Petroleum Research 8 (2023) 118–129

Contents lists available at ScienceDirect

Petroleum Research

journal homepage: http://www.keaipublishing.com/en/journals/petroleum-research/

Predicting the effects of selected reservoir petrophysical properties on bottomhole pressure via three computational intelligence techniques

Emmanuel E. Okoro [a, d, *], Samuel E. Sanni [b], Tamunotonjo Obomanu [c], Paul Igbinedion [d]

[a] Department of Petroleum Engineering, University of Port Harcourt, Rivers State, Nigeria
[b] Department of Chemical Engineering, Covenant University, Ota, Nigeria
[c] Department of Petroleum Engineering, Federal Polytechnic of Oil and Gas, Bonny Island, Nigeria
[d] Department of Petroleum Engineering, Covenant University, Ota, Nigeria

ARTICLE INFO

Article history:
Received 20 December 2021
Received in revised form
1 July 2022
Accepted 5 July 2022
Available online 7 July 2022

Keywords:
Computational intelligence
Bottomhole pressure
Petrophysical properties
Heuristic search optimizer
Volvo field data

ABSTRACT

This study investigates the effects of selected petrophysical properties on predicting flowing well bottomhole pressure. To efficiently situate the essence of this investigation, genetic, imperialist competitive and whale optimization algorithms were used in predicting the bottomhole pressure of a reservoir using production data and some selected petrophysical properties as independent input variables. A total of 15,633 data sets were collected from Volvo field in Norway, and after screening the data, a total of 9161 data sets were used to develop apt computational intelligence models. The data were randomly divided into three different groups: training, validation, and testing data. Two case scenarios were considered in this study. The first scenario involved the prediction of flowing bottomhole pressure using only eleven independent variables, while the second scenario bothered on the prediction of the same flowing bottomhole pressure using the same independent variables and two selected petrophysical properties (porosity and permeability). Each of the two scenarios involved as implied in the first scenario, the use of three (3) heuristic search optimizers to determine optimal model architectures. The optimizers were allowed to choose the optimal number of layers (between 1 and 10), the optimal number of nodal points (between 10 and 100) for each layer and the optimal learning rate required per task/operation. the results, showed that the models were able to learn the problems well with the learning rate fixed from 0.001 to 0.0001, although this became successively slower as the leaning rate decreased. With the chosen model configuration, the results suggest that a moderate learning rate of 0.0001 results in good model performance on the trained and tested data sets. Comparing the three heuristic search optimizers based on minimum MSE, RMSE, MAE and highest coefficient of determination (R^2) for the actual and predicted values, shows that the imperialist competitive algorithm optimizer predicted the flowing bottomhole pressure most accurately relative to the genetic and whale optimization algorithm optimizers.

1. Introduction

Petrophysics is another broad and important field that incorporates engineering and earth science for reservoir characterization and field development. Rock physics is the science of studying the basic chemical and physical properties of porous media, especially reservoir rocks and those fluids found in them (Bai et al., 2021). These include accumulation and flow properties (fractional flow, porosity, and permeability), fluid identification, liquid phase distribution in void space (saturation) and surface interaction forces between rocks and fluids in the pore spaces (Capillary pressure), pressure measurement, pressure conditions, fluid conductivity, and so on (El Sharawy and Gaafar, 2019). These properties and conditions are used to identify and evaluate hydrocarbon reservoirs, host rocks, aquifers and cap rocks (Zheng et al., 2020). Petrophysical properties of any rock, are a set of basic technical parameters, and basic tools in quantifying reservoirs to help petroleum engineers evaluate hydrocarbon reserves, plan oil and gas well completions, and optimize production operations; it continues to be the primary tool for providing reliable and

* Corresponding author. Department of Petroleum Engineering, University of Port Harcourt, Rivers State, Nigeria.
E-mail address: emeka.okoro@covenantuniversity.edu.ng (E.E. Okoro).

https://doi.org/10.1016/j.ptlrs.2022.07.001

trustworthy information about a reservoir (Salimidelshad et al., 2019).

A reliable and accurate method of predicting the pressure drop of multi-phase vertical flow is essential for good completion and proper design of artificial lifting systems and for the optimization and prediction of the production efficiency (Raghavan and Chin, 2004). Due to the complexity of multi-phase flows, empirical or semi-empirical correlations are often developed to predict pressure drop. Many correlations have been developed, most of which were obtained in the laboratory and will be inaccurate when scaled-up to field conditions (Yilmaz et al., 1994). Most researchers agreed that they could not find a single correlation that applied to all the variables with reasonable levels of accuracy. Traditional methods of modeling the flow of reservoirs in vertical and horizontal wells always assume constant reservoir petrophysical properties throughout the reservoir's life. However, when the reservoir is being depleted, the geological rocks expand and the petrophysical properties of the reservoir change (Nguyen et al., 2020).

Pressure drop is a natural phenomenon in the formation after hydrocarbon production has started. One of the important parameters for determining the potential output of wells is the permeability of the reservoir formation (Yasser et al., 2011). As production decreases, artificial lift techniques are commonly used in wells to maintain production before other expensive reservoir simulations such as enhanced oil recovery techniques are taken into account, such as enhanced oil recovery techniques. The basic principle of all artificial lift methods is the application of external energy to the wells so as to reduce the pressure in the flowing bottom hole or increase the drawdown pressure (Tabatabaie et al., 2015). This high pressure drawdown can cause the formation to shrink and reduce its permeability. Therefore, in order to maximize production or recovery factor, it is important to find the optimal drawdown for a particular reservoir production system (Belyadi et al., 2016). However, the effective stress in a rock formation is computed/estimated by subtracting the pore fluid pressure from the overburden stress as defined by the effective stress equation. If the reservoir is depleted or exposed to high pressure drawdown during production, the pore pressure will decrease leading to an increase in the effective stress (Akai et al., 2016); this compresses the formation and reduces its permeability.

A relationship between porosity and pressure can be defined using compressibility factor. Compressibility is the relative change in volume of the rock with respect to changes in pressure. When production is initiated in a gas reservoir, the pore pressure reduces and as a result, the grain size increases depending upon the compressibility of the rock matrix. (Ayub and Ramadan, 2019). The combined effect of overburden stress and expansion of the rock matrix results in reduced pore size. From the Kozeny Carman equation, it is evident that permeability is a function of porosity and is directly proportional to porosity (Kommineni et al., 2017). Hence such reduction in porosity reduces the permeability of the reservoir with production. Several authors have used computational intelligence applications to predict bottomhole pressure due to its ability to handle huge streams of data generated in the field (Tariq et al., 2020). Some authors have highlighted the importance of the petrophysical properties of the reservoir formation in predicting the nonlinearity behavior of reservoir formations.However, the focus has always been on the prediction of bottomhole pressure from readings/measurements of downhole pressure gauges inside the well (Awadalla and Yousef, 2016; Chen et al., 2017).

Jahanandish et al. (2011) predicted the bottomhole flow and pressure drop in vertical multiphase flowing wells using an artificial neural network. Their study used nine input parameters (that is, oil rate, water rate, gas rate, GOR, pipe length, wellhead pressure, oil gravity “API”, surface temperature and bottomhole temperature) but, the field's petrophysical properties were not considered. Ebrahimi and Khamehchi (2015) used well-head and bottomhole pressure data measured by an Amerada gauge to train and develop a robust artificial neural network model for determining the vertical multiphase flowing well pressure drop. The comparison between the best obtained results and the results of the available methods indicated the higher accuracy of the proposed method; however, it did not consider the effects of petrophysical properties. Awadalla et al. (2016) used radial basis and feed-forward neural networks as artificial intelligence techniques for prediction of flowing bottom hole pressure using crude oil production data. these data include tubing head pressure, motor current, liquid production rate, oil production rate, water production rate, gas production, base sediment and water cut, formation gas oil ratio, oil specific gravity, produced water specific gravity, pump intake true vertical depth, and pump discharge pressure, again, petrophysical properties of the reservoir were not considered.

Chen et al. (2017) predicted the flowing bottomhole pressure of a well using support vector machine and random sample selection; although the study proposed a good prediction precision method from the production data used, it did not consider the influence of petrophysical properties. Amar et al. (2018) estimated bottomhole pressure in multiphase scenario using hybridization neural networks and grey wolves’ optimization. Despite feeding in ten parameters as inputs variables (oil flow rate, gas flow rate, water flow rate, oil gravity, gas gravity, depth in, inside pipe diameter, well head temperature, well head pressure, and the gas oil ratio) of the developed ANN models for the prediction of bottom-hole pressure (BHP) in vertical wells; the impact of petrophysical properties of the reservoir was not considered. Firouzi and Rathnayake (2019) used advanced data analytics to predict flowing bottomhole pressure and the data used for the study were from seven sensor production measurements generated within 15–18 months. Again, did not consider the impact or effect of petrophysical properties on the production data. This then informed the need to improve further the accuracy of such predictions for a more practical/economical design of wells and better optimization of production operations. Literature has shown that correlations are basically statistically derived global expressions with limited physical considerations, and thus do not render them a true physical optimization. However, mechanistic models outperform specific flow pattern predictions (Osman et al., 2005).

Knowing the bottom-hole pressure of hydrocarbon wells can help forecast the well's potential during the well's life cycle. In addition, mechanistic models and empirical correlations are practically limited by their specific assumptions, thus, they have failed to provide a satisfactory and reliable tool for estimating or predicting bottomhole pressure. Consequently, porosity and permeability are among the factors that indirectly affect the flowing bottomhole pressure. Therefore, this study investigates the effects of selected petrophysical properties on predicting the flowing well bottomhole pressure of a reservoir. The algorithms adopted in predicting the BHP from available data include the genetic, imperialist competitive and whale optimization algorithms, which is aimed at deducing the ideal computational intelligence model for predicting the flowing bottomhole pressure for field applications. The effectiveness of each intelligence model was identified through trend analysis while also comparing the predicted values with the true values and the average absolute error and the relative prediction error.

Table 1 shows the prior approaches and parameters considered in predicting bottomhole pressure from some selected published literature.

The present work differs from the similar ones in the literature by two main points: (1 unique heuristic optimization techniques

Table 1
Highlights on parameters considered in predicting Bottomhole Pressure.

Author, Year	Method Adopted	Parameters Considered	Remark
Paasche et al. (2011)	A regularized and adaptive Moving Horizon Estimation (MHE)	Flow rate, pressure, and choke parameter	Simulations employing real data deliver good results provided the observer model is carefully tuned or selected parameters are estimated
Li et al. (2012)	The study applied a proposed estimation and adaptive control scheme to the nonlinear ODE model.	Choke opening, bottomhole flow rate	With the estimated parameters used in the feedback path, the adaptive controller achieved uniform performance bounds for system's input and output signals.
Chen et al. (2017)	Support vector machine (SVM) and random samples selection	Production data	The FBHP-SVM method has a good performance because it can be applicable for all ranges of production data and conditions.
Spesivtsev et al. (2018)	Artificial neural network	Wellbore flowing parameters	Although average error increased, the model was capable of capturing the basic features of the noisy transient BHP behavior.
Okoro et al. (2021)	Artificial intelligence - Extra Tree and Feed Forward models	well data, general information/data, depths, hole size, and depths	The overall result shows that the proposed models can serve as a predictive tool for managing and handling of bottom hole pressure while conducting underbalanced drilling. These models gave a significant prediction of the possible spikes in actual bottom hole pressure which are often ignored in other models proposed in literature.

were used in optimizing the algorithm model and (2) effect of selected petrophysical properties were examined in the proposed model.

2. Methodology

2.1. Data input

In this study, stream hours, average downhole temperature, average tubing pressure differential, average annular pressure, average choke size, average well head pressure, pressure differential in chokes, bore oil volume, bore gas volume and bore water volume were used as the independent input variables for the prediction of the average downhole pressure using a Feed Forward Neural Network. Permeability and porosity were the selected petrophysical properties in addition to the independent variables; separately tested these two variables for their impact in predicting the flowing bottomhole pressure. The neural network used in this model is the multilayer perceptron neural network, whose node has a one-directional flow of decisions that move sequentially from input to output without going through loops or cycles. Depending on their structures, neural networks can have different numbers of hidden layers and nodes. Three heuristic optimization techniques were used to optimize the model so as to achieve the desired results.

2.2. Algorithm theory

2.2.1. Genetic algorithm (GA)

The algorithm reflects the process of natural selection, where the fittest individuals are selected for reproduction to produce the next generation of offspring. The genetic algorithm used for this neuron optimization problem, produces a set of possible solutions/architectural models thus, indicating the number of layers, the nodes in each layer, and the learning rate. These structures then recombine and mutate to reproduce new offsprings (structures), a process that is repeated in different generations. Fitness values are assigned to individual model structures based on training and validation loss, at a specified number of data points (Nooraeni et al., 2021). A ranking algorithm is used to determine the structure of the population/model, selected as a parent, to produce the next generation of offsprings. It then uses a crossover process to exchange information randomly between two models of structures with the same level number. Mutation operations can change the value of a randomly selected node, and the process is repeated until certain completion criteria are met (Amar et al., 2020).

2.2.2. Imperialist competitive algorithm (ICA)

The imperialist competition algorithm is a computational method used to solve various optimization problems without using functional scales in the optimization process. It is considered to be the social equivalent of the genetic algorithm. The imperialist competition algorithm simulates the competition process between empires in human society. The algorithm starts with an initial set of randomly generated possible architectural models called countries (indicating the number of layer, the nodes in each layers, and the learning rate). The cost of each country is calculated based on the loss caused by the training and validation of a certain number of data points in the architecture of each country. The country is then divided into imperialists and colonies. The imperialists are the best countries in terms of population and the colonies are the others left (Dhegihan et al., 2021).

The colonies are randomly distributed to the imperialists. The number of colonies that the imperialist receives is proportional to its power. The power of each imperialist is calculated and normalized according to its value. The empire consists of imperialists and its respective colonies.

The next process, called assimilation, takes place in each group of empires as the colony moves to imperialist positions and assumes certain aspects of imperialism. Then, there is a process called revolution in which specific colonies are randomly selected and replaced with new randomly created countries. In the context of assimilation and revolution, when the colonies are better than imperialists, the colonies and the imperialists will exchange roles (Moayedi et al., 2021). Then, in the next step, if all the empires try to take colonies from other empires, a war between the empires will take place.

First, the sum of each empire is calculated and normalized according to the following formula: Where T. C_n and N.T.C_n represent the sum and normalized sum of the nth empire, respectively, and ξ is a slightly positive number whose value defines the role of the colonies in determining the total cost of the empire;

$$T.C._n = Cost\ (imperialist_n) + \xi\ mean\ \{Cost\ (colonies\ of\ empire_n)\} \quad (1)$$

$$N.T.C._n = T.C._n - \max\ \{\ T.C._i\ \} \quad (2)$$

The weakest colony in the weakest empire is then picked out. Other empires try to succeed through competition. The success rate of each empire is given by equation (3),

$$P_{p_n} = \left| \frac{N.T.C._n}{\sum_{i=1}^{N_{imp}} N.T.C._i} \right| \tag{3}$$

Then, if the odds of success are different, the P vector defined in equation (4) is the probability of success. The vector R is created with the same size as P, which is a random number with evenly distributed elements. The vector D is then created by subtracting R from P. The empire with the largest corresponding index in D will be the specified colony at the end.

$$P = \left[P_{p_1}, P_{p_2}, P_{p_3}, \cdots, P_{p_{N_{imp}}} \right] \tag{4}$$

$$R = \left[r_1, r_2, r_3, \cdots, r_{N_{imp}} \right] where\ r_i \sim U(0,\ 1)\ and\ 1 \le i \le N_{imp} \tag{5}$$

$$\begin{aligned} D = P - R &= \left[D_1, D_2, D_3, \cdots, D_{n_{imp}} \right] \\ &= \left[P_{p_1} - r_1, P_{p_2} - r_2, P_{p_3} - r_3, \cdots P_{p_{N_{imp}}} - r_{N_{imp}} \right] \end{aligned} \tag{6}$$

The algorithm terminates when only one empire remains or when the specified maximum number of iterations is reached. ICA was dedicated to the continuous optimization problems (Hosseini and Al Khaled, 2014), but it is currently applied to many complex discrete combinatorial optimization problems such as uncertainties associated with hydrocarbon exploration and production operations. The direct and indirect relationship of some dominant variables in the subsurface during production can be attributed to colonialism, known as neocolonialism. The decision-making process is actualized in this application using scheduling theory which allocates resources over a period to optimize a collection of tasks to improve one or more objectives.

2.2.3. Whale optimization algorithm (WOA)

WOA is a swam-based algorithm that uses the bubble-net hunting maneuver technique of humpback whales to solve complex optimization problems. It has a simple structure, requires few control parameters, has a fast convergence rate and balances between exploration and exploitation stages (Kalananda and Komanapalli, 2021; Du et al., 2021). In the WOA, humpback whale populations seek prey in a multidimensional search space. The position of the humpback whale is expressed as several determinants, but the distance between the humpback whale and the bait corresponds to the objective cost. The time-dependent position of a roll is measured by three operational procedures:

i. **Shrinking encircling prey** - Humpback whales can identify and surround replacement sites (prey), but because they do not know in advance where the best model architecture is, WOA assumes that the current best possible architecture is the target prey (Zeng et al., 2021). It then determines the best search agent and updates its position next to the current best search agent. The function is expressed by equations (7)–(10). Where **X*** is the overall best position, **X** is the selected position, **t** is the last iteration, **a** is a linear reduction from 2 to 0 in the iteration, and **r** is a uniformly distributed random number in the range of [0, 1], and $\vec{A} \cdot \vec{D}$ is the position vector.

$$\vec{D} = \left| \vec{C} \cdot \vec{X^*}(t) - \vec{X}(t) \right| \tag{7}$$

$$\vec{X}(t+1) = \vec{X^*}(t) - \vec{A} \cdot \vec{D} \tag{8}$$

$$\vec{A} = 2\vec{a} \cdot \vec{r} - \vec{a} \tag{9}$$

$$\vec{C} = 2 \cdot \vec{r} \tag{10}$$

ii. **Bubble-net attacking method (exploitation phase) -** To formulate a humpback whale in the form of a bubble-net, a mathematical spiral equation is applied between the whale position and the prey of the humpback whale to simulate the helix-shaped motion of the whale.

ii. **Search for prey (exploration phase)** – The global Optimizer is used here, for instance, if A> 1 or A < −1, the search agent is updated according to random specifications in place of the optimal search agent.

The deep learning neural networks used in this research and heuristic search engine optimization tools used in conducting this investigation were created in Python using the Tensorflow library, which provides state-of-the-art features such as computing tensor and deep neural networks built on tape-based automatic differentiation systems.

The training algorithm used in this method is the gradient decent algorithm, which is an iterative first-order optimization algorithm for finding the minimum value of a variable function. The input variables used for training are: Well Depth, Average Downhole Pressure, Average Downhole Temperature, Average Tubing pressure differential, Average Annulus Pressure, Average Choke Size, Average Well-Head Pressure, Average Well-Head Temperature, Pressure Differential in Chokes, and Bore Oil Volume. The variables were obtained from the Norwegian field (Volvo Field). In the first case, the data sample was fed into the optimizers and the optimal model in the form of "ON_STREAM_HRS, AVG_DOWNHOLE_TEMPERATURE, AVG_DP_TUBING, AVG_ANNULUS_PRESS, AVG_CHOKE_SIZE_P, AVG_WHP_P, AVG_WHT_P, DP_CHOKE_SIZE, BORE_OIL_VOL, BORE_GAS_VOL, BORE_WAT_VOL" which represent a single array containing eleven (11) independent variables with [AVG_DOWNHOLE_PRESSURE] as the dependent output variable.

In the second case study, [PORO_PERM] was added to the independent variables, making the system an array of thirteen (13) independent variables fed to the optimizers and the optimal model to produce separate results. The data input minimum and maximum values of the input element parameters and their confidence intervals, assigned to the three heuristic optimization techniques are tabulated in Table 2. The minimum value in the data set is the smallest mathematical value in the data set, while the maximum value is the largest mathematical value in the data set. The outlier is a value in the dataset that lies outside the given data set. In statistical calculations, missing data and outliers for each variable tested are usually removed. The computed confidence interval shows the probable range of values that somewhat define the confines of the true value lies. The 80%, 90%, and 95% regions show the percentage of probability, or certainty, that the confidence interval contains the true parameter value when random samples are drawn many times. Table 2 shows that it is advisable to stay within/stick to predictions within the 95% confidence level owing to the wider consideration of data it provides.

The next most imperative step in this study's data analysis process was to determine the correlation map of each input parameter. Each square in Fig. 1 shows the correlation between the variables on each axis, and the correlation ranges from −1 to +1.

Table 2
Statistical analysis and data ranges for the input variables.

	AVG Downhole Pressure	Stream HRS	AVG Downhole Temperature	AVG DP Tubing	AVG Annulus Pressure	AVG Choke Size	AVG WHP P	AVG WHT P	DP Choke Size	BORE OIL VOL	BORE GAS VOL	BORE WAT VOL
Mean	174.73	21.31	74.03	151.85	15.21	58.46	43.80	71.94	16.27	1073.53	159014.18	1970.10
Std	113.25	6.97	47.11	78.90	8.16	37.56	21.92	25.08	20.18	1248.77	178066.63	1730.85
Min	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	457.84
25%	0.00	24.00	0.00	49.95	11.46	18.59	31.38	68.01	3.22	204.17	31379.66	119.49
50%	230.51	24.00	102.15	176.84	16.55	60.36	36.13	82.30	7.54	624.80	97038.12	1980.00
75%	254.77	24.00	106.23	206.45	21.41	100.00	52.59	88.56	21.58	1292.61	197499.79	3475.13
Max	397.59	25.00	107.51	345.91	30.02	100.00	137.31	93.51	124.12	5888.69	835981.33	8019.74
Confidence Interval (80%)	173–176	21.2 –21.4	73.3–74.7	151–153	15.1–15.3	57.9–59	43.5 –44.1	71.6 –72.3	16–16.6	1060 –1090	156,000 –162,000	1940 –2000
Confidence Interval (90%)	173–177	21.2 –21.4	73.1–74.9	150–153	15.1–15.4	57.7–59.2	43.4 –44.2	71.5 –72.4	15.9 –16.7	1050 –1100	156,000 –162,000	1940 –2000
Confidence Interval (95%)	172–177	21.2 –21.5	73–75.1	150–154	15–15.4	57.6–59.3	43.3 –44.3	71.4 –72.5	15.8 –16.7	1050 –1100	155,000 –163,000	1930 –2010

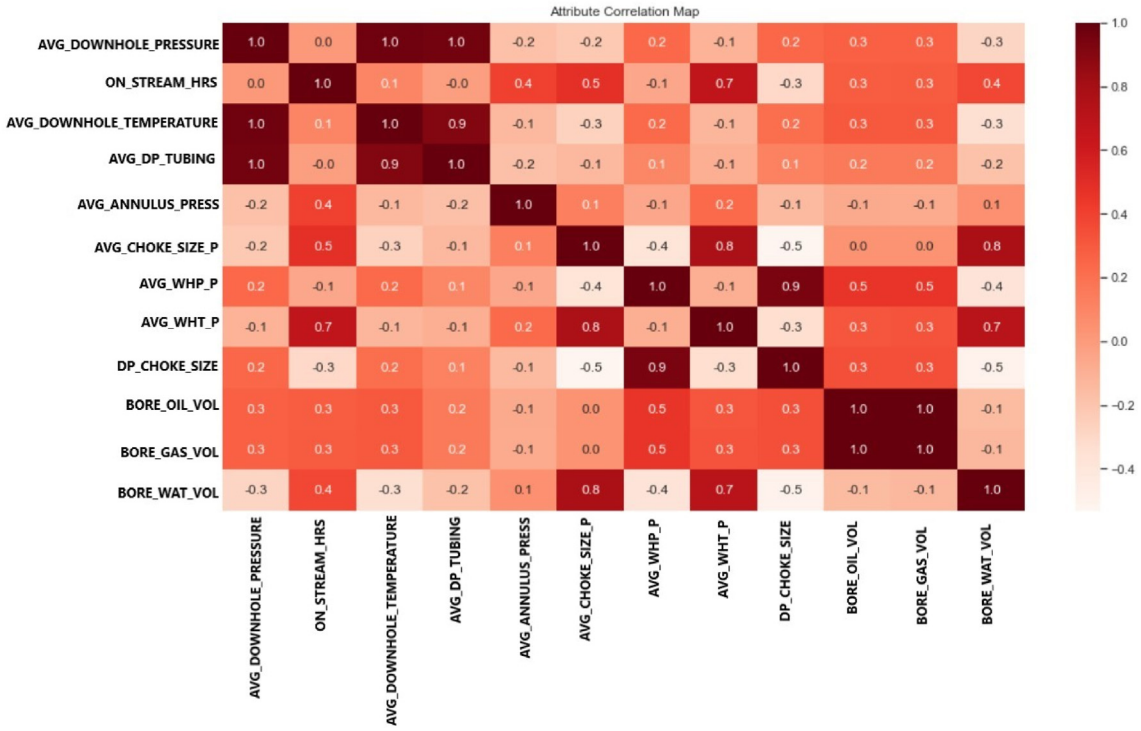

Fig. 1. Correlation map of input parameters.

Values closer to zero infer that no linear trend exists between the two variables. The closer a value is to +1 implies that the correlation is strongly related and more positively correlated; if one increases, so does the other. A correlation closer to −1 is similar but, implies that the parameters are inversely related; that is, as one of the variables increases, the other decreases. Among the parameters considered for this study, average downhole temperature/average tubing pressure differential (0.9), average tubing pressure differential/average downhole temperature (0.9), and average well-head pressure/pressure differential in chokes (0.9) have stronger relationships and are also positively correlated than the rest of the parameters. Although, the average choke size/bore water volume (0.8), average choke size/average well-head temperature (0.8), and on stream hours/average well-head temperature (0.7) also have a positive relationship and are positively correlated.

Figs. 2 and 3 show the graphical representation of information and correlate data in this study. These provide an accessible way to see and understand trends, outliers and patterns in the data used. In the world of Big Data, data visualization is indispensable for analyzing massive amounts of information to make data-driven decisions. Figs. 2 and 3 show the graphical representation of the most correlated variables from the correlation map analysis.

The probability density is the relationship between the observations and their probabilities which can be either high or low. The general form of probability density is called probability distribution, and the probability calculation of outcomes for a particular random variable is performed by the probability density function (PDF). It is helpful to know the probability density function of a data sample because, this will inform whether a particular observation is unlikely or very unlikely to be considered an anomaly or outlier, and if it should be removed. It also helps in choosing the right training method that requires a specific probability distribution for the distribution. It is unlikely that the probability density function of random data is known. Therefore, probability density must be estimated using a procedure called probability density estimation. Fig. 4 shows the Kernel density estimation of the variables based on

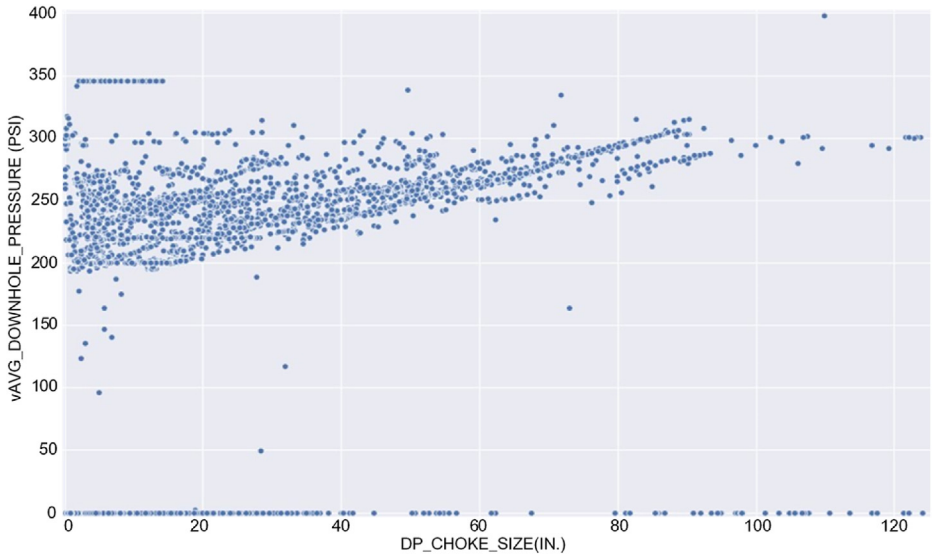

Fig. 2. Choke size vs downhole pressure.

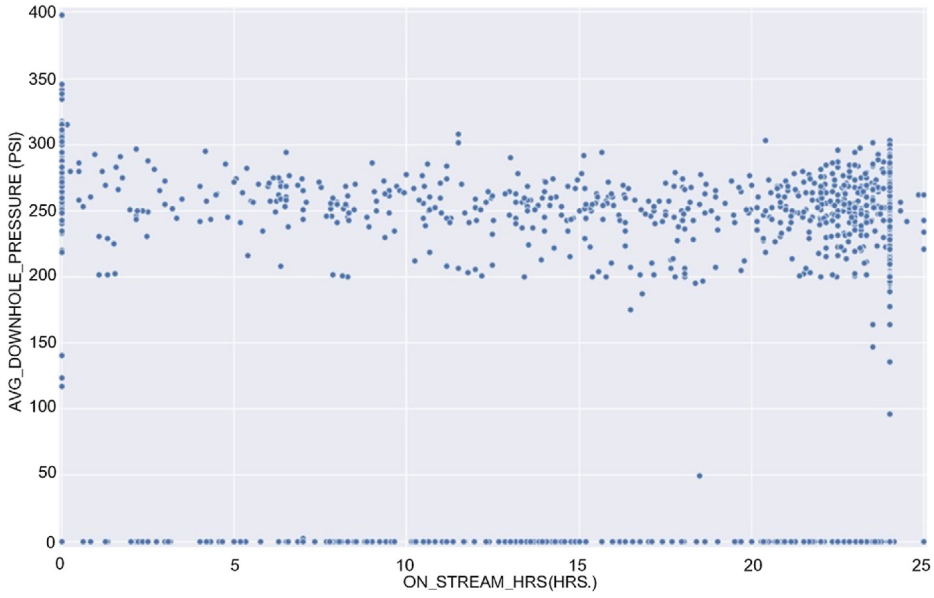

Fig. 3. Time on-stream vs downhole pressure.

the finite data samples. It is a useful non-parametric technique for visualizing the underlying distribution of the variables, and this approach was adopted to estimate the input data class-conditional marginal densities. This approach is unique because, it does not assume any underlying distribution for the variables analyzed. Fig. 4 shows that the input variables have different values of bandwidth. The kernel function evaluates the contribution of observations from a sample of data based on the relation or distance to a specific query data for which probability value is required.

2.3. Data selection and pre-processing

A total of 15,633 data sets were collected from Volvo fields in Norway. To check the validity of the collected data and identification of outliers, empirical correlations, and mechanistic models were used to predict the bottomhole flowing pressures and compared with the measured value. The approach by Jahanandish et al. (2011) was adopted for data quality assurance and control. Data sets which consistently resulted in poor predictions were considered to be invalid and, therefore, removed. After such a screening, 9161 data sets were used to develop the artificial neural network model. These were randomly divided into three groups: training, validation, and testing. Different partitioning ratios were tested (2:1:1, 3:1:1, and 4:1:1), however, the 4:1:1 partitioning rule yielded better training and testing results. The training set is used to develop and adjust the weights in a network; and the validation set is used to ensure the generalization of the developed network during the training phase. The testing set is used to examine the final performance of the network and compare the model performance with field data. Test data do not affect training, so they provide an independent measure of network performance during and after training.

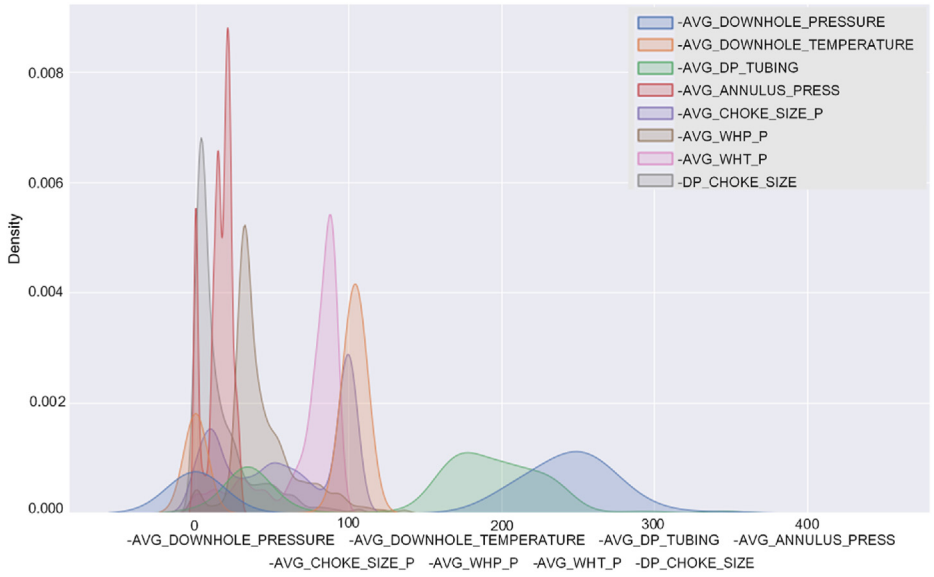

Fig. 4. Input variables kernel density estimation.

Fig. 5 shows the workflow for applying the heuristic search optimizer in developing the model and predicting the results.

3. Results and discussion

Two case scenarios were considered in this study. The first scenario involved the prediction of flowing bottomhole pressure using only eleven production independent variables. In comparison, the second scenario involved two additional variables, that is, two selected petrophysical properties (porosity and permeability). The two scenarios involve using three (3) heuristic search optimizers to determine the best optimal model architecture. The optimizers were allowed to choose the optimal number of layers (between 1 and 10), the optimal number of nodal points (between 10 and 100) for each layer and the optimal learning rate. Each model is a feed forward neural network with its number of hidden layers and fully connected nodal points determined by the heuristic search optimizer. The model is compiled using the Adam optimizer and Mean Squared Error loss function at an optimal learning rate determined by the heuristic search optimizers. The fully connected neural network structure was designed to accept the independent input parameters and return the average downhole pressure.

After training and testing the models using specific parts of the data set, the expected results were compared with the actual measurements taken from the field (Table 3). The logical network topology/configuration describes how nodes and links are set up in relation to each other (Ghimire and Mohan, 2017). The applied layout of the topology was selected to increase performance and effectively allocate resources across the network, so as to improve data efficiency and energy. When predicting the flowing bottomhole pressure for the two scenarios with the heuristic methods, multiple combinations of model parameters were considered, and as a trade-off between algorithmic speed and result precision, Table 2 shows the optimized logical network topology applied in this study.

The amount by which weights are updated during training is called “learning rate”. Specifically, the learning rate is a regulated hyperparameter that is often used to train neural networks with small positive values in the range 0.0–1.0. Learning rate determines how quickly the model adapts to the problem. The slower the learning rate, the more some learning, more time is required and the less weight changes per update. On the other hand, the faster the learning rate, the faster the changes and the lesser the time of learning needed. If the training speed is too fast, the model will converge to an optimal solution very quickly, and if the training speed is too slow, the process may freeze (Scutari et al., 2019). From the results (Table 3), it was observed that the models could learn the problems well with the learning rates of 0.001 and 0.0001 for

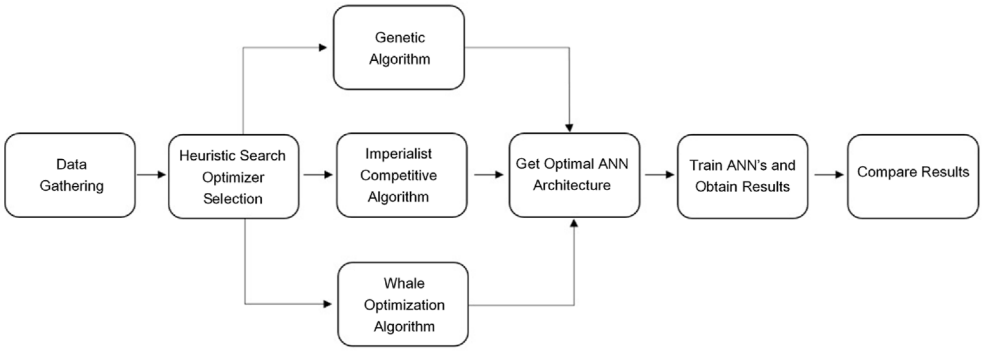

Fig. 5. Workflow of this study.

Table 3
Accuracy of the four heuristic search optimizers as validated with field data.

Study	Heuristic Search Optimizer	Network Topology	Learning Rate	R^2 score	MSE	RMSE	MAE	Variance Score
Study 1	GA	n-70-10-50-80-60-10-90-40-100-10	0.0001	0.9979	0.2973	0.3135	0.0328	0.9988
Study 1	ICA	n-75-25-50-25-100-25-50-25-25-25	0.001	0.9985	0.1190	0.2265	0.0207	0.9987
Study 1	WOA	n-10-10-20-20-10-10-20-10-10-40	0.001	0.9975	0.3003	0.3831	0.0334	0.9981
Study 2	GA	n-10-20-10-10-10-10-10-20	0.001	0.9984	0.2315	0.2573	0.0207	0.9985
Study 2	ICA	n-75-50-100-100-50-100-50	0.0001	0.9989	0.1347	0.1653	0.0191	0.9991
Study 2	WOA	n-60-60-60-50-10-60-60-20-20	0.001	0.9927	0.9018	0.7799	0.0205	0.9931

both cases, although the process became successively slower as the leaning rate decreased. With the chosen model configuration, the results suggest that a moderate learning rate 0.0001 results in good model performance on the trained and test data sets. The challenge of deep learning neural network training is to choose the speed of learning carefully and this is probably the most important parameter of the model (Malik and Yadav, 2021).

For scenarios 1 and 2, considering the coefficient of determination (R^2), GA heuristic search optimizer predicted the flowing bottomhole pressure (FBHP) with R^2 value range of 0.9979–0.9984, ICA with R^2 range of 0.9985–0.9989, and WOA with R^2 value of 0.9975 to 0.9927 respectively. When the petrophysical properties were considered for the second scenario, there was a slight increment in the coefficient of determination for the first two heuristic search optimizers (GA and ICA). But, a decrease in the coefficient of determination was observed for the WOA heuristic search optimizer for the second scenario (Table 3). The mean square error (MSE), root mean squared error (RMSE) and mean absolute error (MAE) were used to estimate the difference between the predicted value and the true field value. The result shows that out of the three heuristic search optimizers used for the two scenarios, ICA had the least MSE values (that is, 0.1190 and 0.1347), and this implies that it is closest to the line of best fit for the two scenarios when compare to those of GA (0.2973 and 0.2315) and WOA (0.3003 and 0.9018) heuristic search optimizers respectively. Since MSE is a risk function corresponding to the expected value of the squared error loss, it can be deduced that the ICA heuristic search optimizer has a lower risk in predicting the FBHP when compared to GA and WOA heuristic search optimizer.

The RMSE measures the error of the three heuristic search optimizers in predicting the FBHP field data. Table 3 shows that the ICA RSMES next prediction of FBHP for the two scenarios will be 0.2265 and 0.1653, respectively. For GA, the model's next prediction will be 0.3135 and 0.2573, the actual values for studies 1 and 2, respectively, while that of WOA will be 0.3831 and 0.7799 for both scenarios, respectively. The MAE result (Table 3) shows that the expected errors from the ICA heuristic search optimizer are lower (that is, 0.0207 and 0.0191) when compared with those of GA (0.0328 and 0.0207) and WOA (0.0334 and 0.0205) heuristic search optimizers for both scenarios respectively. The variance score was estimated because the mean (error) is not equal to zero. Yet, the ICA heuristic search optimizer gave the highest percentage of variance, thus indicating the strongest strength of association for the two scenarios. A comparison between these three heuristic search optimizers based on minimum MSE, RMSE, MAE and highest coefficient of determination (R^2) between actual and predicted values, shows that the ICA heuristic search optimizer predicted the flowing bottomhole pressure most accurately compared to the GA and WOA heuristic search optimizers for the continuous variables considered in this study. Table 4 shows a comparative section for the optimizers calculability and parameters utilized for this study.

Figs. 6–8 show the cross-plots between actual and predicted flowing bottomhole pressures for the first scenario (without the two selected petrophysical properties as input variables). The GA, ICA, and WOA heuristic search optimizers gave 0.9982, 0.9988 and 0.9984 coefficients of determination, respectively. Akinsete and Adesiji (2019) also used the data set from Volvo field to predict bottomhole pressure using an artificial neural network. Their coefficient of determination was also 0.99, but compared to the other three intelligent models (support vector machines, decision tree, and random forest), only random forest gave a 0.99 coefficient of determination for the predicted bottomhole pressure. Ahmadi and Chen (2019) also used machine learning models to predict bottomhole pressure, and out of the five machine learning methods investigated in their study, only the hybrid meta-heuristic optimization method gave a 0.99 coefficient of determination.

Porosity and permeability are among the factors that indirectly affect the flowing bottomhole pressure of a reservoir. Most of the predictions of bottomhole pressure in the literature using the artificial neural network approach did not consider petrophysical properties as independent variables while predicting the flowing bottomhole pressure (Okoro et al., 2021). Wang et al. (2018) highlight permeability dependence on stresses and that permeability is pore pressure dependent. Espinoza et al. (2014) also noted that, instead of the variation of horizontal effective stress with permeability, the direct/indirect influence of petrophysical properties in predicting BHP need be established. Figs. 9 -11 show the cross-plots for the actual and predicted flowing bottomhole pressure for the second scenario (selected petrophysical properties were added as input variables). The predicted BHP using the selected petrophysical properties (porosity and permeability) showed slight accuracy improvement when compared with the first scenario's values.

4. Conclusions

For optimum hydrocarbon field production optimization, accurate prediction of well flowing bottom-hole pressure is necessary for reduction in cost per barrel, and quantification of remedial workover operations. Mechanistic models and empirical correlations are practically limited by their specific assumptions. Thus, they have failed to provide a satisfactory and reliable tool for estimating or predicting bottomhole pressure. In addition, by this study, it can be asserted that:

(1) Porosity and permeability are factors that indirectly affect the flowing bottomhole pressure. In current literature, most of the predictions of bottomhole pressure via the artificial neural network, did not consider petrophysical properties as one/more independent variables while predicting the flowing bottomhole pressure of a well.
(2) This study adequately justifies the effect of porosity and permeability in predicting flowing well bottomhole pressure.
(3) The result shows that out of the three heuristic search optimizers used for the two scenarios, ICA had the least MSE values (0.1190 and 0.1347), and this implies that it is closest to the line of best fit for the two scenarios when compared to the values obtained for the GA (0.2973 and 0.2315) and WOA (0.3003 and 0.9018) heuristic search optimizers respectively.

Table 4
CPU and Parameters used by the optimizing algorithms.

Optimizing Algorithms	Control Parameters	Average computational time (s)
Genetic Algorithm	Population size, number of parents, mutation rate, number of generations.	884
Imperialist Competitive Algorithm	Initial number of countries, initial number of empires, assimilation rate, revolution rate, alpha rate, revolution probability.	485
Whale Optimization Algorithm	Number of generations, a, b	478

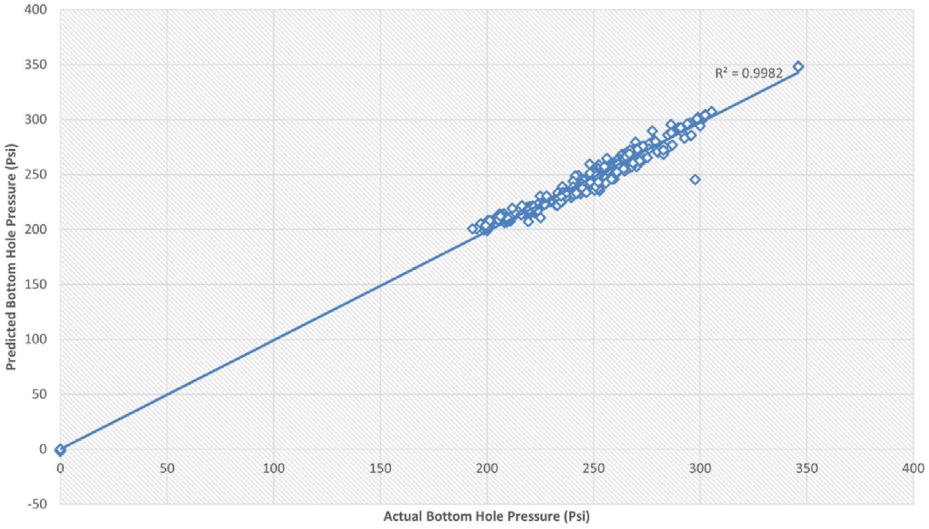

Fig. 6. Actual and predicted BHPs using GA optimal model for first scenario.

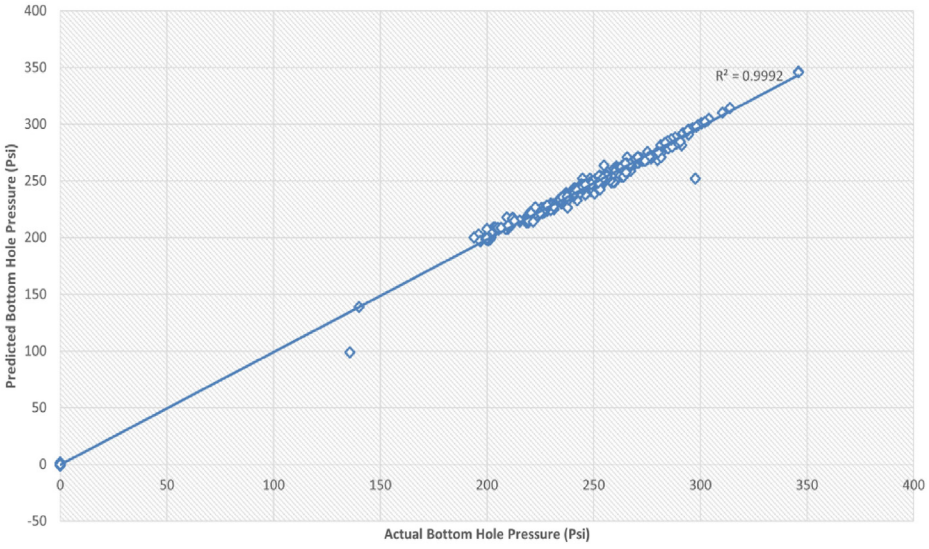

Fig. 7. Actual and predicted BHPs using ICA optimal model for first scenario.

(4) Since MSE is a risk function corresponding to the expected value of the squared error loss, it can be deduced that the application of ICA heuristic search optimizer has a lower risk in predicting the FBHP when compared to GA and WOA heuristic search optimizers.
(5) A comparison among these three heuristic search optimizers shows that the minimum MSE, RMSE, MAE and the highest coefficient of determination (R^2) between the actual and predicted values, shows that the ICA heuristic search optimizer predicted the flowing bottomhole pressure most accurately when compared to GA and WOA heuristic search optimizers for the continuous variables considered in this study.
(6) Based on the findings, the predicted BHP was influenced by the selected petrophysical properties (porosity and permeability) owing to the slight improvement in accuracy obtained for case scenario 2 relative to 1 for the test data.

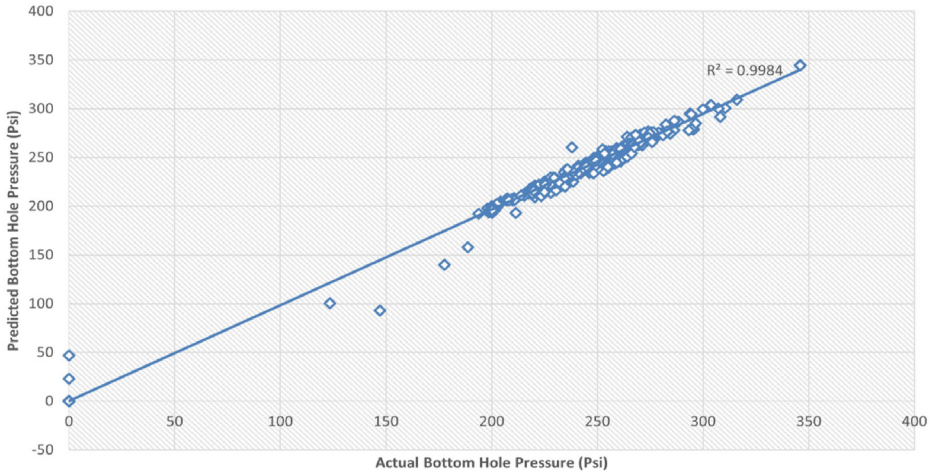

Fig. 8. Actual and BHPs using WOA optimal model for first scenario.

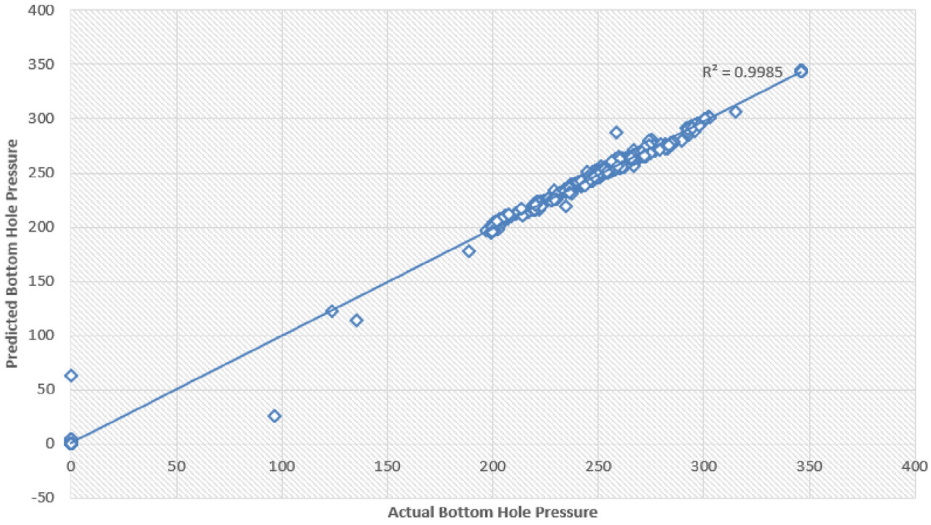

Fig. 9. Actual and predicted BHPs using GA optimal model for second scenario.

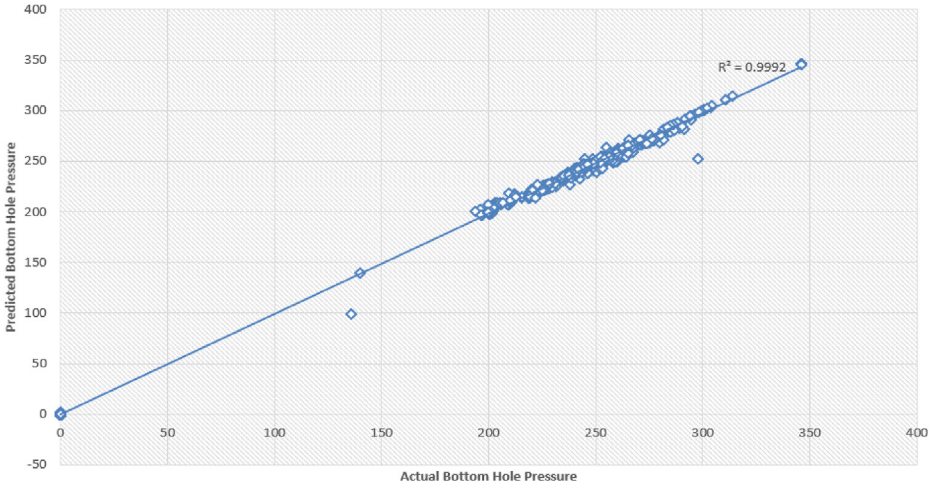

Fig. 10. Actual and predicted BHPs using ICA optimal model for second scenario.

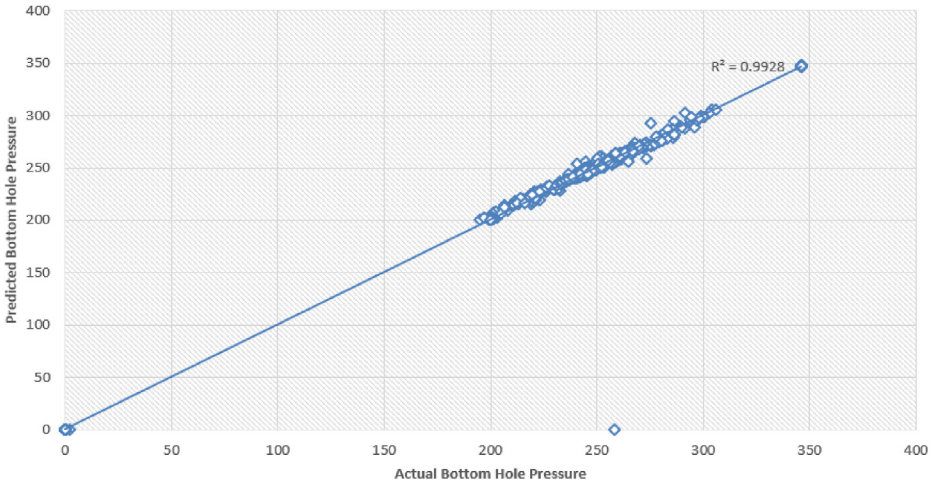

Fig. 11. Actual and BHPs using WOA optimal model for second scenario.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Ahmadi, M.A., Chen, Z., 2019. Machine learning models to predict bottom hole pressure in multi-phase flow in vertical oil production wells. Can. J. Chem. Eng. 97 (11), 2928–2940.
Akai, T., Takakuwa, Y., Sato, K., 2016. Pressure Dependent Permeability of Tight Rocks. SPE Low Perm Symposium, Denver, Colorado, USA.
Akinsete, O., Adesiji, B.A., 2019. Bottom-hole pressure estimation from wellhead data using artificial neural network. In: SPE Nigeria Annual International Conference and Exhibition. SPE.
Amar, M.N., Zeraibi, N., Redouane, K., 2018. Bottom hole pressure estimation using hybridization neural networks and grey wolves optimization. Petroleum 4, 419–429.
Amar, M.N., Zeraibi, N., Jahanbani Ghahfarokhi, A., 2020. Applying hybrid support vector regression and genetic algorithm to water alternating CO_2 gas EOR. Greenhouse Gases: Sci. Technol. 10 (3), 613–630.
Awadalla, M., Yousef, H., 2016. Neural networks for fow bottom hole pressure prediction. Int. J. Electr. Comput. Eng. 6, 1839.
Awadalla, M., Yousef, H., Al-Shidani, A., Al-Hinai, A., 2016. Artificial intelligent techniques for flow bottom hole pressure prediction. International Journal of Computers and Technology 15 (12), 7263–7283.
Ayub, M., Ramadan, M., 2019. Mitigation of near wellbore gas-condensate by CO_2 huff-n-puff injection: a simulation study. J. Petrol. Sci. Eng. 175, 998–1027.
Bai, Y., Zhao, J., Wu, W., 2021. Methods to determine the upper limits of petrophysical properties in tight oil reservoirs: examples from the Ordos and Songliao Basins. J. Petrol. Sci. Eng., 107983
Belyadi, H., Fathi, E., Belyadi, F., 2016. Managed Pressure Drawdown in Utica/point Pleasant with Case Studies. SPE Eastern Regional Meeting, Ohio, USA. Canton.
Chen, W., Di, Q., Ye, F., Zhang, J., Wang, W., 2017. Flowing bottomhole pressure prediction for gas wells based on support vector machine and random samples selection. Int. J. Hydrogen Energy 42, 18333–18342.
Dhegihan, M., Mashayekhi, M., Sharifi, M., 2021. An efficient imperialist competitive algorithm with likelihood assimilation for topology, shape and sizing optimization of truss structures. Appl. Math. Model. 93, 1–27.
Du, W., Zhang, Q., Chen, Y., Ye, Z., 2021. An urban short-term traffic flow prediction model based on wavelet neural network with improved whale optimization algorithm. Sustain. Cities Soc. 69, 102858.
Ebrahimi, A., Khamehchi, E., 2015. A robust model for computing pressure drop in vertical multiphase flow. J. Nat. Gas Sci. Eng. 26, 1306–1316.
El Sharawy, M.S., Gaafar, G.R., 2019. Impacts of petrophysical properties of sandstone reservoirs on their irreducible water saturation: implication and prediction. J. Afr. Earth Sci. 156, 118–132.
Espinoza, D.N., Vandamme, M., Dangla, P., Pereira, J.M., Vidal-Gilbert, S., 2014. Measurement and modeling of adsorptive-poromechanical properties of bituminous coal cores exposed to CO_2 adsorption, swelling strains, swelling stresses and impact on fracture permeability. Int. J. Coal Geol. 134–135, 80–95.
Firouzi, M., Rathnayake, S., 2019. Prediction of the Flowing Bottom-Hole Pressure Using Advanced Data Analytics. SPE/AAPG/SEG Asia Pacific Unconventional Resources Technology Conference, Brisbane, Australia.
Ghimire, R., Mohan, S., 2017. Logical topology optimization of free space optical networks with tracking transceivers. Opt. Switch. Netw. 24, 57–64.
Hosseini, S., Al Khaled, A., 2014. A survey on the imperialist competitive algorithm metaheuristic: implementation in engineering domain and directions for future research. Appl. Soft Comput. 24, 1078–1094.
Jahanandish, I., Salimifard, B., Jalalifar, H., 2011. Predicting bottomhole pressure in vertical multiphase flowing wells using artificial neural networks. J. Petrol. Sci. Eng. 75, 336–342.
Kalananda, V.K.R.A., Komanapalli, V.L.N., 2021. A combinatorial social group whale optimization algorithm for numerical and engineering optimization problems. Appl. Soft Comput. 99, 106903.
Kommineni, M., Kavanamani, R., Kurupati, A., Cherian, T.T., Aborig, A.M., Hossain, M., Miah, M.I., 2017. Effect of varying permeability on inflow performance curve of gas reservoirs. In: SPE Oil and Gas India Conference and Exhibition.
Li, Z., Hovakimyan, N., Kaasa, G.-O., 2012. Bottomhole pressure estimation and L1 adaptive control in managed pressure drilling system. In: Proceedings of the 2012 IFAC Workshop on Automatic Control in Offshore Oil and Gas Production. Trondheim, Norway.
Malik, H., Yadav, A.K., 2021. A novel hybrid approach based on relief algorithm and fuzzy reinforcement learning approach for predicting wind speed. Sustain. Energy Technol. Assessments 43, 100920.
Moayedi, H., Gör, M., Foong, L.K., Bahiraei, M., 2021. Imperialist competitive algorithm hybridized with multilayer perceptron to predict the load-settlement of square footing on layered soils. Measurement 172, 108837.
Nguyen, T.C., Pande, S., Bui, D., Al-Safran, E., Nguyen, H.V., 2020. Pressure dependent permeability: unconventional approach on well performance. J. Petrol. Sci. Eng. 193, 107358.
Nooraeni, R., Arsa, M.I., Projo, N.W.K., 2021. Fuzzy centroid and genetic algorithms: solution for numeric and categorical mixed data clustering. Procedia Comput. Sci. 179, 677–684.
Okoro, E.E., Obomanu, T., Sanni, S.E., Olatunji, D.I., Igbinedion, P., 2021. Application of artificial intelligence in predicting the dynamics of bottom hole pressure for under-balanced drilling: extra tree compared with feed forward neural network model. Petroleum. https://doi.org/10.1016/j.petlm.2021.03.001.
Osman, S.A., Ayoub, M.A., Aggour, M.A., 2005. An artificial neural network model for predicting bottomhole flowing pressure in vertical multiphase flow. In: SPE Middle East Oil and Gas Show and Conference.
Paasche, M., Johansen, T.A., Imsland, L., 2011. Regularized and adaptive nonlinear moving horizon estimation of bttomhole pressure during oil well drilling. In: Proceedings of the 18th World Congress the International Federation of Automatic Control. Milano (Italy).
Raghavan, R., Chin, L.Y., 2004. Productivity Changes in Reservoirs with Stress-dependent Permeability. SPE Reservoir Evaluation & Engineering.
Salimidelshad, Y., Moradzadeh, A., Kazemzadeh, E., Pourafshary, P., Majdi, A., 2019. Effect of hysteresis on petrophysical properties of limestone hydrocarbon reservoir rock. J. Petrol. Sci. Eng. 177, 745–755.
Scutari, M., Graafland, C.E., Gutiérrez, J.M., 2019. Who learns better Bayesian network structures: accuracy and speed of structure learning algorithms. Int. J. Approx. Reason. 115, 235–253.
Spesivtsev, P., Sinkov, K., Sofronov, I., Zimina, A., Umnov, A., Yarullin, R., Vetrov, D., 2018. Predictive model for bottomhole pressure based on machine learning.

J. Petrol. Sci. Eng. 166, 825–841.
Tabatabaie, S.H., Pooladi-Darvish, M., Mattat, L., 2015. Draw-down management leads to better productivity – or does it?. In: SPE/CSUR Unconventional Resources Conference Calgary, Alberta, Canada.
Tariq, Z., Mahmoud, M., Abdulraheem, A., 2020. Real-time prognosis of flowing bottom-hole pressure in a vertical well for a multiphase flow using computational intelligence techniques. J. Pet. Explor. Prod. Technol. 10, 1411–1428.
Wang, L.L., Vandamme, M., Pereira, J.M., Dangla, P., Espinoza, N., 2018. Permeability changes in coal seams: the role of anisotropy. Int. J. Coal Geol. 199, 52–64.
Yasser, M., Metwally, M., Sondergeld Carl, H., 2011. Measuring low permeabilities of gas-sands and shales using a pressure transmission technique. Int. J. Rock Mech. Min. Sci. 48, 1135–1144.
Yilmaz, O., Nolen-Hoeksema, R.C., Nur, A., 1994. Pore pressure profiles in fractured and compliant rocks. Geophys. Prospect. 42, 693–714.
Zeng, F., Nait Amar, M., Mohammed, A.S., Motahari, M.R., Hasanipanah, M., 2021. Improving the performance of LSSVM model in predicting the safety factor for circular failure slope through optimization algorithms. Eng. Comput.
Zheng, D., Pang, X., Zhou, L., You, X., Liu, X., Guo, F., Li, W., 2020. Critical conditions of tight oil charging and determination of the lower limits of petrophysical properties for effective tight reservoirs: a case study from the Fengcheng Formation in the Fengcheng area, Junggar Basin. J. Petrol. Sci. Eng. 190, 107135.