

Neural Networks for Flow Bottom Hole Pressure Prediction

Medhat Awadalla^{1,2}, Hassan Yousef¹

¹ Electrical and Computer Engineering Department, SQU, Oman

² Helwan University, Department of Communications, Electronics, and Computers, Egypt

Article Info

Article history:

Received Feb 10, 2016

Revised May 26, 2016

Accepted Jun 8, 2016

Keyword:

Feed-forward neural network
Flowing bottom-hole pressure
Hole pressure
Radial basis neural network
The empirical modes

ABSTRACT

Installation of down-hole gauges in oil wells to determine Flowing Bottom-Hole Pressure (FBHP) is a dominant process especially in wells lifted with electrical submersible pumps. However, intervening a well occasionally is an exhaustive task, associated with production risk, and interruption. The previous empirical correlations and mechanistic models failed to provide a satisfactory and reliable tool for estimating pressure drop in multiphase flowing wells. This paper aims to find the optimum parameters of Feed-Forward Neural Network (FFNN) with back-propagation algorithm to predict the flowing bottom-hole pressure in vertical oil wells. The developed neural network models rely on a large amount of available historical data measured from actual different oil fields. The unsurpassed number of neural network layers, the number of neurons per layer, and the number of trained samples required to get an outstanding performance have been obtained. Intensive experiments have been conducted and for the sake of qualitative comparison, Radial Basis neural and network and the empirical modes have been developed. The paper showed that the accuracy of FBHP estimation using FFNN with two hidden layer model is better than FFNN with single hidden layer model, Radial Basis neural network, and the empirical model in terms of data set used, mean square error, and the correlation coefficient error. With best results of 1.4 root mean square error (RMSE), 1.4 standard deviation of relative error (STD), correlation coefficient (R) 1.0 and 99.4% of the test data sets achieved less than 5% error. The minimum sufficient number of data sets used in training ANN model can be low as 12.5% of the total data sets to give 3.4 RMSE and 97% of the test data achieved 90% accuracy.

Copyright © 2016 Institute of Advanced Engineering and Science.
All rights reserved.

Corresponding Author:

Medhat Awadalla
Electrical and Computer Engineering Department,
SQU, Oman.
Email: medhatha@squ.edu.om

1. INTRODUCTION

Petroleum engineers are always interested in finding appropriate and reliable tools to predict the productivity of horizontal well as accurate predictions seem very important to conduct technical and economical feasible studies before drilling the wells which is very costly [1]. With the increased utilization and deployment of permanent down-hole gauges, measuring flowing bottom-hole pressure (FBHP) gets relaxed and faster. However, these gauges require continuous maintenance and calibration to avoid erroneous readings. Also, by intervening a well from time to time to measure FBHP is an expensive task, associated with production risk and interruption. For these reasons, the motivation of the prediction of FBHP has been argued.

Flowing bottom hole pressure prediction in gas wells is an old petroleum engineering problem. There is a long history of attempts to develop empirical correlations to predict the pressure drop in pipes. Some of these attempts have produced correlations that provide good prediction in some cases [2]-[3].

However, their general applicability is questionable. Correlations that address only a specific class of problems exist. These types of correlation usually perform better than those which attempt to meet the need of a variety of problems. Usually, the higher the number of variables in the model the lesser the reliability and general applicability of the correlations. This is the result of using methodologies such as conventional regression analysis [4]. In such methodologies, the chances of correctly and completely capturing the relationship between variables decrease as the number of variables increases. Many parameters could be involved in these types of problems, such as gas-oil ratios in two-phase systems, water flow in three phase systems, and inclination angles of the pipe. Models proposed by the investigators are based on empirical wells contained modest amounts of gas and oil production rates correlations developed from laboratory studies.

Most of the existing methods for predicting FBHP require one or more assumptions [5]-[7] (e.g., steady state flow, ideal gas of constant viscosity, small and constant compressibility and constant viscosity fluid) be applied. These methods appear to be subjected to appreciable error unless better limits of applicability are defined.

Neural networks and artificial intelligent approaches have been used in different fields [8]-[10], however these approaches recently have been employed in the petroleum industry [11]-[12], but their potential has not been fully investigated. In areas where a pattern exists between sets of data, a successful correlation can be developed with an artificial neural network (ANN). The pattern recognition capability of ANN makes it a desirable tool to employ under a variety of conditions. When the data contains a relationship that is implicit in nature, a network such as Kohonen, Probabilistic, or Back-propagation may discover that relationship despite the complexity. Most applications of artificial neural networks (ANNs) in multi-phase flow are confined to pipes. Authors [13] found fairly good bed heights estimations using an ANN. Flow pattern and frictional pressure drop were predicted [14]-[16] using an ANN. Neural networks (NNs) estimated the flow pattern Bottom Hole Pressure with less than 5% error and frictional pressure drop with less than 30%. Satisfactory results have been found for three phase relative permeability compared with experiments using adopted a PSO and neurofuzzy models to train the perceptron and to predict pollutant levels in gas wells [17]-[19]. The approach was proved to be feasible and effective by applying to some real air-quality problems and by comparison with a simple back-propagation (BP) algorithm. Support vector machine approach has been also used in training ANNs for predicating flow bottom hole pressure [20]-[21]. Even though the mentioned approaches have obtained reseanoable results, but the standard structure of these approaches have been used. In the paper, an attempt to find the optimum structure parameters that affects the performance especially in case of Artificial neural netowrks will be accomplished. Furthermore, practical sets of data available from an oil field are used for learning, testing, and validating the designed schemes. To prove the effectiveness of the proposed neural networks in estimation of the FBHP, intensive performance analysis is carried out.

2. FLOWING BOTTOM-HOLE PRESSURE IMPORTANCE AND ELECTRICAL UBMERSIBLE PUMP (ESP) WELL SYSTEM

Electric Submersible Pumping (ESP) is one of the most commonly used methods of well production/fluid lifting in the oil and gas industry. It is responsible for the highest amount of total fluids produced (oil and water) by any artificial lift method and an ideal method for high water cut wells. Centrifugal pumps can be single-stage or multi-stage units. Single-stage pumps are mainly used when low to medium discharge pressure is required, while multi-stage pumps are designed to overcome higher discharge pressures. This is the case of ESP used in the petroleum industry where fluids must be lifted from deep formations.

ESP is normally installed at the end of the production tubing string, which is inserted inside a bigger piping called casing. Normally ESP installation depth is shallower than the formation (producing zone) depth. The pressure drop schematic of flowing oil well with ESP is shown in Figure 1 and Figure 2. The pressure drop lines of interest in this study are the lines labeled A, B, and C. The pressure at the top of line A is the well head pressure and the end of it is the pressure of the pump discharge. Line A represents the drop in pressure due to the hydrodynamic multiphase flowing column and the frictional losses in the tubing. Line B is the line difference between the discharge and intake pressures of the pump. Simply it represents the total pressure developed by ESP. Line C represents the drop in pressure between the pump intake and the perforations at the producing formation due to the hydraulic column and frictional loss in the casing below the pump. Its top end is the pump intake pressure and its lower end is the well flowing bottom-hole pressure. It is a normal practice to have online pressure measurements at, well-head, pump discharge and pump intake. In fields of the study, these measurements are recorded every 15 minutes. In this paper re-sampled data for daily records have been used.

The pressure measurements for the pump discharge and intake are normally obtained by permanent pressure gauges installed within ESP assembly. Unfortunately, FBHP at the perforations has no permanent measurements and in our case we almost have no records for FBHP of an ESP well due to difficulties of access and other restrictions. Therefore, the scope of the paper is limited to the estimation of the pressure drop along the lines A and B only.

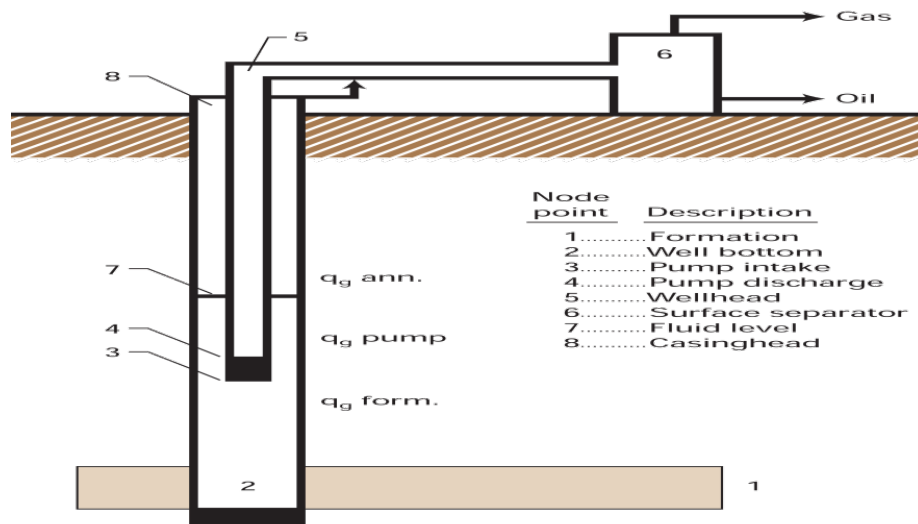


Figure 1. ESP Well System

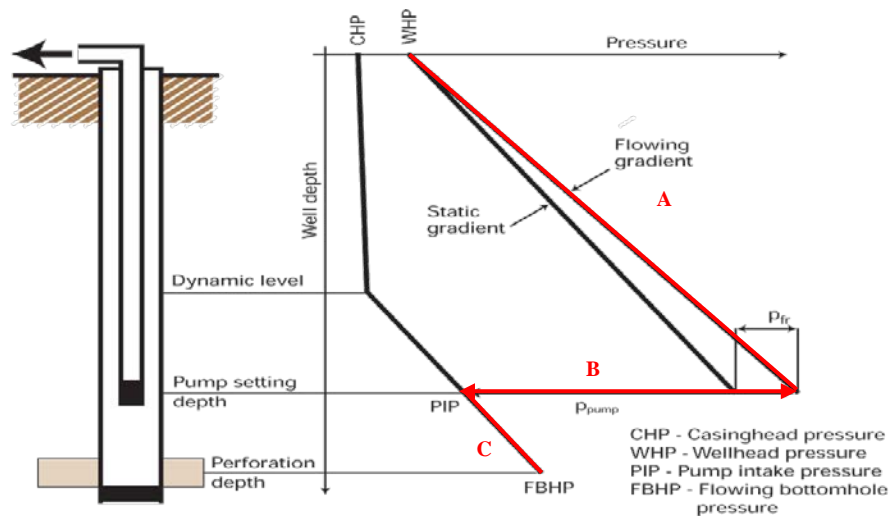


Figure 2. Typical Pressure Drop Profile Diagram

The estimation of the pressure drop along the line C (i.e. estimation of the pressure drop from the FBHP to the pump intake pressure) is less complex compared to drop in pressure along line A due to better homogenous flow and negligible friction. Hence, this pressure drop is out of the paper scope. Also, it would be very difficult to evaluate this estimation due to unavailability of FBHP records.

Flowing bottom-hole pressure of a well is the pressure that is measured or calculated at or near the producing formation at the bottom of the well while the well is flowing or producing hydrocarbons as shown in Figure 3. It is always higher than the flowing pressure at the surface, but lower than the shut in bottom-hole pressure.

Knowing the bottom-hole pressure of an oil well can help forecasting the well potential during the life cycle of the well. In other words, well production monitoring and artificial lifting optimization can be performed, which is a key objective for oil production maximization and operational cost reduction [12].

Bottom-hole pressure data also can be used to provide information on pore pressure that can be calculated for safety while drilling development wells in the area. It is critical for drilling operations especially underbalanced drilling. This also provides valuable data to select accurate kill fluid weight. The data also can be used to improve accurate under- or over-balance before perforation.

Tubing pressures and casing pressures of flowing wells have always been important factors in operating wells and under restricted production their importance is increased. Changes in these pressures, correlated with age or with rate of production, have been considered as giving important information as to the quality of the well, sand conditions, conditions of the bore hole through the sand, and whether the equipment in the hole is operating properly. A general study of bottom-hole pressures throughout an entire field has a direct application to the operation of a particular lease or an individual well. Bottom-hole pressure surveys of the field will provide data which will assist in making a more accurate estimate, much earlier in the life of the field, of the time when wells must be produced by artificial lift and of the amount of fluid that will have to be handled. It is of considerable value to know within reasonable limits when the wells will have to be pumped.

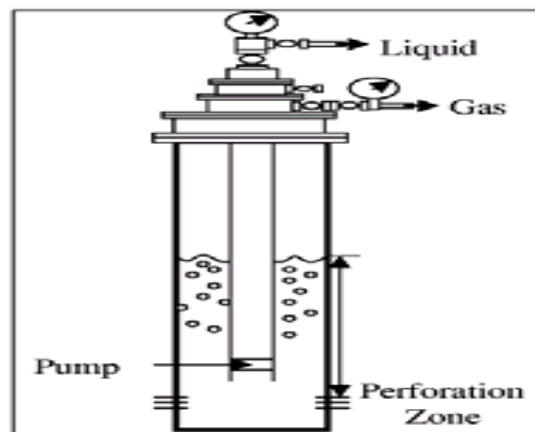


Figure 3. Schematic of Oil Well with ESP [12]

3. THE PROPOSED NEURAL NETWORK FOR FBHP ESTIMATION

Typically, FFNN consists of a set of sensory units (source nodes) that constitute the input layer, one or more hidden layers of computation nodes, and an output layer. The source nodes in the input layer of the network supply respective elements of the activation pattern (input vector), which constitute the input signals applied to the neurons (computation nodes) in the second layer (i.e., the first hidden layer). The output signals of the second layer are used as an input to the third layer, and so on for the rest of the network. The set of output signals of the neurons in the output (final layer) constitute the overall response of the network to the activation pattern supplied by the source nodes in the input (first) layer. A multilayer FFNN with one input, two hidden and one output layers is shown in Figure 4.

The activation function of the artificial neurons in FFNNs implementing the back propagation algorithm is a weighted sum (the sum of the inputs x multiplied by their respective weights w_{ji}):

$$A_j(x, w) = \sum_{i=0}^n x_i w_{ji} \quad (1)$$

The activation function depends only on the inputs and the weights. If the output function would be the identity (output=activation), then the neuron would be called linear. But, it has severe limitations and the most common output function is the sigmoidal function:

$$O_j(x, w) = \frac{1}{1 + e^{A(x, w)}} \quad (2)$$

The output depends only on the activation function, which in turn depends on the values of the inputs and their respective weights. Now, the goal of the training process is to obtain a desired output when certain inputs are given. Since the error is the difference between the actual output $O_j(x, w)$ and the desired

output d_j , the error depends on the weights, and we need to adjust the weights in order to minimize the error. The error function for the output of each neuron can be defined as:

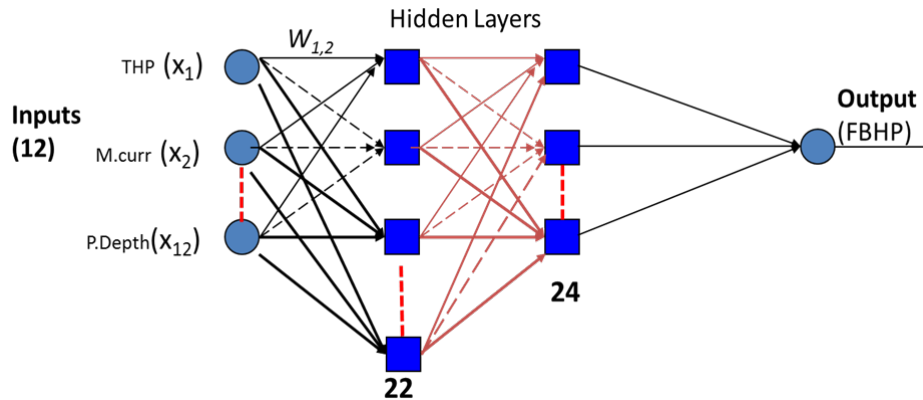


Figure 4. Multi-layer Feed forward neural network

$$E_j(x.w.d) = (O_j(x.w) - d_j)^2 \quad (3)$$

The error of the network will simply be the sum of the errors of all the neurons in the output layer:

$$E(x.w.d) = \sum_j (O_j(x.w) - d_j)^2 \quad (4)$$

Using the gradient descent to minimize the error (4), one can obtain the following adjustment rule for the weights

$$\Delta w_{ji} = -\eta \frac{\partial E}{\partial w_{ji}} \quad (5)$$

Equation (5) is used to get the updated weights as:

$$w_{ji}(k+1) = w_{ji}(k) + \Delta w_{ji} \quad (6)$$

4. NEURAL NETWORKS DEVELOPMENT AND OPTIMIZATION

In order to construct the neural network models that capable of estimating the FBHP, realistic data sets available from Oman oil fields, specifically from oil production wells lifted with ESP, are used. The data were collected from three different fields, Field-A, Field-B and Field-C. All three fields have water injection as reservoir pressure support and all of them have well production with two different artificial lifting; namely, ESP and gas-lift. There are twelve different input variables and one output variable for the ANN to be constructed. List of these variables, their units and their recording frequency are listed in Table 1. The motor current and pressures data are obtained from online-meters measurements that record samples every 15 minutes in a database historian system. The production data rates and ratios are obtained from testing the well production at dedicated three phase test separator equipped with individual phase flow meters. The measurement frequency is every one to three months, depending on the well oil production rate. The data pertaining to fluids properties is obtained from laboratory analysis that is done once for every well or field.

Fifteen-minute frequency data is re-sampled to daily frequency records to reduce number of samples while keeping reasonable representative variation in the data. Since the study scope is to estimate the FBHP, then, the data sets for well static conditions (i.e. when well is not flowing) were excluded. Also, incomplete data sets with readings over the meter range are excluded. Then, monthly production data is aligned with the daily data by replicating the monthly data daily until next month sample. Well (field) fluid data is then aligned with the daily data sets for each well (field). Before training the neural network model, the data is normalized. After training, the results of data sets are de-normalized and reorganized back to wells and timing sequence.

Table 1. ANN Description of Input and Output Variables

	Input	Abbreviation	Unit	Measurement Frequency
1	Tubing Head Pressure	THP	KPa	15 minutes
2	Motor Current	M.Curr	Am	15 minutes
3	Liquid Production Rate	Gross	m ³ /d	1-3 months
4	Oil Production Rate	Oil	m ³ /d	1-3 months
5	Water Production Rate	Water	m ³ /d	1-3 months
6	Gas Production Rate	Gas	m ³ /d	1-3 months
7	Base Sediment & Water (water cut)	BS&W	%	1-3 months
8	Formation Gas Oil Ratio	FGOR	-	1-3 months
9	Oil Specific Gravity	Oil API	-	one per well
10	Produced Water Specific gravity	Water API	-	one per well
11	Pump Intake True Vertical Depth	P.Depth	m	fixed per pump
12	Pump Discharge Pressure	Disch.P	KPa	15 minutes
Output				
1	Pump Intake Pressure	Intake.P	KPa	15 inutes

4.1. FFNN with a single hidden Layer Structure

A very basic structure of FFNN with one hidden layer is considered. To reach to the optimal number of hidden neurons, a test of the network is started with 4 neurons in the hidden layer and then increased the number of neurons in a multiple of four. In each case, the network is trained to a specified training error goal of 0.005. The data used for the neural network model is taken from Field-A. The network performance statistical factors are recorded. The main performance metrics considered are the relative root mean square error (RMSE) and standard deviation (STD). Also, the percentage of the test data attaining 95% and 90% accuracy of FBHP estimation are used as a secondary performance metric. Figure 5 shows the network performance against the number of neurons in the hidden layer.

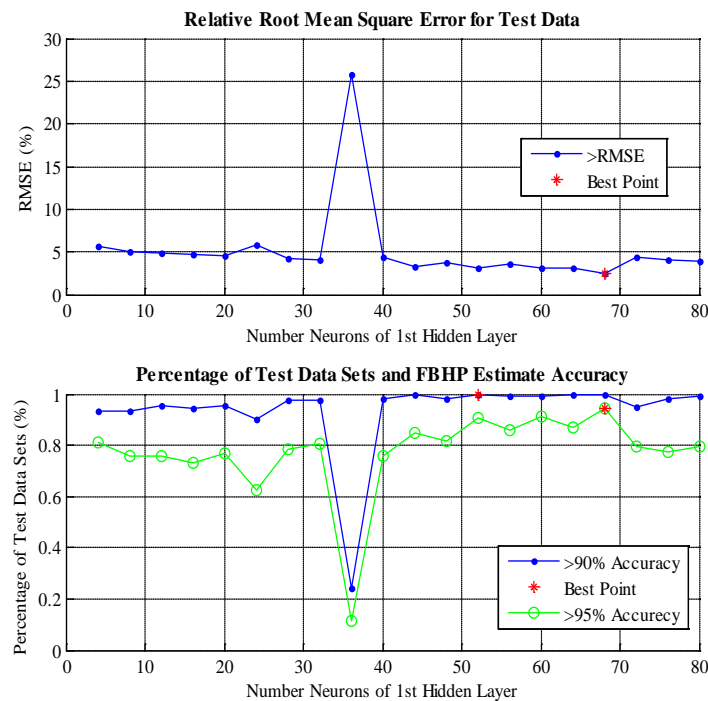


Figure 5. RMSE and Accuracy of FF BPNN vs. Number of Neurons of single Hidden Layer

The results indicate that the optimal number of neurons that achieves minimum RMSE is 68. Also, this result is supported from the overall accuracy of both 95% and 90% accuracy trends. Table 2 summarizes all statistical indicators of the network performance. It indicates that a remarkable performance of single hidden layer is achieved with 68 neurons where the RMSE is 2.53% and the STD is 2.44%. Also the percentage of the test data that showed intake pressure estimations within the 5% and 10% error from the actual intake pressure measurements are 94.4% and 99.8% respectively. The selected structure is addressed

again with different training mean square error goals and the performance is analyzed. The network is trained and tested for twenty different training mean square error goals within the range of 0.0001 to 0.04. Figure 6 illustrates the achieved results in terms of RMSE and accuracy.

Table 2. FFNN with 68 Number of Neurons in single Hidden Layer

	%Relative Error (Test Data)	%Relative Error (Training Data)
Root Mean Square Error	2.5264	1.6958
STD of Error	2.4378	1.6130
Correlation Coefficient	0.9921	0.9953
% of Data >95% Accuracy	0.9437	0.9870
% of Data >90% Accuracy	0.9969	0.9982
ErrorAvg	0.6650	0.5240
Abs ErrorAvg	1.8276	1.0998
Error min	0.0003	0.0004
Error max	16.9147	20.1590

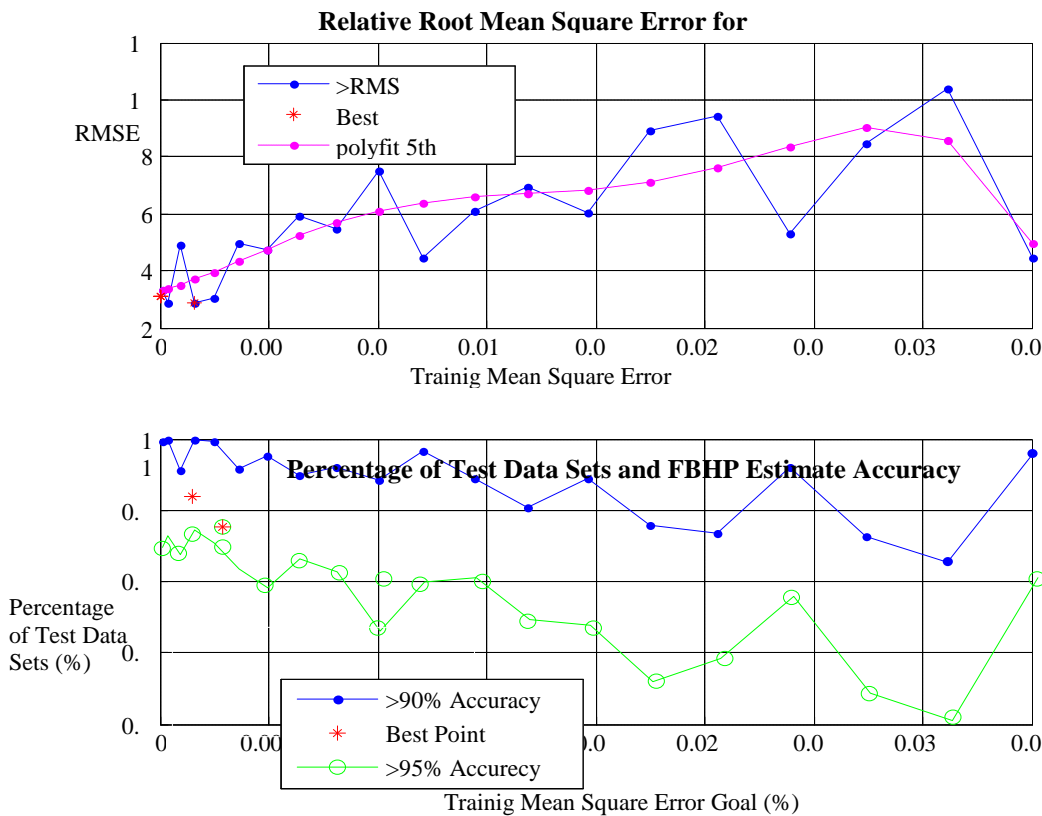


Figure 6. RMSE and Accuracy of FF with BPNN {68} vs. Training Error Goal

4.2. Investigating a FFNN with Two Hidden Layers Structure

In this section, the two hidden layer option is investigated. The objective of the analysis is to select the number of hidden neurons for each hidden layer that will yield the best network performance. To start with, a rule of thumb is used to select 15 neurons for the second hidden layer. Then, the network is trained and tested for different numbers of neurons for the first hidden layer starting from 4 neurons then increasing the number of neurons in multiple of four. Then, the network performance analysis is carried out for each case using the testing data as done with the single hidden layer case. The achieved results have been demonstrated in Figure 7. It is obvious that, the best RMSE is 3.6% occurs at 20 neurons are used in the first hidden layer. This is also supported by good accuracy points of 99% of test data fall within the 10% error and of 85% of test data fall within the 5% error. Then, the number of neurons in the first hidden layer is selected

to be 20 neurons. The network is trained and tested again for different numbers of neurons in the second hidden layer. Performance analysis for different number of neurons in the second hidden layer were carried out starting with 2 neurons for the second hidden layer and then increasing the number of neurons in multiple of two until reached 40 neurons. Again, the obtained results are depicted in Figure 8.

Based on the achieved results in Figure 8, it is clear that, the best RMSE is about 2.6% occurs when 24 neurons are used in the second hidden layer. This is also supported by the good accuracy points of 99% of test data fall within 10% error band and 92% fall within 5% error band. Moreover, the tuning of the number of neurons for the first hidden layer is repeated with 24 neurons for the second hidden layer.

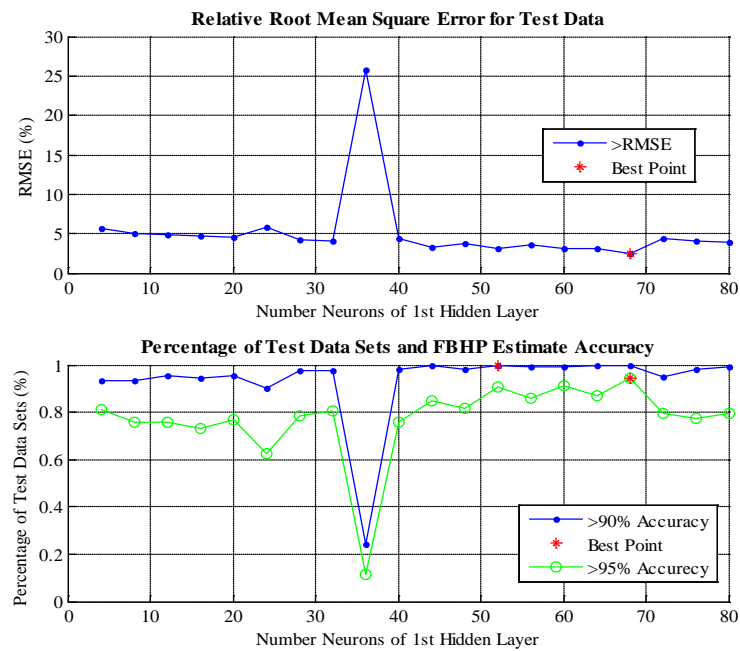


Figure 7. RMSE and Accuracy of FFNN with {x 15} vs. Number of Neurons 1st Hidden Layer

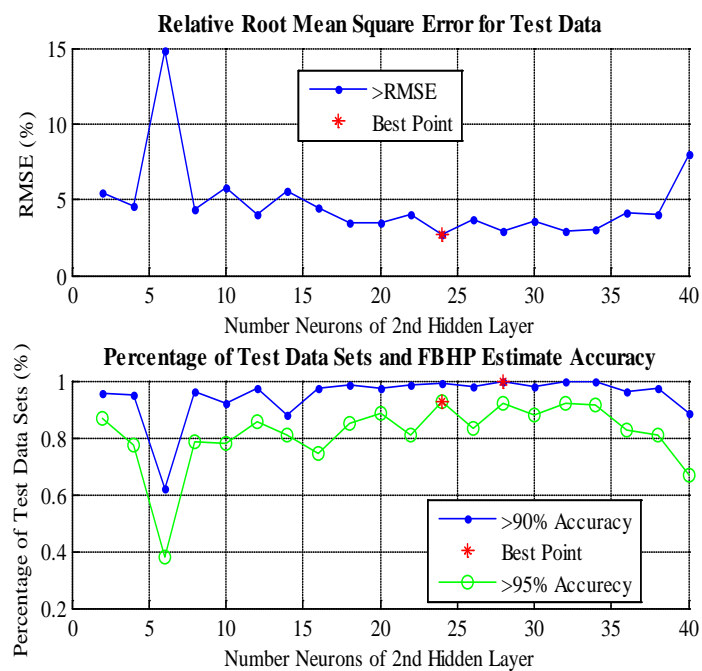


Figure 8. RMSE and Accuracy of FFNN with {20 x} vs. Number of Neurons 2nd Hidden Layer

As illustrated in Figure 9, the best RMSE is about 2.9% occurs for 22 neurons in the first hidden layer. This is also supported by the good accuracy points of 99% of test data fall within 10% error band and 91% fall within 5% error band. Therefore, the selected number of neurons in the two hidden layers are 22 and 24 neurons for the first and second hidden layers respectively. A further improvement to the neural network structure is tried by investigating the best selection of the training mean square error goal as shown in Figure 10.

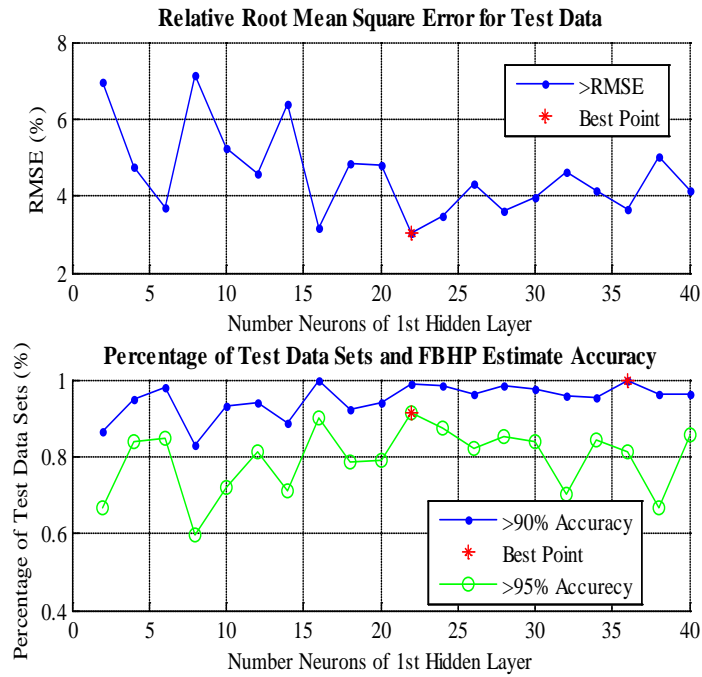


Figure 9. RMSE and Accuracy of FFNN with {x 24} vs. Number of Neurons 1st Hidden Layer

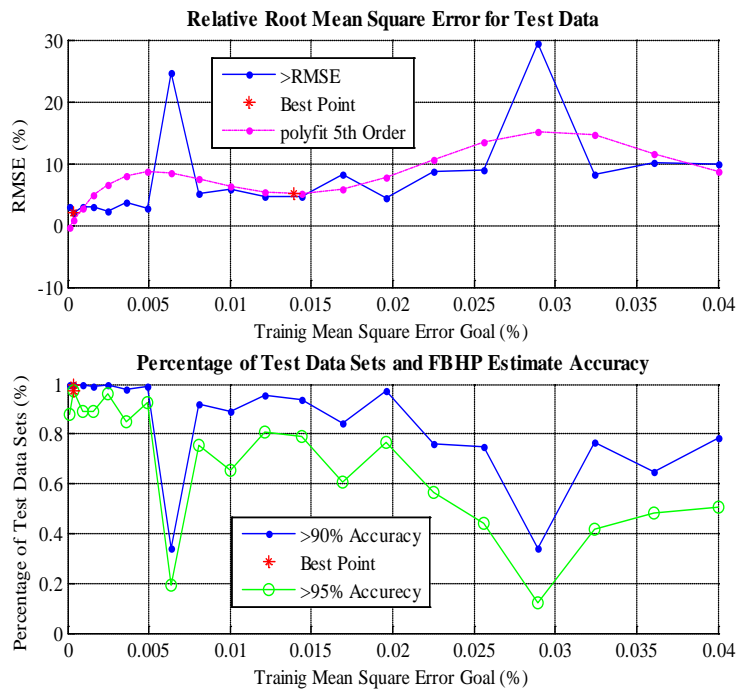


Figure 1. RMSE and Accuracy of FFNN with {22 24} vs. Training Error Goal

Similar to the single hidden layer analysis of the training mean square error goal, it is shown that, the RMSE is proportional to the training mean square error goal. The smaller the goal is the least the RMSE of the test data. As demonstrated in Figure 10, the selection of the training mean square error goal between 0.0001 and 0.005 would result in an acceptable small RMSE when testing the model. For simplicity, the training mean square error goal is selected to be 0.001. Therefore, the final feed-forward neural network structure with two hidden layer is the {22 24} neurons and 0.001 training mean square error goal.

5. EXPERIMENTAL RESULTS AND DISCUSSIONS

In this section, the constructed single-layer and multilayer FFNN are used to estimate the FBHP. The obtained results of the neural networks models and their performance analysis are illustrated.

5.1. Intake Pressure Estimation using FFNN with Two Hidden Layers in Field-A

The data analysis is performed for the data collected from 15 wells in the Field-A, below are the results for the FFNN with two hidden layer with 22 neurons in the first layer and 24 neurons in the second layer {22 24} and 0.001 training mean square error goal. The number of training data sets is 6000 with 60%, 20%, and 20% of training data are used for training, validation and testing respectively and the maximum number of epochs is 100. Table-3 shows the summary of some statistical performance indicators of the model for both the additional test data sets after training and for the training data sets. The main three indicators are RMSE, STD of error and the correlation coefficient of the relative error for the test data sets after training, R. The value of these parameters after the testing phase are 2.16%, 2.07% and 0.993 respectively.

Table 3. Field-A Performance Results Using FFNN {22 24} Model

	Abs Error KPa(Test Sets)	%Relative Error (Test Sets)	Abs Error KPa (Train sets)	%Relative Error (Train Sets)
Root Mean Square Error RMSE	149.554	2.164	121.516	1.872
STD of Error	140.727	2.067	121.290	1.864
Error Avg	50.695	0.643	-7.563	-0.180
Abs Error Avg	103.290	1.481	76.001	1.153
Error min	0.009	0.000	0.004	0.000
Error max	1554.999	19.040	1398.879	19.620
Correlation Coefficient R	-	0.993	-	0.995

The actual intake pressure and the corresponding estimates from FFNN {22 24} model are shown in Figure 11. These data were rearranged into well by well intake pressure sets for the purpose of analysis and clear illustration. It is clear that, the model estimates are superbly close to the actual measurements with minor errors at cases where the intake pressure of the well is fluctuating sharply which could be due to unstable well flow. These values are at data sets number 900-1000 and 2200-2400. Therefore, we could state that the model accuracy slightly decreases under unstable well flow conditions. Figure 12 shows the model estimates of intake pressure versus the actual intake pressure measurements. It is obvious that the best linear fit of the data is almost identical to the estimated values. This is very powerful indicator of the model's worthy accuracy over full measurements range. This is also supported by the good correlation coefficient factor, $R=0.998$. The slight deviation over the 7000-8000 range is explained due to the unstable flow described previously. Figure 13 shows the model estimates error against the percentage of the test data sets and their corresponding errors in the estimate of the intake pressure. It shows that almost 100% (all the 2560) of the test data have returned estimates within $\pm 10\%$ error (within 90% accuracy). Furthermore, the derived calculations show that 96% of the test data sets have their intake pressure estimates with 95% accuracy. Therefore, these results confirm the robust capability of the model to estimate the intake pressure with a remarkable accuracy.

By summing the final hidden neurons weights for each individual input and calculating its percentage of the total inputs weights, then the relevancy of individual inputs to the final output (intake pressure estimate) can be analyzed. Figure 14 shows the individual inputs relevancies to the contribution of calculating the final intake pressure estimate. It is found that, the most relevant input is the pump intake depth followed by the pump discharge pressure. In general, all the twelve inputs have similar relevancy within 6-10 % with the exception of the pump intake depth which have the highest relevancy of about 14%.

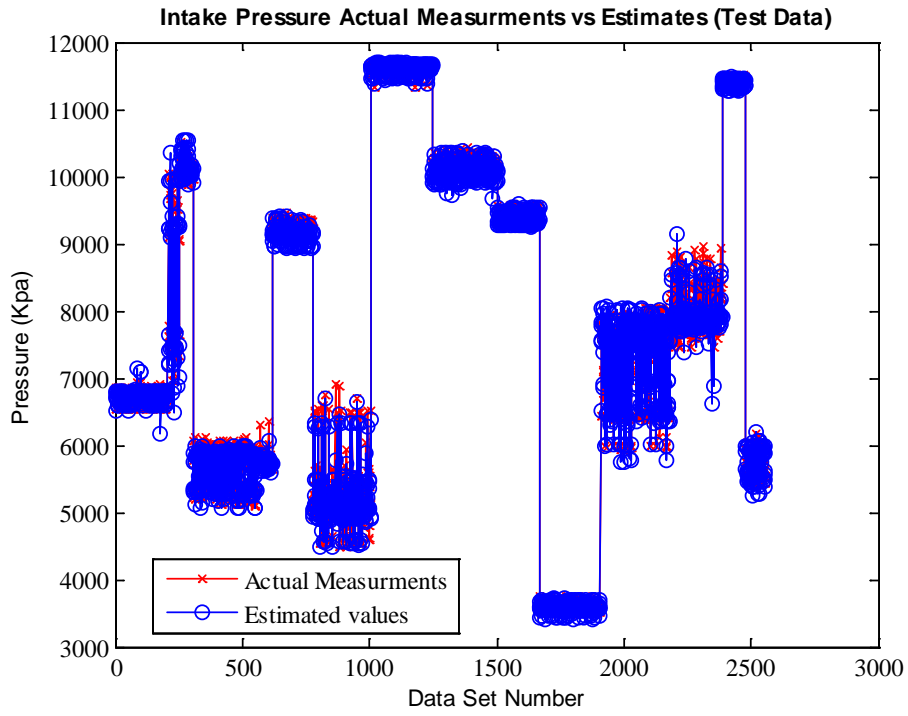


Figure 2. Intake Pressure Measurements and FFNN {22 24} Model Estimations

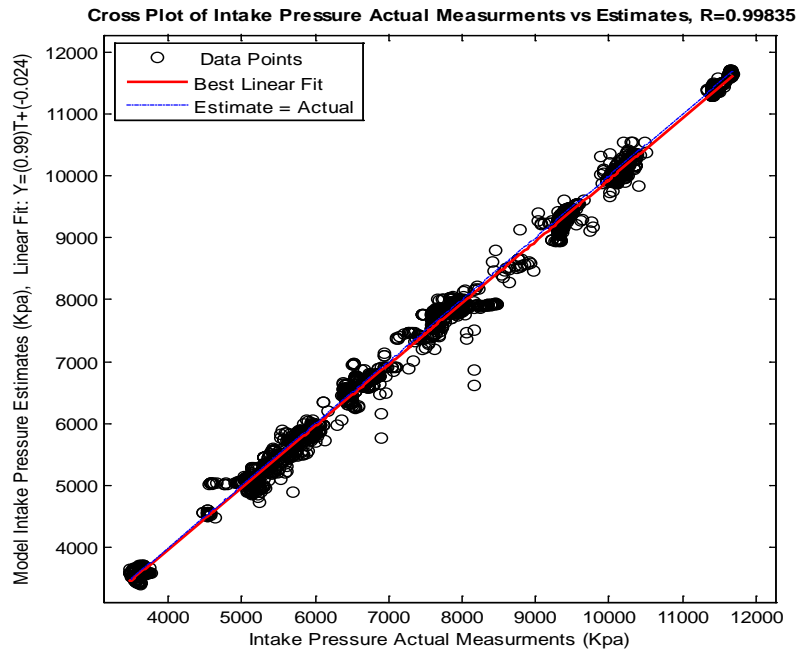


Figure 12. Cross Plot of Intake Pressure Measurements vs. FFNN {22 24} Model Estimations

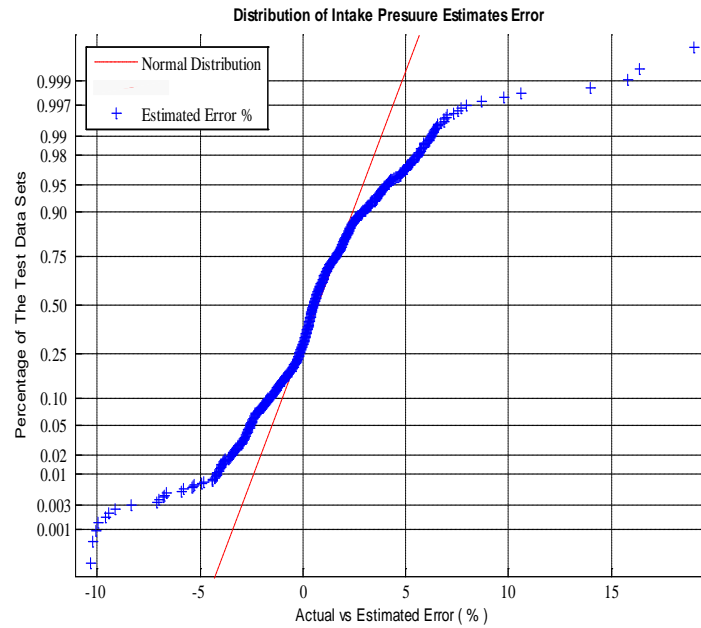


Figure 13. Distribution of Intake Pressure Estimations Error of FFNN {22 24} Model

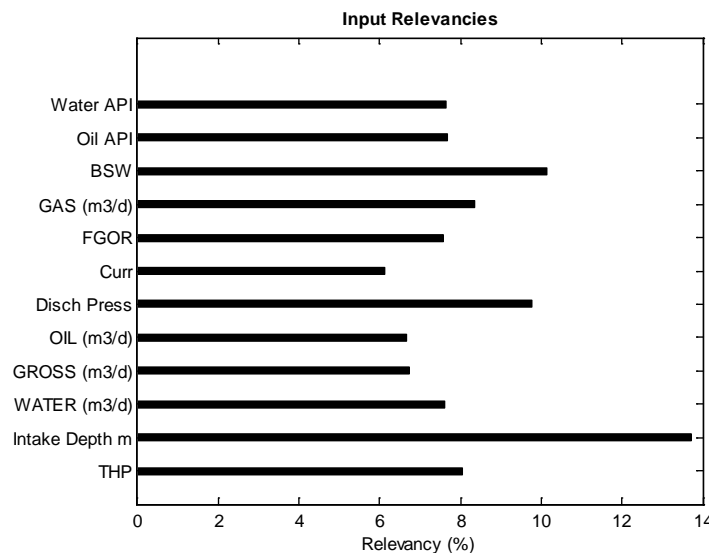


Figure 14. Inputs Relevancy of FFNN {22 24} Model Estimations for Field-A

5.2. Intake Pressure Estimation using FFNN with Single Hidden Layer

Table 4 shows the summary of the statistical performance indicators of the model for both the additional test data sets and for the training data sets. Again, the same three indicators are used - RMSE, STD of error and the correlation coefficient of the relative error for the test data sets, R. They show a respectable performance with 2.3, 2.3 and 0.0993 values respectively. These performance matrices are very close to the results obtained from the two hidden layer model.

Figure 15 shows the actual intake pressure and the corresponding estimates from the model. It is clear that, the model estimates are superbly close to the actual measurements with minor error at cases where the intake pressure of the well is fluctuating sharply which could be due unstable well flow. These values can be seen at data sets number 2200-2400. Therefore, we could state that the model accuracy slightly decreases under unstable well flow conditions. Figure 16 shows the cross plot of the actual intake pressure measurements versus the model estimates. The chart illustrates that the best linear fit of the data is almost

perfectly falls on top of the line where the actual values equal the estimate ones (45 degree). It indicates the remarkable accuracy of the model over the full measurements range. This is also supported by the remarkable correlation coefficient factor, $R= 0.998$. The slight deviation over the 7000-8000 range is again due to the unstable flow described previously. Figure 17 shows the model estimates error distribution. It shows that almost of 2560 test data have returned estimates within $\pm 10\%$ error (90% accuracy). Furthermore, the derived calculations show that 97% of the test data sets have their intake pressure estimates within 95% accuracy. Again, these analyses present the robust capability of the model to estimate the intake pressure.

Figure 18 shows the individual inputs relevancies to the contribution of calculating the final intake pressure estimate. All inputs have similar relevancy within 6-10 % with exception of the pump intake depth which have the highest relevancy of about 12%. Table 3 and Table 4 show that the significant performance of the two models exception for data of the field-C where a slight drop in performance is noticed. Further analysis has been done to address this drop. It might be related to down-hole measurements of intake pressure accuracy of Field-C or the high gas to liquid ratio characteristic of the Field-C. By analyzing the achieved performance indicators, it is clear that FFNN with two hidden layers outperforms FFNN with a single hidden layer.

Table 4. Field-A Performance Results Using FFNN {68} Model

	Abs Error KPa (Test Sets)	%Relative Error (Test Sets)	Abs Error KPa (Train sets)	%Relative Error (Train Sets)
Root Mean Square Error RMSE	142.5	2.3	116.1	1.8
STD of Error	142.6	2.3	91.5	1.5
Error Avg	1.4	0.1	-71.5	-1.1
Abs Error Avg	111.7	1.7	85.2	1.3
Error min	0.1	0	0	0
Error max	1047.6	15.2	1915.2	20.4
Correlation Coefficient R	-	0.9927	-	0.9945

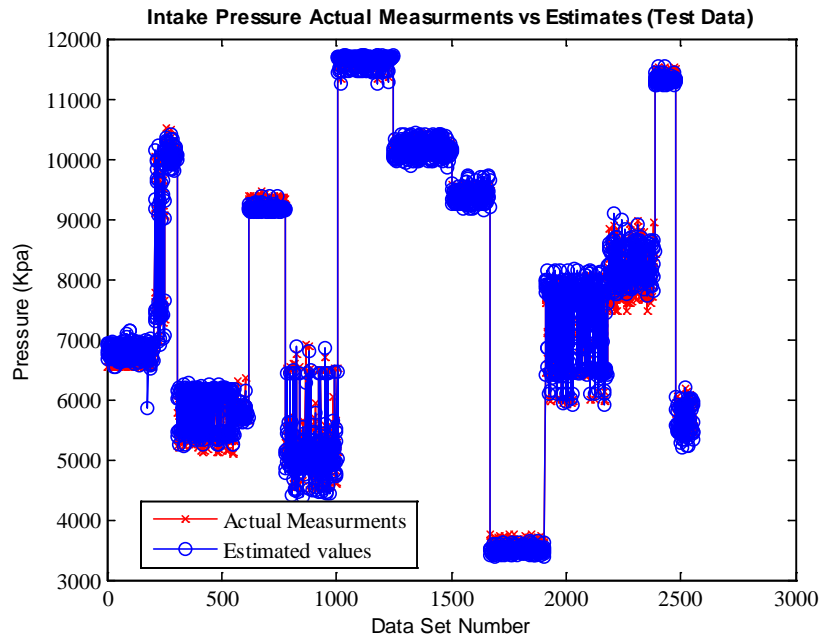


Figure 15. Intake Pressure Measurements and FFNN {68} Model Estimations

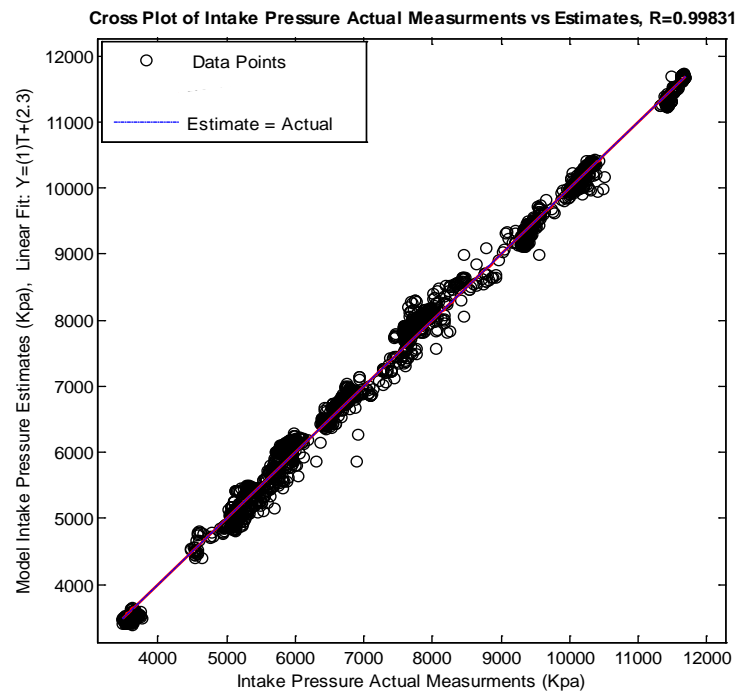


Figure 16. Cross Plot of Intake Pressure Measurements vs. FFNN {68} Model Estimations

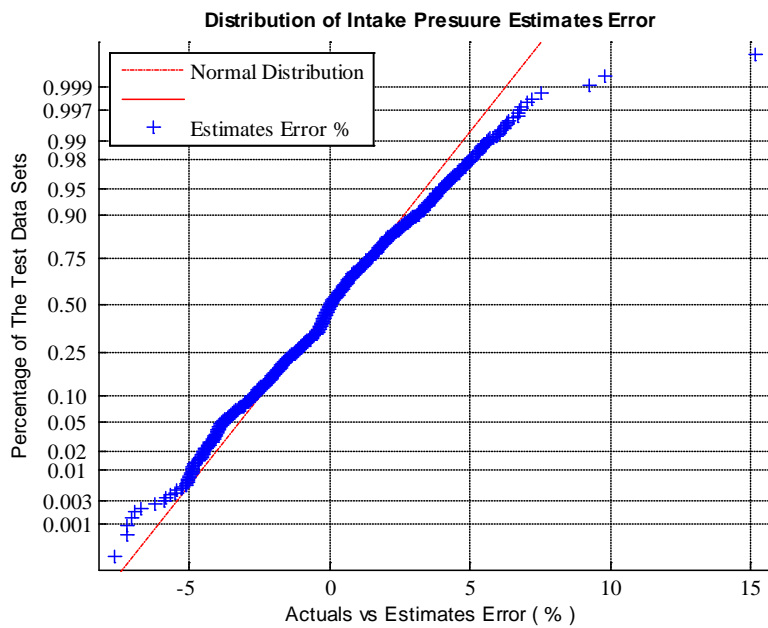


Figure 17. Distribution of Intake Pressure Estimations Error of FFNN {68} Model

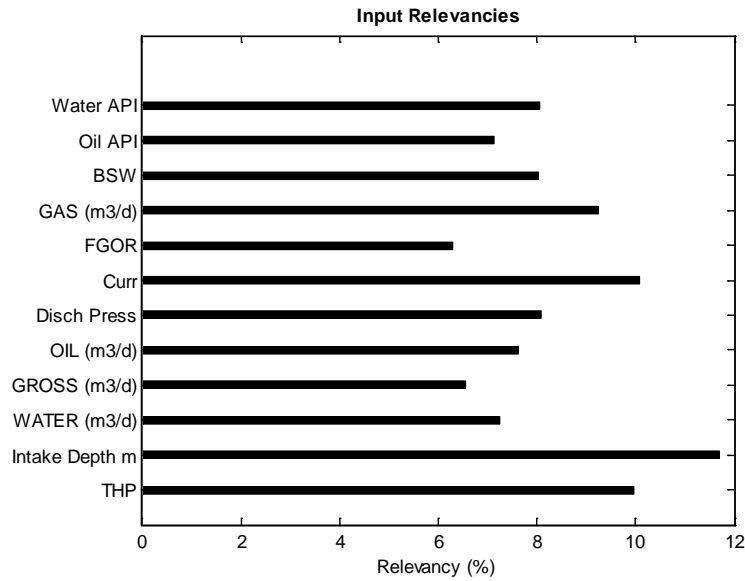


Figure 18. Inputs Relevancy of FFNN {68} Model Estimations for Field-A

5.3. Number of Data Sets Used in Model Training

To determine the minimum sufficient number of data sets used to train a two hidden layer model with {22 24} neurons, the neural network is trained, tested and the performance parameters are recorded. This was repeated for ten different number of data sets increasing from 3 sets to 3000 sets. The achieved results are shown in Figure 19. It is shown that the accuracy of estimating FBHP increases with increasing the number of data sets used in training. However, a small number of training data sets, 375 sets, can still be used for training within an acceptable estimation accuracy. The achieved RMSE is 3.4 while 97% and 93% of the test data achieved 90% accuracy and 95% accuracy respectively. With around 1500 data sets, the model resulted in a RMSE of 2.1 and 99.8% and 98% of the test data achieved 90% accuracy and 95% accuracy respectively.

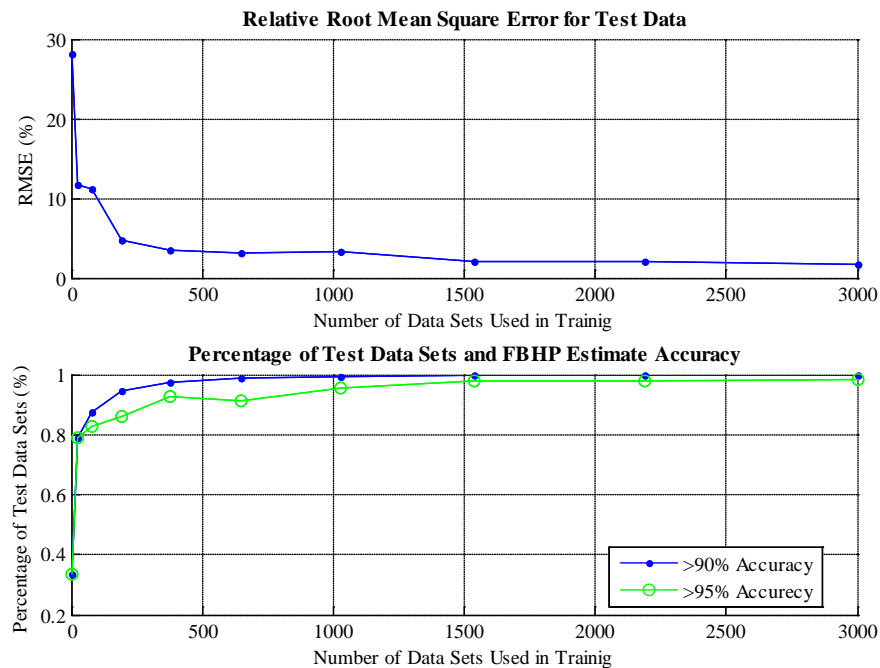


Figure 19. Results Performance of {22 24} Model vs Number of Data Sets used in Training

5.4. Model Generalization

To test ANN models for generalization across different fields, Radial Basis neural network, RB NN, has been developed for the sake of comparison. The models have been tested using data sets from another different field unused to train and develop the model. For example, a {22 24} FFNN model developed (trained) using data from field-A and then tested using data sets from Field-B. This cross field testing showed a poor FBHP estimation accuracy. Although, the same model has shown a good FBHP estimation when tested using data sets from the same field even though that well was not used in the training process. This is a clear illustration and prove of the limitation of ANN model applicability to the range of the used data in training the model. This can be explained as a result of the capability of the ANN to extract and model the hidden complexity of system modeled using its power of feature detection due to parameters' variations and changes. In fact, each field has some unique parameters that are not changing within the field and physically linked to FBHP, however they are not used in ANN modeling inputs. Consequently, ANN model will be able to estimate FBHP well within the same field (these parameters are not changing), but when moving across different fields (these parameters are changing) and ANN logically will fail to estimate FBHP as it has not learned the impact of their variances on FBHP. Figure 20 shows the achieved results of single and two-layer neural network results for A/B/C Fields and gives the comparison between them and Radial Basis model achieved results.

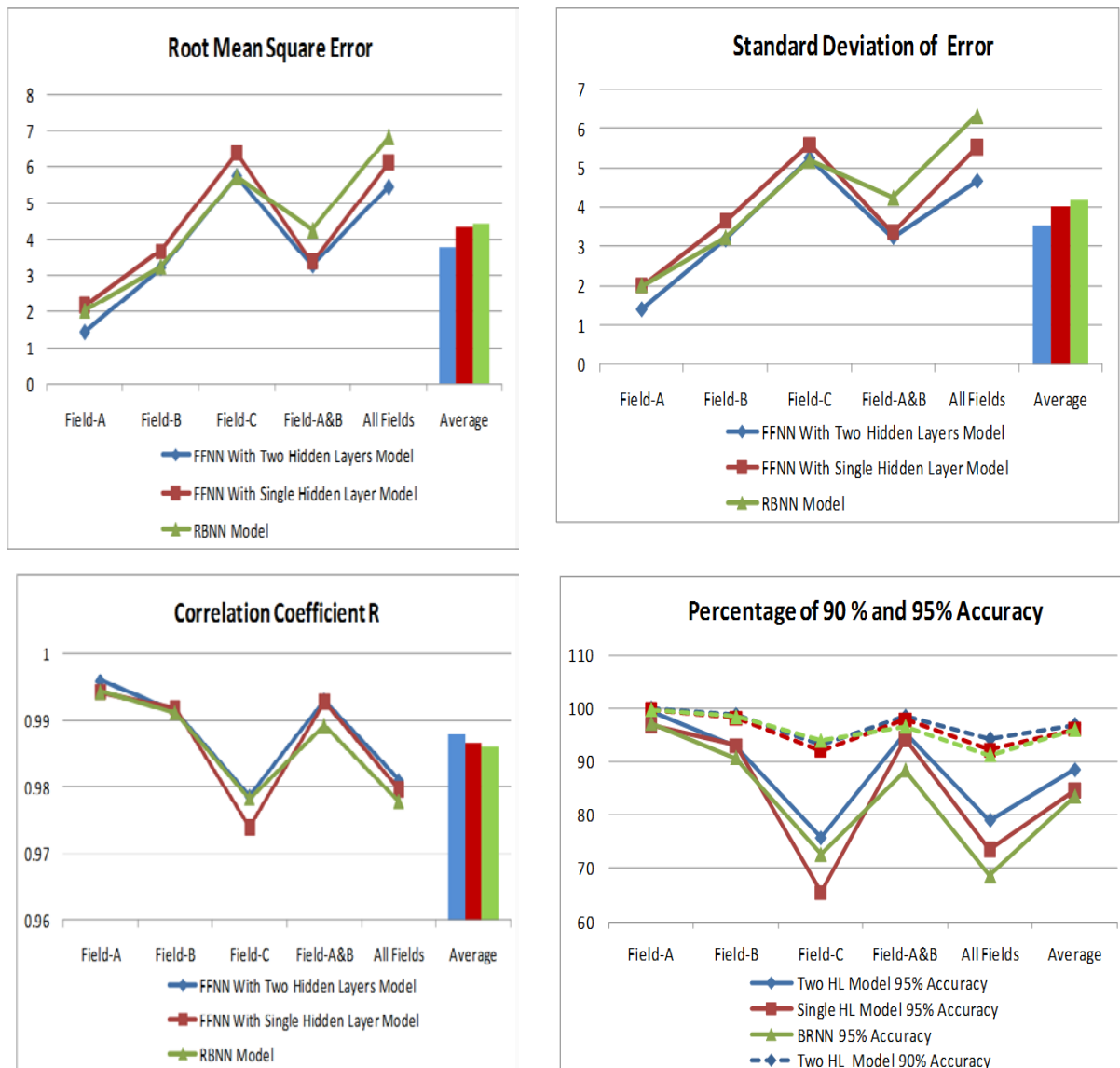


Figure 30. Trends of Single Layer, Multilayer and RBNN Models Results for A/B/C Fields

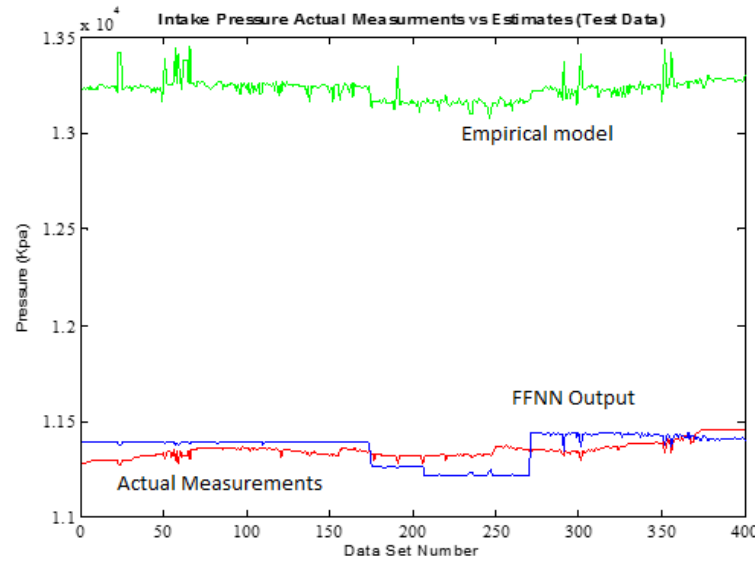


Figure 21. Performance Comparison between FFNN {22 24} and Empirical Models

Figure 20 illustrates that the overall performance of the three models is significant, however overall performance of field-C is slightly less. This drop might come back to the accuracy of down-hole measurements of intake pressure or high gas to liquid ratio of Field-C. based on the achieved results, FFNN with two hidden layers outperforms the others. Furthermore, Figure-21 shows that the performance of the best achieved model, two-layer neural network, outperforms the simplified developed empirical model.

6. CONCLUSION

This paper presented one and two hidden feed forward with back-propagation algorithm artificial neural network models to estimate FBHP of oil wells. The developed neural network models rely on a large amount of available historical data on oil wells. They have shown exceptionally accurate FBHP performance estimation that significantly outperforms the empirical model and Radial Basis neural network. The number of neural network layers and the number of neurons per layer have been developed to find the optimum neural network structure. The paper shows that the accuracy of FBHP estimation using FFNN with two hidden layer model is better than FFNN with single hidden layer model in terms of data set used, mean square error, and the correlation coefficient error. With best results of 1.4 root mean square error (RMSE), 1.4 standard deviation of relative error (STD), correlation coefficient (R) 1.0 and 99.4% of the test data sets achieved less than 5% error. The minimum sufficient number of data sets used in training the ANN model can be low as 375 sets only to give a 3.4 RMES and 97% of the test data achieved 90% accuracy.

7. Compliance with Ethical Standards

The research leading to these results has received Project Funding from Sultan Qaboos University Internal Grant Agreement No [IG/ENG/ECED/16/1]. The authors would like to thank SQU for all the administrative and financial support.

REFERENCES

- [1] M. Ahmadi, *et al.*, "Determination of oil well production performance using artificial neural network (ANN) linked to the particle swarm optimization (PSO) tool," *Petroleum*, vol/issue: 1(2), pp. 118–132, 2015.
- [2] O. Adewale, "Optimization of natural gas field development using artificial neural networks," MSc Paper, The Pennsylvania State University, USA, 2010.
- [3] O. Adekomaya, *et al.*, "Predictive Tool for Bottom-Hole Pressure in Multiphase Flowing Wells," *Petroleum & Coal*, vol/issue: 50(3), pp. 67-73, 2008. ISSN 1337-7027.
- [4] M. Awadalla, *et al.*, "Prediction of Oil Well Flowing Bottom-hole Pressure in Petroleum Fields," *Sixth IEEE International Conference on Industrial Engineering and Operations Management*, Kuala Lumpur, March 8-10, 2016.

- [5] A. Mehta, *et al.*, "A Multi-Layer Artificial Neural Network Architecture Design for Load Forecasting in Power Systems," *International Scholarly and Scientific Research & Innovation*, vol/issue: 5(2), pp. 207-220, 2011.
- [6] M. Awadalla and Sadek M. A., "Spiking neural network-based control chart pattern recognition," *Alexandria Engineering Journal*, vol/issue: 51(1), pp. 27-35, 2012.
- [7] D. Beggs and J. P. Brill, "A study of two-phase flow in inclined pipes," *Journal of Petroleum Technology*, vol/issue: 25(5), pp. 607-617, 1973.
- [8] S. Bikbulatov, *et al.*, "Flowing Bottomhole Pressure Calculation for a Pumped Well under Multiphase Flow," Cornell University Library Web site <http://arxiv.org/abs/physics/0504083v1>, 2005.
- [9] T. Gabor, "Considerations on the Selection of an Optimum Vertical Multiphase Pressure Drop Prediction Model for Oil Wells," *Society of Petroleum Engineers*, pp. 1-5, 2001.
- [10] A. Alalousi, *et al.*, "A Preliminary Performance Evaluation of K-means, KNN and EM Unsupervised Machine Learning Methods for Network Flow Classification," *International Journal of Electrical and Computer Engineering (IJECE)*, vol/issue: 6(2), pp. 778-784, 2016.
- [11] S. Haykin, "Neural and Learning Machines," London, Pearson, 2009.
- [12] I. Jahanandish, *et al.*, "Predicting bottomhole pressure in vertical multiphase flowing wells using artificial neural networks," *Journal of Petroleum Science and Engineering*, vol. 75, pp. 336-342, 2011.
- [13] H. Jeff, "Introduction to Neural Networks," Chesterfield: Heaton Research. Retrieved from <http://www.heatonresearch.com/book/programming-neural-networks-java-2.html>, 2008.
- [14] M. Mohammadpoor, *et al.*, "A New Methodology for prediction of buttonhole flowing pressure in vertical Multiphase flow in Iranian Oil Fields using Artificial Neural Networks (ANNs)," *Society of Petroleum Engineers*, 2010.
- [15] Petroleum Exploration Company, "Products," 2013. Retrieved from <http://www.petex.com/products/?ssi=3>.
- [16] D. Himansu, *et al.*, "A Novel PSO Based Back Propagation Learning-MLP (PSO-BP-MLP) for Classification," *Proceedings of the International Conference on CIDM*, pp. 461-471, 2014.
- [17] M. Awadalla, "Power Efficient Scheduling Scheme based on PSO for Real time systems," *International journal of computer Applications*, vol/issue: 111(4), pp. 24-30, 2015.
- [18] M. Awadalla and A. Elewi, "Enhanced PSO Approach for Real Time Systems Scheduling," *7th International Conference on Computer Engineering and Technology (IC CET 2015)*, France, Paris, 2015.
- [19] A. Draidi and D. Labeled, "A Neuro-fuzzy Approach for Predicting Load Peak Profile," *International Journal of Electrical and Computer Engineering (IJECE)*, vol/issue: 5(6), pp. 1304 -1310, 2015.
- [20] M. A., Ahmadi and A. Bahadori, "Determination of oil well production performance using artificial neural network (ANN) linked to the particle swarm optimization (PSO) tool Petroleum," 2015. <http://dx.doi.org/10.1016/j.petlm>.
- [21] M. A. Ahmadi, *et al.*, "Prediction breakthrough time of water coning in the fractured reservoirs by implementing low parameter support vector machine approach," *Fuel*, vol/issue: 117(2014), pp. 579-589, 2014.