

AMADI, K., IYALLA, I., PRABHU, R., ALSABA, M. and WALY, M. 2023. Development of predictive optimization model for autonomous rotary drilling system using machine learning approach. *Journal of petroleum exploration and production technology* [online], 13(10), pages 2049-2062. Available from: <https://doi.org/10.1007/s13202-023-01656-9>

Development of predictive optimization model for autonomous rotary drilling system using machine learning approach.

AMADI, K., IYALLA, I., PRABHU, R., ALSABA, M. and WALY, M.

2023

© The Author(s) 2023.



Development of predictive optimization model for autonomous rotary drilling system using machine learning approach

Kingsley Amadi¹ · Ibiye Iyalla² · Radhakrishna Prabhu² · Mortadha Alsaba¹ · Marwa Waly¹

Received: 18 September 2022 / Accepted: 19 May 2023 / Published online: 16 June 2023
© The Author(s) 2023

Abstract

The growing global energy demand and strict environmental policies motivate the use of technology and performance improvement techniques in drilling operations. In the traditional drilling method, the effort and time required to optimize drilling depend on the effectiveness of human driller in selecting the optimal set of parameters to improve system performance. Although existing work has identified the significance of upscaling from manual drilling to autonomous drilling system, little has been done to support this transition. In this paper, predictive optimization model is proposed for autonomous drilling systems. To evaluate optimized operating procedure, a comparative study of surface operating parameters using weight on bit (WOB), rotary speed (RPM) versus drilling mechanical specific energy (DMSE), and feed thrust (FET) is presented. The study used a data-driven approach that uses offset drilling data with machine learning model in finding a pair of input operating variables that serves as best tuning parameters for the topdrive and drawwork system. The results illustrate that derived variables (DMSE, FET) gave higher prediction accuracy with correlation coefficient (R^2) of 0.985, root mean square error (RMSE) of 7.6 and average absolute percentage error (AAPE) of 34, whilst using the surface operating parameters (WOB, RPM) delivered an R^2 , RMSE and AAPE of 0.74, 28 and 106, respectively. Although previous researches have predicted ROP using ANN, this research considered the selection of tuning control variables and using it in predicting the system ROP for an autonomous system. The model output offers parameter optimization and adaptive control of autonomous drilling system.

Keywords Autonomous drilling system · Penetration rate prediction · Artificial neural network · Machine learning

Abbreviations

List of symbols

d_b	Diameter of the bit (in.)
D	Depth (m)
DTOR	Downhole torque (Kft.lbs)
WOB	Weight on bit (Klbs)
FD	Footage drilled by bit (ft.)
k	Drillability constant (N)
DMSE	Drilling mechanical specific energy (Kpsi)
Q	Mud flow-in-rate (Gpm)
R	Coefficient of correlation
R^2	Coefficient of determination

ROP	Rate of penetration (ft/hr)
RPM	Revolution per minutes (Rev/min)
UCS	Unconfined compressive strength (Kpsi)
ν	Cutter's radius
Vf	Transfer function
Wif	Connection length
Xf	Input data (in.)
ϕ	Feed rate
y_i	Output function of the ith hidden node (ft/sec)

Abbreviations

Term	Acronyms
AAPE	Average absolute percentage error
ANN	Artificial neural network
ASP	Actual surface parameter
DCV	Derived controllable variable
Dia	Diameter
DMSE	Drilling mechanical specific energy
ELM	Ensemble
FET	Feed thrust
GA	Genetic algorithms

✉ Kingsley Amadi
K.Amadi@au.edu.kw

¹ College of Engineering, Australian University, Mishref, Kuwait

² School of Engineering, Robert Gordon University, Aberdeen, UK

GR	Gamma-ray
ML	Machine learning
MLP	Multilayer perceptron
PDM	Predictive data-driven modelling
QRP	Quantitative real-time prediction
ROP	Rate of penetration
RPM	Revolution per minute
SVR	Support vector regression

Introduction

Hydrocarbons are formed on earth's subsurface by the decomposition of organic sediments deposited several millions of years ago. Upon increasing burial with depth, it becomes subjected to increasing temperature and pressure forming kerogen which later produces hydrocarbons within the pore spaces of the rock. A rock material is a naturally occurring aggregate of minerals, constituting an important part of earth crust. According to Emery (1966), a rock is defined as a composition of granular material and glue. Formation rock is a heterogeneous and anisotropic material and therefore a complex material to study from the mechanical viewpoint (Alfreds 1983). One of the most used classification of rocks is based on their origin, which classified rocks into three types of igneous, sedimentary, and metamorphic rocks Alfreds (1983). Hydrocarbons are commonly found in sedimentary rocks enclosed within a geologic trap. Drilling a borehole is the only way to harness the hydrocarbon deposited several thousands of feet beneath the subsurface. In oil-well drilling, the process involves creating a borehole achieved by simultaneous rotary action of the topdrive and the application of axial force by the drawwork hoist, wherein the former transmits torque to the drill bit via the drill string, and the latter facilitates drill string longitudinal motion, thereby establishing drill bit normal force commonly referred to as weight on bit (WOB) (Akgun 2002). During the drilling operation, the drill bit cuts the rock material and the resulting drill cuttings are removed from the borehole by the circulation of drilling fluid which is pumped into a well through the rotary hose and drill string; Fig. 1 shows the schematic of a topdrive rotary drilling rig.

There are two main modes of rotary drilling system: manual drilling and autonomous drilling system. The traditional manual mode of drilling system is manned by the driller who controls the rotary action of the topdrive and the axial force of the drawwork. The driller determines the operating parameters either by intuition, previous experience, or by trial-and-error approach. Alternatively, the autonomous rotary drilling system is designed to reduce mental and physical workload of human operator. There are several other benefits that could be derived from

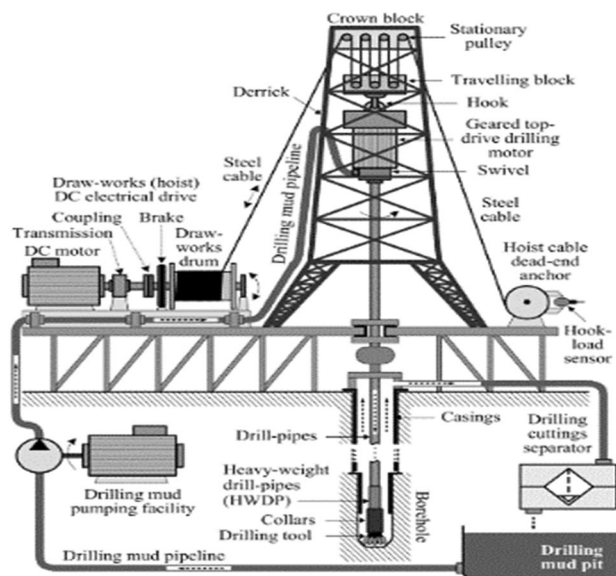


Fig. 1 Schematic of a topdrive rotary drilling rig with bottom-hole assembly (BHA) (Šprljan et al. 2020)

autonomous drilling system including efficiency, health, safety and environmental factors. First, automation responds faster to problems with fast and small corrections versus large corrections or costly remedial actions of conventional system. Also, autonomous drilling system will offer a step change in downhole monitoring of drilling conditions improving operation safety. Economic benefits emanate from operating closer to constraints and reducing well delivery time and cost and improve operational health and safety by elevating the driller from direct involvement in the drilling operation into a supervisory role, thereby eliminating the trial-and-error methods which in turn improve efficiency in terms of manpower and operational cost. This research supports the initiative in modelling and optimizing rate of penetration (ROP) of an autonomous rotary drilling system using predictive data-driven (PDM) modelling.

Literature review

The drive to optimize oil and gas portfolios through automation dates back to the late 90's. According to Norwegian Oil Industry Association (2006), Norway Scientific Council and Statoil developed a program called integrated operation to take the advantage of exploiting deeper subsea fields under a new information and communication technology program. Autonomous drilling system consists of three parts: the real-time monitoring, decision-making, and actuator as presented in Amadi et al. (2022). The real-time monitoring provides the environmental information for decision-making and control execution.

Amer et al. (2017) used the application of control system and information technology to operate both the topdrive and the drawwork system. McKenna et al. (2015). Chiranth et al. (2017) discussed the different established drilling models, which were classified into two groups: physics-based models and data-driven models.

The physics-based empirical models are commonly used in ROP prediction (Soares et al. 2016) including Graham and Muench model, Maurer model, Maurer (1962), Bingham model, Bingham (1965), Warren model, Warren (1987), Bourgoyne and Young model. Bourgoyne and Young (1974), Soares and Gray (2019) showed the disadvantages of the physics-based models, which includes the use of empirical constants such as the bit design constants, mud properties and formation constants, as well as low prediction, accuracy and their conformality to a homogenous formation. In contrast to the physics-based model, the contemporary data-driven models used the surface drilling parameters such as the weight on bit (WOB), rotary speed (RPM) and flowrate to predict ROP, Khoukhi and Ibrahim (2012). The implementation of the data-driven model overcame the limitations of the traditional physics-based models making optimization task less complex. McKenna et al. (2015) investigated the concept of a search logic optimizer to emulate experienced human controller using a typical “hill-climber” algorithm in a statistical model (Buntine 1992) for formation drilled whilst optimizing the setpoint combinations. However, it was reported that the primary challenge to the stochastic modelling systems was the non-homogeneity and the continuous change in the formation being drilled Hedge et al. (2015).

Bilgesu et al. (2017) tested the technology of neural network with the data-driven models to predict ROP and bit wear with different operating parameters. In their study, nine input parameters were used and result achieved a correlation coefficient (R^2) that ranged from (0.902 to 0.982). A similar study was conducted by Moran et al. (2010), using artificial neural network (ANN) to predict ROP with six input parameters including RPM, WOB, mud weight, rock strength, abrasion, and rock type from offset well data. The ROP prediction attained a coefficient of determination of R^2 value of 0.8. Manshad et al. (2017) created a multilayer ANN to model ROP with genetic algorithm (GA) with 332 datapoints and ten input parameters. Several other investigators including Jiang and Samuel (2016); Bhowmik et al. (2017); Arabjamaloei and Shadizadeh, (2011); Bataee and Mohseni (2011) and Manshad et al. (2017) have used ANN to predict ROP with five to ten input parameters. Table 1 shows the summary of penetration rate models and their respective model performance parameters.

Ahmed et al. (2019) performed a ranking of the 19 input parameters used in predicting ROP from the various studies based on parameters with the greatest influence on ROP using the feature extraction rule, which is used to evaluate the effect of input drilling parameters on ROP as shown in Fig. 2. The analysis showed that WOB, RPM and flow rate were the first, second and third most influencing factors. The application of machine learning (ML) for ROP prediction is widely supported by many researchers; out of the 53 reviewed works, 47% of the researchers used ANN followed by the Ensemble at 15% and SVR at 12% as shown in Fig. 3.

Data and methods

This section presents the procedures used in preparing dataset; it shows the defined input–output data relationship using a graphical representation. In addition, calculation of the derived controllable variables using appropriate established correlation and finally steps used in ANNs modelling and sensitivity studies were presented. Well-P05 has a total of 246 datapoints from 7400 to 10630 ft measured depth. The following drilling surface parameters were collated: weight on bit (WOB), rotary speed (RPM), torque (Tor) and penetration rate (ROP) as shown in Table 2.

Data gathering and reprocessing

Data collection and quality checks were the most challenging and time-consuming process in the study. During the study actual field drilling mechanics data from already drilled well, well-P05 was collected and resampled at 30-ft interval to reduce the data density as the data acquisition was done at every 0.5ft. This was done to check the influence of lesser data density on the accuracy of the prediction. Table 3 shows the summary of the statistics of the data.

Calculation of derived controllable variables.

The calculation of derived controllable variables was performed using established empirical relationship that relates the penetration rate with the derived input variables. A brief description of the derived variables and their empirical relationship is presented.

Feed thrust (FET)

This represents the force that keeps the drill bit in contact with the formation. This is controlled by the distance of drilling line released by the drawwork hoisting system. It is

Table 1 Summary of penetration rate models

Models	Prediction details		Method	Input parameters used in the model	Model performance parameters
	Output	# of inputs			
Hegde et al. (2018)	ROP	4	PSO	WOB, RPM, mud weight,	
Jiang and Samuel (2016)	ROP	5	ANN-Ant Colony Optimization	WOB, RPM, flowrate, depth, and gamma-ray	$R = 0.999$
Bataee and Mohseni (2011)	ROP	5	Genetic Algorithms	RPM, WOB, mud weight, Depth and Bit diameter	ANN, RMSE of 14.4 for training and 23.4 for testing using SVR, 27.3 for training and 27.6 for testing
Bilgesu et al. (1997)	ROP& Bit wear	9	ANN	WOB, RPM, torque, flowrate, rotating time, tooth wear, bearing wear, formation abrasiveness and formation drillability	500 dataset records, 90% for training and 10% for testing (R^2) ranged from (0.902) to (0.982)
Moran et al. (2010)	ROP	6	ANN	RPM, WOB, MW, rock strength, rock type formation abrasiveness	500 dataset records, 90% for training and 10% ANN and 23.2 for training and 27.1 R2 equal to 0.8
Jahanbakhshi et al. (2012)	ROP	21	ANN	WOB, RPM, pump pressure, ECD, mud type, YP, PV, Gel, mud pH, solid %, bit type, bit wear, bit hydraulic power, UCS, hole size porosity, hole depth, permeability, rock hardness, differential pressure,	ELM 0.94 for training 0.81 for testing using SVR, 0.74 for training and 0.72 for testing using ANN and 0.82 for training and 0.71 for testing
Arabjamaloei & Shadizadeh (2011)	ROP	7	ANN	RPM, WOB, flow rate, mud density, viscosity, depth, bit size, bit hours, bit efficiency and annulus pressure	330 dataset records, 10- input parameters Training ($R^2 = 0.94$) and testing ($R^2 = 0.74$)
Elkatatny et al. (2017)	ROP	7	ANN	RPM, WOB, flowrate, SPP, torque, drilling fluid density and plastic viscosity	3333 dataset records, 70% training, and 30% for testing. ($R = 0.997$, & $R = 0.993$) for training and testing respectively AAPE = 5.6
Ahmed et al. (2018)	ROP	8	ANN, SVR, ELM	Depth, flow rate, WOB, RPM, TQ standpipe pressure, mud weight and bit size	8869 datasets, 9- input, 70% for training 15% for testing and 15% for Validation $R^2 = 0.95$ AAPE=0.22, SVR: $R = 0.96$,
Bodaghi et al. (2015)	ROP	12	ANN, SVR & others	Pump rate, tooth wear, mud weight (MW), WOB, pump pressure, well deviation, mud viscosity, lithology, bit size, RPM, bit tooth wear, and interval drilled	93 datapoint from 13 wells, 154 points for training, 39 points for testing, Model result: ANN ($R^2 = 0.95$ AAPE=0.22) For SVR ($R = 0.96$, AAPE=0.078)
Shi et al. (2016)	ROP	10	ANN and ELM	RPM, WOB, pump pressure, mud density, mud viscosity, abrasiveness, UCS, bit wear, bit type, and bit size	5000 data records, ANN: $R^2 = 0.90$ RMSE 3.56, SVR: $R = 0.96$, AAPE=0.078

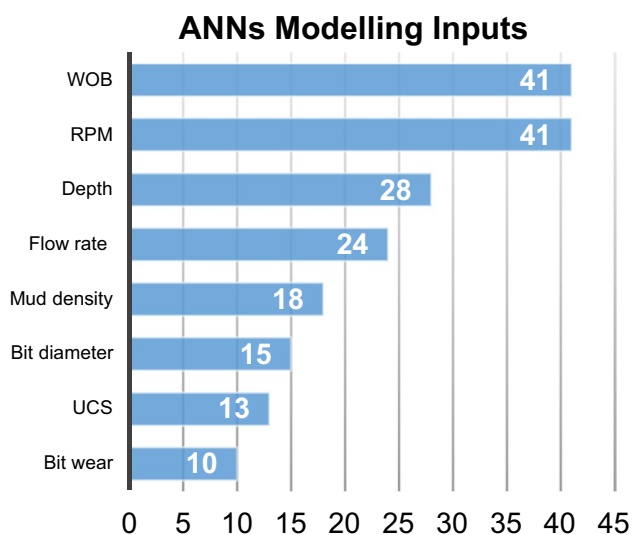


Fig. 2 Frequency of input variables used in ROP prediction

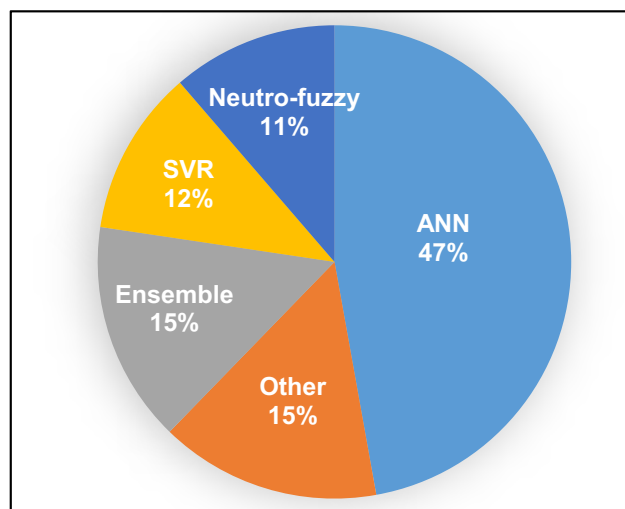


Fig. 3 Distribution of researchers and machine learning methods

Table 2 List of drilling mechanics data used for this study

Surface drilling data	Symbol	Unit
Weight on bit	(WOB)	(klbs)
Rotary Speed	(RPM)	(rev/min)
Rotary Torque	(TOR)	(Kft-lb)
Penetration rate	(ROP)	(ft/hr)

a representation of the weight on bit (WOB) or axial thrust force. Equation (1) used to calculate feed thrust force (FET) was derived from earlier work of Lindqvist (1982) on the indentation force of hemispherical carbide buttons in rock and discussed in Cavanaugh et al (2008)

Table 3 Range of input parameters and statistics (Dataset-P05)

Parameters	Min	Max	Mean	Median
Database counts	246	246	–	–
Depth (ft)	7400	10,630	9308	9345
ROP (ft/hr)	1.03	163.39	26.42	17.06
Torque (Kft-lb)	0	28	21.40	3.7
SWOB (Klbs)	1	70	38.0	38.70
RPM (rev/min)	1	70	119	120

$$FET = \frac{1.5 * Tor}{Dia} \frac{\varphi - 2v}{\sqrt{(\varphi v - v^2)}} \tag{1}$$

where FET is feed thrust force (klbs), Tor is surface torque (Kft.lbs), v is cutter radius (in), φ is penetration rate per revolution (ft), Dia is drill bit or hole diameter (in) and RPM is revolution per minute.

Drilling mechanical specific energy (DMSE)

The concept of drilling mechanical specific energy was first introduced by Teale (1965), defined as the amount of energy required to destroy a unit volume of rock. Equation (2) is the mathematical relationship for DMSE in oilfield units. The DMSE is a representation of the energy from topdrive.

$$DMSE = \left(\frac{480 * TOR * RPM}{Dia^2 * ROP} + \frac{4 * WOB}{\pi * Dia^2} \right) \times \text{Bit factor} \tag{2}$$

where DMSE is drilling mechanical specific energy (Kpsi), Tor is surface torque (Kft.lbs), RPM is revolution per minute, bit factor is 0.125 for PDC bit and 0.25 for roller cone bit, WOB is weight on bit (klbs), ROP is rate of penetration (ft/hr), and Dia = hole diameter (in),

Figure 4 shows the graphical presentation of the data points from well-P05 showing the effect of the respective tuning parameters on ROP. An explicit relationship is observed with the derived variable as seen in tracks [C&D].

In track D, an inverse relationship is observed between DMSE and ROP, whilst in track C, a direct relationship is observed between FET and ROP. No explicit relation was observed in track A and track B.

Statistical analyses

The influence of the input surface drilling parameters on ROP (output) in the drilling system was studied by performing statistical analysis of the dataset. The statistical analysis was useful in identifying outliers (data outside the acceptable limits) which were subsequently deleted from the dataset. Table 3 shows summary of input parameters for well-P05. Understanding the relationship between the

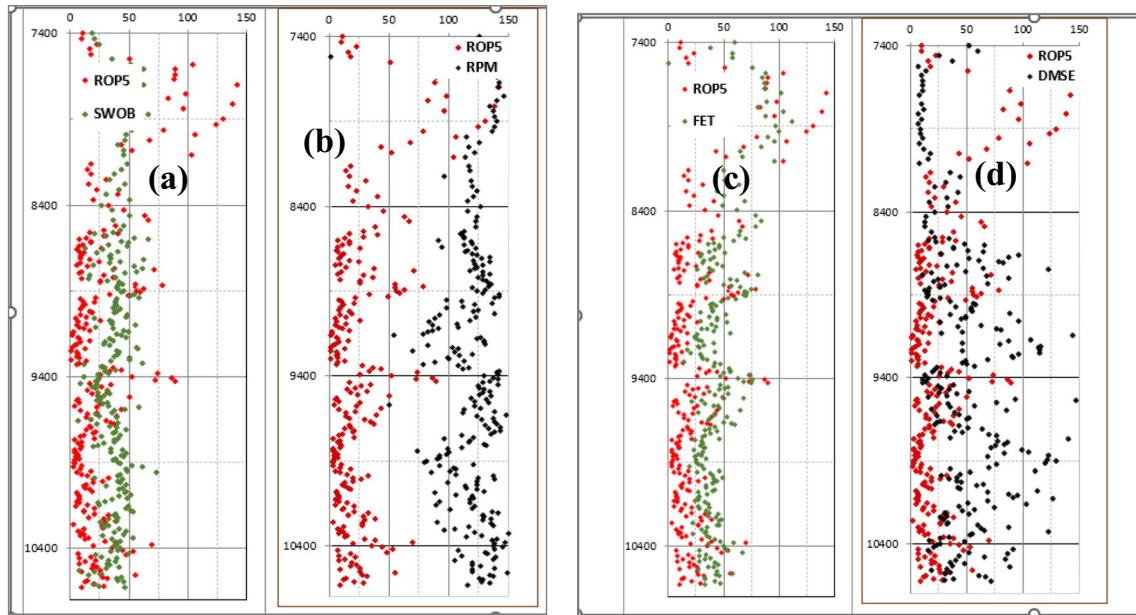


Fig. 4 Plot showing relationship between the selected tuning parameters with ROP **A** SWOB versus ROP **B** RPM versus ROP **C** FET versus ROP and **D** DMSE versus ROP for Dataset of Well P05

input and output variables is of primary importance when developing ANN models.

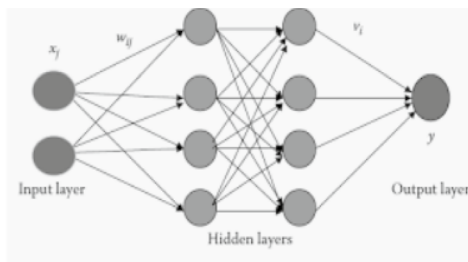
Artificial neural networks

ANNs are information processing structures that are universal function approximator and simplify simulation of a biological learning process with performance features like those of biological neural networks. Neural network consists of groups of nodes, or artificial neurons, that are linked in a way that allows them to be a universal function approximator, which means that given the right combination of nodes and connections, ANNs can simulate any input and output relationship. This is because they are adaptive, parallel information processing systems, which can build associations and mappings between objects or data and have been proven in solving problems

in automation and pattern recognition Shadizadeh et al (2010). A multilayer perceptron (MLP) network can learn the behaviour and trends of earlier events with valid dataset. Therefore, the quality and accuracy of the datasets are very crucial in the accuracy of the prediction and subsequent decisions made by the ANNs technique. The processing elements of ANNs are artificial neurons. These neurons consist of four basic components that include input data (xf), connection lengths (wf), a transfer function (vf), and output value (y) (Ahmed et al 2019) as shown in Fig. 5

Artificial neural networks configuration

Artificial neural networks (ANNs) are interconnected in multilayer network topology that comprise three layers: (1) input layer, (2) one or more hidden layers, and (3)



ANN Parameters	
Number of hidden layers	1
Number of neurons	30
Type of network function	FEED FORWARD
Type of transfer function	PURELIN
Maximum number of iterations	1000
Training Function	Levenberg Marquardt
Minimum performance gradient	15.4
Maximum value for Mu	0.00100

Fig. 5 ANNs structure with 2 hidden layers showing input data (xf), transfer function (vf), and output value (y)

an output layer as shown in Fig. 5. The hidden layer(s) are the coefficients that provide the relationship between the input and output layers. The most common types of ANNs are feed-forward networks, which are the most efficient ones Abbas et al (2018). In the feed-forward ANNs structure, the information will transmit in one direction from input neurons through the transfer function of the hidden neurons to the outputs. Depending on the relative contribution of each unit in the hidden layer to produce the original output, these units will receive only a fraction of the total error signal. The neurons use transfer functions to create their output from the network input. The most used transfer functions are the PURELIN, TANSIG, and LOGSIG. The number of neurons in the input layer is equal to the number of parameters being fed into the network as input and vice versa for the target layer. However, a variation of neutrons in the hidden layers was used in the optimization and detailed discussed provided in the sensitivity analysis section below. Furthermore, the parameters used in the ANN modelling are summarized in Table 4. The model was built using ANNs with two neurons in the input layer and only one neuron in the output layer, the optimized number of thirty (30) neutrons. Different hidden layers of varying numbers of neurons were simulated with the PURELIN transfer functions, whilst the output layer was one neuron. The optimum number of neurons and layers was selected based on an iterative process by performing sensitivity analysis on the number of neutrons that provide the highest accuracy with respect to correlation coefficient (R^2) as recommended by Haykin and Hakin (1998). In the modelling process, the database was randomly divided into two parts: A training dataset was used to develop and adjust the weights in a network, whilst the testing dataset was applied to examine the final performance of the ANNs. A 70:30 ratio was used for training and testing, respectively, as recommended in literature as more datapoint is required for training the model.

ANNs Simulation Cases.

Two simulation cases were studied, the first case “Case 1” was with input variables of WOB and RPM and ROP as

the single output variable. However, in “Case 2” the input variables were DMSE and FET with ROP as the single output parameter. To effectively compare the performance of the input variables in ROP prediction, the same ANNs network configuration, number of hidden layers and transfer function were used in the scenarios modelling.

ANNs model performance assessment criteria

The criteria applied for assessing performance of the two cases are the three commonly used in engineering analysis benchmark to align with the best practice. The predicted performances of neural network models were assessed by correlation coefficient (R^2), root mean square error (RMSE), and absolute average percentage error (AAPE).

Correlation coefficient (R^2) is a measure of the similarity between the actual and the predicted values. The range of value of (R^2) varies between 0 and 1. Whilst the value of 0 suggests no similarity, 1 signifies an excellent correlation between the model output and the actual predicted values. It is mathematically expressed using Eq. (3):

$$R^2 = 1 - \frac{\sum_{i=1}^n (f(xi) - yi)^2}{\sum_{i=1}^n f(xi)^2 - \sum_{i=1}^n f\left(\frac{yi}{n}\right)^2} \quad (3)$$

where yi presents actual data, $f(xi)$ is the predicted data, xi are the input parameters, and n is the total number of records. The higher R^2 shows a close approximation between the actual and predicted values. The range of value of (R^2) varies between 0 and 1. Whilst the value of 0 suggests no similarity, 1 signifies an excellent correlation between the model output and the actual predicted values.

The root means square error (RMSE) is a measure of error between the actual and the predicted values. It is used as an error function for the quality evaluation of the model. It is mathematically expressed in Eq. (4);

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (f(xi) - yi)^2} \quad (4)$$

The average absolute percentage error (AAPE) is a statistical measure of the relative accuracy of the model prediction expressed in percentage. It can be calculated as the ratio of the mean of the absolute error as shown in Eq. (5)

$$AAPE = \frac{1}{n} \sum_{i=1}^n \left[\frac{|xi - yi|}{xi} \right] * 100 \quad (5)$$

Table 4 ANN modeling and tuning parameters

ANN parameters	
Number of hidden layers	1
Number of neutrons	30
Type of network function	Feed-forward
Type of transfer function	PURELIN
Maximum number of iterations	1000
Training function	Levenberg–Marquardt
Minimum performance gradient	15.4
Maximum value for Mu	0.00100

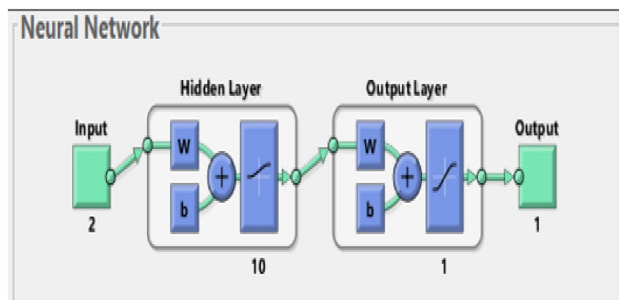


Fig. 6 ANN configuration with two input and an output variable

Methodology

As illustrated in Fig. 7, a step-wise procedure is used in evaluating tuning functionality for autonomous system. Two options investigated include [WOB, RPM] and [DMSE, FET]. All the steps of the methodology were coded in the MATLAB version R2019a. The steps include:

Step 1: Calculation of the derived controllable variables

The feed thrust (FET) and drilling mechanical specific energy (DMSE) were computed using Eqs. (1 and 2), respectively.

Step 2: Statistics and graphical analysis was performed to identify and eliminate outliers from the dataset

Step 3: The dataset was divided in the ratio of 70:30 for training and testing, respectively. A total of 246 datapoints were used for well-P05 and 4185 datapoints for well-W1

Step 4: The ANN model was constructed with input layer, one hidden layer and the output layer. The number of neurons in the hidden layer was varied in the sensitivity analysis to optimize the model result. Figure 6 shows the model configuration

Results

The summary of results of the step-wise procedure shown in Fig. 7 for ROP prediction using the actual surface parameters (ASP) and derived controllable parameters (DCP) for well P05 is shown in Table 5.

Result of prediction using actual surface parameter (Case 1):

The application of two actual surface parameters: (i) the weight on bit (WOB) and (ii) rotary speed (RPM) to predict the actual measured ROP obtained during the drilling process is presented in Fig. 8. The input data gave a poor prediction with R^2 values of 0.74, RMSE of 28, and AAPE of 106. The cross-plot of actual measured ROP versus predicted ROP (a), the error distribution curve (b), and the

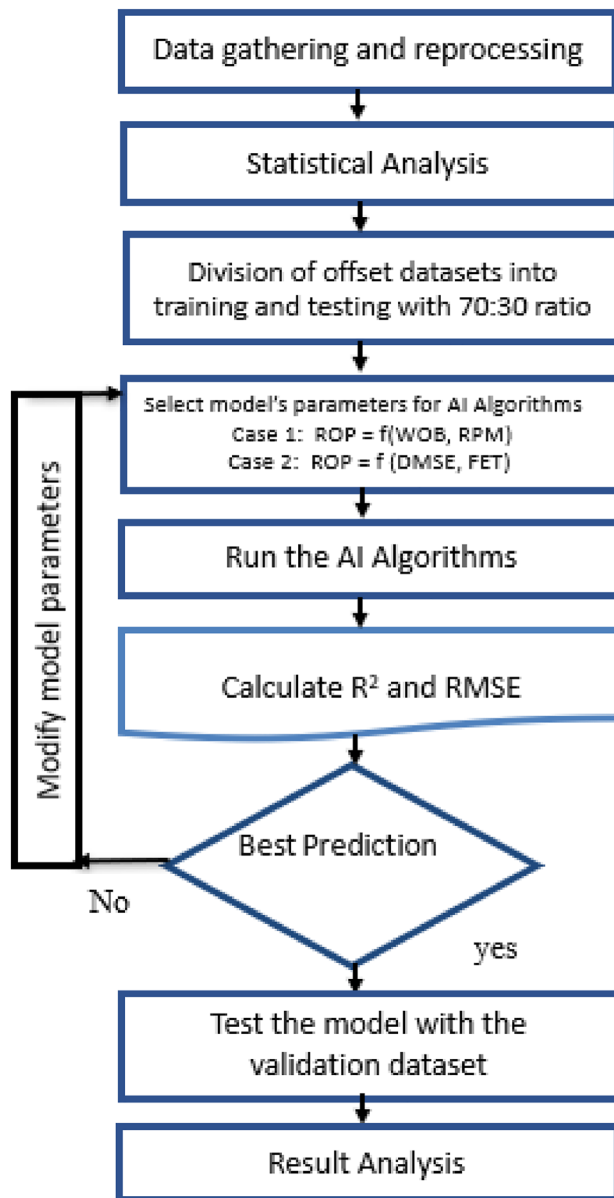


Fig. 7 Research flowchart showing the step-by-step approach

comparison of actual versus predicted plotted along depth (c) are presented in Fig. 8, whilst samples of numerical values comparison of the actual measured ROP and the ANN predicted values are presented in Table 6. This result revealed that as the value of weight on bit increases, the value of ROP increases and vice versa. However, the overall prediction accuracy using surface operating parameters was poor as shown in Table 6. This may be due to impact of downhole condition such wellbore tortuosity and hole cleaning on the effective transfer of weight to the drill bit.

Table 5 Summary of the quantitative prediction and accuracy

Well name		Variable type	Correlation Coefficient (R^2)	Root Mean Square Error (RMSE)	Average Absolute Percentage Error (AAPE)
Well, P05	ASP	[WOB, RPM]	0.74	28	106
	DCV	[DMSE, FET]	0.985	7.6	34
Well, W1	ASP	[WOB, RPM]	0.28	25	115
	DCV	[DMSE, FET]	0.98	3.8	21

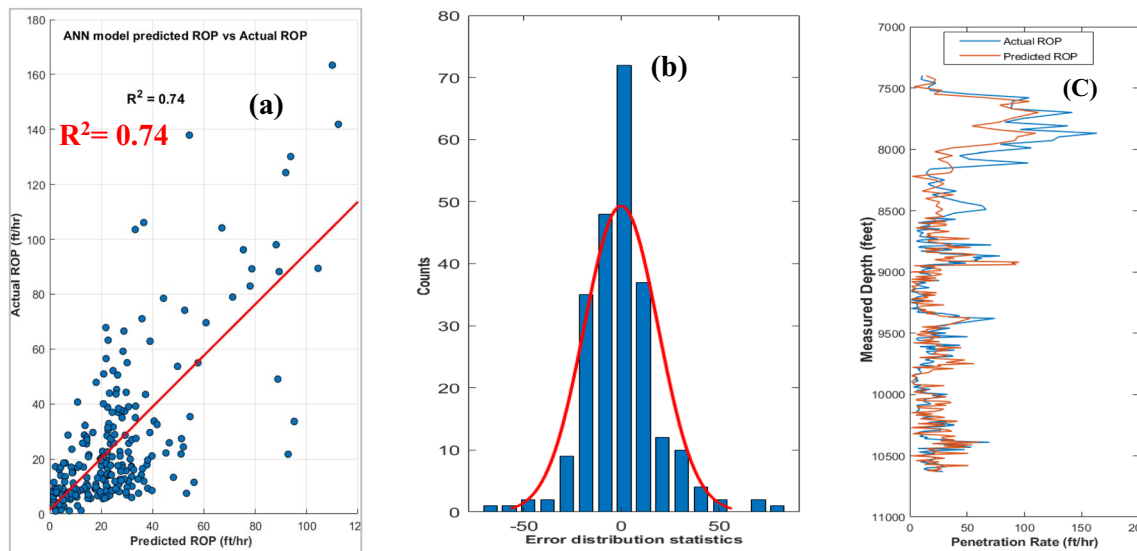


Fig. 8 Plot of **a** cross-plot predicted versus actual data: **b** Error distribution curve **c** Complot with depth with [WOB, RPM]

Table 6 Sample of predicted versus actual ROP data using [WOB, RPM]

Depth (ft)	WOB (Klbs)	(RPM) (rev/m)	Actual (ROP) (ft/hr)	ANN predicted ROP (ft/hr)	Prediction quality
7460	25	115	22.83	11.6	Poor
7550	35	125	50.92	22.92	Poor
7780	48	140	82.96	139.54	Poor
8160	48	115	17.91	23.02	Average
8550	42	115	28.70	15.05	Poor
8670	40	115	7.73	1.05	Poor
8790	36	130	10.06	22.8	Poor
9000	36	100	9.24	18.76	Poor
9400	12	130	51.94	25.23	Poor
9800	20	130	9.85	22.76	Poor
10,310	37	100	25.78	21.62	Average
10,630	46	115	9.50	11.80	Average

Result of prediction using derived controllable variables (case 2)

The use of derived controllable variables to predict ROP was also performed keeping the same ANN configurations

as in case 1. The modelling results delivered an excellent prediction with correlation coefficient (R^2) of 0.985, RMSE of 7.6, and AAPE of 34. The model (a) the cross-plot of the actual versus predicted ROP, (b) error distribution curve, and (c) the comparison plot of actual versus predicted ROP with depth are shown in Fig. 9.

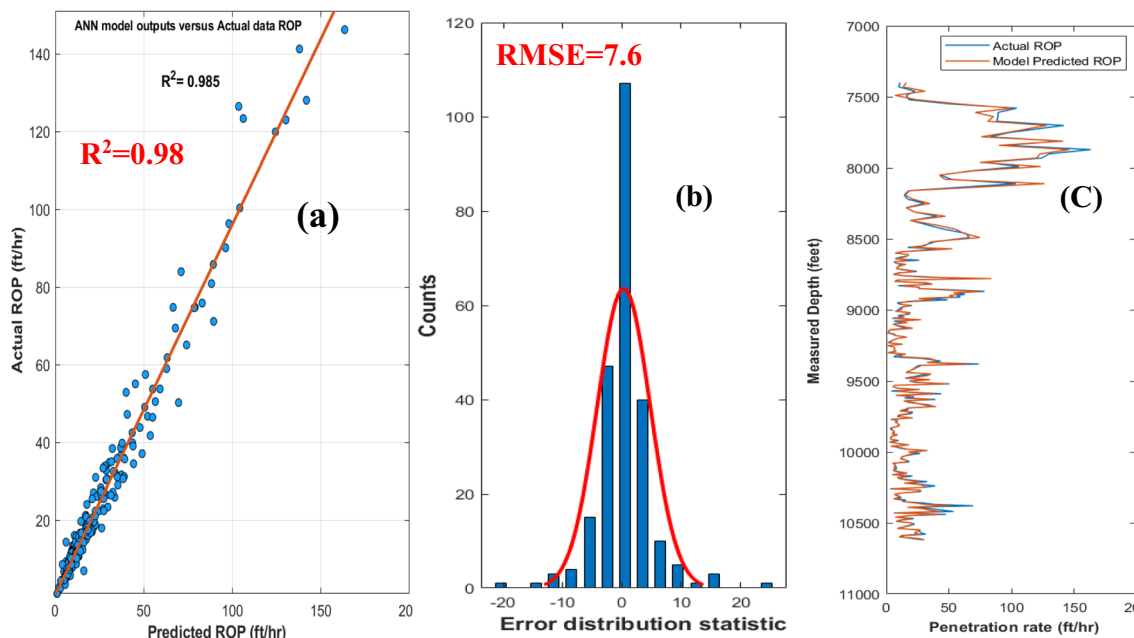


Fig. 9 a Cross-plot predicted versus Actual ROP; b Error distribution c Complot with depth with [DMSE, FET]

Table 7 Sample of predicted versus actual ROP data using [FET, DMSE]

Depth (ft)	FET (Klbs)	MSE (Psi)	Actual ROP (ft/hr)	ANN Predicted ROP (ft/hr)	Prediction quality
7460	0.308	25.85	22.83	31.82	Average
7550	0.42	14.70	50.92	59.10	Average
7780	0.48	12.23	82.96	80.15	Very good
8160	0.18	35.58	17.91	19.15	Very good
8550	0.26	22.00	28.70	29.40	Excellent
8670	0.13	87.60	7.73	7.21	Excellent
8790	0.12	67.78	10.06	9.05	Excellent
9000	0.15	54.66	9.24	11.20	Very good
9400	1.63	16.85	51.94	52.00	Excellent
9800	0.22	75.48	9.85	10.38	Very good
10,310	0.14	32.67	25.78	28.64	Very good
10,630	0.13	71.22	9.50	9.21	Excellent

The numerical values comparison of the actual ROP and ANN predicted ROP using [DMSE, FET] is presented in Table 7. It can be observed that derived controllable variable predicted the ROP within acceptable level of accuracy across the heterogeneous lithologic columns, which implies that the performance of autonomous system can be better modelled using derived variables as a tuning parameter.

Discussion of result

The study used ANN modelling in the determination of appropriate tuning parameters for autonomous rotary

drilling system in a complex heterogeneous formation by performing ROP prediction using proposed tuning variables. The adaptative self-optimizing capability of the system was investigated by evaluating the variation of the tuning parameters with changes in penetration rate across the heterogeneous rock. The following characteristics were maintained: (1) The ANN model was designed only for two input parameters each acting as a tuning parameter for the topdrive and the drawworks, respectively. (2) The predictive optimization model proposed two sets of tuning parameters: the actual surface operating parameter [WOB, RPM] and the derived energy parameters of [DMSE, FET] and tested their effectiveness as a controller variable with

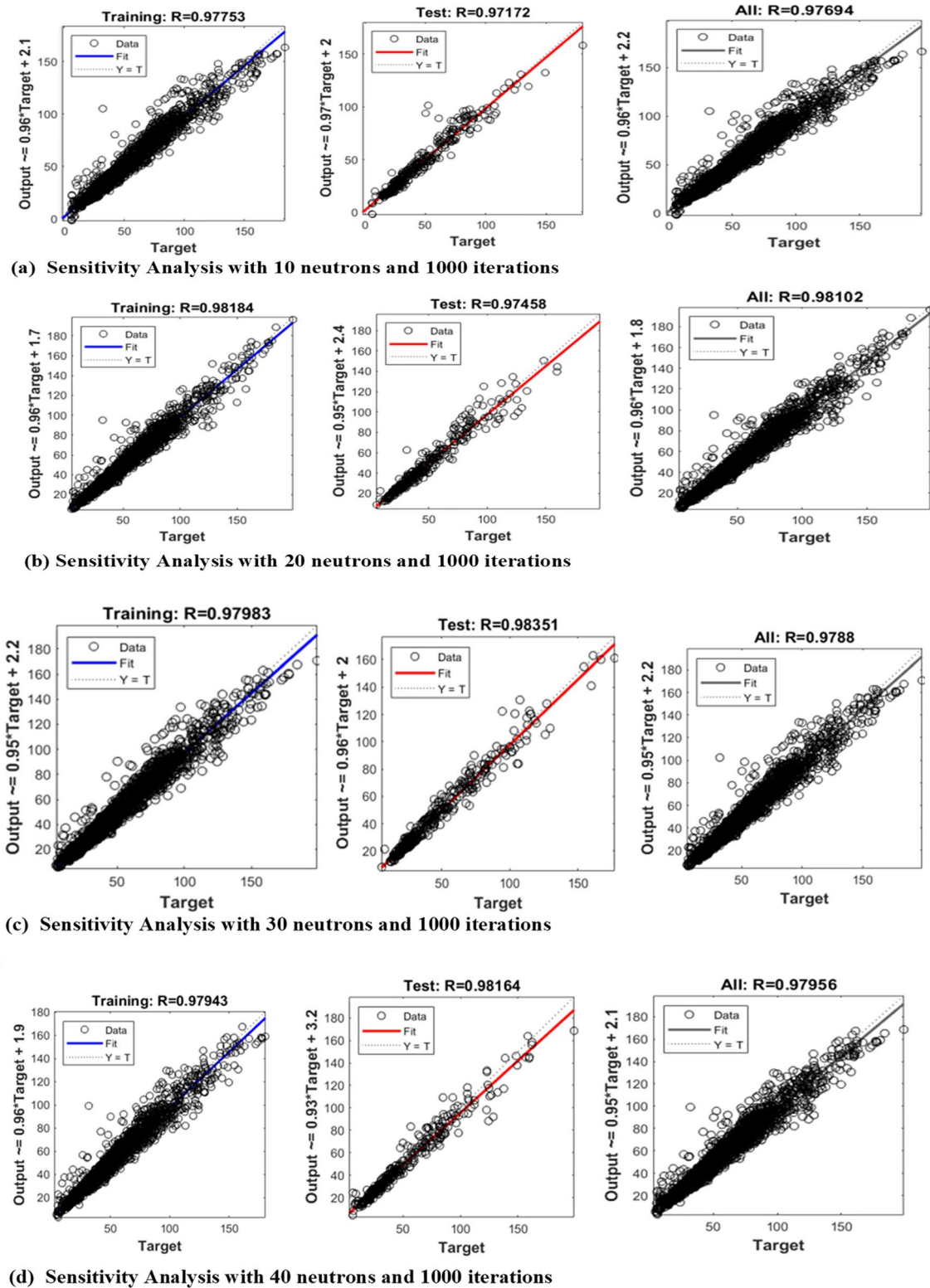


Fig. 10 Sensitivity analysis for the selection of optimal number of neutrons

ROP predictions. (3) The actual surface parameter [WOB, RPM] could not accurately predict the drill rate in a complex heterogeneous formation because of the influence of

downhole conditions such as wellbore tortuosity and torque and drag. Armentia (2008) identified these conditions for inefficient drilling. Therefore, ASP are not effective as an

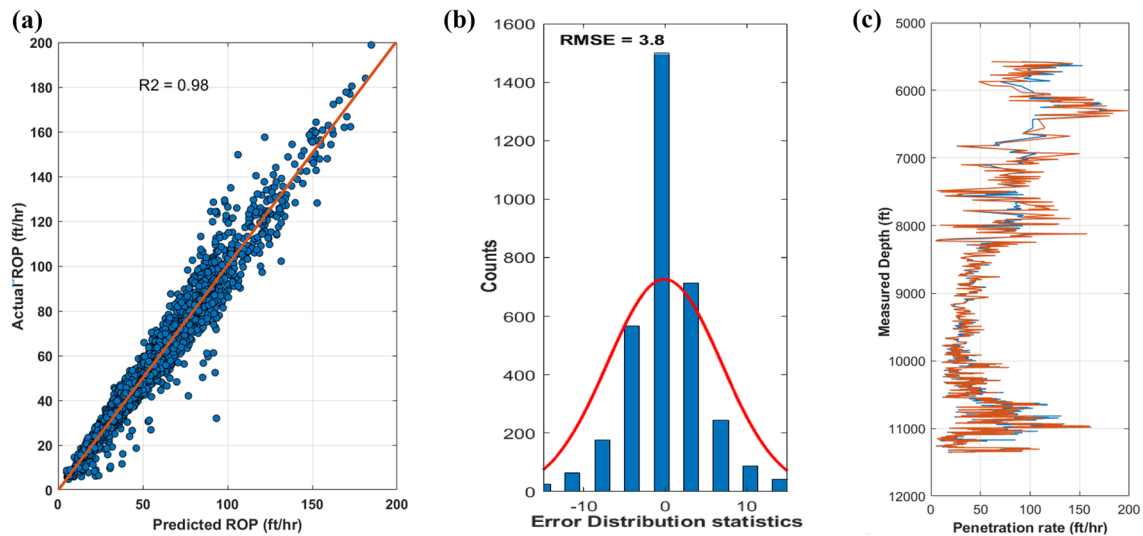


Fig. 11 a Cross-plot predicted versus actual data: b Error distribution c Complot with depth with [DMSE, FET] for Well W1

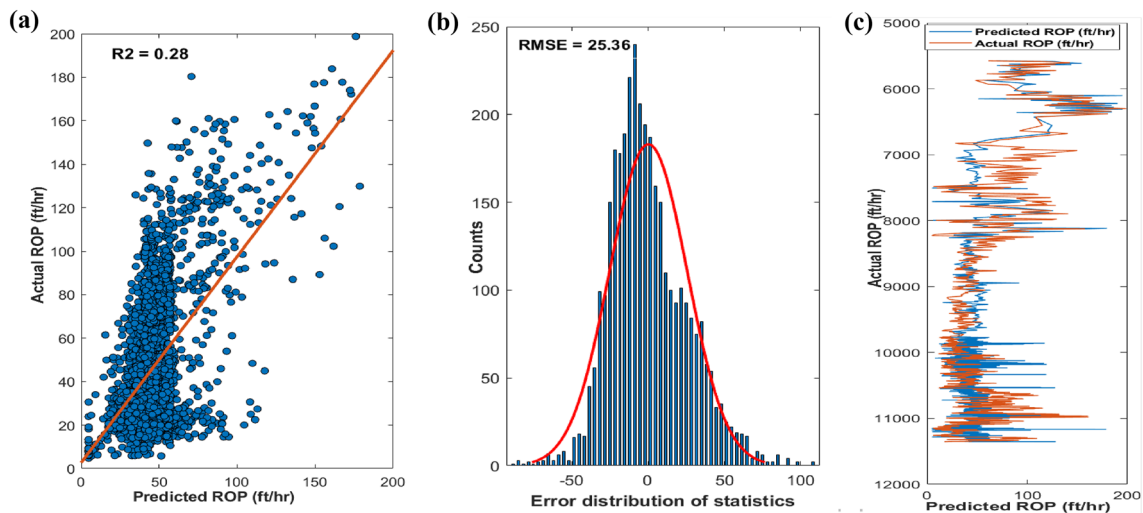


Fig. 12 a Cross-plot predicted versus actual data: b Error distribution c Complot with depth with [WOB, RPM] for Well W1

adaptive input parameter in autonomous system (4) The derived energy variable (DCV) performed excellently in effectively predicting the ROP with a high level of accuracy with a well-established trend with drill rate performance, thus proven to be an effective adaptive input parameter in autonomous system. 5) It was observed that the accuracy of the model depends on the quality and number of datapoints. The accuracy increases with increasing number of datapoint available. 6) The problem of overfitting in the model was avoiding by Splitting the datasets in 70:30 ratio for training and testing, respectively. Also, a new dataset was used in the validation of the model.

Sensitivity analysis

Sensitivity analysis was performed in the selection of optimum number of neurons for the hidden layers. Results showed that the optimum number of 30 neurons gave the highest accuracy which cannot be further improved. Figure 10 shows the results of the sensitivity analysis with number of neurons at 1000 iterations performed using the derived controllable variable (DCV) at 10, 20, 30, and 40 neurons, respectively.

Model verification

Verification of the model was performed with dataset from a different wellbore (well-W1) which has total of 4185 datapoints. The performance result using the DCV and ASP is shown in Figs. 11 and 12, respectively. The prediction error of the DCV variable decreased with higher data density. Similarly, the performance of ASP was worst with higher data density with R^2 value of 0.28 and RMSE of 25 and AAPE of 115, conversely the DCV variables, precisely predicted the actual measured ROP in a complex heterogeneous formation column with R^2 , RMSE, and AAPE values of 0.98, 3.4, and 34, respectively. The overall summary of result is presented in Table 5.

Conclusions

This research aimed at identifying appropriate tuning parameters and their predictive performance for self-optimizing autonomous rotary drilling systems using an artificial neural network with actual surface drilling parameters and derived controllable parameters. The following conclusions are reached based on the study;

1. The direct use of actual surface parameters (WOB, RPM) as tuning parameter for autonomous system produced a poor ROP prediction for drill rate with coefficient of determination (R^2) of 0.74, RMSE of 28, and AAPE of 106 with well-P05 dataset, suggesting the inadequacy of using these two parameters alone.
2. The use of derived variables (DMSE, FET) as tuning parameter gave an excellent prediction accuracy with coefficient of determination (R^2) of 0.98, RMSE of 7.6, and AAPE of 34 using same well-P05 dataset.
3. The result supports the decision to use the data-driven (ANN) model with derived controllable variables in the quantitative prediction of drilling rate for an autonomous drilling system.
4. Establishing a precise quantitative relation between derived variables and rate of penetration will improve parameter optimization, operational efficiency, and equipment reliability.
5. In achieving an optimized operating procedure, the auto-controller should monitor the resulting DMSE for the topdrive rotary speed an adjust parameter with the aim of achieving the lowest DMSE values.
6. Similarly, auto-controller will monitor the resulting feed thrust (FET) from the drawwork hoist and adjust parameter to attain a higher FET value provided all other drilling conditions such as hole cleaning and slick slip are within acceptable range.

Author contributions All authors contributed to the study's conception and design. Material preparation, data collection, and analysis were performed by Kingsley Amadi and Ibiye Iyalla. All authors read and approved the final manuscript. Kingsley Amadi contributed to conceptualization, methodology, software, investigation, writing—original draft, and formal analysis. Mortadha Alsaba contributed in validation, review and editing the manuscript. Prabhu Radhakrishna contributed in the supervision, and project administration and Marwa Waly performed data curation, editing, and visualization.

Funding No funding was received to assist with the preparation of this manuscript.

Declarations

Conflict of interest The authors have no relevant financial or non-financial interests to disclose.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Abbas AK, Rushdi S, Alsaba M (2018) Modeling rate of penetration for deviated wells using artificial neural network. In: Abu Dhabi International Petroleum Exhibition and Conference (ADIPEC), Abu Dhabi, UAE, Nov. 12–15, Paper No. SPE-192875-MS. <https://doi.org/10.2118/192875-MS>
- Ahmed KA, Rushdi S, Alasba M, AlDushaishi MF (2019) Drilling rate of penetration prediction of high angled wells using artificial neural network. *J Petrol Sci Eng* 141:1–11. <https://doi.org/10.1115/1.4043699>
- Ahmed A, Elkhatny S, Ali A, Mahmoud M, Abdurraheem A (2018) New model for pore pressure prediction while drilling using artificial neural networks. *Arabian J. Sci. Eng.* 2018, 44, 6079–6088. <http://doi.org/https://doi.org/10.1007/s13369-018-3574-7>
- Akgun F (2002) How to estimate the maximum achievable drilling rate without jeopardizing safety. In: Proceedings of the SPE-78567 MS Was Presented at Abu Dhabi International Petroleum Exhibition and Conference, Abu Dhabi, UAE, 13–16 October 2002. <https://doi.org/10.2118/78567-MS>
- Alfreds RJ (1983) *Rock mechanics*. Clausthal-Zellerfeld, 2nd edn. Trans Tech Publications, Gulf Pub Co., Federal Republic of Germany
- Amadi K, Iyalla I, Prabhu R, Alsaba M, Waly M (2022) Continuous dynamic drill-off test whilst drilling using reinforcement learning in autonomous rotary drilling system SPE-211723-MS presented at 2022 Abu Dhabi International Petroleum Exhibition and Conference. 31st–3rd Nov. 2023. <https://doi.org/10.2118/211723-MS>
- Amer MM, Dahas AS, El-Sayed AH (2017) An ROP predictive model in Nile delta using artificial neural networks. In:

- SPE-187969-MS Presented at the SPE Kingdom of Saudi Arabia Annual Technical Symposium and Exhibition. <https://doi.org/10.2118/187969-MS>
- Arabjamaloei R, Shadizadeh S (2011) Modeling and optimizing rate of penetration using intelligent systems in an Iranian southern oil field (ahwaz oil field). *Pet Sci Technol* 29:1637–1648. <https://doi.org/10.1080/10916460902882818>
- Armenta M (2008) Identifying inefficient drilling conditions using drilling specific energy. In: SPE Annual Technical conference and Ex. 2008 Denver, Colorado. <https://doi.org/10.2118/116667-MS>
- Bataee M, Mohseni S (2011) Application of artificial intelligent systems in ROP optimization: a case study. In: Proceedings of the SPE-140029-MS was Presented at SPE Middle East Unconventional Gas Conference and Exhibition, Muscat, Oman, 31 January–2 February 2011. <https://doi.org/10.2118/140029-MS>
- Bhowmik S, Panua R, Debroy D, Paul A (2017) Artificial neural network prediction of diesel engine performance and emission fueled with diesel–kerosene–ethanol blends: a fuzzy-based optimization. *ASME J Energy Resour Technol* 139(4):042201. <https://doi.org/10.1115/1.4035886>
- Bilgesu HI, Tetrick LT, Altmis U, Mohaghegh S, Ameri S (1997) A new approach for the prediction of rate of penetration (ROP) values. In: Proceedings of the SPE-39231-MS Was Presented at SPE Eastern Regional Meeting, Lexington, KY, USA, 22–24 October 1997. <https://doi.org/10.2118/39231-MS>
- Bilgesu H, Cox ZD, Elshehabi C, Idowu GO (2017) A realtime interactive drill-off test utilizing artificial intelligence algorithms for DSAT drilling automation university competition 2017 <https://doi.org/10.2118/185730-MS>
- Bingham MG (1965) A new approach to interpreting—rock drillability. Petroleum Pub. Co., Tulsa, OK
- Bodaghi A, Ansari HR, Gholami M (2015) Optimized support vector regression for drilling rate of penetration estimation. *Open Geosci*. <https://doi.org/10.1515/geo-2015-0054>
- Buntine W (1992) Learning classification trees. *Stat Comput* 2(2):63–67. <https://doi.org/10.1007/BF01889584>
- Burgoyne AT, Young FS (1974) Multiple regression approach to optimal drilling and abnormal pressure detection. *J SPE* 14:371–384. <https://doi.org/10.2118/4238-PA>
- Cavanough GL, Kochanek M, Cunningham JB, Gipps ID (2008) A self-optimizing control system for Hard rock percussive drilling. *IEEE/ASME Trans Mechatron* 2008(1):1083–4435. <https://doi.org/10.1109/TMECH.2008.918477>
- Chiranth H, Daigle H, Millwater H, Gray K (2017) Analysis of the rate of penetration (ROP) prediction in drilling using physic-based and data driven Models. *J Petrol Sci Eng* 159:295–306. <https://doi.org/10.1016/j.petrol.2017.09.020>
- Elkatatny SM, Tariq Z, Mahmoud MA, Al-Abdul-Jabbar A (2017) Optimization of rate of penetration using artificial intelligent techniques. In: Proceedings of the ARMA-2017–0429 Was Presented at 51st U.S. Rock Mechanics/ Geomechanics Symposium, San Francisco, CA, USA, 25–28 June 2017
- Emery CL (1966) The strain in rocks in relation to highway design. *Rock Mech* 135:1–9
- Haykin S, Hakin S (1998) The neural networks a comprehensive foundation. Macmillan College Publishing Company, Stuttgart, Germany
- Hedge CM, Wallace SP, Gray KE (2015) Use of regression and bootstrapping in drilling interference and prediction. Presented at SPE Middle East intelligent oil and Gas Conference and Exhibition. SPE. <https://doi.org/10.2118/176791-MS>
- Hegde C, Soares C, Gray K (2018) Rate of penetration (ROP) modeling using hybrid models: deterministic and machine learning. In: Proceedings of the 6th Unconventional Resources Technology Conference. <https://doi.org/10.15530/urtec-2018-2896522>
- Jahanbakhshi R, Keshavarzi R, Jafarnezhad A (2012) Real-time prediction of rate of penetration during drilling operation in oil and gas wells. In: Proceedings of the ARMA-2012–244 Was Presented at 46th U.S. Rock Mechanics Geomechanics Symposium, Chicago, IL, USA, 24–27 June 2012
- Jiang W, Samuel R (2016) Optimization of rate of penetration in a convoluted drilling framework using ant colony optimization. In: Proceedings of the SPE-178847MS Presented at IADC/SPE Drilling Conference and Exhibition, Fort Worth, TX, USA, 1–3 March 2016. <https://doi.org/10.2118/178847-MS>
- Khoukhi A, Ibrahim A (2012) Rate of penetration prediction and optimization using advanced artificial neural networks, a comparative study. In: Proceedings of the 4th International Joint Conference on Computational Intelligence, Barcelona, Spain. 5–7 October 2012
- Lindqvist (1982) Rock fragmentation by indentation and disc cutting some theoretical and experimental studies
- Manshad A, Rostami H, Toreifi H, Mohammad A (2017) An optimization of drilling penetration rate in oil fields using artificial intelligence technique. In: *Heavy Oil: Characteristics, Production and Emerging Technologies*. Nova Science Publishers, Hauppauge, NY, USA, pp. 255–269
- Maurer W (1962) The perfect cleaning theory of rotary drilling. *J Pet Technol* 14:1270–1274. <https://doi.org/10.2118/408-PA>
- McKenna I, Pan R, Koederitz W (2015) The application of real-time stochastic analysis for autonomous drilling optimization. *Oil Mag Eur* 41(1):21–23
- Moran DP, Ibrahim HF, Purwanto A, Osmond J (2010) Sophisticated ROP prediction technology based on neural network delivers accurate results. In: Proceedings of the SPE-132010-MS Was Presented at IADC/SPE Asia Pacific Drilling Technology Conference and Exhibition, Ho Chi Minh City, Vietnam, 1–3 November 2010. <https://doi.org/10.2118/132010-MS>
- Norwegian Oil Industry Association, (2006) Integrated operation enhance value creation. Equinor Norway
- Shadizadeh SR, Karimi F, Zoveidavianpoor M (2010) Drilling stuck pipe prediction in Iranian oil fields: an artificial neural network approach. *Iran J Chem Eng* 7(4):29–41
- Shi X, Liu G, Gong X, Zhang J, Wang J, Zhang H (2016) An efficient approach for real-time prediction of rate of penetration in offshore drilling. *Math Prob Eng* 2016(2016):1–13. <https://doi.org/10.1155/2016/3575380>
- Soares C, Daigle Gray HK (2016) Evaluation of PDC bit ROP models and the effect of rock strength on the model coefficient. *J Nat Gas Sci Eng* 2016(34):1225–1236. <https://doi.org/10.1016/j.jngse.2016.08.012>
- Soares C, Gray K (2019) Real-time predictive capabilities of analytical and machine learning rate of penetration (ROP) models. *J Pet Sci Eng*. <https://doi.org/10.1016/j.petrol.2018>
- Šprljan P, Pavkovic D, Cipek M, Klaić M, Staroveski T, Kolar D (2020) Automation systems design and laboratory prototyping aimed at mature petroleum drilling rig retrofitting. *Tehnicki Vjesnik* 27(1):229–236. <https://doi.org/10.17559/TV-20180924173727>
- Teale R (1965) The concept of specific energy in rock drilling. *Int J Rock Mech Min Sci* 2(1):57–73. [https://doi.org/10.1016/0148-9062\(65\)90022-7](https://doi.org/10.1016/0148-9062(65)90022-7)
- Warren TM (1987) Penetration-rate performance of roller cone bits. *SPE Drill Eng* 1987(2):9–18. <https://doi.org/10.2118/13259-PA>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.