

## Research Article

# Application of Artificial Intelligence Techniques in Predicting the Lost Circulation Zones Using Drilling Sensors

Abdulmalek Ahmed,<sup>1</sup> Salaheldin Elkatatny ,<sup>1</sup> Abdulwahab Ali,<sup>2</sup> Mahmoud Abughaban,<sup>3</sup> and Abdulazeez Abdurraheem<sup>1</sup>

<sup>1</sup>College of Petroleum Engineering and Geosciences, King Fahd University of Petroleum & Minerals, 31261 Dhahran, Saudi Arabia

<sup>2</sup>Center of Integrative Petroleum Research, King Fahd University of Petroleum & Minerals, Dhahran 31261, Saudi Arabia

<sup>3</sup>Saudi Aramco, EXPEC ARC, Saudi Arabia

Correspondence should be addressed to Salaheldin Elkatatny; [elkatatny@kfupm.edu.sa](mailto:elkatatny@kfupm.edu.sa)

Received 13 April 2020; Revised 1 September 2020; Accepted 9 September 2020; Published 22 September 2020

Academic Editor: Rafael Morales

Copyright © 2020 Abdulmalek Ahmed et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Drilling a high-pressure, high-temperature (HPHT) well involves many difficulties and challenges. One of the greatest difficulties is the loss of circulation. Almost 40% of the drilling cost is attributed to the drilling fluid, so the loss of the fluid considerably increases the total drilling cost. There are several approaches to avoid loss of return; one of these approaches is preventing the occurrence of the losses by identifying the lost circulation zones. Most of these approaches are difficult to apply due to some constraints in the field. The purpose of this work is to apply three artificial intelligence (AI) techniques, namely, functional networks (FN), artificial neural networks (ANN), and fuzzy logic (FL), to identify the lost circulation zones. Real-time surface drilling parameters of three wells were obtained using real-time drilling sensors. Well A was utilized for training and testing the three developed AI models, whereas Well B and Well C were utilized to validate them. High accuracy was achieved by the three AI models based on the root mean square error (RMSE), confusion matrix, and correlation coefficient ( $R$ ). All the AI models identified the lost circulation zones in Well A with high accuracy where the  $R$  is more than 0.98 and RMSE is less than 0.09. ANN is the most accurate model with  $R = 0.99$  and  $RMSE = 0.05$ . An ANN was able to predict the lost circulation zones in the unseen Well B and Well C with  $R = 0.946$  and  $RMSE = 0.165$  and  $R = 0.952$  and  $RMSE = 0.155$ , respectively.

## 1. Introduction

The demand for drilling high-pressure, high-temperature (HPHT) wells has become more significant in the petroleum industry. HPHT wells are known for bottom-hole pressures of more than 10,000 psi and bottom-hole temperatures of more than 300°F [1]. The main advantages of drilling HPHT wells are increasing oil production and improving economic success [2]. Drilling these wells involves some challenges and difficulties, mainly the appropriate tools for formation evaluation. These include many areas such as HPHT cement integrity, modified testing procedures and equipment, battery technology, proper zonal isolation, elastomers, alternative sealing agents, electronics/sensors and selection of drilling mud, design of drill string, and bits [3].

Drilling HPHT wells can result in many problems that delay the drilling operation and impact the cost. Loss of circulation is a common problem in drilling these wells. The partial or entire loss of the drilling mud from the wellbore to the formation is called loss of circulation or loss of return [4]. The losses will occur when a path for the flow exists, while the pressure inside the well is greater than the formation pressure [5].

Loss of circulation increases the nonproductive time spent on mitigating the losses [6], besides increasing the total drilling cost due to loss of drilling mud, which represents, in some cases, 40% of the total cost. The oil and gas industry reported more than \$12 billion in the cost of drilling materials and fluids in 2018 [7]. Loss of circulation leads to poor hole cleaning due to the reduction of mud level in the

borehole, which decreases its ability to transfer the cutting outside the wellbore [8]. The decrease in the mud level might reduce the hydrostatic pressure and cause a kick or blowout if the wellbore pressure became less than the formation pressure [9].

A range of methods has been used to overcome the circulation loss. The first method is adjusting the properties of drilling mud to reduce the equivalent circulation density (ECD) and consequently decreasing the quantity of the lost drilling mud [10]. The second method is pumping the lost circulation material (LCM) to seal and plug the losses [11]. Nevertheless, these methods are time-consuming and very expensive [12].

To minimize loss of return, it is essential to identify the lost circulation zones. Although various approaches are available, such as ECD, temperature profile, and resistivity [13, 14], nevertheless, some of these approaches are impractical either due to the high cost or lack of technology or owing to inaccurate prediction of the thief zones.

## 2. Artificial Intelligence (AI)

Artificial intelligence (AI) allows computers to perform tasks that require human intelligence. According to Mohaghegh et al. [15], AI is aimed at building a model or an algorithm, which enables machines to perform duties that need knowledge, understanding, and experience when performed by humans. A broader definition includes problem-solving, language perception, and conscious and unconscious processes [16]. AI is also known as a subfield of computer science involving the use of computers in tasks, which usually needs reasoning, knowledge, learning, and understanding abilities.

*2.1. Artificial Neural Network (ANN).* ANN is an information-processing system, which attempts to imitate the performance features of the biological nervous system. The network is adapted as a computer model, which can advance transformations, associations, or mappings between data [17]. The feature of ANN is that it does not require any physical phenomenon that explains the system under study [18]. Any nonlinear complex function can be approximated by ANN to make a relationship between input and output parameters.

According to Ahmed et al. [19], an artificial neural network is made up of many components such as neurons, hidden layers, transfer function, learning function, training function, and epoch size. Neurons are components that have specific input/output, and they are connected to form a network of nodes that makes the neural networks [20]. Weights and biases are used to handle the input parameters to find a relationship between the neurons and the source, so the performance of the network depends on the selection of those weights and biases. Also, the performance of the estimation model relies on the choice of the number of hidden layers, training function, and number of neurons [21].

*2.2. Fuzzy Logic (FL).* Fuzzy logic (FL) is a method of reasoning where the rules of deduction are estimated rather than precise. FL is valuable for handling information that is

incomplete, inaccurate, or irresponsible. FL is closely similar to the theory of fuzzy groups that belongs to a set of objects with boundaries in which membership is a problem of degree [22].

The fuzzy system is typically used to characterize uncertainty, which is due to the imprecision of the data or insufficient input variables that have an essential effect on the results. A property or an item can be defined by categorizing it under one of the different noncrisp groups and also a degree of membership for every group [23]. The fuzzy set theory proposes that a truth value that is between 0 and 1 should be added when working with noncrisp variables.

A membership function is used to define the relationship between a truth value and its variable. It has a value between 0 and 1, and that describes the “degree” of membership [24]. The membership functions can be represented by different functions such as sigmoid, Gaussian, trapezoidal, or straight lines [25]. When set membership had been defined again in this method, you can explain a reasoning system based on techniques for relating distributions [26].

The fuzzy inference system (FIS) is the procedure of creating a formulated mapping from an input to an output. The system contains logical processes, a set of “if-then” rules, and formulating membership functions. FIS is composed of five main parts: fuzzification interface, rule base, database, decision-making unit, and defuzzification interface, as shown in Figure 1. At first, the fuzzification interface transfers the input data into degrees of a match with linguistic values. Then, the rule applies a number of fuzzy “if-then” rules. Databases are used in the rules for membership function, and the decision-making unit is utilized for the operations of inference. Finally, the defuzzification interface transfers fuzzy output to crisp results.

According to Jyh-Shibg [27], fuzzy if-then rules are the cases where membership functions characterize the following statement if “ $A = x$  then  $B = y$ ” where  $A$  and  $B$  are linguistic variables and  $x$  and  $y$  are linguistic values that are connected with membership functions.

A Sugino-type is also another kind of fuzzy if-then rules, where the premise part contains fuzzy sets only, whereas a nonfuzzy set defines the consequent part. It is also known as an adaptive neurofuzzy inference system (ANFIS) that is a type of fuzzy logic and neural network [27]. It has the ability to extract the advantages of both fuzzy logic and neural network in a single method [28]. It uses the algorithm of back-propagation and the least squared to learn the data to alter the membership function that assists the fuzzy to train the data to be modeled [29].

*2.3. Functional Network (FN).* A functional network (FN) is an extension of an ANN that comprises several layers of neurons linked to each other. Every processing neuron makes an explicit calculation: a scalar usually monotone  $f$  function of a weighted total of inputs. The  $f$  function, combined with the neurons, is constant, and the weights are learned from data utilizing some famous algorithms like the least-square fitting [30].

An FN comprises the input layer of input data, an output layer, single or many computing neuron layers that appraise

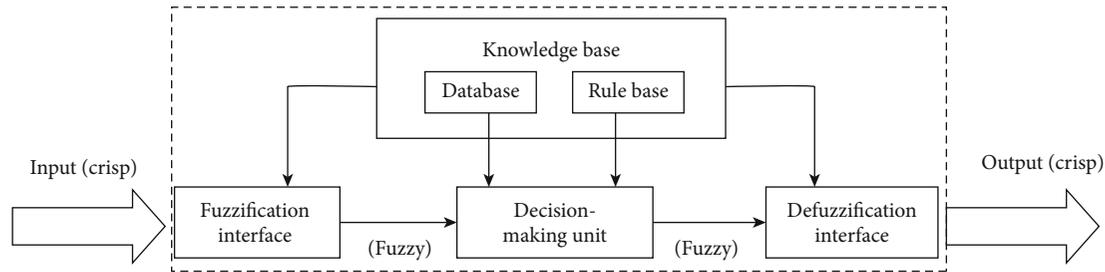


FIGURE 1: Parts of fuzzy logic.

a group of input values, comes from the input layer, and provides a group of output values to the output layer. The computing neuron layers are associated with each other, which mean that the output from one unit is able to work as a portion of the input to another neuron. When the input parameters are provided, the output is found by the type of functional network [31]. The differences between FN and ANN are that the weights in FN do not appear while the weights in ANN can appear. The neuron outputs of ANN are not the same while the neuron outputs of FN are coincident, and the neural functions in ANN are univariate, while in FN, they are multivariate.

Loss of return is influenced by various factors such as fluid properties, formation properties, and several known and unknown parameters. Therefore, it is arduous to predict. Therefore, many researchers applied the artificial intelligence to solve problems related to lost circulation such as Anifowose et al. [32], Castillo [33], Moazzeni et al. [34], Toreifi et al. [35], Efendiyev et al. [36], Far and Hosseini [37], Solomon et al. [38], Manshad et al. [39], Al-Hameedi et al. [40], Alkinani et al. [41], Abbas et al. [42], Cristofaro et al. [43], and Jahanbakhshi and Keshavarzi [44]. All these studies applied a single technique of AI to predict either the type of losses, the amount of losses, or the loss treatment, besides using many input parameters that are difficult to access in every well. None of these studies predicted the zones of the losses or used the real-time mechanical surface drilling parameters in their predictions.

The objective of this study is to predict the lost circulation zones using surface drilling parameters obtained by real-time drilling sensors. Three artificial intelligence techniques, namely, ANN, fuzzy logic (FL), and FN, are compared to achieve the objective.

### 3. Methodology

Figure 2 summarizes the processes of the methodology used in this study to predict the loss zones.

**3.1. Data Acquisition.** Three onshore wells were selected, where the lost circulation records and the mechanical surface drilling parameters were used for this study. The data were acquired on a per-foot basis from real-time sensors. The loss records include the flow out of the well (FLWOUT %) and the depth of the losses. The mechanical surface drilling parameters include the depth ( $D$ ), hook height (HKHT), hook load (HKL), flow pump (FPWPMP), rate of penetra-

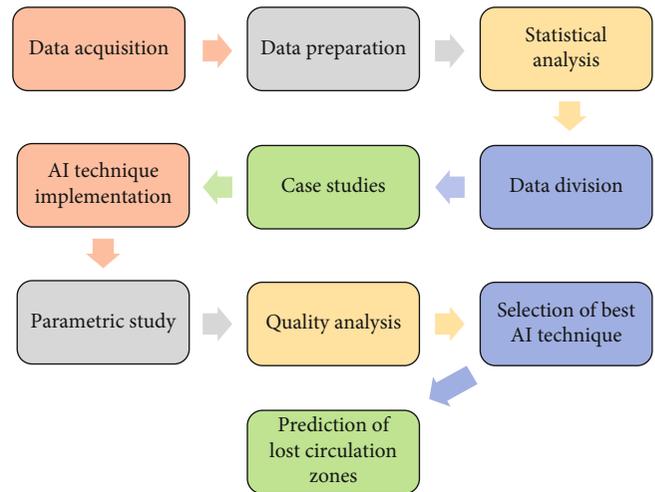


FIGURE 2: Steps of the methodology used.

tion (ROP), string rotary speed (RPM), standpipe pressure (SPP), drilling torque (TORQUE), and weight on bit (WOB). The circulation loss occurred in the three wells, and the drilling was continued until reaching the end of the section without curing the losses. Figure 3(a) shows the collected data of WOB versus the depth in Well A.

**3.2. Data Preparation.** Firstly, the data were collected from all operations involved in the three phases of the overall drilling process, i.e., drilling, tripping, and running the casing. All missing values, such as 999 values, and negative values, were removed. The second step was to include only the data in the drilling phase operation, while the data from the other phases were considered as unwanted. The data from the drilling phase operation were reorganized based on fresh footage, which requires human involvement to mark the minimum and maximum depths reached and eliminate any depth values beyond the maximum depth. Then, any footage values less than the previous were removed and will be considered a tripping operation. Figure 3(b) shows the data of WOB versus the depth from Well A after removing the random values and selecting the drilling phase operation.

The next stage was to further smoothen the data by eliminating the outliers or noise. Many filtration techniques, which have been implemented to allow data automation in the future, were applied to smoothen the data. These techniques included *movmean*, *movmedian*, *Gaussian*, *lowess*, *loess*,

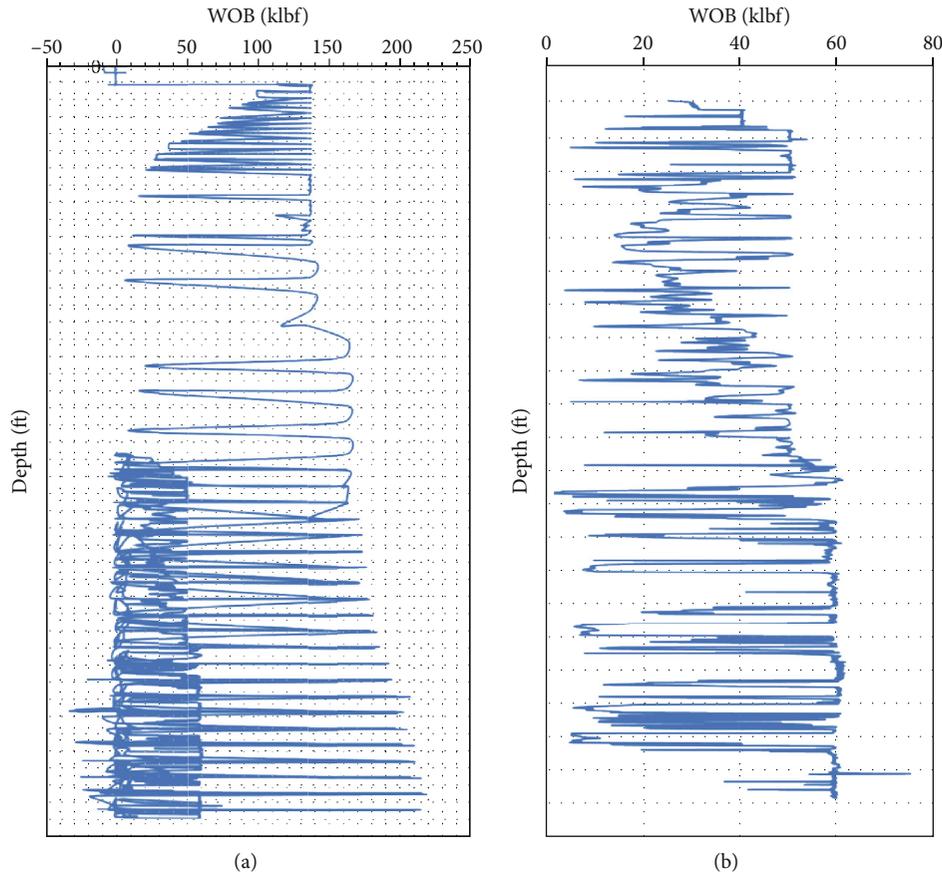


FIGURE 3: The collected WOB vs. depth from Well A (a) before filtration and (b) after filtration.

*rloess*, *rlloess*, and *sgolay*. Figure 4 shows the application of all the techniques on the WOB parameter from Well A.

The best filtration technique is *movmean*, which ensures that most of the data are preserved without significantly altering them. The performance of the *sgolay* filter was also found to be close to that of *movmean*; however, when processing big data in real-time, *movmean* is preferred as it requires less computing power [45]. The *movmean* technique was also applied to filter the WOB parameter from Well A with a span of 2, 4, 6, 8, 10, and 5 to determine the optimum noise reduction while retaining the data structure. The span of 5 is the best for data smoothing.

Regarding the output, the only action taken was to prepare the data in the proper format. As the two relevant conditions for each well section are losses or no losses, the data were arranged, as shown in Table 1, with the corresponding condition identified with 1 or 0.

**3.3. Statistical Analysis.** The best approach to examine the influence of different parameters on the loss of circulation is by performing a statistical analysis. Data diversity was assessed through a comprehensive statistical analysis. Statistical description contains a minimum, maximum, mean, range, mode, variation, kurtosis, skewness, and standard deviation. Table 2 shows the statistical analysis of Well A.

**3.4. Data Division.** Data from Well A were used to build the three AI models, while data from Well B and Well C were used to validate the AI models. The Well A data were randomly divided into two parts: the first part was used to train the model and the second part was used to test its ability to predict the values of the relevant parameters. The percentages of data used for training and testing were selected by trial and error.

Initial ANN runs were conducted using several different percentages of data for training and testing to select the best proportion on a trial and error basis. The previously identified six input parameters of FLWPMP, ROP, RPM, SPP, TORQUE, and WOB were used in these trials. Figure 5 shows the results of all the trials for the selected training and testing data distributions. The results reveal that the distribution of 75% training dataset and 25% testing dataset is the best in terms of both *R* and RMSE.

**3.5. Case Studies.** Several cases were evaluated to examine the impact of the input parameters on the prediction of lost circulation zones and enhance AI accuracy by removing the unnecessary parameters. In each run, the effect of a specific parameter on predicting the lost circulation zones was evaluated, while keeping the other parameters constant. Figure 6 presents the results of all the trials defined in Table 3 for

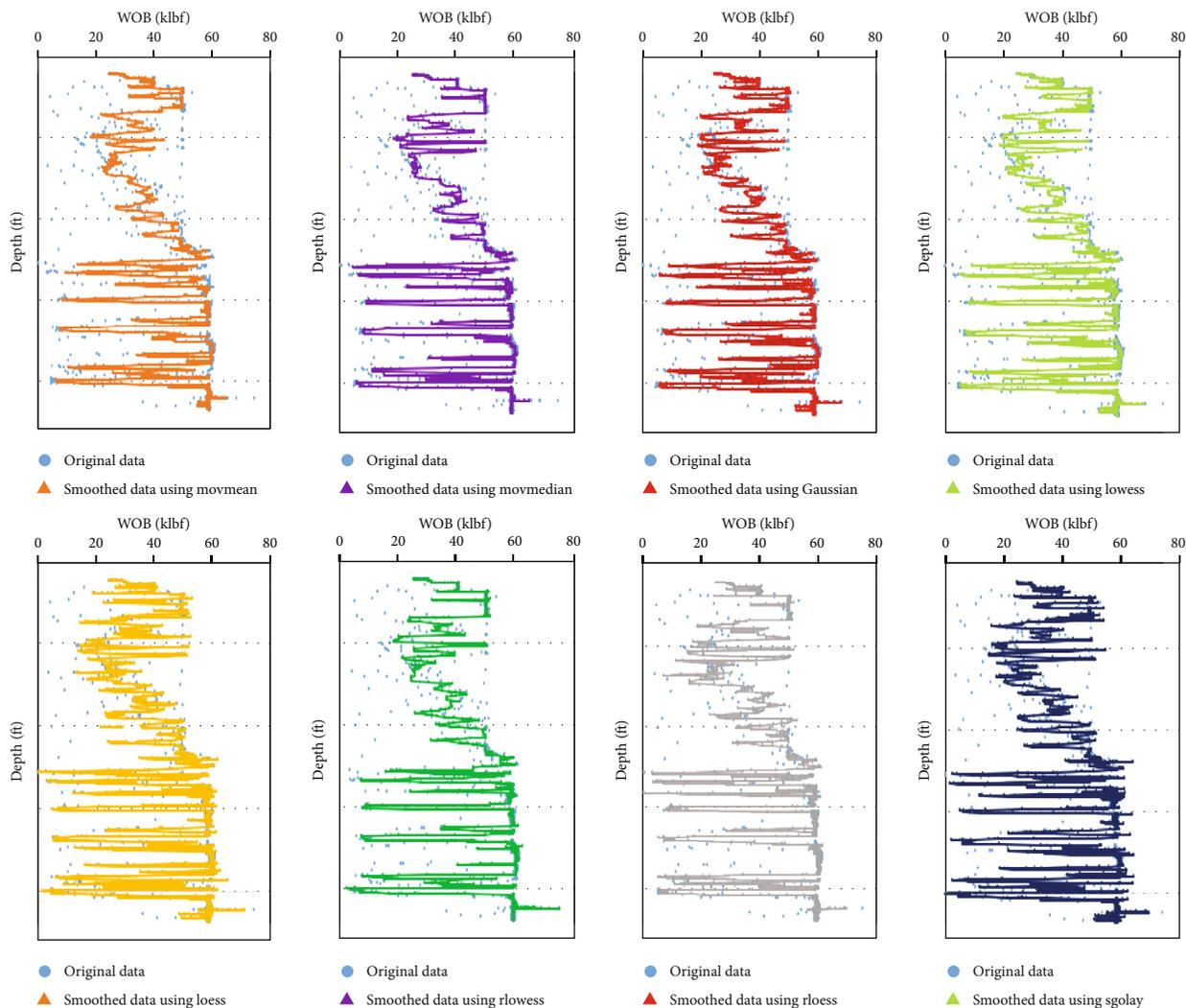


FIGURE 4: Filtration techniques applied to WOB from Well A.

TABLE 1: Input and output setup.

	Depth zone	FLWPMPS (gal/min)	Input				WOB (klbf)	Output	
			ROP (ft/h)	RPM (rpm)	SPP (psi)	TORQUE (klbf-ft)		Losses	No losses
Loss zone	1	1006	135.4	65.8	2451.5	11.9	48.9	0	1
	2	1006	167.7	66	2431.4	11.8	45.5	0	1
	3	970	162.1	64	2159.9	10.3	34.6	0	1
No loss zone	1	850	140.2	60.8	1520.5	7.5	14.7	1	0
	2	790	116.2	59.6	1225.8	6.1	8.9	1	0
	3	769	92.1	58.2	1097.1	5.7	8.2	1	0

selecting the input parameters. Based on the correlation coefficient and the root mean square error, the six input parameters (Trial 1) of FLWPMPS, ROP, RPM, SPP, TORQUE, and WOB were selected to predict lost circulation zones. Trials 1, 4, and 7 are giving the best results based on various combinations of independent parameters. Trial 1 requires six parameters that are one parameter more than each of

Trial 4 and Trial 7. Nevertheless, we believe that Trial 1 includes parameters that are physically important to detect the loss of circulation zones such as RPM and WOB. In contrast, Trials 4 and 7 are missing at least one of them.

The trials focused only on the mechanical surface drilling parameters that are very important in drilling and readily available for each well. All six parameters show significant

TABLE 2: Statistical analysis of all the parameters in Well A.

Parameter	FLWOUT (%)	HKHT (ft)	HKLI (klbf)	FLWPMPS (gal/min)	ROP (ft/h)	RPM (RPM)	SPP (psi)	TORQUE (kft-lbf)	WOB (klbf)
Minimum	0	0.218	160.946	721.000	9.260	49.720	1002.727	4.480	5.128
Maximum	54.480	88.323	249.116	1201.800	200.408	100.280	2753.440	14.344	68.372
Mean	9.664	44.615	191.665	1072.861	54.497	73.125	2193.087	8.686	49.335
Range	54.480	88.105	88.170	480.800	191.148	50.560	1750.713	9.864	63.244
Mode	0.000	63.043	200.476	1006.000	18.060	80.000	2459.280	8.696	58.956
Skewness	1.028	-0.051	0.749	-0.350	1.571	-0.034	-0.794	0.487	-1.305
Standard deviation	15.254	24.063	15.406	126.587	53.358	12.418	368.205	1.703	13.656

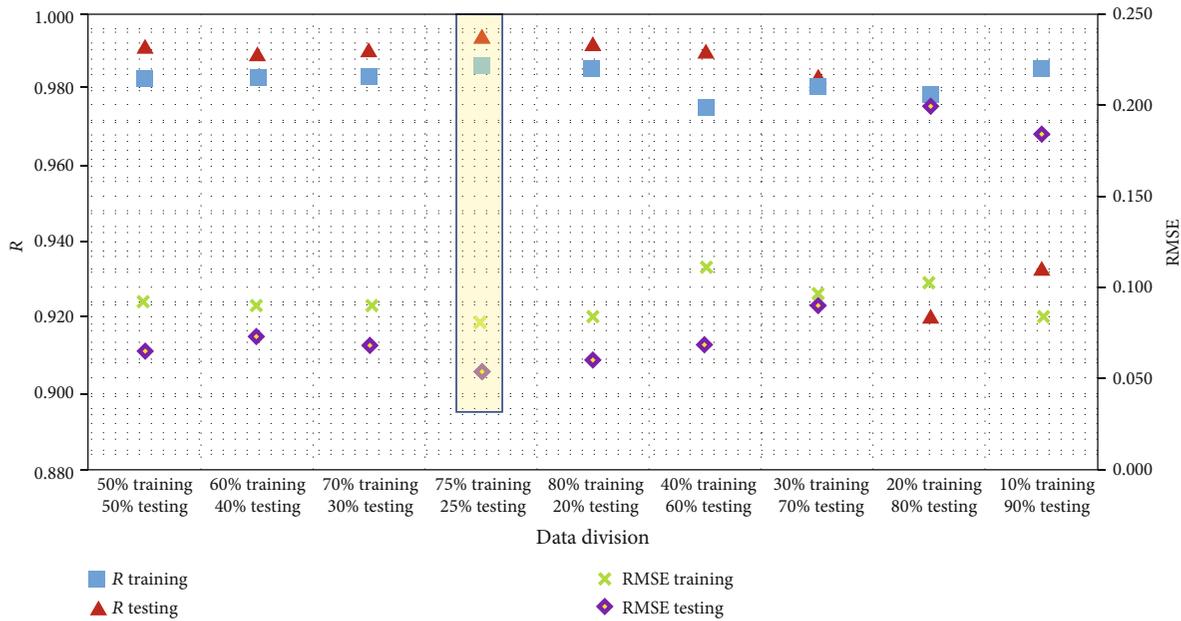


FIGURE 5: The results of the trials conducted for selecting the training and testing data distributions.

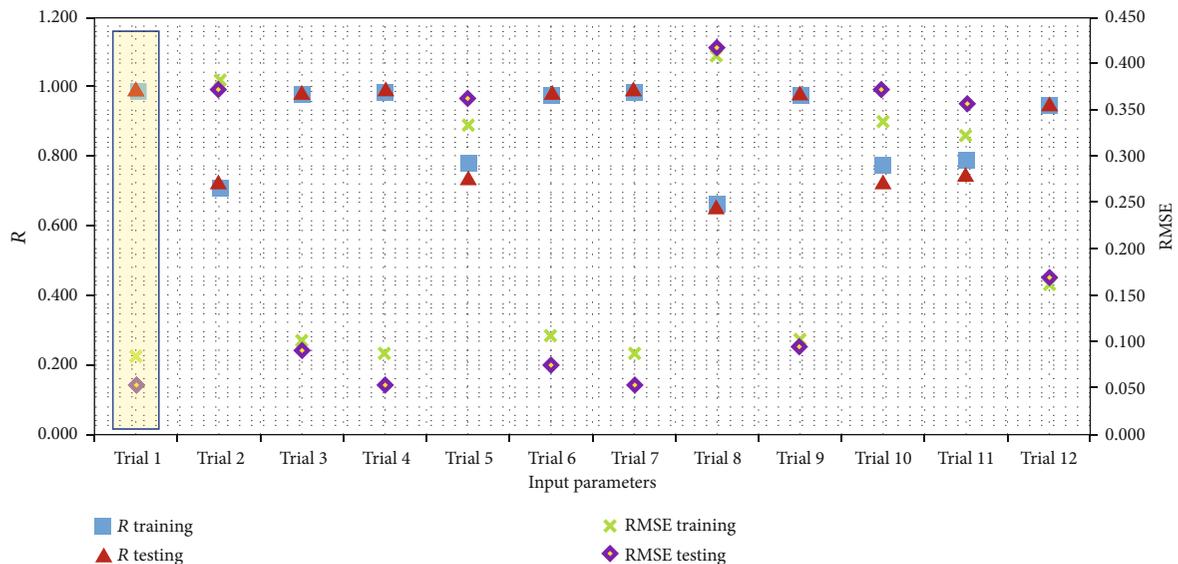


FIGURE 6: The results of the evaluation to select the important input parameters.

TABLE 3: Parameters used for the trials for selecting input parameters.

Trial #	Input parameters
Trial 1	FLWPMPS, ROP, RPM, SPP, TORQUE, WOB
Trial 2	ROP, RPM, SPP, TORQUE, WOB
Trial 3	FLWPMPS, RPM, SPP, TORQUE, WOB
Trial 4	FLWPMPS, ROP, SPP, TORQUE, WOB
Trial 5	FLWPMPS, ROP, RPM, TORQUE, WOB
Trial 6	FLWPMPS, ROP, RPM, SPP, WOB
Trial 7	FLWPMPS, ROP, RPM, SPP, TORQUE
Trial 8	RPM, SPP, TORQUE, WOB
Trial 9	FLWPMPS, SPP, TORQUE, WOB
Trial 10	FLWPMPS, ROP, TORQUE, WOB
Trial 11	FLWPMPS, ROP, RPM, WOB
Trial 12	FLWPMPS, ROP, RPM, SPP

changes in their values after the occurrence of lost circulation, as shown in Figures 7(a)–7(c). Hook loud was excluded from the trials as it did not change after the event of lost circulation and due to the availability of WOB, which is an alternative for hook loud.

**3.6. Implementation of AI Techniques.** After selecting the best input parameters for constructing the AI models, the next step is to apply artificial intelligence techniques. As previously stated, the artificial intelligence tools ANN, FL, and FN were used. MATLAB 2016 software was used to implement the AI methods.

**3.6.1. ANN Implementation.** Many trials were conducted to select the optimum number of neurons, training functions, transfer functions, and the network function using ANN to predict the zones of lost circulation. One layer with different numbers of neurons was used, and the results are shown in Figure 8. The results are very close to each other and are of high accuracy. Increasing the number of neurons will increase the computational time and will result in a large number of weights and biases, which in turn will increase the number of constants in the correlation equation. Therefore, five neurons are selected based on their higher accuracy in the testing part and to keep the network fast and efficient.

The performance of 13 training functions was evaluated to determine the optimum training function. The results presented in Figure 9 indicate that the optimum training function that has the highest accuracy is *Trainbr*. Thirteen transfer functions were also evaluated to observe their impacts on the prediction of lost circulation zones. Figure 10 summarizes the performance of these transfer functions and reveals that *logsig* is the optimum transfer function.

Then, ten network functions were assessed to determine their influence in the prediction of lost circulation zones. The results presented in Figure 11 show that the values are close to each other. Even though *Fitnet*, *Newff*, and *Newfit* produce the same accuracy in the prediction of the lost circu-

lation zones, the *Fitnet* network function is the one selected in the ANN model.

The ANN model was built with five neurons in one internal layer, and with a training function of *Trainbr*, the transfer function of *Logsig*, and network function of *Fitnet*.

**3.6.2. FL Implementation.** Two tools of fuzzy logic were performed to determine their impacts on the prediction of the lost circulation zones. The first tool is Mamdani FIS (*Genfis 1*), and the second tool is Sugeno FIS (*Genfis 2*). Mamdani FIS is not suitable because it has a long processing time. Many trials were performed using *Genfis 2* (Sugeno FIS) to determine the best selection of the epoch size and the radius. The results in Figure 12 indicate that the prediction accuracy does not change after an epoch size of 70, and based on the highest correlation coefficient and the lowest root mean square error, an epoch size of 70 was selected.

Several radii were evaluated at 70 iterations, and the results are shown in Figure 13. A radius of 0.5 produced the highest accuracy. Therefore, the fuzzy logic model was built using Sugeno FIS with an epoch size of 70 and a radius of 0.5.

**3.6.3. FN Implementation.** Several FN trials were run to select the best methods and types of functional networks to predict the zones of lost circulation. Many procedures were evaluated for both linear and nonlinear types to find their effect on the prediction of the lost circulation zones. These procedures include functional network backward-forward method (FNBFM), functional network forward-backward method (FNFBM), functional network backward-exhaustive method (FNBEM), functional network forward-selection method (FNFSM), and functional network exhaustive selection method (FNESM). The results are shown in Figure 14. Based on the highest  $R$  and the lowest RMSE, the Type 3 of FNFBM has the highest accuracy. Therefore, the FN model, with the features of functional network forward-backward method (FNFBM) and nonlinear Type 3 was selected to predict the lost circulation zones.

## 4. Results and Discussion

The ANN implementation showed that an ANN, of five neurons in one internal layer, with the training function of *Trainbr*, the transfer function of *Logsig*, and the network function of *Fitnet*, gives the best performance with the highest correlation coefficient and the lowest root mean square error. The ANN predicted the lost circulation zones in the training part (75% of the data) with  $R$  of 0.987 and RMSE of 0.081, as shown in Figure 15(a). In the testing part (25% of the data), the ANN predicted the lost circulation zones with  $R$  of 0.994 and RMSE of 0.053, as indicated in Figure 16(a).

The results of the prediction are also presented in a confusion matrix. The results for the training set, depicted in Figure 17(a), show that the ANN model was able to predict 735 out of 742 locations of the lost circulation zones correctly, i.e., 99.1%, and the ANN model was not able to accurately predict only 7 locations, representing 0.9%. In the zones where losses do not occur, the ANN model was able

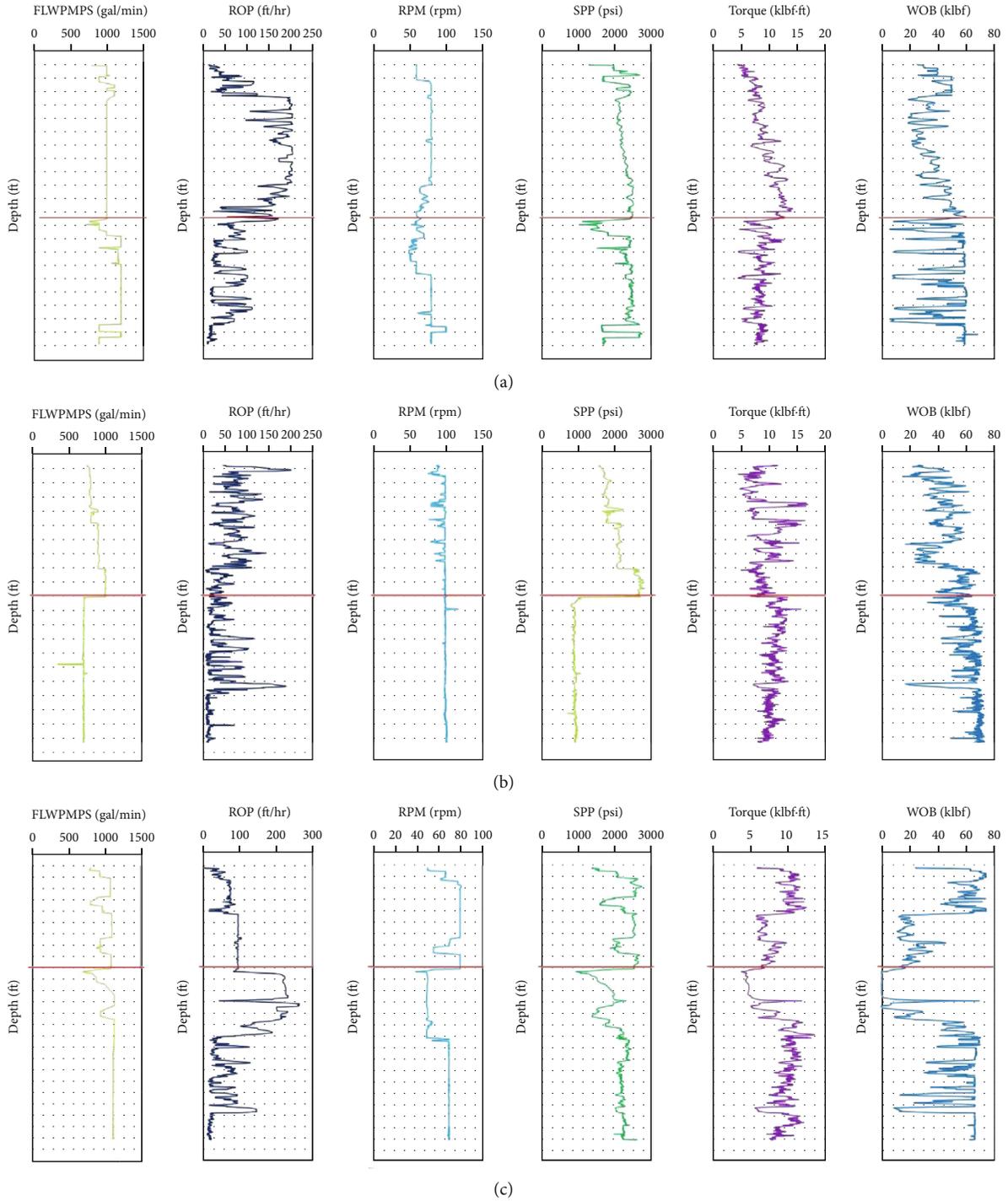


FIGURE 7: (a) The change of the six surface parameters after the occurrence of lost circulation in Well A. (b) The change of the six surface parameters after the occurrence of lost circulation in Well B. (c) The change of the six surface parameters after the occurrence of lost circulation in Well C.

to predict all 321 locations correctly. Considering all the locations of lost circulation in the two zones, the ANN model correctly predicted 99.3% of the locations, and the prediction of only 0.7% of the locations are incorrect. The results for the testing dataset shown in Figure 18(a) reveal that the ANN

model was able to predict 263 out of 264 locations in the lost circulation zones correctly, i.e., 99.6% of the locations, and it could not precisely predict only 1 point, representing 0.4% of the locations. In the zones where losses do not occur, the ANN model was able to predict all 90 locations correctly.

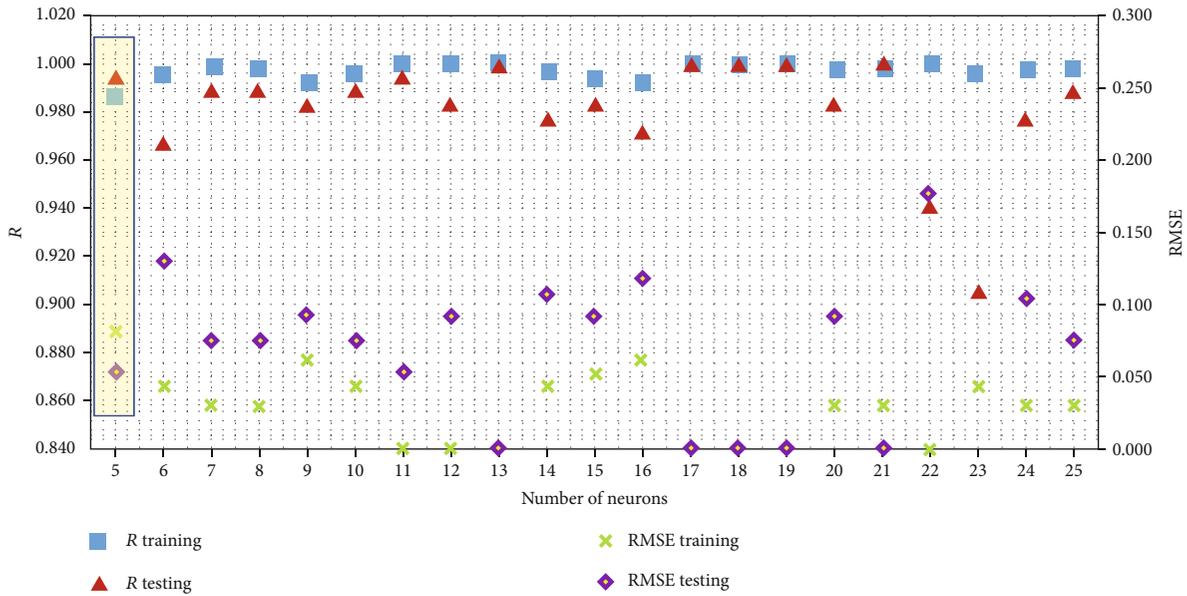


FIGURE 8: Impact of the number of neurons in one layer of the training and testing datasets.

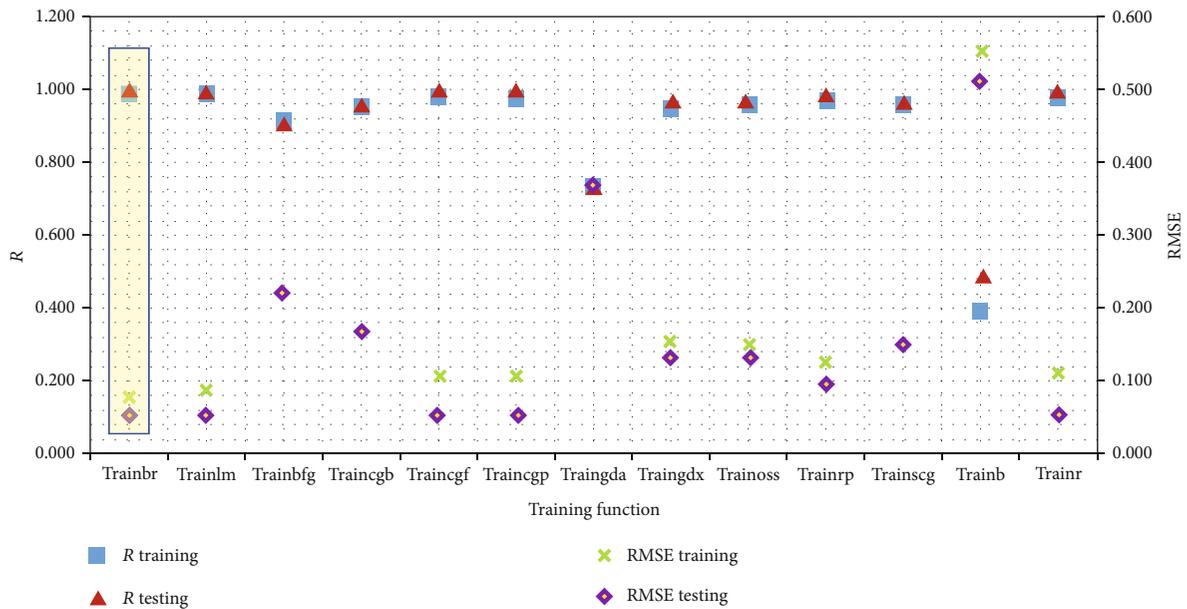


FIGURE 9: Results of using different training functions for the training and testing datasets.

Considering all the locations in the two zones, the ANN model correctly predicted 99.7% of the locations, and the prediction of only 0.3% of the locations is incorrect.

The FL implementation showed that the best performance of the FL that produces the highest correlation coefficient and the lowest root mean square error is with the features of Sugeno FIS, an epoch size of 70, and a radius of 0.5. The FL model predicted the lost circulation zones in the training part (75% of the data) with  $R$  of 0.993 and RMSE of 0.053, as shown in Figure 15(b). In the testing part (25% of the data), the FL predicted the lost circulation zones with  $R$  of 0.993 and RMSE of 0.053, as indicated in Figure 16(b).

A confusion matrix also presents the results of the prediction. The results for the training dataset shown in Figure 17(b) indicate that the FL model was able to predict 739 out of 742 locations in the lost circulation zones correctly, with an accuracy of 99.6%. In comparison, it mispredicted only 3 locations, representing 0.4% of the data. In the zones where losses do not occur, the FL model was able to predict all 321 locations correctly. Considering all the locations in the two zones, FL correctly predicted 99.7% of the locations, with only 0.3% mispredicted. The results for the testing dataset depicted in Figure 18(b) show that the FL model was able to predict 263 out of 264 locations in the lost circulation zones correctly, with an accuracy of 99.6%. In

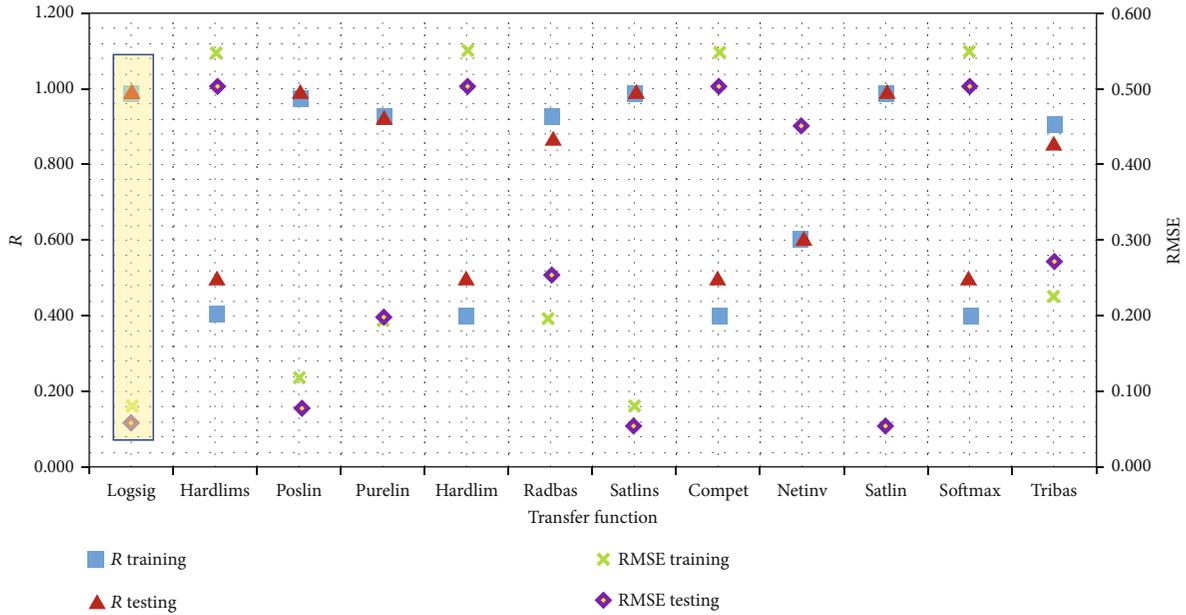


FIGURE 10: Influence of using different transfer functions for the training and testing datasets.

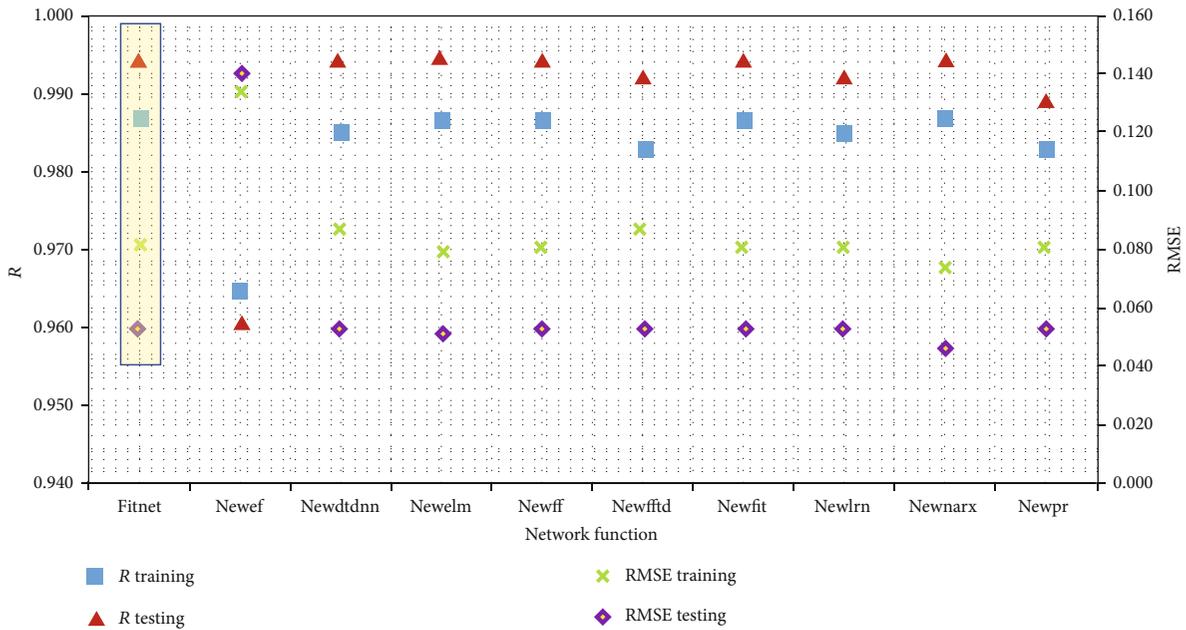


FIGURE 11: Effects of using different network functions for the training and testing datasets.

contrast, it mispredicted only one location, representing 0.3% of the data. In the zones where losses do not occur, the FL model was able to predict all 90 locations correctly. Considering all the locations in the two zones, the FL correctly predicted 99.7% of the locations, with only 0.3% mispredicted.

The FN implementation showed that the best performance of the FN that produces the highest correlation coefficient and the lowest root mean square error is with the features of FNFBM and nonlinear Type 3. The FN model predicted the lost circulation zones in the training part (75% of the data) with  $R$  of 1 and RMSE of 0, as shown in

Figure 15(c). In the testing part (25% of the data), the FN model predicted the lost circulation zones with  $R$  of 0.985 and RMSE of 0.075, as indicated in Figure 16(c).

A confusion matrix also presents the results of the prediction. The results for the training dataset depicted in Figure 17(c) indicate that the FN model was able to predict all 742 locations in the lost circulation zones correctly. In the zones where losses do not occur, the FN model was also able to predict all 321 locations correctly. Considering all the locations in the two zones, FN correctly predicted 100% of the data locations. The results for the testing dataset shown

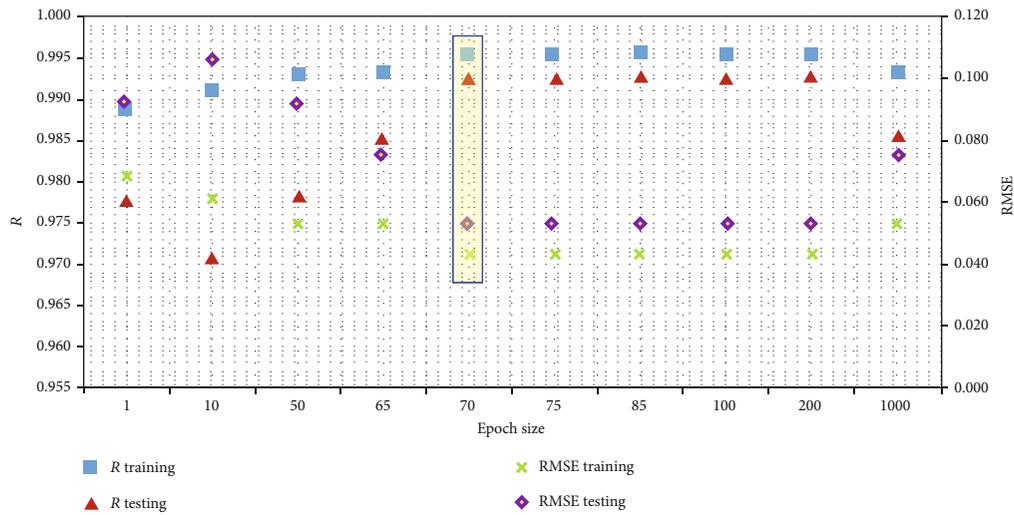


FIGURE 12: Influence of the iteration number in Sugeno FIS for the training and testing datasets.

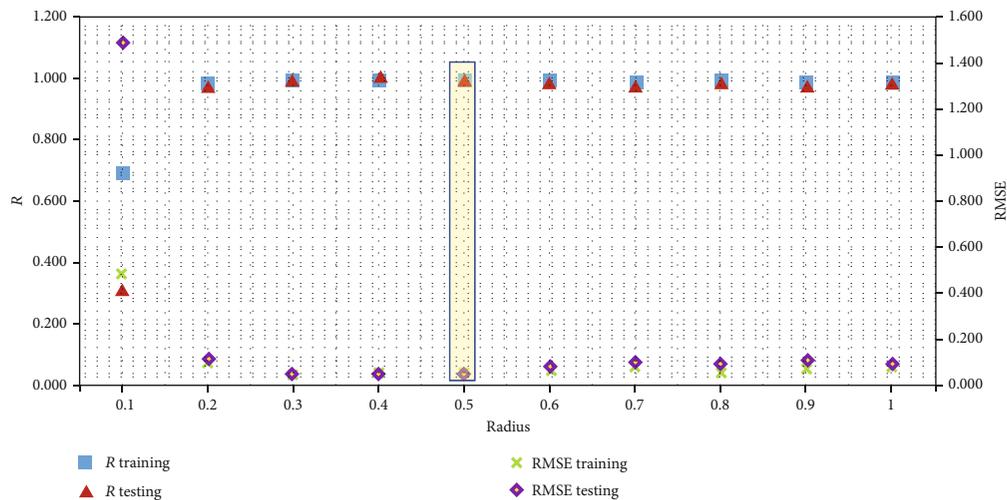


FIGURE 13: Impact of the radius in Sugeno FIS for the training and testing datasets.

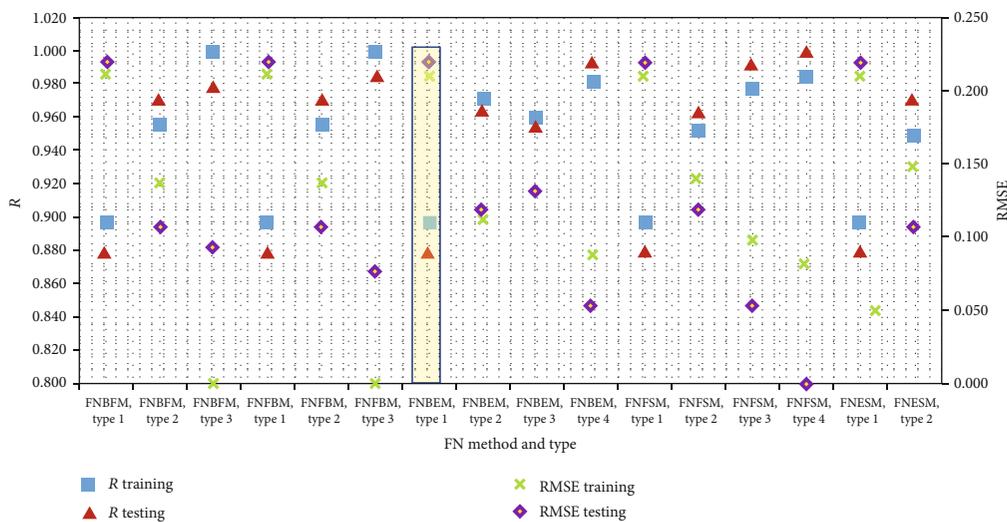


FIGURE 14: Influence of the FN methods and types for the training and testing datasets.

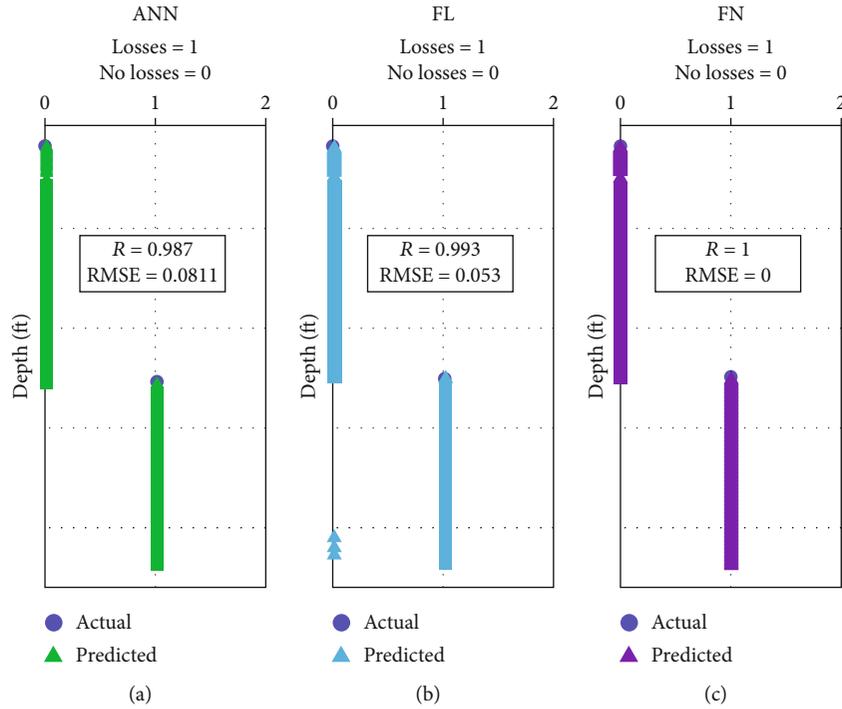


FIGURE 15: Prediction of the lost circulation zones using the five AI models for the training dataset.

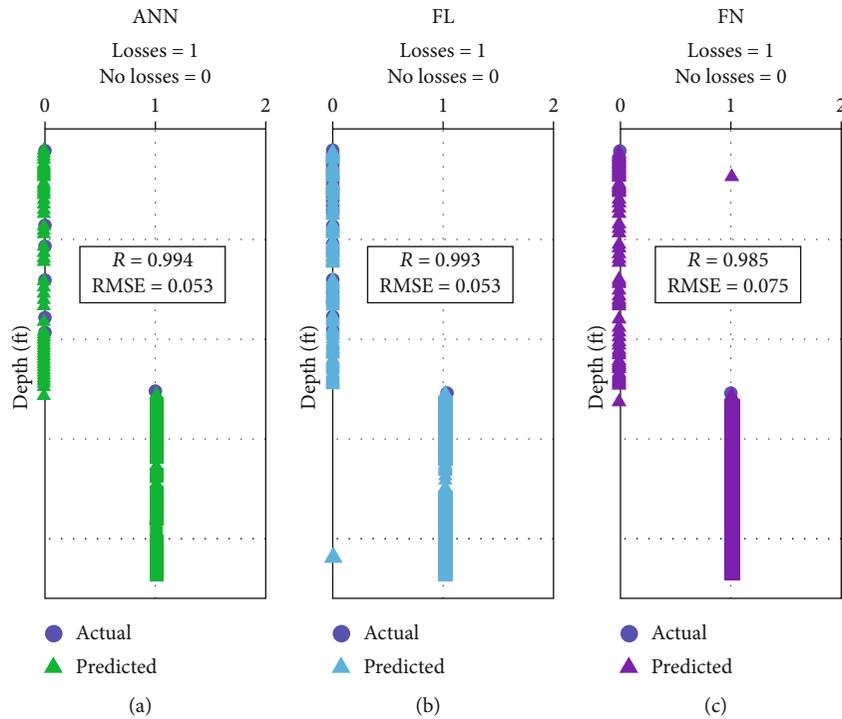


FIGURE 16: Prediction of the lost circulation zones using the five AI models for the testing dataset.

in Figure 18(c) reveal that the FN model was able to predict 263 out of 264 locations in the lost circulation zones correctly, with an accuracy of 99.6%. In the zones where losses

do not occur, the FN model was able to predict 89 locations out of 90 locations correctly, representing 98.8% of the data, and it mispredicted only one location, representing 1.1% of

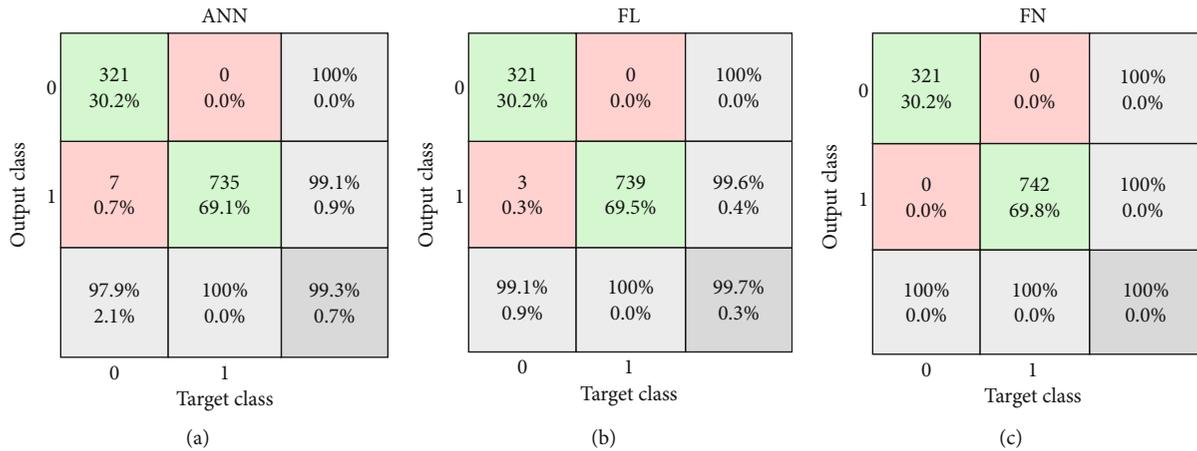


FIGURE 17: Confusion matrix for the prediction of lost circulation zones using the three AI models for the training dataset.

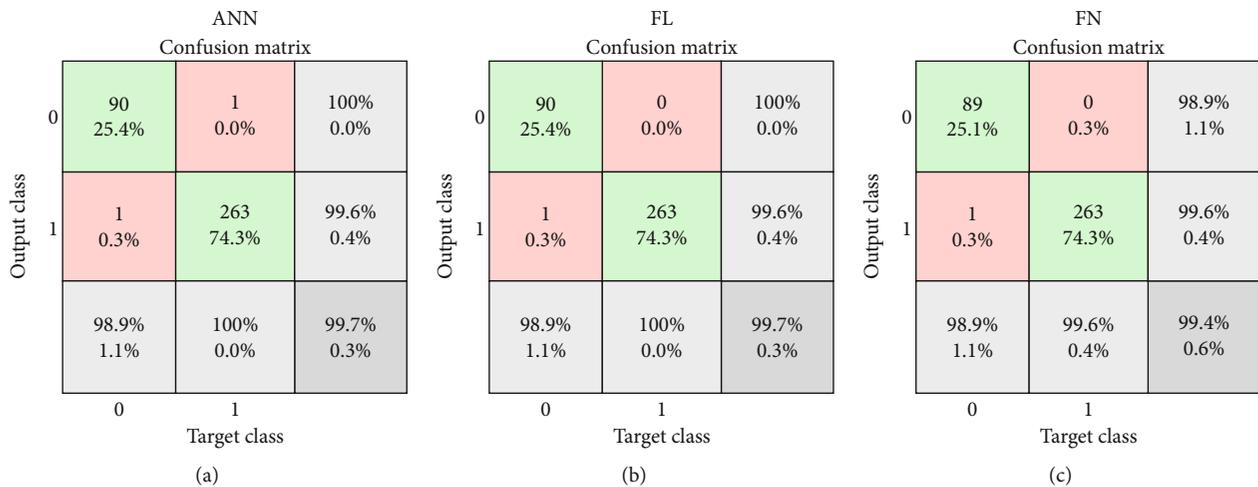


FIGURE 18: Confusion matrix for the prediction of lost circulation zones using the three AI models for the testing dataset.

the data. Considering all locations in the two zones, the FN correctly predicted 99.4% of the data locations, with only 0.6% mispredicted.

### 5. Comparison between AI Techniques

A comparison among the three AI techniques was conducted to select the best AI technique, i.e., the technique that correctly predicts the lost circulation zones with high accuracy, and the results are shown in Figure 19.

With 75% of the data used for training, the actual lost circulation zones were compared with the predicted lost circulation zones. ANN was able to predict the lost circulation zones correctly with  $R$  of 0.987 and RMSE of 0.0811. Under the same conditions, the FL model predicted them with  $R$  of 0.981 and RMSE of 0.097, while the FN model predicted lost circulation zones with  $R$  of 1 and RMSE of 0.

With 25% of the data used for testing, actual lost circulation zones were compared with the predicted lost circulation zones. The ANN model was able to correctly predict the lost circulation zones with  $R$  of 0.994 and RMSE value of 0.053. The FL model predicted the losses zones with  $R$  and RMSE

of 0.993 and 0.053, respectively. The FN model predicted the lost circulation zones with  $R$  of 0.985 and RMSE of 0.075.

The results for the training dataset reveal that FN produces the highest accuracy, while ANN produces the lowest  $R$  and the highest RMSE. For the testing dataset, the results indicate that ANN has the highest  $R$  and the lowest RMSE, while FN produces the lowest  $R$  and the highest RMSE. Therefore, ANN was selected as the best AI method to predict the zones of lost circulation because of its highest accuracy in the testing part. The reason behind this is that ANN has many parameters to optimize such layers, number of neurons, network functions, training functions, and transfer functions. In contrast, the other AI methods, such as FN and FL, have a smaller number of parameters to optimize.

### 6. Validation of the AI Techniques

The AI technique with the best accuracy in predicting the lost circulation zones in Well A was validated by using it to predict the lost circulation zones in Well B and Well C. ANN

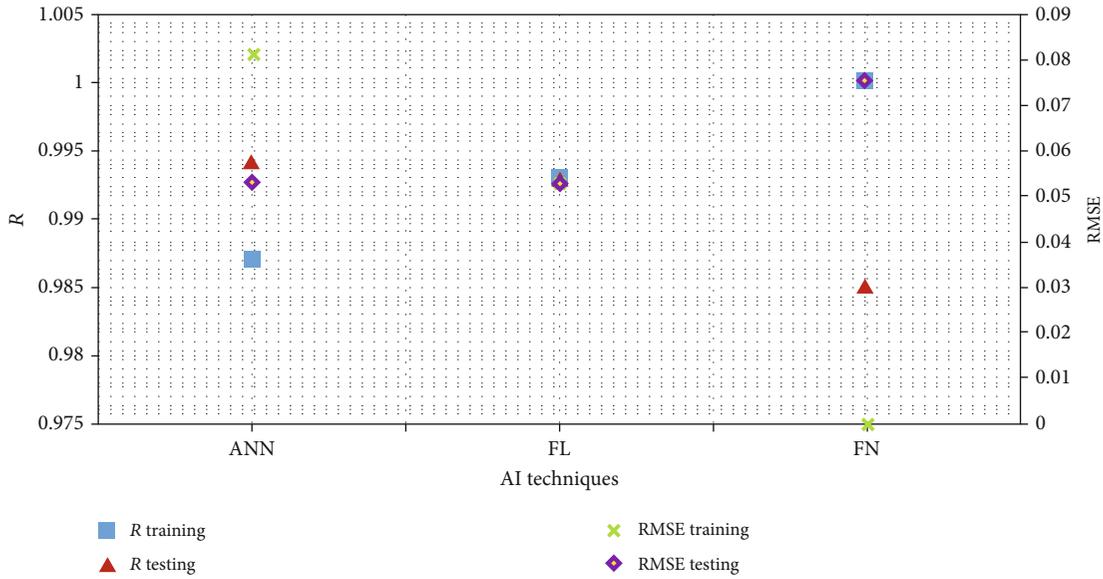


FIGURE 19: Comparison of the five AI techniques for the training and testing datasets.

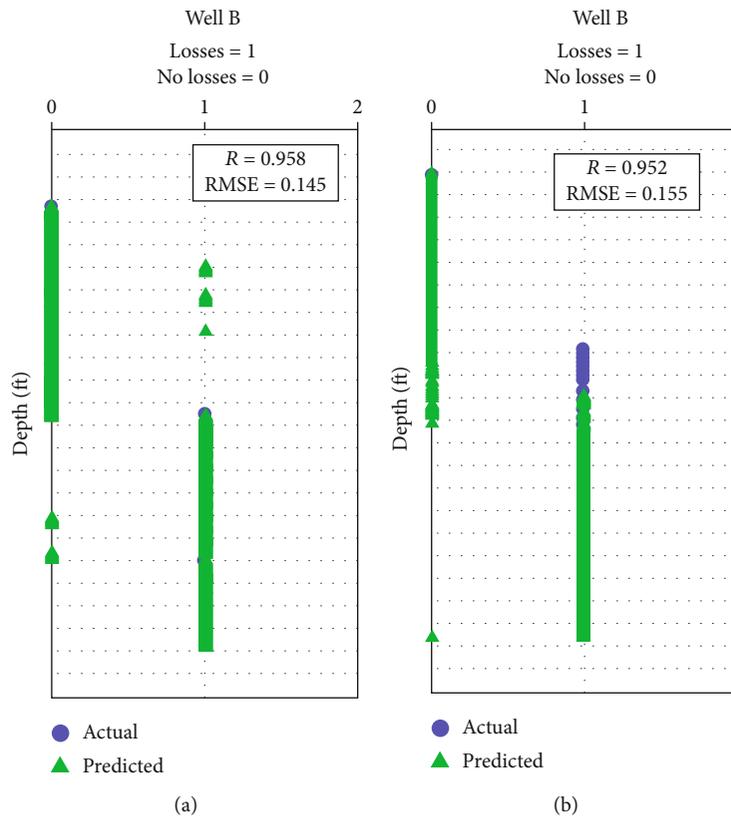


FIGURE 20: Prediction of the lost circulation zones using the ANN model in Well B (a) and Well C (b).

was selected for the validation using the data from Well B and Well C because of its highest accuracy in the testing part.

In Well B, ANN was able to predict the lost circulation zones with high accuracy ( $R = 0.958$  and  $RMSE = 0.145$ ), and the results of the evaluation are shown in Figure 20(a). The confusion matrix shown in Figure 21(a) indicates that the ANN model was able to predict 2067 out of 2103 loca-

tions in the lost circulation zones correctly, with an accuracy of 98.3%, with only 36 locations, representing 1.7%, not predicted correctly. In the zones where losses do not occur, the ANN model was able to predict 745 out of 769 locations correctly, with an accuracy of 96.9%, while only one location, representing 3.1% of the data, was mispredicted. Considering all the locations in the two zones, the ANN model correctly

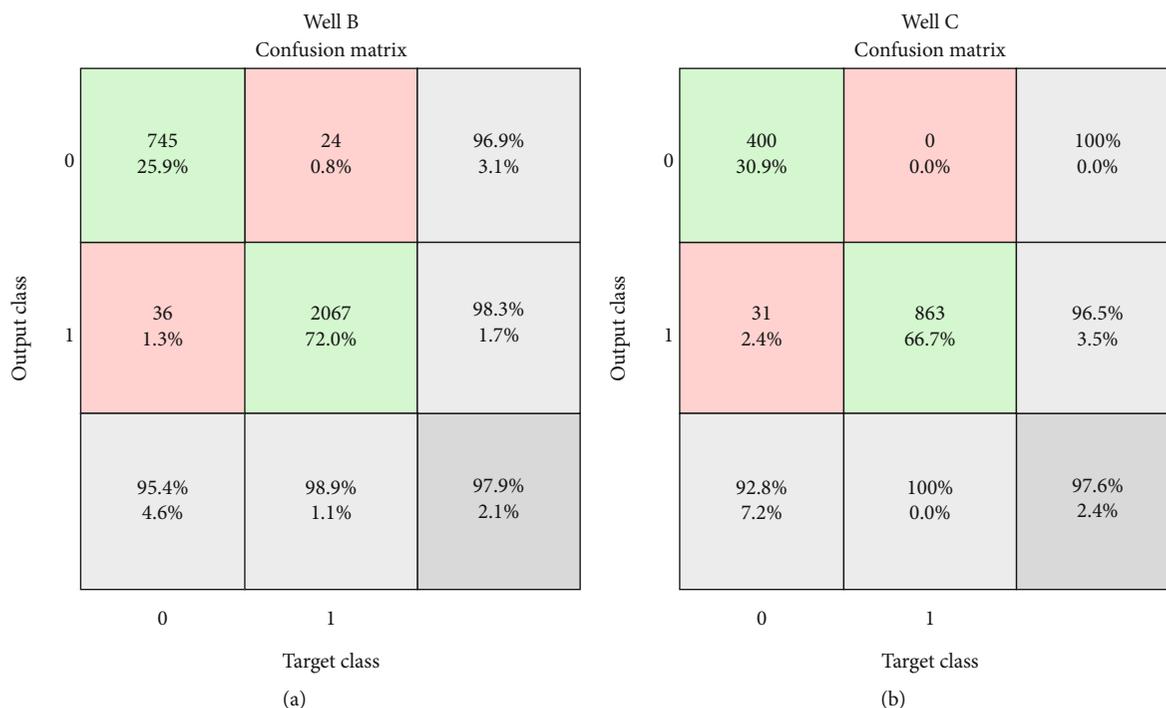


FIGURE 21: Confusion matrix for the prediction of lost circulation zones in Well B (a) and Well C (b) using the ANN technique.

predicted 97.9% of the locations, while only 2.1% was predicted correctly.

In Well C, the ANN model was able to predict the lost circulation zones with high accuracy ( $R = 0.952$  and  $RMSE = 0.155$ ), and the results of the evaluation are shown in Figure 20(b). The confusion matrix shown in Figure 21(b) indicates that the ANN model was able to predict 863 out of 894 locations in the lost circulation zones correctly, with an accuracy of 96.5%, with only 31 locations, representing 3.5% of the data, not predicted correctly. In the zones where losses do not occur, the ANN model was able to predict all 400 locations correctly. However, considering the two zones, the ANN technique correctly predicted 97.6% of the data, while only 2.4% were mispredicted.

## 7. Conclusions

This study evaluated three AI techniques to predict the lost circulation zones based only on six mechanical surface drilling parameters. These techniques are functional networks (FN), artificial neural networks (ANN) and fuzzy logic (FL). The six parameters are real-time measurements of flow pump (FPWPMP), rate of penetration (ROP), string rotary speed (RPM), standpipe pressure (SPP), drilling torque (TORQUE), and weight on bit (WOB). More than 4500 real-field data points from three wells were used in the evaluation. The following conclusions are drawn from the results of this study:

- (i) The AI techniques were trained and tested using the data from Well A to predict the lost circulation zones with high precision
- (ii) The accuracy of the ANN was  $R = 0.987$  and  $RMSE = 0.0811$  for training and  $R = 0.994$  and  $RMSE = 0.053$  for testing
- (iii) The accuracy of the FL was  $R = 0.993$  and  $RMSE = 0.0531$  for training and  $R = 0.993$  and  $RMSE = 0.053$  for testing
- (iv) The accuracy of the FN was  $R = 1$  and  $RMSE = 0$  for training and  $R = 0.985$  and  $RMSE = 0.0752$  for testing
- (v) The ANN was the best technique due to its highest accuracy in the testing. So, its model was validated using data from the second and third wells (Well B and Well C, respectively), which are unseen
- (vi) The ANN was able to identify the lost circulation zones in Well B and Well C with a high performance of  $R = 0.958$  and  $RMSE = 0.145$  and  $R = 0.952$  and  $RMSE = 0.155$ , respectively
- (vii) The main advantage of AI techniques is their simplicity, which allows the prediction of lost circulation zones from only the mechanical surface drilling parameters that are readily available in each well

The future direction of this work is to validate the developed ANN model on several wells and improve the ANN model to be able to predict the amount of the losses using the real-time surface drilling parameters, where the mud properties, formation properties, and fracture's length and width are also major parameters that need to be considered to predict the loss amount.

## Nomenclature

AI:	Artificial intelligence
FN:	Functional network
ANN:	Artificial neural networks
FL:	Fuzzy logic
Trainbr:	Bayesian regularization
Trainlm:	Levenberg-Marquardt backpropagation
Trainbfg:	BFGS quasi-Newton backpropagation
Traincgb:	Powell-Beale conjugate gradient backpropagation
Traincgb:	Powell-Beale conjugate gradient backpropagation
Traincgf:	Fletcher-Powell conjugate gradient backpropagation
Traincgp:	Polak-Ribière conjugate gradient backpropagation
Trainгда:	Gradient descent with adaptive learning rule backpropagation
Trainгдаx:	Gradient descent with momentum and adaptive learning rule backpropagation
Trainoss:	One-step secant backpropagation
Trainr:	Random-order incremental training with learning functions
Trainrp:	Resilient backpropagation (Rprop)
Trainscg:	Scaled conjugate gradient backpropagation
Compet:	Competitive transfer function
Hardlim:	Hard-limit transfer function
Hardlims:	Symmetric hard-limit transfer function
Logsig:	Log-sigmoid transfer function
Netinv:	Inverse transfer function
Poslin:	Positive linear transfer function
Purelin:	Linear transfer function
Radbas:	Radial basis transfer function
Satlin:	Saturating linear transfer function
Satlins:	Symmetric saturating linear transfer function
Softmax:	Softmax transfer function
Tansig:	Hyperbolic tangent sigmoid transfer function
Tribas:	Triangular basis transfer function
Newcwf:	Create cascade-forward backpropagation network
Newtdnn:	Create distributed time delay neural network
Newelm:	Create Elman backpropagation network
Newff:	Create feedforward backpropagation network
Newfftd:	Create feedforward input-delay backpropagation network
Newfit:	Create fitting network
Newlrn:	Create layered-recurrent network
Newnarx:	Feedforward backpropagation network with feedback from output to input
Newpr:	Create pattern recognition network
Fitnet:	Function fitting neural network
Genfis 1:	Mamdani FIS
Genfis 2:	Sugeno FIS
FNFBM:	Functional network forward-backward method
FNBEM:	Functional network backward-exhaustive method
FNFSM:	Functional network forward-selection method
FNESM:	Functional network exhaustive selection method
FLWPMPS:	Flow rate of the pump (gal/min)

ROP:	Rate of penetration (ft/hr)
SPP:	Stand pipe pressure (psi)
WOB:	Weight on bit (klbs)
RPM:	Rotary speed (rpm)
TORQUE:	Drilling torque (klbf)
R:	Correlation coefficient
RMSE:	Root mean square error

## Data Availability

All the data required are included in the text.

## Conflicts of Interest

The author declares no conflict of interest.

## References

- [1] Halliburton, "High pressure high temperature (HP/HT) (n.d.)," 2011, <https://www.halliburton.com/en-US/ps/solutions/high-pressure-temperature/about-hp-ht-ht.html>.
- [2] M. Overholt, "Advantages of offshore oil rigs and drilling," 2017, <https://www.tigergeneral.com/advantages-offshore-oil-rigs-drilling/>.
- [3] A. Shadravan and M. Amani, "HPHT 101: what every engineer or geoscientist should know about high pressure high temperature wells," 2012, <https://www.onepetro.org/conference-paper/SPE-163376-MS>.
- [4] C. R. Miranda, J. L. Oliveira, G. M. S. Cavalcante et al., "Materials for controlling severe lost circulation-laboratory evaluation," in *SPE-185582-MS was presented at SPE Latin America and Caribbean Petroleum Engineering Conference*, Buenos Aires, Argentina, 2017.
- [5] J. Ramasamy and M. Amanullah, "Novel fibrous lost circulation materials derived from deceased date tree waste," in *SPE-187989-MS was presented at SPE Kingdom of Saudi Arabia Annual Technical Symposium and Exhibition*, Dammam, Saudi Arabia, 2017.
- [6] S. P. Almagro, C. Frates, J. Garand, and A. Meyer, "Sealing fractures: advances in lost circulation control treatments," *Oil-field Review*, vol. 26, no. 3, pp. 4–13, 2014.
- [7] A. T. T. Al-Hameedi, H. H. Alkinani, S. Dunn-Norman et al., "Using machine learning to predict lost circulation in the Rumaila field, Iraq," in *SPE-191933-MS was presented at SPE Asia Pacific Oil and Gas Conference and Exhibition*, Brisbane, Australia, 2018.
- [8] J. Abdollahi, I. M. Carlsen, S. Mjaaland, P. Skalle, A. Rafiei, and S. Zarei, "Underbalanced drilling as a tool for optimized drilling and completion contingency in fractured carbonate reservoirs," in *SPE-91579-MS was presented at SPE/IADC Underbalanced Arabian Journal for Science and Engineering 1 3 Technology Conference and Exhibition*, Houston, Texas, 2004.
- [9] G. V. Chilingarian and P. Vorabutr, *Drilling and drilling fluids*, Elsevier, Amsterdam, Netherlands, 1981.
- [10] D. E. Caughron, D. K. Renfrow, J. R. Bruton et al., "Unique crosslinking pill in tandem with fracture prediction model cures circulation losses in deepwater Gulf of Mexico," in *SPE-74518-MS was presented at IADC/SPE Drilling Conference*, Dallas, Texas, 2002.

- [11] A. Lavrov, *Lost Circulation Mechanisms and Solutions*, Gulf Professional Publishing, Amsterdam, Netherlands, 2016.
- [12] R. Jadhav and S. Patil, "Acid-soluble thixotropic cement system for lost circulation challenges," in *SPE-193168-MS was presented at Abu Dhabi International Petroleum Exhibition & Conference*, Abu Dhabi, UAE, 2018.
- [13] T. R. Bratton, I. M. Rezmer-Cooper, J. Desroches, Y.-E. Gille, Q. Li, and M. McFayden, "How to diagnose drilling induced fractures in wells drilled with oil-based muds with real-time resistivity and pressure measurements," in *SPE-67742-MS was presented at SPE/IADC Drilling Conference*, Amsterdam, Netherlands, 2001.
- [14] J. U. Messenger, *Lost Circulation*, PennWell Publishing Company, 1981.
- [15] S. Mohaghegh, R. Arefi, S. Ameri, and D. Rose, "Design and development of an artificial neural network for estimation of formation permeability," *SPE Computer Applications*, vol. 7, 1995.
- [16] H. I. Bilgesu, U. Altmis, S. Ameri, S. Mohaghegh, and K. Aminian, "A new approach to predict bit life based on tooth or bearing failures," in *SPE Eastern Regional Meeting*, Pittsburgh, Pennsylvania, 1998.
- [17] K. Al-Azani, S. Elkatatny, A. Ali, E. Ramadan, and A. Abdurraheem, "Cutting concentration prediction in horizontal and deviated wells using artificial intelligence techniques," *Journal of Petroleum Exploration and Production Technology*, vol. 9, no. 4, pp. 2769–2779, 2019.
- [18] R. P. Lippmann, "An introduction to computing with neural nets," *IEEE ASSP Magazine*, vol. 4, no. 2, pp. 4–22, 1987.
- [19] A. Ahmed, A. Ali, S. Elkatatny, and A. Abdurraheem, "New artificial neural networks model for predicting rate of penetration in deep shale formation," *Sustainability*, vol. 11, no. 22, p. 6527, 2019.
- [20] R. Schalkoff, *Artificial Neural Networks*, McGraw-Hill Education (India), The University of Michigan, 1997.
- [21] A. Ahmed, S. Elkatatny, A. Ali, M. Mahmoud, and A. Abdurraheem, "New model for pore pressure prediction while drilling using artificial neural networks," *Arabian Journal for Science and Engineering*, vol. 44, pp. 6079–6088, 2019.
- [22] S. Alarifi, S. AlNuaimi, and A. Abdurraheem, "Productivity index prediction for oil horizontal wells using different artificial intelligence techniques," in *SPE Middle East Oil & Gas Show and Conference*, Manama, Bahrain, 2015.
- [23] B. Kosko, *Neural Networks and Fuzzy Systems: A Dynamical Systems Approach to Machine Intelligence*, Prentice-Hall Inc., 1992.
- [24] Mathwork, *Fuzzy Logic Toolbox User's Guide*, MathWorks Inc., Natick, MA, USA, 2010, <http://citeseerx.ist.psu.edu/viewdoc/download?>
- [25] S. J. Cuddy and T. W. Putnam, "Litho-Facies and permeability prediction from electrical logs using fuzzy logic," *SPE Reservoir Evaluation & Engineering*, vol. 3, 1998.
- [26] R. B. C. C. Gharbi and G. A. Mansoori, "An introduction to artificial intelligence applications in petroleum exploration and production," vol. 49, 2005 *Journal of Petroleum Science and Engineering*, 2005.
- [27] J.-S. R. Jang, "ANFIS: adaptive-network-based fuzzy inference system," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 23, no. 3, pp. 665–685, 1993.
- [28] P. Tahmasebi and A. Hezarkhani, "A hybrid neural networks-fuzzy logic-genetic algorithm for grade estimation," *Computers & Geosciences*, vol. 42, pp. 18–27, 2012.
- [29] A. A. Anifowose and A. Abdurraheem, "Fuzzy logic-driven and SVM-driven hybrid computational intelligence models applied to oil and gas reservoir characterization," *Journal of Natural Gas Science and Engineering*, vol. 3, no. 3, pp. 505–517, 2011.
- [30] S. M. Elkatatny, T. Zeeshan, M. Mahmoud, A. Abdulazez, and I. M. Mohamed, "Application of artificial intelligent techniques to determine sonic time from well logs," in *50th U.S. Rock Mechanics/Geomechanics Symposium*, Houston, Texas, 2016 <https://www.onepetro.org/conference-paper/ARMA-2016-755>.
- [31] A. Ahmed, S. Elkatatny, A. Abdurraheem, M. Mahmoud, A. Z. Ali, and I. M. Mohamed, "Pore pressure prediction while drilling using fuzzy logic," in *SPE Kingdom of Saudi Arabia Annual Technical Symposium and Exhibition*, Dammam, Saudi Arabia, 2018.
- [32] F. Anifowose, T. Helmy, and A. Abdurraheem, "(PDF) Functional networks as a best-model selector in computational intelligence hybrid models for petroleum reservoir characterization," 2011, [https://www.researchgate.net/publication/215662101\\_Functional\\_Networks\\_as\\_a\\_Best-Model\\_Selector\\_in\\_Computational\\_Intelligence\\_Hybrid\\_Models\\_for\\_Petroleum\\_Reservoir\\_Characterization](https://www.researchgate.net/publication/215662101_Functional_Networks_as_a_Best-Model_Selector_in_Computational_Intelligence_Hybrid_Models_for_Petroleum_Reservoir_Characterization).
- [33] E. Castillo, "Functional networks," *Neural Processing Letters*, vol. 7, no. 3, pp. 151–159, 1998.
- [34] A. R. Moazzeni, M. Nabaei, and S. G. Jegarluei, "Prediction of lost circulation using virtual intelligence in one of Iranian oil-fields," in *SPE-136992-MS was presented at Nigeria Annual International Conference and Exhibition*, Tinapa - Calabar, Nigeria, 2010.
- [35] H. Toreifi, H. Rostami, and A. K. Manshad, "New method for prediction and solving the problem of drilling fluid loss using modular neural network and particle swarm optimization algorithm," *Journal of Petroleum Exploration and Production Technology*, vol. 4, no. 4, pp. 371–379, 2014.
- [36] G. Efendiyev, P. Mammadov, I. Piriverdiyev, and V. Mammadov, "Clustering of geological objects using FCM-algorithm and evaluation of the rate of lost circulation," *Procedia Computer Science*, vol. 102, pp. 159–162, 2016.
- [37] P. Behnoud far and P. Hosseini, "Estimation of lost circulation amount occurs during under balanced drilling using drilling data and neural network," *Egyptian Journal of Petroleum*, vol. 26, no. 3, pp. 627–634, 2017.
- [38] O. Solomon, D. Adewale, and C. Anyanwu, "Fracture width prediction and loss prevention material sizing in depleted formations using artificial intelligence," in *SPE-189068-MS was presented at SPE Nigeria Annual International Conference and Exhibition*, 31, Lagos, Nigeria, 2017.
- [39] A. Manshad, H. Rostami, H. Niknafs, and A. Mohammadi, *In book: Heavy Oil*, pp. 243–253, Nova Science Publishers, Inc., NY, USA, 2017.
- [40] A. T. T. Al-Hameedi, H. H. Alkinani, S. Dunn-Norman et al., "Mud loss estimation using machine learning approach," *Journal of Petroleum Exploration and Production Technology*, vol. 9, no. 2, pp. 1339–1354, 2019.
- [41] H. H. Alkinani, A. T. T. Al-Hameedi, S. Dunn-Norman, M. M. Alkhamis, and R. A. Mutar, "Prediction of lost circulation prior to drilling for induced fractures formations using artificial neural networks," in *SPE-195197-MS was presented at SPE Oklahoma City Oil and Gas Symposium*, Oklahoma City, OKC, USA, 2019.

- [42] A. K. Abbas, N. A. Al-Haideri, and A. A. Bashikh, "Implementing artificial neural networks and support vector machines to predict lost circulation," *Egyptian Journal of Petroleum*, vol. 28, no. 4, pp. 339–347, 2019.
- [43] R. A. L. Cristofaro, G. A. Longhin, A. A. Waldmann et al., "Artificial intelligence strategy minimizes lost circulation non-productive time in Brazilian deep water pre-salt," in *OTC-28034-MS Presented at the Offshore Technology Conference OTC Brasil*, Rio de Janeiro, Brazil, 2017.
- [44] R. Jahanbakhshi and R. Keshavarzi, "Quantitative and qualitative analysis of lost circulation in natural and induced fractured formations: the integration of operational conditions and geomechanical parameters," *European Journal of Environmental and Civil Engineering*, vol. 19, no. 4, pp. 418–444, 2014.
- [45] L. Sheremetov, I. Z. Batyrshin, D. Filatov, J. Martinez, and H. Rodriguez, "Fuzzy expert system for solving lost circulation problem," *Applied Soft Computing*, vol. 8, no. 1, pp. 14–29, 2008.