Journal of Energy Research and Reviews

Volume 15, Issue 2, Page 70-93, 2023; Article no.JENRR.107598
ISSN: 2581-8368

Machine Learning-Based Models for Basic Sediment & Water and Sand-Cut Prediction in Matured Niger Delta Fields

Frank A. Abuh [a,b], Julius U. Akpabio [b] and Anietie N. Okon [b*]

[a] Nigerian Upstream Petroleum Regulatory Commission, Abuja, Nigeria.
[b] Department of Chemical and Petroleum Engineering, University of Uyo, Uyo, Nigeria.

Authors' contributions

This work was carried out in collaboration among all authors. All authors read and approved the final manuscript.

Article Information

DOI: 10.9734/JENRR/2023/v15i2310

Open Peer Review History:
This journal follows the Advanced Open Peer Review policy. Identity of the Reviewers, Editor(s) and additional Reviewers, peer review comments, different versions of the manuscript, comments of the editors, etc are available here: https://www.sdiarticle5.com/review-history/107598

Original Research Article

Received: 10/08/2023
Accepted: 15/10/2023
Published: 19/10/2023

ABSTRACT

Oil production from matured fields in the Niger Delta is characterised by basic sediment and water (BS&W) and sand or sand-cut (S_{cut}) production. The predominant factor for this production is the unconsolidated nature of the formations in the Niger Delta. The available correlations for estimating BS&W and S_{cut} are based more on the intrinsic reservoir properties than controllable wellhead variables during oil production. This study developed neural-based models to predict BS&W and S_{cut} based on multiple-inputs single-output (MISO) and multiple-inputs multiple-outputs (MIMO) networks using 457 datasets from 43 oilfields in the Niger Delta. The performances of the neural-based models with new fields test datasets were determined using some statistical yardsticks: coefficient of determination (R^2), correlation coefficient (R), mean square error (MSE), root mean

**Corresponding author: E-mail: anietieokon@uniuyo.edu.ng;*

square error (RMSE), average relative error (ARE), and average absolute relative error (AARE). The results indicate that the MISO neural-based models had overall R and MSE values of 0.9999 and 2.0698×10^{-5}, respectively, for BS&W and 0.9995 and 2.1529×10^{-6} for S_{cut}. In contrast, the MIMO neural-based model had overall R and MSE values of 0.9997 and 7.5865×10^{-5}. The generalisation performance of the MISO neural-based models with new field test datasets resulted in R^2, R, MSE, RMSE, ARE and AAPRE of 0.97406, 0.98695, 2.08143, 1.44272, -0.00638 and 0.28755, respectively, for the BS&W model and R^2 of 0.89558, R of 0.93544, MSE of 0.01736, RMSE of 0.13177, ARE of 0.01338 and AARE of 0.01759 for the S_{cut} model. Furthermore, the MIMO-based model with new field test datasets resulted in R^2, R, MSE, RMSE, ARE and AAPRE of 0.97317, 0.98650, 2.15293, 1.46729, -0.00713 and 0.25064, respectively, for BS&W, while the S_{cut} model had R^2 of 0.87505, R of 0.93544, MSE of 0.02118, RMSE of 0.14554, ARE of -0.02280 and AARE of 0.02996. Also, the relative importance of the input parameters of the MISO and MIMO neural-based models in predicting BS&W and S_{cut} is $q_o > P_r > P_{wh} > S > \gamma_{API}$. Based on the statistical indicators obtained, the predictions of the developed neural models were close to the actual fields' datasets. Thus, the neural-based models should apply as tools for estimating BS&W and S_{cut} in mature fields in the Niger Delta.

Keywords: Machine learning-based models; basic sediment and water; sand-cut; matured fields; Niger Delta.

1. INTRODUCTION

Hydrocarbon extraction from reservoirs is often accompanied by the production of basic sediment and water (BS&W) and sand, especially in a matured oilfield with unconsolidated formations - like the Niger Delta. In other words, the crude oil drawn out of the reservoir contains suspended particles and water that were present during the reservoir's development. Sediment or mud are the terms used to describe the particles. When it comes to older fields or if a water flooding strategy promotes oil production, the water content can vary widely from field to field and may be present in large amounts [1]. To reduce the amount that needs to be transported or processed further, most of the produced water and silt are typically separated at the field [2]. As a result, BS&W measures the residual amount of these undesirable contaminants in crude oil [3]. As the reservoirs mature, the quantity of the BS&W and sand produced begins to compete with the produced reservoir stream (i.e. oil and gas) [4]. Most times, the cost of producing, handling, and disposing of the produced BS&W and sand is a serious concern to the operating companies [5].

In the Niger Delta fields, most producing wells produce oil with some basic sediment and water (BS&W) in or along the production tubing/flowlines at high flow rates and agitation [6]. Thus, the presence of BS&W and sand leads to the formation of crude oil emulsions. Regrettably, unlike its sand counterparts, limited attention is put into the production and control of BS&W. According to Salahi *et al*. [7], sand production is regarded as a severe production concern that significantly lowers wellbore productivity. A critical operational inefficiency that might cause wells to collapse is producing sand or solids in production operations [8]. In production wells situated in weak or weakly cemented sandstone reservoirs, its production is a regular issue. Along with the Niger Delta formation in Nigeria, fields in Canada, Egypt, the Gulf of Mexico, Indonesia, Malaysia, Trinidad, and Venezuela also have serious sand production issues. According to Osisanya [9], the reservoirs in these formations are between 3,500 and 10,000 feet deep. Until now, much work has gone into creating models that can forecast sand production over the next few decades. According to Chin and Ramos [10], the created models help with field operations like sand management and control, the best well completion design, and production optimisation.

The petroleum sector has been interested in addressing the effects of BS&W and sand production, ranging from economic and safety risks to well productivity. These consequences include downhole and surface equipment erosion, pipeline blockage and leakage, formation collapse, productivity loss, elevated costs, the complexity of intervention [11], prolonged downtime, and other environmental consequences of disposal, particularly in swampy and offshore areas [9]. Each year, these issues cost the oil sector billions of dollars.

Understanding the BS&W and sand production mechanisms is advantageous, and can foresee and control their production rate. Predicting the potential volume and frequency of products (such as sand and BS&W) that could be carried through the wellbore and into the facilities on the surface is crucial.

Strategies are continually being researched to relate their forecast, control, and management to avoid issues related to the production of sediments and sand. Bellarby [12] asserts that determining whether to utilise downhole sand management and the best kind depends on foreseeing when a formation will fail and start to produce sediments and sand. To develop a field optimally, one needs to understand the mechanism of sand production in terms of rate, volume, and sand-producing patterns in the reservoir [11]. According to Osisanya [9], the formation strength, flow stability, viscous drag forces, and pressure drop into the wellbore are all factors in the generation of sediments and sand. As a result, formation strength, in-situ stress, and production rate are crucial variables that influence correct predictions of sand production potential and output [13]. Additionally, there are formation features, pressure drawdown, reservoir pressure, produced fluid kinds and phases, natural permeability, formation cementation, compressibility, the exposed surface to flow, and reservoir depth [14].

Today, it takes a different science to determine whether a reservoir would yield sediment and sand. For predicting sand production, there are several empirical, numerical, and analytical models. The now available models require many rock mechanics characteristics rarely used in fieldwork. When quick-sand control decisions are required, some models' computations could be more practical [15]. On the other hand, the estimation of BS&W is carried out using the well-known bottle test method and the grab sample analysis. The bottle test method is time-consuming and expensive, requires visual phase separation observation, and makes it difficult to choose an effective demulsifier. In addition, as the samples are frequently taken in groups and might not completely represent the flow line, the data produced using this method could be incorrect, necessitating data reconciliation. Therefore, this study will use an artificial intelligence (machine learning) approach to develop neural network-based models for predicting BS&W and sand volume using production test datasets from matured fields in the Niger Delta.

2. METHODOLOGY

2.1 Data Acquisition and Preparation

The datasets, namely choke (beans) size (S), wellhead pressure (P_{wh}), reservoir pressure (P_r), oil gravity (γ_{API}) and oil flow rate (q_o) as input variables, and basic sediment and water (BS&W) and sand-cut (S_{cut}) as output parameters were collected from the matured fields in the Niger Delta region. The datasets collected were based on production test analysis (PTA) conducted on about forty-three (43) fields in the Niger Delta. The statistical description of the acquired datasets and the dependence of the output parameters on each input variable were determined using the correlation coefficient (R) (Equation 1) and presented in correlation plots (i.e. heat map).

$$R = \sqrt{1 - \frac{\sum_{i=1}^{n}\left(x_{inputdata_i} - y_{outputdata_i}\right)^2}{\sum_{i=1}^{n}\left(x_{inputdata_i} - \bar{y}_{outputdata_i}\right)^2}} \qquad (1)$$

where $x_{inputdata_i}$ represents the input variables, $y_{outputdata_i}$ denotes the output variables and $\bar{y}_{outputdata_i}$ is the average value of the output variables.

The acquired datasets were screened to remove or delete incomplete datasets and outliers. To ensure that any possible outliers in the datasets used for the network training were properly eliminated, box and whisker plots of the input datasets were determined. Because of the dataset ranges, these plots are presented in Figs. 1 and 2.

After that, the datasets were normalised using the maximum-minimum normalisation approach (Equation 2). This equation reduces the dataset ranges to between 0 and 1, which improves the learning of the neural network training and prediction (Tugwell and Livinus, [16] Therefore, normalisation brings the datasets to fall within the same range of values. It helps the training process run smoothly and increases the network's performance. According to Khare and Nagendra [17] both show comparable variance if one input has a large number and another small one. The possibility of the network ignoring the

small input in favour of the larger one would be sufficiently high if the data range is not reduced.

$$y_{normal.} = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} \quad (2)$$

where $y_{normal.}$ is the normalised values for input or output variables, x_i is the values of the not normalised variables, $x_{\min}$ and $x_{\max}$ represent the minimum and maximum values of the not scaled variables, respectively.

2.2 Neural Network-Based Models Development

2.2.1 Neural network training and evaluation

Neural network toolbox (nntool) in Matlab R2020a software was used to develop the neural-based models for predicting basic sediment and water (BS&W) and sand-cut (S_{cut}) in the Niger Delta oilfields. The basic settings of the software toolbox for the models' development are presented in Table 1. As indicated in the table, 319 datasets (about 70% of the datasets) were involved in the training of the neural networks, 69 datasets (15% of the datasets) were used for the network testing and another 69 datasets for the networks' prediction (performance) validation. The network training was based on trial and error to establish the number of neurons at the hidden layer. Also, two neural network configurations were considered, namely, multiple-input single-output (MISO) and multiple-input multiple-output (MIMO), for predicting BS&W and S_{cut}. The neural networks learned the input and output datasets using the Leverberg-Marquardt algorithm based on the feedforward back-propagation (FFBP) method. Also, the network used *tansig* and *purelin* as the activation or transfer functions at the hidden and output layers, respectively. Other default settings of the neural networks are the number of epochs (iterations) at 1000, the learning rate at 0.7, the targeted mean square error (MSE) at 10^{-7} and the minimum performance gradient at 10^{-7}. The neural network prediction performance was evaluated using the mean square error (MSE) and regression (i.e. cross) plots obtained from the training, testing, validation and the overall performance of the network predictions.

2.2.2 Determining the generalisation potential of the developed models

It would be challenging to compare the performance of the developed models in this study with the existing ones in the literature. The reason is that the developed machine learning-based models are based on production test analysis (PTA) datasets. In contrast, the existing AI-based models are based on reservoir rock geomechanics properties. In this regard, the generalisation potential of the developed models with new (unseen) datasets provides the needed yardstick to assess the developed models. Thus, 48 datasets from eight fields in the Niger Delta were used to evaluate the generalisation potential of the developed models. The models' input variables in this evaluation were normalised and exported from the Microsoft Excel environment to Matlab software. Using the developed neural-based networks, the input variables S, P_{wh}, P_r, γ_{API} and q_o were used to predict (i.e. unsupervised learning) the output variables BS&W and S_{cut}. Afterwards, the predicted outputs were denormalised and compared with the actual field test datasets.

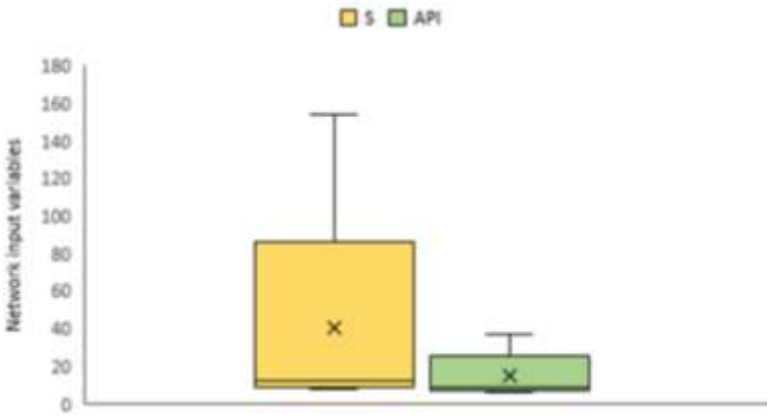

Fig. 1. Box and whisker plots of network input variable - choke size and API gravity

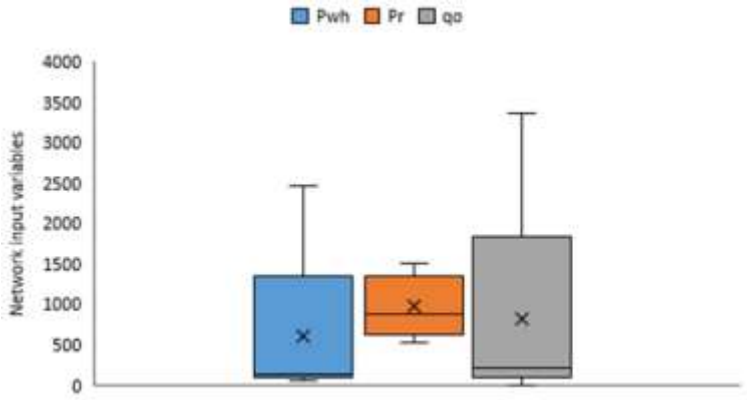

Fig. 2. Box and whisker plots of network input variable – wellhead pressure, reservoir pressure and oil flow rate

Table 1. Basic settings of the neural networks training parameters

Parameters	Values
Training Datasets	319 (70% of the datasets)
Validation Datasets	69 (15% of the datasets)
Testing Datasets	69 (15% of the datasets)
Hidden Layer	1
Hidden Layer Neurons	6, 3, 4
Hidden Layer Activation Function	*tansig*
Output Layer Activation Function	*purelin*
Learning Algorithm	Levenberg-Marquardt
Number of Epochs	1000
Rate of Learning	0.7
Architecture Selection	Trial-and-error
Target Goal MSE	10^{-7}
Minimum Performance Gradient	10^{-7}

2.2.3 Parametric sensitivity analysis

Parametric sensitivity analysis measures the variability of the input variables for the developed model and determines the effects on the output variable. According to Lawson and Marion [18], the analysis identifies the limitations of the developed model. In addition, since the model developed is representative of a problem studied, Okon *et al*. [19] maintained that assessing and predicting the input variable's relative importance on the output is important. Therefore, the parametric sensitivity analysis of the input variables on the output variables was evaluated using Garson's approach in Equation 3;

$$RI = \frac{\sum_{j}^{n_i}\left(\frac{|Lw_{ij}|}{\sum_{k}^{I_m}|Lw_{ij}|}Iw_i\right)}{\sum_{i}^{I_m}\sum_{j}^{n_i}\left(\frac{|Lw_{ij}|}{\sum_{k}^{I_m}|Lw_{ij}|}Iw_i\right)}\times 100\% \quad (3)$$

where Iw_i is the input layer weights, Lw_{ij} represents the hidden layer weights to the output neuron, n_i and I_m denote the numbers of inputs and hidden layer's neurons.

3. RESULTS AND DISCUSSION

3.1 Statistical Description and Correlation Coefficient of the Datasets

Table 2 provides the statistical description of the datasets used for the neural-based model development. The table showed that the output variables BS&W and S_{cut} have maximum, minimum, range, average and standard deviation values of 97.5, 0.005, 97.45, 52.52 and 27.889, respectively, for BS&W while S_{cut} had 61.0, 0.01, 60.99, 1.565 and 4.184. Also, Fig. 3 and 4 present the dependency or correlation among the datasets used for the models' development. From the figures, it is observed that there is less correlation among the datasets; the correlation

coefficients obtained for the datasets are less than 0.4, except for the P_{wh}-γ_{API} datasets with an R-value of 0.51.

3.2 Developed Neural Network Performance Evaluation

3.2.1 Basic Sediment and Water (BS&W) neural network performance

The developed neural network for basic sediment and water (BS&W) prediction is a feed-forward back-propagation network with input, hidden and output layers. The topology (architecture) in Fig. 5 indicated that the best BS&W predictions were obtained with 5-6-1. This outcome implies that the BS&W neural network has five neurons at the input layer, six at the hidden layer, and one at the output layer. Therefore, the developed network in Fig. 5 is a multiple-inputs single-output (MISO) neural network. Table 3 depicts the performance indices: mean square error (MSE) and correlation coefficient (R) values of the network during the training, validation and testing stage of the network development. As shown in Table 3, the BS&W neural network has MSE values of 1.5307×10^{-5}, 2.0698×10^{-5} and 2.2023×10^{-5} for training, validation and testing, respectively, with corresponding R values of 0.99991, 0.99989 and 0.99986 during the network development.

Table 2. Statistical description of the datasets for the development of the neural-based models

Parameters	Maximum	Minimum	Range	Average	Std. Dev.
Bean (Choke) Size, /64 inch	202.0	8.0	194.0	24.485	34.207
Wellhead Pressure, P_{wh}, psia	3045.0	53.65	2991.35	507.574	530.76
Reservoir Pressure, P_r, psia	4835.0	886.22	3948.78	2712.93	828.78
Oil Gravity, γ_{API}	72.7	14.0	58.70	28.61	8.725
Oil Flow Rate, q_o, stb/d	4063.0	4.56	4058.44	526.23	484.95
Basic Sediment & Water, *BS&W*, %	97.5	0.005	97.45	52.52	27.889
Sand-cut, S_{cut}	61.0	0.01	60.99	1.565	4.184

Fig. 3. Correlation plot of the input variables with output - BS&W

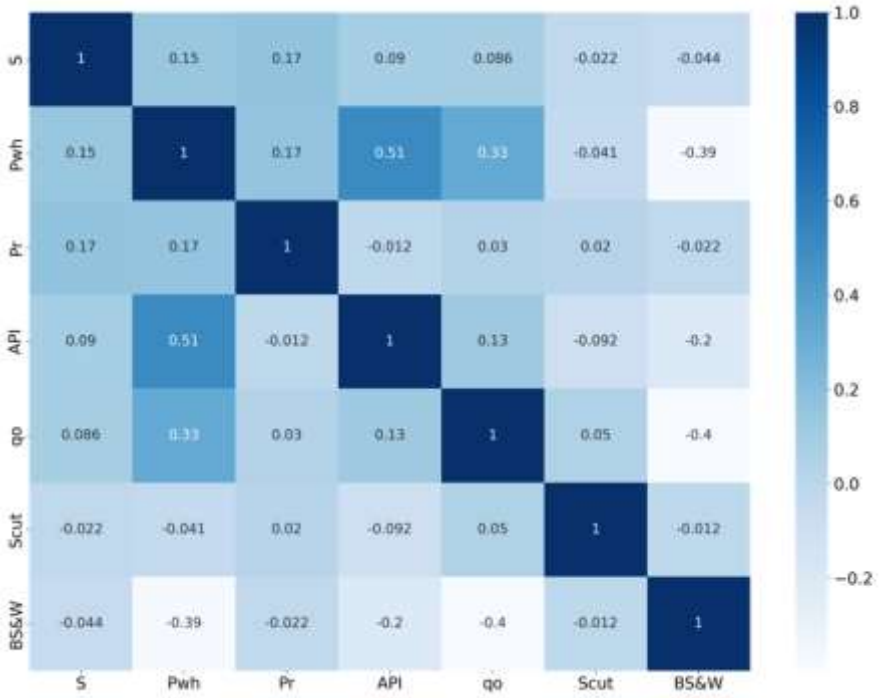

Fig. 4. Correlation plot of the input variables with output - S_{cut}

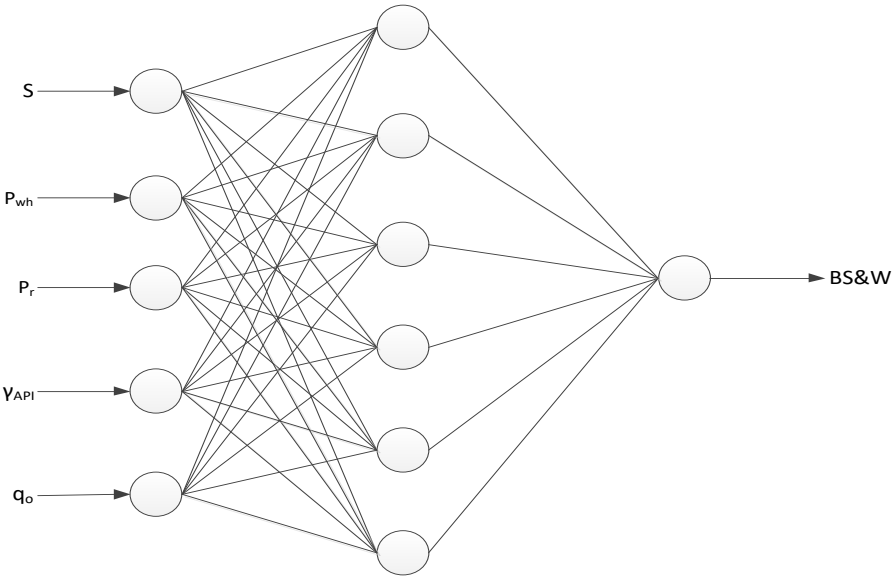

Fig. 5. Artificial neural network topology for BS&W model

Table 3. Performance indices of the BS&W neural network training, validation and testing

	NN-Model	Indices	Training	Validation	Testing	Overall
i.	BS&W model	MSE	1.5307×10^{-5}	2.0698×10^{-5}	2.2023×10^{-5}	2.0698×10^{-5}
		R	0.99991	0.99989	0.99986	0.99990

The overall MSE and R values of the BS&W neural network are 2.0698×10^{-5} and 0.99990, as indicated in Table 3 and Fig. 6. The overall performance of the BS&W network showed that the predictions were close to the fields' BS&W datasets. The observation is because the MSE and R values are within acceptable limits for the network performance. Therefore, based on the R-value obtained, the network can predict the fields' BS&W with 99.0% certainty. Again, the closeness of the network-predicted BS&W with the fields BS&W datasets is noted on the diagonal trend of the output (i.e. network predictions) and target (fields datasets) for the overall performance in Fig. 6. According to Al-Bulushi *et al.* [20] and Okon *et al*. [21] the field and predicted data points aligned along a unit slope imply a good agreement between them, as observed in Fig. 6.

The limitation alluded to most neural network models in the literature is the availability of the essential details of the models for their reproducibility. One such detail is the weights and biases of the neural network [22]. This study's weights and biases of the BS&W network are depicted in Table 4. The neural network equation to represent the BS&W network is expressed in Equation 4.

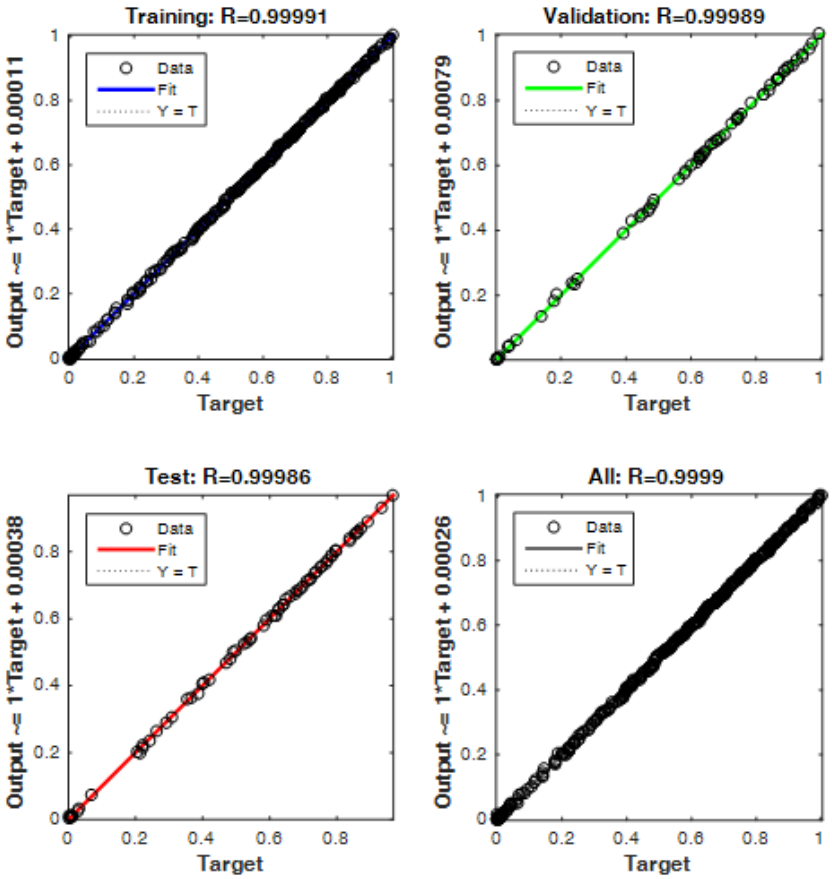

Fig. 6. Regression plot of the BS&W network training, validation, and testing performance

$$(BS\,\&W)_{ANN} = \sum purelin\left\{\sum_{i=1}^{6}\sum_{j=1}^{5} t\,\text{an}\,sig\left[\left(Sj_1 + P_{wh}j_2 + P_r j_3 + \gamma_{API} j_4 + q_o j_5\right)_i + b_i\right]\right\}\ldots$$

$$..\times Lw_{ij} + b_k \tag{4}$$

where $(BS\,\&W)_{ANN}$ is the neural network predicted oil flow rate in normalised form, the variables j_1, j_2, j_3, j_4, and j_5 are the weights of the network inputs: S, P_{wh}, P_r, γ_{API} and q_o, to the hidden

layer neuron; Lw_{ij} represents the hidden layer weights that connect the output layer neuron; b_i and b_k are biases at the hidden and output neurons, respectively.

Table 4. Weights and biases of the BS&W neural network

	Input weights					Hidden biases	Hidden weights	Output bias
i	$(S)\,j_1$	$(P_{wh})\,j_2$	$(P_r)\,j_3$	$(\gamma_{API})\,j_4$	$(q_o)\,j_5$	b_i	Lw_1	b_k
1	-0.334516	0.4749149	-0.339765	1.4812252	1.2542208	1.166288	0.922581	-2.224626
2	0.4004205	-0.069611	-2.022304	1.4259147	2.5346743	2.146209	-2.290278	
3	-4.436704	4.3051040	4.080693	0.3355494	7.4627034	8.835142	0.841016	
4	1.2970319	2.8070231	-13.56891	-7.2452882	3.0478121	-5.727893	-3.169822	
5	-2.288220	-2.236867	12.224117	5.0411335	1.9597688	7.608174	-3.415598	
6	-0.501180	0.1127184	1.4133671	-0.6361507	-1.925401	-2.206243	-4.007471	

3.2.2 Developed sand-cut (S_{cut}) neural network performance

The developed neural network for sand-cut (S_{cut}) prediction is a feed-forward back-propagation network with input, hidden and output layers. The architecture (topology) in Fig. 7 indicated that the best S_{cut} predictions were obtained with 5-3-1. This outcome implies that the S_{cut} neural network has five neurons at the input layer, three at the hidden layer and one at the output layer. Thus, the developed network in Fig. 7 is a multiple-inputs multiple-outputs (MISO) neural network. Table 5 depicts the performance indices: mean square error (MSE) and correlation coefficient (R) values of the network during the training, validation and testing stage of the network development. As shown in Table 5, the S_{cut} neural network has MSE values of 5.7487×10^{-6}, 2.1529×10^{-6} and 2.2330×10^{-6} for training, validation and testing, respectively, with corresponding R values of 0.99953, 0.99799 and 0.99957 during the network development.

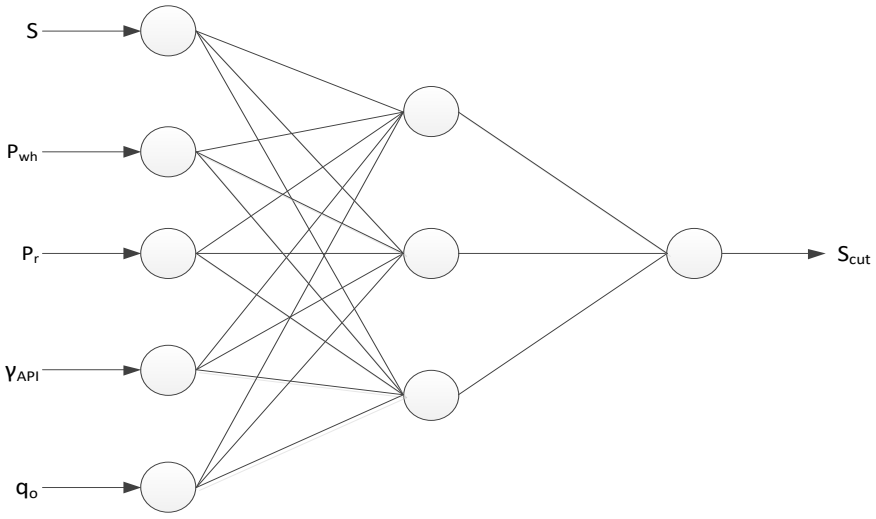

Fig. 7. Artificial neural network topology for S_{cut} model

Table 5. Performance indices of the S_{cut} neural network during training validation and testing

	NN-Model	Indexes	Training	Validation	Testing	Overall
i.	S_{cut} model	MSE	$5.7487\text{x}10^{-6}$	$2.1529\text{x}10^{-6}$	$2.2330\text{x}10^{-6}$	$2.1529\text{x}10^{-6}$
		R	0.99953	0.99799	0.99957	0.99950

The S_{cut} neural network's overall MSE and R values are 2.1529×10^{-6} and 0.99950, as indicated in Fig. 8. The overall performance of the S_{cut} network showed that the predictions were close to the fields' S_{cut} datasets. The observation is because the MSE and R values are within acceptable limits for the network performance. Therefore, based on the R-value obtained, the network can predict the fields' S_{cut} with 99.0% certainty. Again, the closeness of the network-predicted S_{cut} with the fields S_{cut} datasets is noted on the diagonal trend of the output (i.e. network predictions) and target (fields datasets) for the overall performance in Fig.8. According to Al-Bulushi *et al.* [20] and Okon *et al*. [21], when the field and predicted data points aligned along a unit slope, it implies a good agreement between them, as observed in Fig. 8. The S_{cut} neural network weights and biases for its reproducibility are presented in Table 6. Again, the neural network equation to represent the S_{cut} network is expanded in Equation 5.

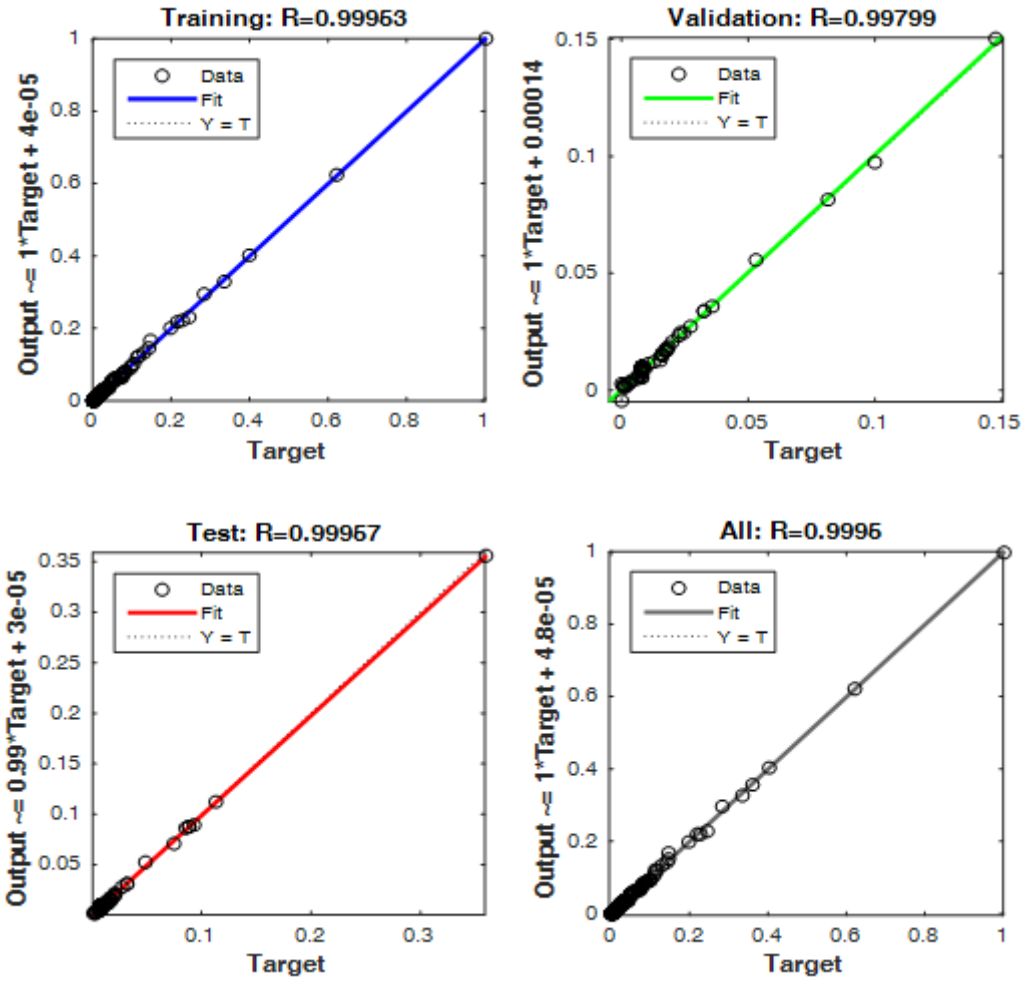

Fig. 8. Regression plot of the S_{cut} network training, validation, and testing performance

Table 6. Weights and biases of the S_{cut} neural network

	Input weights					Hidden biases	Hidden weights	Output bias
i	$(S)\, j_1$	$(P_{wh})\, j_2$	$(P_r)\, j_3$	$(\gamma_{API})\, j_4$	$(q_o)\, j_5$	b_i	Lw_1	b_k
1	-2.429965	-2.337407	-1.217410	2.2068083	-0.6970751	4.1571864	-1.8438429	1.4915672
2	-0.253977	2.5537693	0.9511435	-0.749508	-4.948833	0.3704629	-0.3327475	
3	-0.239506	-0.845081	-0.388827	0.131459	1.3135761	0.9866002	-0.4260308	

$$(S_{cut})_{ANN} = \sum purelin\left\{\sum_{i=1}^{3}\sum_{j=1}^{5} t\,\text{an}\,sig\left[\left(Sj_1 + P_{wh}\,j_2 + P_r\,j_3 + \gamma_{API}\,j_4 + q_o\,j_5\right)_i + b_i\right]\right\}...$$

$$..\times Lw_{ij} + b_k \quad (4)$$

where $(S_{cut})_{ANN}$ is the neural network predicted sand-cut in normalised form, the variables j_1, j_2, j_3, j_4, and j_5 are the weights of the network inputs: S, P_{wh}, P_r, γ_{API} and q_o, to the hidden layer neuron; Lw_{ij} represents the hidden layer weights that connect the output layer neuron; b_i and b_k are biases at the hidden and output neurons, respectively.

3.2.3 Developed Basic Sediment and Water (BS&W) - Sand-cut (S_{cut}) neural network performance

The developed neural network for basic sediment and water (BS&W) - sand-cut (S_{cut}) prediction is a feed-forward back-propagation network with input, hidden and output layers. The architecture (topology) in Fig. 9 indicated that the best BS&W-S_{cut} predictions were obtained with 5-4-2. This outcome implies that the BS&W-S_{cut} neural network has five neurons at the input layer, four at the hidden layer, and two at the output layer. Thus, the developed network in Fig. 9 is a multiple-inputs single-output (MIMO) neural network. Table 7 depicts the performance indices: mean square error (MSE) and correlation coefficient (R) values of the network during the training, validation and testing stage of the network development. As shown in Table 7, the BS&W-S_{cut} neural network has MSE values of 6.6144×10^{-5}, 7.5865×10^{-5} and 9.1633×10^{-5} for training, validation and testing, respectively, with corresponding R values of 0.99971, 0.99967 and 0.99961 during the network development.

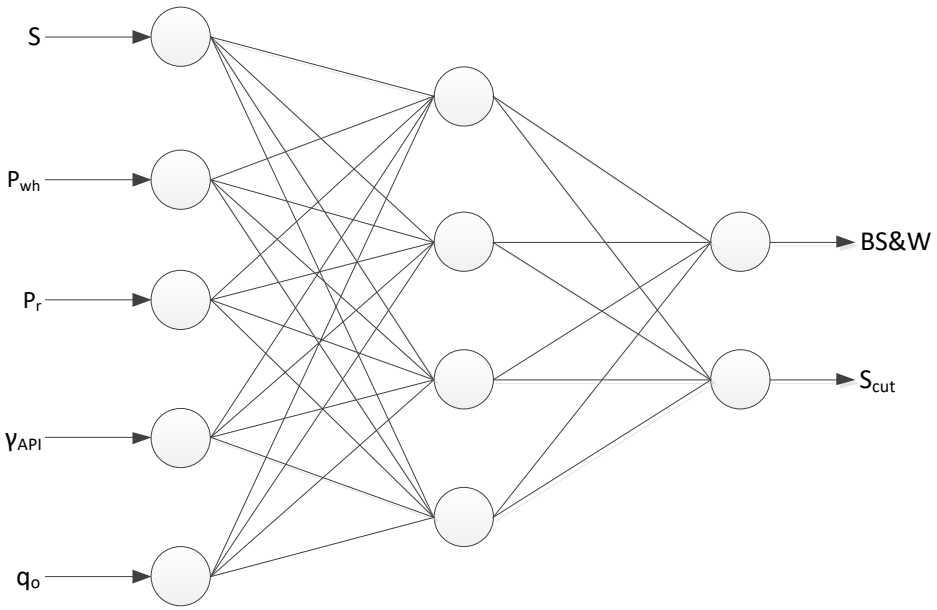

Fig. 9. Artificial neural network topology for BS&W-S_{cut} model

Table 7. Performance indices of the BS&W-S_{cut} neural network during training validation and testing

	NN-Model	Indexes	Training	Validation	Testing	Overall
i.	BS&W-S_{cut} model	MSE	6.6144×10^{-5}	7.5865×10^{-5}	9.1633×10^{-5}	7.5865×10^{-5}
		R	0.99971	0.99967	0.99961	0.99968

Table 8. Weights and biases of the BS&W-S_{cut} neural network

Input weights						Hidden biases	Hidden weights		Output bias
i	$(S)\,j_1$	$(P_{wh})\,j_2$	$(P_r)\,j_3$	$(\gamma_{API})\,j_4$	$(q_o)\,j_5$	b_i	Lw_1	Lw_2	b_k
1	-1.750413	-0.908136	-1.050490	0.289735	-0.358459	3.604541	-0.019242	-2.592108	-0.187536
2	-0.077740	-0.609269	0.537582	0.626393	-1.451739	-1.127728	-0.249118	-0.120460	1.629434
3	-1.173151	1.393986	-6.054674	-1.238638	-0.755157	-4.003239	-0.425717	-0.004425	
4	-0.397865	-0.851939	-0.557050	0.101713	-2.705917	-3.126723	-0.518810	-0.027540	

The overall MSE and R values of the BS&W-S_{cut} neural network are 7.5865×10^{-5} and 0.99968, as indicated in Table 7 and Fig.10. The overall performance of the BS&W-S_{cut} network showed that the predictions were close to the BS&W and S_{cut} datasets fields. The observation is because the MSE and R values are within acceptable limits for the network performance. Therefore, based on the R-value obtained, the network can predict the fields' BS&W and S_{cut} with 99.0% certainty. Again, the closeness of the network-predicted BS&W-S_{cut} with the fields BS&W and S_{cut} datasets is noted on the diagonal trend of the output (i.e. network predictions) and target (fields datasets) for the overall performance in Fig.10. According to Al-Bulushi *et al*. [20] and Okon *et al*. [21] the field and predicted data points aligned along a unit slope imply a good agreement between them, as observed in Fig. 10. The BS&W-S_{cut} neural network weights and biases for its reproducibility are presented in Table 8. Again, the neural network equation to represent the BS&W-S_{cut} network is expanded in Equation 6.

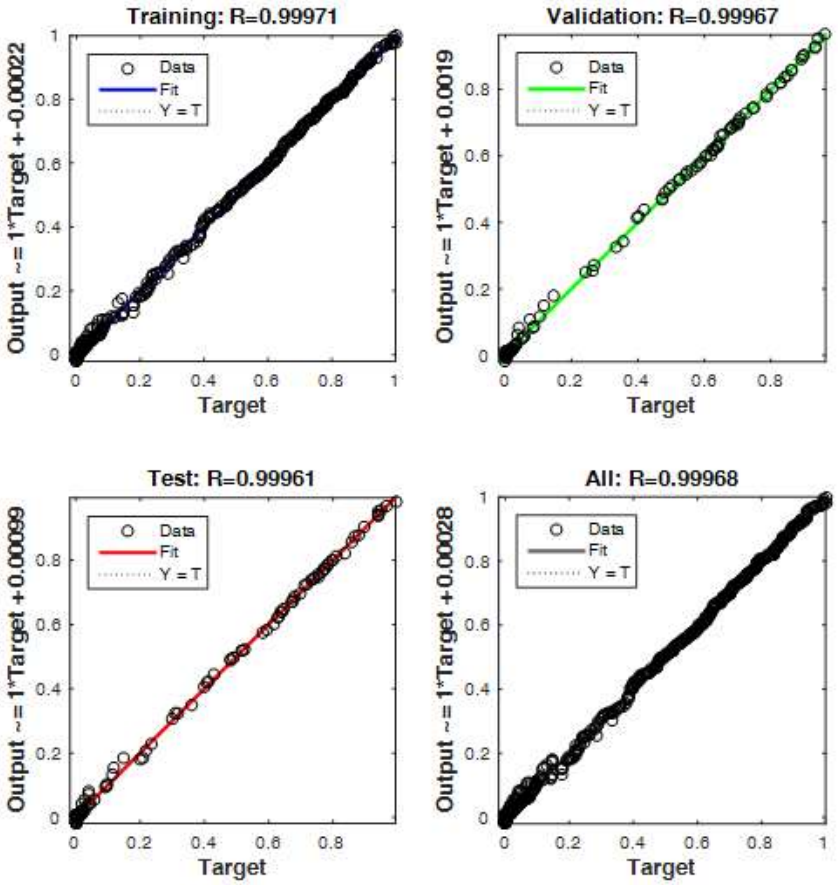

Fig. 10. Regression plot of the S_{cut} network training, validation and testing performance

$$(BS\,\&W - S_{cut})_{ANN} = \sum purelin\left\{\sum_{i=1}^{4}\sum_{j=1}^{5} t\,\text{an}\,sig\left[\left(Sj_1 + P_{wh}j_2 + P_r j_3 + \gamma_{API} j_4 + q_o j_5\right)_i + b_i\right]\right\}... \\ ..\times Lw_{ij} + b_k \quad (3)$$

where $(BS\,\&W - S_{cut})_{ANN}$ is the neural network predicted basic sediment and water (BS&W) and sand-cut in normalised form, the variables j_1, j_2, j_3, j_4 and j_5 are the weights of the network inputs: S, P_{wh}, P_r, γ_{API} and q_o, to the hidden layer neuron; Lw_{ij} represents the hidden layer weights that connect the output layer neuron; b_i and b_k are biases at the hidden and output neurons, respectively.

3.3 Relative Importance of the Developed Neural Networks Input Variables

Tables 9 through 11 present the relative importance (RI) of the neural network input parameters on the predicted outputs: BS&W and S_{cut} using Garson's method (Equation 3) for estimating relative importance (RI). In Table 9, the results obtained indicated that reservoir pressure (P_r), oil flow rate (q_o) and oil gravity (γ_{API}) were more relatively important on the BS&W neural network than wellhead pressure (P_{wh}) and choke (beans) size (S). The RI results showed that P_r had 31.78%, q_o had 28.14%, γ_{API} had 20.48%, S had 10.25%, and P_{wh} had 9.35%. Thus, the BS&W neural network input variables ranking on the output is $P_r > q_o > \gamma_{API} > S > P_{wh}$.

On the other hand, in Table 10, the results obtained for S_{cut} neural network RI depicted that oil flow rate (q_o) and wellhead pressure (P_{wh}) were more relatively important on the S_{cut} neural network than choke (beans) size (S), oil gravity (γ_{API}) and reservoir pressure (P_r). The RI results indicated that q_o had 35.06%, P_{wh} had 27.42%, S had 12.74%, γ_{API} had 12.42%, and P_r had 12.36%. Therefore, the Scut neural network input parameters ranking on the output is $q_o > P_{wh} > S > \gamma_{API} > P_R$.

Furthermore, the relative importance of the BS&W-S_{cut} neural network input variables on the outputs (i.e. BS&W and S_{cut}) showed that oil flow rate (q_o) and reservoir pressure (P_r) were more relatively important on the BS&W-S_{cut} neural network than wellhead pressure (P_{wh}) and choke (beans) size (S); they were substantially significant than oil gravity (γ_{API}). The RI results resulted in 39.31%, 36.50%, 23.63%, 20.73% and 13.16% for q_o, P_r, P_{wh}, S and γ_{API}, respectively (Table 11). Therefore, the BS&W-S_{cut} neural network input variables ranking on the outputs is $q_o > P_r > P_{wh} > S > \gamma_{API}$. In summary, the most predominant input variables of the developed neural networks are q_o, P_r and P_{wh}. This observation implies that the accurate prediction of BS&W and S_{cut} would depend on the quality and reliability of the mentioned parameters datasets.

Table 9. BS&W neural network input variables' relative importance on the network output

	S	**P_{wh}**	**P_r**	γ_{API}	**q_o**
	0.086112314	0.122254494	0.087463722	0.381302932	0.322866538
	0.062052557	0.010787526	0.313393408	0.220971857	0.392794652
	0.215157208	0.208775297	0.197892533	0.016272412	0.36190255
	0.046378784	0.100372484	0.485191893	0.259074311	0.108982528
	0.096345653	0.094183446	0.514697381	0.212257306	0.082516214
	0.109217739	0.024563721	0.308002454	0.138630646	0.419585439
Sum	0.615264254	0.560936969	1.906641391	1.228509465	1.688647921
RI, %	10.25440424	9.348949481	31.77735651	20.47515775	28.14413202
Ranking	**4th**	**5th**	**1st**	**3rd**	**2nd**

Table 10. S_{cut} neural network neural network input variables relative importance of the network output

	S	P_{wh}	P_r	γ_{API}	q_o
	0.273377929	0.262964912	0.136962029	0.248272165	0.078422966
	0.026855289	0.270033538	0.100573161	0.079252363	0.523285649
	0.082066135	0.28956514	0.133230712	0.045043961	0.450094052
Sum	0.382299353	0.822563589	0.370765902	0.372568489	1.051802667
RI, %	12.74331178	27.4187863	12.3588634	12.41894962	35.06008889
Ranking	**3rd**	**2nd**	**5th**	**4th**	**1st**

Table 11. BS&W-S_{cut} neural network neural network input variables relative importance of the network output

	S	P_{wh}	P_r	γ_{API}	q_o
	0.401725932	0.208420396	0.241091036	0.066495072	0.082267564
	0.023538108	0.184474693	0.162769359	0.189659606	0.439558233
	0.11051193	0.13131481	0.570355912	0.116680838	0.07113651
	0.086220872	0.18462285	0.120717806	0.022042072	0.586396399
Sum	0.621996843	0.70883275	1.094934113	0.394877588	1.179358706
RI, %	20.73322809	23.62775832	36.49780376	13.16258628	39.31195688
Ranking	**4th**	**3rd**	**2nd**	**5th**	**1st**

3.4 Explicit Representation of the Neural Network-Based Models

Earlier works by Okon and Ansa [23] and Okon *et al*. [19] reported that several authors have presented neural networks or models in "black box" form. The developed neural-based models are not in a simplified mathematical form as presented in Equations 7 and 21. This drawback limits interested readers to understanding the application of any developed neural network-based model [23].

3.4.1 Explicit BS&W neural-based model

$$BS\&W = 0.005 + 97.45\left(BS\&W\right)_{ANN} \quad (7)$$

where $BS\&W$ is the de-normalised basic sediment and water, $\left(BS\&W\right)_{ANN}$ is the neural network predicted basic sediment and water in normalised form.

Then, based on a multiple-inputs single-output (MISO) neural network, $\left(BS\&W\right)_{ANN}$ is presented in Equation 8;

$$\left(BS\&W\right)_{ANN} = \left[0.92258\sigma\left(z_1\right) - 2.29028\sigma\left(z_2\right) + 0.84102\sigma\left(z_3\right) - 3.16982\sigma\left(z_4\right)...\right.$$
$$\left..-3.41560\sigma(z_5) - 4.00747\sigma(z_6)\right] - 2.22463 \quad (8)$$

Then, $\sigma\left(z_1\right)$ through $\sigma\left(z_6\right)$ in Equation 4.5 is expressed as $\sigma\left(z_1\right) = \frac{2}{1+e^{-2(\beta_1)}} - 1$, $\sigma\left(z_2\right) = \frac{2}{1+e^{-2(\beta_2)}} - 1$, $\sigma\left(z_3\right) = \frac{2}{1+e^{-2(\beta_3)}} - 1$, $\sigma\left(z_4\right) = \frac{2}{1+e^{-2(\beta_4)}} - 1$, $\sigma\left(z_5\right) = \frac{2}{1+e^{-2(\beta_5)}} - 1$, and $\sigma\left(z_6\right) = \frac{2}{1+e^{-2(\beta_6)}} - 1$

where β_1 through β_6 are the computations at the hidden neurons presented in Equations 9 through 14;

$$\beta_1 = \left[-0.33452(S)_n + 0.47492(P_{wh})_n - 0.33977(P_r)_n + 1.48123(\gamma_{API})_n \ldots \right.$$
$$\left. \ldots + 1.25422(q_o)_n\right] + 1.16629 \quad (9)$$

$$\beta_2 = \left[0.40042(S)_n - 0.06961(P_{wh})_n - 2.02230(P_r)_n + 1.42592(\gamma_{API})_n \ldots \right.$$
$$\left. \ldots + 2.53467(q_o)_n\right] + 2.14621 \quad (10)$$

$$\beta_3 = \left[-4.43670(S)_n + 4.30510(P_{wh})_n + 4.08069(P_r)_n + 0.33555(\gamma_{API})_n \ldots \right.$$
$$\left. \ldots + 7.46270(q_o)_n\right] + 8.83514 \quad (11)$$

$$\beta_4 = \left[1.29703(S)_n + 2.80702(P_{wh})_n - 13.56891(P_r)_n - 7.24529(\gamma_{API})_n \ldots \right.$$
$$\left. \ldots + 3.04781(q_o)_n\right] - 5.72789 \quad (12)$$

$$\beta_5 = \left[-2.28822(S)_n - 2.23687(P_{wh})_n + 12.22412(P_r)_n + 5.04113(\gamma_{API})_n \ldots \right.$$
$$\left. \ldots + 1.95977(q_o)_n\right] + 7.60817 \quad (13)$$

$$\beta_6 = \left[-0.50118(S)_n + 0.11272(P_{wh})_n + 1.41337(P_r)_n - 0.63615(\gamma_{API})_n \ldots \right.$$
$$\left. \ldots - 1.92540(q_o)_n\right] - 2.20624 \quad (14)$$

On the other hand, based on multiple-inputs, multiple-outputs (MIMO) neural network $(BS\,\&W)_{ANN}$ is presented in Equation 15;

$$(BS\,\&W)_{ANN} = \left[-0.01924\sigma(z_1) - 0.24912\sigma(z_2) - 0.42572\sigma(z_3) - \ldots \right.$$
$$\left. \ldots - 0.51881\sigma(z_4)\right] - 0.18754 \quad (15)$$

Then, $\sigma(z_1)$ through $\sigma(z_4)$ in Equation 15, are expressed as $\sigma(z_1) = \frac{2}{1+e^{-2(\beta_1)}} - 1$, $\sigma(z_2) = \frac{2}{1+e^{-2(\beta_2)}} - 1$, $\sigma(z_3) = \frac{2}{1+e^{-2(\beta_3)}} - 1$, and $\sigma(z_4) = \frac{2}{1+e^{-2(\beta_4)}} - 1$

where β_1 through β_4 are the computations at the hidden neurons presented in Equations 16 through 19;

$$\beta_1 = \left[-1.75041(S)_n - 0.90814(P_{wh})_n - 1.05049(P_r)_n + 0.28974(\gamma_{API})_n \ldots \right.$$
$$\left. \ldots - 0.35846(q_o)_n\right] + 3.60454 \quad (16)$$

$$\beta_2 = \left[-0.07774\left(S\right)_n - 0.60927\left(P_{wh}\right)_n + 0.53758\left(P_r\right)_n + 0.62639(\gamma_{API})_n \ldots \right.$$
$$\left. \ldots -1.45174(q_o)_n \right] - 1.12773 \quad (17)$$

$$\beta_3 = \left[-1.17315\left(S\right)_n + 1.39399\left(P_{wh}\right)_n - 6.05467\left(P_r\right)_n - 1.23864(\gamma_{API})_n \ldots \right.$$
$$\left. \ldots -0.75516(q_o)_n \right] - 4.00324 \quad (18)$$

$$\beta_4 = \left[-0.39787\left(S\right)_n - 0.85194\left(P_{wh}\right)_n - 0.55705\left(P_r\right)_n + 0.10171(\gamma_{API})_n \ldots \right.$$
$$\left. \ldots -2.70592(q_o)_n \right] - 3.12672 \quad (19)$$

Okon *et al*. [20] presented detailed workings of the neural network to achieve its prediction (output) from the input variables. Following the reported steps in the mentioned source or reference, the basic neural network computations steps of the BS&W neural-based models are as follows:

i. At the input layer, input variables (S , P_{wh} , P_r , γ_{API} and q_o) from the input neurons multiply with input weights (j_1 , j_2 , j_3 , j_4 and j_5), respectively, and are linked to hidden layer neurons;

ii. At the upper hidden layer neuron (i.e. *i* = 1), the input (i.e. $Sj_1 + P_{wh}j_2 + P_r j_3 + \gamma_{API} j_4 + q_o j_5$) from the input layer combined with the upper neuron's bias (b_i) and the sum (i.e. $\sum_{i=1}\left(Sj_1 + P_{wh}j_2 + P_r j_3 + \gamma_{API} j_4 + q_o j_5\right) + b_i$) is transformed by the sigmoid function (Equation 20), to the output neuron;

$$\sigma\left(z_i\right) = \frac{2}{1+e^{-2(\beta_i)}} - 1 \quad (20)$$

where β_i is $\sum_{i=1}\left(Sj_1 + P_{wh}j_2 + P_r j_3 + \gamma_{API} j_4 + q_o j_5\right) + b_i$

iii. The transformed output from the upper hidden neuron (i.e. $\sigma\left(z_i\right)$) multiplied with the upper hidden neuron weight (Lw_i) and linked to the output neuron in the output layer;

iv. At the output neuron, the output from the hidden layer combined with the output neuron's bias (b_k), thus, $\left(\sigma(z_i)\times Lw_i\right) + b_k$

v. steps (i) through (iv) are repeated for values of *i* = 2,..6 for the neurons and at the output neuron, the sum $\sum_{j=1}^{5}\sum_{i=1}^{6}\left[\left(\sigma(z_{ij})\times Lw_{ij}\right) + b_{kj}\right]$ is transformed using the *purelin* function as the network's output. Thus, the predicted values are $purelin\sum_{j=1}^{5}\sum_{i=1}^{6}\left[\left(\sigma(z_{ij})\times Lw_{ij}\right) + b_{kj}\right]$.

The values for the variables j_1 , j_2 , j_3 , j_4 , j_5 , b_i , Lw_{ij} , and b_{kj} are in Tables 4 and 8 for MISO and MIMO-based neural networks, respectively. The output (i.e. BS&W) from the neural network is presented in the normalised form in Equations 4 and 6, which would require de-normalisation to transform the network predictions to the required BS&W values. Thus, the explicit neural network-based model for basic sediment and water (BS&W) prediction is in Equation 7.

3.4.2 Explicit S_{cut} neural-based model

$$S_{cut} = 0.01 + 60.99\left(S_{cut}\right)_{ANN} \tag{21}$$

where S_{cut} is the de-normalised sand-cut, $\left(S_{cut}\right)_{ANN}$ is the neural network predicted sand-cut in normalised form. Then, based on a multiple-inputs single-output (MISO) neural network, it $\left(S_{cut}\right)_{ANN}$ is presented in Equation 22;

$$\left(S_{cut}\right)_{ANN} = \left[-1.84384\sigma\left(z_1\right) - 0.33275\sigma\left(z_2\right) - 0.42603\sigma\left(z_3\right)\right] + 1.49157 \tag{22}$$

Then, $\sigma\left(z_1\right)$ through $\sigma\left(z_3\right)$ in Equation 17, it is expressed as $\sigma\left(z_1\right) = \frac{2}{1+e^{-2(\omega_1)}} - 1$, $\sigma\left(z_2\right) = \frac{2}{1+e^{-2(\omega_2)}} - 1$, and $\sigma\left(z_3\right) = \frac{2}{1+e^{-2(\omega_3)}} - 1$

where ω_1 through ω_3 are the computations at the hidden neurons presented in Equations 23 through 25;

$$\omega_1 = \left[-2.42997\left(S\right)_n - 2.33741\left(P_{wh}\right)_n - 1.21741\left(P_r\right)_n + 2.20681(\gamma_{API})_n \ldots \right.$$
$$\left. \ldots - 0.69708(q_o)_n\right] + 4.15719 \tag{23}$$

$$\omega_2 = \left[-0.25398\left(S\right)_n + 2.55377\left(P_{wh}\right)_n + 0.95114\left(P_r\right)_n - 0.74951(\gamma_{API})_n \ldots \right.$$
$$\left. \ldots - 4.94883(q_o)_n\right] + 0.37046 \tag{24}$$

$$\omega_3 = \left[-0.23951\left(S\right)_n - 0.84508\left(P_{wh}\right)_n - 0.38883\left(P_r\right)_n + 0.13146(\gamma_{API})_n \ldots \right.$$
$$\left. \ldots + 1.31357(q_o)_n\right] + 0.98660 \tag{25}$$

On the other hand, based on multiple-inputs, multiple-outputs (MIMO) neural network $\left(S_{cut}\right)_{ANN}$ is presented in Equation 26;

$$\left(S_{cut}\right)_{ANN} = \left[-2.59211\sigma\left(z_1\right) - 0.12046\sigma\left(z_2\right) - 0.00443\sigma\left(z_3\right) - \ldots \right.$$
$$\left. \ldots - 0.02754\sigma(z_4)\right] + 1.62943 \tag{26}$$

Then, $\sigma\left(z_1\right)$ through $\sigma\left(z_4\right)$ in Equation 26, are expressed as $\sigma\left(z_1\right) = \frac{2}{1+e^{-2(\omega_1)}} - 1$, $\sigma\left(z_2\right) = \frac{2}{1+e^{-2(\omega_2)}} - 1$, $\sigma\left(z_3\right) = \frac{2}{1+e^{-2(\omega_3)}} - 1$, and $\sigma\left(z_4\right) = \frac{2}{1+e^{-2(\omega_4)}} - 1$

where ω_1 through ω_4 are the computations at the hidden neurons presented in Equations 27 through 30;

$$\omega_1 = \left[-1.75041(S)_n - 0.90814(P_{wh})_n - 1.05049(P_r)_n + 0.28974(\gamma_{API})_n \ldots \right.$$
$$\left. \ldots - 0.35846(q_o)_n\right] + 3.60454 \quad (27)$$

$$\omega_2 = \left[-0.07774(S)_n - 0.60927(P_{wh})_n + 0.53758(P_r)_n + 0.62639(\gamma_{API})_n \ldots \right.$$
$$\left. \ldots - 1.45174(q_o)_n\right] - 1.12773 \quad (28)$$

$$\omega_3 = \left[-1.17315(S)_n + 1.39399(P_{wh})_n - 6.05467(P_r)_n - 1.23864(\gamma_{API})_n \ldots \right.$$
$$\left. \ldots - 0.75516(q_o)_n\right] - 4.00324 \quad (29)$$

$$\omega_4 = \left[-0.39787(S)_n - 0.85194(P_{wh})_n - 0.55705(P_r)_n + 0.10171(\gamma_{API})_n \ldots \right.$$
$$\left. \ldots - 2.70592(q_o)_n\right] - 3.12672 \quad (30)$$

The basic network computation steps of the S_{cut} neural-based models are as follows:

i. At the input layer, input variables (S , P_{wh} , P_r , γ_{API} and q_o) from the input neurons multiply with input weights (j_1 , j_2 , j_3 , j_4 and j_5), respectively, and are linked to hidden layer neurons;

ii. At the upper hidden layer neuron (i.e. i = 1), the input (i.e. $Sj_1 + P_{wh}j_2 + P_r j_3 + \gamma_{API} j_4 + q_o j_5$) from the input layer combined with the upper neuron's bias (b_i) and the sum (i.e. $\sum_{i=1}\left(Sj_1 + P_{wh}j_2 + P_r j_3 + \gamma_{API} j_4 + q_o j_5\right) + b_i$) is transformed by the sigmoid function (Equation 31), to the output neuron;

$$\sigma(z_i) = \frac{2}{1+e^{-2(\omega_i)}} - 1 \quad (31)$$

where ω_i is $\sum_{i=1}\left(Sj_1 + P_{wh}j_2 + P_r j_3 + \gamma_{API} j_4 + q_o j_5\right) + b_i$

iii. The transformed output from the upper hidden neuron (i.e. $\sigma(z_i)$) multiplied with the upper hidden neuron weight (Lw_i) and linked to the output neuron in the output layer;

iv. At the output neuron, the output from the hidden layer combined with the output neuron's bias (b_k), thus, $\left(\sigma(z_i) \times Lw_i\right) + b_k$

v. Steps (i) through (iv) are repeated for values of i = 2 and 3 for the neurons, and at the output neuron, the sum $\sum_{j=1}^{5}\sum_{i=1}^{3}\left[\left(\sigma(z_{ij}) \times Lw_{ij}\right) + b_{kj}\right]$ is transformed using the *purelin* function as the network's output. Thus, the predicted values are $purelin\sum_{j=1}^{5}\sum_{i=1}^{6}\left[\left(\sigma(z_{ij}) \times Lw_{ij}\right) + b_{kj}\right]$.

The values for the variables j_1 , j_2 , j_3 , j_4 , j_5 , b_i , Lw_{ij} , and b_{kj} are in Tables 6 and 8 for MISO and MIMO-based neural networks, respectively. The output (i.e. S_{cut}) from the neural network is presented in the normalised form in Equations 5 and 6, which would require de-normalisation to transform the network predictions to the required S_{cut} values. Thus, the explicit neural network-based model for sand-cut (S_{cut}) prediction is in Equation 21.

3.5 Comparison of the MISO and MIMO Predictions Performance

The statistical indices of the multiple-inputs single-output (MISO) and multiple-inputs multiple-outputs (MIMO) neural networks predictions performance are in Table 12. From the table, the MISO neural network predicted BS&W resulted in R^2, R, MSE, RMSE, ARE and AARE of 0.99980, 0.99990, 0.15583, 0.39475, -0.01949 and 0.29544, respectively, while the MIMO network had R^2 of 0.99873, R of 0.99937, MSE of 0.98551, RMSE of 0.99273, ARE of -0.09979 and AARE of 0.72406. On the other hand, the S_{cut} neural network predictions for the MISO and MIMO resulted in R^2, R, MSE, RMSE, ARE and AARE of 0.9990, 0.99950, 0.01739, 0.13187, -0.00109 and 0.08227, respectively, for the MISO network while MIMO network predictions resulted in R^2 of 0.99319, R of 0.99659, MSE of 0.11894, RMSE of 0.34488, ARE of 0.00111 and AARE of 0.19186.

The statistical indices from the MISO and MIMO neural networks predictions indicate that their predictions and the fields test BS&W and S_{cut} are very close. Again, the closeness of the MISO and MIMO neural network predictions is further depicted in the cross plots in Fig. 11 and 12, as the predicted and field test datasets aligned along the diagonal trend in the figures. This observation or results implies that either of the neural networks, MISO or MIMO, can be applied to predict these field parameters (i.e. BS&W and sand-cut) in the Niger Delta oilfields.

Table 12. Performance indices of the MISO and MIMO neural-based models

	Model	Statistical Indices					
		R^2	R	MSE	RMSE	ARE	AARE
BS&W	MISO	0.99980	0.99990	0.15583	0.39475	-0.01949	0.29544
	MIMO	0.99873	0.99937	0.98551	0.99273	-0.09979	0.72406
S_{cut}	MISO	0.99900	0.99950	0.01739	0.13187	-0.00109	0.08227
	MIMO	0.99319	0.99659	0.11894	0.34488	0.00111	0.19186

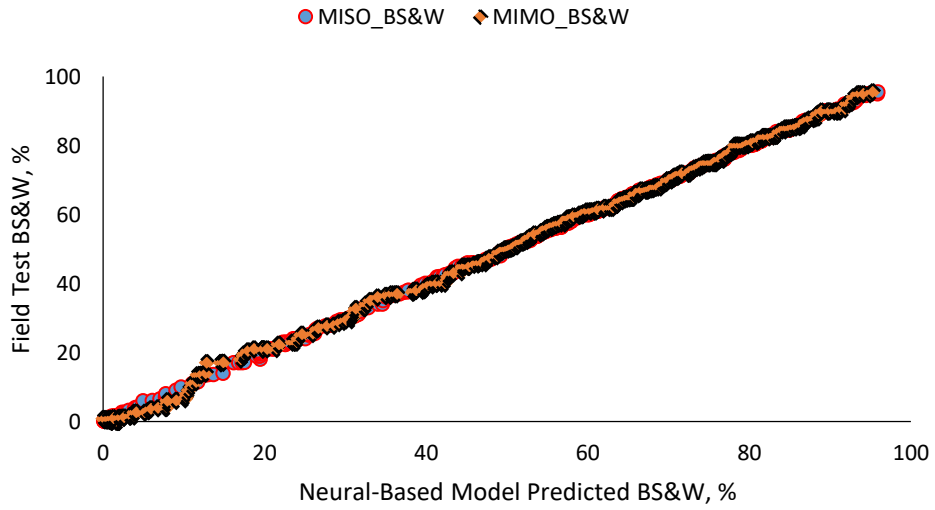

Fig. 11. Comparison of MISO and MIMO predicted BS&W with field datasets

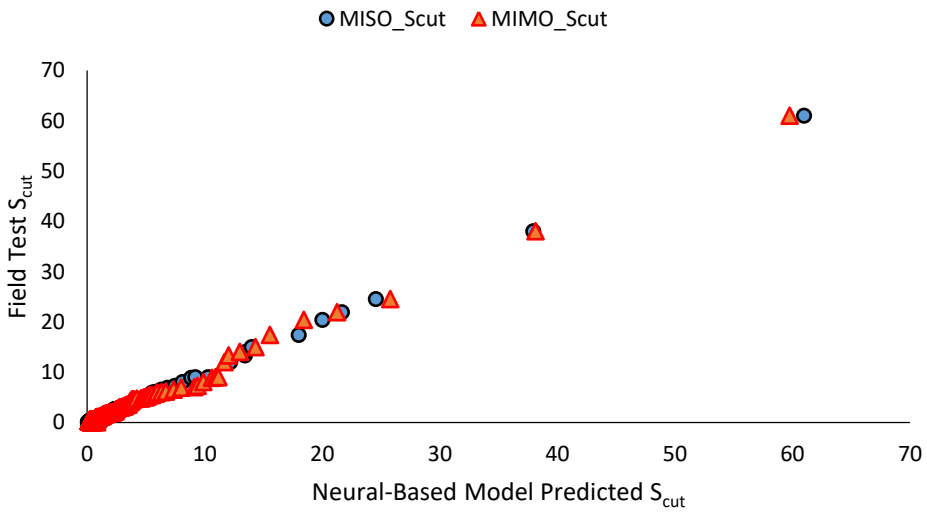

Fig. 12. Comparison of MISO and MIMO predicted S_{cut} with field datasets

Table 13. Statistical description of the input datasets for the models' generalisation test

Parameters	Maximum	Minimum	Range	Average	Std. Dev.
Bean (Choke) Size, /64 inch	36.0	12.0	24.0	22.50	5.694
Wellhead Pressure, P_{wh}, psia	3000.0	90.0	2910.0	762.31	634.12
Reservoir Pressure, P_r, psia	4863.0	2170.0	2693.0	3474.21	686.05
Oil Gravity, γ_{API}	57.77	16.4	41.37	37.51	13.057
Oil Flow Rate, q_o, stb/d	2594.0	29.6	2564.4	748.24	582.045

3.6 Generalisation Performance of the Developed Neural Network-Based Models

As earlier alluded, the models' generalisation performance is the available yardstick to assess their robustness to predict BS&W and S_{cut} from matured fields in the Niger Delta. Table 13 depicts the statistical description of the input variables datasets used to perform the models' generalisation assessment. The statistical indices of the developed multiple-inputs single-output (MISO) and multiple-inputs multiple-outputs (MIMO) neural-based models generalisation predictions performance are in Table 14. From the table, the developed MISO neural-based model predicted BS&W (i.e. new fields test datasets) resulted in R^2, R, MSE, RMSE, ARE and AARE of 0.97406, 0.98695, 2.08143, 1.44272, -0.00638 and 0.28755, respectively. In contrast, the MIMO neural-based model had R^2 of 0.97317, R of 0.98650, MSE of 2.15293, RMSE of 1.46729, ARE of -0.00713 and AARE of 0.25064. Conversely, the S_{cut} neural-based model predictions (i.e. new S_{cut} test datasets) for the MISO and MIMO resulted in R^2, R, MSE, RMSE, ARE and AARE of 0.89558, 0.94635, 0.01736, 0.13177, 0.01338 and 0.01759, respectively, for the MISO neural-based model while MIMO neural-based model predictions resulted in R^2 of 0.87505, R of 0.93544, MSE of 0.02118, RMSE of 0.14554, ARE of -0.02280 and AARE of 0.02996.

The statistical indices from the developed MISO and MIMO neural-based model predictions indicate that their predictions and the new fields test BS&W and S_{cut} are very close. The assertion is observed in the correlation values (i.e. R-values) obtained for the various models. The implication is that the MISO neural-based model can predict BS&W and S_{cut} in the Niger Delta oilfields with 98.7% and 94.6% certainty, respectively, while MIMO neural-based model

would give 98.7% and 93.5%. Also, the closeness of the MISO and MIMO neural-based models' predictions is further depicted in the

regression plots in Fig. 13 and 14. The figures align the predicted and new field test datasets along the diagonal trend. These results imply that the neural-based models can be exploited to predict BS&W and the Niger Delta oilfields.

3.7 Conceptual Application of the Developed Neural-Based Models

Any developed model, be it empirical, analytical or numerical, is to solve some or specific problems in its field of application. Okon *et al*. [20] state that any developed model is only relevant if applicable. Therefore, the developed neural-based models would be useful for predicting basic sediment and water (BS&W) and sand-cut (i.e. sand volume) to evaluate the potential of reservoirs in the Niger Delta oilfields. It is worth mentioning that the available petroleum engineering software, like PIPESIM and PROSPER, calculate produced sand volume using the intrinsic properties of the reservoir rock. One limitation of this method is that the intrinsic properties of the reservoir are readily available for estimation. Therefore, the approach to apply machine learning using performance test analysis (PTA) datasets to determine the reservoir's BS&W and produced sand volume in a new direction.

The conceptual flowchart of the developed neural-based models for petroleum engineering software to estimate BS&W and S_{cut} using PTA datasets is presented in Fig.15. The software inputs would be the choke (beans) size (S), wellhead pressure (P_{wh}), reservoir pressure (P_r), oil gravity (γ_{API}) and oil flow (production) rate (q_o). It then normalises the input parameters (i.e. $(S)_n$, $(P_{wh})_n$, $(P_r)_n$, $(\gamma_{API})_n$ and $(q_o)_n$). The user chooses the neural network approach (i.e. MISO or MIMO) to apply in the estimation of BS&W and S_{cut}. The computer program assesses the appropriate model input weights and biases to compute β_i , ω_i and $\sigma(z_i)$ to determine $(BS\&W)_{ANN}$ and $(S_{cut})_{ANN}$ if the MIMO network approach were chosen, $(BS\&W)_{ANN}$ or $(S_{cut})_{ANN}$ would be estimated if the MISO network method was implemented. Finally, the estimated $(BS\&W)_{ANN}$ and $(S_{cut})_{ANN}$ are denormalised to establish the BS&W and S_{cut} values.

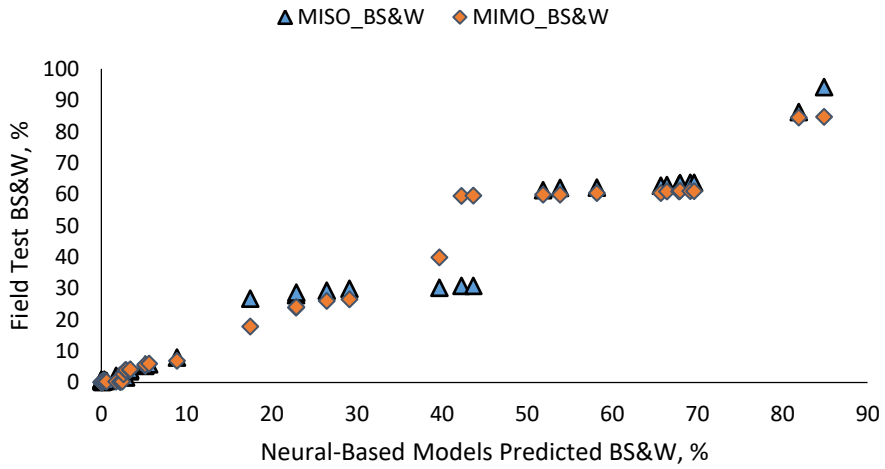

Fig. 13. Generalization Performance of MISO and MIMO models BS&W with field test datasets

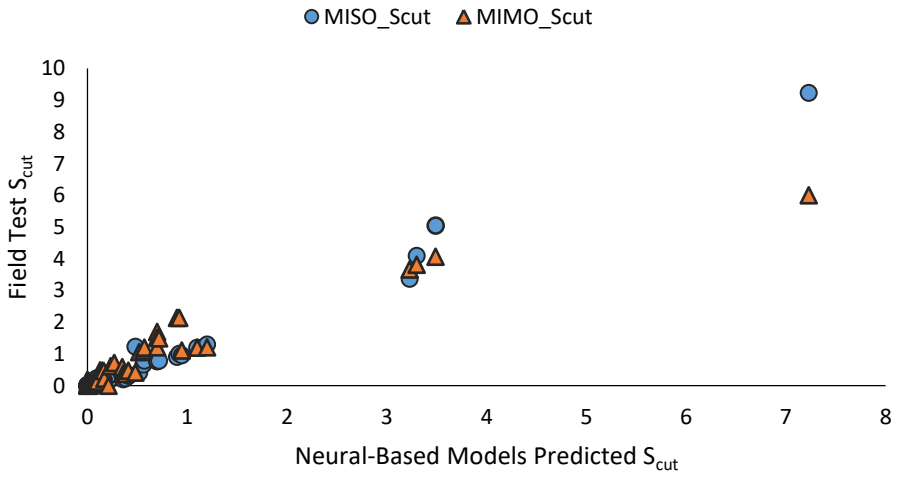

Fig. 14. Generalization Performance of MISO and MIMO models predicted S_{cut} with field test datasets

Table 14. Generalization Performance of the MISO and MIMO Neural-based Models

	Model	Statistical Indices					
		R^2	R	MSE	RMSE	ARE	AARE
BS&W	MISO	0.97406	0.98695	2.08143	1.44272	-0.00638	0.28755
	MIMO	0.97317	0.98650	2.15293	1.46729	-0.00713	0.25064
S_{cut}	MISO	0.89558	0.94635	0.01736	0.13177	0.01338	0.01759
	MIMO	0.87505	0.93544	0.02118	0.14554	-0.02280	0.02996

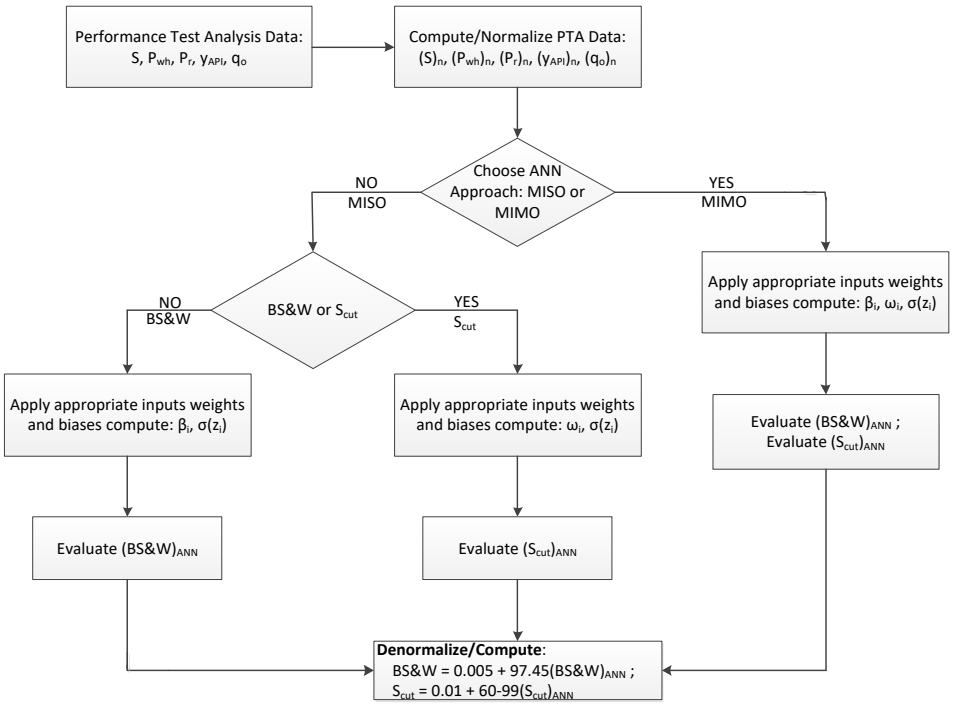

Fig. 15. Software application flowchart of the developed neural-based models

4. CONCLUSION

The following are the conclusions drawn from this study:

i. The multiple-inputs single-output neural networks predicted basic sediment and water (BS&W) and sand-cut (S_{cut}) were close to the fields test datasets with overall mean square error (MSE) and correlation coefficient (R) values of 2.0698×10^{-5} and 0.9999 for BS&W network, and 2.1529×10^{-6} and 0.9995 for S_{cut} network;

ii. The multiple-inputs multiple-outputs (MIMO) neural network predicted basic sediment and water (BS&W) and sand-cut (S_{cut}) agrees with the actual fields test datasets with overall MSE and R values of 7.5865×10^{-5} and 0.9997;

iii. The relative importance (RI) of the MISO networks input variables S, P_{wh}, P_r, γ_{API} and q_o resulted in 10.25%, 9.35%, 31.78%, 20.48% and 28.14%, respectively, for the BS&W neural network and 12.74%, 27.42%, 12.36%, 12.42% and 35.06% for the S_{cut} neural network. Thus, the network inputs RI ranking was $P_r > q_o > \gamma_{API} > S > P_{wh}$ for the BS&W network and $q_o > P_{wh} > S > \gamma_{API} > P_r$ for the S_{cut} network;

iv. Also, the RI for the MIMO neural network had 20.73% of S, 23.63% of P_{wh}, 36.50 of P_r, 13.16% of γ_{API} and 39.31% of q_o with RI ranking of $q_o > P_r > P_{wh} > S > \gamma_{API}$. The overall RI ranking for the MISO and MIMO neural networks to predict BS&W and S_{cut} is $q_o > P_r > P_{wh} > S > \gamma_{API}$;

v. the generalisation performance of the MISO neural-based models with the test datasets resulted in R^2, R, MSE, RMSE, ARE and AAPRE of 0.97406, 0.98695, 2.08143, 1.44272, -0.00638 and 0.28755, respectively, for BS&W neural-based model and R^2 of 0.89558, R of 0.93544, MSE of 0.01736, RMSE of 0.13177, ARE of 0.01338 and AARE of 0.01759 for S_{cut} neural-based model; and

vi. finally, the generalisation performance of the MIMO neural-based model with the fields test datasets resulted in R^2, R, MSE, RMSE, ARE and AAPRE of 0.97317, 0.98650, 2.15293, 1.46729, -0.00713 and 0.25064, respectively, for BS&W and R^2 of 0.87505, R of 0.93544, MSE of 0.02118, RMSE of 0.14554, ARE of -0.02280 and AARE of 0.02996 for S_{cut}.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. Okon AN, Olagunju DT, Akpabio JU. Water coning control: A comparison of downhole water sink and downhole water loop technologies. Journal of Scientific and Engineering Research, 2017;4(12):137-148
2. Waldschläger K, Brückner MZM, Almroth BC, Hackney CR, Adyel TM, Alimi OS, Belontz SL, Cowger W, Doyle D, Gray A, Kane I, Kooi M, Kramer M, Lechthaler S, Michie L, Nordam T, Pohl F, Russell C, Thit A, Umar W, Valero D, Varrani A, Warrier AK, Woodall LC, Wu N. Learning from natural sediments to tackle microplastics challenges: A multidisciplinary perspective. Earth-Science Reviews. 2022;228:104021. Available:https://doi.org/10.1016/j.earscirev.2022.104021.
3. Abatai MC, Akpabio JU, Okon AN, Etuk BR. Demulsification of crude oil emulsion in Well X in a Niger Delta field. Engineering and Applied Sciences. 2020;5(5):81-91.
4. Iradukunda P, Bwambale E, Ahsan A. Reservoir sedimentation and its effect on storage capacity – A case study of Murera reservoir, Kenya, Cogent Engineering. 2021;8:1. Available:https://doi.org/10.1080/23311916.2021.1917329.
5. Oyeneyin BM. Total sand management solution for guaranteed flow assurance in subsea development. Paper presented at the Society of Petroleum Engineers Nigeria Annual International Conference and Exhibition, Lagos, Nigeria; 2014.
6. Ikhumetse AA, Abioye OP, Ijah UJJ, Bankole MT. A critical review of oil spills in the Niger Delta aquatic environment: causes, impacts, and bioremediation assessment. Environmental Monitoring and Assessment. 2022;194:816. Available:https://doi.org/10.1007/s10661-022-10424-x
7. Salahi A, Dehghan AN, Sheikhzakariaee SJ, Davarpanah A. Sand production control mechanisms during oil well production and

construction. Petroleum Research. 2021;6(4):361-367.
8. Aroyehun ME, Oko FN, Onyeanusi O, Oguntade T, Kabara A, Dimkpa B. Comparative study of sand control methods in selected Niger-Delta sandstone reservoirs. Paper presented at the Society of Petroleum Engineers Nigeria Annual International Conference and Exhibition, Lagos; 2018.
Available:https://doi.org/10.2118/193526-MS
9. Osisanya SO. Practical guidelines for predicting sand production, in: Society of Petroleum Engineers - Nigeria Annual International Conference and Exhibition 2010, NAICE; 2010
Available:https://doi.org/10.2118/136980-ms
10. Chin LY, Ramos GG. Predicting volumetric sand production in weak reservoirs. Paper presented at the SPE/ISRM Rock Mechanics Conference, Irving, Texas; 2002.
11. Aborisade OM. Practical approach to effective sand prediction, control and management. MSc. Thesis, African University of Science and Technology, Nigeria; 2011
12. Bellarby, J. (2009). *Well completion design*. Elsevier's Publications, Oxford, U.K.
13. Christiana AH, Dosunmu A, Ogbonna J. Numerical sand prediction model analysis for sand onset, sand volume and sanding rate. International Journal of Scientific and Research Publications. 2019;9(8): 129-134.
14. Eyvazov J, Guliyeva M, Guliyev U. The effect of sand production to the well drainage area. In: Lin, J. (eds) Proceedings of the 2022 International Petroleum and Petrochemical Technology Conference, IPPTC 2022. Springer, Singapore; 2023.
Available:https://doi.org/10.1007/978-981-99-2649-7_14
15. Khamehchi E, Kivi IR. Akbari MA novel approach to sand production prediction using artificial intelligence. Journal of Petroleum Science and Engineering. 2014;123:147–154.
Available:httsps/doi.org/10.1016/j.petrol.2014.07.033.
16. Tugwell KW, Livinus A. Predictive models for oil in place for oil rim reservoirs in the Niger Delta using machine learning approach. Petroleum & Petrochemical Engineering Journal. 2023;7(3):000361.
17. Khare M, Nagendra SMS. Artificial neural networks in vehicular pollution modelling. Springer Berlin, Heidelberg; 2007.
Available:https://doi.org/10.1007/978-3-540-37418-3
18. Lawson D, Marion G. An introduction to mathematical modeling; 2008.
Available:https://people.maths.bris.ac.uk/~madjl/course_text.pdf
Accessed: 25th June 2019
19. Okon AN, Effiong AJ, Daniel DD. Explicit neural network-based models for bubble point pressure and formation volume factor prediction. Arabian Journal for Science and Engineering. 2023;48: 9221-9257.
20. Al-Bulushi N, King PR, Blunt MJ. Kraaijveld M. Development of artificial neural network models for predicting water saturation and fluid distribution. Journal of Petroleum Science and Engineering. 2009;68:197-208.
21. Okon AN, Adewole SE, Uguma EM. Artificial neural network model for reservoir petrophysical properties: Porosity, permeability and water saturation prediction. Modeling Earth Systems and Environment, 2020;7(4):2373-2390.
22. Effiong AJ, Etim JO, Okon AN. Artificial intelligence model for predicting formation damage in oil and gas wells. Paper presented at the Nigeria Annual International Conference and Exhibition, Lagos, Nigeria; 2021.
23. Okon AN, Ansa IB. Artificial neural network models for reservoir-aquifer dimensionless variables: influx and pressure prediction for water influx calculation. Journal of Petroleum Exploration and Production. 2021;11(4):1885-1904.

Peer-review history:
The peer review history for this paper can be accessed here:
https://www.sdiarticle5.com/review-history/107598