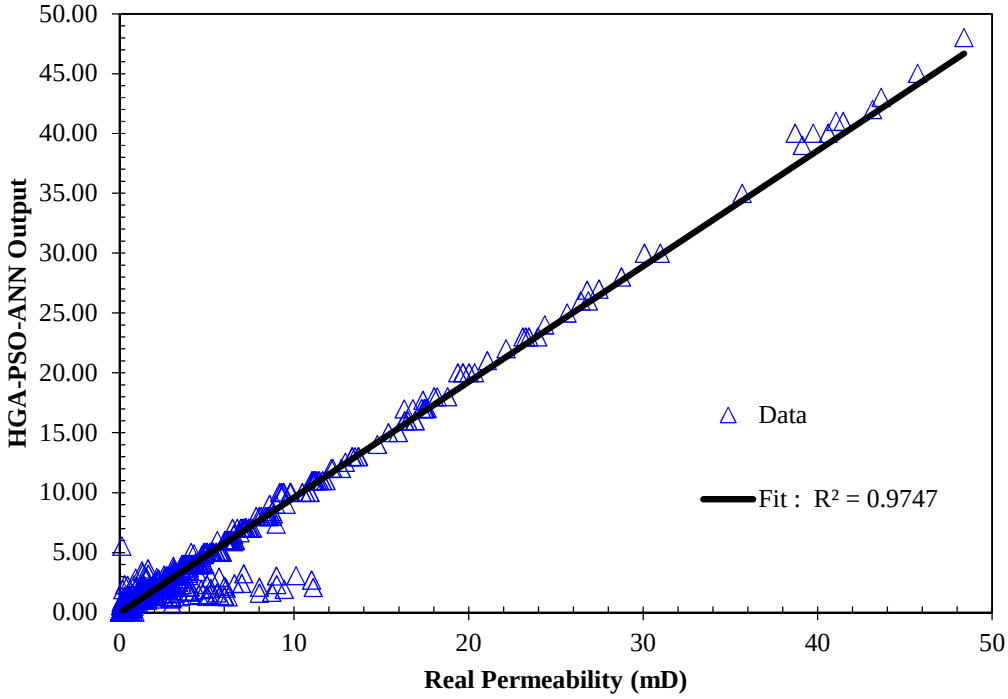


Figure 19: Regression plot of HGAPSO-ANN output versus relevant real permeability data samples

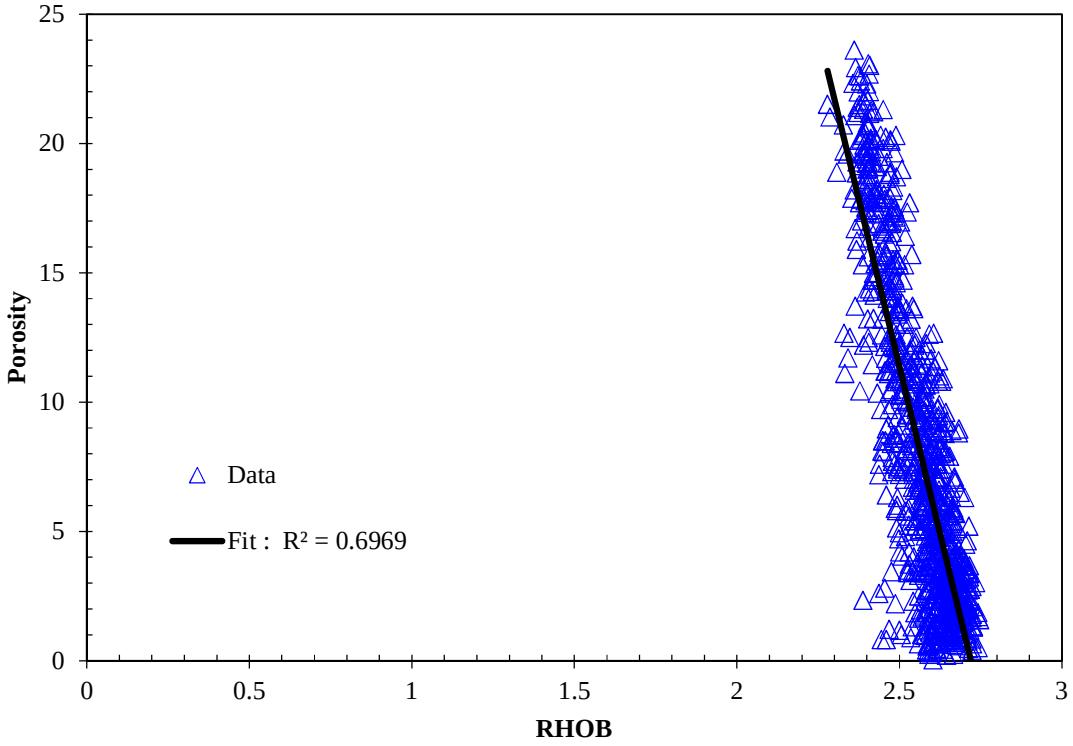


Figure 20: Dependency of porosity RHOB log data samples

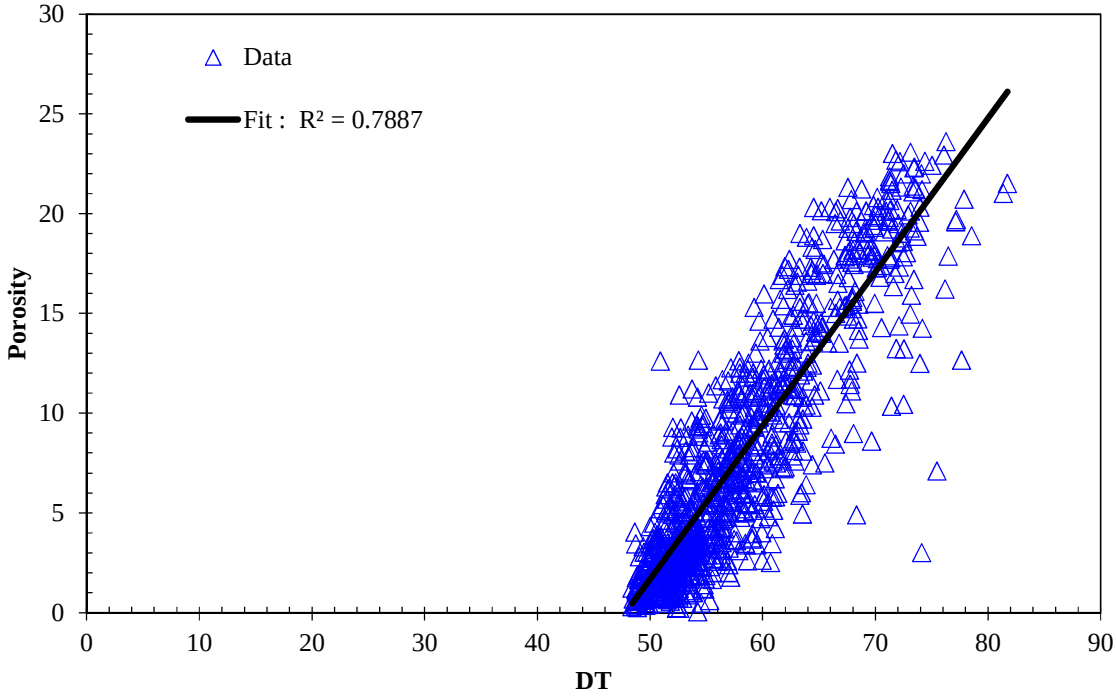


Figure 21: Dependency of porosity DT log data samples

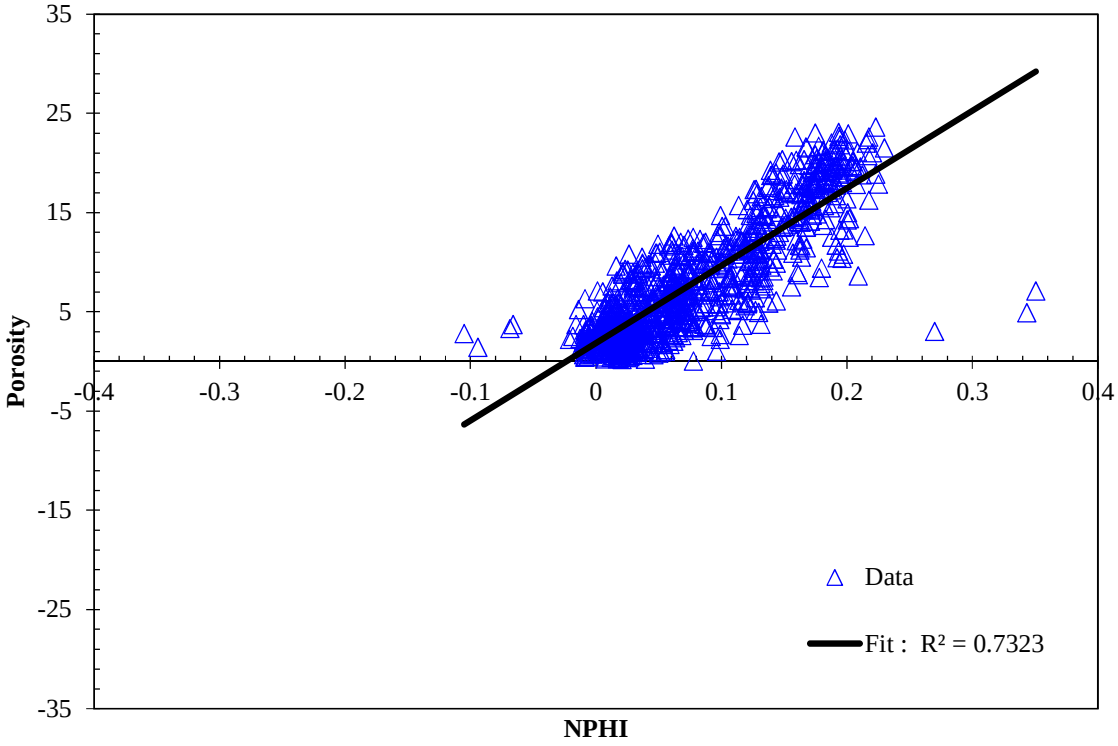


Figure 22: Dependency of porosity NPHI log data samples

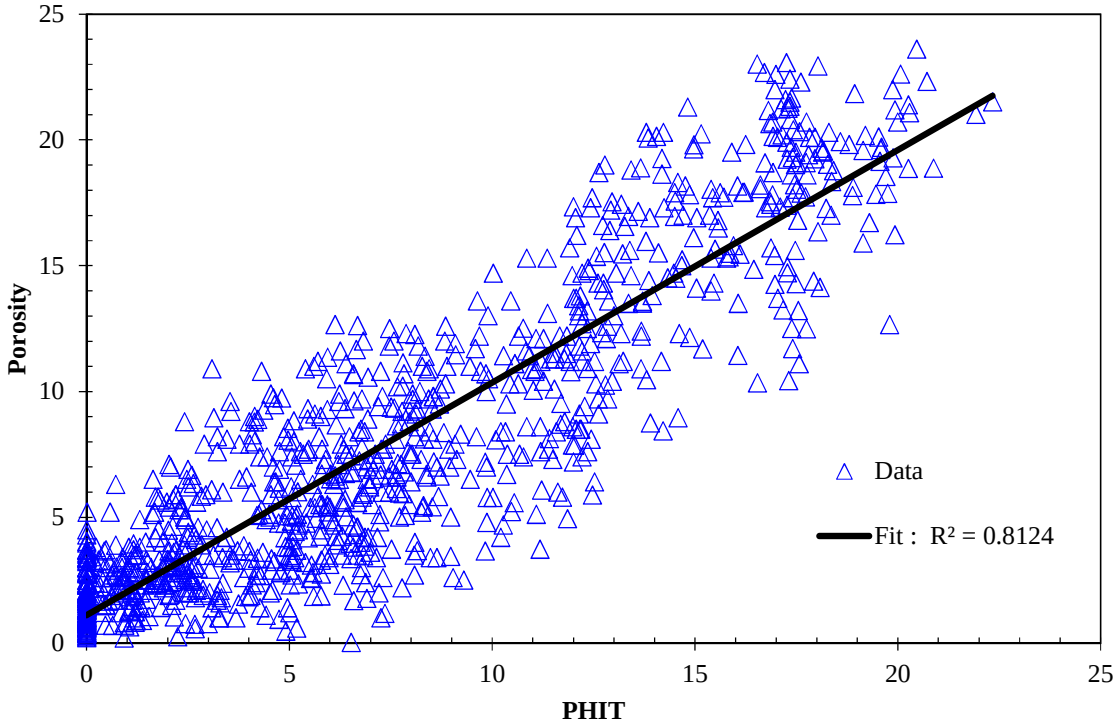


Figure 23: Dependency of porosity PHIT log data samples

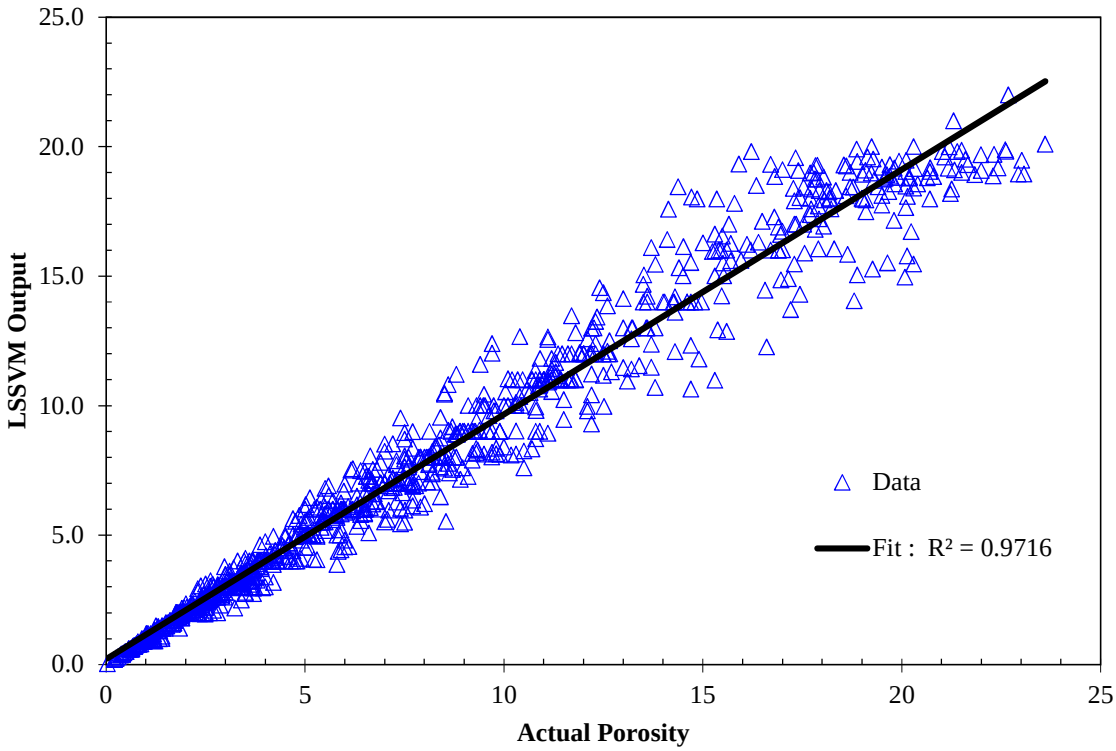


Figure 24: Regression plot of LSSVM output versus relevant real porosity data samples

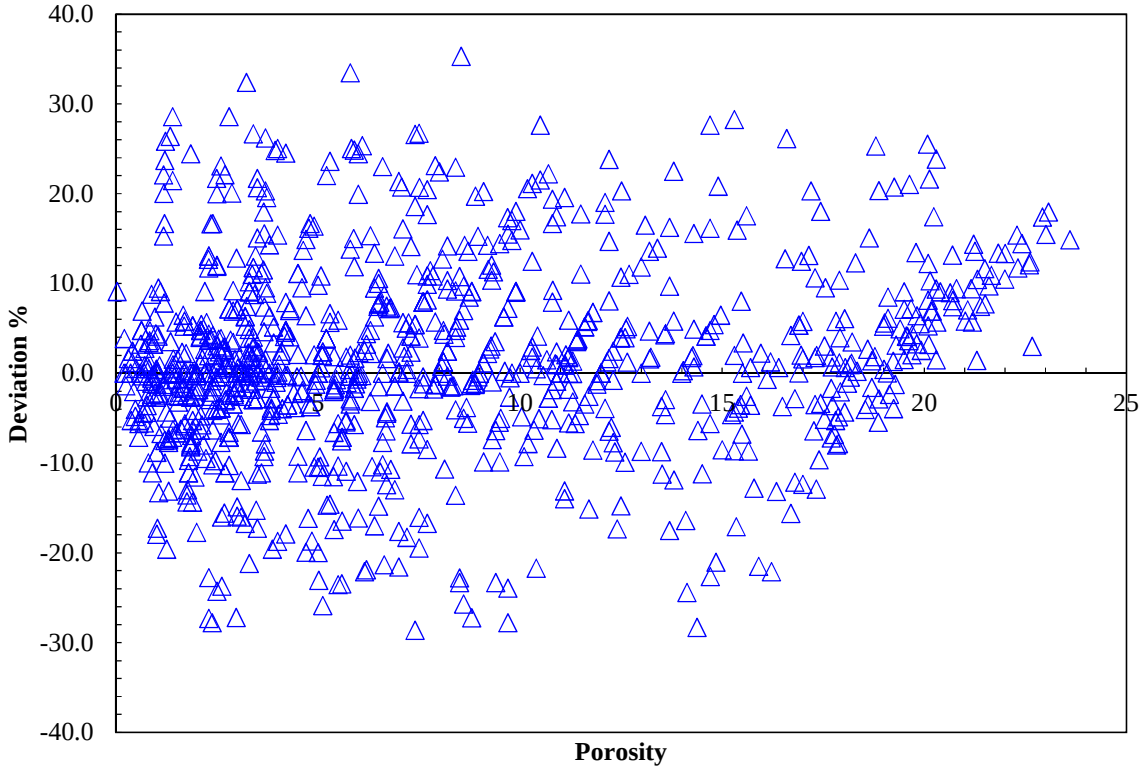


Figure 25: Relative deviation of developed approach for porosity prediction versus relevant porosity data samples

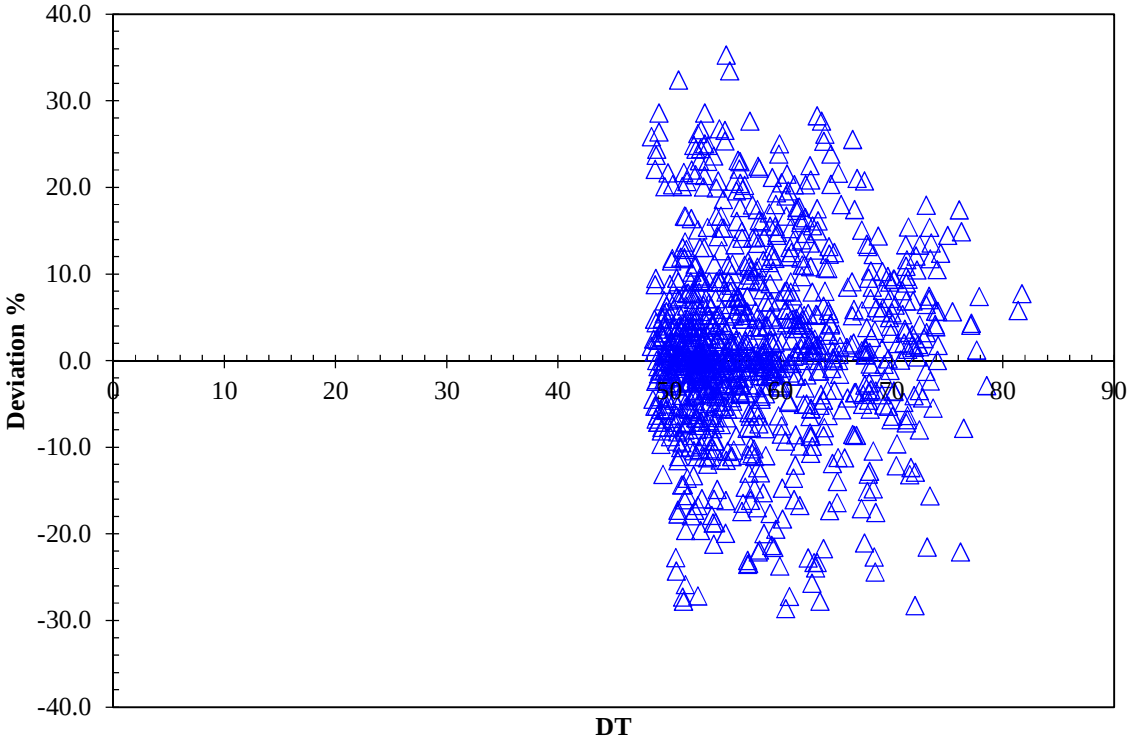


Figure 26: Relative deviation of developed approach for porosity prediction versus relevant DT log data samples

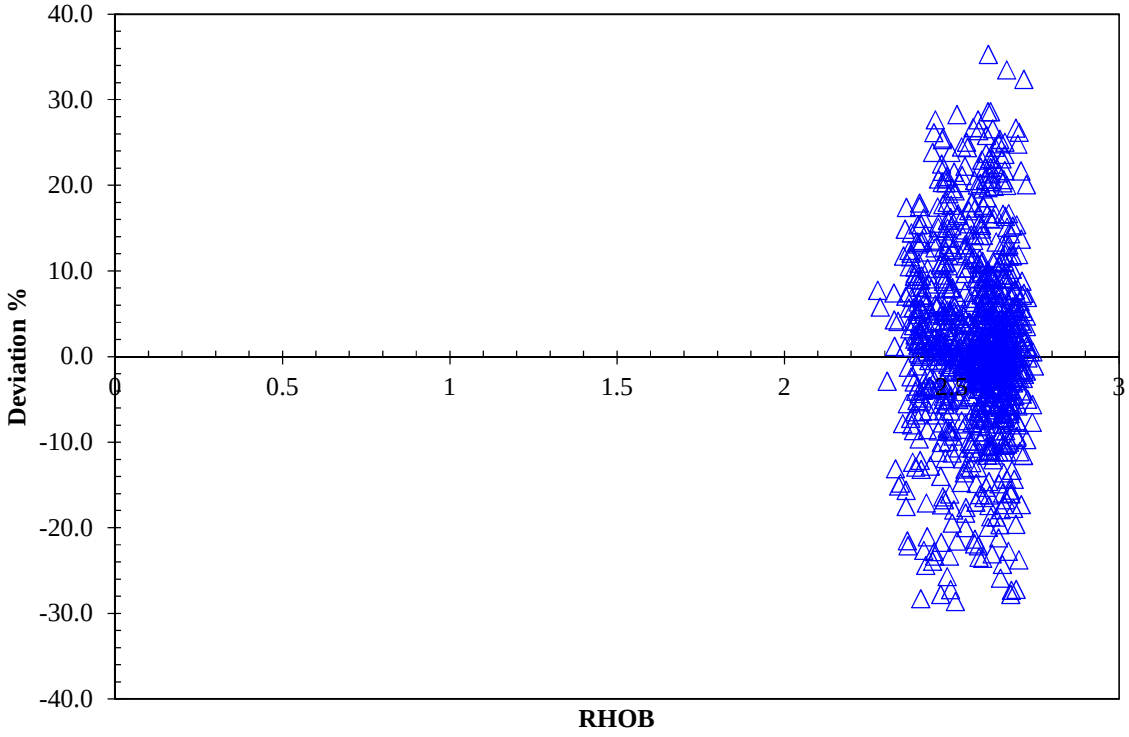


Figure 27: Relative deviation of developed approach for porosity prediction versus relevant RHOB log data samples

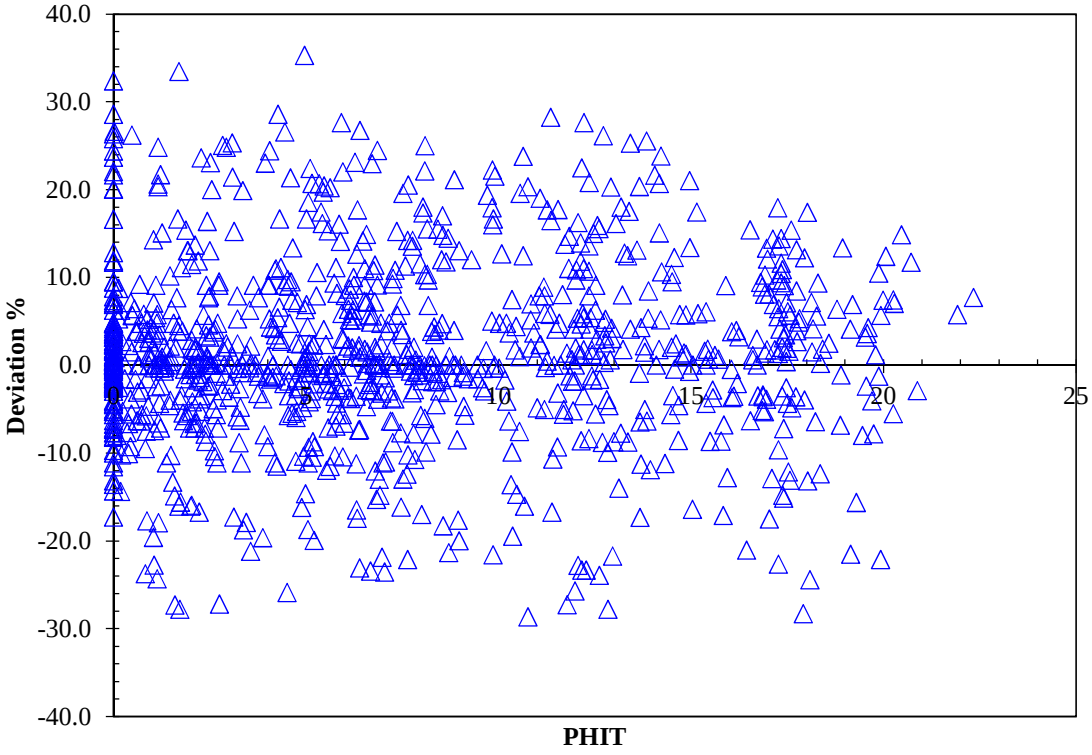


Figure 28: Relative deviation of developed approach for porosity prediction versus relevant PHIT log data samples

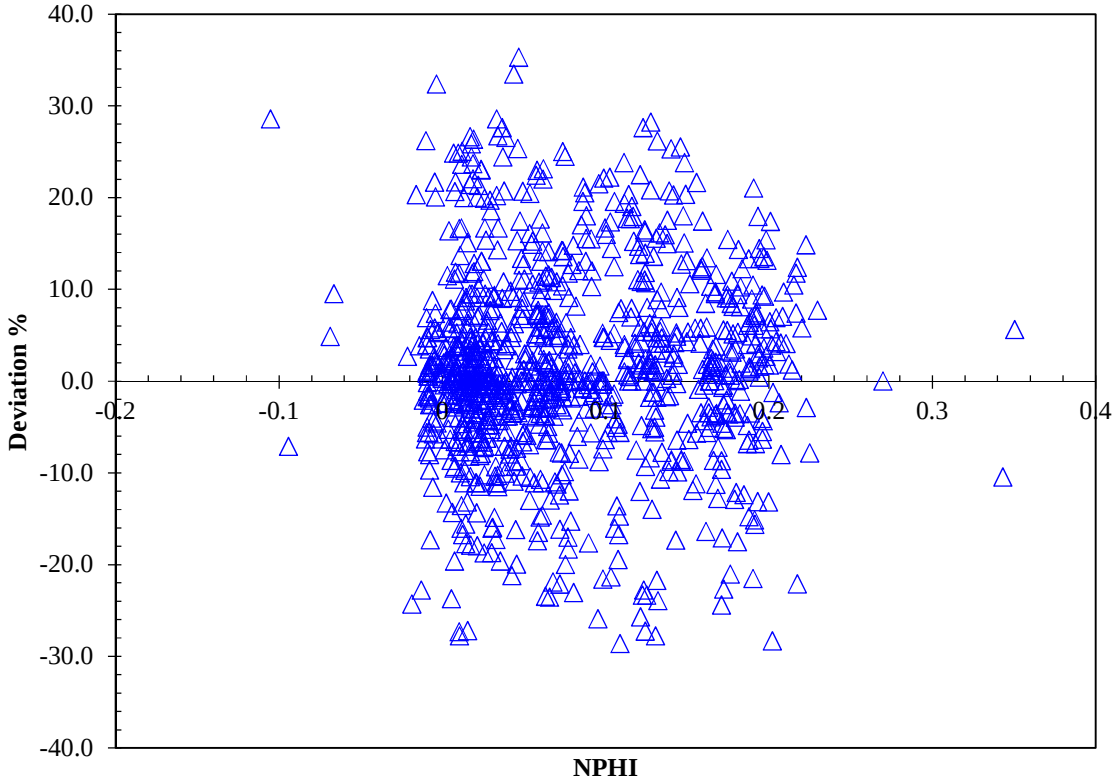


Figure 29: Relative deviation of developed approach for porosity prediction versus relevant NPHI log data samples

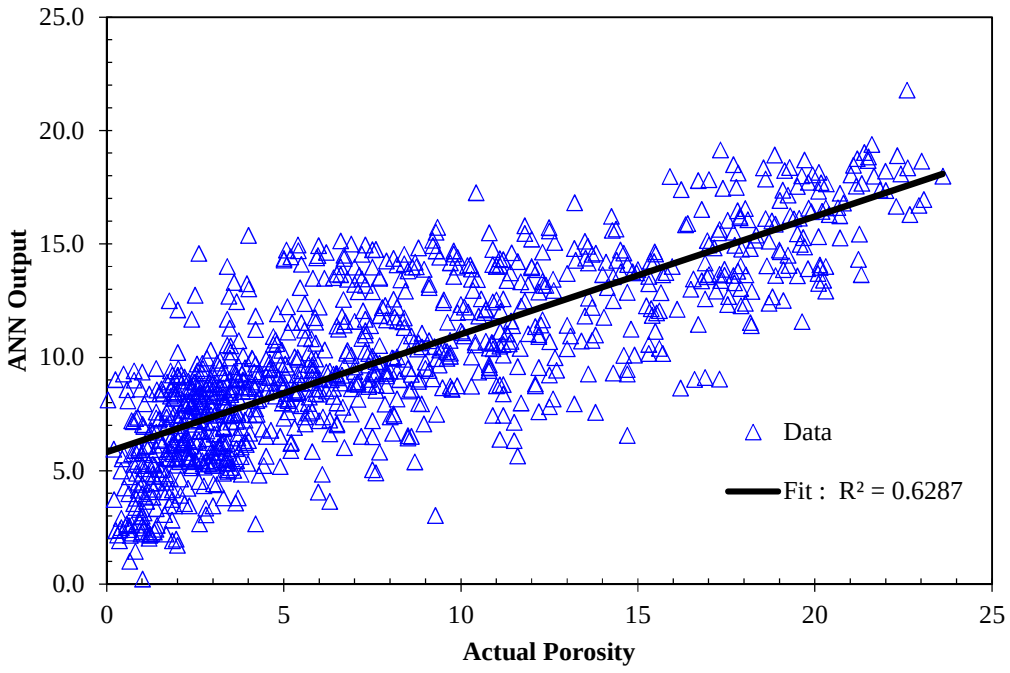


Figure 30: Regression plot of ANN output versus relevant real porosity data samples

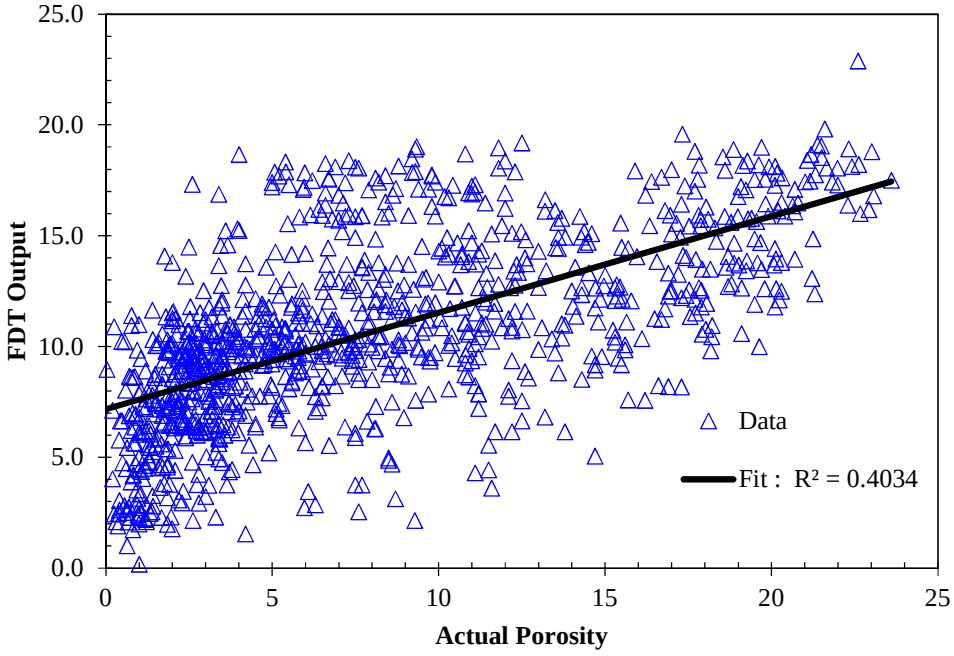


Figure 31: Regression plot of FDT output versus relevant real porosity data samples

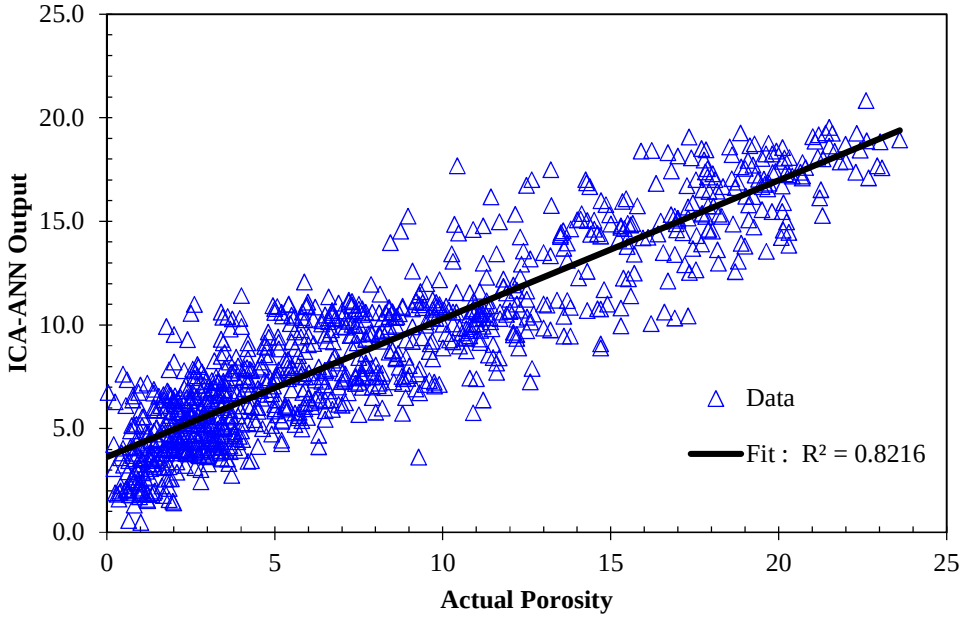


Figure 32: Regression plot of ICA-ANN output versus relevant real porosity data samples

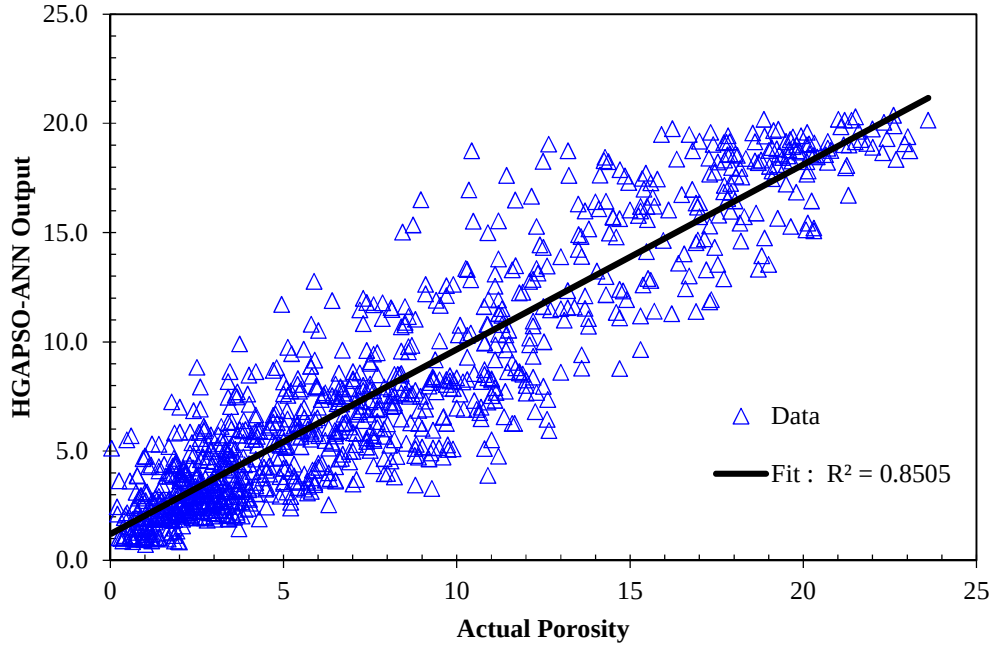


Figure 33: Regression plot of HGAPSO-ANN output versus relevant real porosity data samples