International Journal of Computer Information Systems and Industrial Management Applications.
ISSN 2150-7988 Volume 5 (2013) pp. 413-426

Prediction of Petroleum Reservoir Properties using Different Versions of Adaptive Neuro-Fuzzy Inference System Hybrid Models

Fatai A. Anifowose[1], Jane Labadin[2] and Abdulazeez Abdulraheem[3]

[1] Faculty of Computer Science and Information Technology, Universiti Malaysia Sarawak,
94300 Kota Samarahan, Sarawak, Malaysia
fanifowose@gmail.com

[2] Faculty of Computer Science and Information Technology, Universiti Malaysia Sarawak,
94300 Kota Samarahan, Sarawak, Malaysia
ljane@fit.unimas.my

[3] Department of Petroleum Engineering, King Fahd University of Petroleum and Minerals,
Dhahran 31261, Saudi Arabia
aazeez@kfupm.edu.sa

***Abstract*: This paper presents a comparative study of the performance of three versions of Adaptive Neuro-Fuzzy Inference System (ANFIS) hybrid model and two innovative hybrid models in the prediction of oil and gas reservoir properties. ANFIS is a hybrid learning algorithm that combines the rule-based inferencing of fuzzy logic and the back-propagation learning procedure of Artificial Neural Networks. Functional Networks-Support Vector Machines (FN-SVM) and Functional Networks-Type-2 Fuzzy Logic (FN-T2FL) were proposed to improve the performance of the stand-alone SVM and T2FL models respectively. The FN component of the FN-T2FL hybrid model automatically extracts the most relevant attributes from the input data using the least square fitting algorithm as an improvement over the individual Functional Networks and Type-2 Fuzzy Logic models. The former is more promising as it combines two existing techniques that are very close in performance and well known for their computational stability and fast processing. The FN-SVM hybrid model also benefits from the excellent performance of the least-square-based model-selection algorithm of Functional Networks and the non-linear high-dimensional feature transformation capability that is based on structural risk minimization and Tikhonov regularization properties of SVM. Training and testing the SVM component of the hybrid model with the best and dimensionally-reduced variables from the input data resulted in better performance with higher correlation coefficients, lower root mean square errors and less execution time than the traditional SVM model. A comparison of FN-SVM and FN-T2FL with the three versions of ANFIS showed the superiority of the FN-SVM model over the others. The three ANFIS models still proved to be good in solving real industrial problems due to their speed of execution especially in dense data conditions.**

Keywords: hybrid models, computational intelligence, porosity, permeability, functional networks, support vector machines, ANFIS**.**

I. Introduction

A recent study [1] reported that Computational Intelligence (CI) and Machine Learning techniques such as Functional Networks (FN), Type-2 Fuzzy Logic System (T2FLS) and Support Vector Machines (SVM) have shown to be effective for a wide range of real-world applications. However, the “No Free Lunch theorem” [2] applies since each of these techniques has its limitations and constraints that would not make it appropriate to solve all problems in different data and operational scenarios. This calls for the need to hybridize these techniques so that one of them would complement the limitations and weaknesses of others to ensure increased performance in various challenging real-world scenarios. Thus, hybridization of CI techniques can boost their individual performance and make them achieve much success in dealing with large-scale, complex problems [1].

The concept of hybridizing existing CI techniques is especially useful in oil and gas exploration and production where a little improvement in accuracy of the prediction of various petroleum reservoir properties could lead to a very high increase in the exploration and production of more energy. The two important properties of oil and gas reservoirs that are focused in this study are porosity and permeability. They are the fundamental reservoir properties that relate to the amount of fluid contained in oil and gas reservoirs and its ability to flow. They are frequently measured in the laboratory on rock plugs extracted from the core of wells drilled for oil and gas exploration through the process of Logging While Drilling (LWD) or Measurement While Drilling (MWD) and serve as standard indicators of reservoir quality in the oil and gas industry. The measured properties from the laboratory are then depth-matched with the log data taken from the actual wells in the field. Well logs, measured from a probe lowered into the borehole at the end of an insulated cable, provide a wealth of information about the wells that have been drilled. The measurements are recorded graphically or digitally as a function of depth and are commonly known as geophysical well logs, petro-physical logs, or simply well logs [3, 4].

Some of the uses of well logs in the oil and gas industry include identifying potential reservoir rocks, determining reservoir bed thickness, locating hydrocarbons; estimating water saturation, quantifying amount of hydrocarbons, estimating the type and rate of reservoir fluid production, estimating reservoir formation pressure, determining porosity, estimating permeability, and identifying fracture zones in rocks [5]. When the well logs are combined with the core measurements, the new entity serves as a great resource in the determination of subsurface properties for the location, exploration, production and exploitation of a new well to be drilled.

Since both the laboratory and field measurements are usually costly and time-consuming, CI techniques [6 - 8] as well as hybrid methodologies [9 - 12] have been successfully applied in the prediction of these properties to an acceptable degree of accuracy. Due to the limitations of each CI technique that would not make its application desirable in certain operational conditions such as small dataset scenarios [13 - 15] and high dimensionality of data conditions [15, 16], hybrid CI has caught the attention and interest of researchers and practitioners in the oil and gas industry and has become increasingly popular [10]. A good number of hybrid models have been reported to be successfully applied in petroleum engineering [17 - 20].

Since FN partly uses a least-square algorithm that selects the best subset of features from a set of input data and SVM is known for its capability to conveniently handle data of high dimensionality with its insensitivity to data size, combining these two techniques in a way that they complement each other has shown to be a welcomed development [1]. Similarly, an improvement in the prediction performance of T2-FLS was shown in [21] with the use of FN. The importance of feature selection in the improvement of prediction and classification accuracies has also been studied [22].

This proposed study implements three versions of Adaptive Neuro-Fuzzy Inference System (ANFIS) and compares the results with those of FN-SVM and FN-T2F in the prediction of porosity and permeability of oil and gas reservoirs. ANFIS was selected for this comparative study because it is the only hybrid CI technique that is available as a standard toolbox in MATLAB software [23]. Our major motivations for this study are the continued discovery of various CI techniques with common denominators that are suitable for hybridization and the consistent quest for better techniques in the prediction of petroleum reservoir properties for the production of more energy.

II. Survey of Literature

A. Oil and Gas Reservoir Characterization

Oil and gas reservoir characterization is a process for quantitatively describing various reservoir properties in spatial variability using available field and laboratory data [6]. Reservoir characterization plays a crucial role in modern reservoir management. It helps to make sound reservoir decisions and improves the asset value of the oil and gas companies. It maximizes integration of multidisciplinary data and knowledge, and hence improves the reliability of reservoir predictions. The ultimate goal is a reservoir model with realistic tolerance for imprecision and uncertainty [11]. Reservoir characterization focuses on modeling each reservoir unit, predicting well behavior, understanding past reservoir performance, and forecasting future reservoir conditions.

Different borehole and reservoir formation conditions may require different tools to measure the same property. There are many important subsurface properties that need to be detected or measured but porosity and permeability are the most important properties since they jointly serve as a major indicator of petroleum reservoir quality and economic viability. The data acquired from these datasets are used for the estimation of porosity, permeability and other reservoir properties such as rock types, the thickness of rock layers; the amount of hydrocarbons; and water salinity.

Porosity is an important consideration when attempting to evaluate the potential volume of hydrocarbons contained in a reservoir as it is a measure of the percentage of voids and open spaces in a rock. These voids and spaces are potential receptacles for oil and gas. Permeability, on the other hand, is a key parameter in the characterization of any hydrocarbon reservoir as it is a measure of how interconnected the individual voids and spaces are in a rock. In fact, many petroleum engineering problems cannot be solved accurately without having an accurate value of permeability [3, 4]. In view of the importance of porosity and permeability in oil and gas exploration and production, this study focuses on these two properties.

B. Adaptive Neuro-Fuzzy Inference System (ANFIS)

ANFIS, proposed by [23], is a result of an intelligent combination of the learning capabilities of Artificial Neural Networks (ANN) and the reasoning capabilities of Fuzzy Logic as a hybrid intelligent system. It is a network structure that implements Fuzzy Inference System (FIS), a knowledge representation where each fuzzy rule describes a local behavior of the system, and employs hybrid learning rules for training. It is a class of adaptive networks which enjoys many of the advantages claimed by ANN and the linguistic interpretability of FIS while combining the gradient descent and the least-squares method technologies. The neuro-fuzzy model of ANFIS is a multilayer neural network-based fuzzy system. Its basic architecture is shown in Figure 1 with the system having a total of five layers [24].

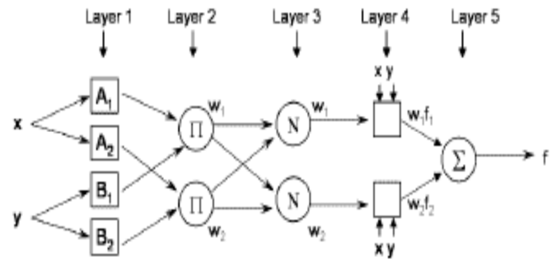

Figure 1. The Basic Architecture of ANFIS [24]

The main objective of ANFIS is to integrate the best features of fuzzy systems, through the representation of prior knowledge into a set of constraints (network topology) to reduce the optimization search space; and ANN through the adaptation of the back-propagation to the structured network for automating the fuzzy system's parametric tuning. ANFIS is one of the best tradeoff between neural and fuzzy systems as it provides smoothness due to the fuzzy system's interpolation capability; and adaptability due to the ANN's back-propagation algorithm [25]. However, the major weakness of ANFIS lies in its strong computational complexity restrictions [25, 26]. Some of the advantages and disadvantages of ANFIS have been presented in [27].

The role of each layer of the ANFIS hybrid system is as follows [24]:

- Layer 1: This takes input from the data. Every node is an adaptive node. The input parameters handled by this layer are called premise parameters.
- Layer 2: Every node is fixed. The output is the product of all the incoming signals. Each node output represents the firing strength of a rule. This forms the core part of the Fuzzy Logic System component of the hybrid network.
- Layer 3: Every node is fixed. Each calculates the ratio of the i^{th} rule's firing strength. Thus the outputs of this layer are called normalized firing strengths.
- Layer 4: Every node is an adaptive node. Parameters are referred to as consequent parameters. This is the core part of the back-propagation algorithm of ANN in the hybrid system.
- Layer 5: The single node in this layer is a fixed node. It computes the overall output as the summation of all incoming signals.

In ANFIS, there are two information passes in the hybrid algorithm: forward pass and backward pass. In the forward pass, the node outputs go forward up to layer 4 while the consequent parameters are identified by the least squares method of the back-propagation algorithm. In the backward pass, the error signals propagate backwards down to layer 2 where the premise parameters are updated by gradient descent algorithm of the Fuzzy Logic component. More details about the architecture of ANFIS can be found in [24].

C. Hybrid Computational Intelligence in Oil and Gas

Some of the successful applications of Hybrid Computational Intelligence (HCI) are found in the bioinformatics, science, technology and engineering. These include the selection of winding material in electric power transformers using a hybrid of decision trees for attribute selection and neural networks for winding material classification for the calculation of the performance characteristics of each considered design [28]; and a hybrid of genetic programming approach and a heuristic rule-based scheme for the classification between different types of aphasia, a human syndrome, often due to brain damage [29]. These as well as others were discussed in [1]. Other applications of the advances of HCI include [30 - 32].

In petroleum engineering, HCI has been successfully applied in many areas such as seismic pattern recognition; porosity and permeability predictions; identification of sandstone lithofacies; drill bit diagnosis; and analysis and improvement of oil and gas well production.

While appreciating the increased application of Functional Networks and Type-2 Fuzzy Logic as stand-alone CI techniques in studies such as [33] and [34], reference [21] proposed a FN-T2FL hybrid model that utilized the capability of the least-square fitting algorithm of FN to extract the most relevant attributes for the T2FL component. Their results, in addition to confirming that the FN-T2FL hybrid model performed better than the individual FN and T2FL techniques used individually, also demonstrated that there is more potential in the hybridization of existing techniques for better accuracy and increased characterization of oil and gas properties. Their work was motivated by the successful trials of various hybrid models such as [7] that produced 2-D fracture intensity and fracture network maps in a large block of field using Artificial Neural Network (ANN) and Fuzzy Logic (FL); [35] that used ANN to predict permeability from petrographic data while using Fuzzy Logic to screen and rank the predictor variables with respect to the target variable; and [36] that proposed a new method for the auto-design of ANN based on Genetic Algorithm.

In [1], the authors built upon their experience on the success of FN-T2FL hybrid model to propose a FN-SVM hybrid model. They argued that the combination of Fuzzy Logic and Artificial Neural Networks (ANN) for the prediction of permeability by means of Flow Zone Index in [37]; the use of Genetic Algorithm (GA) to tune the parameters of ANN in [36]; the combination of Fuzzy Logic and GA for the optimization of gas production operations in [38]; and the use of any of GA, Fuzzy Logic and ANN in other studies [39 - 41] could not have been the best idea based on the following reasons:

- Though, GA is a very robust optimization algorithm that is based on an exhaustive search paradigm, it is well known for its long execution time, its need for high processing power due to its computational complexity and sometimes inefficiency as it gets cut up in some local optima [42].
- Fuzzy Logic becomes complex and time-consuming when applied on high-dimensional data [17] and performs poorly when applied on datasets of small size [22].
- ANN is also known to suffer from many deficiencies such as having no general framework for the design of its appropriate network for a specific task and its frequent requirement of large number of parameters to fit a good network structure [43, 44].

ANFIS has been widely used in various application areas most especially bioinformatics [45, 46], environmental sciences [47], image processing [48], manufacturing engineering [49 - 51], finance [52, 53], and mining [54]. In the prediction of permeability of oil and gas reservoirs, no study was found in literature that compares ANFIS as a hybrid model with any other proposed hybrid system for investigating possible improvements in performance accuracy. Few available studies compared ANFIS with other techniques such as the conventional porosity-permeability transform [39], multi-linear regression technique [41], ANN and conventional empirical transformation [55].

A closely related study to this work is [11]. However, the authors used three hybrid components comprising of FN, T2FL and SVM to propose two hybrid models. In the analysis of the three-component hybrid models, some redundancies were introduced by the presence of T2FL and SVM components as their inclusion in the hybrid system could not be adequately justified. The presence of the three components also made the entire hybrid system too complex to analyze.

This paper further seeks to investigate the capability and robustness of the FN-T2F and FN-SVM models in comparison with ANFIS, the first and the only standard hybrid model (to the best of our knowledge) to be technically and commercially available in the MATLAB Toolbox [23] which is the most widely used tool for most technical computing applications in both industry and the academia.

III. Description of Data, Experimental Design and Model Framework

A. Description of Data

In order to establish a strong basis for a fair comparison, the same sets of porosity and permeability datasets from six wells (three for porosity and three for permeability) that were used for the testing and evaluation of FN-T2FL [22] and FN-SVM [1] models were also used in this study. Site 1, a heterogeneous platform that is made up of carbonate and dolomite, contains six predictor variables for porosity while site 2, majorly of carbonate and sandstone formations, contains twelve predictor variables for permeability. Hence, the datasets are representative of the major oil-bearing geological formations found in most parts of the oil-producing world.

B. Experimental Design

The hybrid computational intelligence and machine learning approaches form the basis of the methodology employed in this study.

1) Design of FN-T2F and FN-SVM Frameworks

For the purpose of avoiding repetition of methodologies in this paper, readers are referred to [1] and [21] for the detailed design framework, methodologies and optimized parameters employed in the design of the FN-SVM and FN-T2F hybrid models respectively.

As shown in Figure 2, the FN-T2F hybrid model is composed of two blocks containing, respectively: FN and T2FL. In the FN block, the training procedure that uses the least-squares fitting algorithm was incorporated to select the best variables from the input data. The output of this block, the most relevant attributes of the input data, were then divided into training and test sets. The training set was used to train the T2FL model's Gradient-Descent approach and a Gaussian membership function while optimizing the model parameters. The trained hybrid model was then evaluated using the testing data subset.

Figure 3 shows the design framework of the FN-SVM hybrid model. Similar to the FN-T2F, the FN component was used as a best-subset selector for the SVM component. The most relevant attributes from the input data outputted by the FN block were divided into training and testing subsets. The training subset was used to train the SVM component which was later used to predict the target variable in the testing subset with the actual values hidden from the system to test its generalization capability.

2) Design of the ANFIS Hybrid Model

The framework of the ANFIS hybrid model used in this study was the one proposed by [23], which is available in the MATLAB Toolbox [26], but extracted, customized and combined with other functions. The general behavior of ANFIS is described in [44] as:

Rule 1: IF ($x_{11} = A_{11}$), ($x_{12} = A_{12}$), ... , and ($x_{1m} = A_{1m}$) THEN ($f_1 = p_{1x1} + q_{1x1} + r_1$)

Rule 2: IF ($x_{21} = A_{21}$), ($x_{22} = A_{22}$), ... , and ($x_{2m} = A_{2m}$)THEN ($f_2 = p_2x + q_2y + r_2$)

.
.
.

Rule n: IF ($x_{n1} = A_{n1}$), ($x_{n2} = A_{n2}$), ... , and ($x_{nm} = A_{nm}$) THEN ($f_n = p_nx + q_ny + r_n$)

where x_i are the inputs, A_{nm} are the fuzzy sets and f_i are the outputs within the fuzzy region specified by the fuzzy rule, p_n. q_n, and r_n are the design parameters that are determined during the training process.

Three versions, based on the training algorithm, of the ANFIS hybrid model were used in this study viz. Grid Partitioning, Subtractive Clustering and Fuzzy C-Means Clustering. The ANFIS with Grid Partitioning (ANFIS-GP) was used to generate a single-output Sugeno-type fuzzy inference system (FIS) using a grid partition on the data (no clustering). The outputs of the adaptive nodes in layer 1 are fuzzy membership grade of the inputs, which are generally given by:

$$O_i^1 = \mu_{Ai}(x) \; i = 1, 2 \qquad (1)$$

and

$$O_i^1 = \mu_{\boldsymbol{B}_{i-2}}(y) \; i = 3, 4 \qquad (2)$$

where $\mu_{\boldsymbol{A}_i}(x)$ and $\mu_{\boldsymbol{B}_{i-2}}(y)$are fuzzy membership functions.

Several input Membership Functions (MFs) were tried with *linear* and *constant* output MFs. All the MFs with *linear* output MF were found to be over-fitting (Figure 4). However, *Gaussian* and *bell-shaped* MFs with constant output MF were found to be highly competitive in performance. A further comparative investigation showed that the *Gaussian* MF is optimal for this problem, especially the porosity data (Figure 5). This agrees with literature [56] that presents the *Gaussian* MF as the best for most applications. The *Gaussian* input MF is given by:

$$\mu_{\boldsymbol{A}_i}(x) = exp\left(-\frac{(c_i - x)^2}{2\sigma_i^2}\right) \qquad (3)$$

where c_i and σ_i^2 are the centre and width of the i^{th} fuzzy set $\boldsymbol{A}_i$ respectively.

The fixed layer 2 nodes serve the role of a simple multiplier. The outputs of this layer, the firing strengths of the rules, can be represented as:

$$O_i^2 = w_i = \mu_{A_i}(x)\mu_{B_i}(y) \quad i = 1, 2 \tag{4}$$

The fixed layer 3 nodes play a normalization role to the firing strengths from the previous layer 2. The outputs of this layer, the normalized firing strengths, can be represented as:

$$O_i^3 = \omega_i = \frac{w_i}{w_1 + w_2} \quad i = 1, 2 \tag{5}$$

In layer 4, the output of each adaptive node in this layer is simply the product of the normalized firing strength and a first order polynomial (for a first order Sugeno model). This is given by:

$$O_i^4 = \omega_i f_i = \omega_i(p_i x + q_i y + r_i) \quad i = 1, 2 \tag{6}$$

In the layer 5, the only one single fixed node performs the summation of all incoming signals. Hence, the overall output of the model is given by:

$$O_i^5 = \sum_{i=1}^{2} \omega_i f_i = \frac{(\sum_{i=1}^{2} \omega_i f_i)}{w_1 + w_2} \tag{7}$$

The ANFIS with Subtractive Clustering (ANFIS-SC) was used to generate a FIS by first applying subtractive clustering on the data. This is accomplished by extracting a set of rules that models the data behavior by first using the *subclust* function to determine the number of rules and antecedent membership functions and then using linear least squares estimation to determine each rule's consequent equations. Different optimal radii were found for the porosity and permeability datasets. For the porosity datasets, 0.7 is optimal (Figure 6, 7, 8) while for the permeability datasets, 1.6 was found to be optimal (Figure 9, 10, 11). It would be noted that Figure 8 shows that ANFIS, similar to Type-2 Fuzzy Logic is also vulnerable to overfitting when handling small datasets [16].

The ANFIS with Fuzzy C-Means Clustering was used to generate a FIS using FCM clustering by extracting a set of rules that models the data behavior. The rule extraction method first uses the *fcm* function to determine the number of rules and membership functions for the antecedents and consequents. Different number of clusters was also found for the porosity and permeability datasets. This is summarized in Table 1 and shown respectively for each well in Figure 12 - 17.

To further ensure a solid basis for fair comparison, the ANFIS models were run in 50 iterations and the average values of the evaluation criteria were recorded.

Table 1. Optimal Number of Clusters for Datasets.

Data Set	Optimal Number of Clusters
Site 1 Well 1	4
Site 1 Well 2	4
Site 1 Well 3	2
Site 2 Well 1	2
Site 2 Well 2	2
Site 2 Well 3	2

C. Model Evaluation Criteria

Similar to the previously published FN-T2F and FN-SVM; and in order to ensure a fair comparison with ANFIS, the performance of the models was evaluated using the correlation coefficient (CC), root mean-squared error (RMSE) and execution time (ET). CC measures the statistical correlation between the predicted and actual values. RMSE is one of the most commonly used error measures of success for numeric prediction as it computes the average of the squared differences between each predicted value and its corresponding actual value. ET is simply the total time taken for a technique to run from the beginning to its end using the CPU time.

IV. Experimental Results and Discussion

A. Experimental Results

The ANFIS-GP model only worked with the porosity datasets and not with the permeability datasets. This is due to the inability of the model to handle such datasets of 12 attributes. However, the results of the comparison of ANFIS-SC and ANFIS-FCM with FN-T2F and FN-SVM are shown in Figure 18 - 23. Since the main focus of any machine learning task is to investigate its generalization capability and in order to reduce the volume of figures presented in this paper, only the comparative results of testing are presented. The comparative result of the CC and RMSE for all the 3 porosity wells are shown in Figure 18 and 19 respectively while their ET comparison for training and testing are shown in Figure 20. Similarly, the comparative result of the CC and RMSE for all the 3 permeability wells are shown in Figure 21 and 22 respectively while their ET comparison for training and testing are shown in Figure 23.

B. Discussion of Results

The results showed that, in terms of CC, the FN-SVM hybrid model outperformed all the other hybrid models with the highest accuracy as shown in Figure 18. In terms of RMSE, FN-SVM has the lowest value. This is conversely equivalent to the CC result (Figure 19). In terms of ET, FN-T2F took the most time for training and testing (Figure 20). The 3 versions of ANFIS equally demonstrated their capabilities by showing competitive performances and at faster speed of execution than FN-T2F and FN-SVM. It would be noted that all the predictions are not perfect since real-life operational field datasets were used.

For permeability, the FN-SVM hybrid model has the highest CC and the lowest RMSE demonstrating its superior predictive capability over the other two models (Figure 21 and 22). The more the correlation between the actual and the predicted target variables, the less the root mean square error is expected to be. Also, in terms of execution time, FN-T2F took the most time for training and testing (Figure 23).

The excellent performance of FN-SVM can be attributed to the reduced dimensionality of the data fed into the SVM block of the hybrid model. It is further due to the role of the least square fitting algorithm of the FN block in the extraction of the most relevant input variables for the training and testing of the SVM block. This ensures that the SVM block used only the best of the input variables and hence is not corrupted by the redundant and irrelevant variables from the original datasets.

The dimensionally-reduced dataset that was used by the SVM block from the output of the FN block also ensured that the data matrix used in the execution of the SVM block is less complex despite than SVM is originally “light-weight”. In addition to the reports that SVM scales relatively well with high dimensional data [8, 10, 15], we further argue that SVM can be improved by reducing the dimension of the input data through best subset selection. This will have the double effect of reducing the training time and further increasing the accuracy of the prediction process.

Based on the result of this study, FN-T2F did not perform as well as FN-SVM. A question could be asked here: Why did T2F not derive as much benefit from the FN block as SVM? We argue that this is due to the peculiar qualities of SVM such as its ability to handle small data and scalability to high-dimensional data which T2F do not possess.

It can be said that despite the superiority of FN-SVM, all the other models are very competitive in their performance, especially with the porosity datasets. Again, we argue that this is partly due to the efficiency of the grid partition, subtractive clustering and Fuzzy C Means algorithms. The “No Free Lunch Theory” [2] still holds true here despite the excellent performance of FN-SVM such that we may not conclude that the FN-SVM hybrid model is absolutely the best. Other hybrid models might perform better or equally good in some other data scenarios such as in the case of permeability datasets where the grid partitioning algorithm of ANFIS did not work.

V. Conclusion and Future Work

A detailed comparative study of Functional Networks-Type-2 Fuzzy Logic, Functional Networks-Support Vector Machines and 3 versions of Adaptive Neuro-Fuzzy Inference System with grid partition, subtractive clustering and fuzzy C-means is presented in this paper. The comparisons were based on the prediction of porosity and permeability of oil and gas reservoirs from 6 different datasets obtained from diverse fields with different lithological and geological formations. The results showed that the FN-SVM hybrid model showed superior performance with the highest correlation coefficients, lowest root mean square errors but taking longer time to execute than the 3 ANFIS algorithms.

Based on the results of this study, we conclude as follows:

- ✓ The superior performance of the FN-SVM hybrid model is due to the dual role of FN to select the best and most relevant input variables; and the consequent reduction in the dimensionality of the data that was used by the SVM block.
- ✓ The subset selection process performed by the FN block contributed to the further improvement in the performance accuracy of the hybrid model while the reduced dimensionality of the input data reduced the time and space complexity of the SVM block, thereby reducing the overall processing time than that of the FN-T2F model.
- ✓ Despite that SVM has proven to be robust with small datasets and scalable with high-dimensional data, it can be further improved by reducing the dimension of the input data through best subset selection.
- ✓ FN-T2F took the most time for training and testing due to the inherent complexity of the gradient descent algorithm of the T2F component [16].
- ✓ ANFIS with grid partitioning performed competitively with the porosity datasets with 5 attributes but could not handle the complexity introduced by the 12 attributes of permeability datasets. This is in line with what has been reported about ANFIS [24, 25].
- ✓ The 3 versions of ANFIS used in this study are equally good and demonstrate competitive capabilities due to the excellent performance of the grid partitioning, subtractive clustering and the fuzzy C-means algorithms.
- ✓ ANFIS, like Type-2 Fuzzy Logic, could not perform well in cases of small datasets and high dimensional data [16].
- ✓ On the average, the outperformance of FN-SVM over all the other models might be marginal. However, a little improvement in the accuracy of the prediction of oil and gas reservoir properties may result in the increased exploration, production and exploitation of more energy and huge increase in the capital base of the oil industry.
- ✓ The main contributions of this study are to prove the potentials of computational intelligence hybrid models in the petroleum industry and to demonstrate that other models, though not optimal, also have their possible positions in solving real industry problems.

Acknowledgment

The authors would like to thank the Universiti Malaysia Sarawak and King Fahd University of Petroleum and Minerals for providing the resources used in the conduct of this study.

References

[1] F. Anifowose, J. Labadin, A. Abdulazeez. “A Hybrid of Functional Networks and Support Vector Machine Models for the Prediction of Petroleum Reservoir Properties". In *Proceedings of 11th International Conference on Hybrid Intelligent Systems*, IEEExplore, pp. 85 - 90, 2011.

[2] D.H. Wolpert, W.G. Macready. "No Free Lunch Theorems for Optimization". *IEEE Transactions on Evolutionary Computation*, 1(1), pp. 67 - 82, 1997.

[3] Schlumberger Oil Field Glossary. “Porosity”. Available online: http://www.glossary.oilfield.slb.com/Display.cfm?Term=porosity. Accessed on April 20, 2011.

[4] Schlumberger Oil Field Glossary. “Permeability”. Available online: http://www.glossary.oilfield.slb.com/Display.cfm?Term=permeabilit. Accessed on April 20, 2011.

[5] R.E. Wyman. "Well logging". in AccessScience, ©McGraw-Hill Companies, 2008, http://www.accessscience.com, Acessed March 16, 2012.

[6] V. Ravi1, Ramanuj Lal, N. Raj Kiran. "Foreign Exchange Rate Prediction using Computational Intelligence Methods". *International Journal of Computer Information Systems and Industrial Management Applications*, 4, pp. 659-670, 2012.

[7] H.M. Goda, H.R. Behrenbruch. “Use of Artificial Intelligence Techniques for Predicting Irreducible Water Saturation in Australian Hydrocarbons Basins”. In *Proceedings Society of Petroleum Engineers Asia Pacific Oil & Gas Conference and Exhibition*, Jakarta, Indonesia, 2007.

[8] J. Taboada, J.M. Matías, C. Ordóñez, P.J. García. “Creating a Quality Map of a Slate Deposit using Support Vector Machines”. *Elsevier Journal of Computational and Applied Mathematics*, 20(4), pp. 84-94, 2007.

[9] X. Chen, Y. Li, R. Harrison, Y.Q. Zhang. “Type-2 Fuzzy Logic Based Classifier Fusion for Support Vector Machines”. *Applied Soft Computing*, 8(3), pp. 1222-1231, 2008.

[10] T. Helmy, F. Anifowose, K. Faisal. “Hybrid Computational Models for the Characterization of Oil and Gas Reservoirs”. *Elsevier International Journal of Expert Systems with Applications*, 37, pp. 5353-5363, 2010.

[11] F. Anifowose, A. Abdulraheem. “Fuzzy Logic-Driven and SVM-Driven Hybrid Computational Intelligence Models Applied to Oil and Gas Reservoir Characterization”. *Journal of Natural Gas Science and Engineering*, 3, pp. 505-517, 2011.

[12] T. Helmy, F. Anifowose. “Hybrid Computational Intelligence Models for Porosity and Permeability Prediction of Petroleum Reservoirs”. *International Journal of Computational Intelligence and Applications*, 9(4), pp. 313–337, 2010.

[13] D. Kaviani, T. D. Bui, J. L. Jensen, C. L. Hanks. “The Application of Artificial Intelligence Neural Networks with Small Data Sets: An Example for Analysis of Fracture Spacing in the Lisbourne Formation, Northeastern Alaska”. *SPE Journal of Reservoir Evaluation and Engineering*, 11(3), pp. 598-605, 2008.

[14] F. Anifowose, A. Abdulraheem. “How Small is a Small Data?,” In *Proceedings of the 2nd Saudi Conference on Oil & Gas Exploration & Production, Saudi Arabia*, 18-20 December, 2010.

[15] X. Peng, Y. Wang. “A Normal Least Squares Support Vector Machine (NLS-SVM) and its Learning Algorithm”. *Neurocomputing*, 72, pp. 3734-3741, 2009.

[16] J. Mendel. “Type-2 Fuzzy Sets: Some Questions and Answers”. *IEEE Connections, Newsletter of the IEEE Neural Networks Society*, 1, pp. 10-13, 2003.

[17] S. Abe. "Fuzzy LP-SVMs for Multiclass Problems". In *Proceedings of the European Symposium on Artificial Neural Networks*, Belgium, pp. 429-434, 2004.

[18] X. Deyi, W. Dave, Y. Tina, R. San. "Permeability Estimation using a Hybrid Genetic Programming and Fuzzy/Neural Inference Approach". In *Proceedings of the Society of Petroleum Engineers Annual Technical Conference and Exhibition*, Dallas, Texas, U.S.A., 2005.

[19] S. Chikhi. "A Fuzzy ART versus Hybrid NN-HMM Methods for Lithology Identification in the Triasic Province". In *Proceedings of the 2nd International Conference on Information and Communication Technologies*, 1, pp. 1884-1887, 2006.

[20] Y. Wang. "Fuzzy Clustering Analysis by using Genetic Algorithm". *ICIC Express Letters, ICIC International*, 2(4), pp. 331-337, 2008.

[21] F. Anifowose, A. Abdulraheem. “A Functional Networks-Type-2 Fuzzy Logic Hybrid Model for the Prediction of Porosity and Permeability of Oil and Gas Reservoirs”. In *Proceedings of the 2nd International Conference on Computational Intelligence, Modeling and Simulation*, IEEEXplore, pp. 193-198, 2010.

[22] S. F. Pratama, A. K. Muda, Y. Choo, N. A. Muda. "A Comparative Study of Feature Selection Methods for Authorship Invarianceness in Writer Identification". *International Journal of Computer Information Systems and Industrial Management Applications*, 4, pp.467-476, 2012.

[23] J.-S. R. Jang. “ANFIS: Adaptive-Network-based Fuzzy Inference Systems”. *IEEE Transactions on Systems, Man and Cybernetics*, 23(3), pp. 665-685, 1993.

[24] C. Ulas, H. Ahmet, E. Sami. "An Adaptive Neuro-Fuzzy Inference System (ANFIS) Model For Wire-EDM". *Expert Systems with Applications*, 36(3), pp. 6135-6139, 2009.

[25] H. Esen, M. Inalli, A.R Sengur, M. Esen. “Predicting Performance of a Ground-Source Heat Pump System using Fuzzy Weighted Pre-Processing-Based ANFIS”. *Building and Environment*, 43(12), pp. 2178-2187, 2008.

[26] J.-S.R. Jang, N. Gulley. “Fuzzy Logic Toolbox". *Reference Manual*, The Mathworks Inc., Natick, MA., 1996.

[27] P.P. Bonissone. “Adaptive Neural Fuzzy Inference Systems (ANFIS): Analysis and Applications”. http://www.rpi.edu/~bonisp/fuzzy-course/Papers-pdf/anfis.rpi04.pdf. Accessed March 16, 2012.

[28] E.I. Amoiralis, P.S. Georgilakis, T.D. Kefalas, M.A. Tsili, A.G. Kladas. “Artificial Intelligence Combined with Hybrid FEM-BE Techniques for Global Transformer Optimization". *IEEE Transactions on Magnetics*, 43(4), pp. 1633-1636, 2007.

[29] A. Tsakonas, G. Dounias, D. Graf Von Keyserlingk, H. Axer. “Evolving Rule-Based Systems in Two Medical Domains using Genetic Programming”. *Journal of Artificial Intelligence in Medicine*, 32(3), pp. 195-216, 2004.

[30] M. Khashei, F. M. Rafiei, M. Bijari, S. R. Hejazi. "A Hybrid Computational Intelligence Model for Foreign Exchange Rate Forecasting". *Journal of Industrial Engineering International*, 7(15), 2011.

[31] S. Bhardwaj, S. Srivastava, J.R.P Gupta, A. Srivastava. "Chaotic Time Series Prediction Using Combination of Hidden Markov Model and Neural Nets". *International Journal of Computer Information Systems and Industrial Management Applications*, 4, pp. 236 -243, 2012.

[32] M. Danziger, F. B. Neto. "A Hybrid Approach for IEEE 802.11 Intrusion Detection Based on AIS, MAS and Naïve Bayes". *International Journal of Computer Information Systems and Industrial Management Applications*, 3, pp. 193-201, 2011.

[33] E. Castillo, J.M. Gutiérrez, A.S. Hadi, B. Lacruz. “Some Applications of Functional Networks in Statistics and Engineering”. *Technometrics*, 43, pp. 10–24, 2001.

[34] J.H. Fang, H.C. Chen. "Fuzzy Modeling and the Prediction of Porosity and Permeability from the Compositional and Textural Attributes of Sandstone". *Journal of Petroleum Geology*, 20(2), pp. 185-204, 1997.

[35] A. Maqsood, C. Adwait. “Using Artificial Intelligence to Predict Permeability from Petrographic Data”. *Computers & Geosciences*, 26, pp. 915 – 925, 2000.

[36] S. Mohsen, A. Morteza, Y.V. Ali. “Design of Neural Networks using Genetic Algorithm for the Permeability Estimation of the Reservoir,” *Journal of Petroleum Science and Engineering*, 59, pp. 97–105, 2007.

[37] M.B. Shahvar, R. Kharrat, R. Mahdavi. “Incorporating Fuzzy Logic and Artificial Neural Networks for Building Hydraulic Unit-Based Model for Permeability Prediction of a Heterogenous carbonate Reservoir”. In *Proceedings of the International Petroleum Technology Conference*, Doha, Qatar. 7-9 December 2009.

[38] H.J. Park, J.S. Lim, U. Roh, J.M. Kang, B.H. Min. “Production-System Optimization of Gas Fields using Hybrid Fuzzy-Genetic Approach”. *SPE Journal*, 15(2), pp. 417- 425, 2010.

[39] L. Ali, S. Bordoloi, S.H. Wardinsky. “Modelling Permeability in Tight Gas Sands Using Intelligence and Innovative Data Mining Techniques”. In *Proceedings of the SPE Annual Technical Conference and Exhibition*, Denver, Colorado, September 2008.

[40] F. Zarei, A. Daliri, N. Alizadeh. “The Use of Neuro-Fuzzy Proxy in Well Placement Optimization”. In *Proceedings of the SPE Intelligent Energy Conference and Exhibition*, Amsterdam, The Netherlands, 25-27 February, 2008.

[41] A. Al-Anazi, I. Gates, J. Azaiez. “Innovative Data-Driven Permeability Prediction in a Heterogeneous Reservoir”. In *Proceedings of the SPE EUROPEC/EAGE Annual Conference and Exhibition*, Amsterdam, The Netherlands, June 2009.

[42] R.R. Bies, M.F. Muldoon, B.G. Pollock, S. Manuck, G. Smith, M.E. Sale. "A Genetic Algorithm-Based, Hybrid Machine Learning Approach to Model Selection". *Journal of Pharmacokinetics and Pharmacodynamics*, 33(2), pp. 195-221, 2006.

[43] J.B. Petrus, F. Thuijsman, A.J. Weijters. *Artificial Neural Networks: An Introduction to ANN Theory and Practice*, Springer, UK, 1995.

[44] C. Rusu, V. Rusu. "Artificial intelligence in theory and practice", in *International Federation for Information Processing*, M. Bramer (Ed.), Springer, Boston, 217, pp. 119-128, 2006.

[45] K. Polat, S. Gunes. “An Intelligent Diagnosis System For Diabetes On Linear Discriminant Analysis And Adaptive Network Based Fuzzy Inference System: LDA-ANFIS”. *Digital Signal Processing*, 17(4), pp: 702-710, 2009.

[46] A.O. Ozkan, S. Kara, A. Salli, M. E. Sakarya, S. Gunes. “Medical Diagnosis of Rheumatoid Arthritis Disease from Right and Left Hand Ulnar Artery Doppler Signals using Adaptive Network Based Fuzzy Inference System (ANFIS) and MUSIC Method”. *Advances in Engineering Software*, 41(12), pp. 1295-1301, 2010.

[47] M. Shibli. "A Novel Approach to Predict Earthquakes using Adaptive Neural Fuzzy Inference System and Conservation of Energy-Angular Momentum". *International Journal of Computer Information Systems and Industrial Management Applications*, 3, pp. 371-390, 2011.

[48] I. De1, J. Sil. "ANFIS Tuned No-Reference Quality Prediction of Distorted/Decompressed Images featuring Wavelet Entropy". *International Journal of Computer Information Systems and Industrial Management Applications*, 3, pp. 298-305, 2011.

[49] R. A. Ghani, A. Mohamed, H. Shareef. “ANFIS Approach for Locating Precise Fault Points with Coordinated Geometries in a Test Distribution System”. *European Journal of Scientific Research*, 35(3), pp.461-473, 2009.

[50] A. Samandar. "A Model of Adaptive Neural-Based Fuzzy Inference System (ANFIS) for Prediction of Friction Coefficient in Open Channel Flow". *Scientific Research and Essays*, 6(5), pp. 1020-1027, 2011.

[51] S. Ravi, M. Sudha, P. A. Balakrishnan. “Design of Intelligent Self-Tuning GA ANFIS Temperature Controller for Plastic Extrusion System”. *Modelling and Simulation in Engineering*, pp. 1-8, 2011.

[52] M.A. Boyacioglu, D. Avci. “An Adaptive Network-Based Fuzzy Inference System (ANFIS) for the prediction of stock market return: The case of the Istanbul Stock Exchange”. *Expert Systems with Applications*, 37(12), pp. 7908-7912, 2010.

[53] H.M. Jiang, C.K. Kwong, W.H. Ip, T.C. Wong. Modeling Customer Satisfaction for New Product Development using a PSO-based ANFIS Approach". *Applied Soft Computing*, 12(2), pp. 726–734, 2012.

[54] P. Tahmasebi, A. Hezarkhani. “Application of Adaptive Neuro-Fuzzy Inference System for Grade Estimation; Case Study, Sarcheshmeh Porphyry Copper Deposit, Kerman, Iran”. *Australian Journal of Basic and Applied Sciences*, 4(3), pp. 408-420, 2010.

[55] T. Weldu, S. Ghedan, O. Al-Farisi. “Hybrid AI and Conventional Empirical Model for Improved Prediction of Log-Drived Permeability of Heterogeneous Carbonate Reservoir”. In *Proceedings of the Society of Petroleum Engineers Production and Operations Conference and Exhibition*, Tunis, Tunisia, 8-10 June, 2010.

[56] C. Quintana, V. Kreinovich, Leon Reznik. "Gaussian Membership Functions are Most Adequate in Representing Uncertainty in Measurements". In *Proceedings North American Fuzzy Information Processing Society (NAFIPS'92) Conference*, 2, pp 618-624, 1992.

Author Biographies

Fatai A. Anifowose is a PhD student in the Faculty of Computer Science and Information Technology, Universiti Malaysia Sarawak. Presently, he works as a Research Engineer in the Center for Petroleum and Minerals, Research Institute, King Fahd University of Petroleum and Minerals (KFUPM), Saudi Arabia. He obtained his Bachelor of Technology (B. Tech.) degree in Computer Science from Federal University of Technology, Akure, Nigeria in 1999 and a Master degree in Information and Computer Science from KFUPM in 2009. His major research interests include the application of hybrids and ensembles of AI techniques in the characterization of oil and gas reservoir properties. He is a member of professional societies such as Society of Petroleum Engineers, type2fuzzylogic.org, UK, International Association of Computer Science and Information Technology and Nigeria Computer Society.

Jane Labadin is currently an Associate Professor at the Faculty of Computer Science and Information Technology, Universiti Malaysia Sarawak (UNIMAS). She received her Ph.D. in Computational Mathematics specializing in Fluid Dynamics from the Imperial College of Science, Technology and Medicine, London, UK in 2002. Her Bachelor degree in Applied Mathematics was from the same university in 1995. Her Master in Computation obtained in 1997 was from the University of Manchester, Institute of Science and Technology, UK. Her research interest is in computational modeling of dynamical systems.

Abdulazeez Abdulraheem is an Associate Professor in the department of Petroleum Engineering, King Fahd University of Petroleum and Minerals, Saudi Arabia. He earned a Ph.D. degree in Civil Engineering with specialization in Rock Mechanics from the University of Oklahoma, Norman, USA in 1994, a Master degree in Civil Engineering from the Indian Institute of Science, Bangalore, India, in 1985, and a Bachelors degree in Civil Engineering from the Osmania University, Hyderabad, India, in 1983. His areas of research interests include theoretical and experimental rock mechanics, constitutive (material behavior) modeling, and numerical simulation using finite element and finite difference, soil mechanics, and structural dynamics.

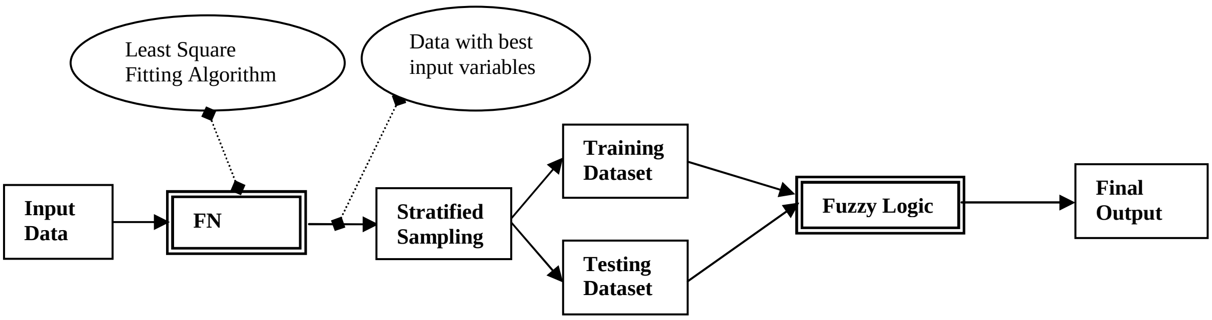

Figure 2. Conceptual Framework of the FN-T2FL Hybrid Model [22]

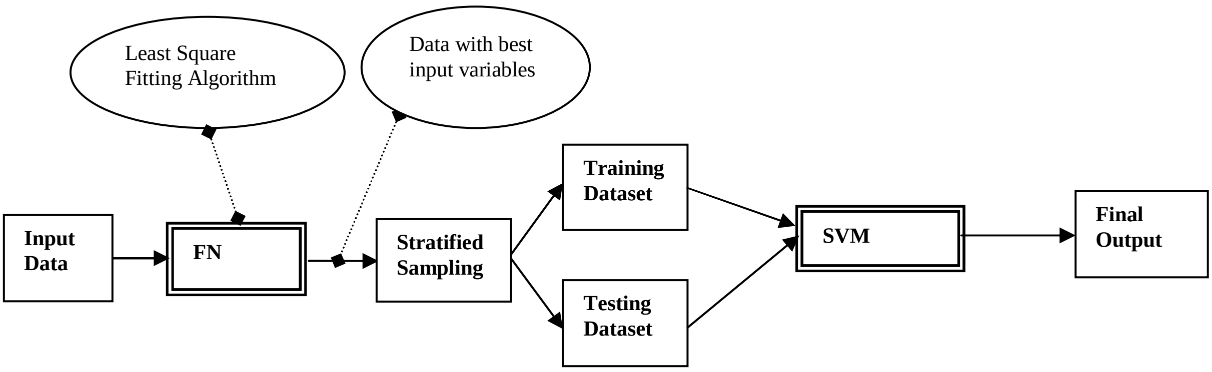

Figure 3. Conceptual Framework of the FN-SVM Hybrid Model [1]

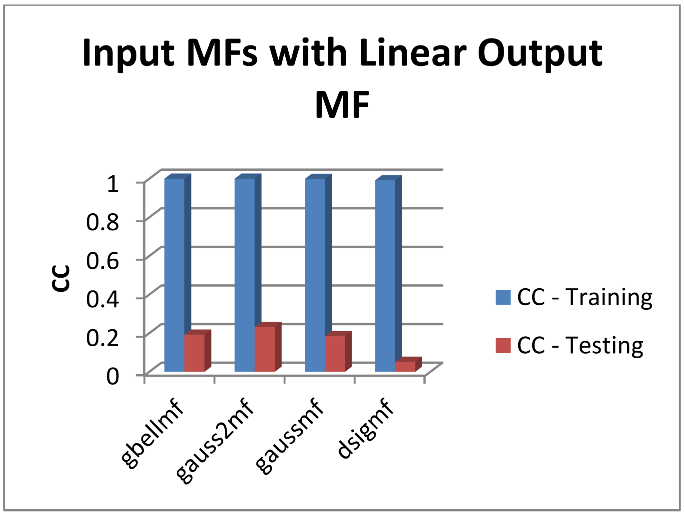

Figure 4. Input MFs with Linear Output MF showing Over-fitting.

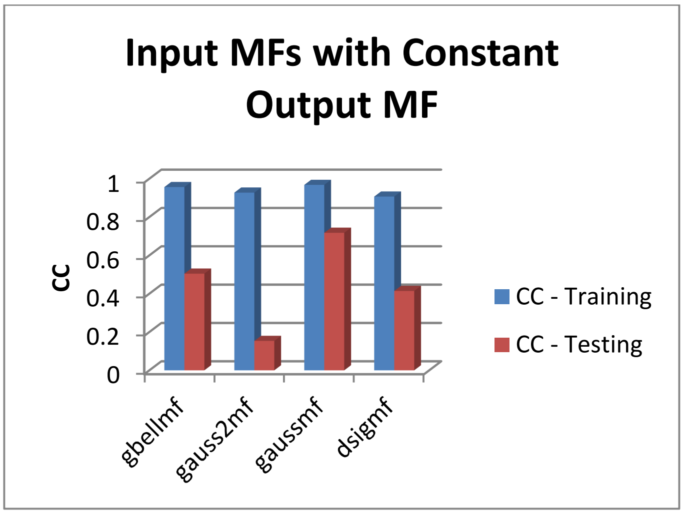

Figure 5. Comparison of Input MFs with Constant Output MF.

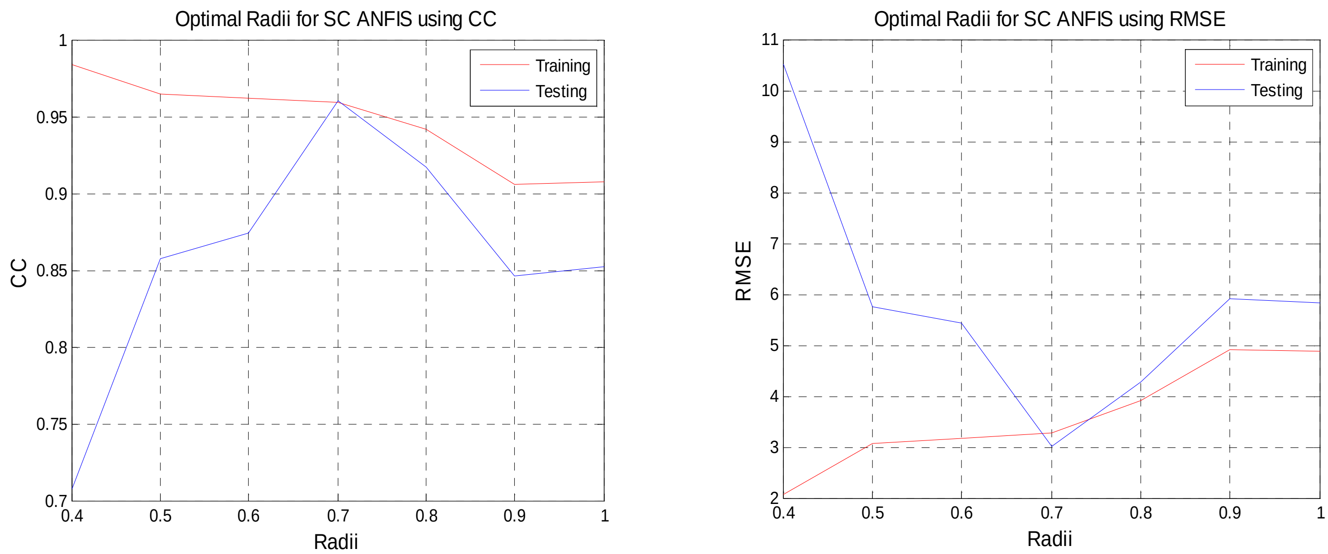

Figure 6. Optimal Radii for ANFIS-SC with Site 1 Well 1 Dataset

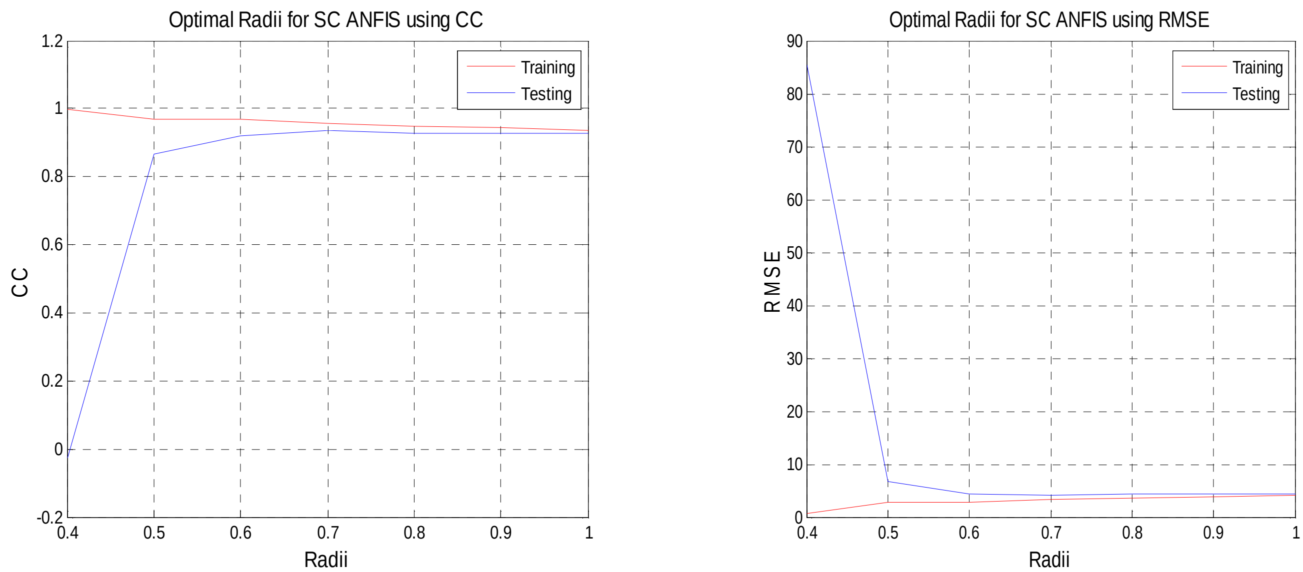

Figure 7. Optimal Radii for ANFIS-SC with Site 1 Well 2 Dataset

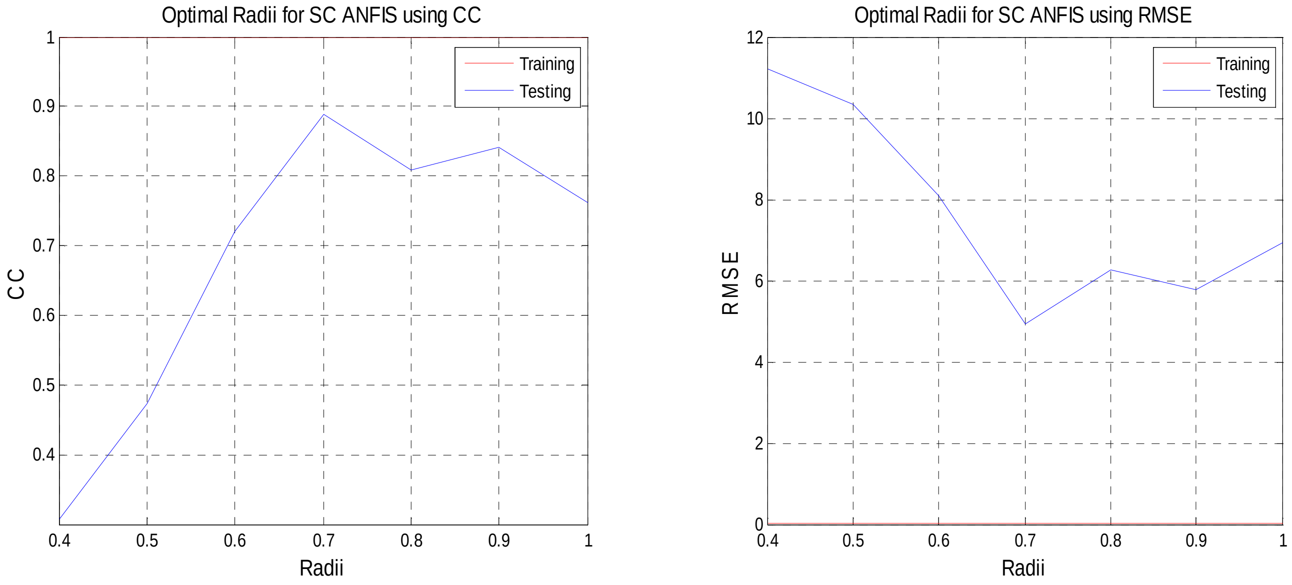

Figure 8. Optimal Radii for ANFIS-SC with Site 1 Well 3 Dataset

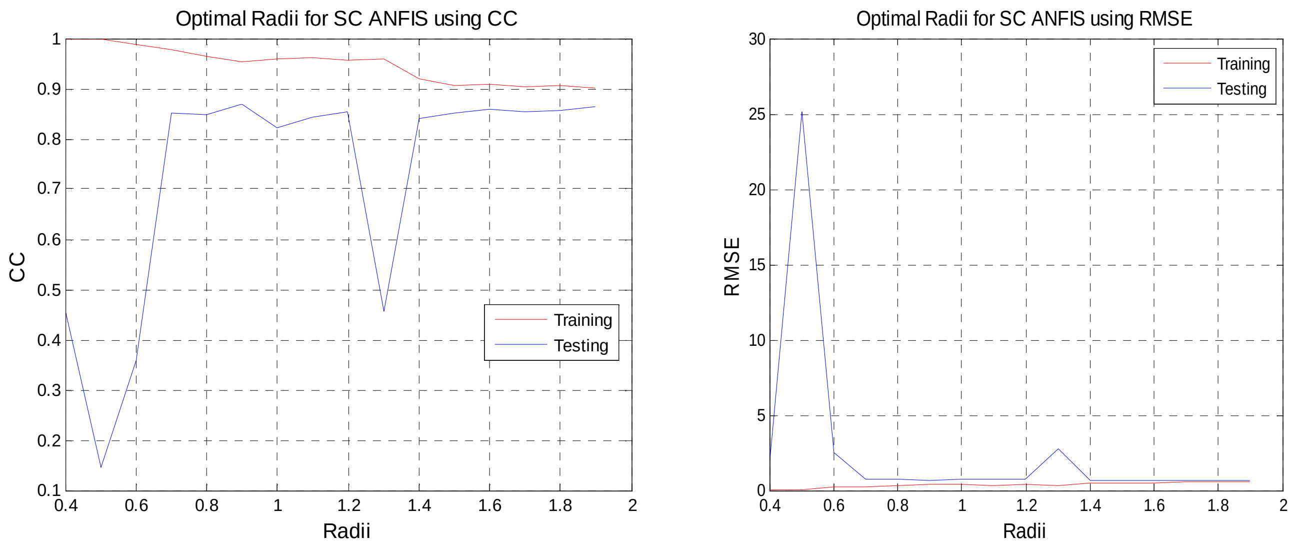

Figure 9. Optimal Radii for ANFIS-SC with Site 2 Well 1 Dataset

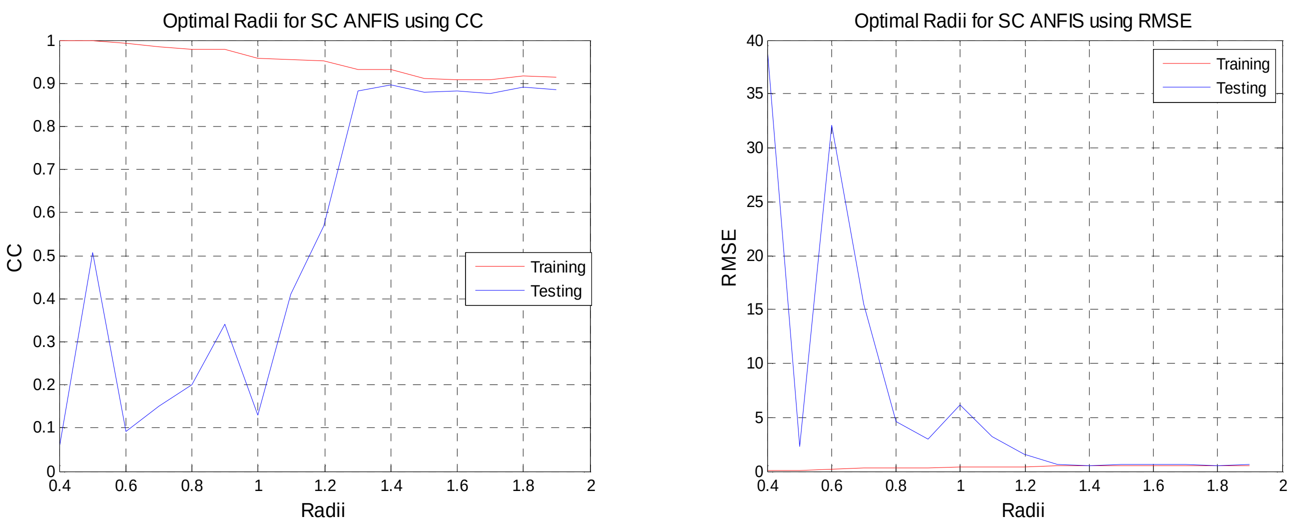

Figure 10. Optimal Radii for ANFIS-SC with Site 2 Well 2 Dataset

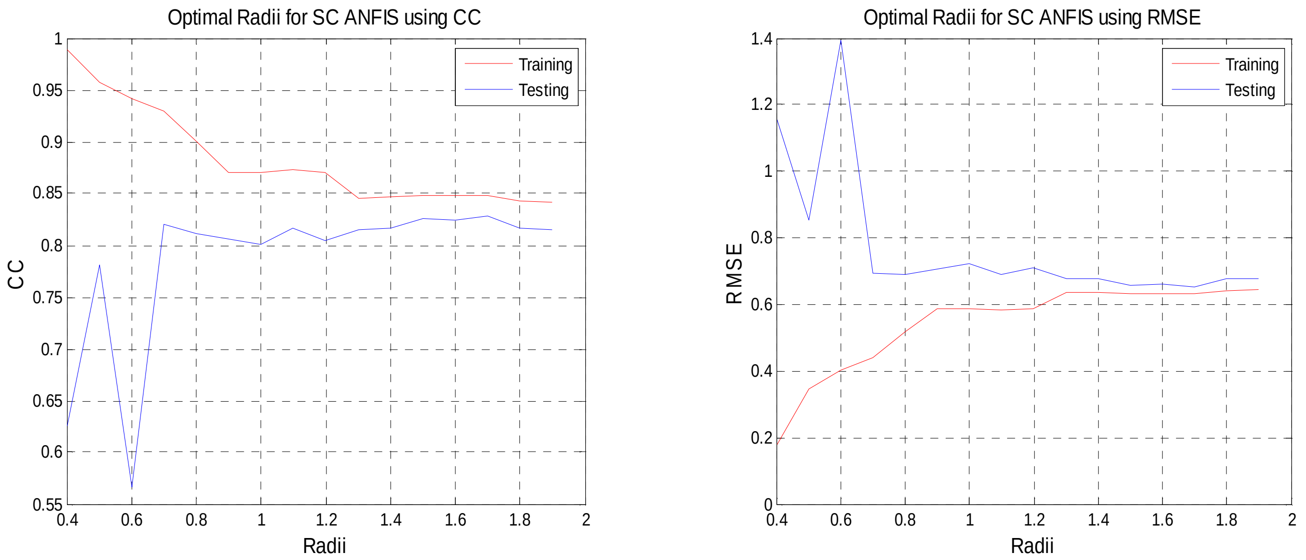

Figure 11. Optimal Radii for ANFIS-SC with Site 2 Well 3 Dataset

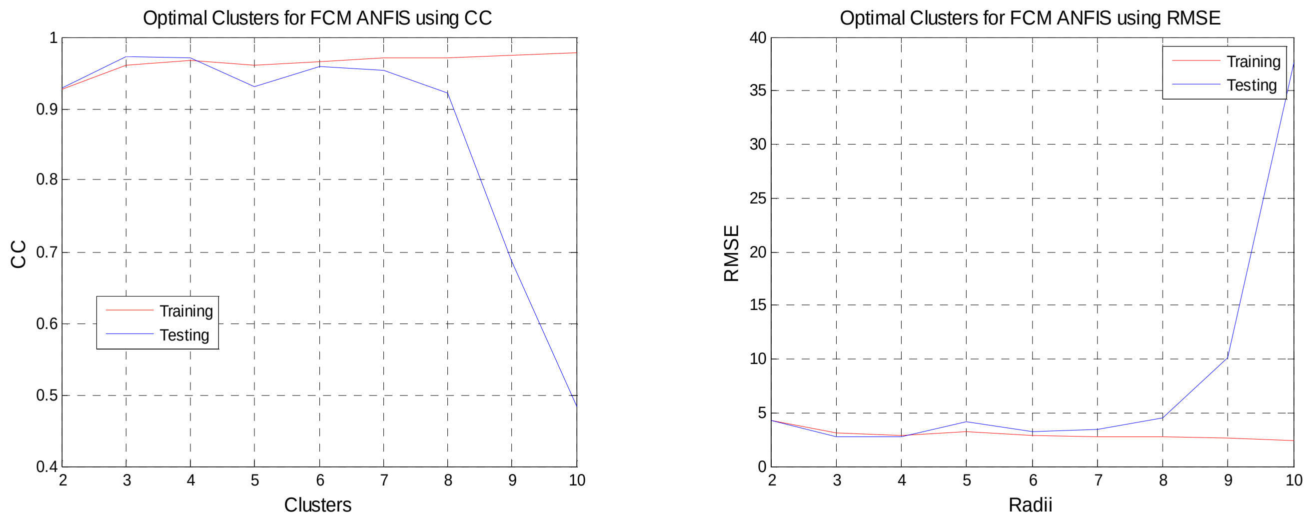

Figure 12. Optimal Clusters for ANFIS-FCM with Site 1 Well 1 Dataset

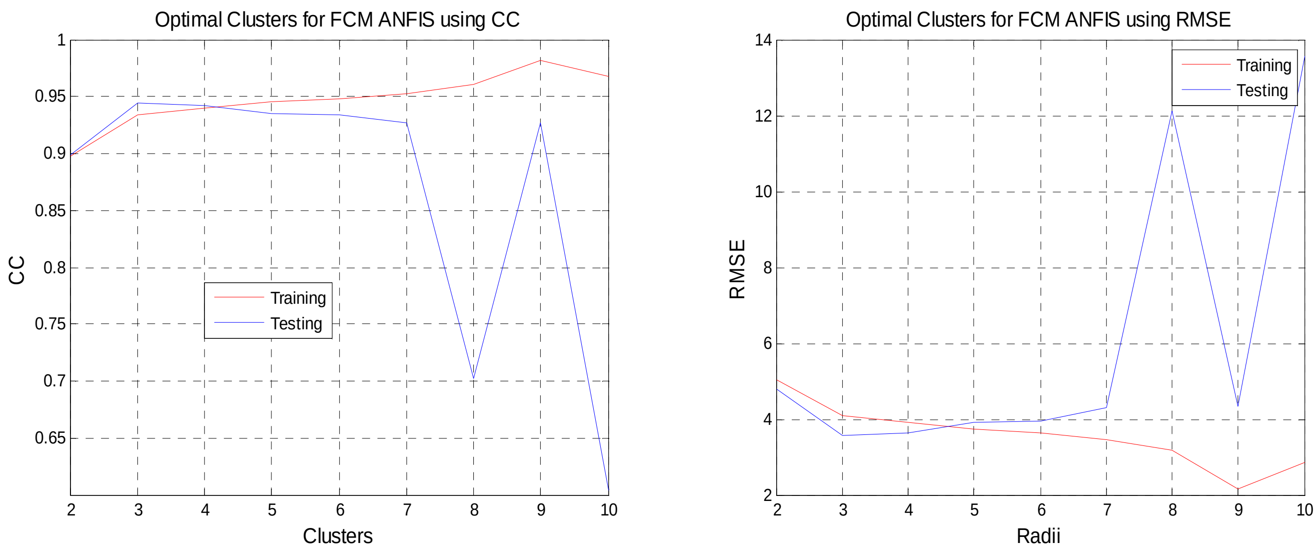

Figure 13. Optimal Clusters for ANFIS-FCM with Site 1 Well 2 Dataset

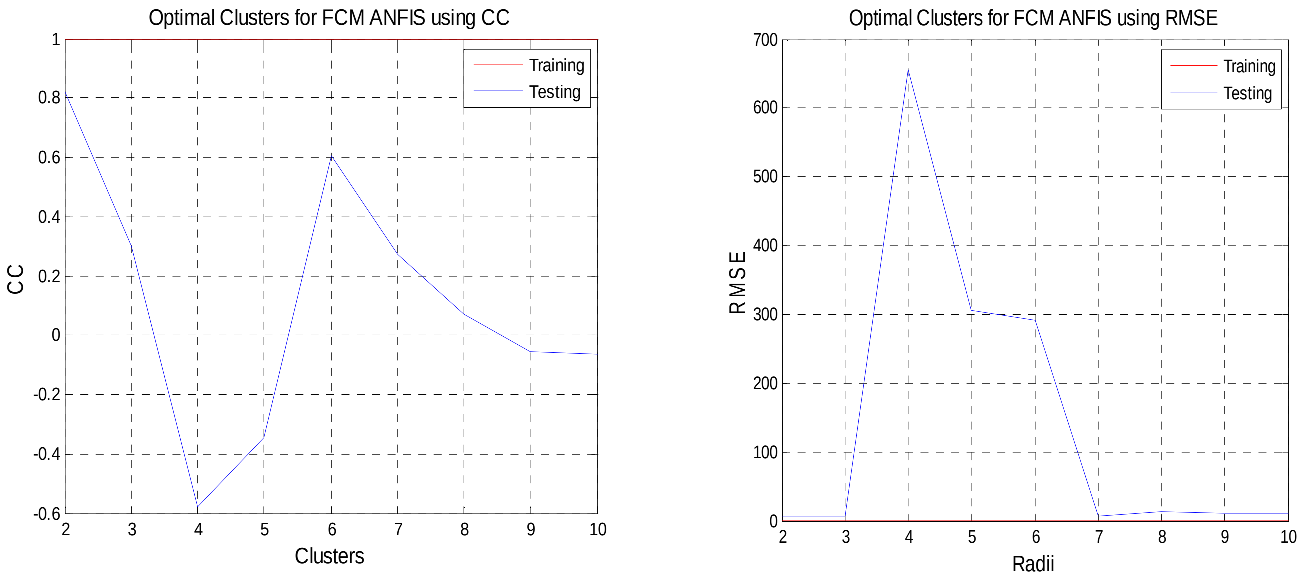

Figure 14. Optimal Clusters for ANFIS-FCM with Site 1 Well 3 Dataset

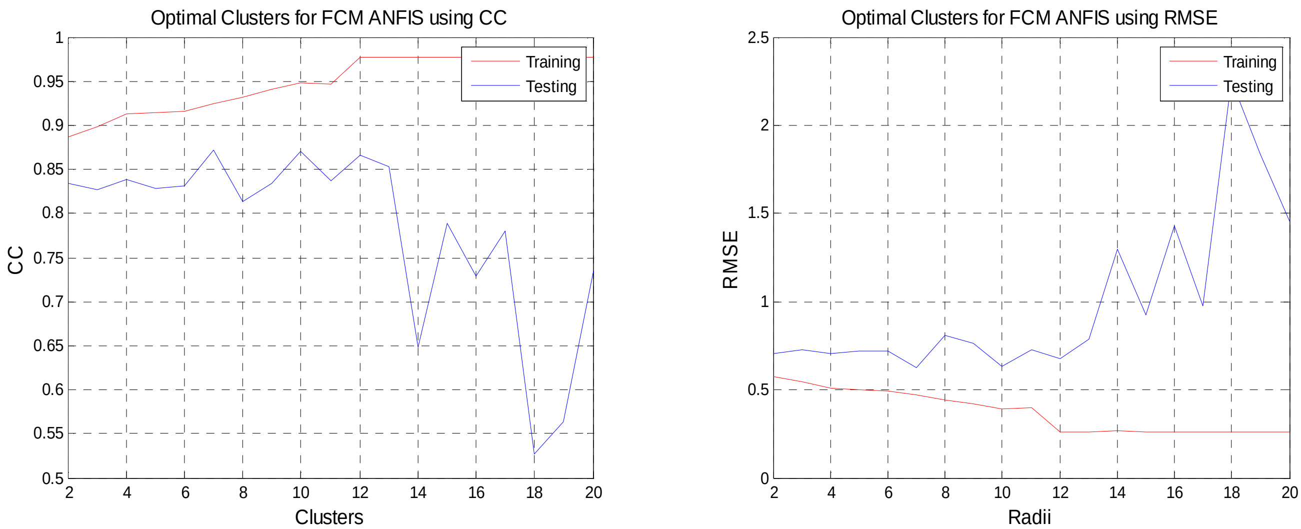

Figure 15. Optimal Clusters for ANFIS-FCM with Site 2 Well 1 Dataset

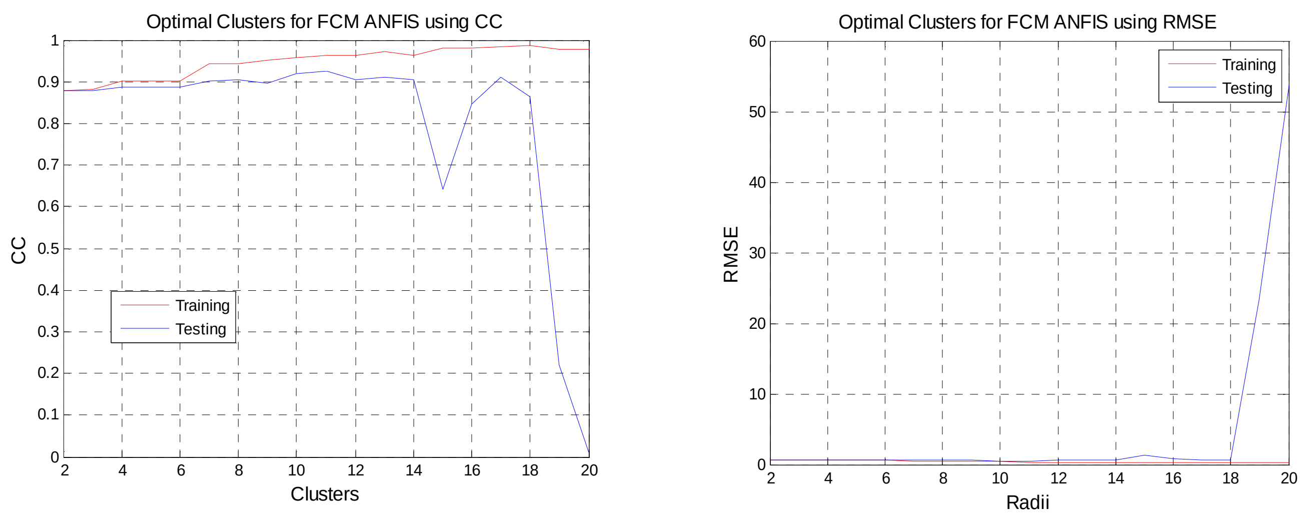

Figure 16. Optimal Clusters for ANFIS-FCM with Site 2 Well 2 Dataset

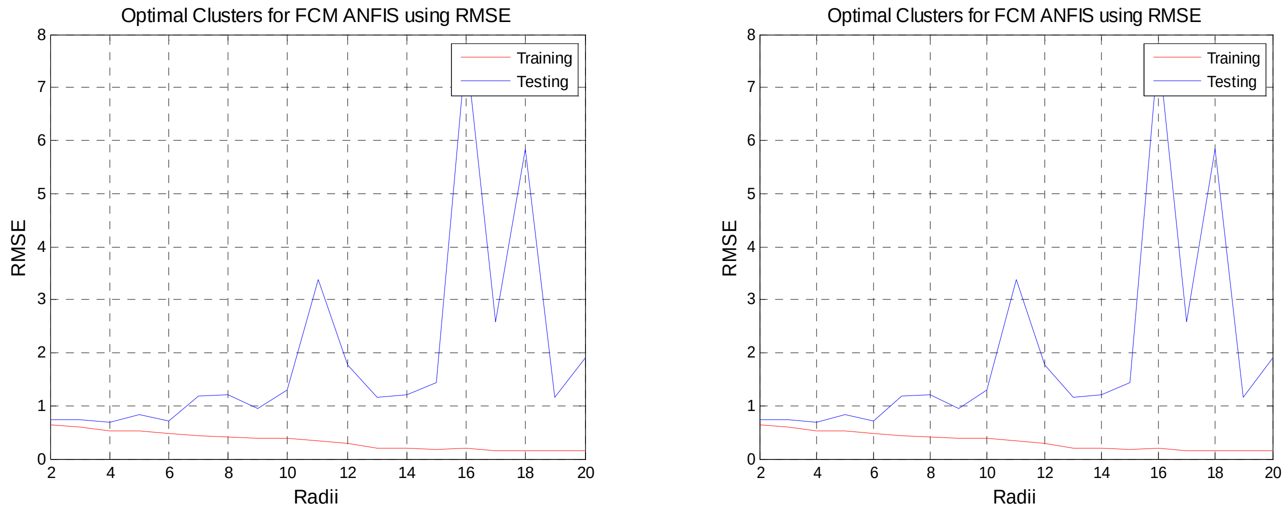

Figure 17. Optimal Clusters for ANFIS-FCM with Site 2 Well 3 Dataset

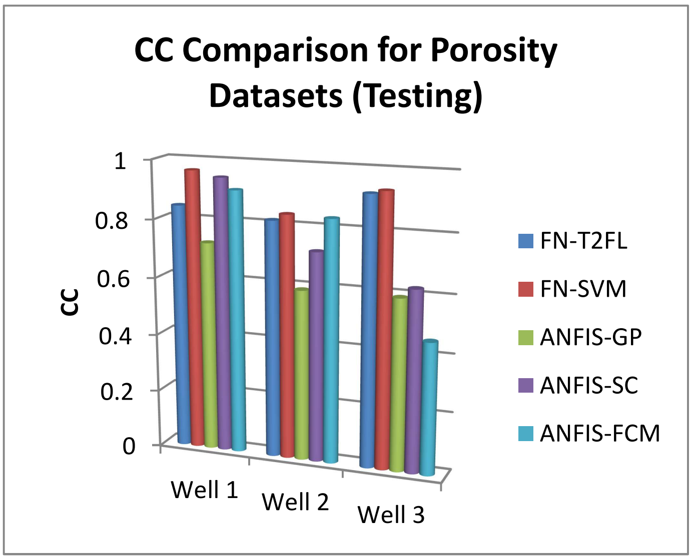

Figure 18. Testing CC Comparison for Porosity Datasets

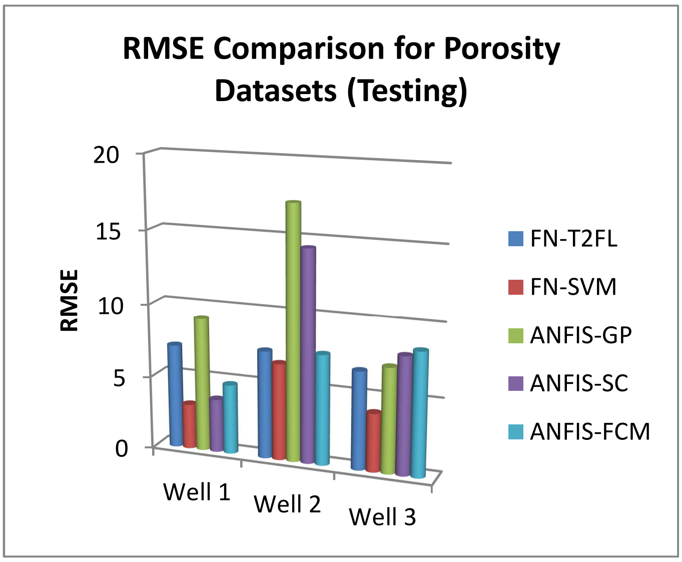

Figure 19. Testing RMSE Comparison for Porosity Datasets

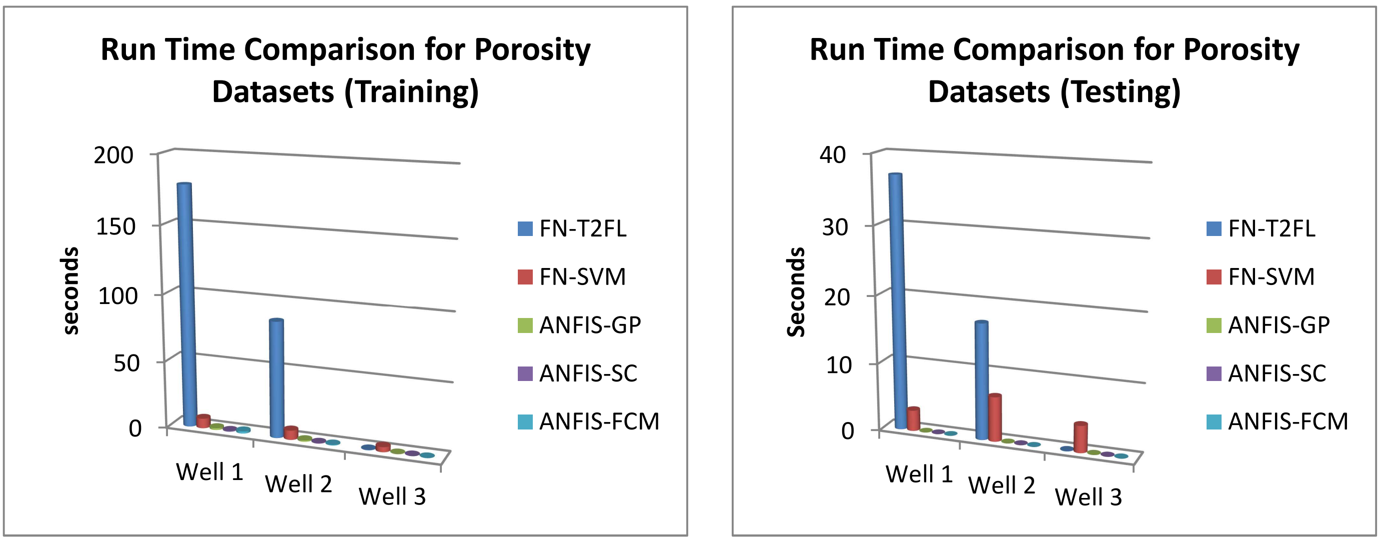

Figure 20. Training and Testing Run Time Comparison for Porosity Datasets

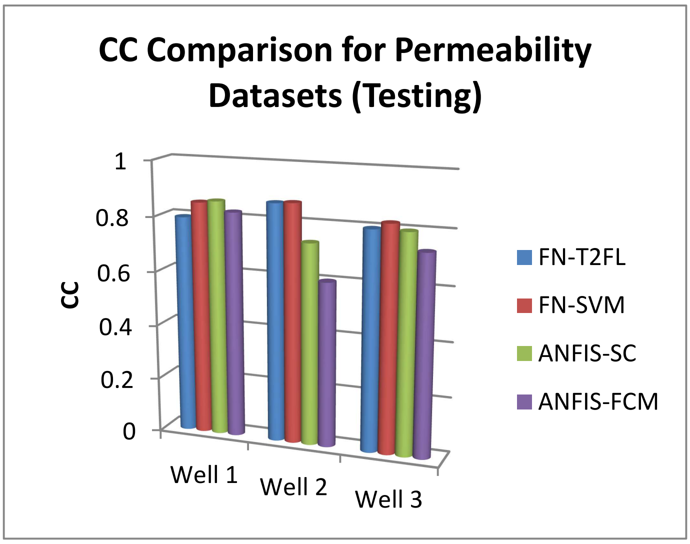

Figure 21. Testing CC Comparison for Permeability Datasets

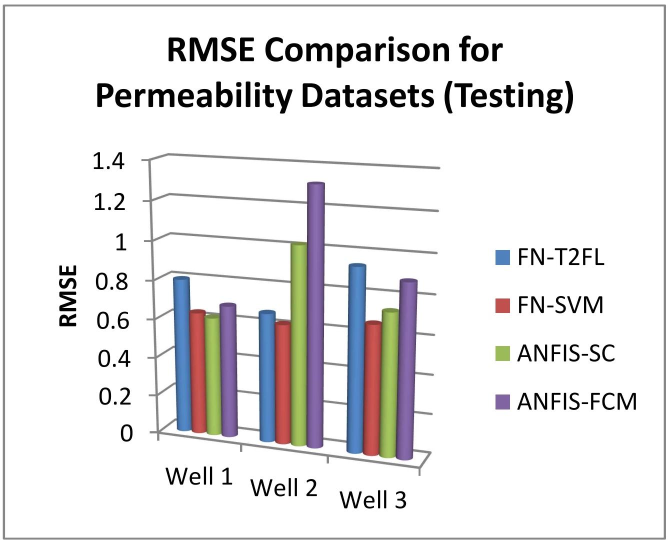

Figure 22. Testing RMSE Comparison for Permeability Datasets

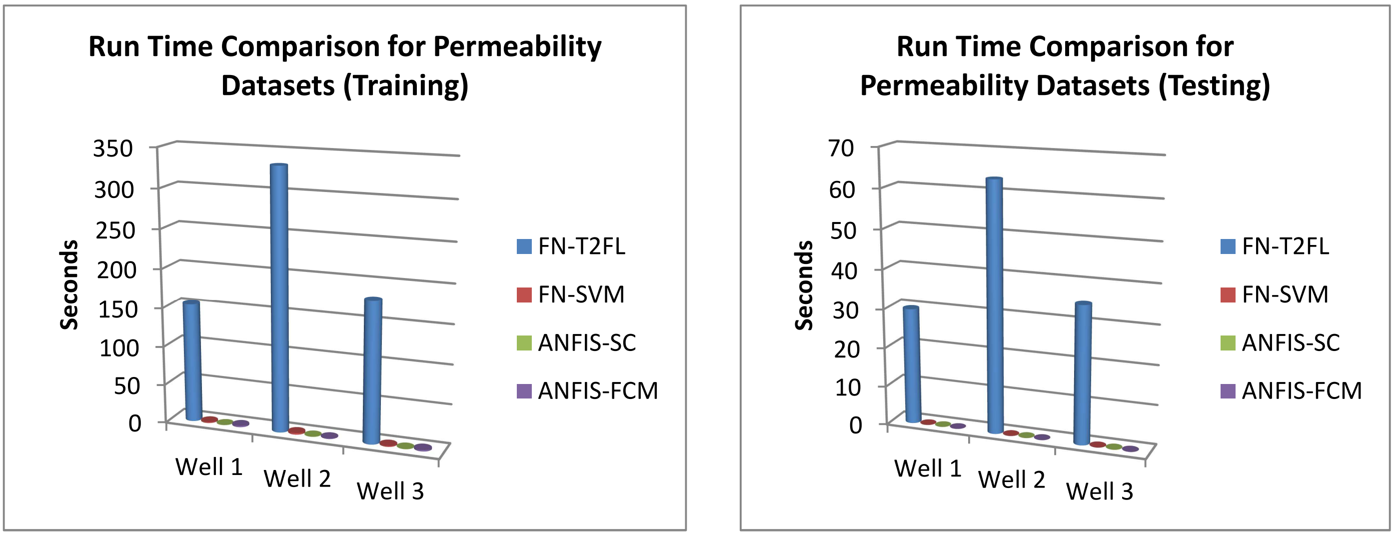

Figure 23. Training and Testing Run Time Comparison for Permeability Datasets