PREDICTION OF FLOW ZONE INDICATORS IN CARBONATE RESERVOIRS USING HYBRID MODELING

BY

Ahmed Yousif Alsahaf

A Thesis Presented to the
DEANSHIP OF GRADUATE STUDIES

KING FAHD UNIVERSITY OF PETROLEUM & MINERALS

DHAHRAN, SAUDI ARABIA

In Partial Fulfillment of the
Requirements for the Degree of

MASTER OF SCIENCE

In

PETROLEUM ENGINEERING

May 2015

KING FAHD UNIVERSITY OF PETROLEUM & MINERALS

DHAHRAN- 31261, SAUDI ARABIA

DEANSHIP OF GRADUATE STUDIES

This thesis, written by Ahmed Yousif Alsahaf under the direction of his thesis advisor and approved by his thesis committee, has been presented and accepted by the Dean of Graduate Studies, in partial fulfillment of the requirements for the degree of **MASTER OF SCIENCE IN PETROLEUM ENGINEERING.**

THESIS COMMITTEE

Dr. Abdulazeez Abdulraheem
(Advisor)

Dr. Abdullah S. Sultan
(Department Chairman)

Dr. Mohamed Mahmoud
(Member)

Dr. Salam A. Zummo
(Dean of Graduate Studies)

Prof. Korvin Gabor
(Member)

10/6/15

Date

DEDICATION

For my parents, lovely wife, and daughters

Without them, this would not be possible

ACKNOWLEDGMENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AI : Artificial Intelligence

ANFIS : Adaptive Neuro Fuzzy Inference System

ANN : Artificial Neural Network

CC : Correlation Coefficient

CT : Computation Time

DT : Decision Trees Algorithm

FFNN : Feedforward Backpropagation Neural Network

FIE : Fuzzy Information Entropy

FL : Fuzzy Logic

FN : Functional Networks Algorithm

FR : Fuzzy Ranking Algorithm

FZI : Flow Zone Indicator

H : Information Entropy

HU : Hydraulic Units

I_E : Information Entropy

I_G : Information Gain

K	:	Permeability
MAE	:	Mean Absolute Error
n	:	Number of data points
NMR	:	Nuclear Magnetic Resonance
RMSE	:	Root Mean-Squared Error
RQI	:	Reservoir Quality Index
SVM	:	Support Vector Machine
T_1	:	Computation start time
T_2	:	Computation end time
T2FL	:	Type-2 Fuzzy Logic
v	:	Weights assigned to features prior to input layer
w	:	Weights between input layer and hidden layers
Φ	:	Porosity
$φ_e$	:	Effective Porosity
$φ_z$	:	Pore Volume to Solid Volume Ratio
y	:	Target Values
$\hat{y}$	:	Predicted target values

F_s : Shape factor

τ : Tortuosity

S_{gv} : Surface area per unit grain volume, μm^{-1}

ABSTRACT

Full Name : Ahmed Yousif Alsahaf

Thesis Title : [PREDICTION OF FLOW ZONE INDICATORS IN CARBONATE RESERVOIRS USING HYBRID MODELING]

Major Field : Petroleum Engineering

Date of Degree : May, 2015

The oil and gas industry has an increasing volume of geological and petrophysical data that can be utilized to enhance reservoirs simulation models. Due to their inherent spatial heterogeneities, the description of carbonate reservoirs poses great challenges. The most accurate way for evaluating their critical parameters such as permeability is through coring and laboratory experiments, but this process requires extensive labor and high costs.

This study presents the utilization of hybrid Artificial Intelligence (AI) to address heterogeneity challenges and predict flow zone indicator (FZI) in carbonate reservoirs. The proposed model will use eight conventional and 20 nuclear magnetic resonance (NMR) logs data as input. The input datasets have 487 data points for each log. First, five non-hybrid systems were developed using different AI techniques. The five AI techniques are Feedforward Neural Networks (FFNN), Radial Basis Function Neural Networks (RBFNN), Generalized Regression Neural Networks (GRNN), Support Vector Machines (SVM), and Type-2 Fuzzy Logic (T2FL). Next, SVM, T2FL, and the best performer of the three neural networks techniques, FFNN, were considered for hybridization. Nine hybrid systems were then developed, and the one with the best

performance were selected. All the hybrid systems consisted of two stages in sequence. The first stage acts as feature selection from input data. Three different algorithms were used and compared for performance, namely Functional Networks, Decision Tree, and Fuzzy Information Entropy. The second stage employed the three selected techniques, which used the selected parameters, to predict the target values for flow zone indicator. All nine developed hybrid systems were benchmarked with their respective non-hybrid models. Fusion of additional inputs was utilized to further improve the FZI prediction models. To select the best model, the performance of all developed models was compared using both statistical error and graphical analysis. Finally, super hybrid prediction models for permeability were built utilizing predicted values of FZI.

Results showed that flow zone indicator prediction models can be further improved by hybridization and Data Fusion. This improvement in FZI prediction also led to improvement in permeability prediction. This work provides more accurate prediction models for flow zone indicators and permeability in carbonate reservoirs.

<div dir="rtl">

ملخص الرسالة

الاسم الكامل: أحمد يوسف الصحاف

عنوان الرسالة: التنبؤ بمؤشرات مناطق التدفق في المكامن الكربونيه بواسطة نماذج الذكاء الاصطناعي المهجنة

التخصص: هندسة بترول

تاريخ الدرجة العلمية: 2015

مع التزايد المستمر في حجم البيانات الجيولوجية والبتروفيزيائية المتاحة في صناعة النفط والغاز ، أصبحت الحاجة المستمرة لتعزيز نماذج المحاكاة للمكامن بارزة. نظرا لطبيعتها المكانية المتغيرة ، تواجه عملية المحاكاة للمكامن الكربونيه تحديات كبيرة. الطرق الأكثر دقة لتقييم المعلمات الحرجة مثل النفاذية تتم من خلال الحفر ومختبر التجارب، ولكن هذه الطرق تتطلب العمل المكثف وارتفاع التكاليف.

تقدم هذه الدراسة استخدام الهجين في الذكاء الاصطناعي لمعالجة هذه التحديات والتنبؤ بمؤشر مناطق التدفق في المكامن الكربونيه. النموذج المقترح سوف يستخدم سجلات تقليدية و سجلات للرنين المغناطيسي النووي كبيانات إدخال. عدد البيانات لكل سجل تتكون من 487 نقطة. أولا، سيتم وضع خمسة أنظمة غير مهجنه باستخدام تقنيات ذكاء اصطناعي مختلفة. التقنيات الخمسة المستخدمة هي: الشبكات العصبية مسبقة التغذية ، الشبكات العصبية ذات الوظيفة شعاعية الأساس ، الشبكات العصبية ذات الانحدار العام ، آليات المتجهات المدعمة ، ونوع-2 للمنطق الضبابي . بعد ذلك، سيتم النظر في آليات المتجهات المدعمة ، نوع-2 للمنطق الضبابي، والنموذج صاحب أفضل أداء من تقنيات الشبكات العصبية الثلاثة الغير مهجنه. سيتم بعد ذلك وضع تسعة أنظمة هجينة، وسيتم اختيار النموذج ذو الاداء الأفضل. جميع النظم الهجينة ستتكون من مرحلتين في التسلسل. ستبدأ المرحلة الأولى بإختيار السجلات المميزة من بيانات الادخال عبر استخدام ثلاثة خوارزميات مختلفة هي: الشبكات الوظيفية ، شجرة القرار ، ومعلومات ضبابية الانتروبيا . ستبدأ المرحلة الثانية بتوظيف التقنيات الثلاث المختارة ، والتي تستخدم المعلمات المحددة، للتنبؤ القيم المستهدفة لمؤشر منطقة التدفق. وسوف يتم إجراء مقارنة معيارية للتسعة الأنظمة الهجينة المتقدمة مع نماذجها غير الهجينه. ايضاو سيتم استخدام مدخلات إضافية لزيادة تحسين نماذج التنبؤ .لتحديد أفضل نموذج، سيتم مقارنة أداء جميع النماذج المطورة باستخدام كل من الخطأ الإحصائي والتحليل البياني. وأخيرا، سيتم بناء نموذج التنبؤ الهجين المتطور للتنبؤ بالنفاذيه وبإستخدام القيم المتوقعة لمؤشر مناطق التدفق.

أظهرت النتائج أن نماذج التنبؤ بمؤشر مناطق التدفق يمكن زيادة تحسينها بواسطة التهجين و إضافة مدخلات البيانات الإضافيه. أدى هذا التحسن في التنبؤ بمؤشر مناطق التدفق إلى التحسن في التنبؤ بالنفاذية. يوفر هذا العمل نماذج تنبؤ أكثر دقة لوصف للمكامن الكربونيه.

</div>

CHAPTER 1

INTRODUCTION

1.0   Introduction

Reservoir characterization is a process for determining reservoir regionalized properties through the integration of available data. These properties vary spatially from one region to another and include porosity, permeability, hydrocarbon content, water saturation, lithology, and thickness. Available data utilized for this process can be geological information, core data, well-log data, seismic data, and engineering data. Each data source differs in acquisition tools, resolution, reservoir coverage, and in whether they provide direct or indirect measurements.

Knowledge of the rock characteristics for hydrocarbon-bearing reservoirs is vital for the successful initial assessment and development of a field. Porosity and permeability are the two most important properties in reservoir characterization and simulation. Porosity is defined as the fraction of void space in a rock volume. It governs the capacity of a rock to store water or hydrocarbons. Permeability is defined as the capacity of the rock to pass through fluids. It determines how much hydrocarbon can be recovered (recovery factor) and assists in well placement optimization. Porosity can be obtained from well logs but proper evaluation of permeability has always been a challenge, especially in carbonate reservoirs, due to the lack of direct evaluation. The most accurate measurements for both

properties are obtained in labs using cores. This process is labor-intensive, time-consuming, and costly.

In recent decades, Artificial Intelligence (AI) techniques have gained considerable attention in the oil and gas industry. This popularity is attributed to their capabilities in providing solutions to non-conventional problems. Their common utilization falls mostly in the area of petrophysical and geological parameters prediction. This work aims at providing a better prediction model for the Flow Zone Indicator (FZI) in carbonate reservoirs using several AI techniques.

The report is composed of eight chapters. Chapter 2 focuses on the literature review that briefly describes the concepts of hydraulic rock typing. Chapter 3 presents the basic concepts of AI techniques and their use in the oil and gas industry. In Chapter 4, the development of non-hybrid prediction models for FZI with their results is presented. Performance improvement of developed models through hybridization of AI systems and fusion of data are discussed in Chapter 5 and 6 respectively. Chapter 7 provides results and analyses permeability prediction improvements as a result of improved FZI prediction. Lastly, the conclusions and recommendations are provided in Chapter 8.

1.1   Statement of the Problem

In carbonate reservoirs, it is always a challenging task to evaluate properties that best characterize the rocks. This difficulty is attributed to the various depositional and digenetic controls over different parameters including reservoir porosity, permeability, and pore geometry. The inherent heterogeneity makes it impossible to correlate porosity

and permeability in these rocks. AI techniques and algorithms provide very useful tools in addressing such non-linear problems.

1.2　Study Objectives

There are five objectives for this study:

- Predict flow zone indicators in carbonate reservoirs using non-hybrid models.
- Utilize feature selection algorithms that produce a best subset, of all the available features in the acquired data, which will best predict flow zone indicators.
- Use selected subsets to improve the predictions of the flow zone indicator. The second and third objectives will result in developing hybrid models.
- Enhance the hybrid models with the fusion of extra data.
- Predict permeability utilizing predicted FZI values as well as other input data.

1.3　Proposed Approach

In total, 14 artificial intelligence models were developed in this study to predict flow zone indicator in carbonate reservoirs using MATLAB software. Nine of them were hybrid. Core and log data were collected from several wells in a giant carbonate reservoir in the Middle East. Log data consisting of both conventional and nuclear magnetic resonance (NMR) logs were used as input while measured core data, flow zone indicator and permeability, were used as target. All hybrid developed models used feature selection algorithms in their first stage and utilize selected features to build the prediction model in the second stage using Feedforward Neural Networks (FFNN), Support Vector Machines

(SVM), and Type-2 Fuzzy Logic (T2FL). Each system will use a different algorithm viz. Functional Networks (FN), Decision Tree (DT), and Fuzzy Information Entropy (FIE). Fusion of additional input data was utilized to further enhance the hybrid models. Five non-hybrid AI models were developed for performance comparison to determine whether any of the feature selection algorithms improves the prediction model using performance measurements, correlation coefficient (CC), root mean-squared error (RMSE), and mean absolute error (MAE). Lastly, super hybrid prediction models for permeability were developed utilizing predicted FZI values as well as other input data.

CHAPTER 2

LITERATURE REVIEW

2.1   Introduction

Reservoir description is very important in understanding hydrocarbon reservoirs. Advances in porosity and permeability predictions contribute significantly to improving reservoir simulations. Carbonate reservoirs are inherently heterogeneous. Rock typing helps define a hydrocarbon-bearing formation as a limited set of volumes or elements, each having unique characteristics. There are three rock typing categories: depositional, petrographic and hydraulic. This chapter discusses the main concepts and equations for two hydraulic rock typing methods in reservoir description.

2.2   Hydraulic Flow Units

Amaefule (1993) defined the hydraulic unit (HU) as the representative elementary volume (REV) of total reservoir rock within which geological and petrophysical properties that affect fluid flow are internally consistent and predictably different from properties of other rock volumes. They are defined by geological attributes of texture, mineralogy, sedimentary structures, bedding contacts and nature of permeability barriers and by petrophysical properties of porosity, permeability and capillary pressure. Each hydraulic unit has distinctive fluid-flow characteristics. Davis *et al.* (1996) defined hydraulic units as a mappable interval of rock with the following characterization: (1) Sufficient thickness and areal extent to be recognized on logs and mapped across the field

(2) Similar averages of rock properties that influence fluid flow (3) All fluids in hydrodynamic communication. Hearn *et al.* (1984) defined hydraulic units as a reservoir zone that is laterally and vertically continuous, and has similar permeability, porosity, and bedding characteristics. Bear (1972) defined the hydraulic unit as the representative elementary volume of the total reservoir rock within which the geological and petrophysical properties of the rock volume are the same. Ebank (1987) defined it as a mappable portion of the reservoir within which the geological and petrophysical properties that affect the flow of fluid are consistent and predictably different from the properties of other reservoir rock volume. Gunter *et al.* (1997) defined the flow unit as stratigraphically continuous interval of similar reservoir process that honors the geologic framework and maintains the characteristics of the rock type.

Slatt and Hopkins (1988) stated that the flow unit model provides the most complete reservoir description since it allows for the interpretation of many of the geological and petrophysical properties into the reservoir description leading to improved recovery and reservoir management. Hydraulic units provide superior porosity-permeability transforms, compared to facies, geologic zones, Winland's technique, and pore geometry transforms, to determine permeability from well log-derived porosity in uncored oil and gas wells (Shenawi, 2009).

In the petroleum literature, there are a number of proposed techniques for the identification of flow units. Testerman (1962) proposed a statistical reservoir zonation technique to identify naturally occurring zones in a reservoir. This technique starts with dividing the reservoir into two zones and then subdividing these zones until they have minimum internal permeability variation and maximum variation between the zones,

externally. Cant (1984) proposed arbitrarily slicing the sedimentary intervals but these arbitrary slices may cut across depositional units and facies. Other methods such as graphical probability, nonlinear regression, Ward's analytical algorithm, and a Bayesian-based probabilistic approach were covered by Abbaszadeh *et al.* (1996). Amaefule described an analytical technique for identifying flow units as follows: If a bundle of straight capillary tubes were used to simulate a porous medium, rock permeability can be obtained by combining both Darcy's law for flow in porous media and Poiseuille's law for flow in tubes:

$$K = \frac{r^2}{8} \varphi_e \tag{2.1}$$

where K is the permeability, r is the tube radius, and φ_e is the effective porosity. Equation 2.1 shows that a pore characteristic of sedimentary rock, pore radius, is an important factor that relates permeability to porosity. Kozeny-Carmen added a tortuosity factor and used mean pore radius in their generalized equation:

$$K = \frac{\varphi_e^3}{(1 - \varphi_e)^2} \frac{1}{F_s \tau^2 S_{gv}^2} \tag{2.2}$$

where, K is the permeability in μm^2 and S_{gv} is specific surface in μm^{-1}. F_s is the shape factor and it is 2 for circular cylinder. Kozeny constant is defined as $F_s \tau^2$. The geological aspect of the HU is captured in the $F_s \tau^2 S_{gv}^2$ term. The ability to determine or discriminate this term is the central part of the HU classification technique. Amaefule *et al.* (1993) developed this technique to identify and characterize HUs. It was based on both the Kozeny-Carmen modified equation and the mean hydraulic radius.

$$0.0314 \sqrt{K/\varphi_e} = \frac{\varphi_e}{1 - \varphi_e} \frac{1}{\sqrt{F_s}\tau S_{gv}} \tag{2.3}$$

$$FZI = \frac{1}{\sqrt{F_s}\tau S_{gv}} \tag{2.4}$$

The equation indicates that for any hydraulic unit, a log-log plot of a Reservoir Quality Index (RQI) versus the pore volume to solid volume ratio (φ_z) should yield a straight line with a unit slope.

$$RQI = 0.0314 \sqrt{K/\varphi_e} \tag{2.5}$$

$$\varphi_z = \frac{\varphi_e}{(1 - \varphi_e)} \tag{2.6}$$

where, K is in millidarcies (0.0314 is the conversion factor from μm^2 to md) and φ is the effective porosity in fractions. Flow Zone Indicator (FZI) is the intercept of the slope unit line (45°) with φ_z=1, which is a unique parameter for each hydraulic unit, Figure .

$$FZI = \frac{RQI}{\varphi_z} \tag{2.7}$$

Knowledge of both porosity and FZI leads to the possibility of evaluating permeability for that HU using Equation 2.8.

$$K = 1014 \, (FZI^2)\varphi_R \tag{2.8}$$

where,

$$\varphi_R = \frac{\varphi_e^3}{(1 - \varphi_e)^2} \tag{2.9}$$
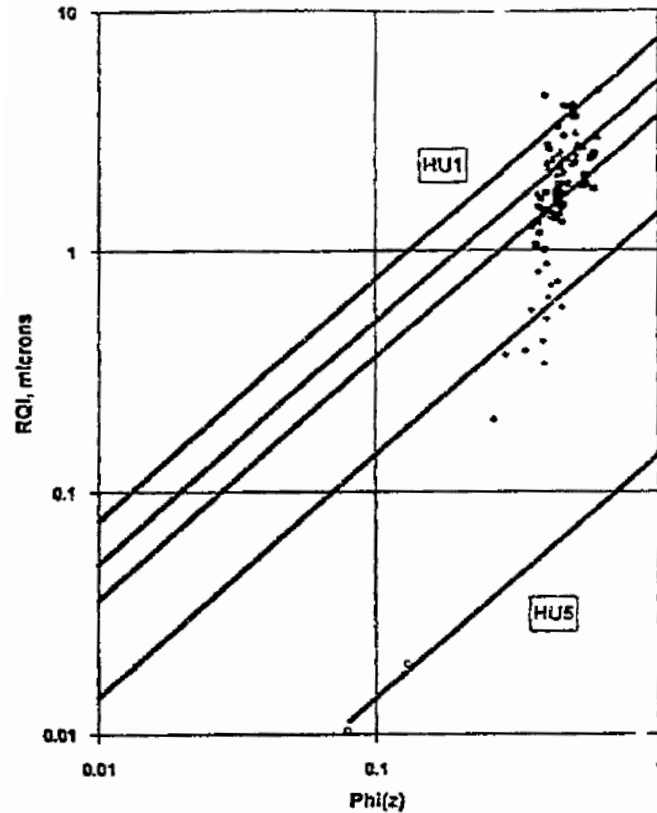
Figure 2.1 - Plot of RQI vs. Porosity, FZI is the Y-value when Porosity Φ=1 (Source: Amaefule et al. (1993))

Flow zonation is independent of depth ordering such that volumes of similar hydraulic properties can appear far apart (Kazeem Lawal, 2005). Desouky (2003) used the hydraulic flow unit (HFU) approach for the identification of six rock types. To achieve this, he used statistical analysis on core data. HFUs helped him to developed better permeability correlation. Desouky also reported that applying the pore level heterogeneity index concept proved the existence of the identified six flow units. Syed Shujath *et al.* (2013) used Support Vector Machines (SVM), Functional Networks, and Adaptive Network Fuzzy Inference Systems (ANFIS) to estimate hydraulic units from 5 conventional well logs. SVM outperformed the other two. They concluded that predicting HUs directly from logs yielded better classification results than estimating it from

9

predicted porosity and permeability. Cumulative prediction error of the porosity and permeability was suspected to be the reason. Guangming Ti *et al.* (1995) demonstrated the application of using flow units for reservoir description. They proposed the use of transmissibility, storability, and the net-to-gross ratio as quantitative indicators for the identification of flow units. Peralta (2013) also demonstrated the application of flow units as one of the rock typing methods in both static and dynamic reservoir modelling. He provided a workflow in utilizing identified flow units in the upscaling process, a part in building a reservoir simulation model. Haghighi *et al.* (2011) improved the permeability prediction in an Iranian tight gas reservoir using hydraulic flow unit based statistical models. In their work, they converted the continuous FZI values to discrete rock typing (DRT) using Equation 2.10.

$$DRT = Round\ [2\ \ln(FZI) + 10.6] \qquad (2.10)$$

Izadi and Ghalambor (2012) proposed a new approach for the determination of hydraulic flow units. They modified the reservoir quality index (MRQI) by multiplying it by $\frac{1}{1-S_{wir}}$, accounting for the connate water in the pore network, Equation 2.11. They reported that the new approach gave better results for petrophysical properties distribution.

$$MRQI = 0.0314\ \sqrt{K/\varphi_e} \times \frac{1}{1 - S_{wir}} \qquad (2.11)$$

Davis *et al.* (1996) utilized integrated pore geometry attributes with wireline logs for the identification of hydraulic units in a mature and heterogeneous carbonate reservoir. They reported improved permeability and permeability predictions in partially-cored and in adjacent un-cored wells. Aminian *et al.* (2003) developed an unsupervised artificial neural networks classification model to predict flow units using only well logs as input.

They used the plot of cumulative flow capacity versus cumulative storage capacity as the primary tool to identify flow units for the model target data. The deflection points on this plot are potential indicators of flow units' boundaries (Gunter, 1997). Hussain Baker et al. (2013) utilized statistical regression to correlate flow zone indicator with well logs in a carbonate reservoir. They then used FZI values to compute permeability. Good agreement between actual and computed permeability values was reported.

Flow units identification techniques rely on the availability of permeability and porosity from core analysis. However, not all wells are cored. Porosity can be obtained from well logs. Permeability is the challenging parameter since there is no direct way to measure it from logs. This limits the identification of hydraulic units statistically.

2.3   Global Hydraulic Element

Corbett *et al.* (2004) suggested the term "petrotyping" to define a specific set of petrophysical rock types. The term is based on pre-defined set of Global Hydraulic Elements (GHE). Figure 2.2 shows the GHE "basemap" template. GHE1 is at the base of the template while GHE10 is at the top. For a range of porosities, Equation 2.12 is used to define the lower boundaries of GHEs. Table 2.1 lists the 10 FZI values that are used for each GHE.

$$K = \left(\frac{FZI \times \left(\frac{\varphi}{1 - \varphi} \right)}{0.0314} \right)^2 \qquad\qquad (2.12)$$

Table 2.1 FZI values used for 10 Global Hydraulic Elements (GHEs) lower boundaries

FZI	GHE
48	10
24	9
12	8
6	7
3	6
1.5	5
0.75	4
0.375	3
0.1875	2
0.0938	1
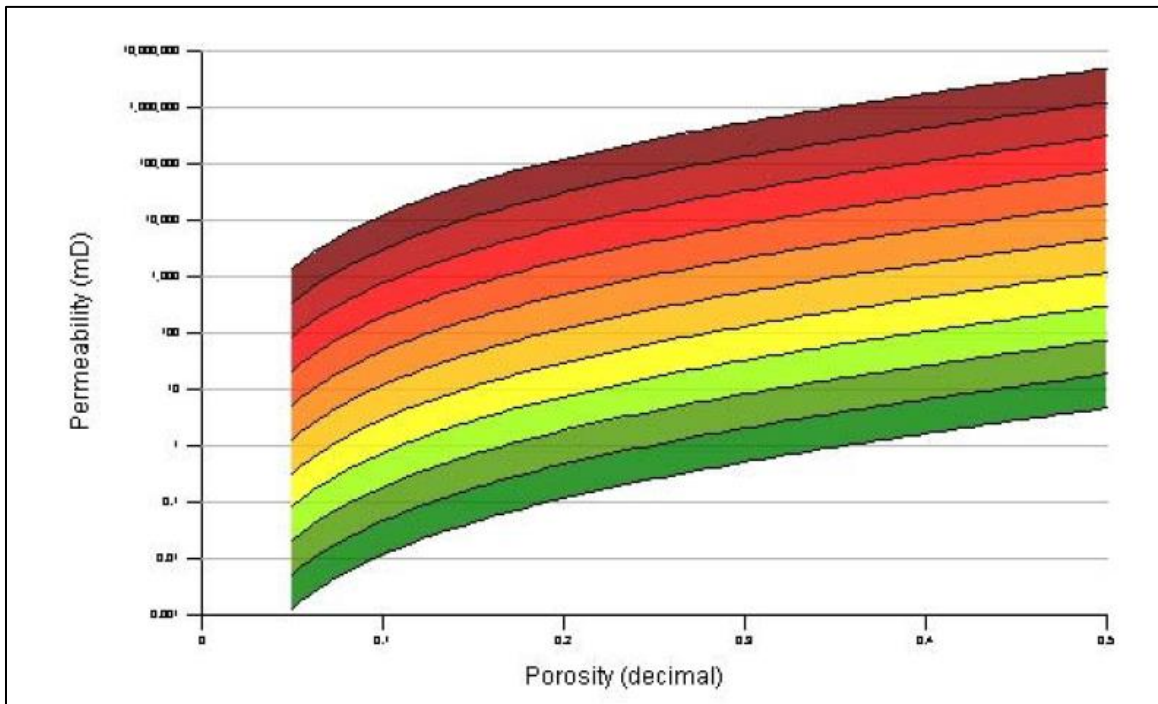
Figure 2.2 - Global Hydraulic Element "basemap" Template (Source: Corbett et al. (2004))

Corbett *et al.* (2004) reported the following advantages for plotting petrophysical data on the GHE basemap:

- It makes it easy to identify data trends

- It gives a constant reference frame for comparisons between reservoirs, wells, fields, core data and simulation data

- It helps in selecting a minimal and representative training data

The 10 GHEs are used as reservoir quality indicators. Salaheddin et al. (2010) reported that GHE1, GHE2, and GHE3 indicate poor reservoir quality. GHE4 and GHE5 indicate medium reservoir quality. Good reservoir quality is indicated by GHE6 and GHE7 whereas GHE8, GHE9 and GHE10 point towards excellent reservoir quality.

Corbett and Mousa (2010) demonstrated how GHE method can be used to select representative data for saturation exponent screening from a limited dataset. Astudillo and Porlles (2010), in a case study, also used GHE basemap to select representative samples as part of a multidisciplinary workflow for reservoir characterization in a heavy oil field. Salaheddin *et al.* (2010) presented a case study where they utilized GHEs in their reservoir characterization of Ordovician sandstones in Libya. They concluded that the GHE method is a good permeability tracer due to its adherence to grain-size distribution. Elarouci *et al.* (2012) utilized the global hydraulic elements method as part of their integrated evaluation of an Algerian tight gas reservoir for permeability prediction.

In this work, Equations 2.5 to 2.7 were used to compute continuous flow zone indicator values from core porosity and permeability data. Calculated Flow zone indicator values

were then used as model targets. Also, porosity and permeability data from cores were used to classify each data points using the global hydraulic elements basemap. The numerical part of each class was then used as additional input to enhance models performance.

CHAPTER 3

ARTIFICIAL INTELLIGENCE

3.1 Introduction

Artificial Intelligence is the theory of developing computer systems to be able to perform tasks that normally require human intelligence, such as visual perception, speech recognition, decision-making, and translation between languages. One of its subfields, machine learning, deals with the construction and study of systems that can learn from data, rather than follow explicitly programmed instructions. For machine learning, data must be available. It becomes convenient when there is no direct and obvious relationship between the available data. With recent advances of the computation powers, algorithms are now more optimized to build systems for forecasting and prediction, such as the Artificial Neural Networks (ANNs), Fuzzy Logic Systems (FLS), and Support Vector Machines (SVM). Other algorithms help identify those data features contributing the most in the model prediction process. Examples of the later algorithms are the Functional Networks (FN), Decision Trees (DT), and Fuzzy Information Entropy (FIE).

Chapter 2 covered the fundamental information about the concepts and formulizations of flow zone indicator. This chapter provides technical review for the five artificial intelligence techniques that were used in this work.

3.2   Artificial Neural Networks

Artificial Neural Networks (ANNs) are computational models inspired by biological nervous systems (in particular the brain) which are capable of machine learning as well as pattern recognition. "The central idea is to extract linear combinations of the inputs as derived features, and then model the target as a nonlinear function of these features" (Hastie *et al.*, 2011). The basic architecture of ANNs consists of three layers. The first layer, named the input layer, takes in the feature dataset from which the algorithm will learn. The second one, called the set of hidden layers, is where the model computation occurs. The third layer, called the output layer, produces the predicted parameters. Figure 3.1 shows the basic architecture for ANNs.
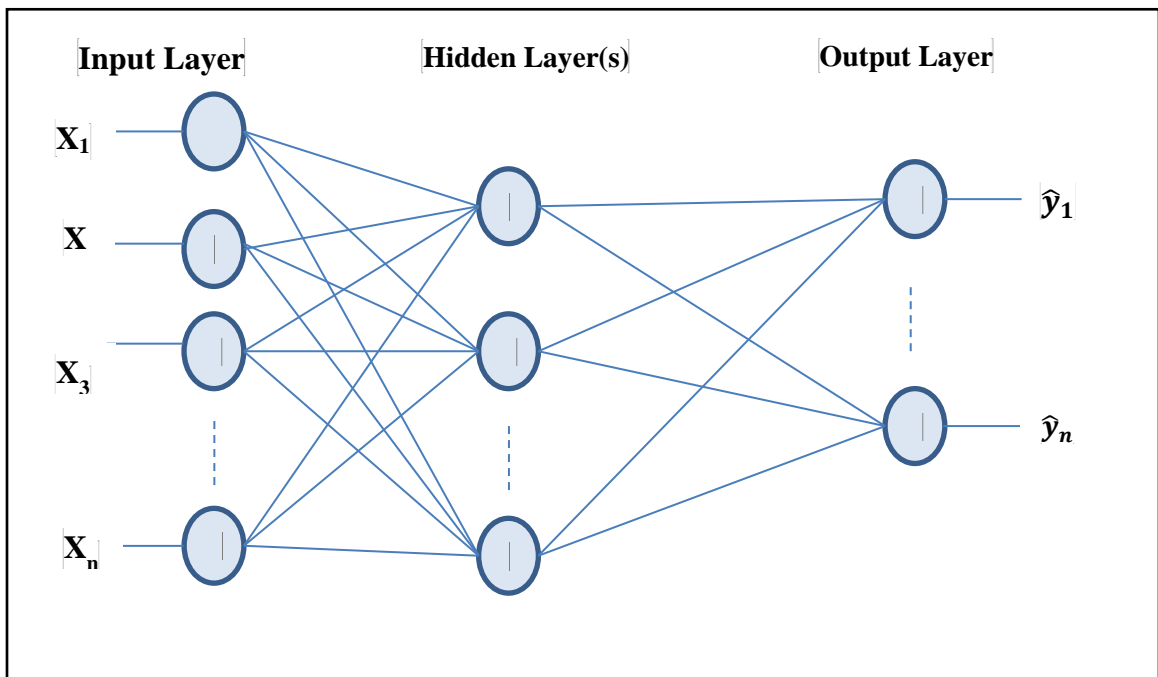
Figure 3.1 - Basic ANN Architecture

The fundamental building block in each layer is the neuron. Each neuron is connected to all other neutrons in the adjacent layers. Neuron activation requires three functional operations. The first function, the weight function, multiplies the single R-element input vector by the weight vector using the dot product. The second function, the net input function, adds the scalar value of bias to the weighted input vector. The last function, the transfer function, takes the net input and produces the scalar output for the neuron. Linear transfer function and log-sigmoid function are the two most commonly used transfer functions. Different Neural networks can be constructed using different weight functions, net input functions, and transfer functions in their neutron layers. Figure 3.2 and Equation 3.1 show the three functional operations that are used to calculate a neuron output.

$$a = f(\boldsymbol{Wp} + b) \qquad\qquad (3.1)$$

where,

$\boldsymbol{Wp}$ **:**   Weight function

$\boldsymbol{Wp} + b$:   Net input function

$f(\ \ )$:   Transfer function

a :   Neuron output
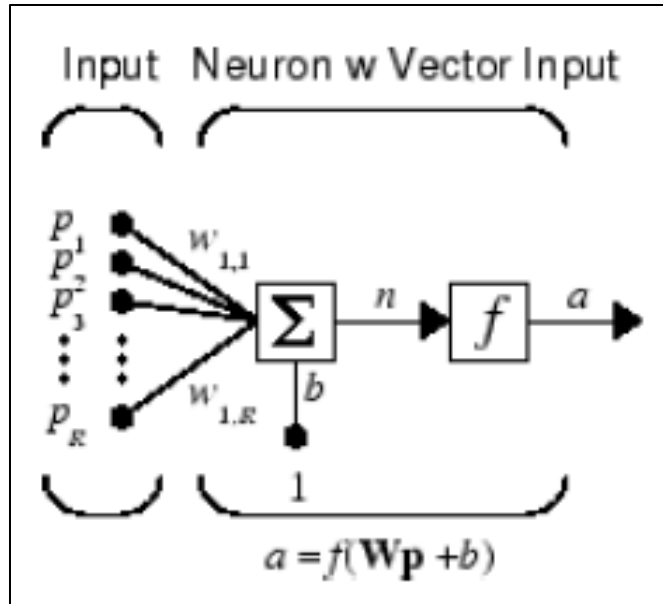
Figure 3.2 - Fundamental Operations in a neuron (Source: Matlab Documentations)

To optimize neural networks, the number of hidden layers, the number of neutrons in each layer, the bias value assigned to each layer, and the initial weights assigned to each connection need to be evaluated. The number of neutrons in the input and output layers equals the number of input features and desired outputs, respectively.

3.2.1   Feedforward Neural Networks

Feedforward Neural Networks (FFNN) consist of one or more hidden layers with sigmoid neurons followed by an output layer that has linear neurons. Its ability to learn non-linear relationships between input and output vectors is attributed to having multiple layers of non-linear transfer functions neurons. The linear output layer is mostly used in function fitting and non-linear regression applications. Figure 3.3 shows a plot of the sigmoid function.

$$\sigma(v) = \frac{1}{(1 + e^{-v})} \tag{3.2}$$
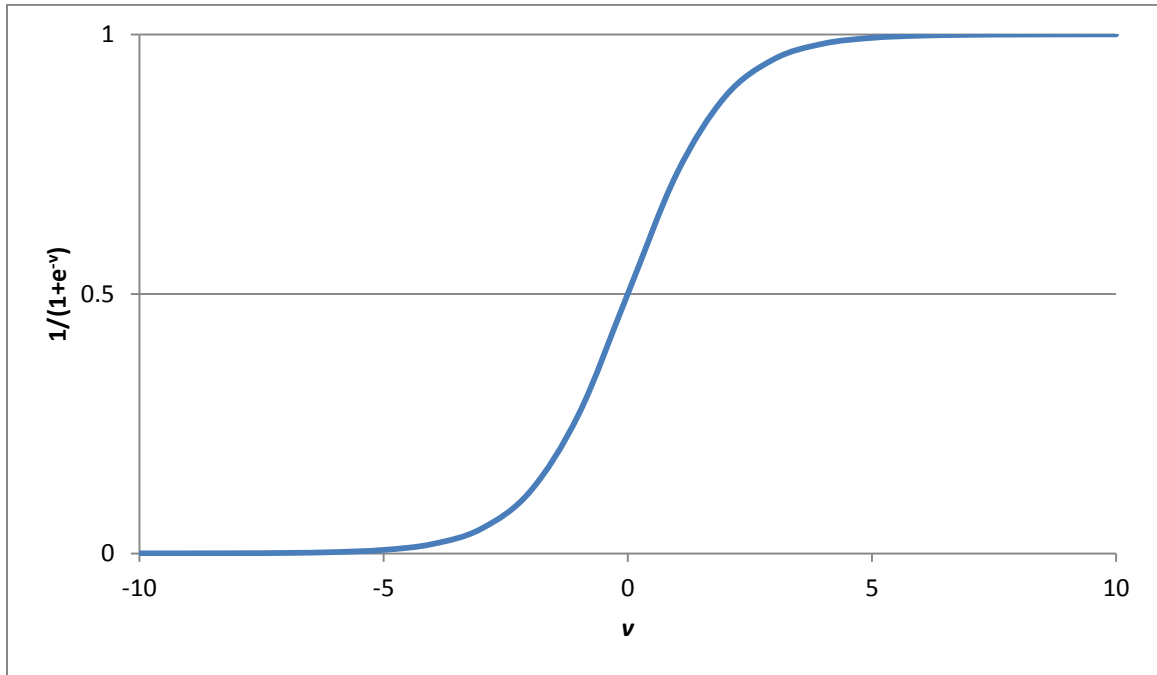
Figure 3.3 - Plot of Sigmoid Function

For a set of inputs x where $x=\{x_1, x_2, ..., x_{n1}\}$, output set y ,and a neural network that has n_2 neutrons in its hidden layer, connection weights (*w* and *v*) will be iteratively calculated. In each iteration, the error which is the difference between the target value *y* and the calculated one *y'*, is determined. This iteration process will continue and connection weights will be adjusted until the minimum error value is reached. The mathematical representation of the function used in the back-propagation neural network is:

$$y = f\left[w_0 + \sum_{j=1}^{n_2} w_j f_j \left(v_{0j} + \sum_{i=1}^{n_1} v_{ji} x_i\right)\right] \tag{3.3}$$

3.2.2 Radial Basis Function Neural Networks

Developing Radial Basis Function Neural Networks (RBFNN) often requires more neurons than standard feedforward back-propagation network. Their best performance is obtained when many training vectors are available.

The radial basis neutron uses the vector input between its input vector p and its weight vector w, all multiplied by the bias b for the calculation of its net input function. It then uses the radial basis transfer function (radbas) to generate the neutron output from the net input. Figure 3.4 shows a plot of the radbas function.

$$radbas(n) = e^{-n^2} \tag{3.4}$$
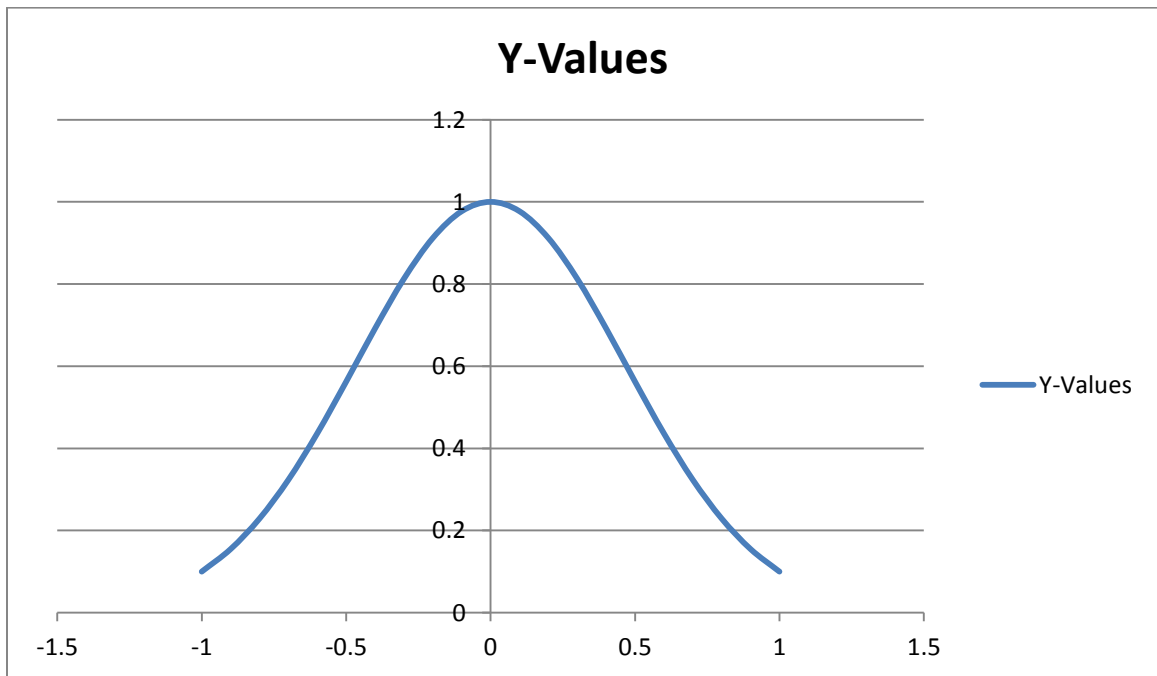
Figure 3.4 - Plot of radbas Function

The basic architecture for this type of networks consists of a hidden radial basis layer and an output linear layer.

3.2.3 Generalized Regression Neural Networks

Generalized Regression Neural Networks (GRNN) has a similar architecture to the radial basis network except it has a special linear layer. In the output layer, its net input function consists of multiplying the input vector, which is the output of the radial layer, by the weights vector using the dot product. Then, all values are normalized using the sum of the elements of the input vector.

Since GRNN has the same RBFNN architecture in the first layer, there are as neutrons as many the number of input vectors in the dataset.

3.3   Type-2 Fuzzy Logic Systems

Fuzzy Logic is about the relative importance of precision. It deals with the trade-off between significance and precision. It maps an input space to an output space using If-Then statements called rules. During its execution, all rules are evaluated in a parallel manner. So, the order of these rules does not matter. These rules are important because they refer to the system variables and the adjectives that describe them. Variables are either input variables or output ones depending on the values to which they refer. They are based on the concept of fuzzy sets. A fuzzy set is defined as a set without a crisp and clear boundary. Its members have a partial degree of membership.

Fuzzy Inference Systems (FIS) are systems that interpret the values in the input vector, input space, and utilize sets of rule-based statements to assign values to the output vector, output space. They consist of three primary components:

- **Membership Functions (MF)**

A membership function is a curve that defines how each value or point in the input space is mapped to a membership value. It is often designated by μ. Membership values must be between zero and one. This is the only condition membership functions need to satisfy. Points that have zero membership value are completely excluded from the set. However, points that have a membership value of 1.0 are completely included in the set. Any value in-between represents partial membership.

Mathematically, if Z is the input space and its elements are denoted by z, then a fuzzy set F in Z is defined as a set of ordered pairs:

$$F = \{z, \mu F(z) | z \in Z\} \tag{3.5}$$

Here, $\mu F(z)$ is the membership function of z in F

- **Fuzzy logic operators**

For its reasoning, Fuzzy Logic systems use three sets of operations: fuzzy intersection, fuzzy union, and fuzzy complement. Fuzzy intersection corresponds to the Boolean AND operator. Its general function for two fuzzy sets A and B is defined as:

$$\mu_A \cap B^{(X)} = \mu_A(X) \otimes \mu_B(X) \tag{3.6}$$

Where, $\otimes$ is a binary operator. Fuzzy union corresponds to the Boolean OR operator. Its general function for a two fuzzy sets A and B is defined as:

$$\mu_A \cup B^{(X)} = \mu_A(X) \oplus \mu_B(X) \tag{3.7}$$

Where, $\oplus$ is also a binary operator. Lastly, the fuzzy complement corresponds to the Boolean NOT operator which negates the input to produce its output.

- **If-Then rules**

If-Then rules are what make fuzzy logic useful. Their general form is:

If x is A, then y is B

The If-part of the statement is called the antecedent or promise, while the second part, Then-part, is called the consequent or conclusion. The antecedent is an interpretation that returns a single number between 1 and 0, whereas the consequent is an assignment that assigns the entire fuzzy set B to the output variable y.

Fuzzy inference systems go through a five-step process. This process is depicted in Figure 3.5.
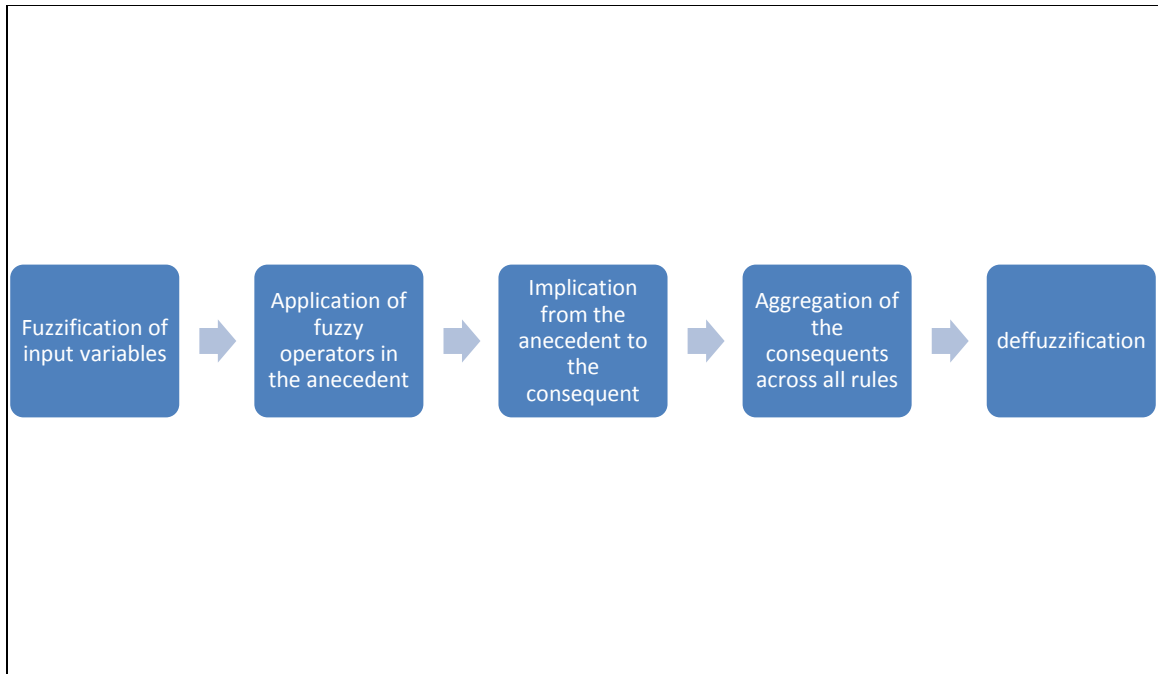
Figure 3.5 - Process Steps for Fuzzy Inference System

3.4 Support Vector Machines

Support Vector Machines (SVM) is a supervised machine learning methodology. It uses kernel functions to explore input data and recognize patterns. It was first proposed by Vapnik (Vapnik *et al.*, 1995). It has been used in many applications, including pattern recognition, regression estimation, and other areas.

Support Vector Machines classify data by finding the best hyperplane that separates data from one class from those of the other class. Data points closer to the separating hyperplane are called support vectors. The best hyperplane is one with the largest margin between the different classes, Figure 3.6.
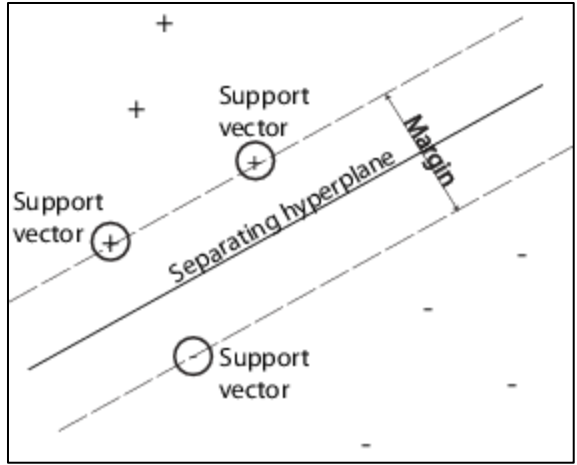
Figure 3.6 - Separating Hyperplane in Support Vector Machine (Source: www.mathworks.com)

CHAPTER 4

MODEL DEVELOPMENT AND RESULTS

4.1   Introduction

In this work, the model development went through three phases: Data pre-processing, Model building, and Model improvement. Each phase consists of three stages. The data pre-processing phase covers the required stages to ensure that the data are ready for the planned models. The model building stage covers the work from deciding on which measurements parameters to use to having the initial models optimized by parametric analysis. The first two stages of the third phase, model improvement, cover the additional efforts that were taken to further improve the initial models. The last stage examines whether all the previously completed stages led to improvements in permeability prediction. Figure 4.1 shows all three phases with their stages.

Figure 4.1 - Stages of Model Development

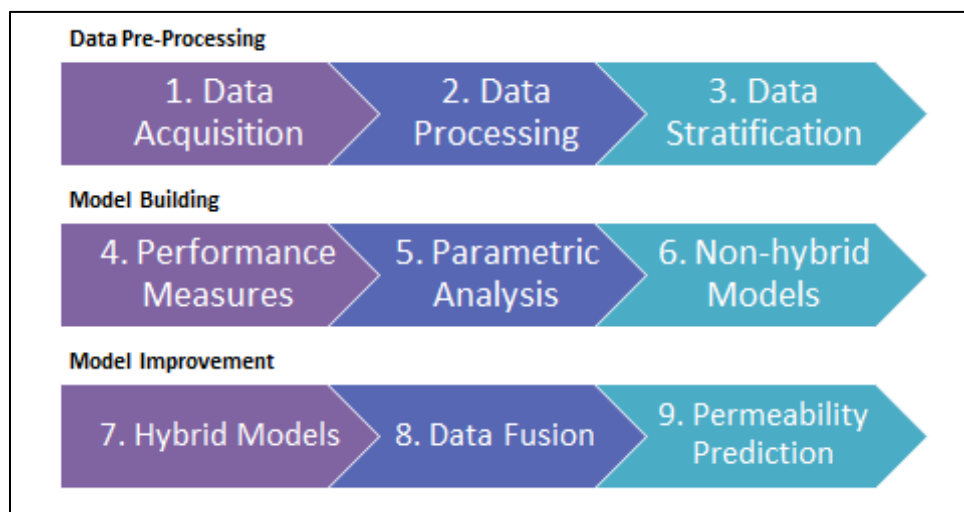In Chapter 3, the technical aspects of five AI techniques were reviewed. In this chapter, the first two phases are discussed in details, which cover the non-hybrid implementation of these techniques. Statistical and graphical comparisons of all models are provided. Then, each stage in the third phase will be discussed in a new chapter.

4.2   Data Acquisition

AI techniques are considered to be data-driven. This means that the performance of models built on them depends on the quality of utilized data. As a result, the collection and pre-processing of these data are vital steps to the success of any AI model development. For this work, conventional and NMR well logs data from several wells in a giant carbonate reservoir were collected. Also, the values for both porosity and permeability were obtained from core data. The resulting dataset included 487 data points. Eight conventional logs with their indices and statistics are compiled in Table 4.1 and 20 NMR parameters logs are listed in Table 4.2. All 28 logs were used as input dataset.

Continuous flow zone indicator values were then computed from equations 2.5 to 2.7. Flow zone indicator values were used as model targets, with their statistics compiled in Table 4.3.

Table 4.1 Statistical Description of Conventional Logs for Input

Index	1	2	3	4	5	6	7	8
Log Name	Gamma Ray (GR)	Laterolog Deep Resistivity (LLD)	Laterolog Medium Resistivity (LLM)	Laterolog Shallow Resistivity (LLS)	Neutron (NEUT)	Spontaneous Potential (SP)	Acoustic Log (DT)	Bulk Density (DRHO)
Units	GAPI	Log(OHMM)	Log(OHMM)	Log(OHMM)	NAPI/MAX(NAPI)	MV	(US/F)/MAX(US/F)	G/C3
Minimum	11.95	-0.16	-1.54	-0.19	0.45	44.82	0.69	-0.01
Maximum	34.53	2.91	2.57	2.01	1.00	82.99	1.00	0.35
Arithmetic Average	20.22	1.13	0.80	0.93	0.61	60.88	0.74	0.04
Median	19.93	1.19	1.07	1.09	0.56	60.20	0.72	0.01
Variance	14.28	1.00	1.84	0.41	0.02	51.63	0.00	0.00
Standard Deviation	3.78	1.00	1.36	0.64	0.12	7.19	0.06	0.06
Skewness	0.88	0.22	-0.18	-0.08	1.05	0.18	2.17	3.10
Kurtosis	2.17	-1.48	-1.63	-1.44	0.16	-0.12	4.36	9.57

Table 4.2 NMR logs

Index	Log Name	Index	Log Name
9	Bound Fluid Volume	19	Computed Regularization Parameter Gamma
10	CMR Free Fluid	20	Gamma Ray
11	CMR Porosity with T2 values greater than 3 ms	21	High Resolution Gamma Ray
12	CMR Porosity using Maximum T1/T2 Ratio	22	Integrated Permeability
13	CMR Porosity using Minimum T1/T2 Ratio	23	Integrated Porosity
14	CMR Porosity using T1/T2 Ratio of Zero	24	Signal Phase
15	Environmentally Corrected Gamma Ray	25	Irreducible Water Saturation
16	R Component of Spin Echoes	26	Computed T1/T2
17	X Component of Spin Echoes	27	T2 Logarithmic Mean
18	High Resolution Corrected Gamma Ray	28	Total CMR Porosity

Table 4.3 Statistical Description of Core Values for Targets

	Minimum	Maximum	Arithmetic Average	Median	Variance	Standard Deviation	Skewness	Kurtosis
FZI	0.000	18.796	1.920	1.394	5.309	2.304	3.120	14.392

4.3   Data Preparation and Processing

Quality measures were performed on all acquired data to ensure that they were complete. Outliers in target data were removed, using 3.0 standard deviations as a cut-off value. Around 20 Nuclear Magnetic Resonance (NMR) logs were included in the input data set. The primary reason for adding such extra logs was to add some degree of lithology independence to the other eight conventional logs. Another reason for their addition was to examine NMR logs contribution to FZI prediction and determine if improvements occurred.

By examining the logs that were selected by the different feature selection algorithms during models runs, some NMR logs were part of the selected input datasets in hybrid models. Table 4.4 summarizes the names of selected NMR logs when all logs were used.

Table 4.4 Selected NMR Logs by Feature Selection Algorithms

Feature Selection Algorithm	Selected NMR Logs
Functional Networks	- Integrated Permeability - Irreducible Water Saturation
Decision Trees	None
Fuzzy Information Entropy	- High Resolution Gamma Ray - Integrated Permeability - T2 Logarithmic Mean

Also, to reduce any effect of large data variances, resistivity logs were converted to log-base and Gamma ray, Neutron and Acoustic logs were normalized.

4.4   Data Stratification

Data stratification refers to how the available data is divided into training and testing datasets. Training data sets, commonly around 70% of the original data, is used for training AI prediction models. Then, these trained models utilize the remaining unseen data, around 30% of original data, to test model generalization.

Table 4.5 - Data Stratification

Data Sets	Percentage of Original Data	Number of Data Points	Total Number of Data Points
Training	70 %	341	487
Testing	30 %	146	

There are different ways to stratify data. Random stratification was used in this study. All data points for training and testing sets were selected randomly from the original one. This randomness helped generalize the developed models and mitigate the effects of any data bias.

4.5   Measures of Performance

Different statistical error measurements were utilized to optimize the performance of all AI models and to compare their performance. The first measurement was the correlation coefficient (CC). It is a calculated statistics between -1 and 1 to represent the linearity of

dependence between two variables or datasets. The second measurement, root mean-squared error (RMSE), also known as root-mean-square deviation (RMSD), is a measure of the difference between a model-predicted values and actual observations. The last measurement, the mean absolute error (MAE), is an average of the absolute differences between predicted values and desired outcomes.

$$CC = \frac{n \sum y\hat{y} - (\sum y)(\sum \hat{y})}{\sqrt{n(\sum y^2) - (\sum y)^2}\sqrt{n(\sum \hat{y}^2) - (\sum \hat{y})^2}} \tag{4.1}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{n}} \tag{4.2}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|\hat{y}_i - y_i| \tag{4.3}$$

where,

      n : Number of data points

      y : Set of target values

      $\hat{y}$: Set of predicted values

4.6   Parametric Analysis

Performance optimization for the AI models was achieved by examining variations of their design parameters. Values that yielded the highest correlation coefficients were then selected as the optimum design parameters. All 28 input logs were utilized in the parametric analysis. Table 4.6 provides a summary of all optimization parameters used in this work for prediction models.

4.6.1 Feedforward Neural Network Model Optimization

Feedforward Neural Networks models are optimized by evaluating the number of hidden neurons parameter, in the hidden layer, at different values. Correlation coefficient values for both training and testing were plotted against the number of hidden neurons. Then, the optimum value was identified at the point where both curves overlapped with the highest CC.

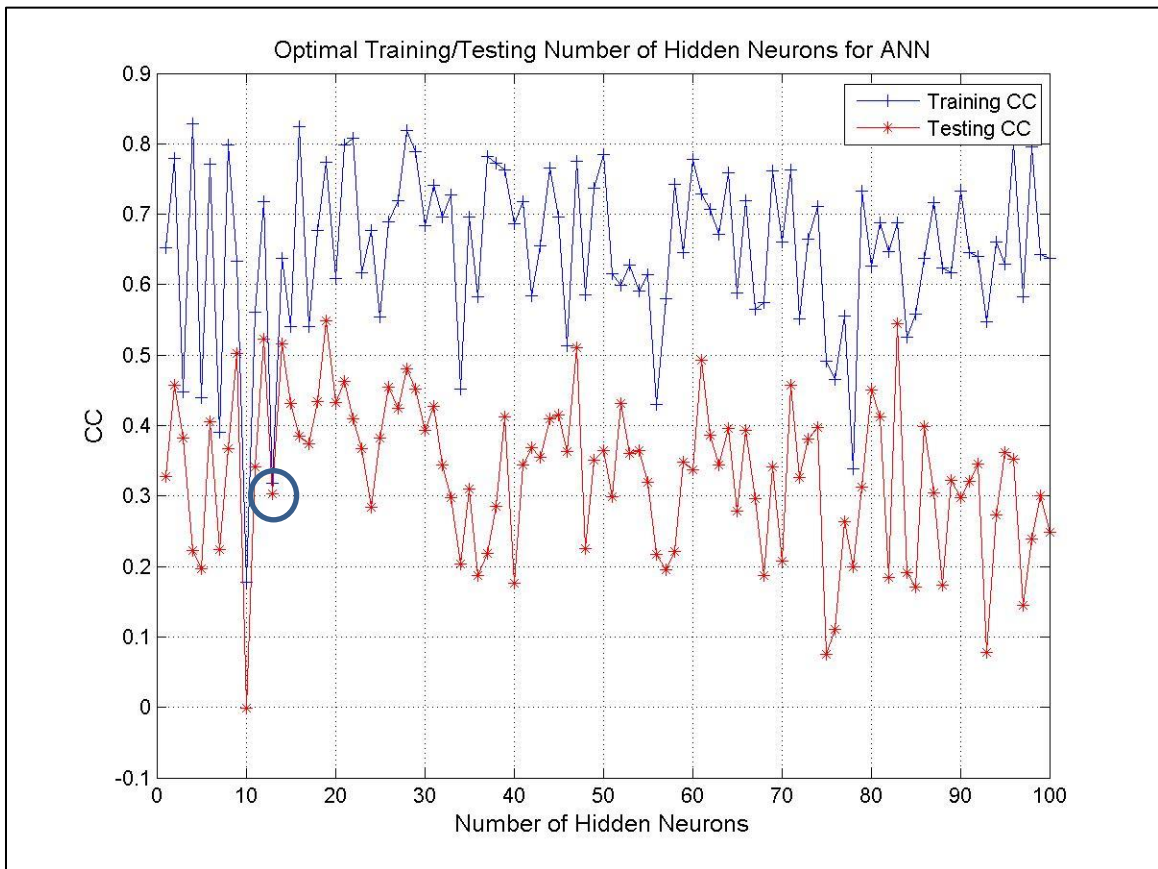Figure 4.2 shows the optimum value for the number of hidden neurons parameter, 13.

Figure 4.2 - Optimal Training/Testing Number of Hidden Neurons parameter for FFNN

4.6.2 Radial Basis Function Neural Network Model Optimization

RBFNN models are optimized using two design parameters: GOAL and SPREAD. To optimize the RBFNN model, the default value for the GOAL parameter was initially used while evaluating the model at different SPREAD values, ranging from 0.5 to 20.0. Figure 4.3 shows that SPREAD parameter = 3.0 yielded the highest correlation coefficient (CC). Then, spread parameter was assigned the value of 3.0 and the goal parameter was varied from 0 to 1. Goal value of 0.1 was identified to be the optimum, Figure 4.4.
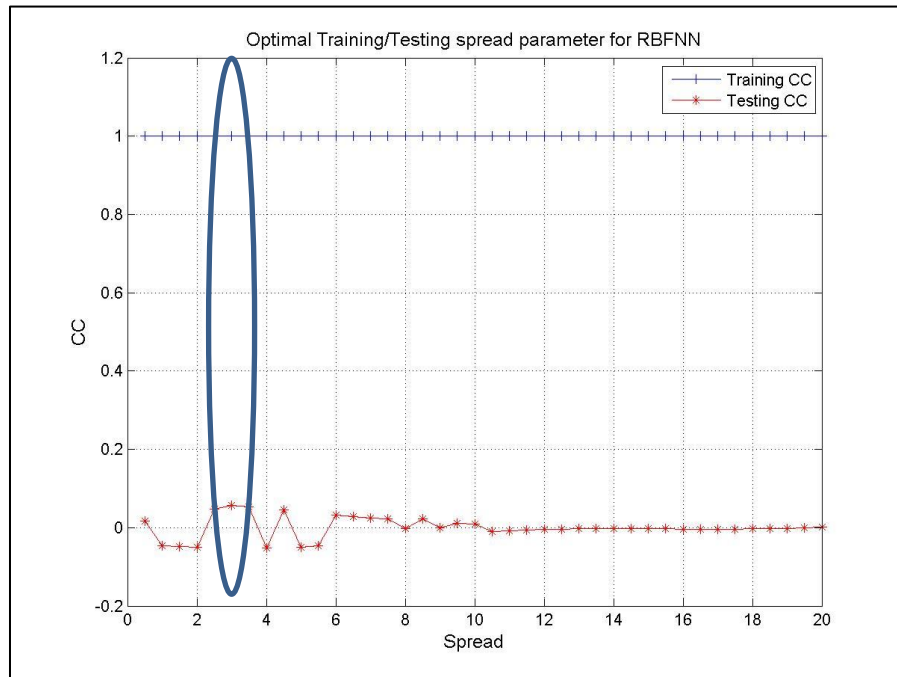
Figure 4.3 - Optimal Training/Testing spread parameter for RBFNN
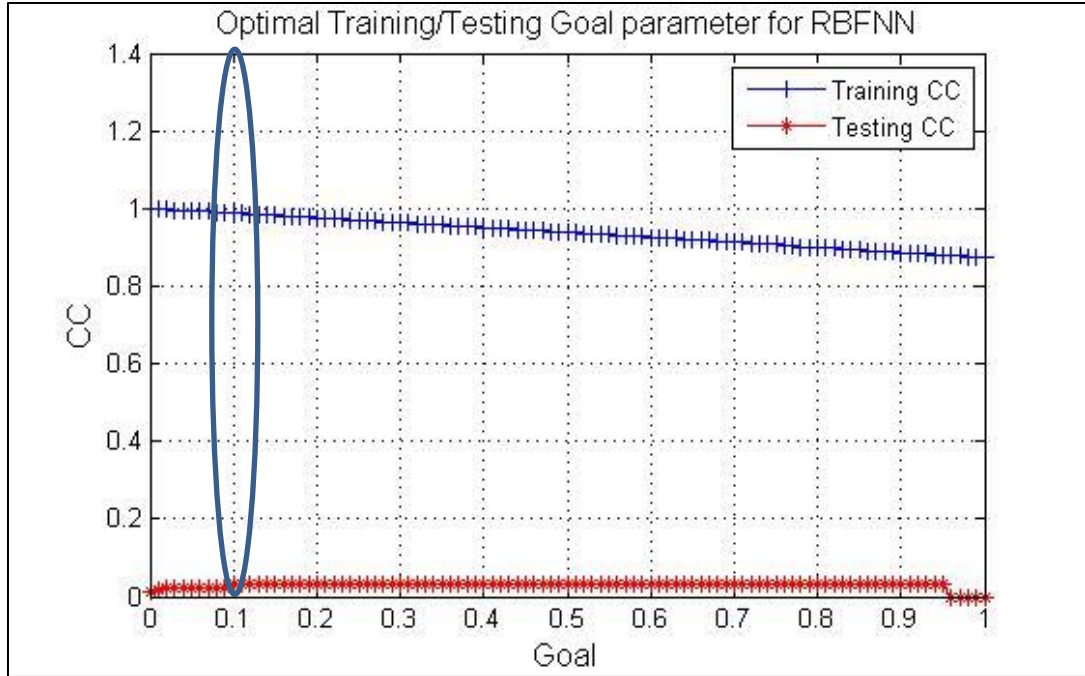
Figure 4.4 - Optimal Training/Testing goal parameter for RBFNN

4.6.3 Generalized Regression Neural Network Model Optimization

GRNN models are optimized using the SPREAD parameter. To optimize the GRNN model for this work, it was evaluated at different SPREAD values ranging from 0.5 to 500.0. Figure 4.5 shows that SPREAD parameter = 50.0 gives the best performance based on correlation coefficient (CC).
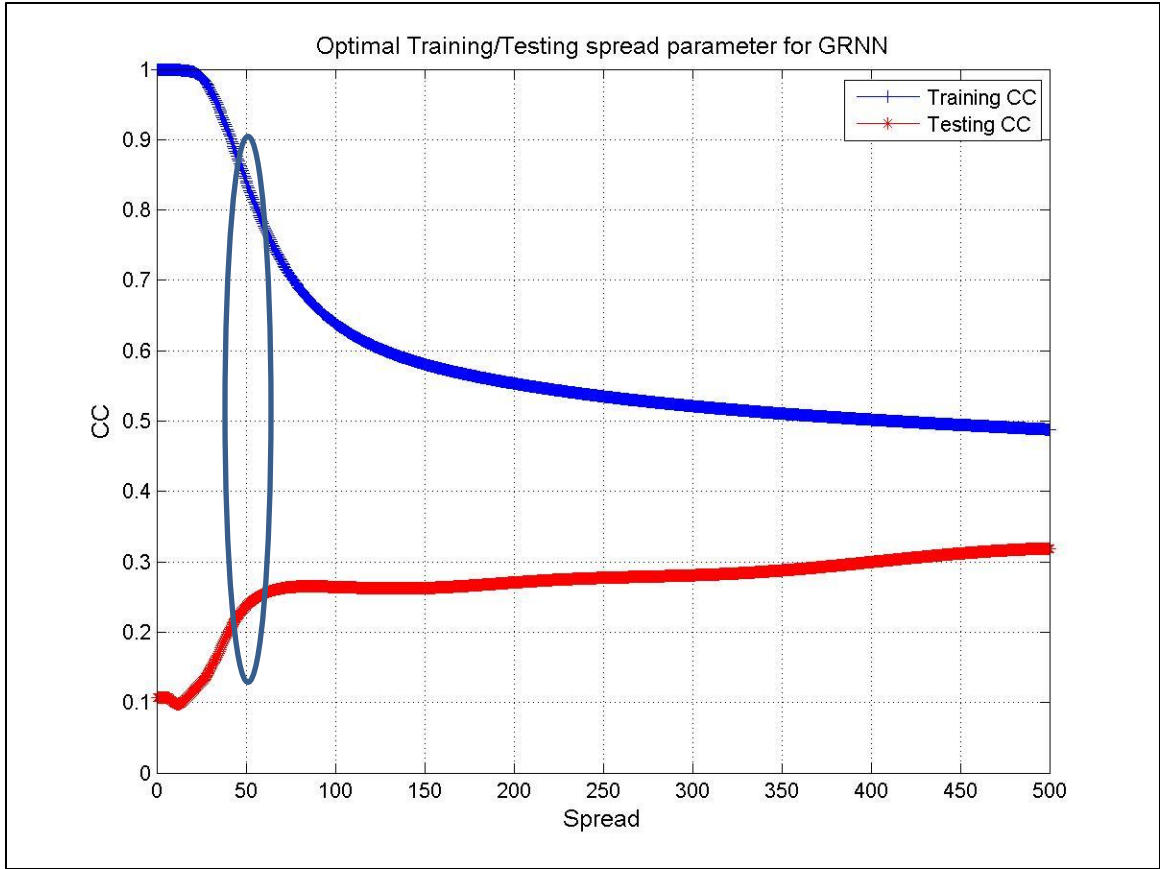
Figure 4.5 - Optimal Training/Testing spread parameter for GRNN

4.6.4   Type-2 Fuzzy Logic Model Optimization

Type-2 Fuzzy Logic Systems (T2FL) are optimized by evaluating the model at different learning rates, α. Optimum rates are selected on the base that the prediction model gives the highest correlation coefficient for the testing. Figure 4.6 indicates that the optimum learning rate for Flow Zone Indicator prediction was 1.75. This value had the smallest difference between training and testing and with the highest CC. This value was then used in all the Type-2 Fuzzy Logic prediction models.
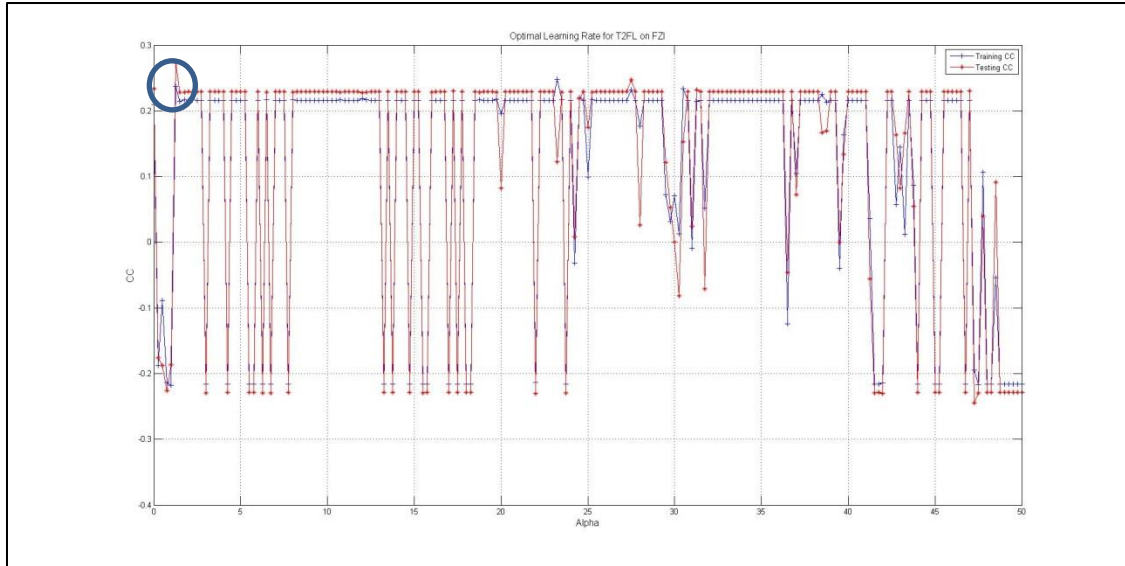
Figure 4.6 - Optimization of T2FL Learning Rate for Flow Zone Indicator Models

Table 4.6 Summary of Optimization Parameters for AI Techniques

AI Model	Optimization Parameters
Feedforward Neural Networks	Number of hidden neurons = 13
Radial Basis Function Neural Networks	Goal = 0.1 Spread = 3.0
Generalized Regression Neural Networks	Spread = 50.0
Type-2 Fuzzy Logic	$\alpha = 1.75$
Support Vector Machines	Type of Kernel = Gaussian Kernel Step Size = 4.35 Verbose = 2.0 Error Allowance, Lambda = 1e-7 Regularization Parameter, C = 10000 Penalty of Over fitting, epsilon = 0.1

4.7   Flow Zone Indicator Prediction Models Development

MATLAB software provides a large library of functions and techniques for building most of the AI models. For this reason, it was used for designing and optimizing the proposed AI models. Five non-hybrid AI models were developed for the prediction of Flow Zone Indicator. Each model uses a different AI technique: FFNN, RBFNN, GRNN, SVM, and T2FL. They use the optimum parameters that were obtained from the parametric analysis performed earlier. They also use all the available 28 data logs.
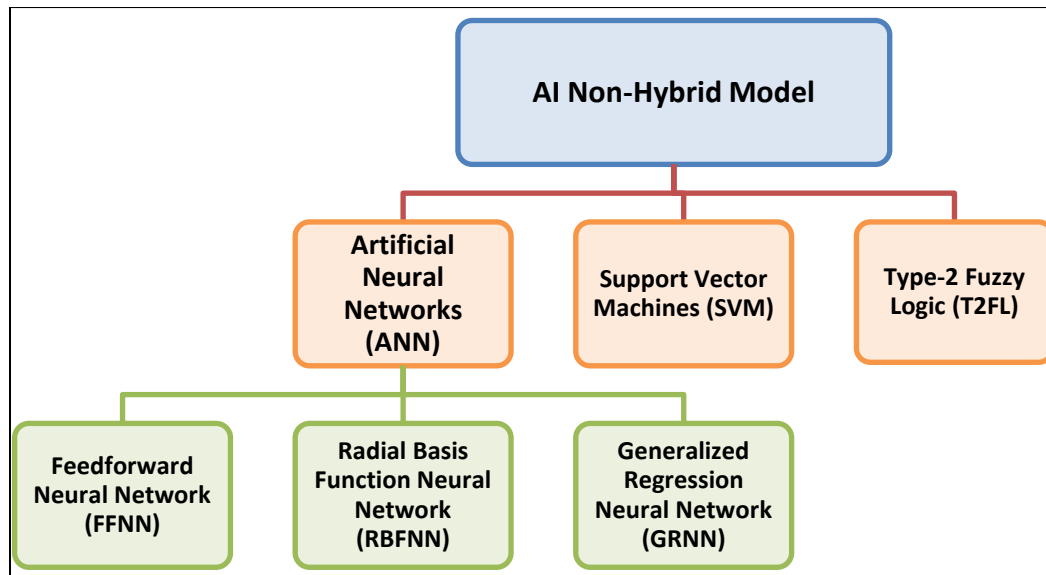
Figure 4.7 - Tree Representing the Non-hybrid AI Models

After designing all the five models, data was stratified into training and testing datasets. The models were trained using the training dataset, and the rest were for testing. To compare the performance of all non-hybrid AI models, results were compared statistically using error analysis and graphically using cross-plots.

4.7.1 Performance of Neural Networks Models

Statistically, FFNN outperformed both RBFNN and GRNN. It gave the best results for prediction in terms of CC, RMSE, and MAE. RBFNN had the lowest CC value. It might be attributed to its bad training performance. GRNN model was the second best performer between all non-hybrid models. Figures 4.8 – 4.10 shows the statistical error comparison for all non-hybrids neural networks models.
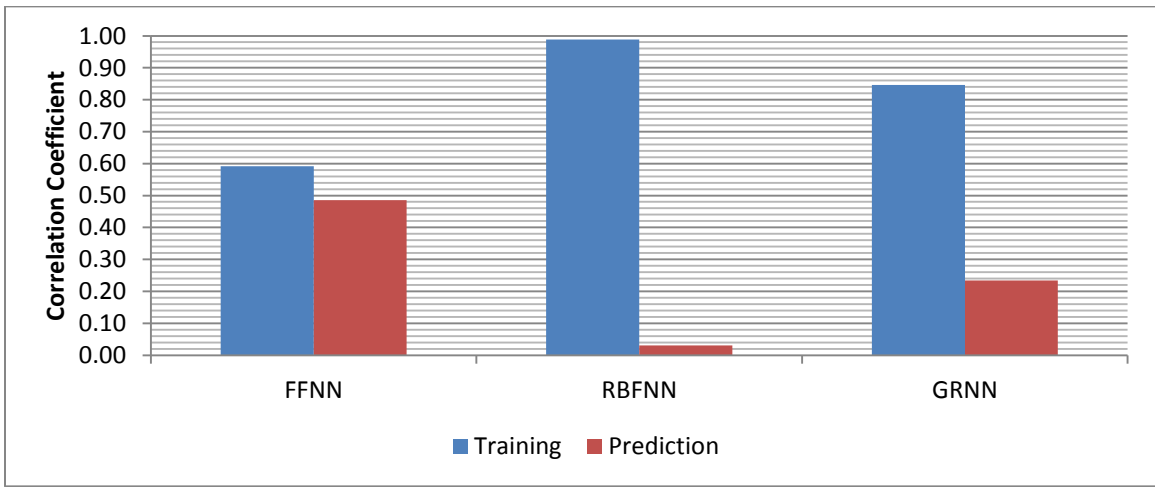
Figure 4.8 - Correlation Coefficient Comparison for Non-hybrid Neural Networks Models
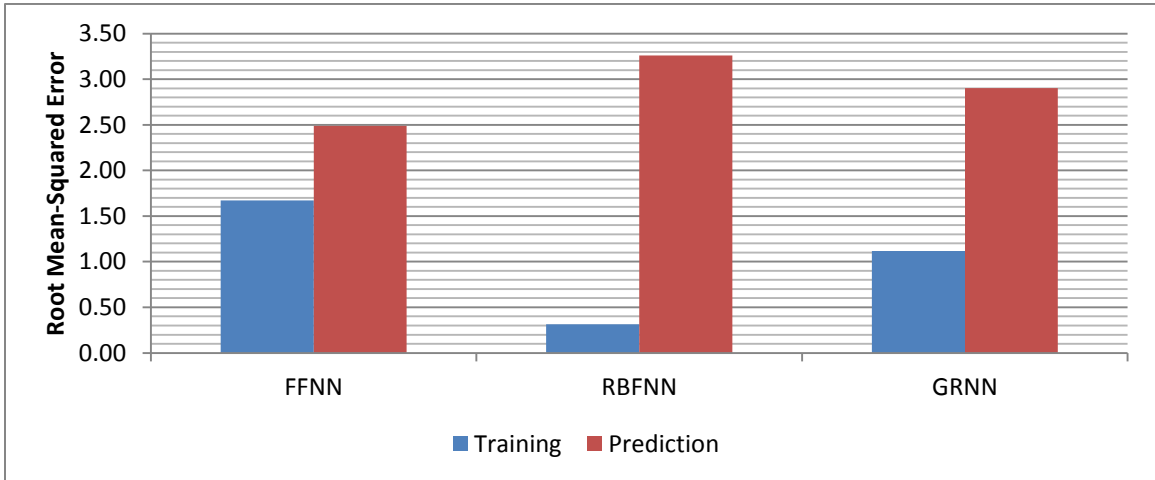
Figure 4.9 - Root Mean-Squared Error Comparison for Non-hybrid Neural Networks Models
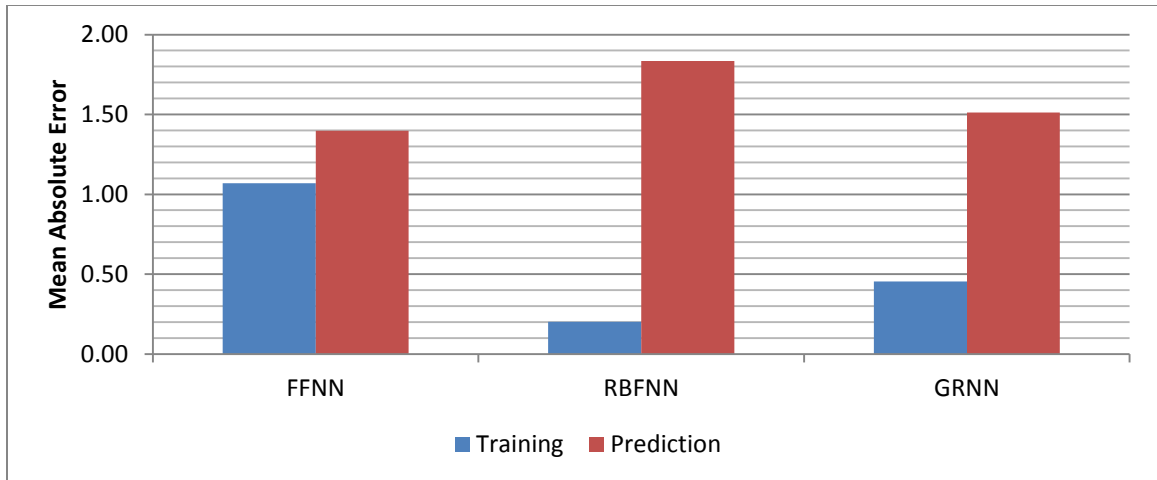
Figure 4.10 - Mean Absolute Error Comparison for Non-hybrid Neural Networks Models

Graphically, Figures 4.11 and 4.13 show the FZI prediction crossplots for training. Figures 4.12 and 4.14 show similar cross-plots for testing. For FFNN and GRNN, it can be seen that for FZI values smaller than 4, the correlation between the values are good. But, all models tend to underestimate the FZI values starting from 4 upwards. Bad correlation was obvious for RBFNN.
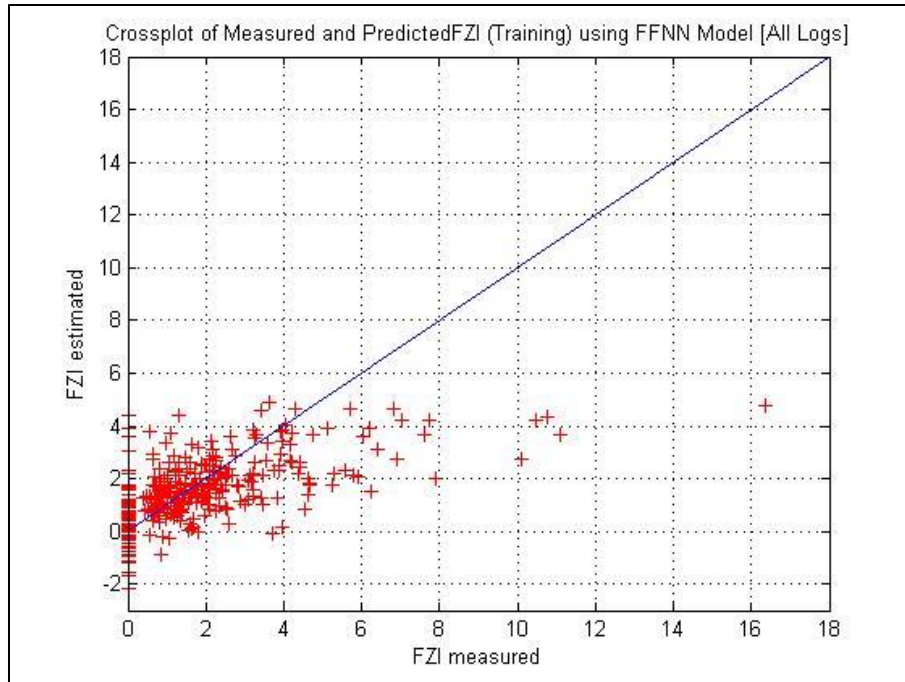
Figure 4.11 - FZI Prediction Crossplot using FFNN (Training)
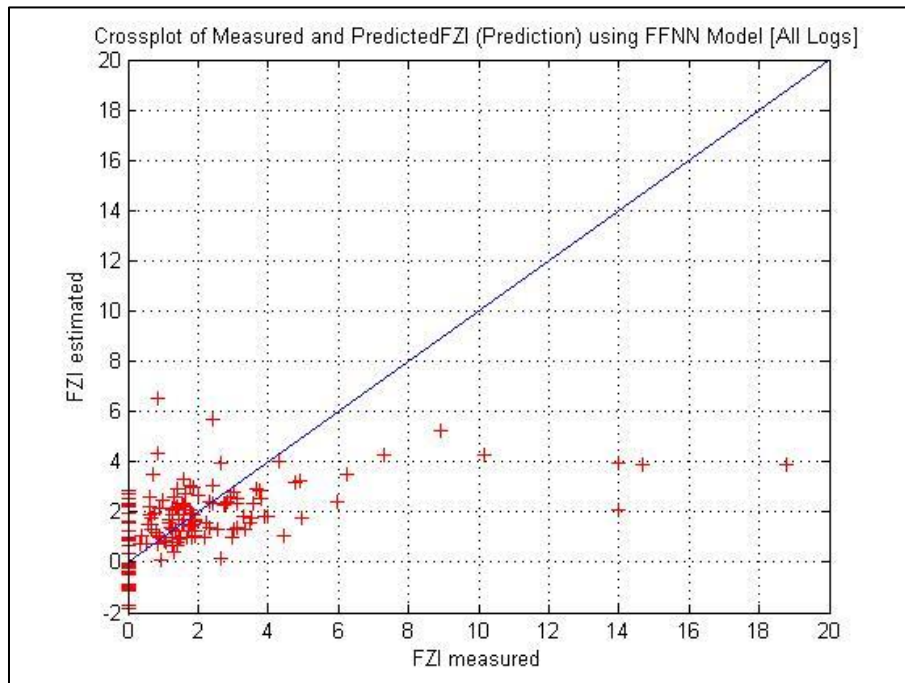
Figure 4.12 – FZI Prediction Crossplot using FFNN (Testing)

41
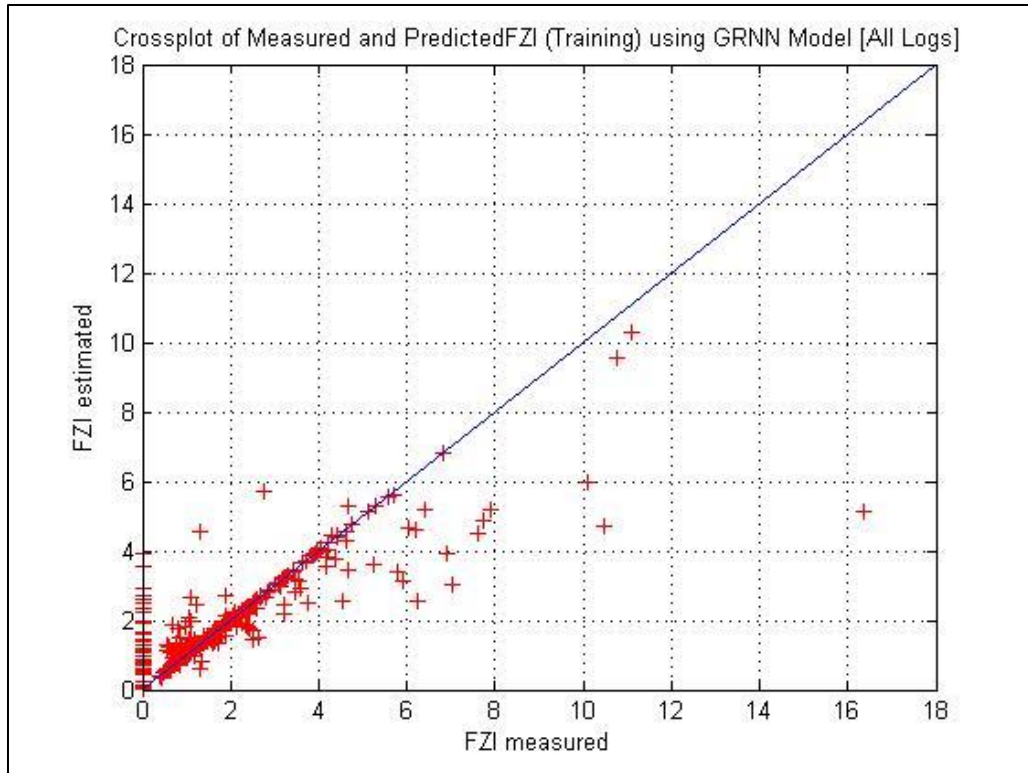
Figure 4.13 - FZI Prediction Crossplot using GRNN (Training)
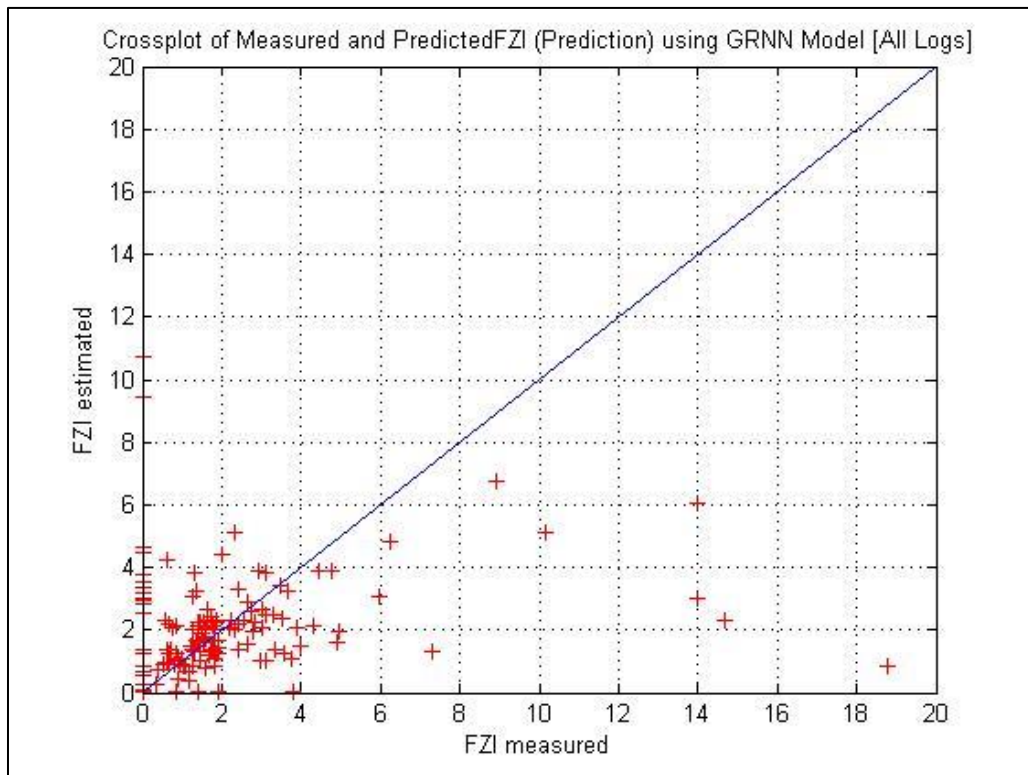
Figure 4.14 – FZI Prediction Crossplot using GRNN (Testing)

42

4.7.2 Performance of Support Vector Machine Models

SVM performance results showed that the model was not better than the one obtained with the FFNN model. By looking at the correlation coefficient for SVM compared to the other models, Figure 4.17, it can be seen that it was the second worst performer. It was due to its bad generalization in the training. Figures 4.18 and 4.19 show RMSE and MAE results for both training and testing. Finally, SVM was the fastest model to compute.

4.7.3 Performance of Type-2 Fuzzy Logic Models

The Type-2 Fuzzy Logic model was the third best performer, with a correlation coefficient of 0.2 for testing. This was also approximately true for RMSE and MAE values. T2FL model had the longest computation time. In addition, Figures 4.15 and 4.16 show that the T2FL followed the same trend in training and testing, where results for FZI below 4 are better correlated than the ones with higher FZI values.
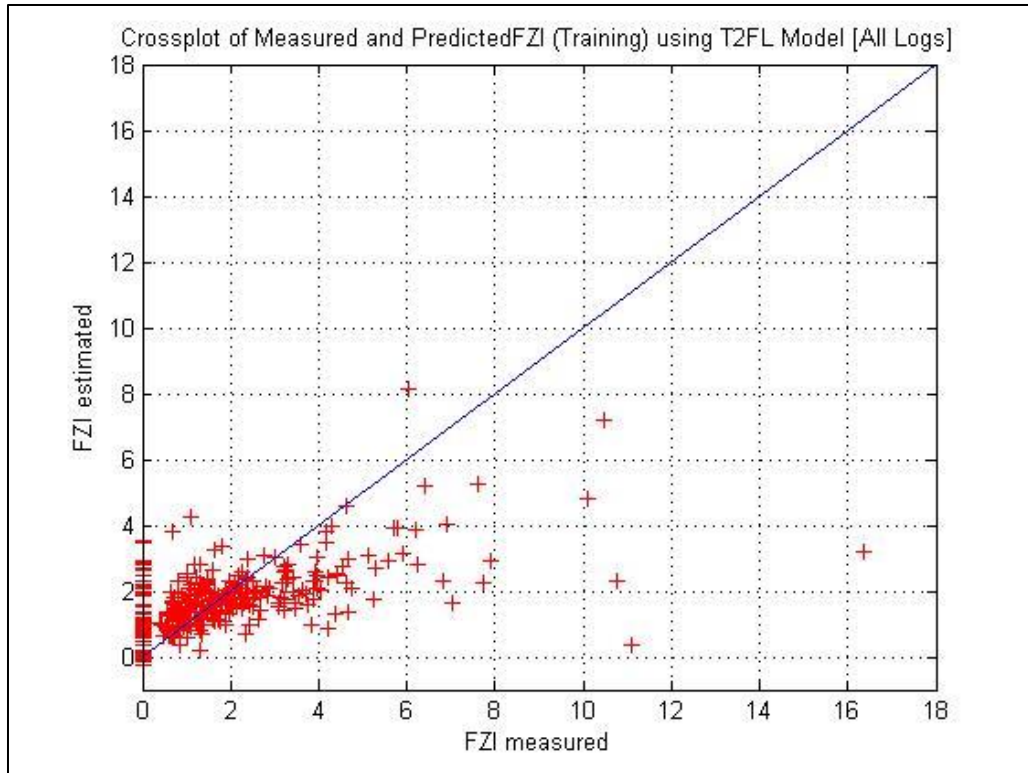
Figure 4.15 - FZI Prediction Crossplot using T2FL (Training)
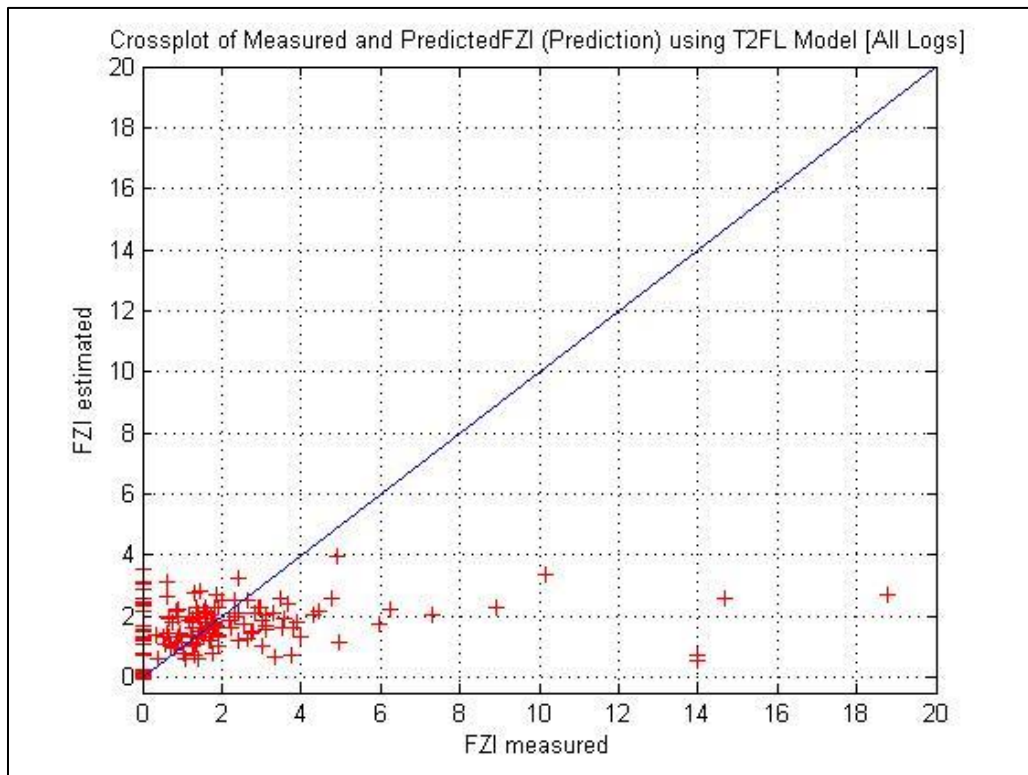
Figure 4.16 – FZI Prediction Crossplot using T2FL (Testing)

44

Figures 4.17 – 4.19 shows statistical error comparison for all non-hybrid AI models. The FFNN model had the best performance among them all, while the RBFNN model had the worst.

For further performance improvements, the best neural networks model, FFNN, will be considered with the SVM and T2FL models.
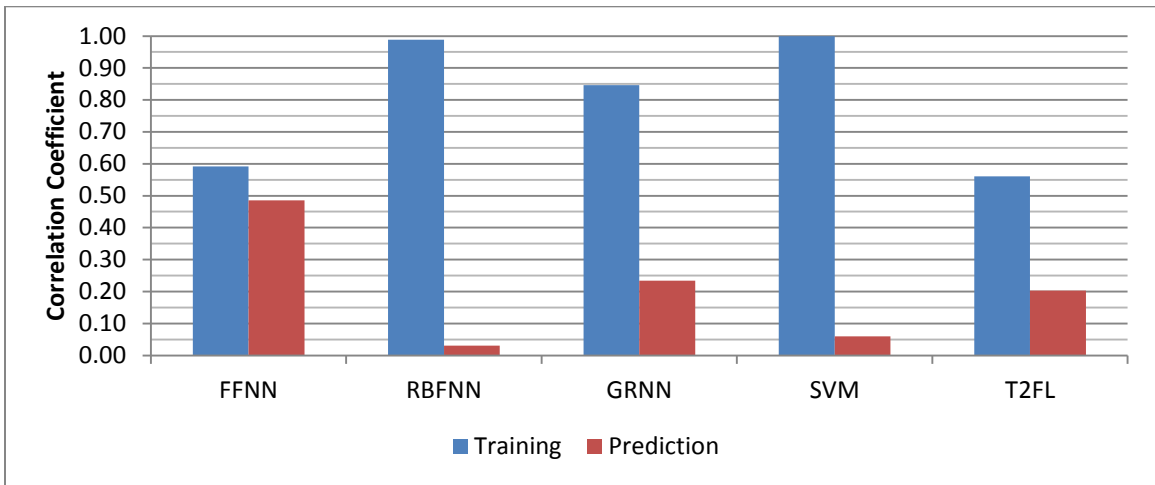

Figure 4.17 - Correlation Coefficient Comparison for Non-hybrid AI Models
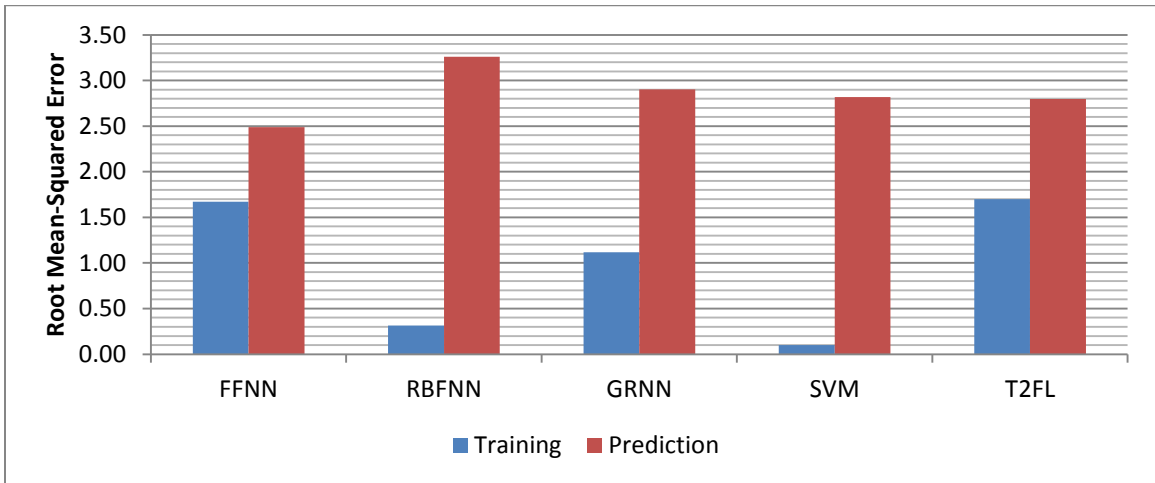
Figure 4.18 - Root Mean-Squared Error Comparison for Non-hybrid AI Models
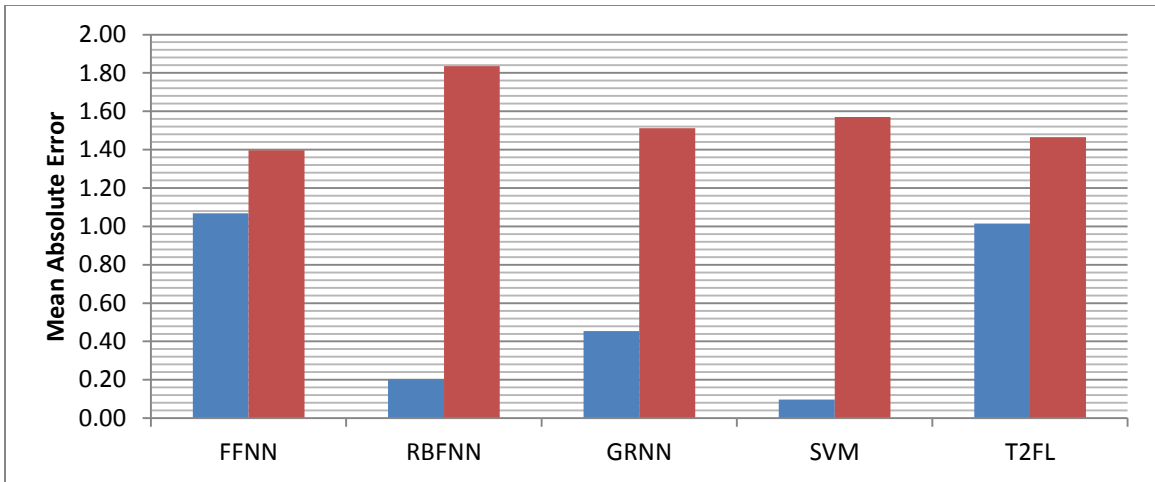
Figure 4.19 - Mean Absolute Error Comparison for Non-hybrid AI Models

CHAPTER 5

HYBRIDIZATION OF ARTIFICIAL INTELLIGENCE

5.1   Introduction

In Chapter 4, the development of five non-hybrid AI models for FZI prediction was covered. Models performance and results were discussed and compared with each other. For performance improvement using hybridization, three of them were selected as predictive models: FFNN, SVM, and T2FL. In this chapter, the fundamentals of AI hybrid systems are presented. Then, feature selection algorithms are introduced. Next, the development of nine hybrid AI models with their results is discussed in details. Lastly, performance comparison is presented for the determination of performance improvement.

5.2   Hybrid Systems

Hybridization is the integration of more than one technique into one. The primary reason for hybridization is to balance one technique's weakness with the other ones' strength. Hybrids are believed to be more versatile and robust than their individual techniques (Anifowose and Abdulazeez, 2010). An example is the Adaptive Neuro Fuzzy Inference System (ANFIS), consisting of both Artificial Neural Networks (ANN) and Fuzzy Logic (FL).

Hybrids common practice in AI applications is to integrate the feature selection algorithms such as FN, DT, and FIE with the machine learning tools such as ANN. The original input dataset is passed through the feature selection algorithm first. Then, the

selected features are divided into training, validation, and testing subsets. Next, the machine learning algorithm is trained using the training subset and validated using the testing subset for model generalization. Performance measurements are finally used to select the final predictive model which yields the most accurate results. Figure 5.1 shows a schematic for a hybrid AI system process.
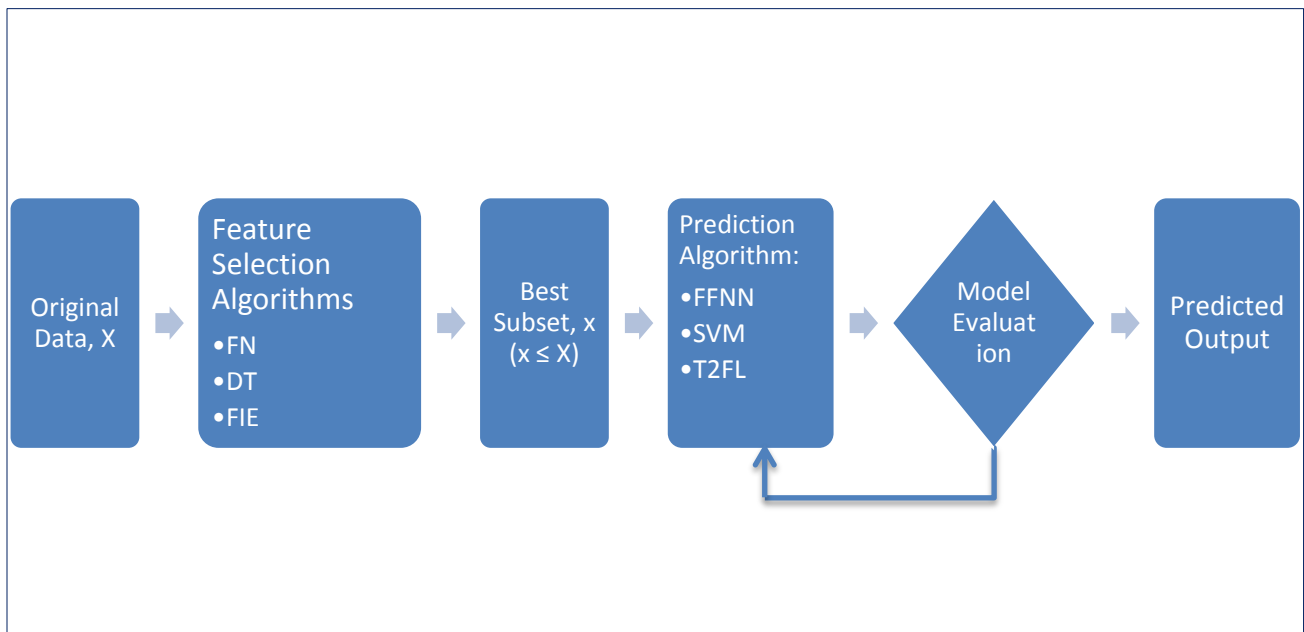

Figure 5.1 - Hybrid AI System Process

5.3 Features Selection Algorithms

Feature selection is the process of selecting subgroup of related features for the use of building predictive models. This process is used in both machine learning and statistics, with the general assumption that the original dataset has irrelevant features. These redundant features do not add more value than the one provided by the selected ones.

Feature selection algorithms or techniques are best used in fields where data have many features compared to the number of available data points. Main advantages of utilizing these algorithms include:

- Shortened execution time as a result of reduced input dataset

- Improved model performance due to the removal of redundancy in data

- Increase resolution of overfitting by enhancing the model generalization

Functional Networks, Decision Trees, and Fuzzy Information Entropy are three examples of feature selection algorithms.

5.3.1 Functional Networks

Functional Networks are considered an extension of ANNs. Their structure consists of different layers having neurons connected together. Each neutron has a scalar fix function which sums the weighted inputs. Functions are learned from the input dataset while suppressing the weights using algorithms such as the least-squares fitting. To select the best subset from an input data, Functional network uses minimum description length (MDL) algorithm. The selection criterion is based on the best non-linear relationship between the input and the target data.

Functional Networks differ from neural networks by not only in learning from the data about the problem but also in using knowledge about the problem to derive its topology. Another difference between them is that during learning, only weights are adapted and the network structure is fixed in neural networks while neural functions are also learnt in functional networks. Also, functions in functional networks can be multivariate while activation functions in neural networks have only one argument. Finally, functional networks are feedforward networks only, they do not have back-propagation.

5.3.2 Decision Trees

Decision Trees (DT) can be used as predictive models to map observations about a problem to conclusions about the problem's target outcomes. In these tree structures, leaves represent class labels, and branches represent conjunctions of features that lead to those class labels. There are two types of decision trees: classification trees and regression trees. Classification trees are used when the prediction targets are classes of variables. Regression trees are used when the predicted outcomes are real numbers. One of the main differences between the two types is in splitting procedure (where to split). Gini Impurity and Information Gain are two of the algorithms used to construct decision trees. Gini impurity is based on squared probabilities of membership for each target category in the node. It reaches its minimum (zero) when all cases in the node fall into a single target category.

$$I_G(i) = 1 - \sum_{j=1}^{m} f(i,j)^2 = \sum_{j \neq k} f(i,j)f(i,k) \tag{5.1}$$

The information gain is based on the concept of entropy used in information theory as expressed in the equation:

$$I_E(i) = 1 - \sum_{j=1}^{m} f(i,j)^2 \log_2 f(i,j) \tag{5.2}$$

During the generation of a decision trees model, the relative importance of each feature in the dataset is determined. If the dataset is divided using features values, a finite number of subsets will be obtained. In each subset, the information value I_i is computed so that I_i < I and the difference $(I - I_i)$ is a measure of how well the parameter has discriminated between different subsets. Then, the value that maximizes this difference is selected. This

50

process is carried on until the subset at a node has the same value as the target variable, or when splitting no longer adds value to the predictions (Yohannes and Webb, 1999).

5.3.3 Fuzzy Information Entropy

Fachao et al. (2008) described the Fuzzy Information Entropy (FIE) algorithm in the following manner:

Let $U = \{x_1, x_2, ..., x_n\}$ be a non-empty universe, R is a fuzzy equivalence relation on U, and let $[x_i]_R$ be the fuzzy equivalence class containing xi generated by R, it follows that:

$$|[x_i]_R| = \sum_{j=1}^{n} R(x_i, x_j) \tag{5.3}$$

which is called the cardinality of $[x_i]_R$. A further relation is given by:

$$H(R) = -\frac{1}{n} \sum_{i-1}^{n} \log_2 \frac{|[x_i]_R|}{n} \tag{5.4}$$

where H is called the information entropy of R. It extracts features in the input dataset (such as well logs) that have strong fuzzy relations (i.e. cardinality) with the target variables (such as porosity and flow zone indicator in the case of this study). However, the main drawback of the information entropy algorithm is its sensitivity to the dimensionality of the input data (i.e. the number of attributes) (White and Liu, 1994).

5.4   Hybrid AI Models Development

Nine hybrid AI models were developed for the prediction of Flow Zone Indicator. For these hybrid models, the first stage utilized three feature selection algorithms: Functional Networks, Decision Trees, and Fuzzy Information Entropy. For prediction models in their

second stage, Feedforward Neural Networks, Support Vector Machines, and Type-2 Fuzzy Logic techniques were used.
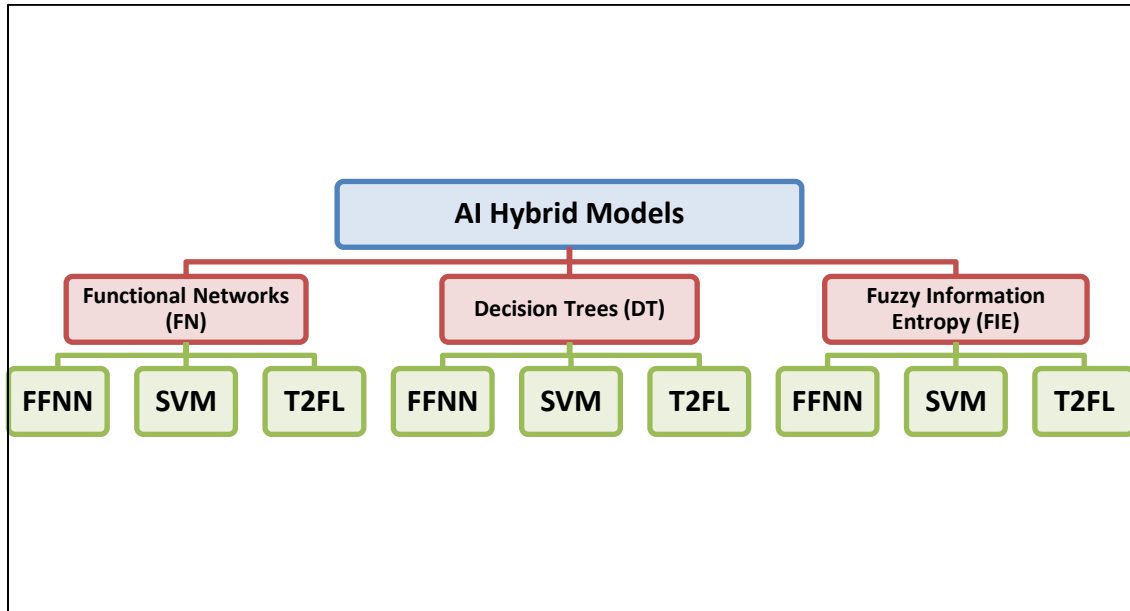
Figure 5.2 - Tree Representing the Hybrid AI Models

5.4.1 Selected Features

The nine hybrid AI models for predicting the flow zone indicator are a combination of the three feature selection algorithms and the three prediction models. This combination is depicted in Figure 5.2. Based on its algorithm, each feature selection method selected a subset of the original 28 data logs and fed the selected ones to the three prediction models.

Algorithm	Functional Networks	Decision Trees	Fuzzy Information Entropy
Selected Features	• Neutron (NEUT) • Bulk Density (DRHO) • CMR Free Fluid • Computed Regularization Parameter Gamma • Integrated Permeability • Irreducible Water Saturation • Total CMR Porosity	• Laterolog Deep Resistivity (LLD) • Neutron (NEUT) • Laterolog Medium Resistivity (LLM) • CMR Porosity using T1/T2 Ratio of Zero • Laterolog Shallow Resistivity (LLS)	• T2 Logarithmic Mean • Spontaneous Potential (SP) • High Resolution Gamma Ray • Integrated Permeability
Total	7	5	4

Looking at Table 5.1, the contribution of NMR data can be easily identified. Five out of seven features selected by Functional Networks algorithm were from NMR. For Decision Trees and Fuzzy Information Entropy algorithms, NMR data contributions to the selected features were at 20% and 75% respectively.

5.4.2 Performance of FFNN Hybrid Models

FFNN was selected as the prediction model for the hybrid neural network because it outperformed both the RBFNN and GRNN models.

Figures 5.3, 5.5, and 5.7 show that training was improved in all hybrid models. In Figures 5.4, 5.6, and 5.8, good correlation for testing between predicted and measured FZI data was clearly observed for FZI values less than 4. Models tend to underestimate FZI for values greater than 4. Figure 5.9 shows that all the three feature selection algorithms

improved the prediction of the FFNN. The highest correlation coefficient for testing, 0.64, was obtained by using the Decision Trees algorithm. Relative improvement was observed in RMSE and MAE for all hybrid models, Figure 5.10 and Figure 5.11.
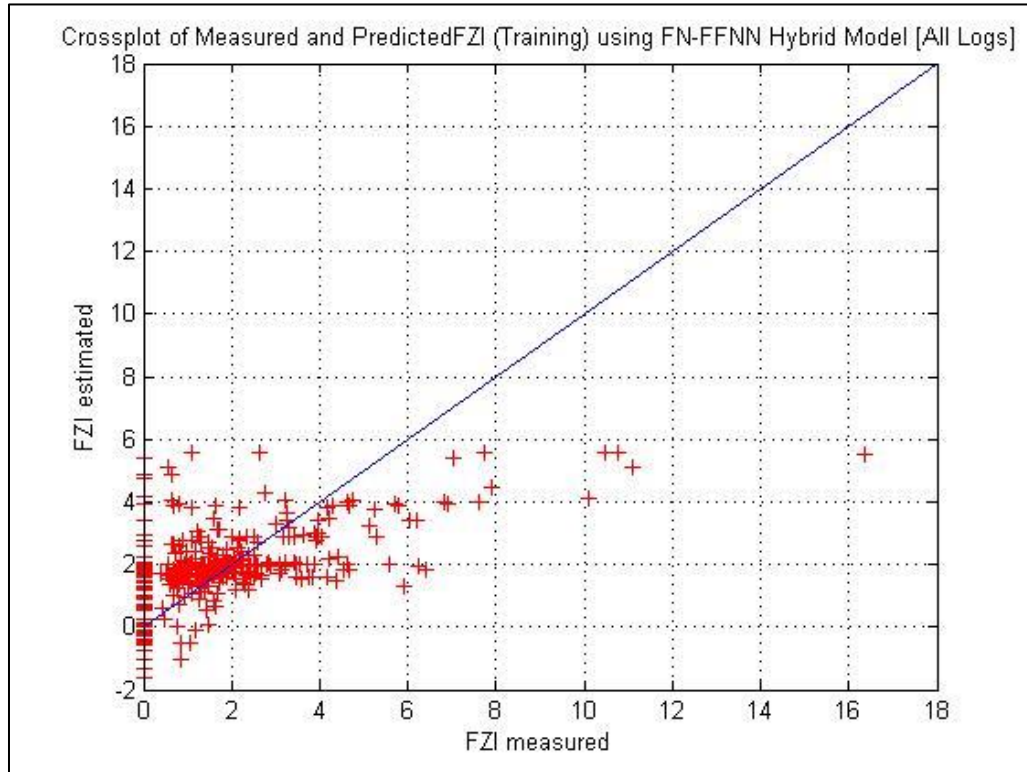
Figure 5.3 - FZI Prediction Crossplot using FN-FFNN (Training)
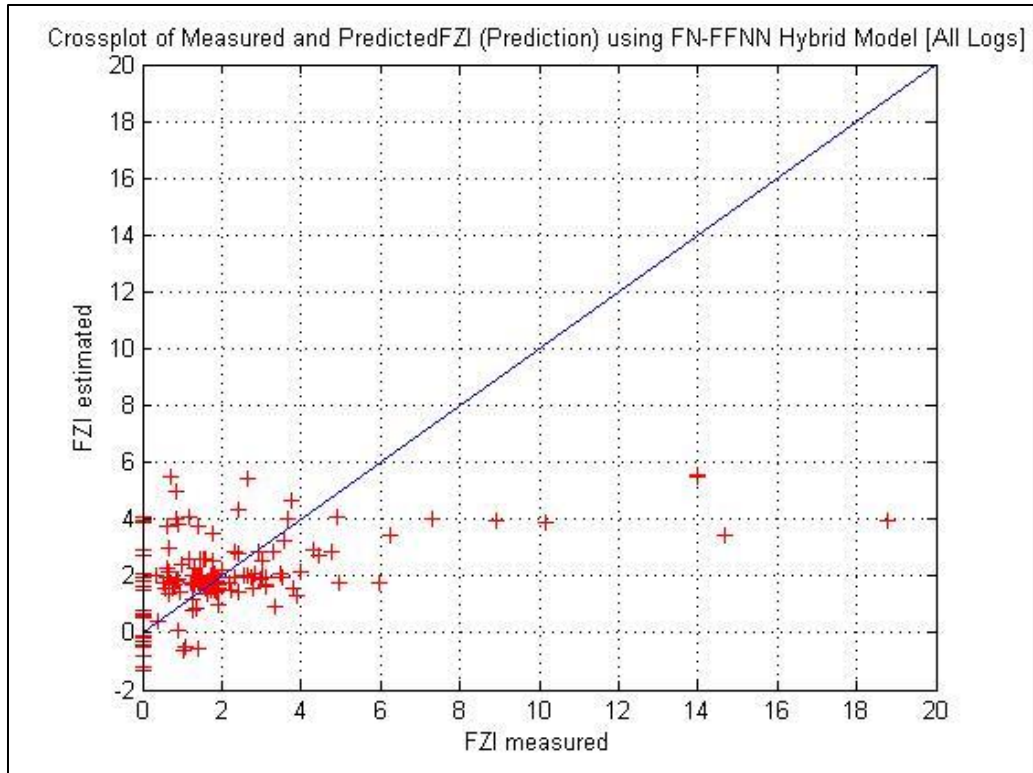
Figure 5.4 – FZI Prediction Crossplot using FN-FFNN (Testing)
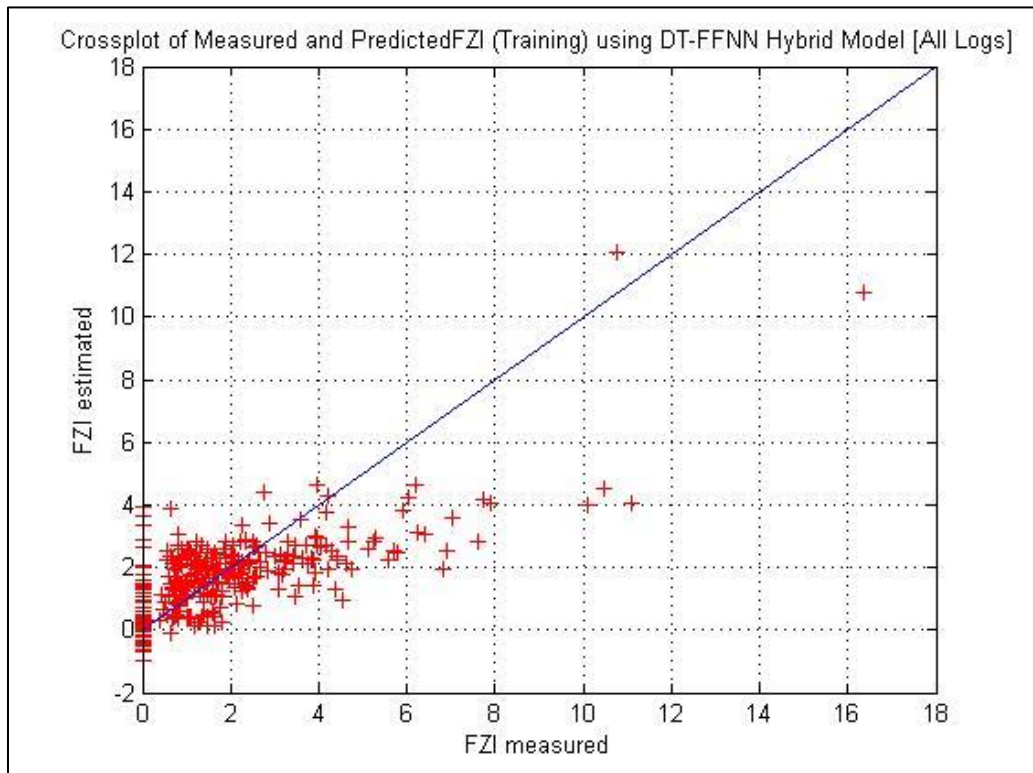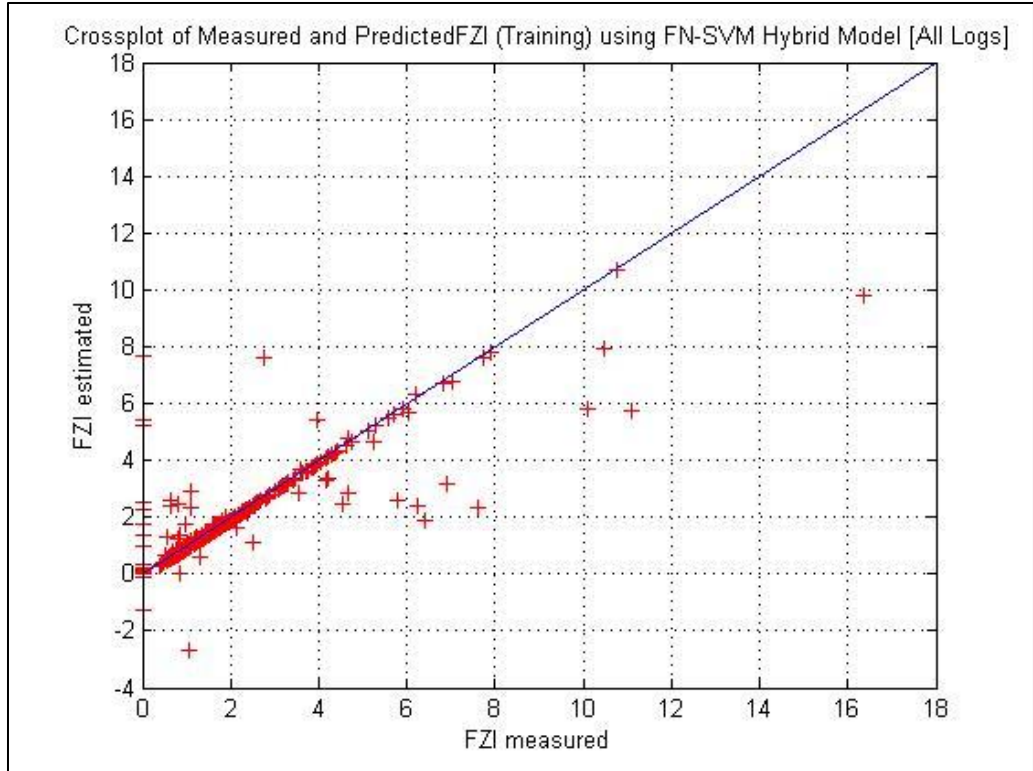

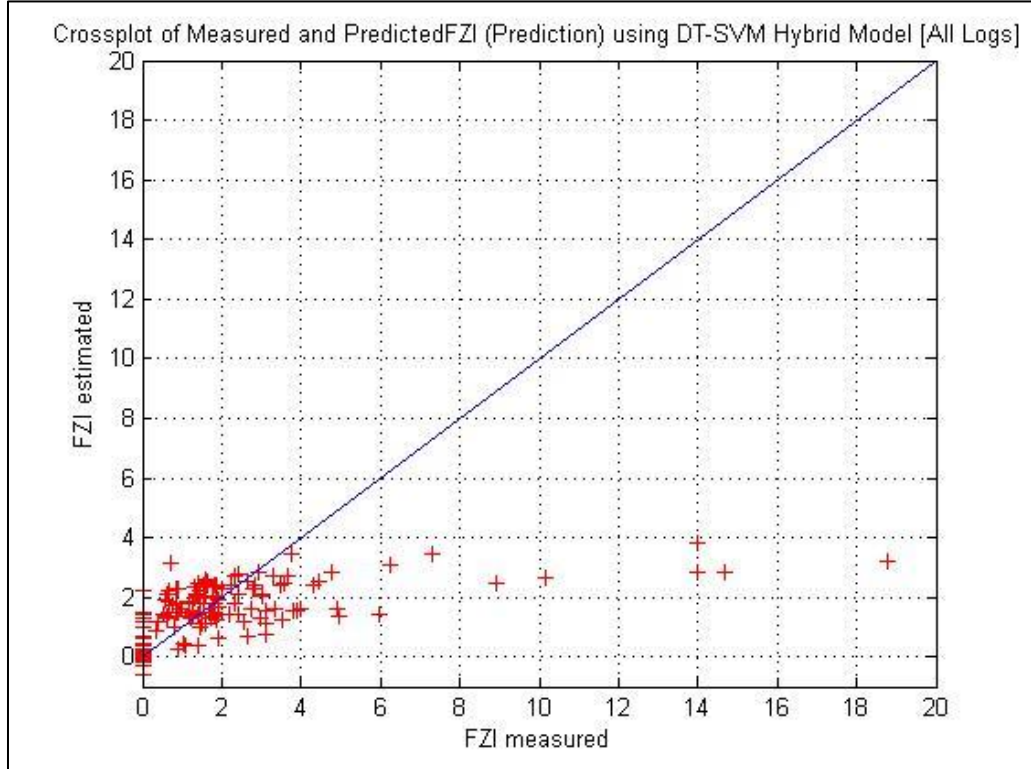**Figure 5.5 - FZI Prediction Crossplot using DT-FFNN (Training)**
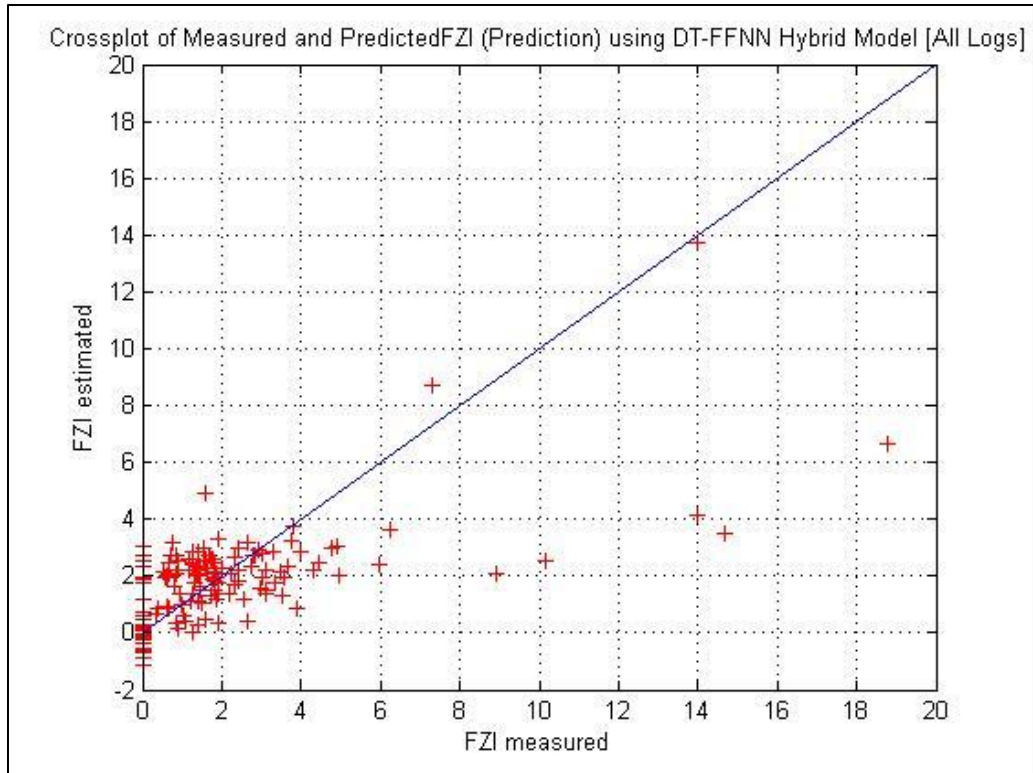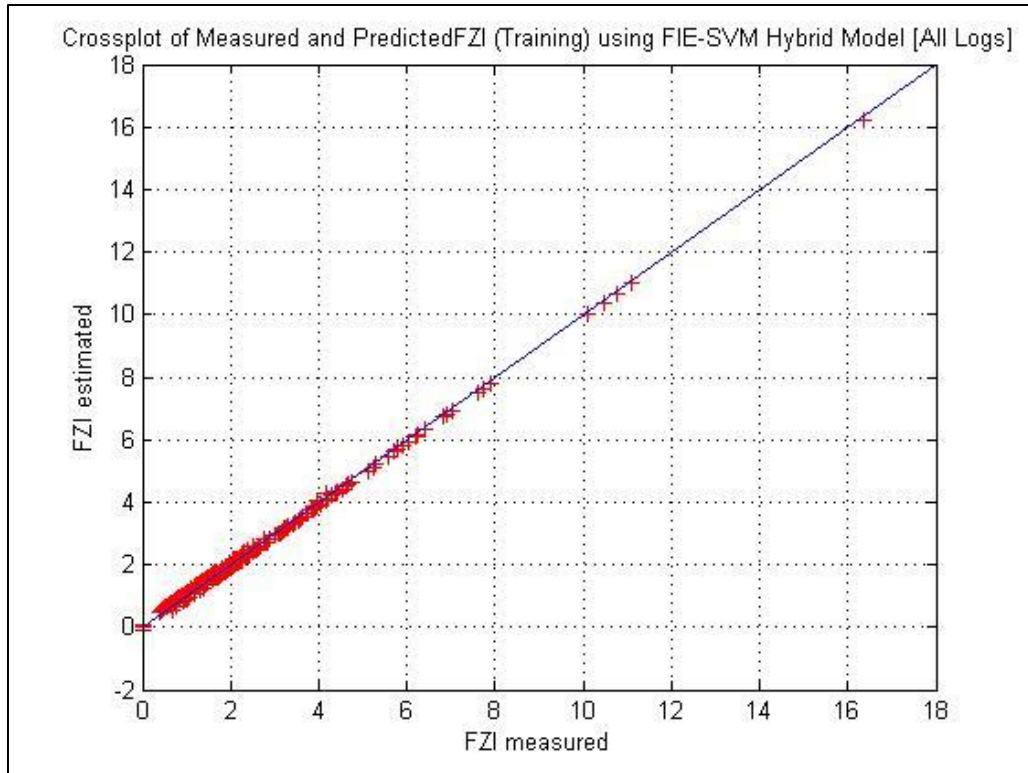
Figure 5.6 – FZI Prediction Crossplot using DT-FFNN (Testing)
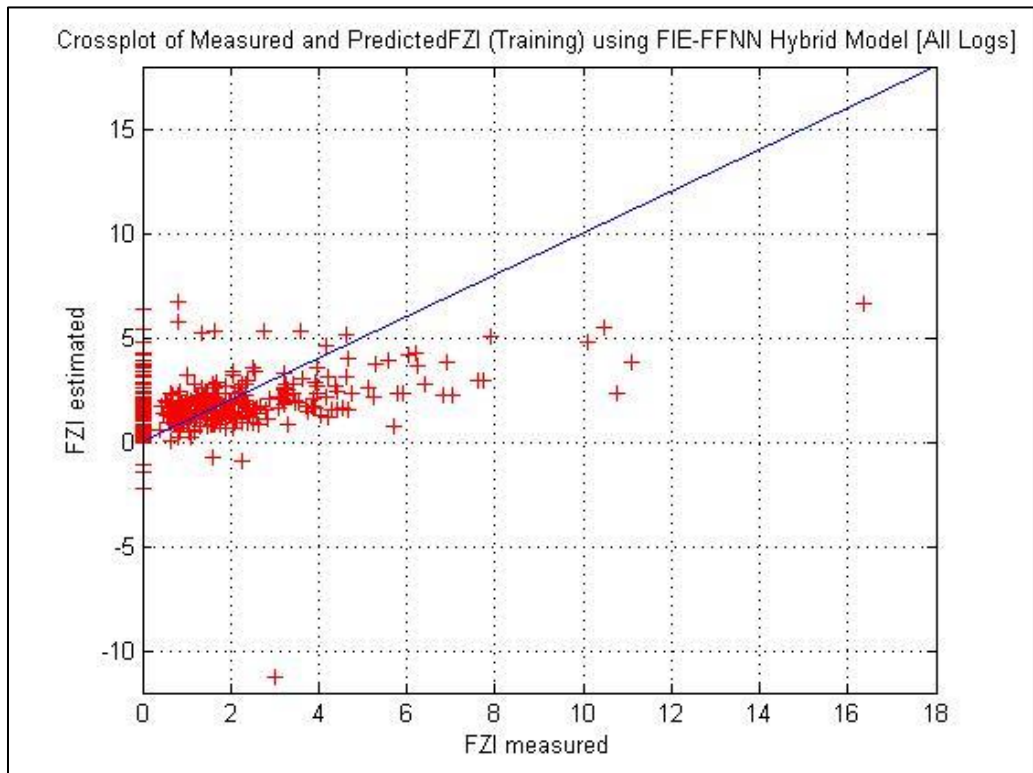
Figure 5.7 - FZI Prediction Crossplot using FIE-FFNN (Training)
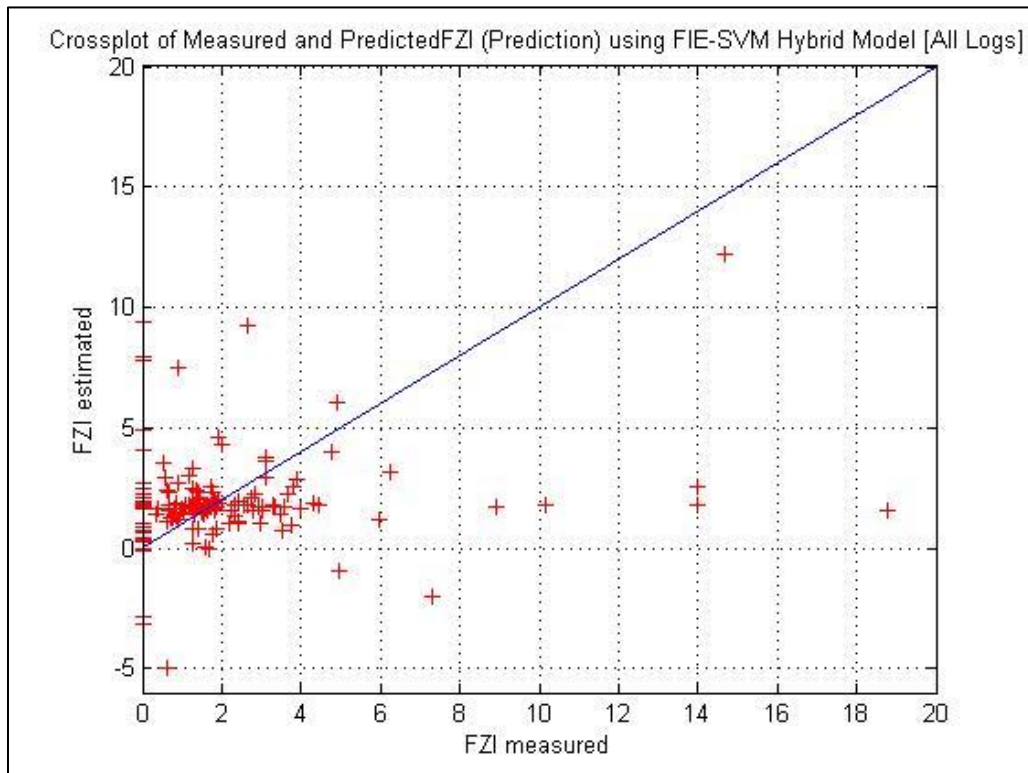
56
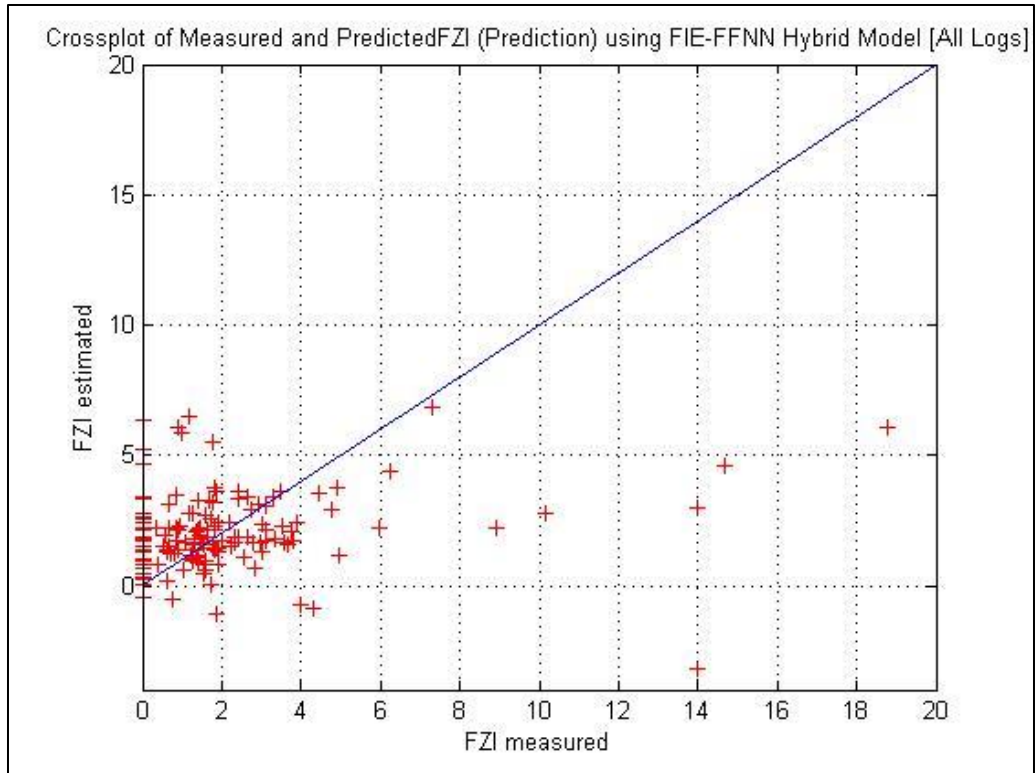
Figure 5.8 – FZI Prediction Crossplot using FIE-FFNN (Testing)
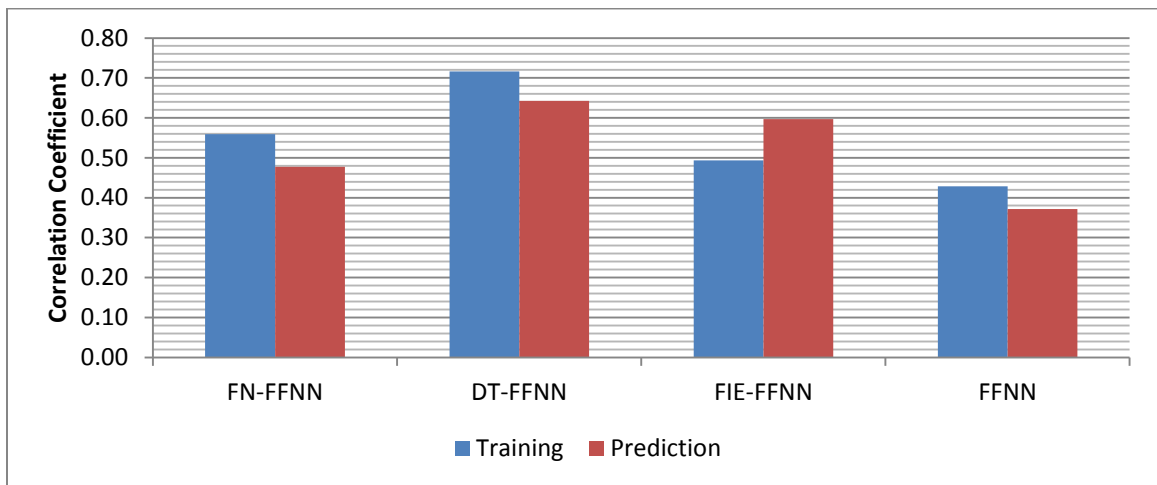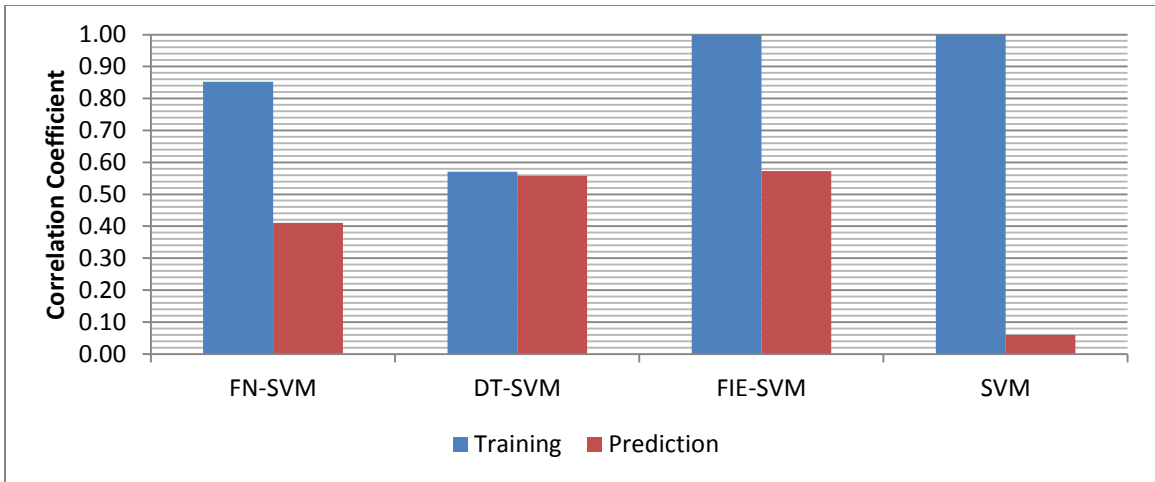

Figure 5.9 - Correlation Coefficient Comparison for FFNN Hybrid AI Models
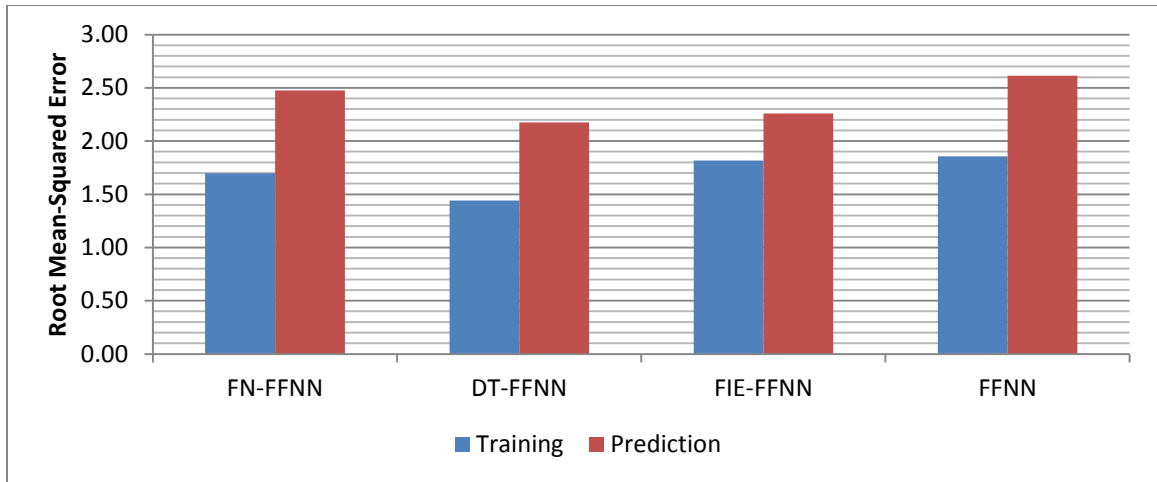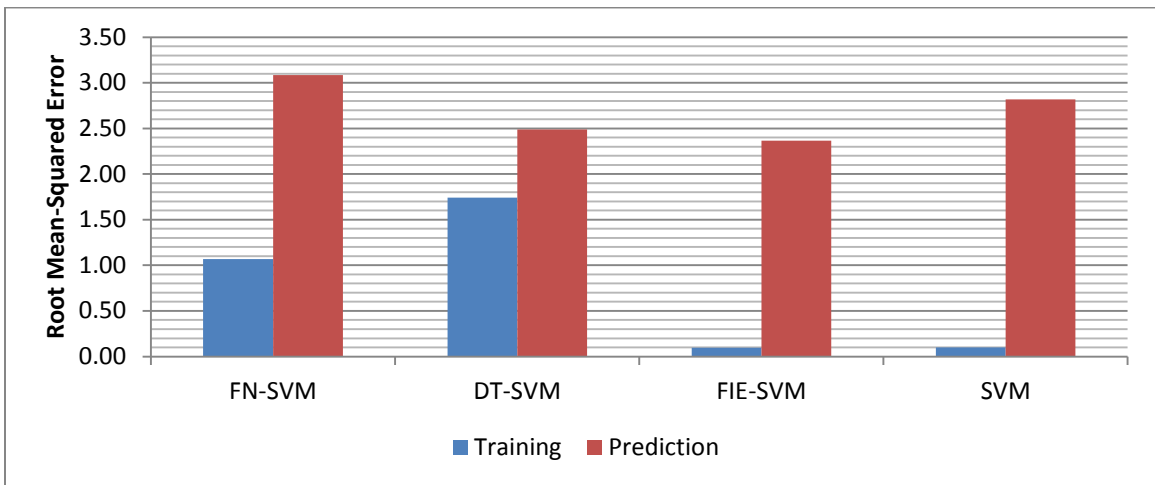
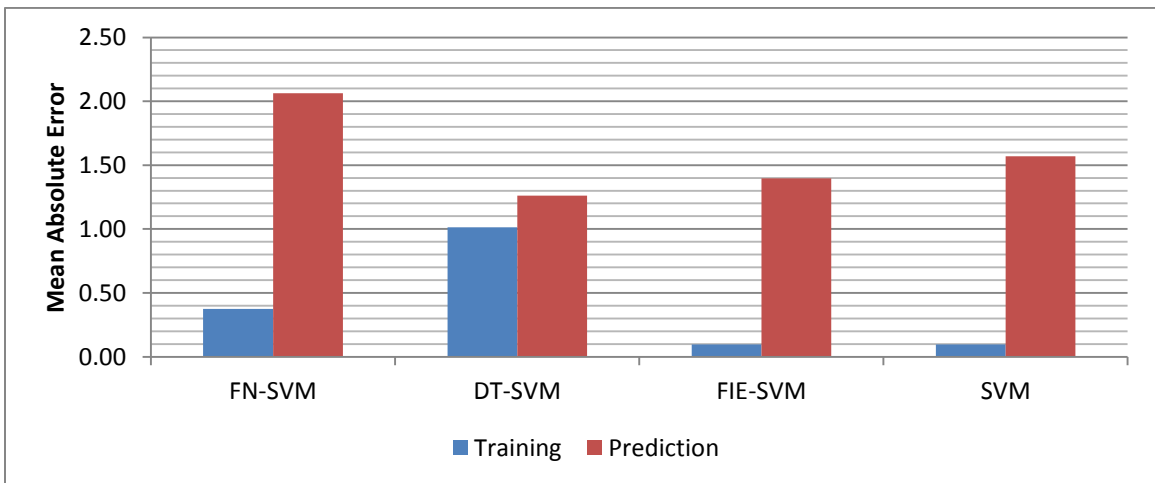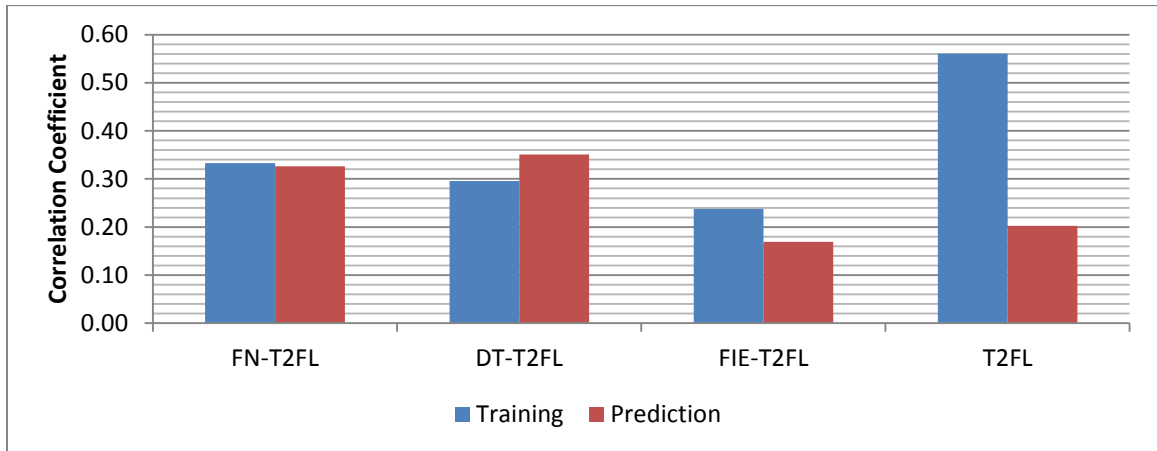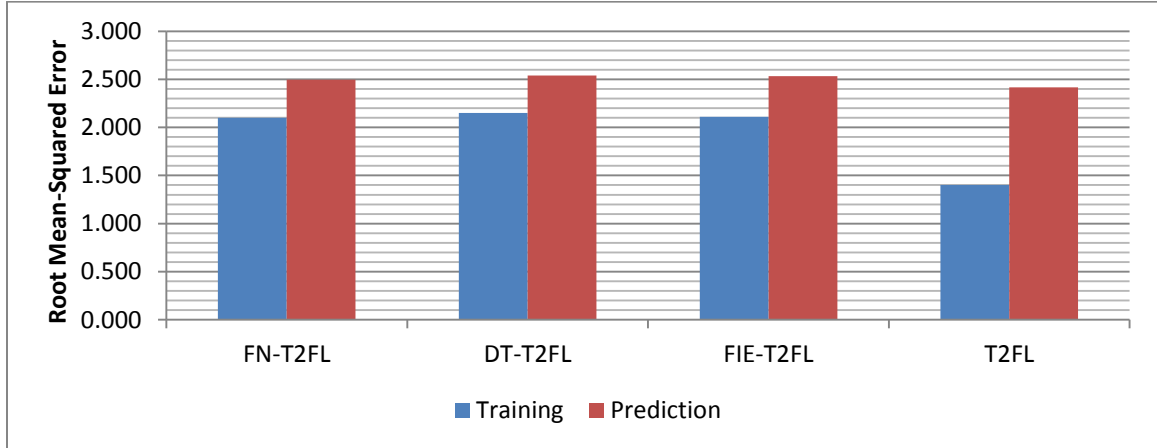**Figure 5.10 - Root Mean-Squared Error Comparison for FFNN Hybrid AI Models**
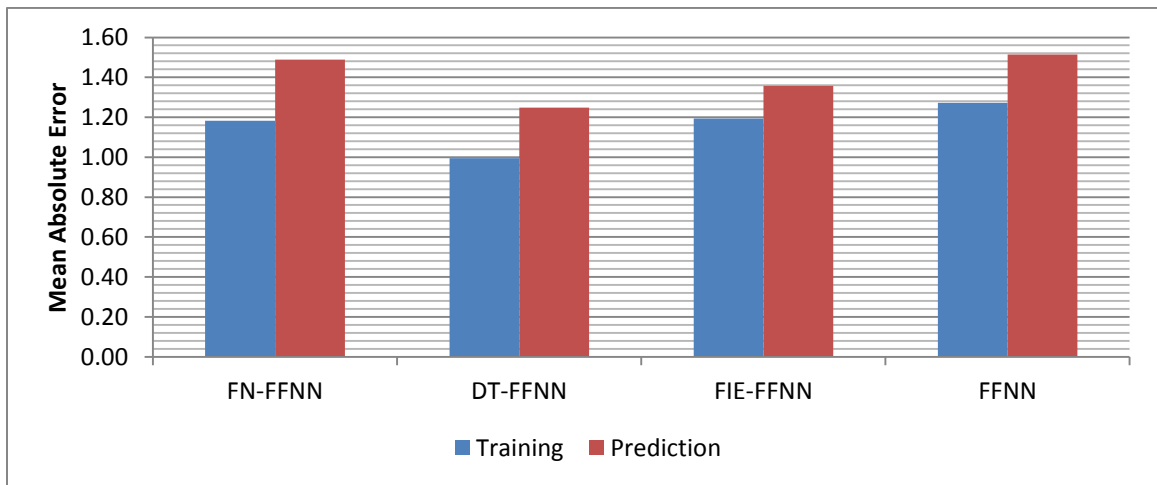
Figure 5.11 - Mean Absolute Error Comparison for FFNN Hybrid AI Models

5.4.3 Performance of SVM Hybrid Models

Figures 5.12, 5.14, and 5.16 show that training was improved in hybrid models except for the FIE-SVM. Figures 5.13, 5.15, and 5.17 show that hybridization greatly improved the prediction of the SVM model. Statistically, the correlation coefficients in Figure 5.18 also confirmed the same observation. Figures 5.19 and 5.20 show relative improvement in RMSE and MAE for Decision Trees and Fuzzy Information Entropy hybrid models.

Figure 5.12 - FZI Prediction Crossplot using FN-SVM (Training)
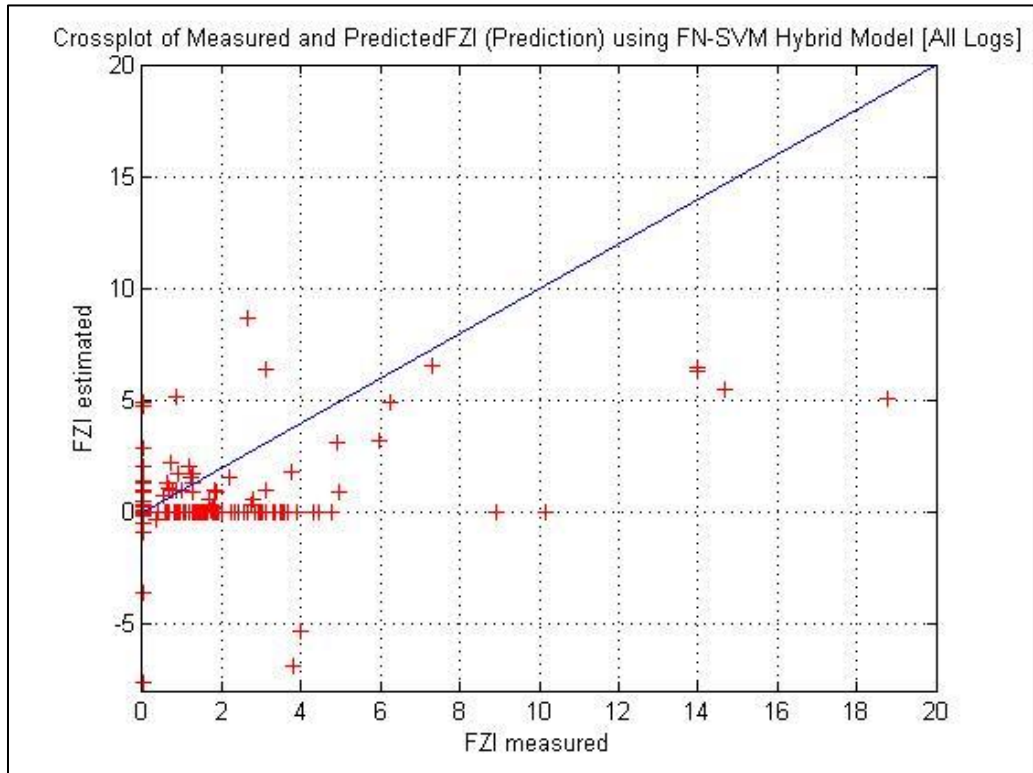
Figure 5.13 - FZI Prediction Crossplot using FN-SVM (Testing)
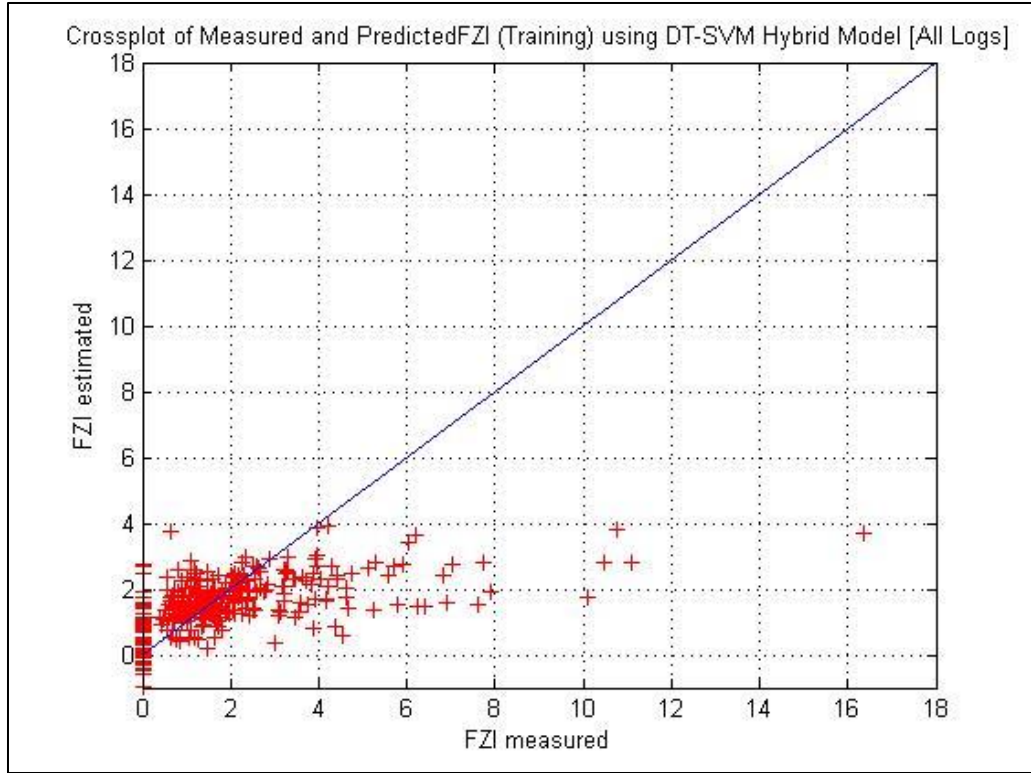
Figure 5.14 - FZI Prediction Crossplot using DT-SVM (Training)

Figure 5.15 - FZI Prediction Crossplot using DT-SVM (Testing)

60

Figure 5.16 - FZI Prediction Crossplot using FIE-SVM (Training)

Figure 5.17 – FZI Prediction Crossplot using FIE-SVM (Testing)

61

Figure 5.18 - Correlation Coefficient Comparison for SVM Hybrid AI Models

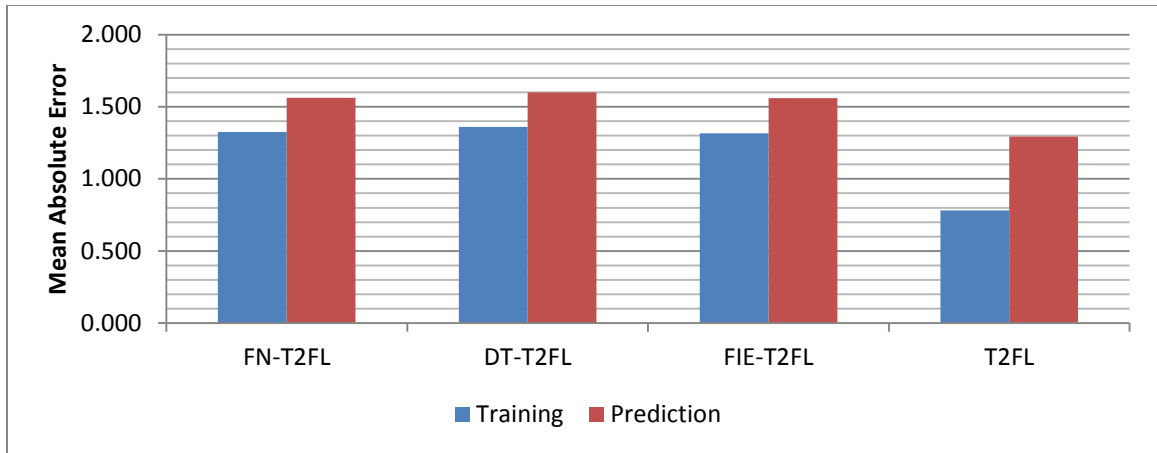**Figure 5.19 - Root Mean-Squared Error Comparison for SVM Hybrid AI Models**

Figure 5.20 - Mean Absolute Error Comparison for SVM Hybrid AI Models

5.4.4 Performance of T2FL Hybrid Models

Hybridization did not work well with Type-2 Fuzzy Logic models. Training performance was not good and prediction was not satisfactory. Correlation coefficients, RMSE, and MAE comparisons are shown in Figures 5.21, 5.22, and 5.23, respectively.
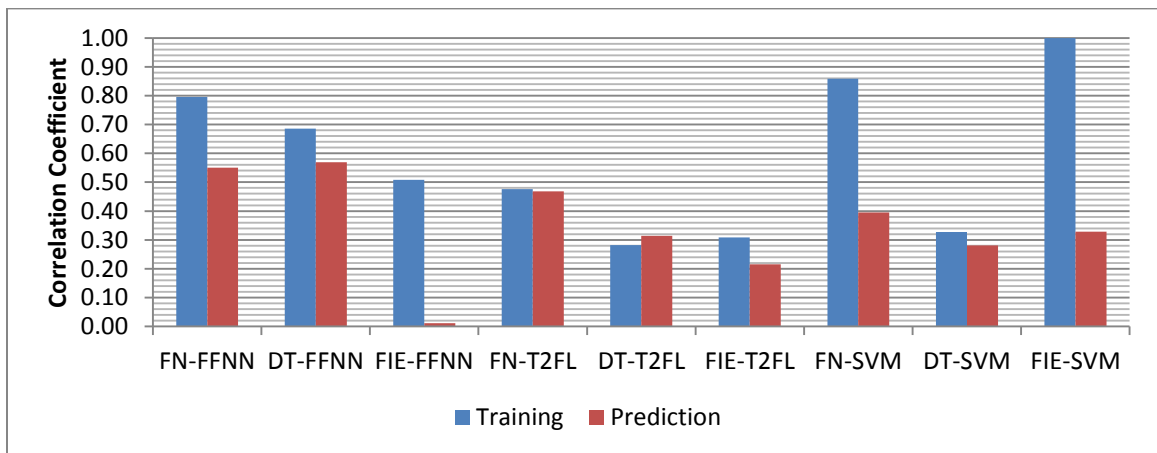

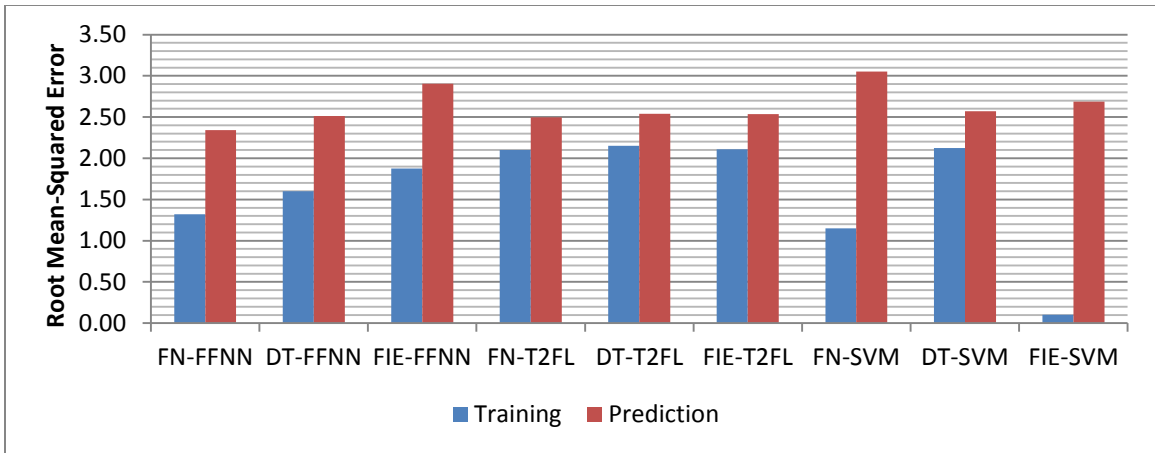**Figure 5.21 - Correlation Coefficient Comparison for T2FL Hybrid AI Models**

**Figure 5.22 - Root Mean-Squared Error Comparison for T2FL Hybrid AI Models**

Figure 5.23 - Mean Absolute Error Comparison for T2FL Hybrid AI Models

5.4.5 Performance Comparison of all Hybrid Models

The comparison for correlation coefficients in Figure 5.24 shows that in general feature selection algorithms improve the performance of the FFNN and SVM predictive models. The DT-FFNN yielded the highest value between the nine hybrid models while the FIE-T2FL had the lowest. Figures 5.25 and 5.26 show the relative variances observed in RMSE and MAE between all hybrid models.

Figure 5.24 - Correlation Coefficient Comparison for all Hybrid AI Models

Figure 5.25 - Root Mean-Squared Error Comparison for all Hybrid AI Models
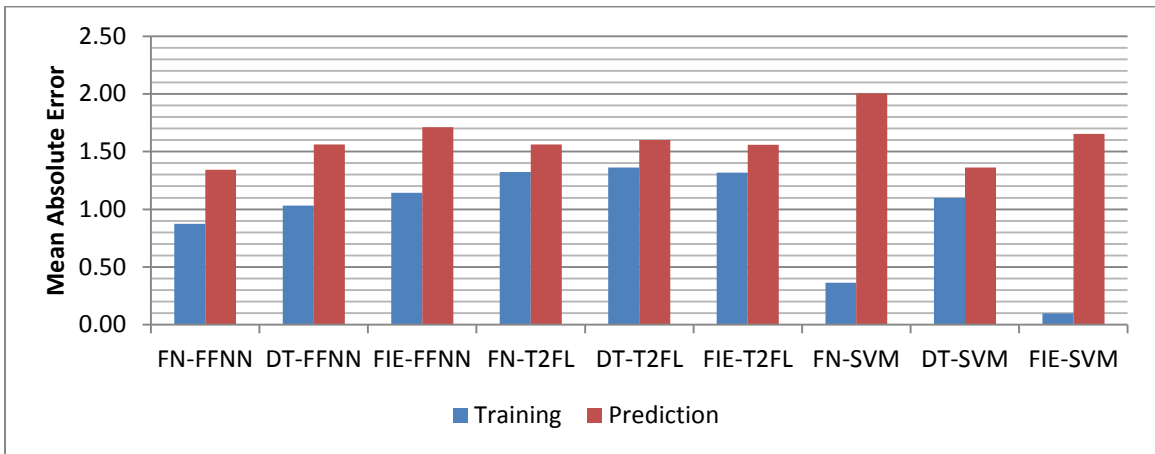
Figure 5.26 - Mean Absolute Error Comparison for all Hybrid AI Models

CHAPTER 6

Model Improvement through Data Fusion

6.1   Introduction

Data Fusion is the process of utilizing combinations of readily-available data and related concepts as additional inputs to improve existing models. Three different combinations were utilized for the sake of improving the previously developed hybrid models.

In Chapter 5, the development of nine hybrid AI models for FZI prediction was covered. Models performance and results were discussed and compared with each other. For performance improvement using data fusion, FFNN was selected as the predictive model. This chapter presents and discusses the results of data fusion application to the FFNN hybrid models.

6.2   Global Hydraulic Elements (GHEs) Classification as Input

Core values for porosity and permeability were utilized to put all data points on the GHE basemap, as shown in Figure 6.1. Then, code was developed to classify each data point by assigning the GHE numerical value to it. The newly obtained information was then added to the input dataset for all nine hybrid models.

Results showed that the new data column was selected by all the features selection algorithms. The cross-plots in Figures 6.2 − 6.10 show that better correlations were obtained for almost all models. This improvement was more obvious in models with Functional Networks and Decision Trees features selection algorithms. Figures 6.11 and

6.12 show the performance comparison for training and testing respectively, in terms of correlation coefficient, between the new hybrid models with Fusion AI and the previous ones without it. It confirmed the same observation from the cross-plots.
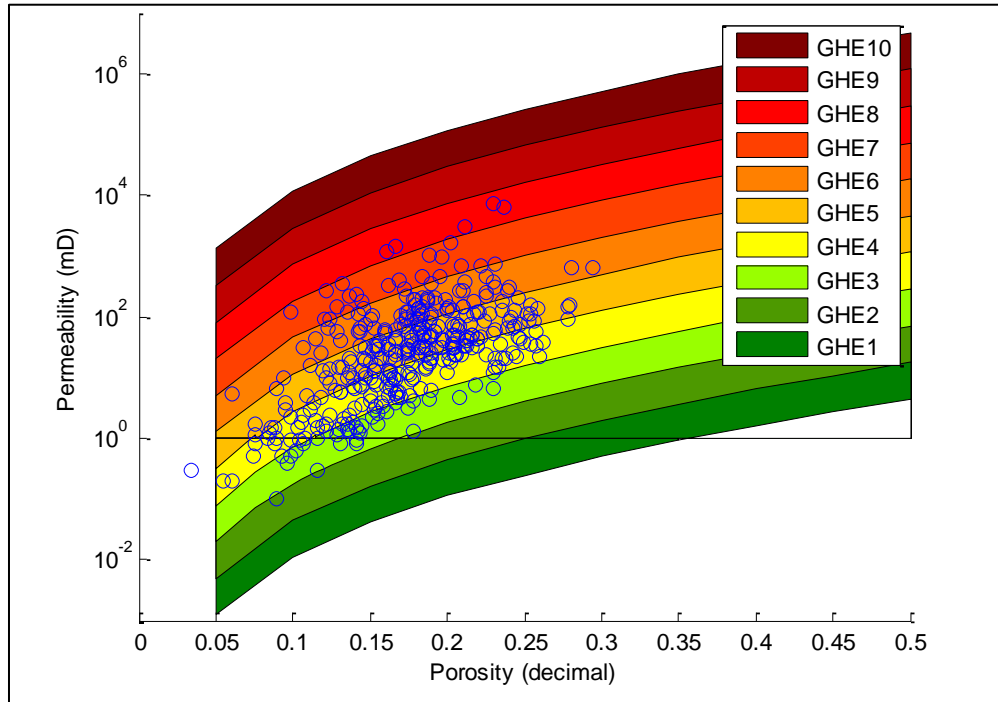

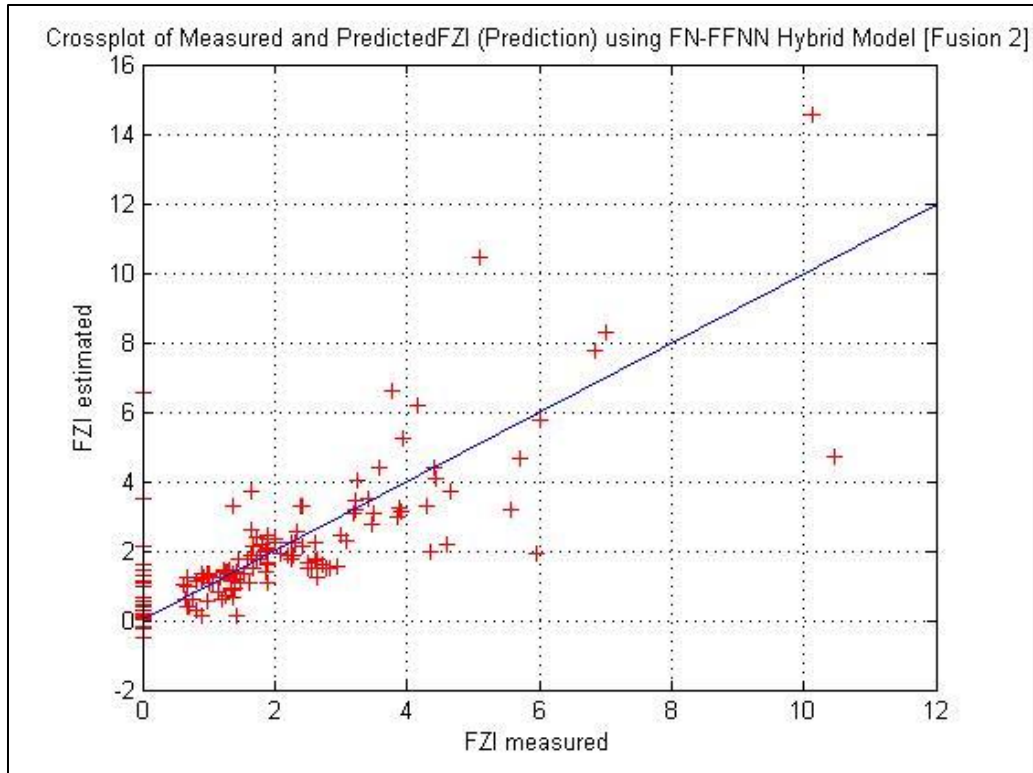**Figure 6.1 - Data Points on GHE Basemap**

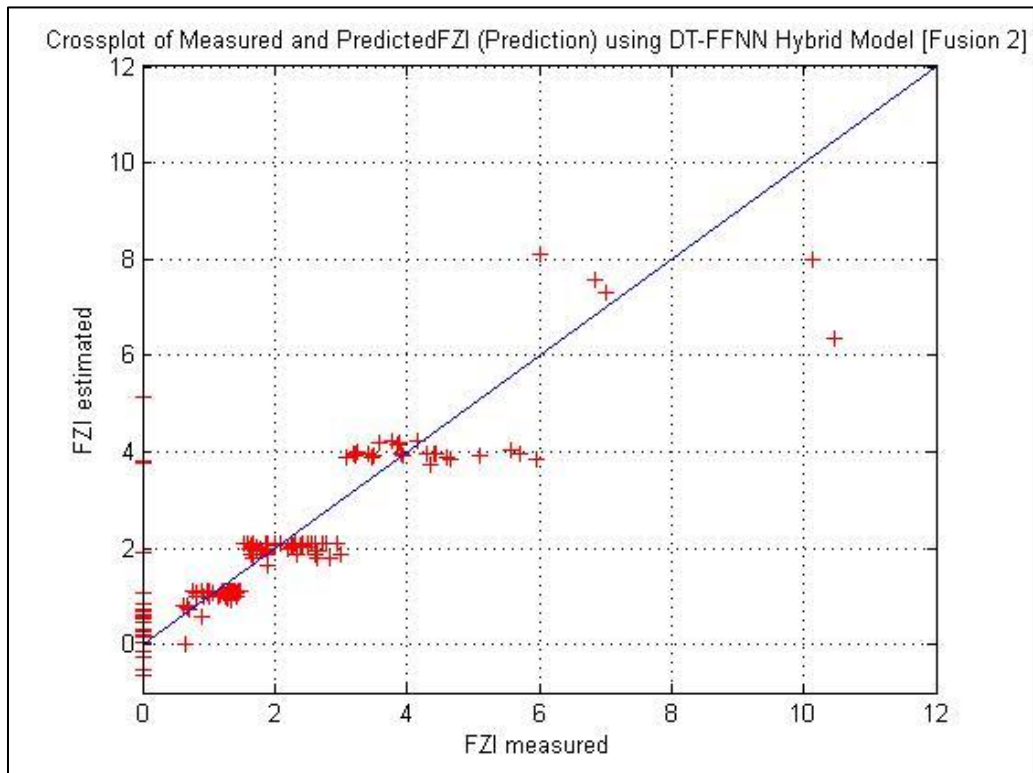**Figure 6.2 - FZI Prediction Crossplot using FN-FFNN (GHE Fusion AI)**

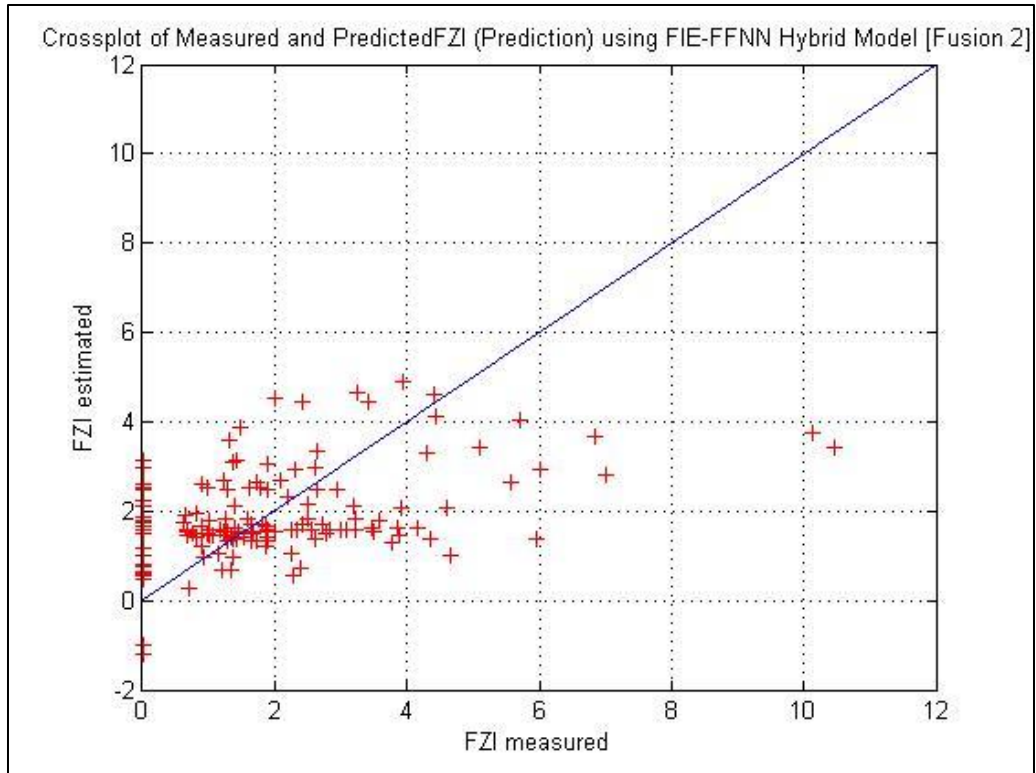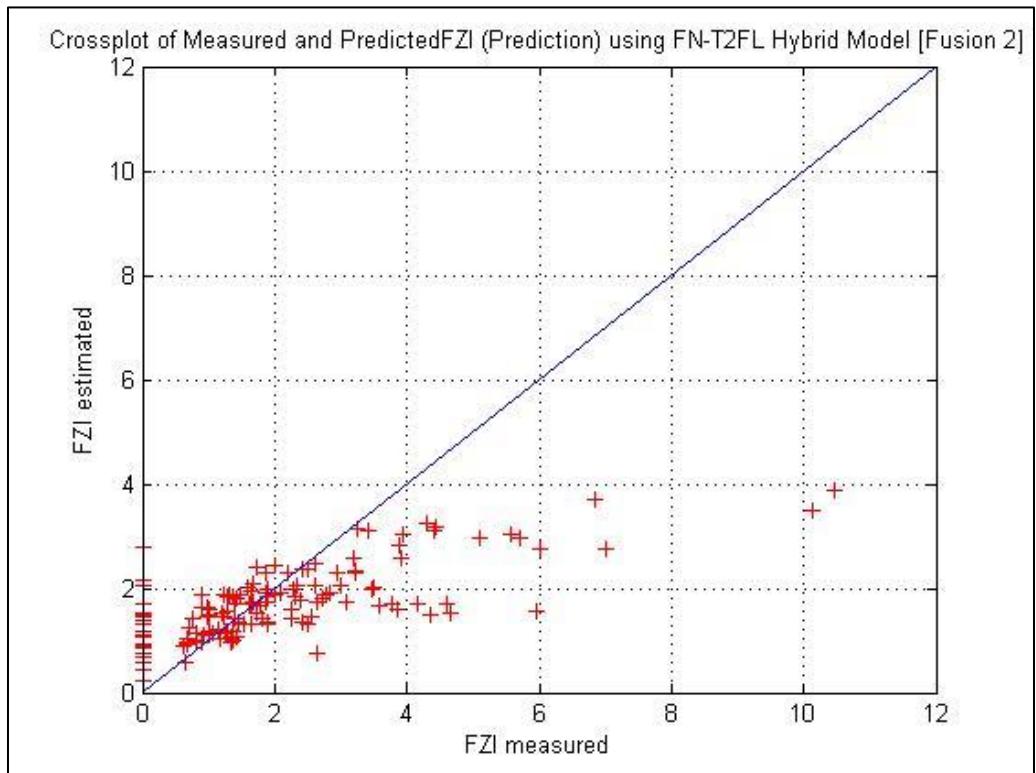**Figure 6.3 - FZI Prediction Crossplot using DT-FFNN (GHE Fusion AI)**

68

Figure 6.4 - FZI Prediction Crossplot using FIE-FFNN (GHE Fusion AI)

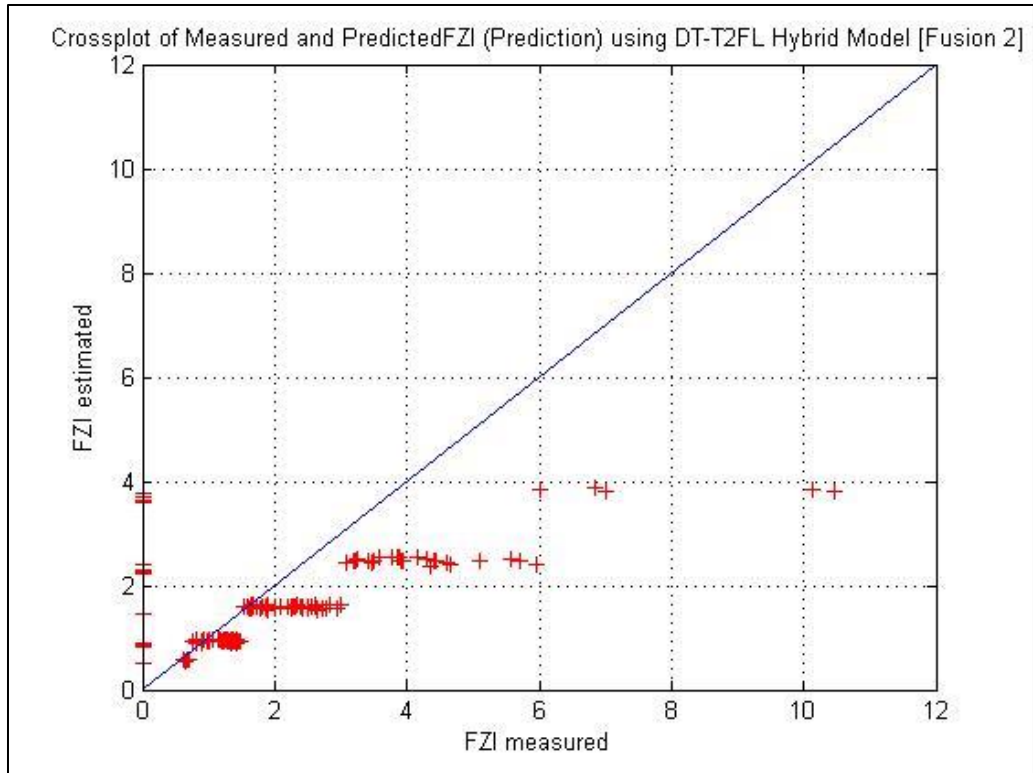**Figure 6.5 - FZI Prediction Crossplot using FN-T2FL (GHE Fusion AI)**

69

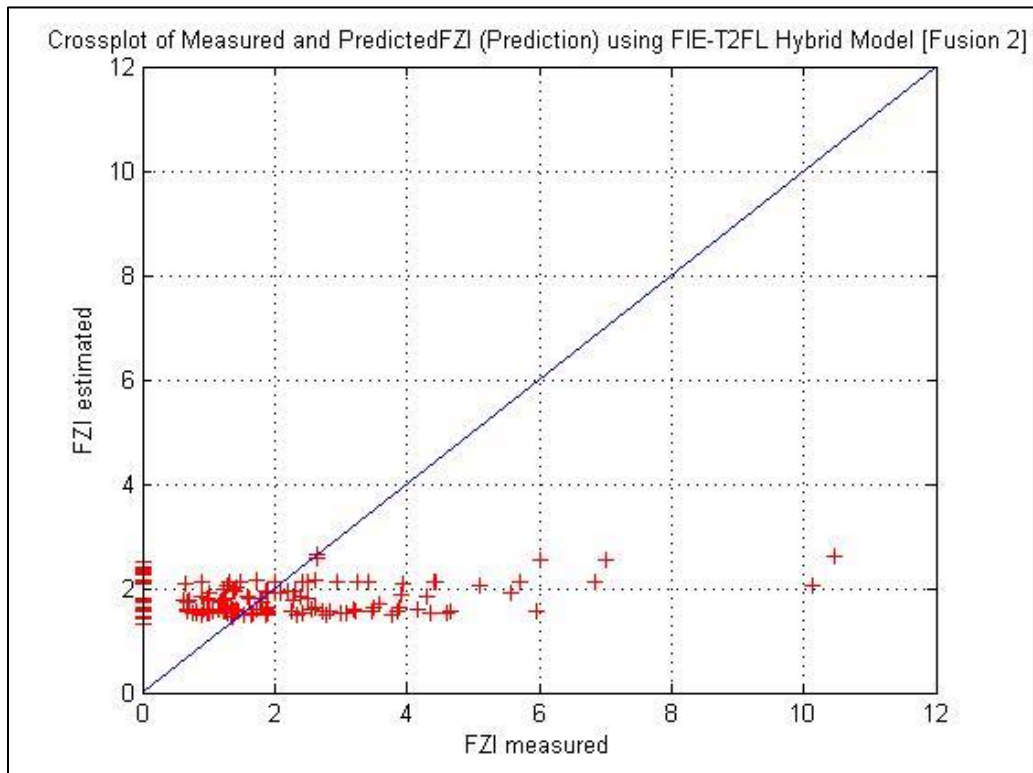**Figure 6.6 - FZI Prediction Crossplot using DT-T2FL (GHE Fusion AI)**

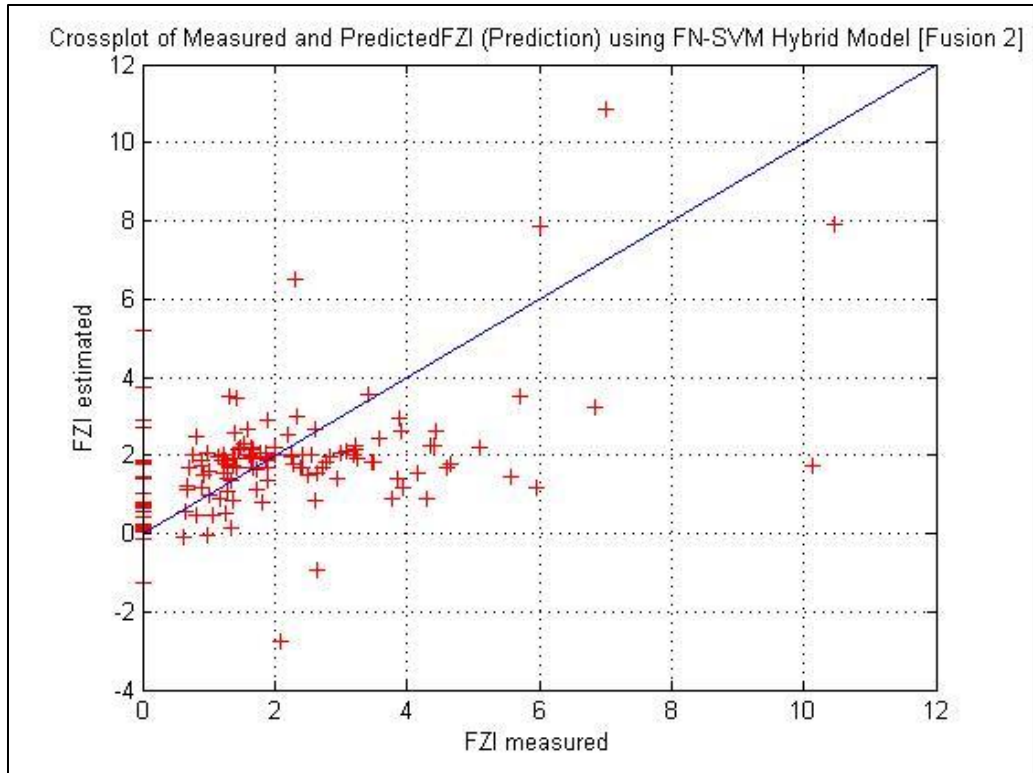**Figure 6.7 - FZI Prediction Crossplot using FIE-T2FL (GHE Fusion AI)**

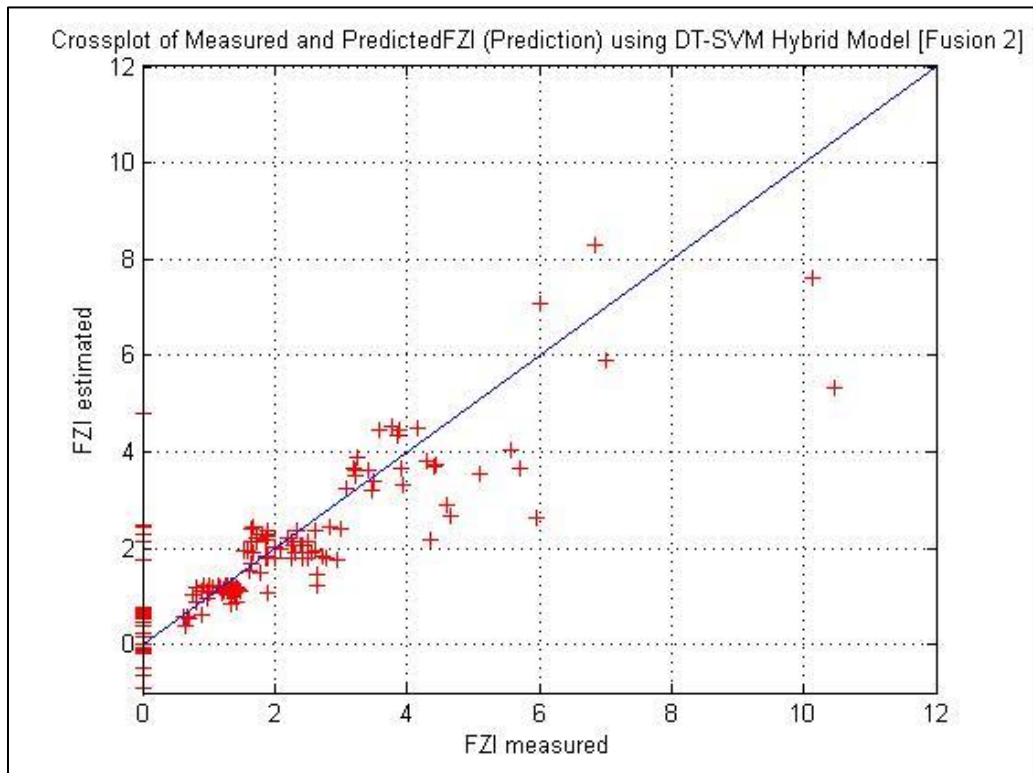**Figure 6.8 - FZI Prediction Crossplot using FN-SVM (GHE Fusion AI)**

Figure 6.9 - FZI Prediction Crossplot using DT-SVM (GHE Fusion AI)
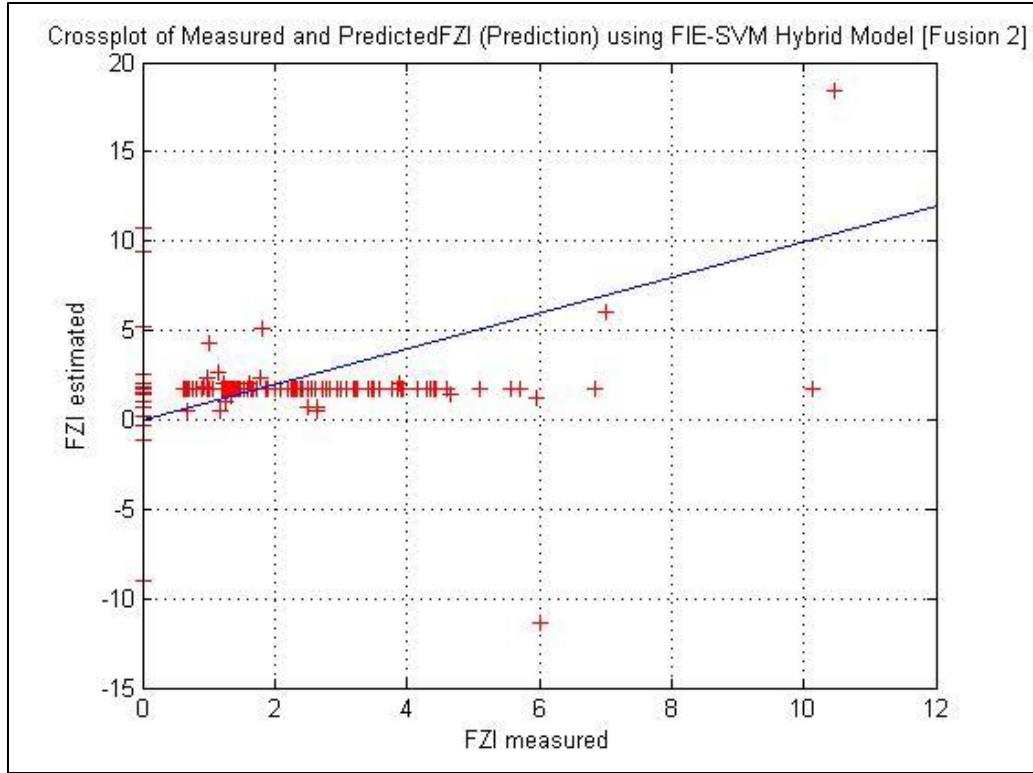
71

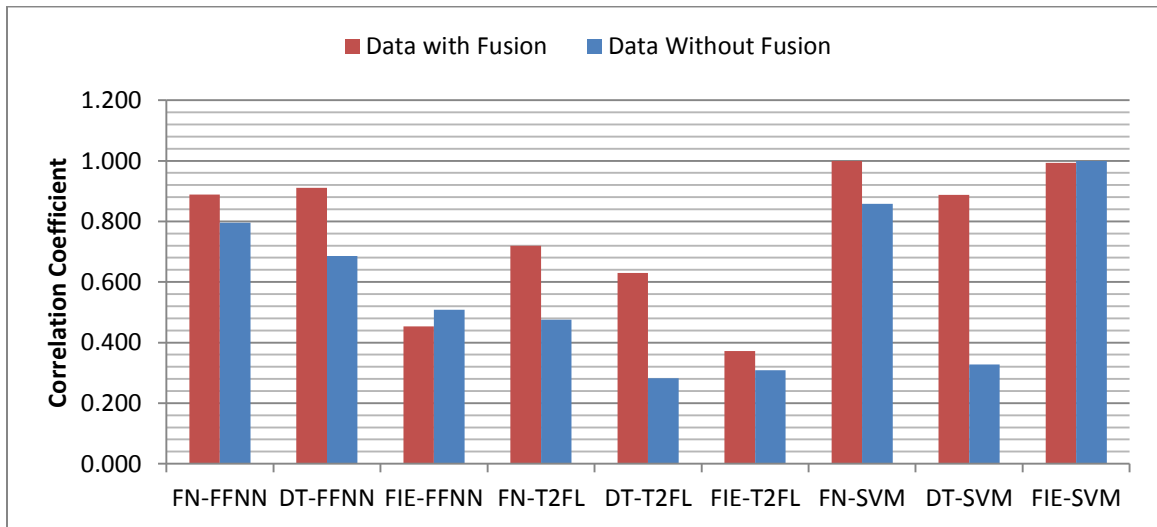**Figure 6.10 - FZI Prediction Crossplot using FIE-SVM (GHE Fusion AI)**

Figure 6.11 - Correlation Coefficient Comparison for all Hybrid AI Models (GHE Fusion AI) [Training]
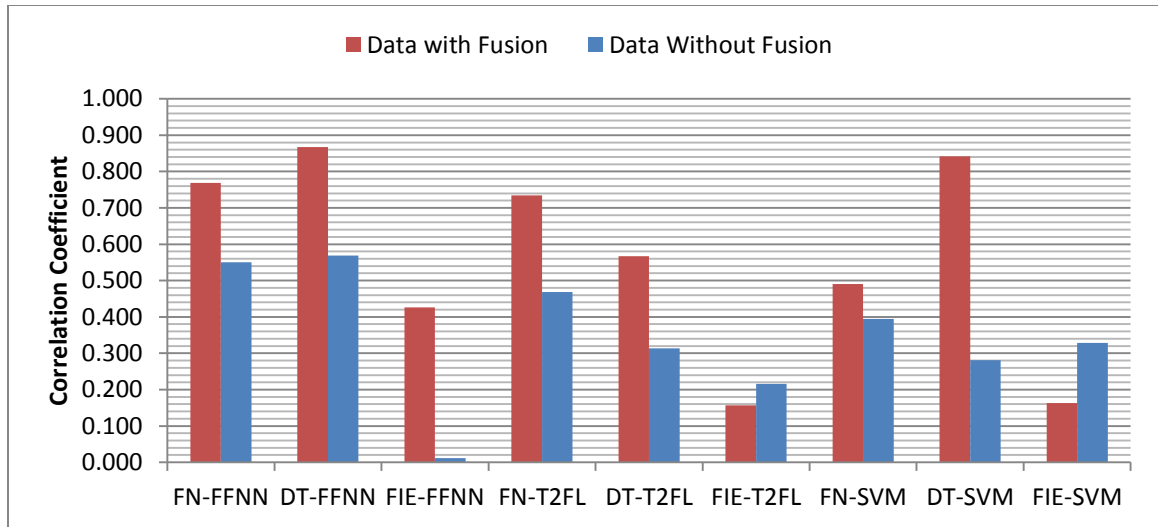
Figure 6.12 - Correlation Coefficient Comparison for all Hybrid AI Models (GHE Fusion AI) [Testing]

6.3 Flow Zone Indicator Components as Input

Equation 2.7 has two components: reservoir quality index (RQI) and pore volume to solid volume ratio. Core values for porosity and permeability were used to compute them without conversion factors or constants. The two additional input parameters were $(\sqrt{K/\varphi})$ and $(\frac{\varphi}{1-\varphi})$.

Functional Networks algorithm selected both parameters while the other two algorithms only selected the $(\sqrt{K/\varphi})$ parameter. Figures 6.13 – 6.21 show that the inclusion of the two parameters improved the correlation between the predicted and measured values. Figures 6.22 and 6.23 show that for all models there were improved results for the correlation coefficient.
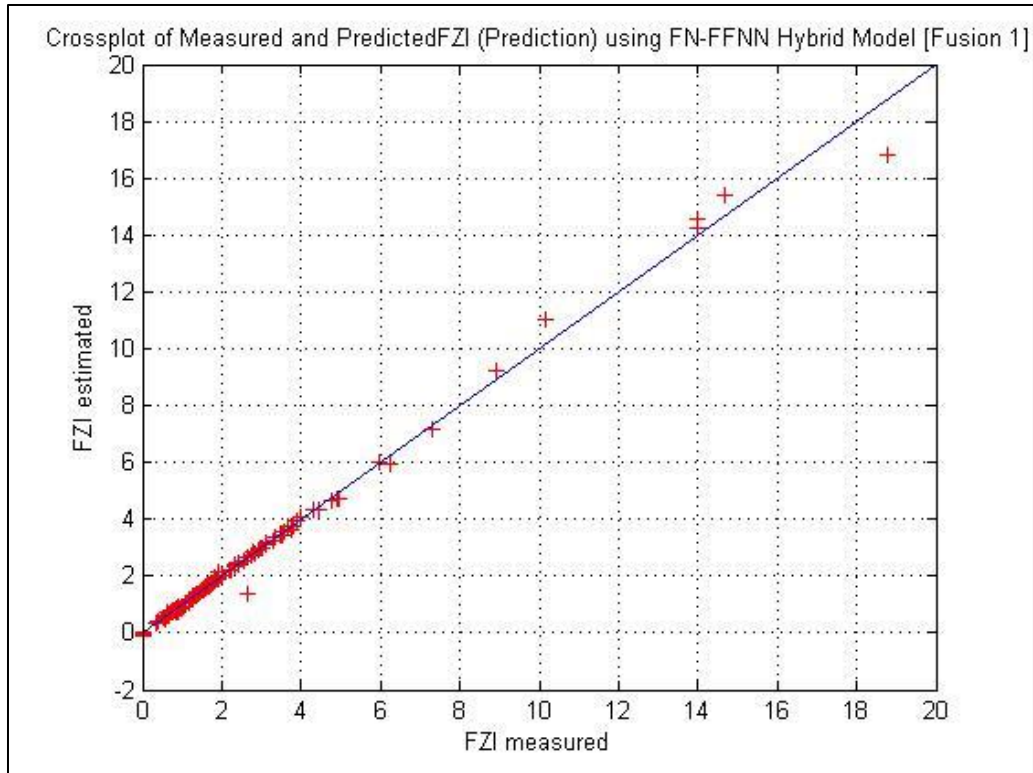
73

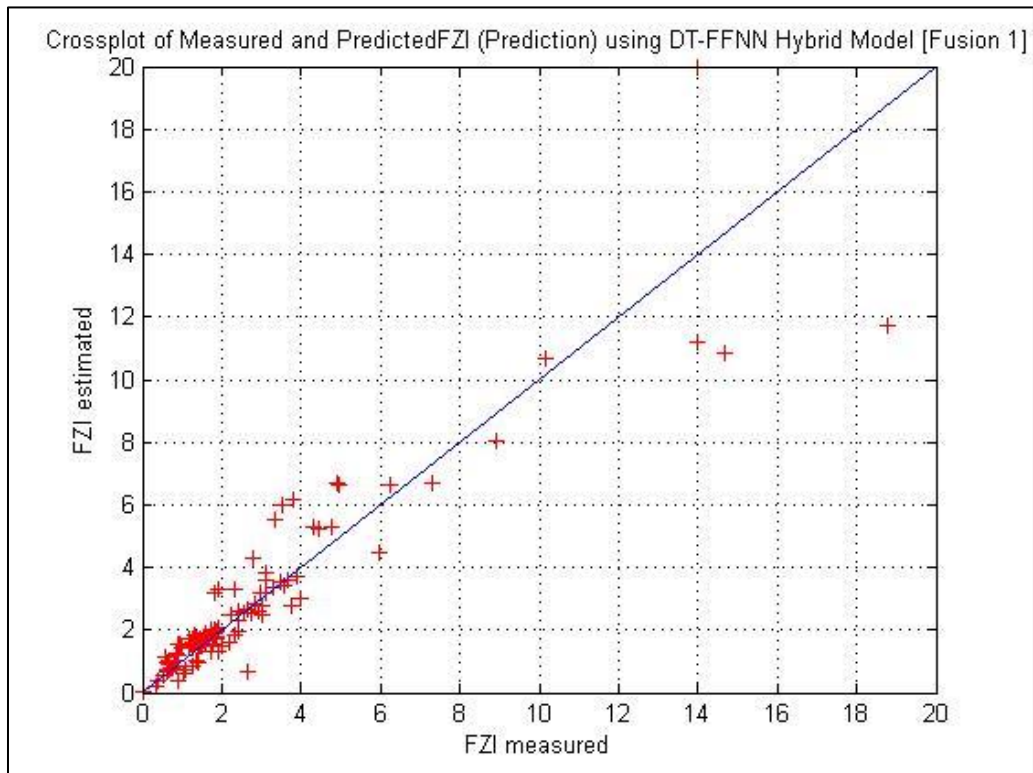**Figure 6.13 - FZI Prediction Crossplot using FN-FFNN (FZI Components Fusion AI)**

Figure 6.14 - FZI Prediction Crossplot using DT-FFNN (FZI Components Fusion AI)
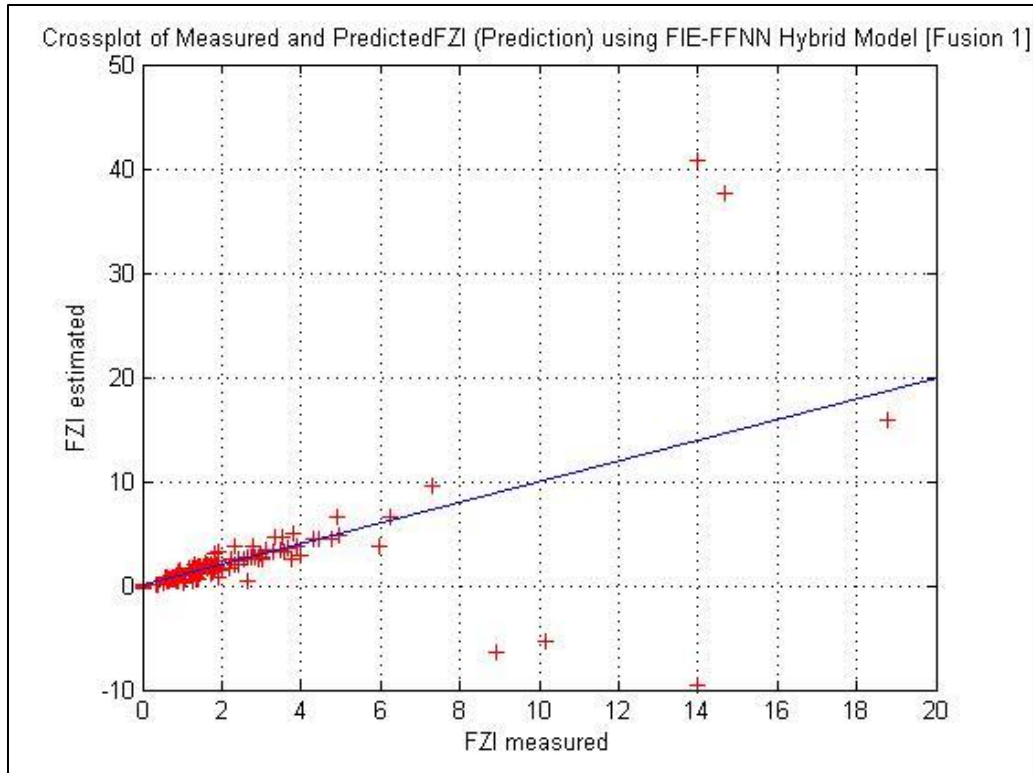
74

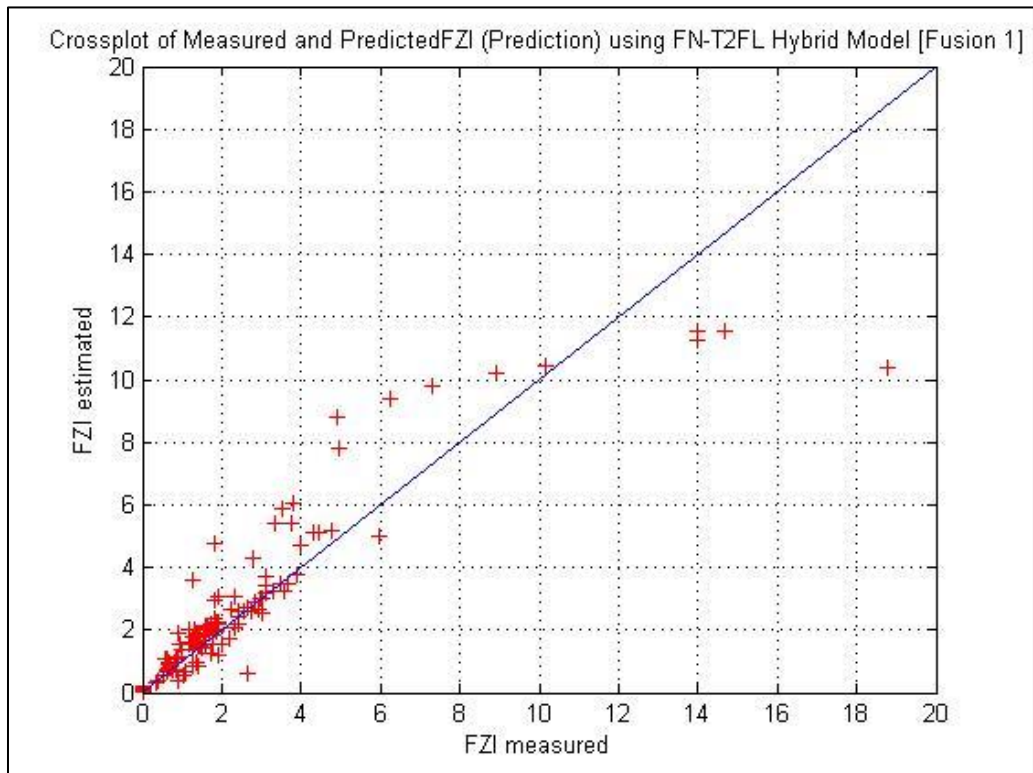**Figure 6.15 - FZI Prediction Crossplot using FIE-FFNN (FZI Components Fusion AI)**

Figure 6.16 - FZI Prediction Crossplot using FN-T2FL (FZI Components Fusion AI)
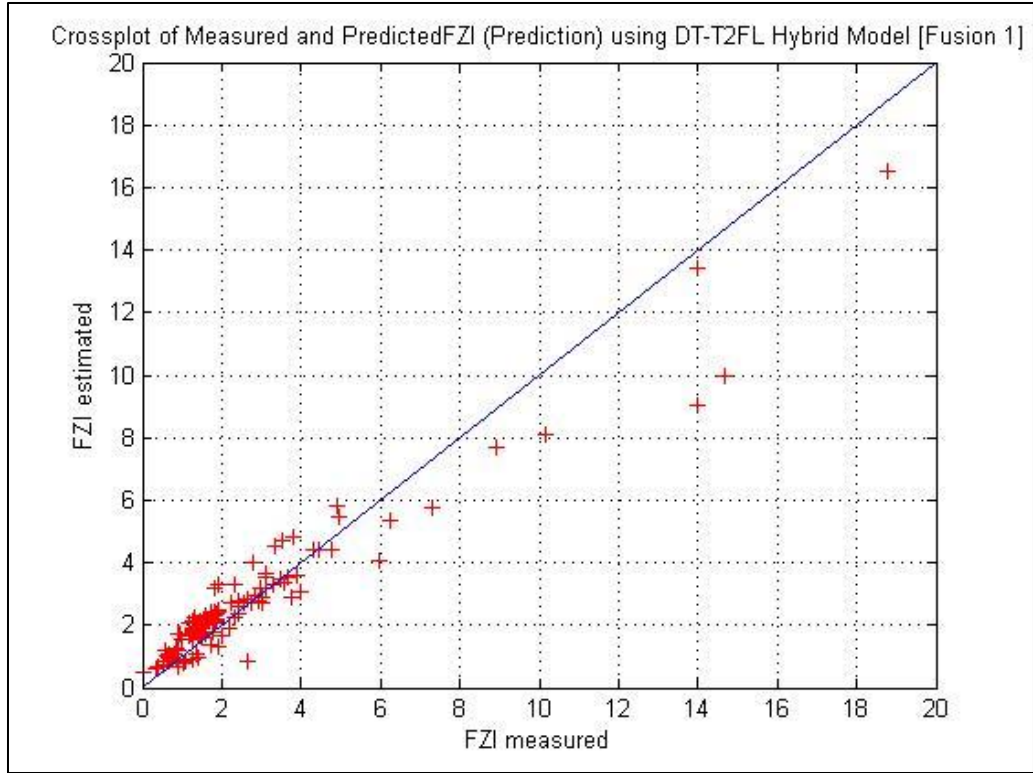
Figure 6.17 - FZI Prediction Crossplot using DT-T2FL (FZI Components Fusion AI)

Figure 6.18 - FZI Prediction Crossplot using FIE-T2FL (FZI Components Fusion AI)

Figure 6.19 - FZI Prediction Crossplot using FN-SVM (FZI Components Fusion AI)

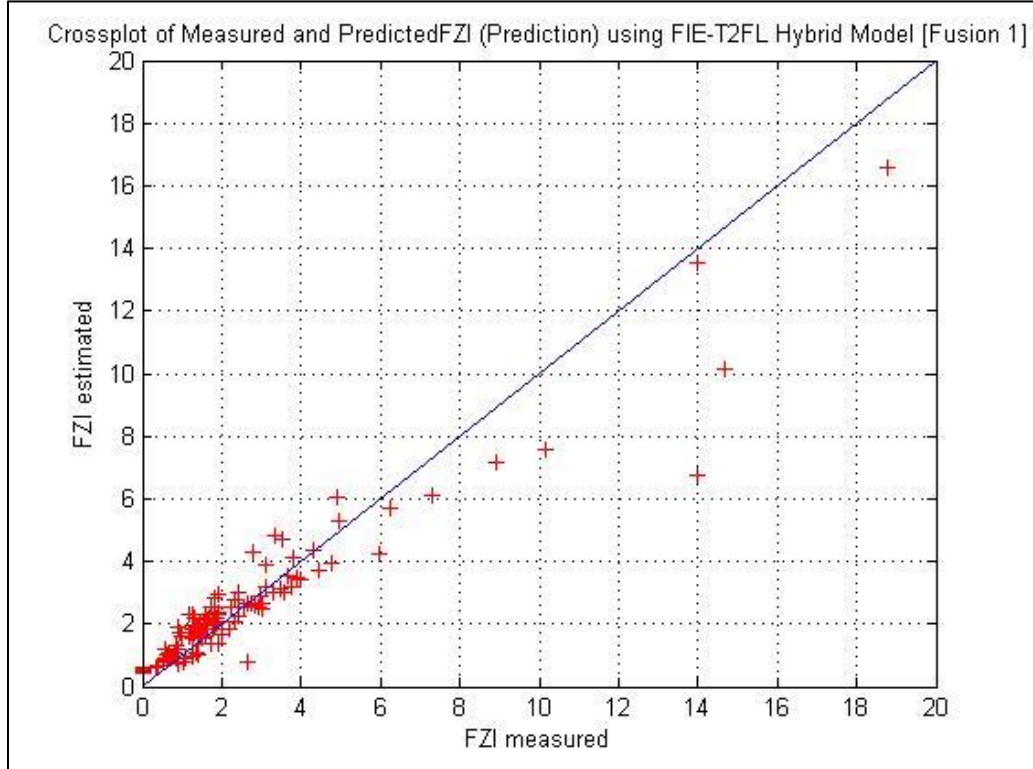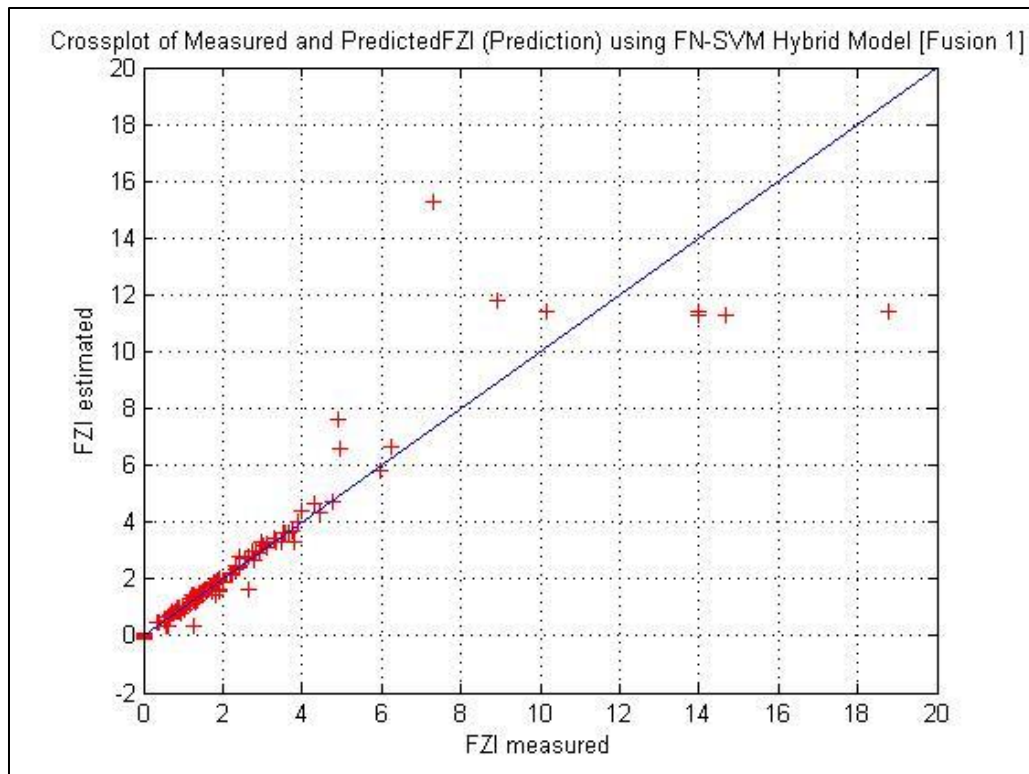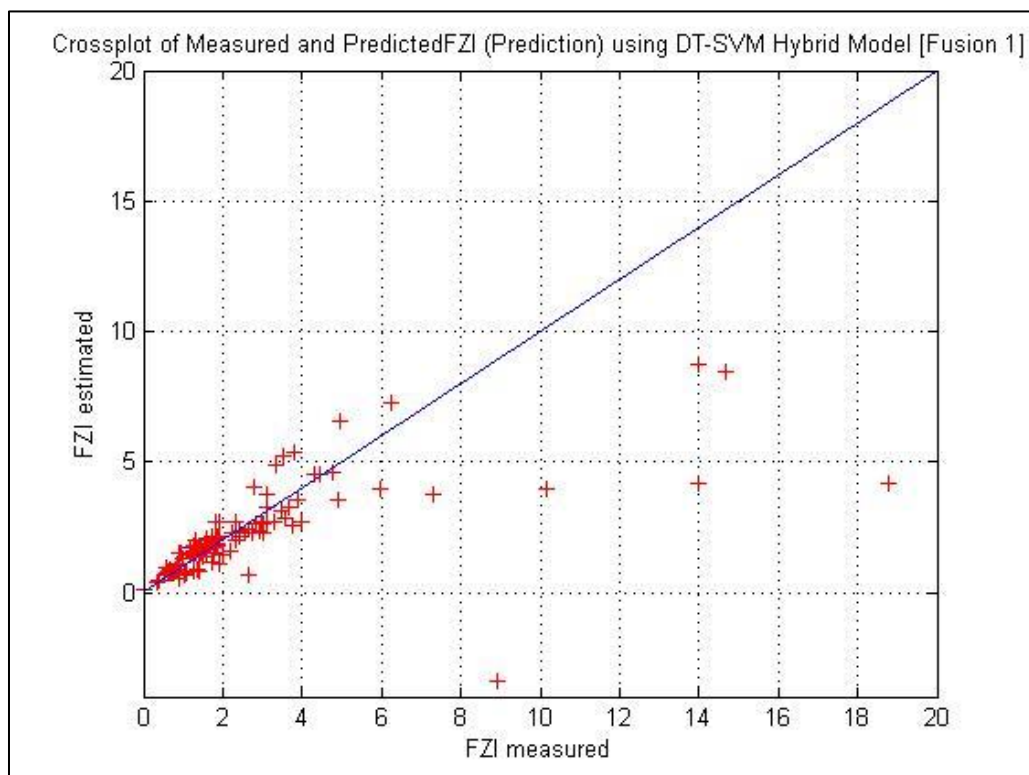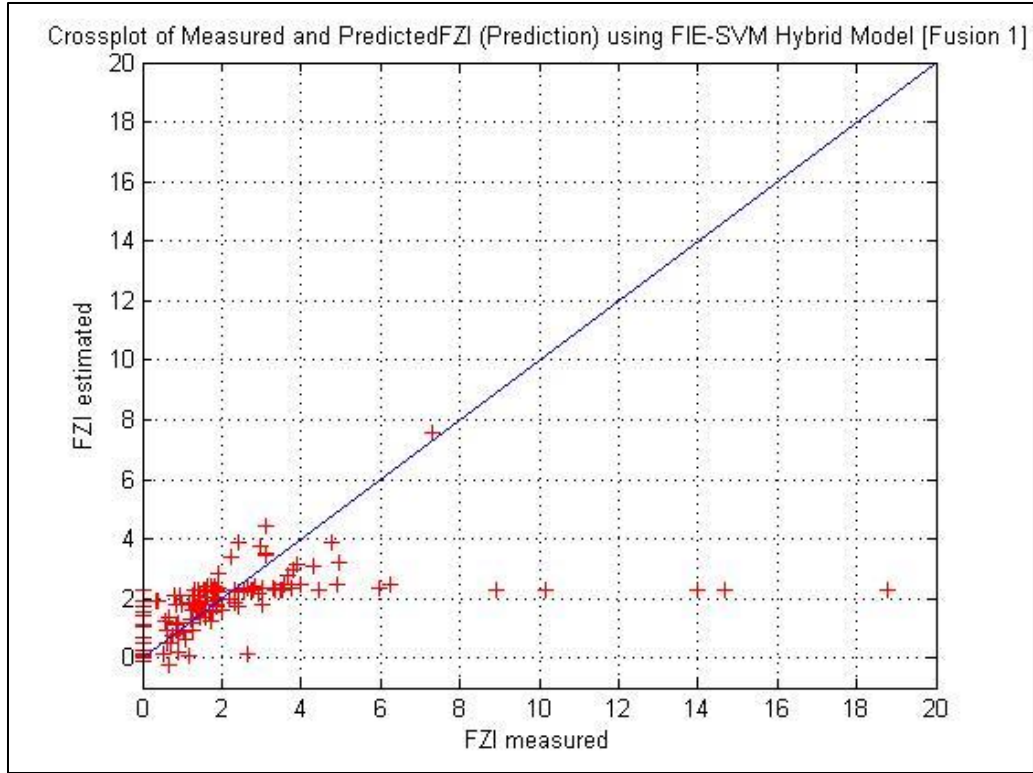**Figure 6.20 - FZI Prediction Crossplot using DT-SVM (FZI Components Fusion AI)**

Figure 6.21 - FZI Prediction Crossplot using FIE-SVM (FZI Components Fusion AI)
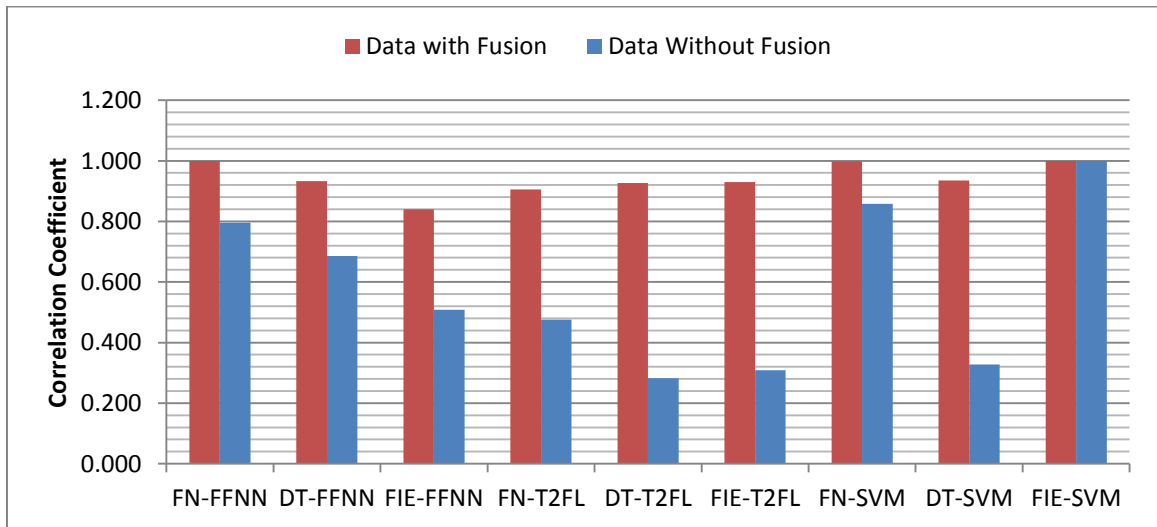

Figure 6.22 - Correlation Coefficient Comparison for all Hybrid AI Models (FZI Components Fusion AI) [Training]
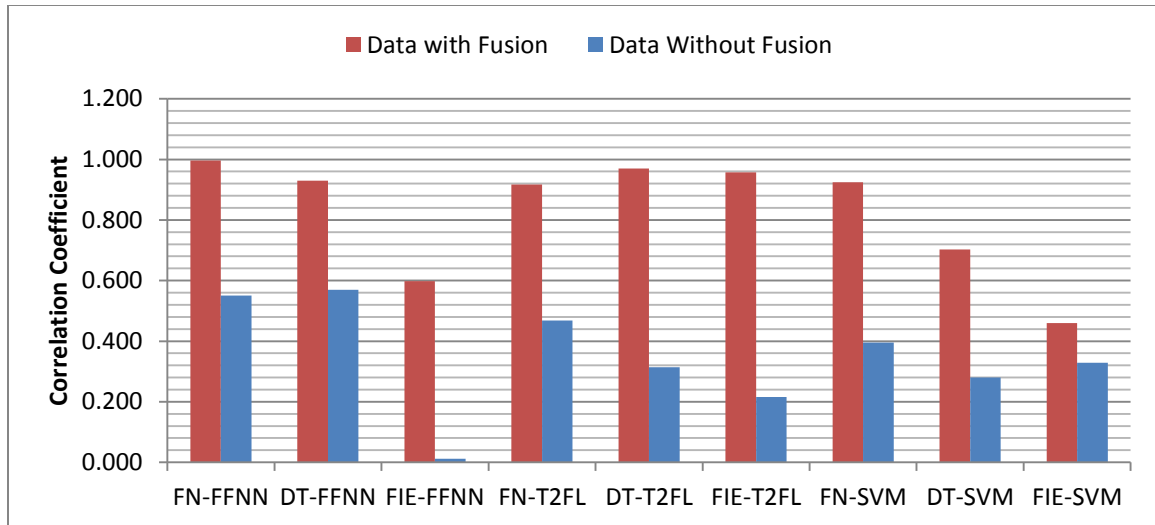
Figure 6.23 - Correlation Coefficient Comparison for all Hybrid AI Models (FZI Components Fusion AI) [Testing]

6.4 GHEs and Flow Zone Indicator Components as Input

A third option was to combine all three additional parameters with the input dataset. This made the total number of input parameters 31. Functional Networks selected all three new parameters while the other two algorithms only selected the $(\sqrt{K/\varphi})$ parameter. This confirmed that this parameter, between the three, had the greatest predictive power. There were also big improvements in all models correlation, Figures 6.24 – 6.32. Figure 6.34 shows that the Feedforward Neural Networks hybrid model with Functional Networks had the highest correlation coefficient, 0.971. On the average, the performance of the FFNN hybrid systems is better than the performance of the hybrids for T2FL and SVM. For this reason, only FFNN hybrid models will be considered for the work in Chapter 7.

Figure 6.24 - FZI Prediction Crossplot using FN-FFNN (GHE & FZI Components Fusion AI)
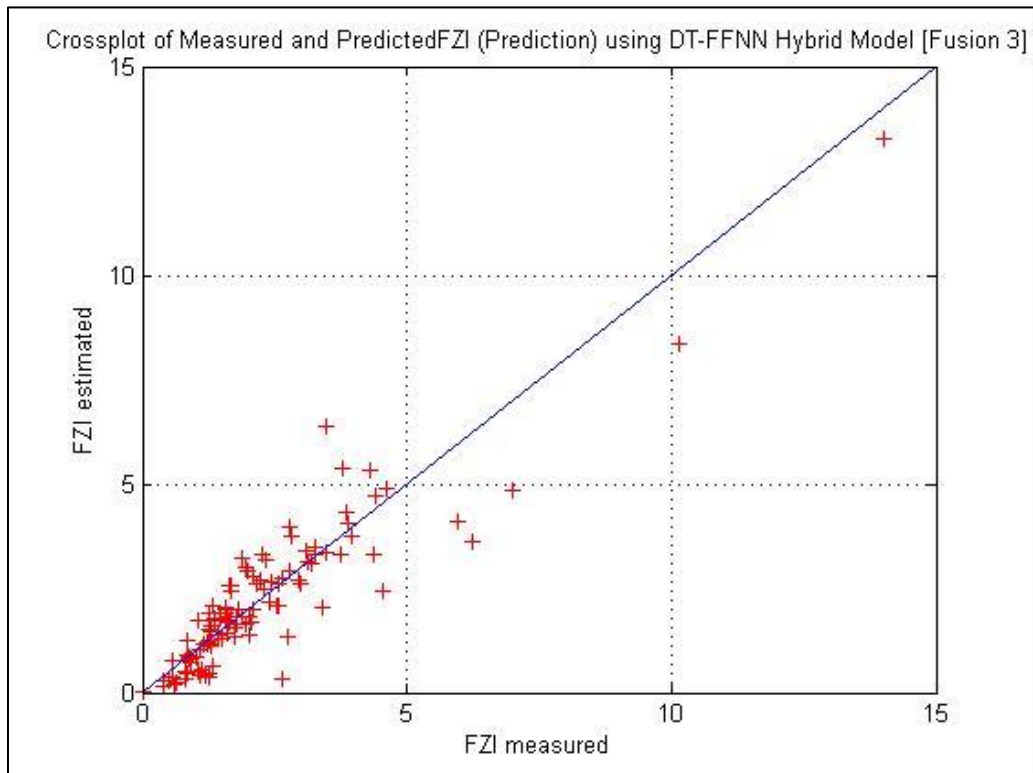
Figure 6.25 - FZI Prediction Crossplot using DT-FFNN (GHE & FZI Components Fusion AI)

Figure 6.26 - FZI Prediction Crossplot using FIE-FFNN (GHE & FZI Components Fusion AI)

Figure 6.27 - FZI Prediction Crossplot using FN-T2FL (GHE & FZI Components Fusion AI)

Figure 6.28 - FZI Prediction Crossplot using DT-T2FL (GHE & FZI Components Fusion AI)

Figure 6.29 - FZI Prediction Crossplot using FIE-T2FL (GHE & FZI Components Fusion AI)
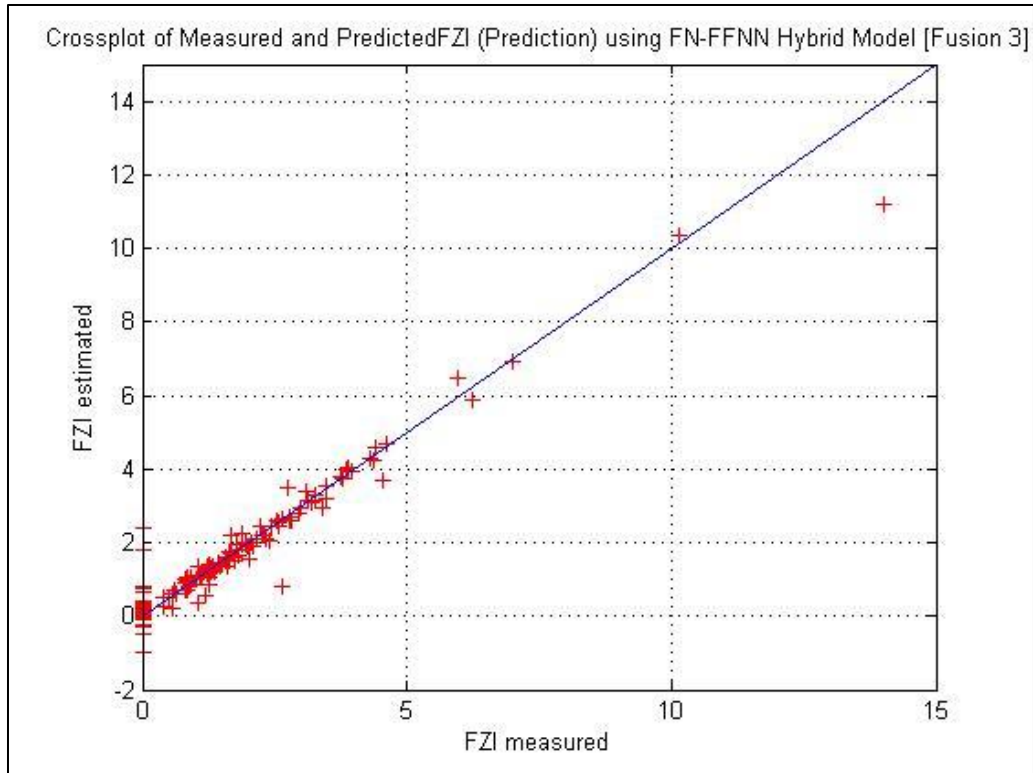
Figure 6.30 - FZI Prediction Crossplot using FN-SVM (GHE & FZI Components Fusion AI)

Figure 6.31 - FZI Prediction Crossplot using DT-SVM (GHE & FZI Components Fusion AI)
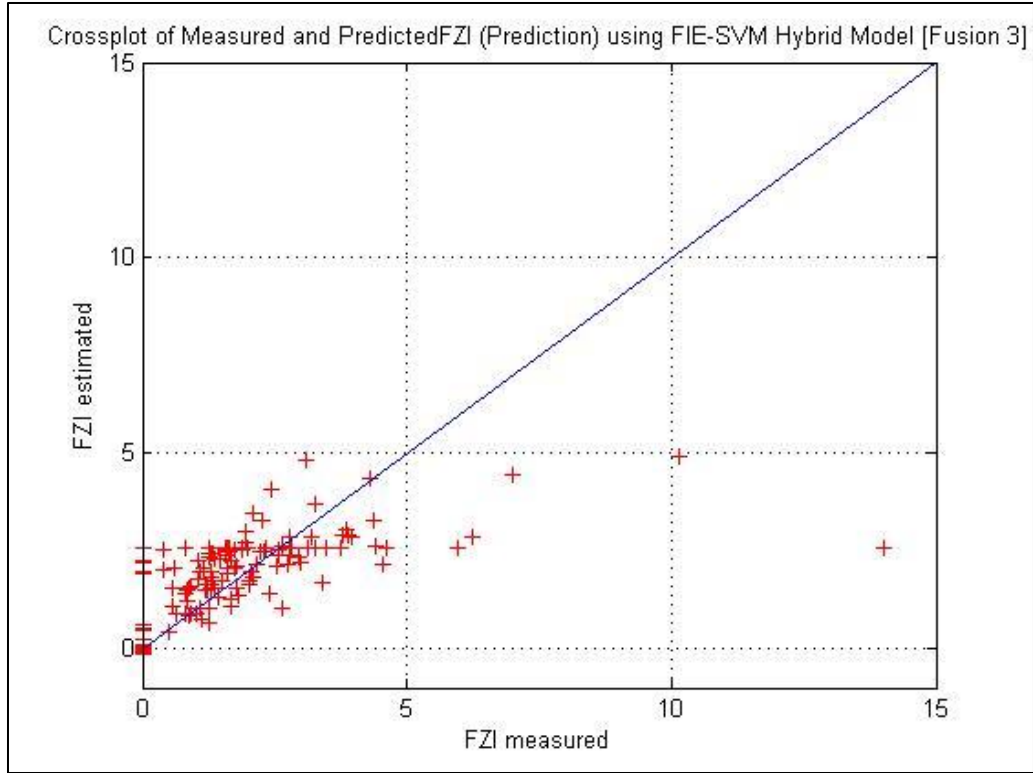
Figure 6.32 - FZI Prediction Crossplot using FIE-SVM (GHE & FZI Components Fusion AI)
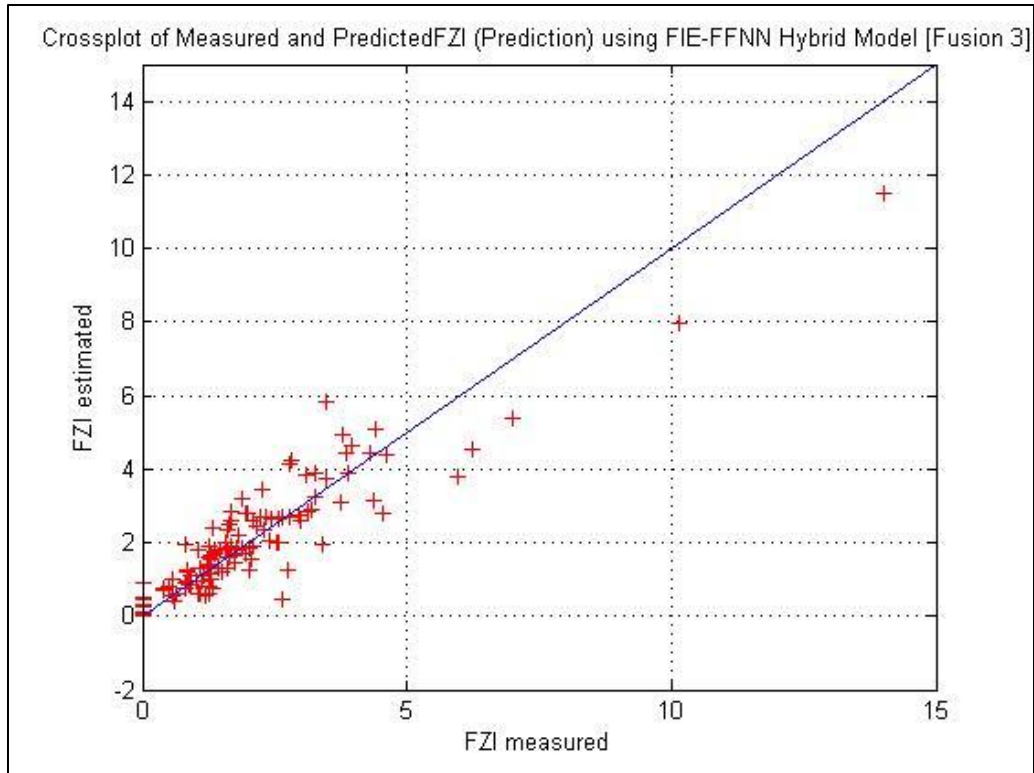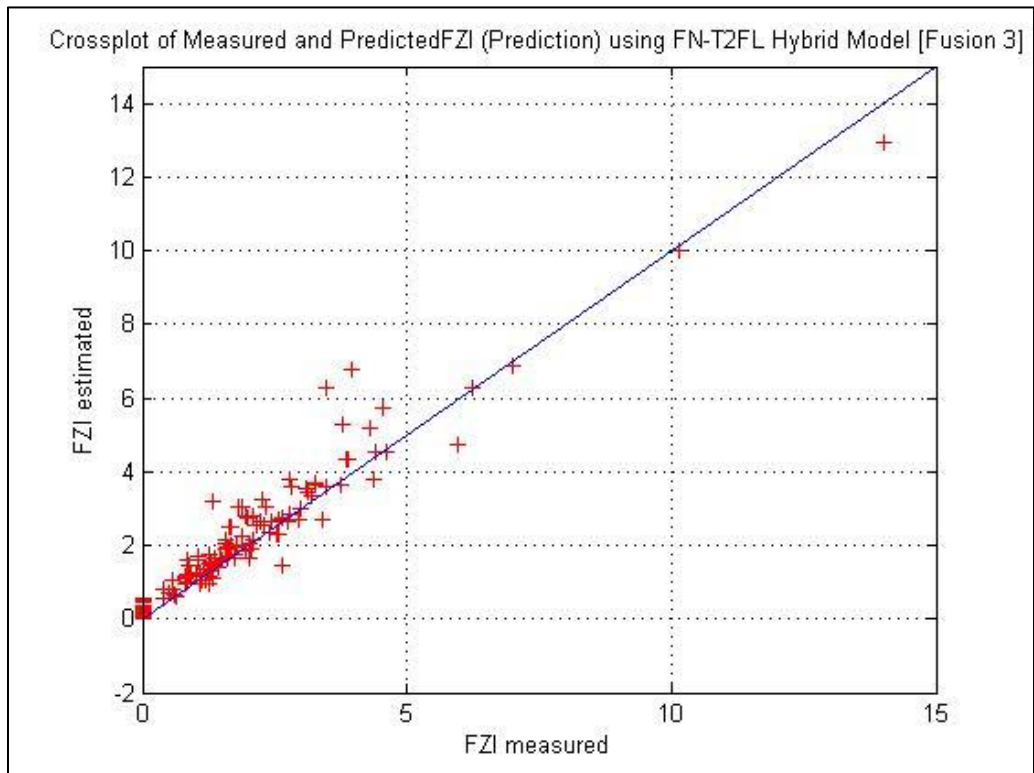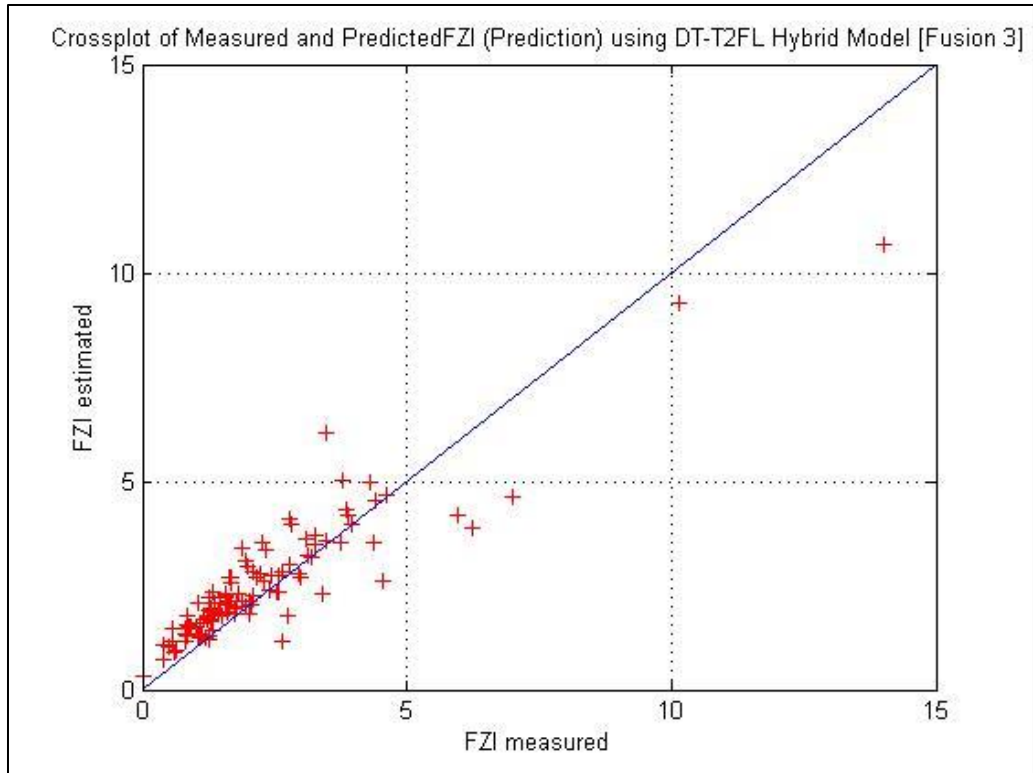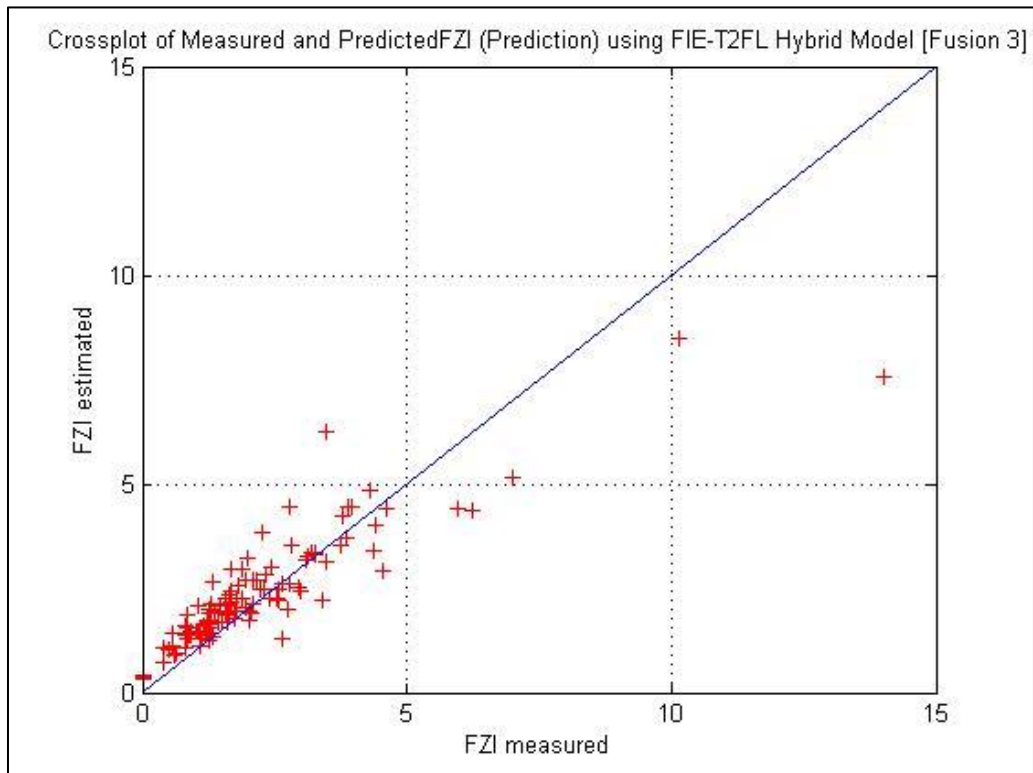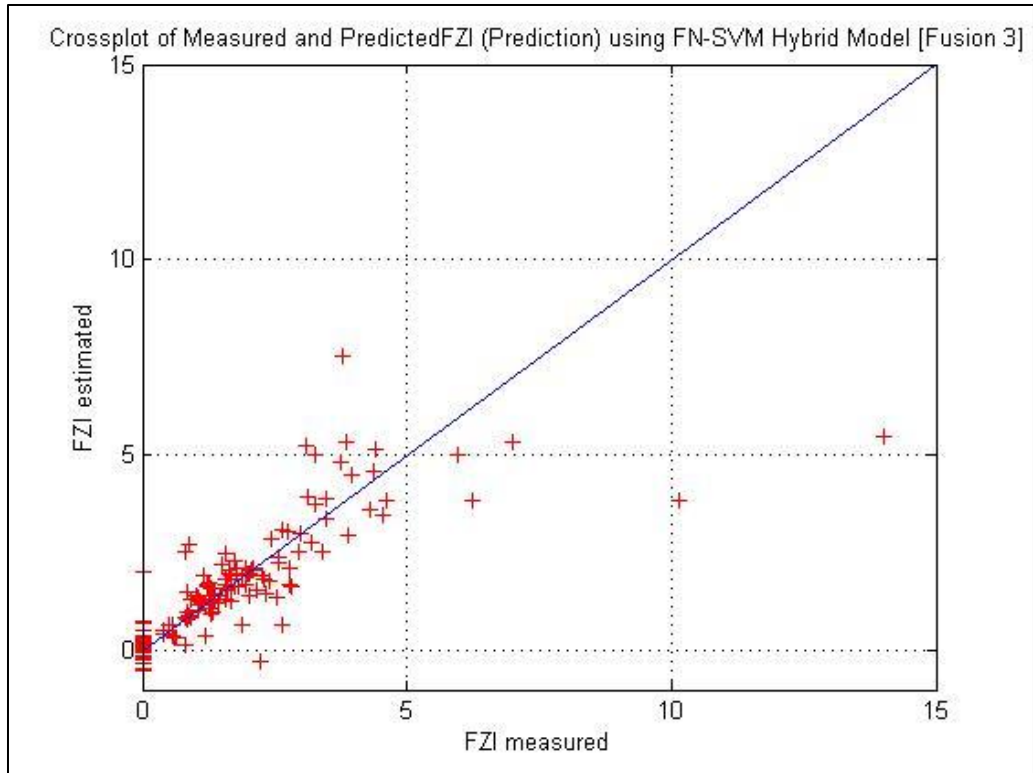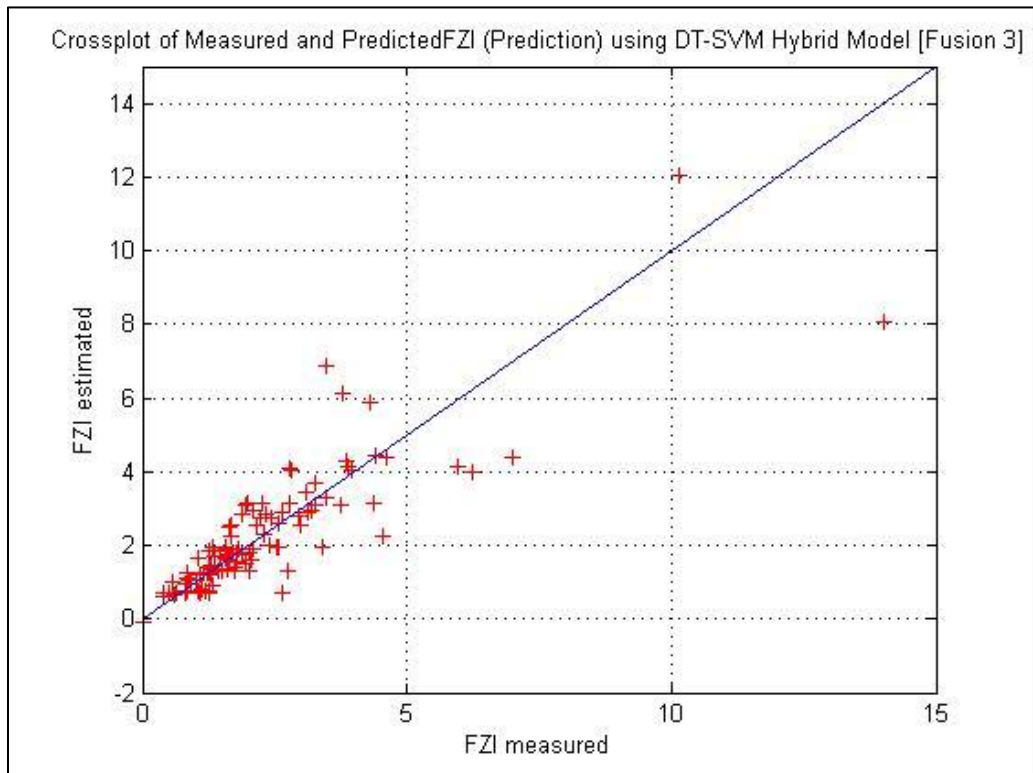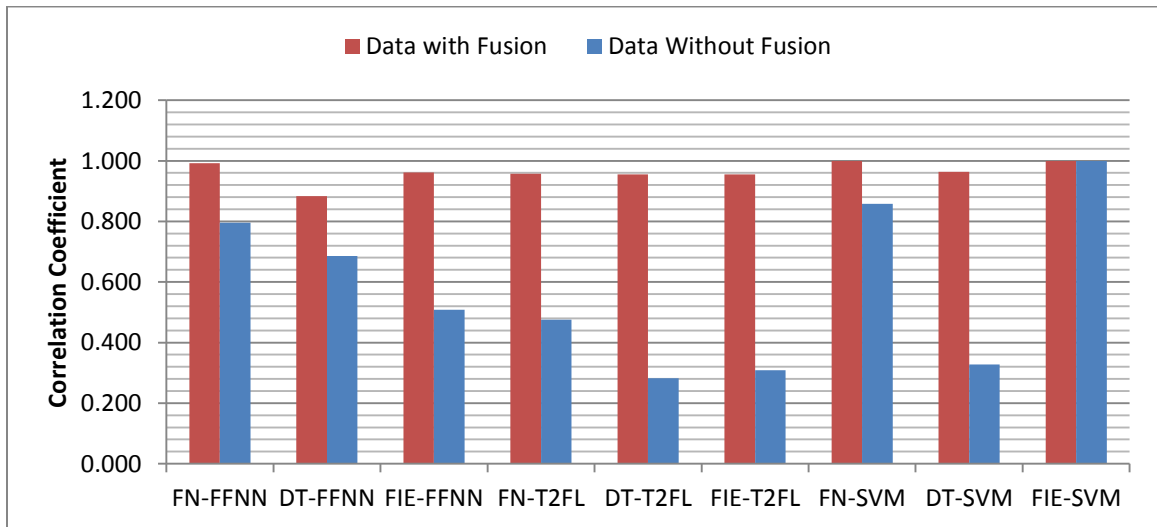

Figure 6.33 - Correlation Coefficient Comparison for all Hybrid AI Models (GHE & FZI Components Fusion AI) [Training]
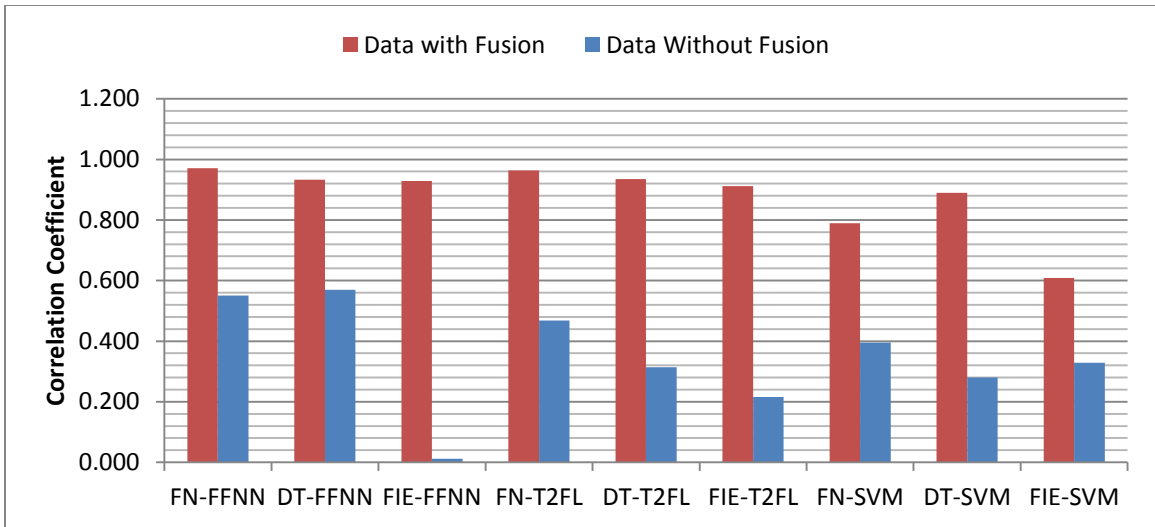
Figure 6.34 - Correlation Coefficient Comparison for all Hybrid AI Models (GHE & FZI Components Fusion AI) [Testing]

CHAPTER 7

Permeability Prediction Improvements

7.1   Introduction

Permeability is one of the most important petrophysical parameters in reservoir characterization. It defines how easy it is for hydrocarbons to flow through the reservoir. Flow zone indicator in the two rock typing methods, hydraulic flow units and global hydraulic elements, relates permeability to porosity, by Equations 2.8 and 2.12 respectively. Chapter 6 covered the improvement of flow zone indicator hybrid models using data fusion. This chapter examines whether one could also improve the prediction of permeability through super hybridization. Only Feedforward Neural Networks (FFNN) hybrid models with Data Fusion were considered here.

7.2   Development of Permeability Prediction Models

First, core permeability data was selected as target for the models. A non-hybrid FFNN was coded and run with the original 28 logs as the input dataset. Then, three FFNN super-hybrid models were developed to predict permeability. Each super-hybrid model uses one of the three features selection algorithms for its prediction of the flow zone indicator: FN-FFNN(FZI), DT-FFNN(FZI), and FIE-FFNN(FZI). Data Fusion was implemented in each one through using both GHE classification and FZI components as additional inputs. Predicted flow zone indicator values were also added to the input dataset for all the FFNN(K) models. Lastly, permeability values were predicted. Figure 7.1 depicts the

super-hybrid modeling with Data Fusion process that was used to predict permeability values. Figures 7.2 and 7.3 show cross-plots of predicted and measured permeability values using the non-hybrid FFNN model for training and testing, respectively. Figures 7.4 and 7.5 show that the model missed many permeability values along the depth.
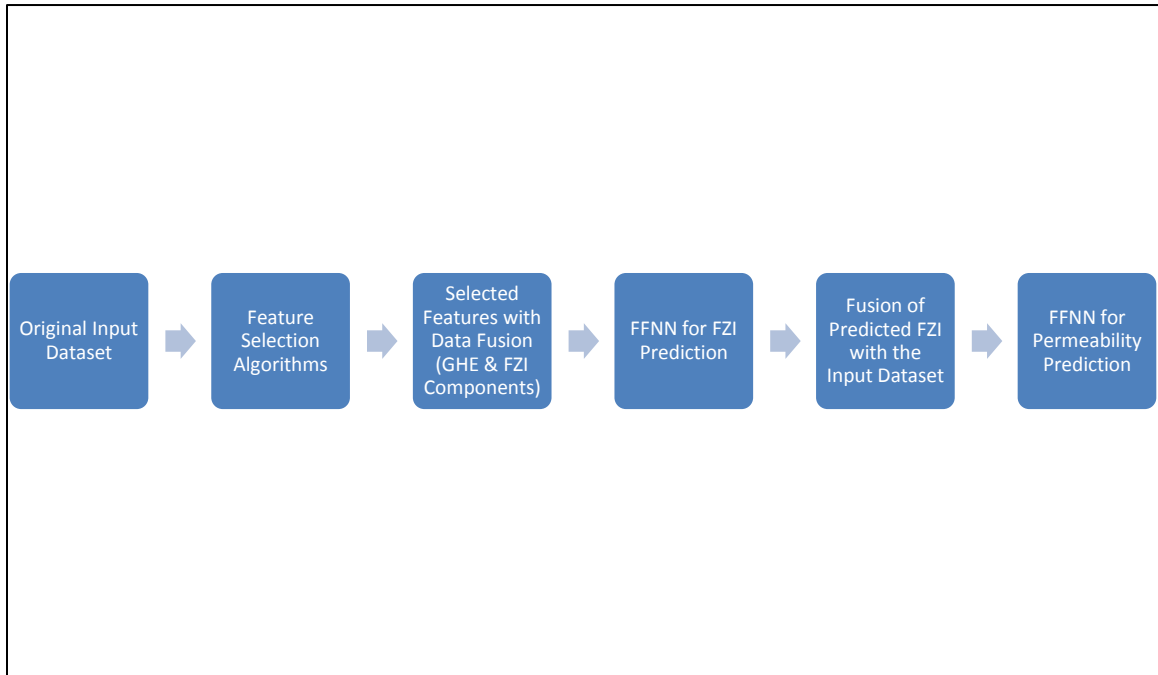
Figure 7.1 - Process of Super-Hybrid Modelling with Data Fusion for Permeability Prediction
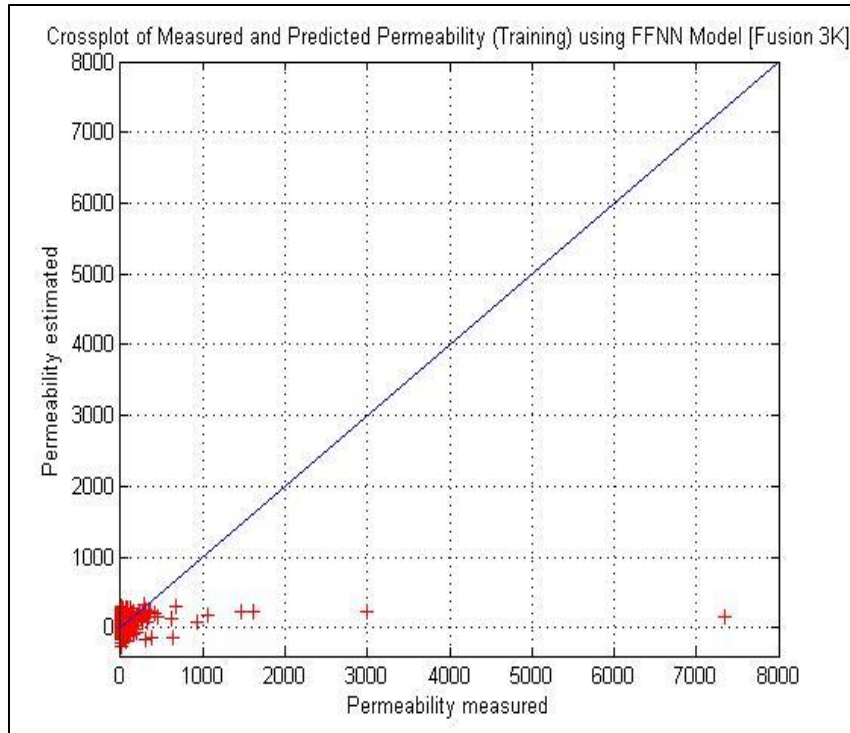
Figure 7.2 – Permeability Prediction Crossplot using FFNN (Training)
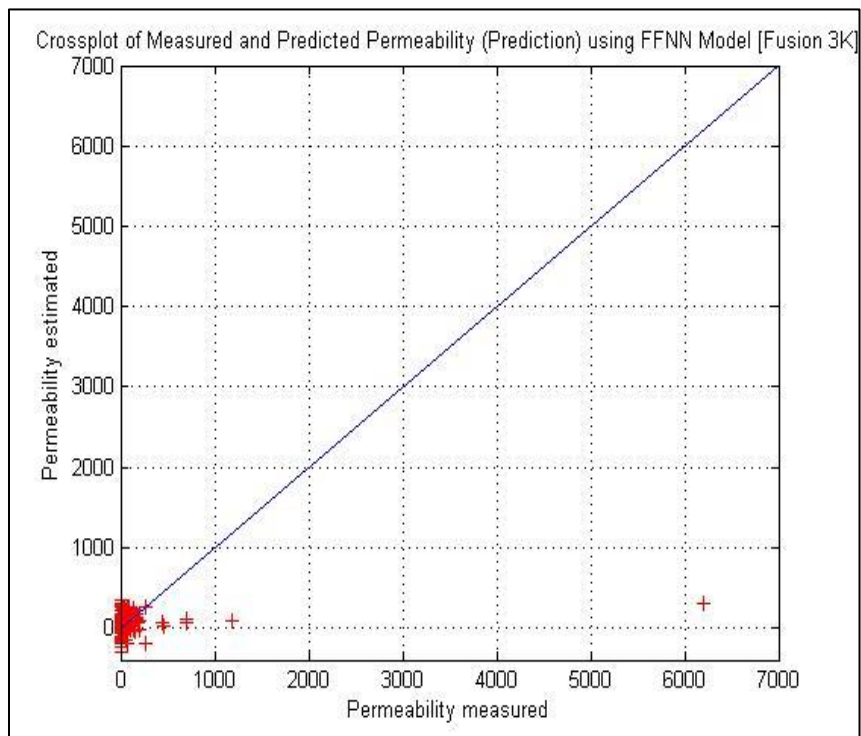
Figure 7.3 - Permeability Prediction Crossplot using FFNN (Testing)
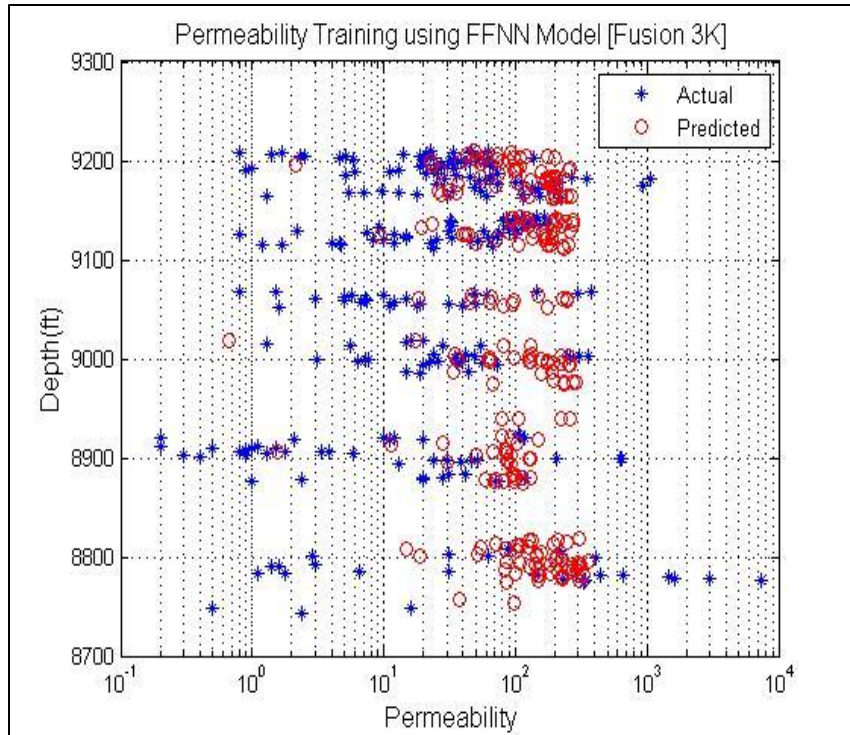
Figure 7.4 - Actual and Predicted Permeability versus Depth Using FFNN Model (Training)
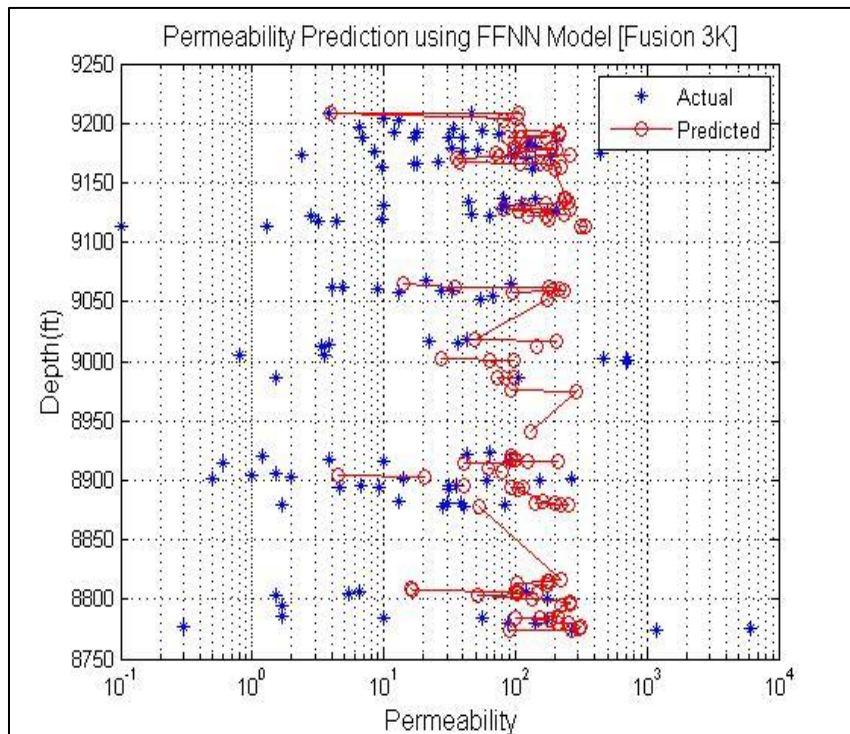
Figure 7.5 - Actual and Predicted Permeability versus Depth Using FFNN Model (Testing)

7.3   Performance of FN-FFNN(FZI)-FFNN(K) Super-Hybrid Model

Figures 7.6 and 7.7 show improved cross-plots for training and testing with few uncorrelated points. This improvement was also obvious when the predicted and measured data were plotted along the depth in Figures 7.8 and 7.9.
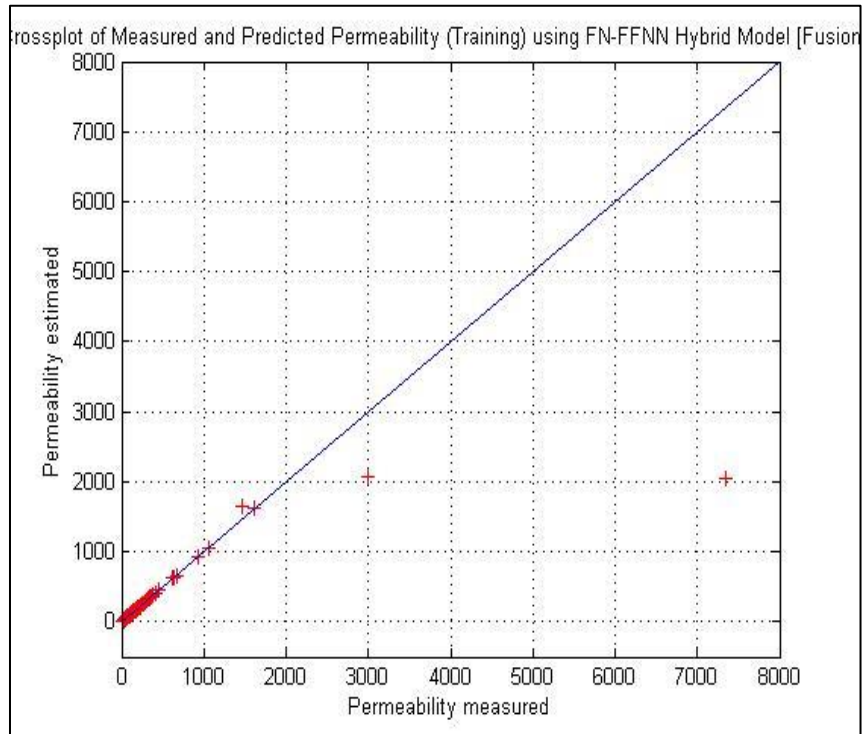
Figure 7.6 - Permeability Prediction Crossplot using FN-FFNN (GHE & FZI Components Fusion AI) [Training]
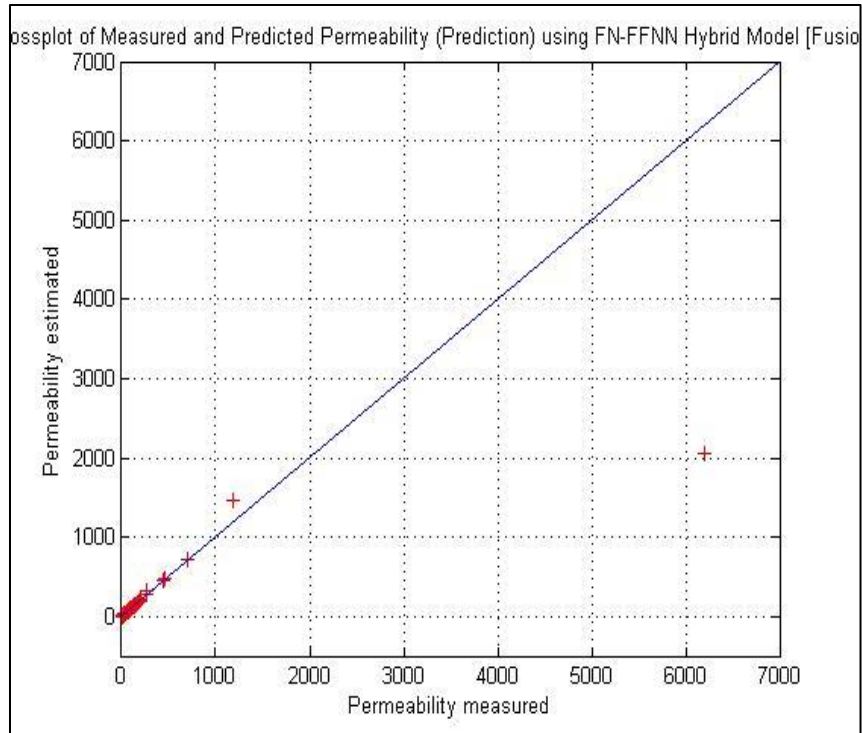
Figure 7.7 - Permeability Prediction Crossplot using FN-FFNN (GHE & FZI Components Fusion AI) [Testing]
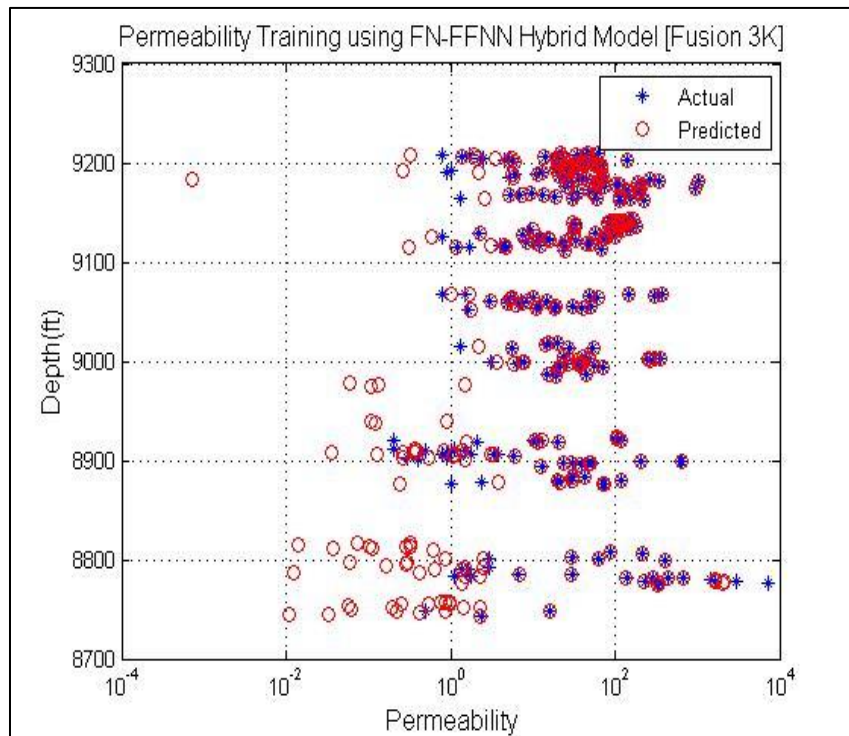
Figure 7.8 - Actual and Predicted Permeability versus Depth Using FN-FFNN Model (GHE & FZI Components Fusion AI) [Training]

91
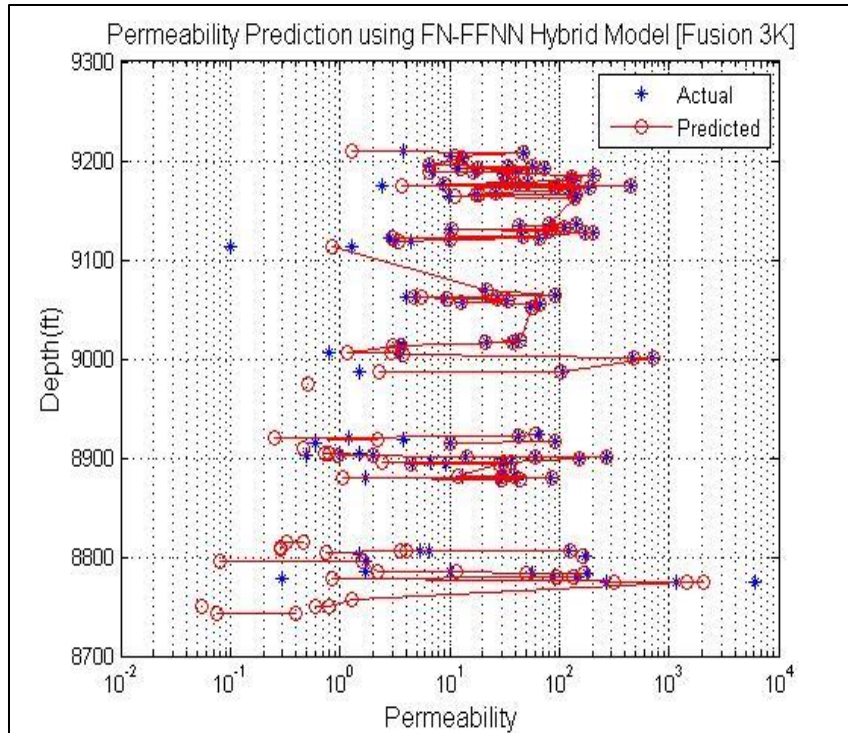
Figure 7.9 - Actual and Predicted Permeability versus Depth Using FN-FFNN Model (GHE & FZI Components Fusion AI) [Testing]

7.4   Performance of DT-FFNN(FZI)-FFNN(K) Super-Hybrid Model

Figures 7.10 and 7.11 also show improved cross-plots for training and testing with even fewer uncorrelated points. Figures 7.12 and 7.13 show improved plots of the predicted and measured permeability data points along the depth.

92
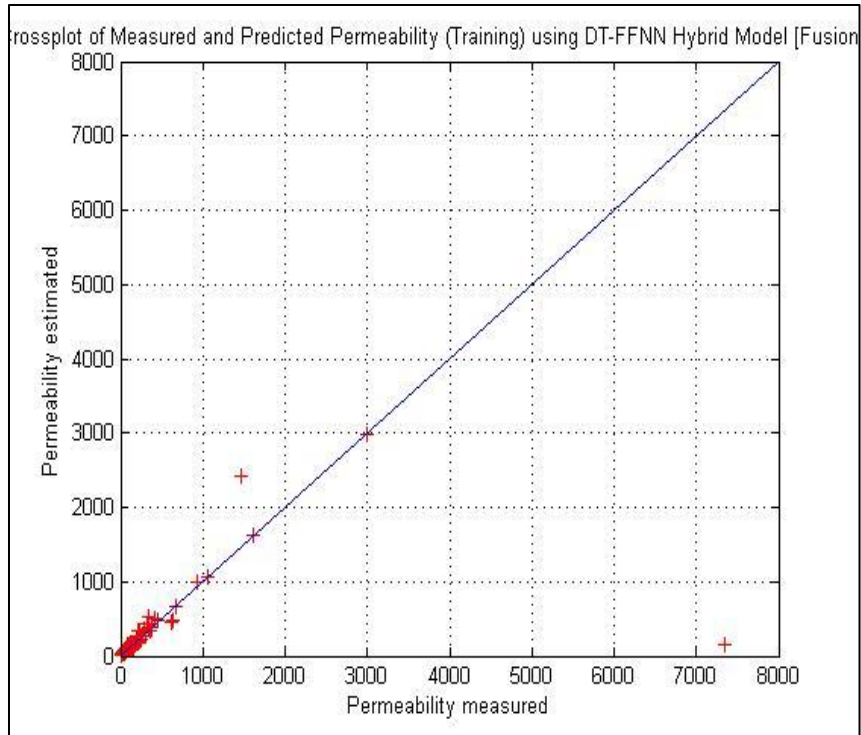
Figure 7.10 - Permeability Prediction Crossplot using DT-FFNN (GHE & FZI Components Fusion AI) [Training]
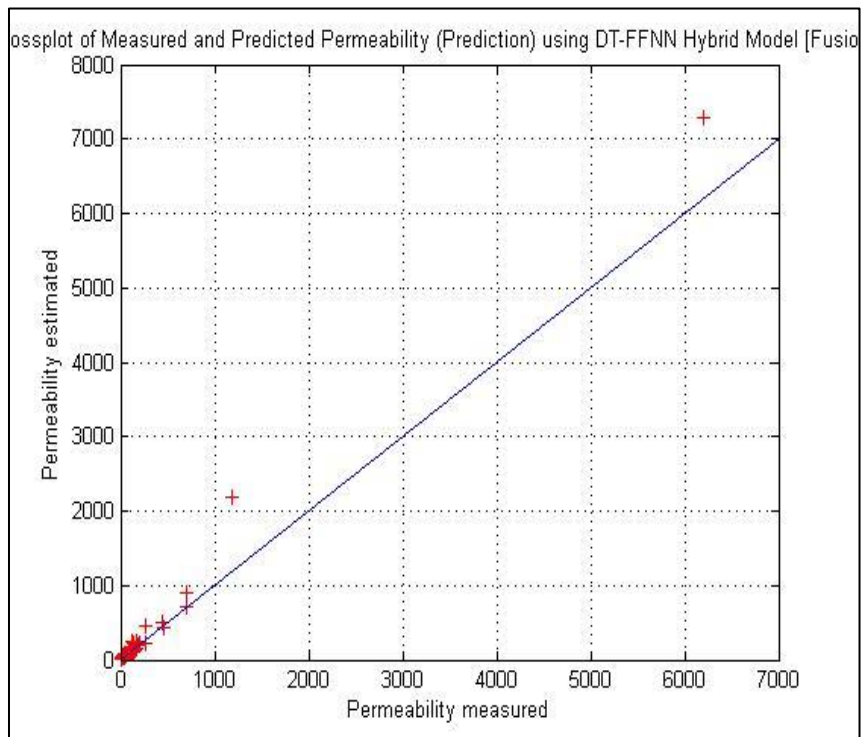
Figure 7.11 - Permeability Prediction Crossplot using DT-FFNN (GHE & FZI Components Fusion AI) [Testing]
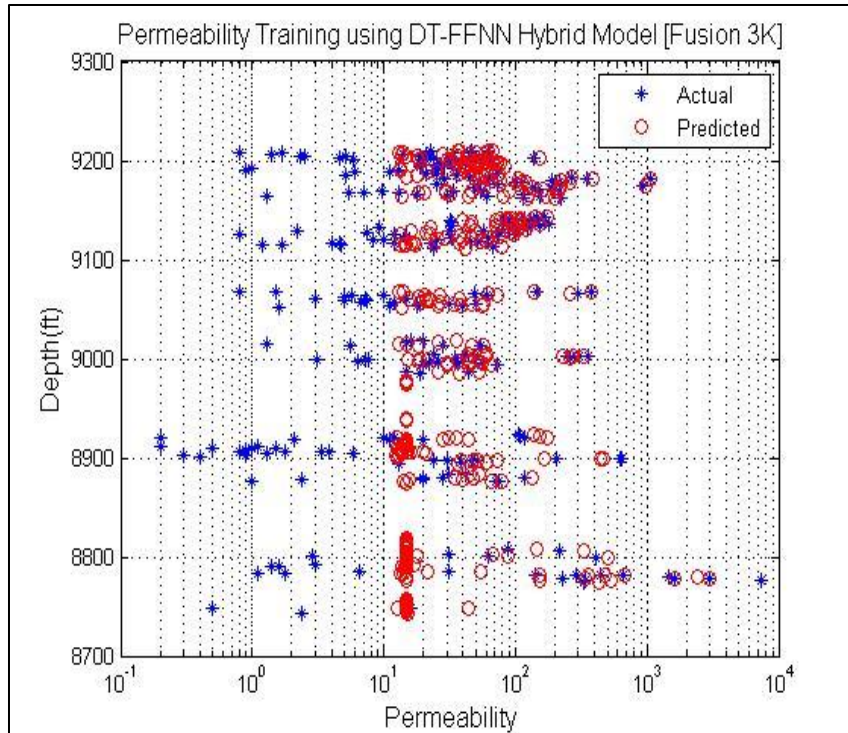
Figure 7.12 - Actual and Predicted Permeability versus Depth Using DT-FFNN Model (GHE & FZI Components Fusion AI) [Training]
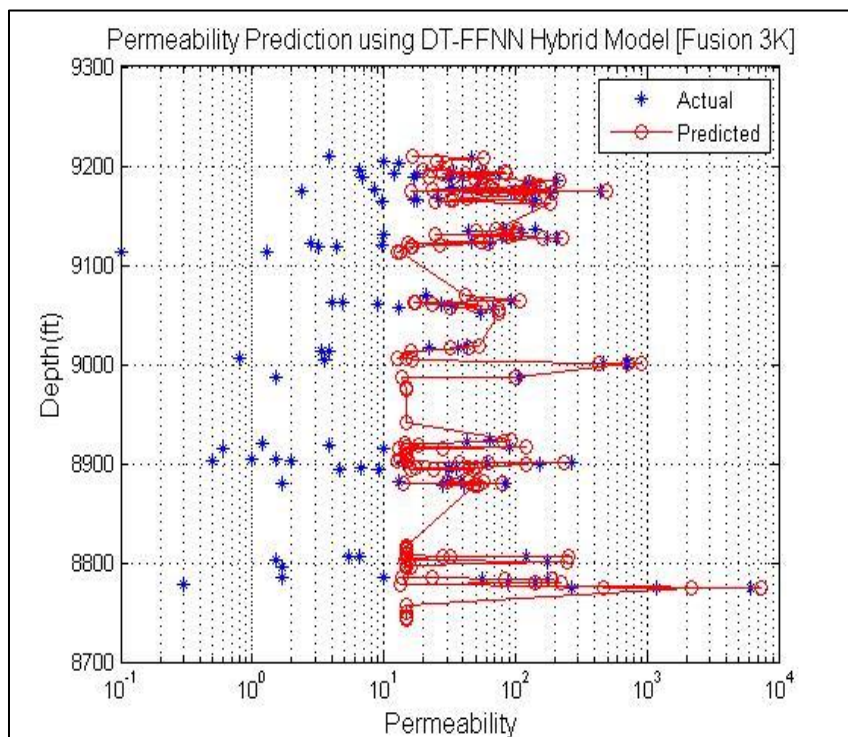
Figure 7.13 - Actual and Predicted Permeability versus Depth Using DT-FFNN Model (GHE & FZI Components Fusion AI) [Testing]

7.5   Performance of FIE-FFNN(FZI)-FFNN(K) Super-Hybrid Model

Figures 7.14 and 7.15 show improved cross-plots for training and testing with few uncorrelated points. This improvement was also obvious when the predicted and measured data were plotted along the depth in Figures 7.16 and 7.17.
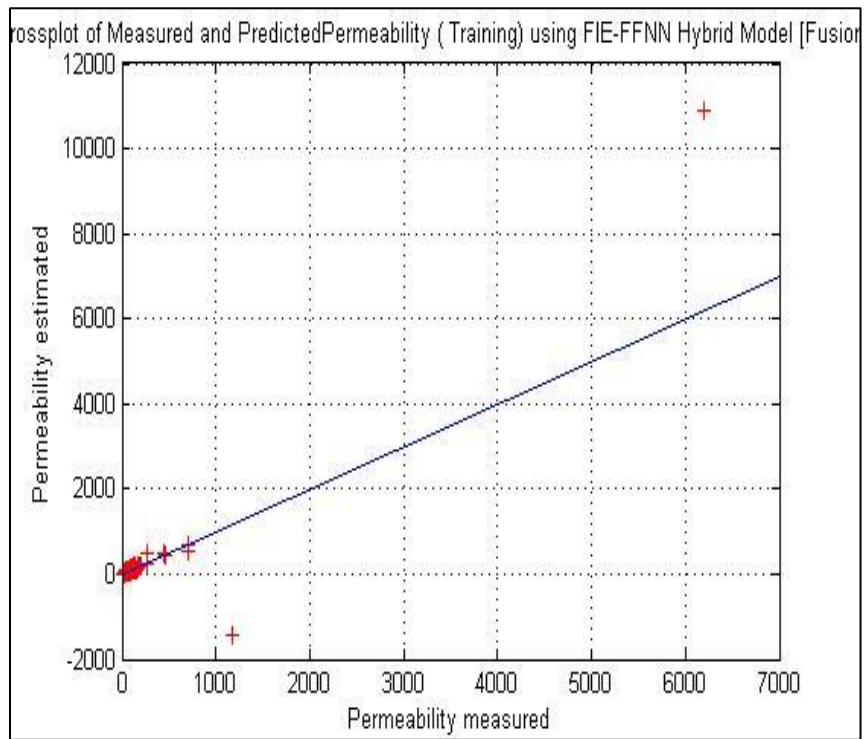
Figure 7.14 - Permeability Prediction Crossplot using FIE-FFNN (GHE & FZI Components Fusion AI) [Training]
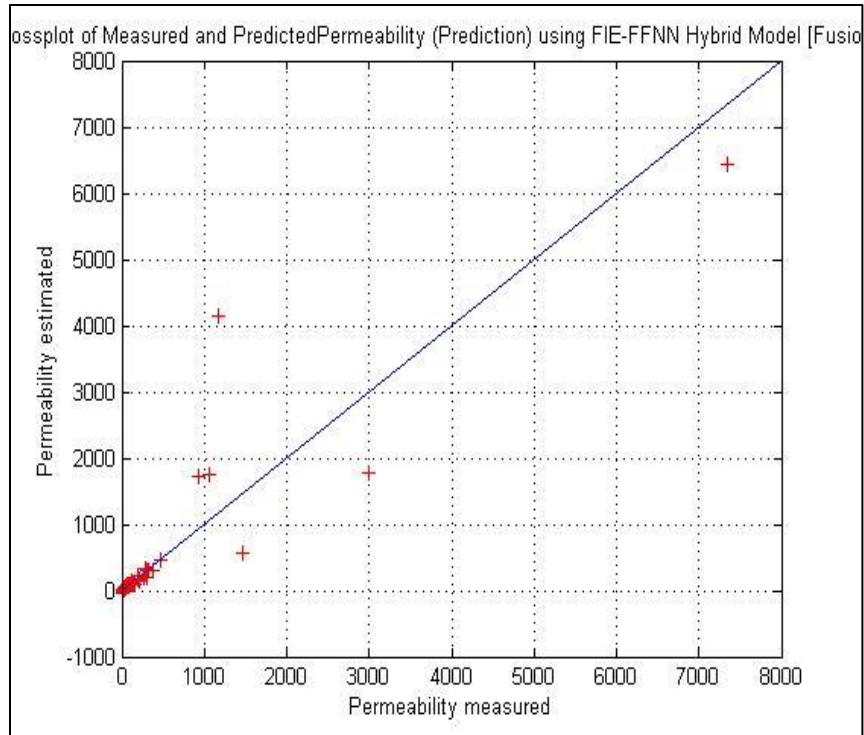
Figure 7.15 - Permeability Prediction Crossplot using FIE-FFNN (GHE & FZI Components Fusion AI) [Testing]
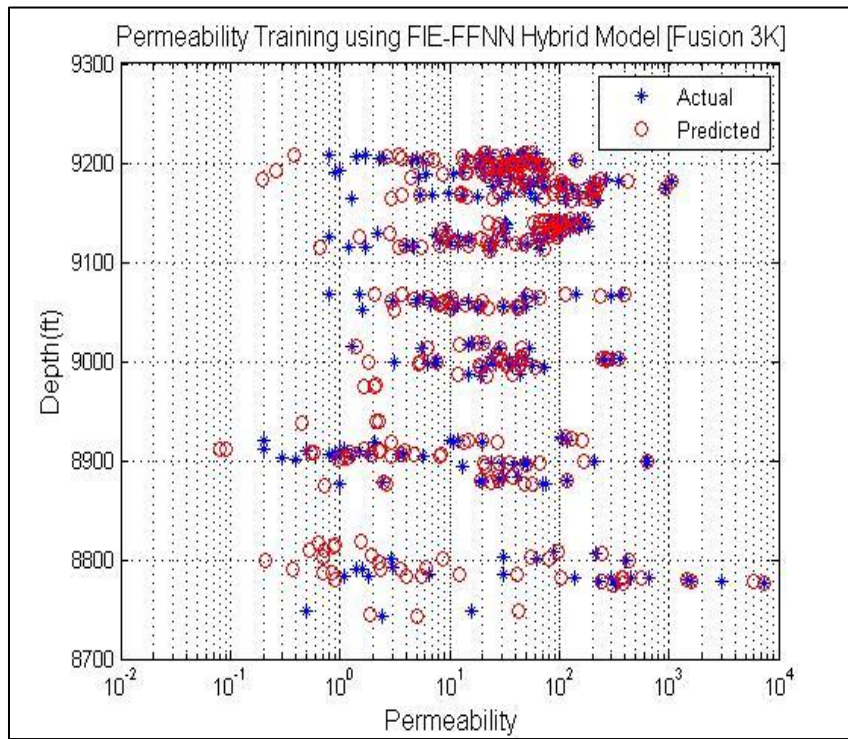
Figure 7.16 - Actual and Predicted Permeability versus Depth Using FIE-FFNN Model (GHE & FZI Components Fusion AI) [Training]
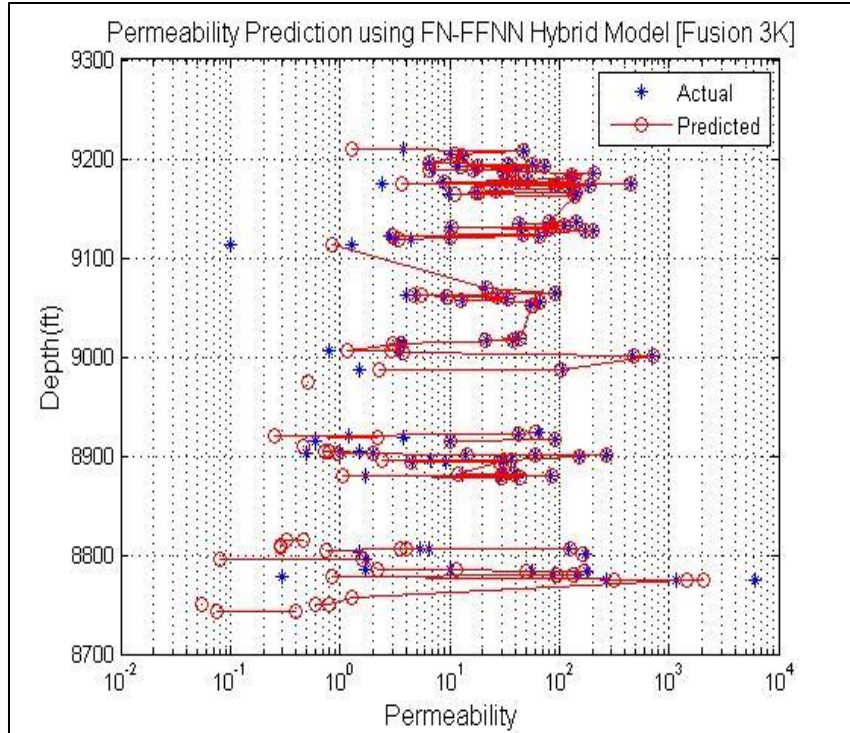
Figure 7.17 - Actual and Predicted Permeability versus Depth Using FIE-FFNN Model (GHE & FZI Components Fusion AI) [Testing]

7.6   Performance Comparison for Permeability Prediction Models

Figure 7.18 shows the comparison of correlation coefficients between the non-hybrid model and the other three super-hybrid ones. Performance improvement is obvious in all super-hybrid models. In the prediction, the correlation coefficient jumped from 0.144 for non-hybrid model to 0.878 for FN-FFNN(FZI)-FFNN(K), to 0.994 for DT-FFNN(FZI)-FFNN(K), and to 0.947 for FIE-FFNN(FZI)-FFNN(K). Figure 7.19 shows the root mean-squared error comparisons for all four models. Table 7.1 provides statistical summary for permeability prediction models.

97

Table 7.1 - Statistical Results for Permeability Prediction Models

Models	Training			Prediction		
	CC	RMSE	MAE	CC	RMSE	MAE
FN-FFNN(FZI)-FFNN(K)	0.840	291.541	19.399	0.878	343.63	32.01
DT-FFNN(FZI)-FFNN(K)	0.501	394.192	41.991	0.994	127.09	34.32
FIE-FFNN(FZI)-FFNN(K)	0.959	159.494	15.333	0.947	446.20	60.62
FFNN(K)	0.105	459.431	155.046	0.144	523.70	178.51
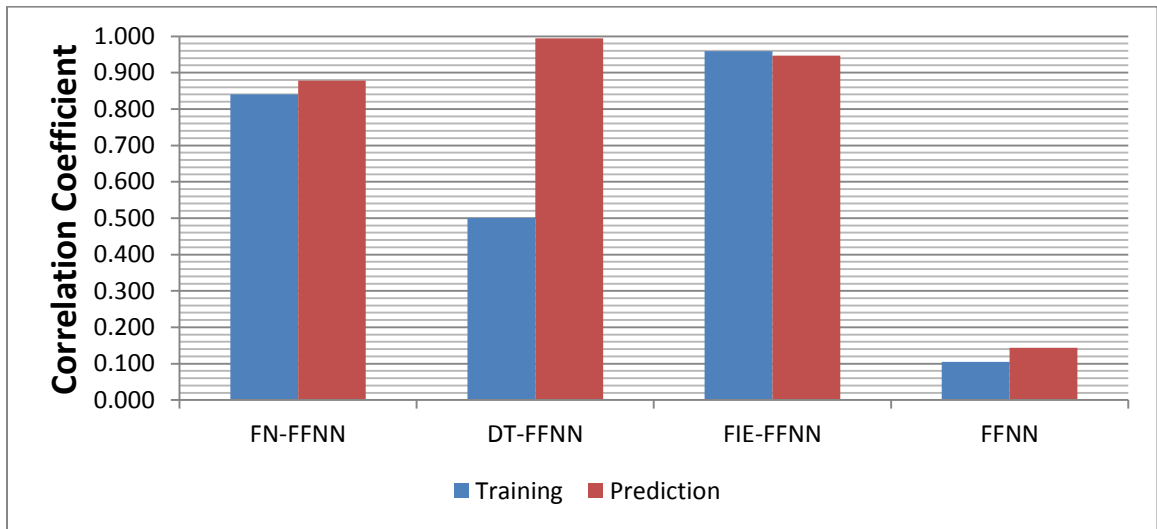
Figure 7.18 - Correlation Coefficient Comparison for all Permeability AI Models
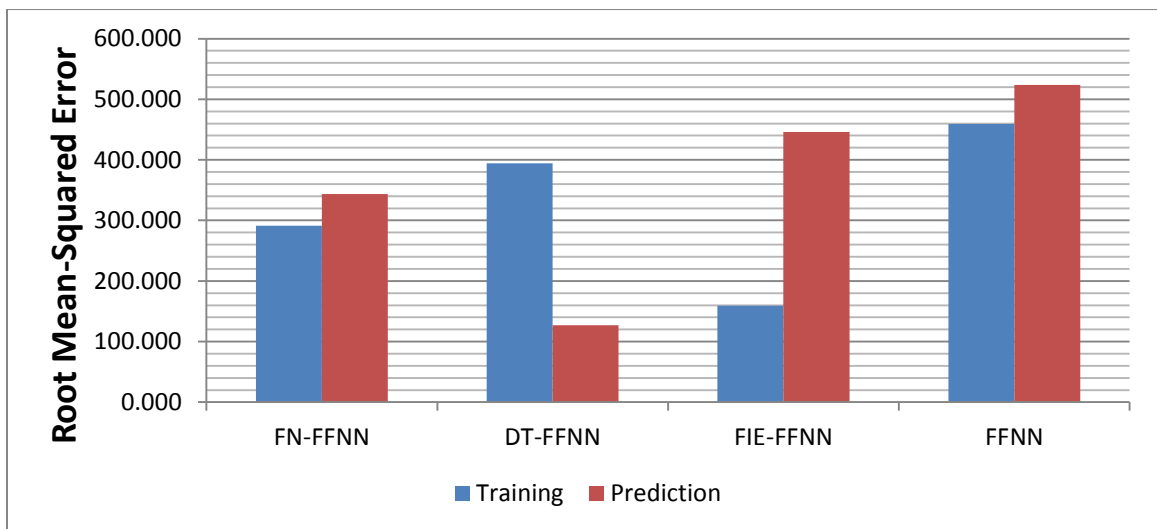
Figure 7.19 - Root Mean-Squared Error Comparison for all Permeability AI Models

98

CHAPTER 8

CONCLUSIONS AND RECOMMENDATIONS

8.1   Conclusions

1. NMR logs were found to be contributing to the prediction of flow zone indicator (FZI). They accounted for approximately 70% of the selected logs by functional networks and fuzzy information entropy. For decision trees algorithm, only one out of five selected logs was from NMR (20%). The work also showed that no common NMR log was selected by the three feature selection algorithms.

2. Feedforward neural network (FFNN) non-hybrid models outperformed the other non-hybrid models in the prediction of FZI with a correlation coefficient (CC) of 0.49. Radial basis function neural networks model was the worst non-hybrid performer with only 0.03 in correlation coefficient for testing.

3. Between the three features selection algorithms, Functional Networks selected the largest number of logs (7) while Fuzzy Information Entropy selected the least number (4)

4. The biggest improvement in prediction performance due to hybridization was observed in the support vector machines models. The correlation coefficient jumped from 0.06 for non-hybrid model to 0.41 for FN-SVM, to 0.56 for DT-SVM, and to 0.57 for FIE-SVM.

5. Moderate improvement was observed in feedforward neural network and type-2 fuzzy logic hybrid models.

6. Generally, FFNN hybrid models were the best in terms of correlation coefficient for prediction and T2FL were the worst performers. In terms of the correlation coefficient for prediction, the DT-FFNN model was the best model with 0.64 and the FIE-T2FL model was the worst with 0.17

7. Generally, data fusion helped improving all the developed hybrid systems. The largest improvement was observed with the fusion of all the three new data, GHE numerical classification and the two FZI components, together with the previously selected logs. All FFNN and T2FL models had correlation coefficients for prediction greater than 0.9

8. When the ($\sqrt{K/\varphi}$) FZI component was fused with the input dataset, it was consistently selected by all feature selection algorithms. This indicated that it had the largest contribution to the improvement of FZI prediction by data fusion.

9. Improvement in flow zone indicator prediction also led to improved permeability prediction. The non-hybrid FFNN model had only 0.144 as its prediction correlation coefficient. The super-hybridization with the fusion of the additional input data (GHE numerical classification, FZI components, as well as predicted FZI values) made the correlation coefficient jump to greater than 0.87

10. The DT-FFNN(FZI)-FFNN(K) super-hybrid model for permeability prediction was the best model in terms of prediction correlation coefficient 0.994.

8.2   Recommendations

The following recommendations can be made to extend the work of this study.

1. The developed models can be further enhanced by adding more data points with wider ranges since AI techniques are data-driven.

2. The approach may be tested for other formations and rock types, such as naturally fractured and vuggy, with suitable code modifications.

3. The approach should also be tested with additional AI tools such as ANFIS

4. Hybridization of other AI tools such as RBFNN and GRNN can also be explored

5. Use of AI classification tools to predict GHEs directly from conventional logs

6. The application of the developed models for the prediction of permeability in tight rocks especially for shales can be explored

7. Depending on data availability, other hydraulic units identification methods may be considered as model targets for permeability prediction improvement

References

A. Kohli; P Arora;, " Application of Artificial Neural Network for Well Logs," *Paper IPTC 17475*, Jan. 2014

Amaefule, J. O; Altunbay M; Henry Ohen; Kersey David; Lane Peter;, " A Hydraulic (Flow) Units-Based Approach for Predicting Formation Damage in Uncored Interval/Wells Using Core/Log Data," *Presented at the International Symposium on Formation Damage Control*, 9-10 Feb. 1994

Bear, J.;, " Dynamics of Fluids in Porous Media," *Elsevier*, New York, 1972

C. L. Hearn;, " Geological Factors Influencing Reservoir Performance of the Hartzog Draw Field," *Journal of Petroleum Technology*, Aug. 1984

Cant, D. J.;, " Subsurface Facies Analysis," *Facies Models*, Geoscience Canada Reprint Series 1, Geological Assn. of Canada Publications, Business & Economic Service, Toronto, 1984

Cesar Bravo; Luigi Saputelli; Francklin Rivas; Anna Gabriela Perez; Michael Nikolaou; Georg Zangl; Shahab Mohaghegh; Gustavo Nunez;, " State of the Art of Artificial Intelligence and Predictive Analytics in the E&P Industry: A Technology Survey," *Paper SPE 150314*, Mar. 2012

D. Astudillo; J. Porlles;, " Multidisciplinary Workflows Applied in Reservoir Characterization: A Heavy Oil Field Case Study During Early Stages of the Asset Life-Cycle," *Paper SPE 139236*, 1-3 Dec. 2010

D. K. Davies; R. K. Vessell;, " Identification and Distribution of Hydraulic Flow Units in a Heterogeneous Carbonate Reservoir: North Robertson Unit, West Texas," *Paper SPE 35183*, 27-29 Mar. 1996

Ebanks, W. J.;, " Flow Unit Concept – Integrated Approach for Engineering Projects," Abstract, AAPG Annual Convention, 1987

Fatai A. Anifowose; Abdulazeez Abdulraheem; Abdullatif Al-Shuhail;, " Improved Permeability Prediction from Seismic and Log Data using Artificial Intelligence Techniques," *Paper SPE 164465*, Mar. 2013

Fatai Anifowose; Abdulazeez Abdulraheem;, " Prediction of Porosity and Permeability of Oil and Gas Reservoirs using Hybrid Computational Intelligence Models," *Paper SPE 126649*, Feb. 2010

Fethi Elarouci; Nicolae Mihail; Iguer Adnan; Hicham Benaissa; Mokrani Nabil; Bouillouta Fouzi;, " Tight Reservoir Permeability Prediction by Integrated Petrophysical Evaluation, Core Data, and Micro-frac Wireline Formation Tester Analysis," *Paper SPE 150611*, 20-22 Feb. 2012

G. Gunter; J. Finneran; D. Hartmann; J. Miller;, " Early Determination of Reservoir Flow Units Using an Integrated Petrophysical Method," *Paper SPE 38679*, Oct. 1997

Guangming Ti; D. O. Ogbe; Walt Munly; D. G. Hatzignatiou;, " Use of Flow Units as a Tool for Reservoir Description: A Case Study," *Paper SPE 26919-PA*, 1995

Hasan A. Nooruddin; Fatai Anifowose; Abdulazeez Abdulraheem;, " Applying Artificial Intelligence Techniques to Develop Permeability Predictive Models using Mercury Injection Capillary-Pressure Data," *Paper SPE 168109*, May. 2013

Hussain Ali Baker; Sameer Noori Al-Jawad; Zainab Imad Mutadha;, " Permeability Prediction in Carbonate Reservoir Rock Using FZI," *Iraqi Journal of Chemical and Petroleum Engineering.* 14. p. 49-54, Sep. 2013

J. K. Ali;, " Neural Networks: A New Tool for the Petroleum Industry?," *Paper SPE 27561*, Mar. 1994

J.P. Salazar; P.A. Romero;, " NMR Measurements in Carbonates Samples and Core-Logs Correlations Using Artificial Neural Nets," *Paper SPE 71701*, Jan. 2001

Jong-Se Lim; Jungwhan Kim;, " Reservoir Porosity and Permeability Estimation from Well Logs using Fuzzy Logic and Neural Networks," *Paper SPE 88476*, Oct. 2004

Jude O. Amaefule; Mehmet Altunbay;, " Enhanced Reservoir Description: Using Core and Log Data to Identify Hydraulic (Flow) Units and Predict Permeability in Uncored Intervals/Wells," *Paper SPE 26436*, Oct. 1993

K. Aminian; S. Ameri; A. Oyerokun; B. Thomas;, " Prediction of Flow Units and Permeability Using Artificial Neural Networks," *Paper SPE 83586*, 19-24 May 2003

K. Salaheddin; S. Laksana; M. Schöbel;, " Rock Typing approach for Reservoir Characterization of Ordovician Sandstones, Field Case Study, Concessions NC115/NC186, Murzuq Basin, Libya," *Paper SPE 128825*, 14-17 Feb. 2010

Kazeem A. Lawal; Mike O. Onyekonwn;, " A Robust Approach to Flow Unit Zonation," *Paper SPE 98830*, 1-3 Aug. 2005

M. Altunbay; D. Georgi; H. M. Takezaki;, " Permeability Prediction for Carbonates: Still a Challenge?," *Paper SPE 37753*, Mar. 1997

M. Sitouah; M. Salmeen; S. Oyemakinde; F. Anifowose; O. Abdullatif;, " Permeability Prediction from Specific Area, Porosity and Water Saturation Using Extreme Learning Machine and Decision Tree Techniques: A Case Study from Carbonate Reservoir," *Paper SPE 164161*, Mar. 2013

Maghsood Abbaszadeh; Hikari Fujii; Fujio Fujimoto;, " Permeability Prediction by Hydraulic Flow Units – Theory and Applications," *Paper SPE 30158-PA*, 1996

Mehdi Bagheripour Haghighi; Mehdi Shabaninejad; Khalil Afsari;, " A Permeability Predictive Model Based On Hydraulic Flow Unit for One of Iranian Carbonate Tight Gas Reservoir," *Paper SPE 142183*, 31 Dec. - 02 Feb. 2011

Mohammad Izadi; Ali Ghalambor;, " A New Approach in Permeability and Hydraulic Flow Unit Determination," *Paper SPE 151576*, 15-17 Feb. 2012

Oscar Osorio Peralta;, " Rock Types and Flow Units in Static and Dynamic Reservoir Modeling: Application to Mature Fields," *Paper SPE 122227*, 31 May - 03 Jun. 2009

P.W.M. Corbett; D.K. Potter;, " Petrotyping: a Basemap and Atlas for Navigating Through Permeability and Porosity Data for Reservoir Comparison and Permeability Prediction," *SCA2004-30*, 5-9 Oct. 2004

Patrick W. M. Corbett; Noreddin I. A. Mousa;, " Petrotype-based Sampling Applied in a Saturation Exponent Screening Study, Nubian Sandstone Formation, Sirt Basin, Libya," *Petrophysics.* 5. p. 264-270, Aug. 2010

Riyaz Kharrat; Ramin Mahdavi; Hashem Bagherpour; Shahab Hejri;, " Rock Type and Permeability Prediction of a Heterogeneous Carbonate Reservoir Using Artificial Neural Networks Based on Flow Zone Index Approach," *Paper SPE 120166*, Mar. 2009

S. E. M. Desouky;, " Predicting Permeability in Un-Cored Intervals/Wells Using Hydraulic Flow Unit Approach," *OMC-2003-118*, 2003

Shahab Mohaghegh; Reza Arefi; Samuei Ameri;, " A Methodological Approach for Reservoir Heterogeneity Characterization Using Artificial Neural Networks," *Paper SPE 28394*, Sep. 1994

Shamsuddin Shenawi; Hisham Al-Mohammadi; Muteb Faqehy;, " Development of Generalized Porosity-Permeability Transforms by Hydraulic Units for Carbonate Oil Reservoirs in Saudi Arabia," *Paper SPE 125380*, 19-21 Oct. 2009

Shawket Ghedan; Tadesse Weldu; Omar Al-Farisi;, " Hybrid Permeability Prediction Model for Heterogeneous Carbonate Reservoirs with Tarmat Layers Considering Different Levels of Cutoffs," *Paper SPE 138685*, Nov. 2010

Slatt, R. M.; Hopkins, G. L.;, " Scales of Geological Reservoir Description for Engineering Applications: North Sea Oilfield Example," *Journal of Petroleum Technology*, Feb. 1990

Syed Shujath Ali; Enamul Hossain; Md. Rafiul Hassan; Abdulazeez Abdulraheem;, " Hydraulic Unit Estimation from Predicted Permeability and Porosity Using Artificial Intelligence Techniques," *Paper SPE 164747*, 15-17 Apr. 2013

Testerman, J.;, " A Statistical Reservoir Zonation Technique," *Journal of Petroleum Technology*, 1962

Trevor Hastie; Robert Tibshirani; Jerome Friedman, The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Second Edition. Springer, 2011.

Vitae

Name   : Ahmed Yousif Alsahaf

Nationality   : Saudi Arabia

Date of Birth   : 9/6/1984

Email   : SAHAFAY@GMAIL.COM

Address   : P.O.BOX 8462, DHAHRAN 31311, SAUDI ARABIA

Academic Background   : I earned a Bachelor degree in Electronics and Computer
Engineering from the University of Nottingham (UK) in 2007.

Professional Background   : I joined Saudi Aramco in 2002. For the period of 2007-
2010, I worked as web systems developer. Since then, I'm involved in the
corporate Knowledge Transfer initiative under the Training & Development
organization.