Journal of Petroleum Science and Engineering 183 (2019) 106332

Contents lists available at ScienceDirect

Journal of Petroleum Science and Engineering

journal homepage: www.elsevier.com/locate/petrol

Machine learning methods applied to drilling rate of penetration prediction and optimization - A review

Luís Felipe F.M. Barbosa[a,*], Andreas Nascimento[b,c,d,**], Mauro Hugo Mathias[e], João Andrade de Carvalho Jr.[a]

[a] São Paulo State University (Unesp), School of Engineering, Department of Energy, Guaratinguetá, Brazil
[b] Universidade Federal do Espírito Santo - Department of Engineering and Technology – Petroleum Engineering Programs (UFES/DETEC), São Mateus, Brazil
[c] International Institute for Applied System Analysis (IIASA/ASA), Laxenburg, Austria
[d] Montanuniversität Leoben - Department Petroleum Engineering - Chair of Drilling and Completion Engineering (MUL/DPE/CDC), Leoben, Austria
[e] São Paulo State University (Unesp), School of Engineering, Department of Mechanics, Guaratinguetá, Brazil

ARTICLE INFO

Keywords:
Data-driven model
Artificial intelligence
Multi-objective optimization
ROP
Drilling optimization

ABSTRACT

Drilling wells in challenging oil/gas environments implies in large capital expenditure on wellbore's construction. In order to optimize the drilling related operation, real-time decisions making have been put in place, so that prediction of rate of penetration (ROP) with accuracy is essential. Despite many efforts (theoretical and experimental) throughout the years, modeling the ROP as a mathematical function of some key variables is not so trivial, due to the highly non-linearity behavior experienced. Therefore, several researches in the recent years have been proposing to use data-driven models from artificial intelligence field for ROP prediction and optimization.

This paper presents an extensive review of the literature on ROP prediction, especially, with machine learning techniques, as well as how these models can be used to optimize the drilling activities. The ROP models are classified as traditional models (based on physics-models), statistical models (e.g. multiple regression), or machine learning methods. This review enables to see that machine learning techniques can potentially outperform in terms of ROP-prediction accuracy on top of traditional or statistical models. Throughout this work, an extensive analysis of different ways of obtaining ROP models is carried out, concluding with different strategies adopted in literature to perform data-driven model optimization.

Despite the saving potential which can be achieved with real-time optimization based on data-driven ROP models, it is noticeable that there is a lack of implementation of those techniques in the industry, as per literature review. To take a step forward in real implementations, the petroleum industry must be aware that yet no rule of thumb already exists on this specific area, but still, good and very reasonable results can be achieved by following the best practices identified in this review. In addition, the modern practices of machine learning provide promising guidelines for implementing projects in oil and gas industry.

1. Introduction

The share of offshore hydrocarbon extraction is globally increasing. Due to high rates of drill rigs, the cost of exploration and well development in offshore reserves is a deterministic factor for the hydrocarbon production in those areas (Skjerpen et al., 2018). Most of well drilling cost is not product cost dependent, but time dependent (Lyons and Plisga, 2004). Therefore, one of the main goals of drilling optimization is to reduce the total time, maintaining the risks as low as possible. One way to achieve it is through selecting of optimum drilling variables prior a run (e.g. selecting a suitable drill-bit and drilling fluid type). Another approach relies on real-time analysis in order to optimize operational parameters (e.g. bit weight, rotary speed) while drilling (see e.g. Eren and Ozbayoglu, 2010; Payette et al., 2017).

To formulate an optimization problem for the drilling activities, it is required to have accurate predictive models. Such models have the goal to assess how some important variables (e.g. drill bits, rotating speed, weight on bit) affect the drilling performance, which can be measured e.g. by rate of penetration (ROP) or other metrics, including the specific energy proposed by Teale (1965), highlighted in the paper from

* Corresponding author.
** Corresponding author.
E-mail addresses: luis-felipe.barbosa@unesp.br (L.F.F.M. Barbosa), andreas.nascimento@ufes.br (A. Nascimento).

https://doi.org/10.1016/j.petrol.2019.106332
Received 31 March 2019; Received in revised form 1 August 2019; Accepted 2 August 2019
Available online 07 August 2019
0920-4105/

Dupriest and Koederitz (2005). The prediction of rate of penetration (ROP) as function of drilling variables paves the way to formulate the optimization problem as maximization of ROP, or minimization of total time or cost per feet drilled. For that, the accuracy of ROP model is crucial (Soares and Gray, 2019). However, understanding how the drilling variables really affect the ROP is an open question in drilling engineering (Mitchell and Miska, 2011).

Despite many efforts (theoretical and experimental), modeling the ROP as a mathematical function of some variables is not so trivial, because this is highly non-linear problem. This was shown recently by Soares et al. (2016), who exposed limitations of traditional ROP modeling based on analytical equations - inclusive the well-known traditional model proposed by Bourgoyne and Young (1974). Therefore, many researchers have started employing machine learning techniques (e.g. artificial neural networks, support vector machines, random forests) to predict ROP due to their well-known capability as universal function approximators (Hornik et al., 1989). Most studies (e.g. Arabjamaloei and Shadizadeh, 2011; Amar and Ibrahim, 2012; Gandelman, 2012; Bataee et al., 2014; Hegde et al., 2017) that compared traditional models with machine learning techniques concluded that a higher accuracy could be achieved in ROP prediction with intelligent techniques. Inclusive, it was shown the theoretical feasibility of implementation artificial intelligence for real-time drilling optimization by Gandelman (2012) and in the series of works from Hegde and colleagues (Hegde and Gray, 2017, 2018; Hegde et al., 2018a).

One of the first applications of Machine Learning (ML) for drilling parameters prediction found in OnePetro was reported by Arehart (1990), who utilized artificial neural networks (ANN) to predict an important drill bit parameter, namely the bit wear. Later, Bilgesu et al. (1997) published the first work that applied ANN to predict the ROP. Both laboratory data and field data were employed for the training process and evaluation. After a long period without further development in this area, researchers have been publishing several works since 2010. This can be seen in Fig. 1, which depicts the chronological distribution of the 61 publications reviewed in the current paper, that used ML methods for ROP prediction. All these publications were grouped in a total of 53 different works, since some works were a continuation from previous contributions.

Intelligent methods can be also applied to predict other drilling parameters, such as weight on bit (Khosravanian et al., 2016b), pump

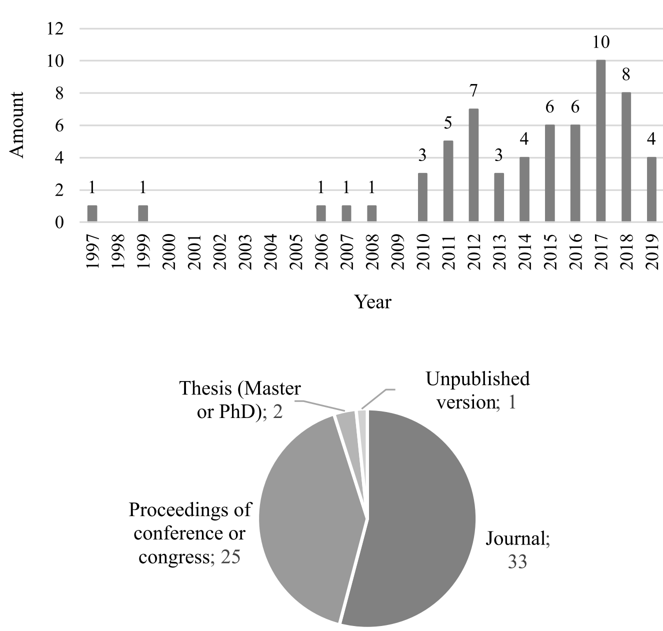

Fig. 1. Yearly distribution and publication source of reviewed works that used ML methods for ROP prediction.

pressure of the drilling fluid (Wang and Salehi, 2015), drilling fluid density (Ahmadi, 2016), or to select the optimum drilling parameters, e.g. the optimum bit (Yılmaz et al., 2002). The applicability of artificial intelligence (AI) methods in drilling engineering can also been seen in some recent reviews. Bello et al. (2016) reviewed the employment of AI techniques in several fields related to oil and gas industry, for example: reservoir simulation, seismic pattern recognition, reservoir characterization, permeability and porosity prediction, drill bit diagnosis, well production optimization etc. The review from Agwu et al. (2018) covered the use of AI in drilling fluid engineering. Noshi and Schubert (2018) summarized data-driven models and workflows for failure patterns recognition and remedial actions recommendations. In another review, Rahmanifard and Plaksina (2018) analyzed several techniques of AI employed in oil and gas industry (e.g. oil production rate, minimum miscibility pressure, and volume of CO2 sequestration), giving a special attention to heuristic optimization methods, such as genetic algorithm, particle swarm and differential evolution. For a broad review on data-driven models applied to oil and gas industry, the readers can refer to Balaji et al. (2018).

In the open literature, it seems that no try has been attempted so far to review exclusively the current progress of ML techniques for ROP prediction. Therefore, this is the main objective of the current paper, which tries to answer some important questions related to prediction of ROP with intelligent techniques. For example, is it really necessary to feed ML models with downhole variables as inputs (e.g. rock strength, drill bit wear), which may not be so easy available as the surface variables? In order to know, which method (ML or traditional models) is more suitable for estimating ROP, this review analyzes those works that compared different approaches in ROP prediction. Another objective is to analyze how these predictive models can be used for optimization purpose. A preliminary version of this study was presented in a conference by the authors (Barbosa et al., 2018). That paper reviewed partially the use non-traditional models to predict the ROP; only 11 studies were analyzed. In addition, that work was restricted to ROP modeling. Now the complete version of the review is presented, showing also how the ROP models can be employed to optimize controllable drilling variables.

1.1. Review outline

After the introduction section, section 2 covers the methods for ROP prediction, presenting three possible approaches for that: traditional models, statistical models and ML models. In section 3, a discussion on ROP modeling with ML techniques is performed. It covers some important decisions to be made in such studies, including the selection of most important inputs, techniques to perform feature selection, some approaches in data partition, and methods for handle possible problems in datasets. Additionally, a comparison between different techniques in ROP modeling is carried out, highlighting those works that compared ML models with traditional equations in ROP modeling.

Since the main goal of ROP prediction is to optimize drilling activities, section 4 analyzed those works that recently utilized ROP models for optimization purposes. The current review shows also that most works employed a single-objective framework, considering normally maximization of ROP as objective function. However, this approach may not improve the overall drilling time. During the development of the current review, some sophisticate optimization methods were found, which consisted of formulating multi-objective optimization problems by considering two or more objectives at the same time. Besides, a discussion about challenges of implementing drilling optimization systems is covered in section 5. Finally, the work is concluded in section 6, by summarizing the most important findings and trends observed in this review.

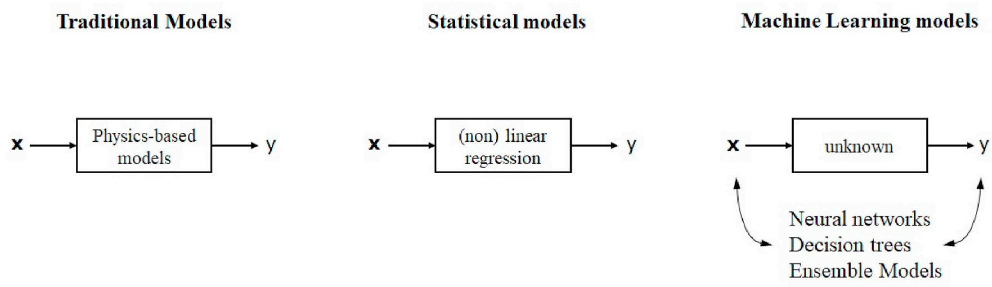

Fig. 2. Approaches for ROP modeling. Adapted from Breiman (2001b).

2. Methods used in ROP prediction

To predict the rate of penetration (ROP), it is necessary to have a model able to assess how the drilling variables affect ROP. In a general framework, obtaining a ROP model can be thought of as a regression problem. Following the notation adopted by Friedman (2006), the prediction of ROP can be formulated as

$$\hat{y} = f(\boldsymbol{x}, \Theta) \tag{1}$$

where $\hat{y}$ represents the estimation for the output variable (i.e. ROP), $\boldsymbol{x}$ represents the vector of inputs or predictors (e.g. drilling parameters, such as drill bit type, torque, WOB, mud flow), $f(\cdot)$ the regression function, and Θ the regression function's parameter set.

The function $f(\cdot)$ maps the input space $\mathscr{X}$ into the output space $\mathscr{Y}$, i.e. $f: \mathscr{X} \to \mathscr{Y}$. The main goal of a regression task is to produce a predictive model $f(\cdot)$ with high accuracy (Friedman, 2006). For that, different methods can be applied to obtain a predictive model for ROP. Hegde et al. (2017) classified the ROP models into two groups: traditional (physics-based) models and data-driven models. The authors grouped the linear regression models and machine learning methods into the group of data-driven models. In the current work, both techniques are not considered as belonging to the same group. In addition, there is the possibility of combining different approaches (e.g. neural networks with traditional models) into a unique predictor, resulting in so-called hybrid models. Thus, the ROP models are classified into four main groups[1], namely:

- traditional (physics-based) models,
- statistical models (e.g. linear regression), and
- machine learning models (e.g. artificial neural networks);
- hybrid models (e.g. combining traditional models with machine learning models).

The resulting classification is illustrated in Fig. 2. As Breiman (2001b) discussed about the two cultures in the use of statistical modeling, a similar comparison can be carried regarding the different cultures regarding ROP modeling. The first culture attempts to understand the physical nature of well drilling, describing the drilling activities with analytical equations. Normally, those equations are deducted to mathematically model how the most important drilling parameters affects the speed at each a hole is being drilled. This approach results in the traditional models, which are commonly found in drilling engineering books, such as the textbook authored by Mitchell and Miska (2011). Similar equations can also be obtained with statistical regression methods (e.g. with multivariate regression). Another approach starts with data, which serves the basis for obtaining the predictive models $f(\cdot)$ with ML methods (Breiman, 2001b). ML algorithms are able to learn how to map drilling operational parameters space into the output space (ROP), capturing the complexity of this activity. No model needs to be set prior the training phase. However, there is not a general ROP model for all situations (Soares and Gray, 2019).

Before presenting each approach in ROP modeling, the supervised-learning paradigm is described, which was employed in ROP studies. In the next Section 2.1, an overview of traditional models is first given, showing also some different methods employed to obtain ROP models based on analytical equations. Then, the non-traditional methods for ROP are presented in Section 2.2 (statistical approach, e.g. multivariate regression) and Section 2.3 (ML methods).

Supervised learning. The learning paradigm for regression problems is the supervised learning. The use of previous examples from a specific problem enables to teach a predictive model how to map the input space $\mathscr{X}$ to output space $\mathscr{Y}$. The dataset employed in learning phase is called as "training" dataset, where the output y has been jointly measured with the inputs variables $\boldsymbol{x}$ (Friedman, 2006).

In order to know whether the model $f(\cdot)$ is good or not, evaluation metrics are employed to assess the predictive model accuracy or its lack of accuracy. According to Friedman (2006, p. 176), it is employed "a 'loss' criterion that reflects the cost of mistakes: $L(y, \hat{y})$ is the loss or cost of predicting a value $\hat{y}$ for the response when its true value is y". The goal of learning is to define the set of parameters Θ from the predicting function by minimizing the loss functions (Hastie et al., 2009). For regression-type problems, the absolute error $L(y, \hat{y}) = |y - \hat{y}|$ is a common cost function, as well as the squared-error $L(y, \hat{y}) = (y - \hat{y})^2$, which yields to much simpler algorithms of minimization (Friedman, 2006).

The studies of ROP prediction consist basically of fitting a model $f(\cdot)$ for drill curves available, regardless of the method employed. In the learning phase, the set of regression parameters Θ is obtained by a minimizing the selected loss function. The set of parameters to be optimized are:

- the empirical coefficients of the analytical equations from the traditional models,
- the regression coefficients from statistical models,
- the internal parameters from ML model which store the knowledge about the problem.

2.1. Traditional models

Here, those ROP models which try to establish mathematical equations among the drilling variables are called as traditional models. This is because most of these models appeared in the initial phase of the scientific research of drilling optimization (see Eren and Ozbayoglu, 2010). Those models do not rely solely on the drilling data, as the ML models do. As consequence, it is possible to find out other researchers calling them as physics-based models (Hegde et al., 2017), or as simple as drilling models (Hareland and Hoberock, 1993). Recently, Soares and Gray (2019) called these models as analytical models.

[1] To the best of our knowledge, the hybrid approach has not been yet investigated, but only mentioned by Yavari et al. (2018). Therefore, we consider three first approaches in ROP modeling after this point in article (traditional models, statistical models and ML models).

2.1.1. Models overview

The goal of present subsection is to briefly introduce the most common traditional models. For more details, the readers can also refer to other works that provided an extensive review about the traditional ROP models (Eren, 2010; Eren and Ozbayoglu, 2010; Nascimento, 2016; Soares et al., 2016; Soares and Gray, 2019). There are many models describing the effects of several parameters on ROP. Some of traditional models which are worth of citing, because they are used in some papers for comparison purpose with data-driven models or due to their importance in the industry, are the following: Graham and Muench (1959); Maurer (1962); Bingham (1965); Young (1969); Bourgoyne and Young (1974); Warren (1987); Detournay and Defourny (1992); Hareland and Rampersad (1994). In the PhD thesis from Eren (2010), an overview of the history of ROP models and drilling optimization until 2010 is given.

Some initial ROP models are known also as R-W-N (rate of penetration, weight-on-bit, rotary speed), because ROP is written as a function mainly of both weight-on-bit and rotary speed (Nascimento et al., 2015). An example is the Maurer model, which assumes a condition of perfect bottomhole cleaning and incomplete bit tooth penetration. Maurer developed the following theoretical equation for rolling cutter bits (Maurer, 1962, *apud* Bourgoyne et al. (1986), p. 226):

$$ROP = \frac{K}{S^2}\left[\frac{W}{d_b} - \frac{W_o}{d_b}\right]^2 N \tag{2}$$

where K is the constant of proportionality, S is the compressive rock strength, W is the bit weight, W_o is the threshold bit weight, d_b is the drill-bit diameter, and N is the rotary speed.

Another R-W-N model was proposed by Bingham (Bingham, 1965, *apud* Bourgoyne et al. (1986), p. 227):

$$ROP = K\left(\frac{W}{d_b}\right)^{a_5} N \tag{3}$$

where K is the constant fo proportionality, including the effect of rock strength, and a_5 is the weight on bit exponent (Bourgoyne et al., 1986).

Bourgoyne and Young (1974) developed one of the most important ROP models, which is widely employed in the industry (Eren and Ozbayoglu, 2010). According to Soares and Gray (2019), the Bourgoyne and Young model (BYM) includes eight parameters and can be written as:

$$\frac{dD}{dt} = \exp\left(a_1 + \sum_{j=2}^{8} a_j x_j\right) \tag{4}$$

where D is the well depth, t is the time, the coefficient a_1 is related to the formation strength parameter, a_2 to the formation compaction, a_3 to pore pressure, a_4 to differential pressure, a_5 to the weight on bit exponent, a_6 to the rotary drilling (N), a_7 to the drill-bit tooth wear, a_8 to the bit hydraulic jet impact. Later, the authors Bourgoyne et al. (1986) proposed an adaptation to their original ROP model:

$$ROP = (f_1)*(f_2)*\cdots*(f_8) \tag{5}$$

where the functions f_1 to f_8 encompasses the empirical coefficients a_1 to a_8. According to Soares and Gray (2019), the main difference between both formulations (Bourgoyne and Young (1974); Bourgoyne et al. (1986)) is in the last function. The first BYM employs Eckel's hydraulics Reynolds number, while the updated version uses a power law function of the hydraulic jet impact force. The BYM equations represent all important aspects of drilling, but some parameters required in the model are not easily measured in real-time (e.g. drill bit wear, differential pressure) (Soares and Gray, 2019).

Based on cutter-rock interaction, Hareland and Rampersad (1994, *apud* Soares et al. (2016), p. 1229) proposed a general drag bit model:

$$ROP = \frac{14.14.\ N_c.\ N.\ A_v}{d_b} \tag{6}$$

where Nc is the number of cutters, A_v is the area of rock compressed ahead of a cutter, which assumes a different formulation according to the drill-bit type. The readers can refer to the works from Soares et al. (2016) for more details.

As already mentioned, the actual relationship among the drilling variables is very complex, and not well understood (Mitchell and Miska, 2011). Therefore, some efforts (Motahhari et al., 2010; Deng et al., 2016; Al-abduljabbar, 2019) have been recently made to better understand the relationship among the drilling variables and how they affect the ROP. For example, Motahhari proposed a ROP model for PDC bit (Motahhari et al., 2010, *apud* Soares et al. (2016), p. 1230)

$$ROP = W_f\frac{G.\ N^{\gamma}.\ W^{\alpha}}{d_b.\ S} \tag{7}$$

where W_f represents the wear, G is a coefficient related to bit-rock interactions and bit geometry, γ and α are coefficients of ROP model, and S is the confined rock strength.

Deng et al. (2016) proposed a theoretical model for determining the ROP for roller cone bit, and this model was validated with lab drilling results. The authors used the rock dynamic compressive strength instead of static compressive strength, what increased the accuracy of the theoretical model. Al-abduljabbar (2019) proposed a new ROP model based on regression analysis:

$$ROP = 16.96\frac{W^a\ .\ N\ .\ T\ .\ SSP\ .\ Q}{d_b^2\ .\ \rho\ .\ PV\ .\ UCS^b} \tag{8}$$

where 16.96 is the unit conversion factor employed by the authors, ρ is the mud density, T is the torque, SSP is the standpipe pressure, Q is the flow rate, PV is the plastic viscosity, UCS is the uniaxial compressive strength. To obtain the coefficients (a and b), nonlinear regression was adopted. According to the authors, this new model outperformed other traditional models (Bingham, BYM, and Maurer).

It is interesting to note that ROP modeling is not restricted to drilling of oil and gas wells. Basarir et al. (2014) mentioned other ROP models, applied to rock excavation and tunnel boring, showing the correlation between the rock properties and ROP, based on previous studies from Howarth et al. (1986); Kahraman (2002). Inclusive, it is possible to find out works using AI methods to predict ROP for tunnel boring machines (Hedayatzadeh et al., 2010; Mahdevari et al., 2014). However, the current works is limited to the prediction of ROP from rotary drilling systems, applied to oil and gas industry.

2.1.2. Fitting the models

The wish of obtaining ROP models for some specific regions is not new. For example, Bizanti (1989) developed ROP models for Lousiana formations, based on R-W-N equations. To obtain the ROP models in general, it is necessary to fit the empirical coefficients Θ from the ROP models, tailored to drill curves available. This fitting process is actually an optimization problem, whose goal is to determine the coefficients that minimizes the loss function. Bourgoyne and Young had originally proposed the employment of multiple regression to determine the eight coefficients ($\Theta = a_1, a_2, \ldots, a_8$) (Bourgoyne and Young, 1974, *apud* Bourgoyne et al. (1986), p. 227). This approach was applied in some works (Eren and Ozbayoglu, 2010; Gandelman, 2012; Soares et al., 2016). Besides, some alternatives in the fitting process were reported in the literature, including adaptations to the original BYM or the use of optimization techniques (e.g. genetic algorithms) to determine the optimum set Θ.

It is important to note that the BYM describes completely the drilling process, but this equation relies on normalization constants for depth, W, N and Q, which were derived for drill bits from the 1970s (Soares and Gray, 2019). Therefore some authors proposed adaptations to the original BYM equations. Nascimento et al. (2015) and Kutas et al. (2015) showed that using different normalization factors of BYM equation and allowing a wider range of applicable drillability coefficients could decrease the relative error of the ROP prediction. Different normalization factors was also utilized later by Soares and Gray (2019).

Rahimzadeh et al. (2011) investigated also another alternative for the multiple regression when using BYM. The authors developed an own iterative process in finding the optimum coefficients, based on progressive stochastic method. This method was compared against the regression and thrust-region. The fitting process with the progressive stochastic method could outperform other fitting techniques in ROP prediction. In another study (Soares and Gray, 2019), different optimization techniques to fit the empirical coefficients were compared (trust region reflective, Basin Hopping, Particle Swarm Optimization) and the authors selected the trust region for fitting the traditional models utilized in their work. To model the coefficient of BYM, Bahari et al. (2008) used a well-known meta-heuristic optimization techniques, called genetic algorithm. In the work from Hasan et al. (2011), the genetic algorithm was employed to first determine the set of parameters from BYM. Then, a General Regression Neural Network was employed in a hierarchical way to map the eight parameters from BYM in ROP prediction.

Other works explained also the fitting process for other models different than BYM. For example, Awotunde and Mutasiem (2014) explained also the fitting process for the modified Warren model (Hareland and Hoberock, 1993). To determine two of three bit coefficients, the authors employed dimensionless groups. The third bit coefficient was obtained by calling back the ROP model. Hankins et al. (2015) employed the Hareland and Rampersad (1994) model in his optimization study. In the series of works from Soares et al. (2016); Soares and Gray (2019), several traditional models were compared against each other.

All the previous studies performed a point estimation. Another way to estimate is called probabilistic prediction, which consists of estimating also the uncertainty around the predicted value. In the study from Formighieri and Filho (2015), the BYM coefficients were not determined as if they were fixed values. Instead, the authors determined the probabilistic distributions of the coefficients, as a natural approach to handle noises in data. Additionally, the uncertainties present in the data was passed to the model itself. The authors estimated the BYM coefficients with Markov Chain Monte Carlo simulation. To the best authors' knowledge, this was the only try to perform probabilistic prediction for ROP.

2.2. Statistical models

The approach of using statistical models in ROP prediction has some similarities with the approach of traditional model, but also some differences. The main similarity is necessity of pre-selecting a model for ROP as function of drilling variables. The main difference is the statistical models does not (normally) attempt to represent the physics of drill bit mechanism and the interaction between rock formation and bit, as the traditional models do.

An example of this approach is the work from Seifabad and Ehteshami (2013), who tested several regression equations for each formation in Ahvaz oil field. The goal was to obtain a general ROP model, tailored to each formation from Ahvaz field. Another example was performed by Moraveji and Naderi (2016). A full quadratic form of multiple regression (i.e. containing linear, quadratic and interaction coefficients) was applied in ROP prediction. By minimizing the error sum of squares, the authors determined the regression coefficients. The authors modeled the ROP as function of six drilling parameters: depth, W, N, bit jet impact force and two fluid properties (yield point to plastic viscosity ratio and 10 min to 10 s gel strength ratio).

Another interesting try was reported in Hegde et al. (2015b). The authors compared several techniques from statistics, including least squares, regularization techniques (rigde and lasso), principal component analysis (PCA) regression and bootstrap. The accuracy in ROP prediction could be improved with regularization techniques, which constrain the coefficients of regression models. The authors showed also an advantage of bootstraping technique, by providing confidential interval for ROP prediction. For more details about the regression methods employed by the authors, the textbook of Hastie et al. (2009) is a good source.

There are also other works (Arabjamaloei and Shadizadeh, 2011; Mantha and Samuel, 2016; Eskandarian et al., 2017; Hegde et al., 2017; Ashrafi et al., 2019) that employed multiple regression, with either linear coefficients or both linear and non-linear coefficients. Their purpose, however, was to compare different techniques of modeling the ROP (e.g. multiple regression with machine learning techniques), or to select the most important features in ROP prediction.

2.3. Machine learning applied to rate of penetration

The previous approaches (traditional models and statistical models) started preselecting a specific model. Differently, ML techniques are able to learn complex patterns during the training (or learning) phase, without having to specify a ROP model. After the learning phase, the trained model is able to make predictions given novel inputs. A brief overview about the most common ML methods employed in ROP prediction is given in section 2.3.1, which serves the basis for classifying the 53 reviewed articles (in section 2.3.2), that employed ML for ROP prediction.

2.3.1. ML overview

The 53 reviewed works that used ML for ROP prediction could be classified into five *methods*: artificial neural networks (ANN), support vector machines (SVM), fuzzy inference systems, neuro-fuzzy, and ensemble models. For each *method*, an additional information is given in the field of *detail*, as seen in Table 1.

ANN. Neural networks differ from each other normally in their architectures, types of unit activation functions (e.g. sigmoid, linear) and training algorithms. Those information was normally reported in the reviewed articles that employed ANN in ROP prediction. A very common type of network structure is feedforward neural network with multilayer perceptron (MLP). Fig. 3 illustrates a single-hidden-layer network, or a two-layer network as recommended by Bishop (2006, p. 229), since the network properties are determined by its amount of adaptive weights layers. In this case, the first layer of adaptive weights, $w^{(1)}$, connects each input, $x_1, \ldots, x_D$, to each hidden unit $z_1, \ldots, z_M$. The input vector belongs to a space $\mathbb{R}^D$. The adaptive weights of second layer, $w^{(2)}$, connect the output of each hidden unit to each outputs $y_1, \ldots, y_K$. The unit x_0 represents the bias term for each hidden unit, and the unit z_0 represents the bias term for each outputs units.

According to Bishop (2006), the overall network function (considering the network from Fig. 3) takes the form

$$\hat{y}_k(\boldsymbol{x}, \boldsymbol{w}) = \sigma\left(\sum_{j=1}^{M} w_{kj}^{(2)} h\left(\sum_{i=1}^{D} w_{ji}^{(1)} x_i + w_{j0}^{(1)}\right) + w_{k0}^{(2)}\right) \tag{9}$$

where $\hat{y}_k$ is the output (estimation) of kth unit, $\boldsymbol{w}$ are the weights and biases to be determined in the learning phase, and $\sigma(\cdot)$ and $h(\cdot)$ are activation functions of the output units and hidden units respectively. A common function is the sigmoidal $\sigma(a) = 1/(1 + exp(-a))$ (Bishop, 2006).

Another class of ANN is the radial basis functions (RBF). It can be thought of as a feedforward network with a single-hidden-layer units, whose activation functions are RBF instead of e.g. logistic. RBF measures only the distance from its input (e.g. $\boldsymbol{x} \in \mathbb{R}^D$) to the origin or some point, $\boldsymbol{c}$ (called as the center point). The distance is measured by $\varphi = (||\boldsymbol{x} - \boldsymbol{c}||)$, where $||\ldots||$ represents a norm imposed in the space $\mathbb{R}^D$ (Broomhead and Lowe, 1988).

The training phase consists basically of adapting the weights and biases (or centers for RBF), so that the neural network minimizes the loss function adopted in the training algorithm. A well-known training approach is the so-called *back-propagation* (BP), which consists of

Table 1
Classification of machine learning methods employed in ROP prediction.

Method	Detail	Reference
ANN		
	according to type/architecture	
	MLP: multilayer perceptrons	
	RBF: radial basis function	Powell (1987); Broomhead and Lowe (1988)
	other	
	according to training algorithm	
	own developed training algorithm by author(s)	
	ELM: extreme learning machine	Huang et al. (2004, 2006)
	BR: bayesian regularization	Buntine and Weigend (1991); MacKay (1991)
	LM-BP: Levenberg-Marquardt back-propagation	Hagan and Menhaj (1994)
	BP: back-propagation	Werbos (1982); Rumelhart et al. (1986); LeCun (1988)
SVM		
	SVR: support vector regression	Drucker et al. (1997)
	LS-SVR: least-squares support vector regression	Suykens et al. (2002)
	own developed training algorithm by author(s)	
Fuzzy Inference System		
	Mamdani	Mamdani and Assilian (1975)
	Sugeno	Sugeno (1985); Takagi and Sugeno (1985)
	hybrid	
Neuro-Fuzzy		
	ANFIS: adaptive neuro-fuzzy inference system	Jang (1993)
	DENFIS: dynamic evolving neural-fuzzy	Kasabov and Song (2002)
Ensemble		
	with homogeneous base learners	
	Bagging	Breiman (1996)
	RF: random forests	Breiman (2001a)
	GBM: gradient boosting machines	Friedman (2001)
	with heterogeneous base learners	

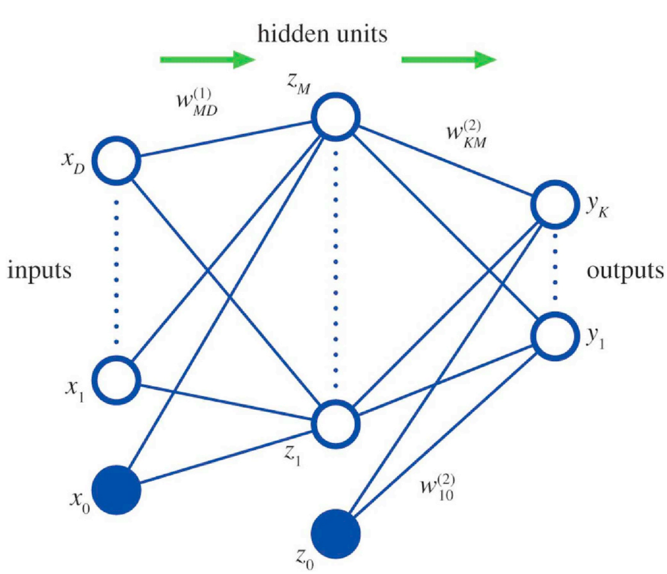

Fig. 3. Feed for war network with multilayer perceptron. Reprinted from Bishop (2013) licensed under CC BY 3.0.

propagating backwards the error between actual output and predicted output. In 1980's, some works (Werbos, 1982; Rumelhart et al., 1986; LeCun, 1988) reported the application of the already existing BP method as learning procedure for neural networks. Several optimization algorithms can be incorporated to the BP (e.g. conjugate gradient, gradient descent, quasi-Newton, bayesian regularization). A specific algorithm was frequently reported in the reviewed articles and, therefore, added as another *detail*, namely the Levenberg-Marquardt back-propagation (LM-BP) (Hagan and Menhaj, 1994). For a historical recapitulation of BP algorithms, the readers can refer to the review from Schmidhuber (2015).

Differently from BP algorithms, there are ANN whose weights and biases are drawn at random, leading to non-iterative algorithms (Cao et al., 2018). Some type of ANN with random weights are Random Neural Networks (Gelenbe, 1989), Random Vector Functional Link Networks (Pao and Takefuji, 1992; Pao et al., 1994), and Extreme Learning Machines (ELM) (Huang et al., 2004, 2006). ELM was specially reported in the ROP studies, being therefore considered as another training procedure in this review. In ELM, only the last layer is adapted to the training dataset, fitting the model to the problem. By assuming random weights in first layer and the respective biases of the hidden units, the training speed increases substantially, since the interactive steps from BP algorithms are avoided.

For more details about neural networks, the readers can refer to some textbooks (e.g. Bishop, 1995; Haykin, 2009).

SVM. Support vector machines (SVM) are very popular ML algorithms. Vapniks and colleagues were responsible for developing the fundamental basis of SVM (Cortes and Vapnik, 1995), culminating in a new learning theory, called the statistical learning theory, described in Vapnik's textbooks (Vapnik, 1998, 2000) or in other books (Bishop, 2006; Hastie et al., 2009; Deng et al., 2012). SVMs have high generalization ability, work well in very high-dimension feature space, as well as with small data samples. SVMs were developed as classification machines. Derivations of SVMs were performed for regression problems, resulting in so-called support vector regression (SVR) (Drucker et al., 1997). A comprehensive tutorial on SVR was written by Smola and Schölkopf (2004). Several adaptations of SVR were reported in the literature. A popular adaptation of SVM is the least squares support vector machines (LS-SVM), which was described in the book from Suykens et al. (2002) and can be used for regression purposes as well.

Fuzzy System Inference. The fuzzy logic set, developed by Zadeh (1965), provides a sophisticated way to deal with vagueness and imprecise information. The fuzzy logic can be employed to map nonlinear processes and patterns of a feature vector into a scalar output, as explained in the tutorial on fuzzy logic systems from Mendel (1995). Three common methods of deductive inference for fuzzy systems based

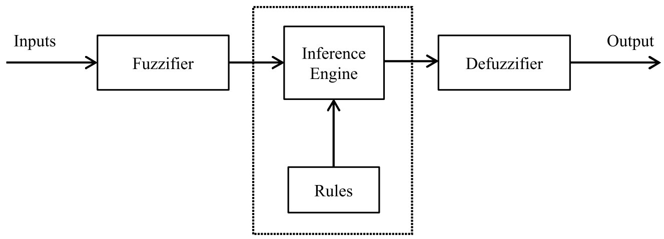

Fig. 4. Block diagram of fuzzy inference system (FIS) controller. Adapted from Mathur et al. (2016, its Fig. 2), under license of CC BY 4.0.

on linguistic rules are: Mamdami (Mamdani and Assilian, 1975) and Sugeno (Sugeno, 1985; Takagi and Sugeno, 1985), and Tsukamoto (Tsukamoto, 1979, *apud* Ross (2010), p. 148). In Table 1, two methods of FIS are shown, since the third method was not reported in the reviewed articles of ROP prediction. All the mentioned inference methods resembles an experienced-human operator who uses linguistic terms for control purposes. Such type of controller is based on *if-then* rules and the fuzzy set theory provides a mathematical way to compute the vagueness information and linguist terms. The rules can be provided by experts or extracted from numeric data (Ross, 2010). The main difference between both FIS is the output membership function (Mathur et al., 2016): Sugeno-type employs membership functions either linear or constant. Fig. 4 illustrates a FIS controller. The first step consists of transforming a crisp value[2] into fuzzy levels (fuzzifier). The inference engine, based on the *if-then* rules, employs the fuzzy logic to determine the output, which is finally transformed into crisp value in the defuzzifier. For more details about FIS, the readers can refer to the textbook from Ross (2010).

Neuro-Fuzzy. Neuro-fuzzy is a combination of two soft-computing techniques: ANN and fuzzy logic. To be more precise, FIS benefits from ANN mainly in the step of determining the membership functions of fuzzy levels (Suparta and Alhasa, 2016). The Adaptive-Network-based Fuzzy Inference System (ANFIS), proposed by Jang (1993), is a very representative method of neuro-fuzzy algorithms. ANFIS architecture has five layers, where the first and fourth layers contain an adaptive node, i.e. without any weight value (Suparta and Alhasa, 2016). To avoid some possible problems from BP (slow convergence rate and the possibility of being stuck in local minima), Jang has proposed a hybrid learning algorithm, which combines the gradient descent methods from BP with recursive least square estimator (Suparta and Alhasa, 2016). Another neuro-fuzzy reported in ROP studies was the Dynamic Evolving Neural-Fuzzy Inference System (DENFIS) (Kasabov and Song, 2002). DENFIS has the ability to effectively learn complex temporal sequences in an adaptive way. It can be employed in dynamic time series prediction.

Ensemble models. Ensemble models can convert the prediction of many "weak learners" (i.e. slightly better than random guessing) into strong learners (Hastie et al., 2009; Sagi and Rokach, 2018). The idea behind of ensemble methods is that the error of a single model (also called as base learner) can be compensated by other models. The theory of ensemble methods is bias-variance-covariance decomposition and a key element of ensembles is the diversity in several forms: data diversity, parameter diversity and structural diversity (Ren et al., 2016).

The ensemble models can be classified according to the type of their base learners. When an ensemble has all learners of the same type, this ensemble is said to be homogeneous. When the base learners are obtained from different learning techniques and combined into a single predictor (e.g. a combination of neural networks and decision trees), then it yields a heterogeneous ensemble (Mendes-Moreira et al., 2012; Ren et al., 2016).

[2] A crisp set is the antonym of fuzzy set. That is, in a crisp set, an element whether belongs to a set or not. In a fuzzy set, an element can belong to multiple sets at the same time.

Two classical homogeneous ensembles were reported in the ROP studies: random forests (RF) (Breiman, 2001a) and gradient boosting machines (GBM) (Friedman, 2001). Both ensemble models have as base learners tree models (see Hastie et al., 2009, Section 9.2). The main difference between them is the learning procedure of the trees as base learners. While the learning phase is performed in a parallel approach based on bagging procedure for RF (Breiman, 1996), the sequential learning adopted by GBM ensures that each base learner focus on the deficits of its previous learner. The differences between both methods and other ensemble techniques are covered in the reviews from Ren et al. (2016) and Sagi and Rokach (2018).

The employment of different base learners to form a strong predictor was reported in some ROP studies. Such studies were not formally classified by the authors as ensemble models. However, the current review classifies them as *ensemble (heterogeneous)*, because the authors combined several different base learners into a single one predictive model. Some example of such studies are Garavand and Esmaeilian (2015) and Ansari et al. (2017), which have developed a new algorithm/architecture for ROP modeling.

2.3.2. Reviewed articles

Table 2 summarizes all 53 works that employed ML techniques for ROP prediction in a chronological order. In the first column, the reference of each work is given. For some cases, a single work encompasses more than one reference, since such studies were either similar to each other or a continuation from previous papers. After the first column, some information about the predictive model for ROP is given, including the predicted output(s), amount of inputs feeding each ROP model, and the *method* and *detail* in parentheses. Table 2 covers also if ML techniques were compared with traditional ROP models. The last column informs whether each work showed how ROP models could be used in optimization tasks or not. If yes, then these studies are explained in Section 4.

3. Discussion on machine learning methods for ROP prediction

3.1. Methods comparison

In Table 2, we can observe that 18 out of 53 works compared machine learning with other techniques (traditional models). Most of them showed that the use of learning algorithms provided a more-accurate ROP prediction than the use of traditional models. The reason for that is the capability of those models to capture non-linear relationship among the variables. Only the work from Rahimzadeh et al. (2010) concluded that there was no overall winner in his comparison of ANN against BYM and Maurer model aided with genetic algorithm.

From those comparative studies, we would like to highlight the recent publication from Soares and Gray (2019). The authors performed an extensive analysis, testing a total of four traditional models (BIN, modified-BYM, HAR and MOT) and tree ML models (RF, SVM and ANN). In general, ML ROP models could predict significantly more accurately than traditional models. One interesting fact observed by the authors was a greater accuracy of ML models in comparison to traditional ROP models when as few as ten data points were available for model training, i.e. lower than expected.

It is possible to find out other authors suggesting that traditional ROP models can be employed simultaneously with ML models if the sample for training is small (Yavari et al., 2018). The authors stated that the traditional models require less data for training than the ML models and always provide a reasonable ROP prediction. For this reason, both different approaches could be employed at the same time when only few points for training are available. Combining traditional models and ML models into a single one predictive model can be achieved e.g. through aggregation functions, the core of ensemble models. To the best of our knowledge, there is no publication which investigated this combination.

Table 2
Reviewed publications using non-traditional model for ROP prediction and drilling optimization (chronological order).

Work	Prediction Method			Compared with Traditional Mod.			Optimized
	Output[a]	Inputs	Method (Detail)[b]	Yes?	To which one[c]?	The Best?	Yes?
Bilgesu et al. (1997)	ROP, bit life	10–6	ANN	No	–	–	No
Dashevskiy et al. (1999)	ROP*	4	ANN	No	–	–	Yes
Fonseca et al. (2006); Mendes et al. (2007)	ROP*	3	ANN (LM)	No	–	–	Yes
Akin and Karpuz (2008)	ROP, WOB, RPM	4	ANN (BP)	No	–	–	No
Moran et al. (2010)	ROP, bit wear	6	ANN (BP)	No	–	–	No
Moradi et al. (2010)	ROP	8	Fuzzy (hybrid)	Yes	GA-BYM (Bahari and Baradaran, 2007)	Fuzzy	No
Rahimzadeh et al. (2010)	ROP	8	ANN	Yes	BYM, M (GA)	no winner	No
Bataee and Mohseni (2011)	ROP	5	ANN (LM)	No	–	–	Yes
Arabjamaloei and Shadizadeh (2011)	ROP	7	ANN	Yes	BYM	ANN	Yes
AL-Rashidi (2011)	ROP, wear rate	5	ANN (BP)	No	–	–	No
Gidh et al. (2011, 2012)	ROP	NI	ANN	No	–	–	Yes
Amar and Ibrahim (2012); AlArfaj et al. (2012)	ROP	7	ANN (ELM, RBF)	Yes	BYM (Eren and Ozbayoglu, 2010)	ANN	No
Arabjamaloei and Karimi Dehkordi (2012)	ROP	11	ANN (BP), ANFIS, MLR	Yes	Mod-M	ANN	No
Esmaeili et al. (2012)	ROP	18–6	ANN (other)	No	–	–	No
Gandelman (2012)	ROP, WOB, RPM	7	ANN (BP, CG)	Yes	BYM, M, Y	ANN	Yes
Monazami et al. (2012)	ROP	13	ANN (LM)	No	–	–	No
Jahanbakhshi et al. (2012)	ROP	21	ANN (BP)	No	–	–	No
Jacinto et al. (2013)	ROP	4	DENFIS, BNI	No	–	–	No
Sui et al. (2013)	ROP*	6	Kalman filter	Yes	BYM	Kalman	Yes
Ning et al. (2013)	ROP	9	ANN (other)	No	–	–	No
Bataee et al. (2014)	ROP	5	ANN (LM)	Yes	BYM, BIN, Y	ANN	Yes
Zare and Shadizadeh (2014)	ROP	11	ANN (BP)	No	–	–	No
Basarir et al. (2014)	ROP	4	ANFIS, MR	No	–	–	No
Wallace et al. (2015)	ROP	NI	unclear	Yes	NI	ML	Yes
Hegde et al. (2015a)	ROP	unclear	Tree, Ensemble (Bagging, RF)	No		–	No
Duan et al. (2014, 2015)	ROP	6	ANN (BP, own)	No	–	–	Yes
Garavand and Esmaeilian (2015)	ROP	6	own ensemble (het.)	No	–	–	No
Bodaghi et al. (2015)	ROP	11	ANN (BP), SVR (own)	No	–	–	No
Valisevich et al. (2015)	ROP, wear	NI	ANN	No	–	–	Yes
Mantha and Samuel (2016)	ROP	4	ANN, Ensemble (GBM, RF), SVR, KNN	No	–	–	No
Shi et al. (2016)	ROP	10	ANN (LM, ELM, other)	No	–	–	No
Kahraman (2016)	ROP	5–4	ANN (LM), LR	No	–	–	No
Jiang and Samuel (2016)	ROP	5	ANN (BR)	unclear	Mod-M	unclear	Yes
Khosravanian et al. (2016a,c)	ROP	7	Fuzzy (unclear)	Yes	BYM	Fuzzy	No
Hegde et al. (2017); Hegde and Gray (2017)	ROP	4	Ensemble (RF), MR	Yes	BIN, MOT, HAR	RF	Yes
Tewari and Dwivedi (2017)	ROP	8	ANN (BP)	Yes	BYM	ANN	No
Ansari et al. (2017)	ROP	8	own ensemble (het.)	No	–	–	No
Eskandarian et al. (2017)	ROP	6	Ensemble (RF), ANN (MLP), LR	No	–	–	No
Diaz et al. (2017b,a)	ROP*	unclear	ANN (on-line)	No	–	–	No
Bezminabadi et al. (2017)	ROP	9–5	ANN (LM), MNR	No	–	–	No
Amer et al. (2017)	ROP	20	ANN (BP)	No	–	–	No
Ayoub et al. (2017)	ROP	5	ANFIS	Yes	BIN, BYM	ANFIS	No
Elkatatny (2018)	ROP	5	ANN (LM)	Yes	M, BIN, BYM	ANN	No
Hegde and Gray (2018)	ROP, TOB, MSE	4	Ensemble (RF)	No	–	–	Yes
Hegde et al. (2018a)	ROP	4	Ensemble (RF)	No	–	–	Yes
Diaz et al. (2018)	ROP	8	ANN (BP-resilient)	Yes	BYM	ANN	No
Yavari et al. (2018)	ROP	7	ANFIS	Yes	BYM, HAR	ANFIS	No
Momeni et al. (2018)	ROP	7	ANN (LM)	No	–	–	Yes
Anemangely et al. (2018)	ROP	5	ANN (BP-optimized)	No	–	–	No
Abbas et al. (2018)	ROP	17	ANN (BP)	No	–	–	No
Ahmed et al. (2019)	ROP	8-5	ANN (ELM, LM), SVR, LS-SVR	No	–	–	No
Soares and Gray (2019)	ROP	4	ANN (BP), SVM, Ensemble (RF)	Yes	4	RF	No
Ashrafi et al. (2019); Sabah et al. (2019)	ROP	8	ANN (BP, RBF, others), MR	No	–	–	No

[a] ROP: prediction; ROP*: forecasting; MSE: mechanical specific energy; TOB: torque on bit.
[b] Abbreviations of ML methods see Table 1 **b** MR: multivariate regression; MLR: multivariate linear regression; LR: linear regression; MNR: multivariate nonlinear regression.
[c] BIN: Bingham; BYM: Bourgoyne and Young's Model; HAR: Hareland and Rampersad; M: Maurer; Mod-M: modified Maurer; MOT: Motahhari; W: Warren; Y: Young.

3.2. Preferred ML methods applied to ROP prediction

Some of 53 reviewed works (Table 2) assessed different ML methods in their ROP study (e.g. Mantha and Samuel, 2016; Ahmed et al., 2019; Soares and Gray, 2019), resulting in a total of 70 ML models obtained. ANNs were employed 39 times, followed by homogeneous ensembles (e.g. RF, GBM), which were used 9 times. Neuro-fuzzy methods, especially ANFIS, appeared seven times in ROP prediction studies and SVM tailored for regression problems was applied four times. Other methods (including fuzzy-inference system and heterogeneous ensembles) were

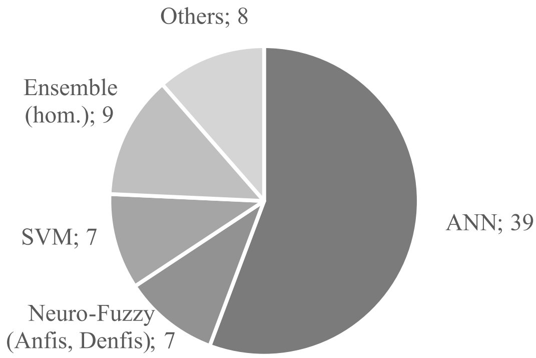

Fig. 5. ML methods employed in ROP studies.

Table 3
Most common softwares/programming language employed in ROP studies based on data-driven models.

Software/Programming Language	Amount
N.I. or unclear	26
Matlab	17
Python	4
R	2
Others	4

reported eight times, as shown in Fig. 5. It is also interesting to analyze the most employed software/programming language in 53 works, as summarized in Table 3.

3.3. Improving the conventional training algorithms for neural networks

As seen in Table 2, back-propagation (BP) algorithms are often used to adjust weights and biases of ANNs and can achieve in general good results. However, such algorithms can have some drawbacks, especially slow error convergence and the possibility of becoming trapped in local minima (Lee et al., 1991; Haykin, 1994; Bishop, 2006; Sabah et al., 2019). Some alternatives were proposed in the literature.

One example was performed by Duan et al. (2014) who improved the traditional BP algorithm by adapting some parameters employed in the learning algorithms. In this work, momentum factor and the dynamical learning rate were introduced. This approach could improve the speed converging and the generalization capability of ANNs trained for ROP prediction.

Another alternative to the traditional BP methods is the use of evolutionary optimization algorithms (Anemangely et al., 2018; Sabah et al., 2019; Ashrafi et al., 2019), such as genetic algorithm or particle swarm optimization. To understand this approach, the readers can refer e.g. to Ashrafi et al. (2019), who performed an extensive comparison between traditional BP algorithms and so-called "hybrid artificial neural networks" (hybrid-ANNs). These hybrid-ANNs consisted of traditional ANNs (e.g. MLP or RBF) trained with evolutionary algorithms as optimization method to adapt the weights and biases. A total of four evolutionary methods were employed: genetic algorithm (GA), particle swarm optimization (PSO), biogeography-based optimizer (BBO) and imperialist competitive algorithm (ICA). The authors concluded that hybrid-ANNs outperformed the ANNs trained with BP algorithms.

Those examples indicate the possibility of tuning the predictive model by adapting the traditional learning procedures of ANNs. If desired, it is possible to improve the accuracy of ROP models by giving a special attention to the learning methods. This idea can be also extended to other ML methods. Bodaghi et al. (2015) used also evolutionary optimization algorithms to optimize SVMs for ROP prediction.

3.4. Alternatives to black-box models

In general, the use of ANN resulted in predictive models with good generalization capability. However, these models are usually complex. For that reason, ANNs are commonly referred as *black-boxes*, because it is not easy to understand them. As alternatives, some works utilized techniques from which understandable rules could be extracted. For example, Basarir et al. (2014); Yavari et al. (2018) modeled ROP with the neuro-fuzzy technique, exploiting the advantage of fuzzy logic to mathematically describe linguistic terms. With this method, the authors could understand some simple rules about how the drilling variables affected the ROP. Jacinto et al. (2013) used the neuro-fuzzy methods, called DENFIS, in order to obtain understandable rules. The fuzzy inference systems were also reported as an alternative to obtain understandable rules (Khosravanian et al., 2016a,c).

The extraction of rules could also be obtained with a different ML method. One of techniques applied by Eskandarian et al. (2017) to model ROP was random forests. This method, similar to ANN, results in predictive model not easy to understand. However, Eskandarian et al. (2017) used a R package, called "inTrees" (Deng, 2018), which allowed them to extract some general rules, splitting the ROP values into three levels: low, medium and high.

The rule extraction is one way to assess the influence of the drilling variables on the ROP. Another way is to perform sensitive analysis, i.e. varying some drilling variables, while others remain unchanged. In the study from Eskandarian et al. (2017), this approach was also described, which helped the authors to understand the relationship among the variables. By analyzing sensitive plots, the authors determined the range of controllable drilling parameters at which ROP is near to the maximum point. Other studies (Akin and Karpuz, 2008; Arabjamaloei and Shadizadeh, 2011; Zare and Shadizadeh, 2014) have performed similar analysis as well.

3.5. The most common inputs

Fig. 6 depicts the amount of inputs used in data-driven models for ROP prediction. As seen, there is a tendency of simplifying the predictive models by selecting a sub-set of inputs, instead of all drilling variables available. Such sub-sets were in general smaller than those required by some traditional models (e.g. BYM). Shallow ML models with three or four inputs were obtained in 10 works, and models with five or six inputs were reported in 12 works. Only four works reported the use of ROP models with 13 or more inputs. It is important to note that some works investigated several different combinations of inputs, being therefore grouped in a particular class of "tried more than one combination".

It was also possible to compute the most common variables selected as inputs for ML models. For that, those works that not informed or clearly reported the variables used could not taken into account in this analysis, as well as those works that investigated more than one combination of inputs. As consequence, we computed the most common

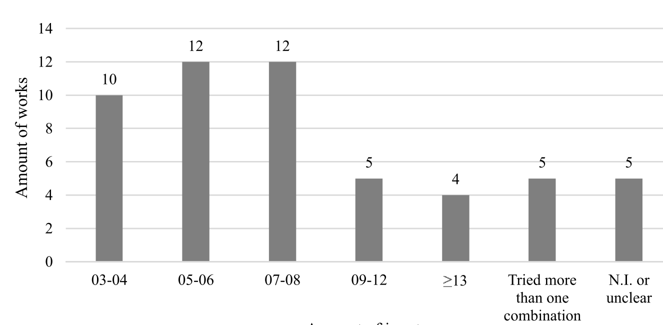

Fig. 6. Amount of inputs employed to feed ROP data-driven models.

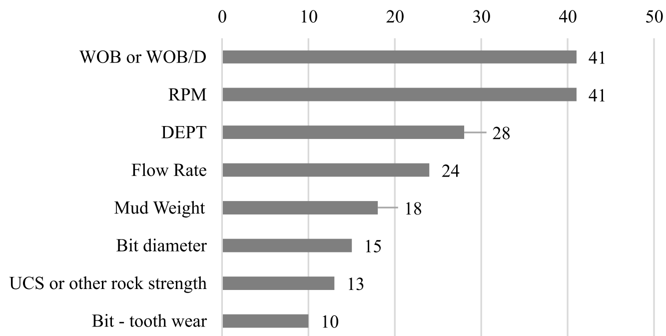

Fig. 7. Frequency of inputs employed to feed ROP data-driven models, considering 43 from all 53 reviewed works.

inputs considering 43 from all 53 reviewed works. Fig. 7 depicts the eight most common inputs obtained in this analysis.

Similar from the traditional model, the WOB and RPM were used in almost all data-driven ROP models. The drilling fluid and hydraulics play also a very important role in hole cleaning efficiency, and on progress of drilling activities itself. Therefore, several works fed their data-driven models to represent the fluids, especially the flow rate and mud weight. Other important parameters, e.g. bit tooth wear and rock formation strength, were reported in ROP studies, but not as so often as the mentioned surface drilling variables. A possible reason for that is the challenge associated with measuring and monitoring in real-time downhole parameters in comparison to drilling variables measured on surface.

3.6. Approaches in using the lithology and rock properties in ROP prediction

It was found four possible approaches regarding the use of petrophysical properties of geological formation (including rock mechanical properties) on ROP prediction. The first approach does not take into account any parameter related to geological formation or lithology. In the second approach only part of the inputs are related somehow to the drilled formation. The third approach emphasizes the importance of the formation on ROP prediction. In the fourth approach, drilling data is split for each geological formation or lithology.

Bilgesu et al. (1997) reported the possibility to obtain reliable predictive ROP models by only using surface parameters (weight on bit, drill bit rotation, mud flow rate, and so on). The authors trained two networks on the data from rig floor simulator. In one case, the first layer of ANN had 10 inputs, and, in another case, 6 inputs after excluding bit tooth, bearing wear, formation abrasiveness and drillability. Even without any parameter related to the drilled formation or to bit status, the use of neural networks resulted in a good prediction capability for the ROP values. Other studies predicted the ROP also without any parameter related to formation (Bataee et al., 2014; Eskandarian et al., 2017).

Many examples of the second approach can be found (Moran et al., 2010; Arabjamaloei and Shadizadeh, 2011; Amar and Ibrahim, 2012; Ansari et al., 2017; Hegde et al., 2017; Bezminabadi et al., 2017; Diaz et al., 2018). In this approach, some parameters from downhole are used as inputs to predict ROP values, including variables related to formation properties, lithology or drill-bit condition (e.g. tooth wear). A combination of surface drilling data and petrophysical logs provides accurate ROP models. However, the variables related to rock formation (e.g. the rock strength or other formation properties) are so not emphasized, as in the third approach, where the use of petrophysical logs is highlighted. The third approach is also often reported in the literature (Akin and Karpuz, 2008; Basarir et al., 2014; Anemangely et al., 2018; Ashrafi et al., 2019). A good example from this third approach is the work from Akin and Karpuz (2008). The authors trained an ANN for ROP prediction by using only four inputs, which were: depth, formation type, rock quality designation and discontinuity frequency index.

The fourth approach is based on splitting the drilling data for each formation. In this case, there is a unique ROP model for each lithology. Gandelman (2012) built neural networks to predict ROP with high accuracy by having a unique trained model for each lithology. A similar approach was employed by Hegde et al. (2017), but the authors employed random forests to predict the ROP instead of neural networks. Another difference between both studies is the data partition. While Gandelman (2012) randomly partitioned the drilling data for each formation into training and testing sets, Hegde et al. (2017) used the initial phase of drilling to train the ROP model. The rest of drilling data of each section was used to validate the ROP models; this process is able to predict ROP in real-time operations, and is, therefore, detailed in following Section 3.8.

3.7. Feature selection

Feature selection is a process of selecting a subset of features (input variables) to feed a machine learning algorithm. By selecting a small subset of relevant features, the following benefits can be achieved: faster learning process, simpler model and better accuracy (Xue et al., 2016). However, the main drawback is that the search space can be very huge, so that finding the optima solution may not be computational feasible. Therefore, there are different techniques to overcome this challenge, by transforming the selection of subset of features into an optimization problem.

One of these techniques is called greedy search. Guyon and Elisseeff (2003) stated that greedy search strategies (forward selection and backward elimination) can be computationally advantageous and robust against overfitting. The forward selection incorporates progressively the variables that improve the accuracy of model. In backward elimination, the model starts with all inputs, and then eliminates progressively the least promising inputs. Some studies in ROP prediction employed sequential methods for feature selection. For example, Anemangely et al. (2018) employed the sequential method, called "plus-l-take-away-r".

There are also methods able to rank the features, providing the importance of each input in the output estimation. Eskandarian et al. (2017) showed a comprehensive way of assessing the most important parameters for modeling the ROP. In that case, the authors used an open-source package called "fscaret" (Szlek and Mendyk, 2018) (written in R-programming language), which allowed them to find out the importance and ranking of input parameters. As consequence, the authors assessed the most important drilling parameters in ROP prediction. Another similar example was performed by Hegde et al. (2017) who also assessed the variable importance of each drilling variable with random forests.

Besides the greedy search strategies and ranking methods, it is possible to employ evolutionary computation approaches (e.g. genetic algorithm or particle swarm optimization) to obtain an optimum subset of features (Xue et al., 2016). These approaches have the ability to reach the global optima, or to obtain a result that is near to the optimum point. In the studies for ROP prediction, genetic algorithm was applied in order to determine the best subset of inputs (Ashrafi et al., 2019; Sabah et al., 2019). Expect for these tries, no other works employed evolutionary algorithms for selecting the features of ROP models.

Differently from feature selection based on optimization algorithms, another approach commonly found in the reviewed articles was a preselection of inputs, based on drilling knowledge. An example is the work from Shi et al. (2016), who performed an extensive discussion on drill bit mechanisms in order to select the inputs for ANN. In their study, a total of 10 inputs were selected, including surface measurements, bit properties, hydraulics variables and formation properties. It is also possible to find other works that approached similarly in the selection of inputs (Jahanbakhshi et al., 2012; Ning et al., 2013).

Sometimes, traditional ROP models served also as a basis for selecting the inputs. For example, Tewari and Dwivedi (2017) developed ANN with the same eight inputs from BYM.

3.8. Approaches for real-time prediction

A common approach of those works from Table 2 was to select training and testing dataset with conventional approaches (e.g. employing holdout validation or k-fold cross validation). The holdout validation consists basically in splitting the drilling data into two datasets: one for training (when required, part of this training set was used in the validation process during the training) and another for testing, which assesses the generalization capability of the predictive model. In general, this data partition is randomly carried out. The k-fold cross validation (Browne, 2000; Berrar, 2019) is a more sophisticated method than the holdout method. This procedure was applied by Eskandarian et al. (2017) in a comprehensive study on ROP prediction. In the cross-validation approach, the data is split into k parts. In the first round, one of k-folds is selected as testing dataset and all others are used for the training. In the second round, a different k-fold is used as testing dataset and, as consequence, the left-out folds are employed as training set. This process goes until all k-folds were selected as testing set. Finally, it is possible to average the evaluation metric(s) from each round. This makes the cross-validation approach more robust than the holdout method. It is a common practice to select k = 10, resulting in so-called 10-fold cross-validation, as illustrated in Fig. 8. When the amount of folds is equal to the sample size, then we have a special case, called as leave one out (Wong, 2015).

Differently from the conventional data partition methods, Hegde et al. (2017) proposed a novel way to partition the data into training and testing dataset, so that it can be used in a real-time environment. In their method, the trained model can be employed to optimize in real-time the drilling activities by finding out the optimum values of the controllable drilling variables, such as weight on bit and bit rotation speed, which maximize the ROP (Hegde and Gray, 2017).

The proposed method by Hegde et al. (2017) is based on splitting the drilling data into several datasets for each lithology. In the initial phase of a formation or lithology section, the drilling data are used to build the ROP model, i.e., the training dataset is the initial data of this section. After training an expert model based on ML models (e.g. ANN or RF) for this respective formation, this model is able to predict the ROP for the rest of the length of this section with better accuracy than the traditional ROP models. When a new formation is encountered, then a new model is trained until it reaches a good prediction capability for the section ahead, and this process goes on. When using RF, it is possible to predict the generalization error of the model by using out-of-bag error prediction. An illustration of this process is given in Fig. 9. Recently, a detailed study about this approach was published by Soares and Gray (2019).

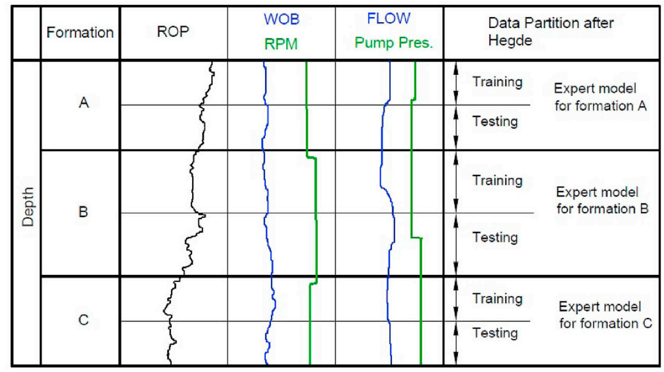

Fig. 9. Data Partition developed by Hegde and colleagues. Reproduced after Hegde et al. (2017).

The previous works developed at Hegde and colleagues employed off-line learning. Another approach which enables to obtain a ROP model in real-time environment is through on-line learning as briefly shown by Diaz et al. (2017a,b). Except for these works, the employment of on-line learning was not reported in the literature.

3.9. Drilling data source and its influence on selecting ML methods

It was possible to observe that the most of ROP studies was carried out with field data, which normally comes from: drilling daily reports (Ansari et al., 2017; Eskandarian et al., 2017) or real-time drilling data (mud logging and well logging with petrophysical information) (Gandelman, 2012; Hegde et al., 2017; Diaz et al., 2018; Ashrafi et al., 2019). A not very common strategy was the use of experimental data from laboratory drilling tests (Kahraman, 2016) or mini-rig facility (Esmaeili et al., 2012).

The drilling data from daily reports (or from drill bit reports) are important source of information. However, the downside when using daily reports is that only few observations are available to build predictive models based on machine learning. For example, if it is required a dataset with more than hundred points, it will not always be possible to analyze drilling daily reports from an individual well. While Diaz et al. (2018) used real-time drilling data from a drilled well with length of 4.6 km in South Korea with a total of 7043 observations, Ansari et al. (2017) used drilling daily reports and needed to gather information

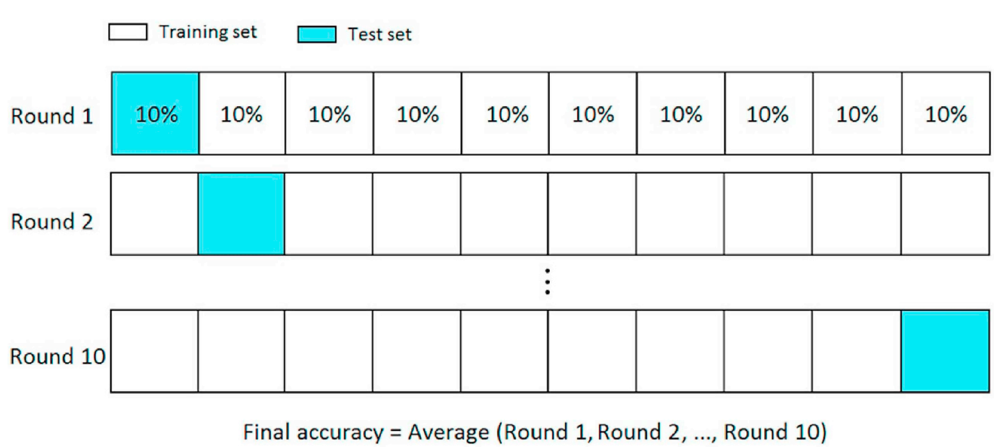

Fig. 8. Schematic of 10 cross-fold validation. Reprinted from Eskandarian et al. (2017), with permission from Elsevier.

from 19 wells drilled in Persian Gulf in order to have available 248 points. The difference in the amount of observations has a direct impact on training the predictive ROP model.

The selection of ML method for obtaining ROP models depends on the amount of observations available for the training phase. As consequence, the type of drilling data available influences on the selection of ML method. For datasets with small amount of observation, support vector machines can be a solution (Bodaghi et al., 2015; Ansari et al., 2017; Agwu et al., 2018). The neuro-fuzzy techniques (ANFIS) was also reported as an alternative for modeling ROP with small datasets (Yavari et al., 2018).

The real-time drilling data (RTDD) can be stored in time domain or in depth domain. The depth domain is the common way of geologists to plot the logs against the depth. Using the depth as index simplifies the data preparation, since only observations while drilling are used in modeling. Some sampling rates found of RTDD when stored in depth domain were 0.25 ft (Hegde et al., 2017; Hegde and Gray, 2017) or 0.5 ft (Nascimento, 2016). Recently, Ashrafi et al. (2019) employed mud logging and well logging in a study for ROP prediction with ANN. The authors reported different sampling rates for each logging. The mud logging, which encompassed operational parameters (WOB, RPM, pump pressure etc.), was recorded at a sampling rate of 50 cm and the well logging (gamma ray, wave velocity and other petrophysical parameters) was recorded at a sampling rate of 0.5 ft (15.24 cm). For this reason, the mud logging was compressed from 15.24 cm to 50 cm.

When using drilling data in time domain, the raw recording can also have data stored with all drilling activities (e.g. tripping, making connection, ream and wash, circulating, drilling in rotating or sliding) Thonhauser (2004). Since the ROP prediction aims to model how some drilling variables affect the drilling performance, it is important to identify the main state: drilling. To properly identify the main drilling activities (e.g., rotary drilling, sliding, tripping connection), some works (Tavares, 2006; Mathis et al., 2007) reported so-called automated operation recognition systems.

One of the few works that used RTDD in time domain to predict drilling parameters was carried out by Fruhwirth et al. (2006), who obtained predictive models for the pump pressure. The need of recognizing different drilling operations was reported by the authors. However, deeper details about operation recognition were not provided in that work. Gandelman (2012) reported also the use of RTDD in time domain in his study of ROP prediction and optimization. The author detailed the data pre-treatment process employed in order to validate the drill-curves, but he did not mention about the existence of non-productive time in the original datasets. Some sampling rates of drill curves employed in the previous works were 1 s (Fruhwirth et al., 2006), 5 s (Donne, 2017), or 15–30 s (Gandelman, 2012).

3.10. Handling measurement errors

Some researches (Arnaout et al., 2013; Otalvora et al., 2016) reported some common measurement errors in drilling data, and proposed methods to determine quality indexes in real-time of streaming data. For our purpose, the best description of the real-time measurement problems was given by Arnaout et al. (2013). The authors classified the measurement errors in the three categories:

- Time problems: missing timestamp, invalid time format, wrong time zone, incorrect or no time synchronization;
- Depth Problems: bit depth/hole depth resets, heave compensation (floating rigs);
- Data Channel Problems: wrong channel description, wrong units, calibration, gaps (missing values and null values), different frequencies, outliers and drifting values.

Most of the reviewed works drew their attention to two types of problems from the category "Data Channel Problems" which are: outliers[3] and gaps. To treat the outliers, it is required first to identify them, and then to apply a suitable treatment for them (i.e. replace the identified outliers by a suitable value or not consider the whole observation with an outlier). The problem of the gaps (missing values or null values) can be very challenging to be treated, because the original values is unknown and any attempt to treat or impute the data has high risk (Little and Rubin, 2002).

Noises (i.e. error in data), or considering a broader concept of outliers (i.e. discordant), make difficult the task of obtaining machine learning models, and increases the training time (Quinlan, 1986; Garćia et al., 2013). Therefore, the reduction of noises in data brings together benefits to the learning process of data-driven models (Garćia et al., 2013). There are many methods to identify and treat both outliers and noises (Salgado et al., 2016; Aggarwal, 2013). Some different approaches to treat measurement errors in drilling datasets were found in the literature, which are:

- no treatment - raw recording is employed to obtain ROP models;
- manual approach - e.g. exclusion of incomplete observation and visible outliers;
- automatic approach - identification and treatment of outliers and missing values with systematic methods (statistical methods or digital signal processing).

In this review, it was found that some works mentioned how outliers were treated. This type of error can be handled: with a manual approach (Arabjamaloei and Shadizadeh, 2011; Gandelman, 2012; Hegde and Gray, 2017), or with filter to smooth the recording, eliminating noises (Diaz et al., 2018; Anemangely et al., 2018; Ashrafi et al., 2019). A manual approach is subject to human interpretation, and can be very time-consuming job, especially when analyzing many drill curves. However, it can be suitable when analyzing a small dataset, because those observations which are visible outliers can be simple removed from the dataset, as done by Hegde and Gray (2017).

Instead of the manual approach, the automate approach employs robust techniques to identify whether an observation is indeed an outlier or not. Few works applied filter to smooth drill recording, reducing noises in data, as consequence some outliers. Some filters employed were low-pass parabolic filter (Diaz et al., 2018), and Savitzky-Golay filter (Anemangely et al., 2018; Ashrafi et al., 2019). Momeni et al. (2018) employed recurrent neural network to remove noise effects and detect possible problems in drilling data.

Another important issue is the completeness of drilling data. It is not rare to encounter drilling data with "gaps" in their recordings. A gap is when a failure in data transmission occurs, so that measurements of one or more drilling variables are not transmitted for a period of time (Arnaout et al., 2013). Having this in mind, it is not difficulty to suppose that some of 53 works had faced this type of problem when training a ROP model. In general, this problem was omitted. One of few works that reported the problem of missing data was from Gandelman (2012), who adopted the approach of complete-cases. The downside of this approach is to cause a substantial loss of information (Little and Rubin, 2002). In addition, Gandelman (2012) detailed the data pretreatment process employed to validate the drill recordings, eliminating possible error measurements (missing values, outliers and invalid observations). To validate the observations, rules were employed based on

[3] Sometimes, the terms noises and outliers can be confusing. The readers can refer to Salgado et al. (2016) in order to understand the difference between them. To sum up, noises are mislabeled examples or errors in the values of attributes. Outlier is a broader concept, because this term includes errors and discordant data (also called *abnormalities, discordants, deviants and anomalies*). Such discordant data is not necessarily an error in the measurement, but it can be a deviation from a population.

knowledge of drilling engineering. For example, one rule employed by the author was the following: ROP could only be above zero, if, and only if, the WOB and RPM were both above zero. Otherwise, this observation would not be considered as a valid one. The author mentioned the reduction of dataset from 43 524 observations (200 h of drilling activities) to 23 949 valid observations (160 h).

4. Drilling optimization based on predictive models

4.1. Optimization overview

A common approach of drilling optimization studies is to determine the optimum combination of e.g. bit weight and rotary speed that maximizes the ROP. In this case, maximization of ROP is the only objective function, leading to a single-objective optimization problem. If multiple drilling performance indicators are taken into account, e.g. maximizing ROP and maximizing bit-life, a multi-objective optimization needs to be formulated. This is because more than one objective is considered, and they may be contradictory to each other.

Here, we present the general frameworks of the single-objective optimization problem and the multi-objective optimization problem. After that, a discussion is performed on the works from Table 2 which employed the predictive models for drilling optimization.

4.1.1. Single-objective optimization

Depending on the problem, we seek to either maximize or minimize an objective function $f(\boldsymbol{x})$. Cui et al. (2017) showed that the general problem of single-objective optimization can be defined as a minimization problem, because, with the transform $\max f(\boldsymbol{x}) \Leftrightarrow \min(-f(\boldsymbol{x}))$, it is possible to transform a maximization problem into a minimization problem and vice verse.

According to Chiandussi et al. (2012), a general single-objective optimization can be defined as the minimization objective function $f(\boldsymbol{x})$, subject to inequality constraints $g_i(\boldsymbol{x}) \leq 0$, $i = \{1,2,\ldots,p\}$ and equality constraints $h_j(\boldsymbol{x}) = 0$, $j = \{1,2,\ldots,q\}$. The universe Ω denotes the space of all possible values of the decision variable $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$, respecting the constraints.

Let us suppose that we wish to determine the optimum solution $\boldsymbol{x}^*$ that minimizes a given objective function *f*. This function may have several local minima, so that the employed optimization algorithm may achieve a local minimum point of *f* instead of converging to the global optima. The Global Optimization methods try to find the global optimum solution, avoiding the problem of being trapped in local minima (Chiandussi et al., 2012).

4.1.2. Multi-objective optimization

A multi-objective problem aims to optimize simultaneously multiple objective functions, and can be formulated by the equation (Zhou et al., 2011):

$$\begin{aligned} \text{minimize} \quad & F(\boldsymbol{x}) = [f_1(\boldsymbol{x}), f_2(\boldsymbol{x}), \ldots, f_m(\boldsymbol{x})]^T \\ \text{s.t.} \quad & \boldsymbol{x} \in \Omega \end{aligned} \tag{10}$$

where the decision variable $\boldsymbol{x}$ belongs to the decision space Ω. In the case of m-objective functions, the objective space belongs to a $m-$dimensional vector space $\mathbb{R}^m$.

The objectives in equation (10) are often contradictory to each other, i.e. an improvement of one objective may lead to deterioration of another. As consequence, there is no single optimum solution able to optimize all objectives simultaneously. In multi-objective problems, a set of optimal solutions can be obtained instead of a single one solution. This set is called *Pareto optimal solutions* (Zhou et al., 2011). Several works explain the concept of Pareto optimality (Zhou et al., 2011; Chiandussi et al., 2012; Antonio and Coello, 2017; Cui et al., 2017).

A general multi-objective optimization problem can be subject to inequalities and equalities (Antonio and Coello, 2017; Cui et al., 2017; Zhou et al., 2011). In this case, the search space Ω can be formulated as follows (Zhou et al., 2011):

$$\Omega = \begin{cases} g_i(\boldsymbol{x}) \leq 0, & i = \{1,2,\ldots,p\} \\ h_j(\boldsymbol{x}) = 0, & j = \{1,2,\ldots,q\} \\ x_i^{min} \leq x_i \leq x_i^{max}, & i = \{1,2,\ldots,n\} \end{cases} \tag{11}$$

where Ω is a $n-$dimensional search space for the decision variable $\boldsymbol{x}$, determined by the upper bound $\boldsymbol{x}^{max} = [x_1^{max}, x_2^{max}, \ldots, x_n^{max}]^T$ and the lower bound $\boldsymbol{x}^{min} = [x_1^{min}, x_2^{min}, \ldots, x_n^{min}]^T$, p inequalities $g_i(\boldsymbol{x}) \leq 0$, $i = \{1,2,\ldots,p\}$, and q equalities $h_j(\boldsymbol{x}) = 0$, $j = \{1,2,\ldots,q\}$. These constraints lead to two regions: a feasible region and infeasible region. For the special case $p = q = 0$, the multi-objective optimization problem is said to be unconstrained (Cui et al., 2017).

4.1.3. Decision making techniques and searching methods

In multi-objective optimization problem, the decision maker needs to decide what kind of trade-offs are more suitable. There are several techniques for that. One of them is the simple idea of combining several criteria into a single-one. Such aggregation techniques (e.g. global criterion method, sum weighted or ε-constraint) are normally used in optimization problems in engineering. The main advantage of these aggregation techniques is the simplicity in finding an optimum solution, since a multi-objective problem is transformed into a single-objective. As consequence, any single-objective solver can be employed. The main drawback is the difficulty in giving appropriate importance for each criterion (Chiandussi et al., 2012).

It is important to note, that the solution of optimization problems involves two main steps: one is the mapping the decision space into the objective space and another is the process of decision making (especially for multi-objective problems). The mapping process can be done with either analytical method or numerical method. While analytical method relies on mathematically proof and gradient information, numerical methods can be applied for black-box problems (Cui et al., 2017). Some searching techniques are: random search, genetic algorithm (GA), particle swarm optimization (PSO), differential evolution algorithm (DE). For more details about the (dis)advantages of some classical intelligent optimization methods, the readers can refer to a comprehensive review from Cui et al. (2017).

4.2. Reviewed articles - single-objective optimization

Attempts to formal optimize drilling variables are not new. Tansev (1975) employed multiple logarithmic regression to model ROP and bit life. With both prediction models, the controllable variables could be optimized in order to minimize the cost per foot, subject to controllable variables' bounds. Drilling optimization still attracts the attention of researchers and oil industry. In the last decade, some works employed predictive models to formulate optimization problems. A summary about 14 works is given in Table 4. It is interesting to analyze which drilling variables have been selected to be optimized in these studies (Table 4). Most of these works have selected three controllable variables to be optimized. Normally, the WOB and RPM were selected to be optimized, and the third variable was related to bit hydraulics (e.g. mud flow rate) or drilling fluid properties (e.g. mud weight).

A common objective function was the maximization of ROP. From 12 works that clearly presented the objective function, maximization of ROP were solely employed in seven works, and in other two in comparative studies (Awotunde and Mutasiem, 2014; Hegde et al., 2018b). Differently, Bahari and Seyed (2009) studied the optimization the controllable drilling variables (WOB and RPM) by minimizing the cost per foot drilled, subject to operating ranges recommended by manufacturing companies and limited to the maximum applicable mechanical energy. A similar approach was adopted by Hankins et al. (2015).

Hegde and Gray (2018) investigated which objective function would be more appropriate for drilling optimization. For that, the authors compared different objective functions in a single-objective

Table 4
The use ROP models to optimize the drilling variables, considering a single-objective problem (chronological order).

Work	ROP model	Obj. Function	Method	Opt. Parameters
Bahari and Seyed (2009) [a]	BYM	min Cost	trust region	WOB, RPM
Arabjamaloei and Shadizadeh (2011) [b]	ANN	max ROP	genetic algorithm	WOB, RPM, bit hydraulic
Bataee and Mohseni (2011); Bataee et al. (2014)	ANN	max ROP	genetic algorithm	WOB, RPM, mud weight
Gidh et al. (2011, 2012)	ANN	max Bit Life	unclear	unclear
Awotunde and Mutasiem (2014) [c]	Warren	max ROP min Total Time	differential evolution	WOB, RPM, mud flow rate
Duan et al. (2015)	ANN	max ROP	particle swarm optimization	WOB, RPM
Hankins et al. (2015)	HAR	min Cost	unclear	WOB, RPM, hydraulics, bit
Valisevich et al. (2015)	ANN	unclear	unclear	WOB, RPM
Wallace et al. (2015)	unclear	unclear	max ROP	unclear
Jiang and Samuel (2016)	ANN	max ROP	ant colony optimization	WOB, RPM, mud flow rate
Hegde and Gray (2017)	RF	max ROP	brute force	WOB, RPM, mud flow rate
Hegde and Gray (2018) [c]	RF	max ROP min Torque min SE	particle swarm optimization	WOB, RPM, mud flow rate
Hegde et al. (2018a)	RF	max ROP	Eyeball Random Search Simplex particle swarm optimization differential evolution	WOB, RPM, mud flow rate
Momeni et al. (2018)	ANN	max ROP	genetic algorithm	Bit

[a] considering the optimization problem of operating conditions WOB and RPM. The authors studied also the optimization of hydraulics and bit tooth wear.
[b] Only added those variables whose optimization were detailed.
[c] These studies investigated different objective functions in the single-objective optimization framework.

optimization framework: minimization of torque, minimization of the specific energy, or maximization of ROP. The authors concluded that the minimizing the specific energy was a better approach, because it could find a tradeoff between increasing the ROP and improving the drilling efficiency. It is interesting to describe how Hegde and Gray (2018) computed the influence of varying the controllable drilling variables (WOB, RPM, mud flow) on the (mechanical) specific energy (MSE). The authors called their approach as coupled one. It consisted of building predictive model for ROP and torque, and then coupling these predictive models with the analytical equation from Teale (1965). This coupled MSE model has been earlier adopted by Gandelman (2012), as explained in the next Section 4.3.

Awotunde and Mutasiem (2014) carried out also a comparative stud. In this case, two objectives functions in single-optimization framework were compared against each other: maximization of ROP and minimization of total-time (drilling time, tripping and bit-change time). At shallow depths, the maximization of ROP yields to lowest total time, but, at deeper depths, the minimization of total time yields to the lowest overall time.

Another interesting work was developed by Hegde et al. (2018a). Several optimization techniques for the problem of maximization of ROP were analyzed, including: eyeball method, random search, simplex (amoeba) method, differential evolution, and PSO. In this study, the goal was to assess the computational effort and ability in finding the global optima from five optimization methods. The authors concluded that the Simplex algorithm presented the best tradeoff between running time and the ability in finding a good solution. PSO was better in finding optimum solutions than other methods, however the computational effort was higher. However, only the studies from Hegde et al. (2018a) study and Gandelman (2012) compared at least two optimization techniques against each other. There is clearly a need to investigate which optimization algorithm is more suitable to drilling optimization purposes.

4.3. Reviewed articles - multi-objective optimization

As described by Lyons and Plisga (2004), the task of drilling optimization is very complex, because it involves not only reducing the total drilling time. Others aspects should be considered, namely, production capacity and health, safety and environment (HSE). Therefore, a multi-objective optimization approach seems interesting, because it makes possible to optimize a very complex activity which requires several metrics for its monitoring. However, not many works proposed a multi-objective optimization problem for the drilling activities, expect for the following works: Gandelman (2012); Guria et al. (2014); Payette et al. (2017).

Gandelman (2012) developed neural networks for different Brazilian rock formations (12 in total), to be used latter in real-time optimization of mechanical drilling variables (WOB and RPM). His optimization approach considered two objective functions at the same time: min SE and min E, where MSE represented the mechanical specific energy, and E a function error between the predicted $ROP_{predict} = f(WOB, RPM)$ and a desired ROP_{set}, which the driller could set. The goal of this optimization was to determine the optimum WOB and RPM that achieved a desired ROP, trying, at the same time, to spend as less as possible energy in the process. This method could be classified as ε-constraint optimization, where the goal was to minimize the MSE subject to a constraint $(ROP_{predict} - ROP_{set}) \leq \varepsilon$, where the decision variables $\boldsymbol{x} \in \Omega$ were within the design limits of drill equipment.

Among other things, Gandelman (2012) tested two optimization methods: particle swarm optimization, and an own developed exhaustive search. This last approach combined *if-then* rules with the grid search in the feasible region Ω of the decision variables. The author selected the bit weight and rotary speedy as the variables to be optimized. For each iteration, 195 000 combinations of WOB and RPM were tested. The author concluded that this method was the only one able to determine the optimum combination of WOB and RPM.

In the drilling advisory presented by Payette et al. (2017), no ROP model was actually obtained, so that this work actually goes beyond the scope of this review. However, it is worth of mentioning the adopted strategy to simplify the multi-objective optimization problem. Up to three different objectives (ROP, SE and stick-slip risk) are aggregated into a single scalar function. It facilitates the optimization task. This report differs from all works, because it was the only one to publish the results from implementing real-time optimization of controllable drilling variables, especially WOB, RPM and mud flow.

The works from Gandelman (2012); Payette et al. (2017) employed decision making techniques, which transform a multi-objective optimization problem into a single-objective problem. Differently, Guria et al. (2014) employed the non-dominated sorting genetic algorithm (NSGA-II) as optimization technique to determine the Pareto front. That is, instead of finding a unique optima solution, the NSGA-II finds a population of several solutions, considered as optima. The authors employed the Bourgoyne and Young formulation for ROP prediction and tooth wear. With these predictive models, a multi-objective optimization involving conflicting objectives was developed. The objective functions were: (*i*) maximization of the drilling depth, (*ii*) minimization of the drilling time, (*iii*) minimization of the drilling cost. The controllable drilling-variables used in this optimization study were four: equivalent circulation mud density, drill bit rotation speed, weight on bit, and Reynolds number of circulating mud through drill bit nozzles.

The intelligent optimization methods for multiple objectives are state-of-art solutions to approximate the Pareto Front in several complex problems. Some algorithms for multi-objective problems are Non-dominated Sorting Genetic Algorithm (NSGA-II), Multiple Objective Particle Swarm Optimization (MOPSO), Non-dominated Neighbor Immune Algorithm (NNIA) (Zhou et al., 2011; Chiandussi et al., 2012; Antonio and Coello, 2017; Cui et al., 2017). Such algorithms are global optimization methods. However, in multi-objective optimization, the term of global solution is unclear, because there is usually a set of optimum solutions (Pareto Set).

The possibility of having several optimum solutions paves the way to implement a more sophisticate selection of optimum drilling variables. For the case of drilling, it is normally desired to optimize the drilling variables, without applying too oft changes. Abrupt changes in drilling controllable variables are neither possible nor desired (Gandelman, 2012). A more sophisticate selection of optimum drilling variables can avoid oft changes in magnitude of drilling variables. However, it seems that this problem has not yet been addressed in the works published in the open literature.

4.4. Predictive model control

From all 61 papers that modeled ROP with AI method, four of them employed predictive model control approach (Dashevskiy et al., 1999; Fonseca et al., 2006; Mendes et al., 2007; Sui et al., 2013). This approach differs from all other optimization studies, because the ROP is estimated (i.e. forecast) for an instance in the future, based on the previous and current data. The main goal of these studies were to investigate the feasibility of implementing predictive control techniques, having as controllable variables, especially, WOB and RPM. In this case, forecasting models were used to determine the best input(s) which could improve the drilling performance for a next instance. This approach, however, seems not to have been investigated very often in the open literature.

5. Discussion

5.1. Challenges in implementing optimization techniques

Hajizadeh (2019) identified several threats which can weaken the implementation of ML projects in the oil and gas industry. However, the main goal of that paper was not only to point out the weaknesses, but also to show how ML projects can be conducted inside the companies from this sector of the industry. For that, the author performed a Strengths, Weaknesses, Opportunities, Threats (SWOT) analysis, in which had been pointed out that a critical weakness is originated from the organizational culture side. As Hajizadeh (2019) stressed out, the risk-averse culture leads to slow adoption of new technologies and the waterfall management (top-down development of projects) contributes even more for this slow adoption. In addition, the industry suffers from organization resistance to change, especially in adopting ML practices because of the fear and concern about job security.

Nevertheless, possible solutions for successful implementation of ML were presented as well, such as the use of the modern ML pipeline (see section 5.2), guaranteeing agile and collaborative teams of ML scientists and domain experts, and retaining long term talents in such fields (Hajizadeh, 2019). If these weaknesses are overcome, one has to account also for concerns about optimization systems. The key advantage of using predictive models is the possibility to apply formal optimization methods for the assurance of efficiency enhancement. Although the driller's experience plays an unquestionable important role in selecting the best practice in drilling activities, including variation of WOB, RPM and flow rate (Hegde and Gray, 2017), this approach relies clearly on subjective analysis, and is not as much ideal for real-time optimization as predictive models can be. Besides, Wang and Salehi (2015) used the term *set-it-and-forget-it* for describing typical approaches in determining optimum values of controllable parameters while drilling (ex. WOB, RPM and flow rate). However, this mindset may not be ideal given the fact that implementation of this optimization techniques relies on varying frequently magnitudes of these mentioned controllable drilling variables while drilling. More often implementation of pre-operational drilling tests, such as drill-off and drill-rate tests (Dupriest and Koederitz, 2005; Nascimento et al., 2016; Payette et al., 2017), is facilitating the task of searching the best optimum set of controllable parameters for a better drillability. In the drilling advisory system presented by Payette et al. (2017), the rig crew members are encouraged to change the controllable parameters within an operational range, by carrying out regular drill-off tests while drilling a specific formation or section.

Another considerably important consideration for a successful implementation of ML in such environments is the development of user-friendly interfaces/softwares (Nascimento et al., 2016). This in specific was a concern highlighted by Payette et al. (2017) during the development of their systems and methods for real-time drilling optimization. The authors employed a so-called human-machine-interaction strategy in the development of optimization system, aiming at creating simple interfaces useful for minimal trainings. When thinking about real-time drilling optimization, special care should be taken into account in order not to foreseen the safe ranges when drilling (e.g. sudden decrease mud weight to a considerable under-operational level). Due to high risks involved in the drilling activities and given its complexity, implementation of new techniques must be carefully tested before gaining industrial application (Balaji et al., 2018).

5.2. Modern ML pipeline for implementations

One challenge faced when implementing AI techniques is the requirement of deep knowledge in algorithms (sometimes) to obtain good results. In addition, the AI and ML related discipline can be challenging for newcomers whose expertise lies in other scopes (Bishop, 2013). There are already automatic techniques (e.g. Swearingen et al., 2017) to select the most suitable ML method for each specific problem and application. Automated ML enables to fill another absentee in the industry through Continuous Integration/Continuous Deployment (CICD) practices in ML applications, as highlighted by Hajizadeh (2019, p. 662). Since the implementation process of ML projects can experience a trial and error approach, a modern CICD provides suitable control mechanisms for that, as illustrated in Fig. 10. The modern ML pipeline makes possible also the continuous update of predictive models over time due to changes in systems.

5.3. Other intelligent techniques for drilling optimization

This specific review focuses on the use of ML for ROP prediction and application of predictive models for drilling optimization. Different fields from ML (e.g. reinforcement learning, automated ML, deep learning, among others) and AI (e.g. case-based reasoning, among others) can be employed in drilling optimization.

For example, reinforcement learning consists of an agent that can learn behavior through trial-and-error (Kaelbling et al., 1996). It can be very suitable in for dynamic environments and applied in process control applications (Badgwell et al., 2018). Just a few researches in the literature have already explored reinforcement learning applied in the

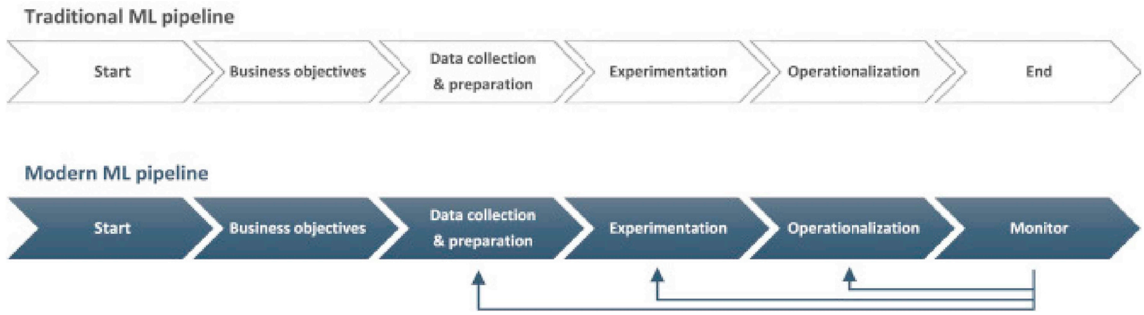

Fig. 10. A modern ML pipeline with Continuous Integration/Continuous Deployment (CICD). Reprinted from Hajizadeh (2019), with permission from Elsevier.

drilling engineering discipline (De Togni, 2018; Pollock et al., 2018).

Another technique from artificial intelligence, called case-based reasoning, can have a great impact in detecting drilling problems, and proposing the respective solutions by combing data analysis of historical and real-time data. This approach enables the identification and mitigation of events that lead to non-productive time (Skalle et al., 2013). For that, knowledge is extracted from past events by integrating data from historical information and rig crew experience. It is then possible to identify the most likely causes of the drilling problems as well as respective solutions and subsequent actions. Yuan et al. (2009) also detailed in his research the possibility in using case-based reasoning for choosing the drilling most suitable for optimization, such as the optimum drill-bit and what values of drilling controllable variables may allow maximizing the ROP and drillability. A detailed review of case-based reasoning applied to drilling engineering is given by Shokouhi et al. (2014).

Finally, a technique that has gained attention in last years is the deep learning. Deep learning models are composed of multiple processing layers, which are able to learn complex tasks that required multiple levels of abstraction (e.g. speech recognitions, visual object detection) (LeCun et al., 2015). For the learning process, a large amount of data is required to train the complex structure of deep learning models. Since the volume of data collected in the oil and gas industry grows exponentially, deep learning is a promising concept for models in drilling (e.g. ROP models) and many other disciplines (Soares and Gray, 2019).

6. Conclusions

This paper provides an extensive review of ML for ROP prediction, comparing intelligent techniques with traditional models (analytical equations). ML methods are normally more accurate than traditional models in ROP prediction. One of key aspects of successful in obtaining predictive data-driven models is the process of selecting the inputs (feature engineering). It was observed a preference in exploiting the flexibility of ML methods, by selecting normally small sub-sets of inputs (up to six variables). Additionally, it was observed that the eight most common inputs used to feed the predictive models reflect the drill bit mechanism theory from the traditional models (Fig. 7). This review classified the ROP studies in four different approaches regarding the use of lithology/rock formation properties; all of them could provide accurate ROP models.

Some classical machine learning algorithms (e.g. neural networks or SVMs) are normally black-boxes models, but, depending on the algorithm employed, it is possible to extract rules that represents the relationship among the drilling variables. These rules can help the drill crew to select the optimum drilling variables. Most of works carried out historical analysis for wells drilled in a similar region, what can be employed as post-analyses or prior to drilling a similar well. However, few tries obtained ROP-models while the drilling (Hegde et al., 2017; Hegde and Gray, 2017; Soares and Gray, 2019). It was also seen that not many works tried to formally treat possible measurement errors in the dataset.

This reviews covers also how the predictive models could be used for optimization purposes. Maximizing ROP was a common approach adopted. However, this may be not always the best approach, because drilling dysfunctions can occur at higher drilling rate. Therefore, other metrics should be taken into account, yielding a multi-objective optimization problem, which seems to be a more reasonable optimization approach due to the complexity of drilling process. Applying decision making techniques for multiple-objective functions (e.g. global criteria method, ε-constraint method) enables an easier implementation of a multi-objective optimization problem, as one of few implementations of real-time optimization reported in the literature (Payette et al., 2017).

Clearly more efforts should be considered across theory and practical aspects when applying machine learning within the petroleum industry. Although the current review provides a deep overview of ROP models (Section 3) and respective related drilling optimization rooms (Section 4), it had not been an intention to address here simple and straight forward rules of thumb for ML implementation in the drilling discipline, aiming optimization and/or efficiency enhancement. It is very likely that the implementation processes of ML in the petroleum industry can undergo trial and error experiences, considering the new approach and novelty the ML embraces. To be more effective, this trial/error approach can be facilitated by adopting the modern ML pipeline which provides suitable control mechanism for ML models (see Section 5.2). Besides, automated systems based on ML methods could enable engineers to develop and monitor in real-time systems, and consequently effectively aim for optimization and efficiency enhancement. For that, the weaknesses and threats concerning the petroleum industry (e.g. as the resistance of the human-industry behavior to changing) should be previously overcome, otherwise ML methods and related projects will not be implementable in a positive manner (Hajizadeh, 2019).

Acknowledgements

This study was financed in part by the *Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil* (CAPES) - Finance Code 001.

List of Acronyms

AI	artificial intelligence
ANFIS	adaptive neuro-fuzzy inference system
ANN	artificial neural networks
BIN	Bingham (1965) ROP model
BP	back-propagation training algorithm
BR	bayesian regularization training algorithm
BYM	Bourgoyne and Young (1974) ROP model
CICD	Continuous Integration/Continuous Deployment
DE	differential evolution algorithm
DENFIS	dynamic evolving neural-fuzzy
ELM	extreme learning machine
FIS	fuzzy inference system

GBM	gradient boosting machines
HAR	Hareland and Rampersad (1994) ROP model
LM-BP	Levenberg-Marquardt back-propagation
LR	linear regression
LS-SVR	least-squares support vector regression
M	Maurer (1962) ROP model
ML	machine learning
MLP	multilayer perceptrons
MLR	multivariate linear regression
MNR	multivariate nonlinear regression
Mod-M	modified Maurer (1962) ROP model
MOT	Motahhari et al. (2010) ROP model
MR	multivariate regression
MSE	mechanical specific energy
PDC	polycrystalline diamond compact
PV	plastic viscosity
Q	mud flow
RBF	radial basis functions
RF	random forests
ROP	rate of penetration
RPM	revolutions per minute
RTDD	real-time drilling data
SSP	standpipe pressure
SVM	support vector machines
SVR	support vector regression
TOB	torque on bit
UCS	uniaxial compressive strength
W	Warren (1987) ROP model
WOB	weight on bit

References

Abbas, A.K., Rushdi, S., Alsaba, M., 2018. Modeling rate of penetration for deviated wells using artificial neural network. In: Abu Dhabi International Petroleum Exhibition & Conference, 12-15 November, Abu Dhabi. UAE, Society of Petroleum Engineers. https://doi.org/10.2118/192875-MS.

Aggarwal, C.C., 2013. Outlier Analysis. Springer New York, New York, NY. https://doi.org/10.1007/978-1-4614-6396-2.

Agwu, O.E., Akpabio, J.U., Alabi, S.B., Dosunmu, A., 2018. Artificial intelligence techniques and their applications in drilling fluid engineering: a review. J. Pet. Sci. Eng. 167, 300–315. https://doi.org/10.1016/j.petrol.2018.04.019.

Ahmadi, M.A., 2016. Toward reliable model for prediction Drilling Fluid Density at wellbore conditions: a LSSVM model. Neurocomputing 211, 143–149. https://doi.org/10.1016/j.neucom.2016.01.106.

Ahmed, O.S., Adeniran, A.A., Samsuri, A., 2019. Computational intelligence based prediction of drilling rate of penetration: a comparative study. J. Pet. Sci. Eng. 172, 1–12. https://doi.org/10.1016/j.petrol.2018.09.027.

Akin, S., Karpuz, C., 2008. Estimating drilling parameters for diamond bit drilling operations using artificial neural networks. Int. J. Geomech. 8, 68–73. https://doi.org/10.1061/(ASCE)1532-3641(2008)8:1(68).

Al-abduljabbar, A., 2019. A robust rate of penetration model for carbonate formation. J. Energy Resour. Technol. 141, 0429031–042903–9. https://doi.org/10.1115/1.4041840.

AL-Rashidi, A.F.A.B., 2011. Application of Neural Networks to Evaluate Factors Affecting Drilling Performance. Doctor Thesis. Imperial College London.

AlArfaj, I., Khoukhi, A., Eren, T., 2012. Application of advanced computational intelligence to rate of penetration prediction. In: 2012 Sixth UKSim/AMSS European Symposium on Computer Modeling and Simulation, pp. 33–38. https://doi.org/10.1109/EMS.2012.79.

Amar, K., Ibrahim, A., 2012. Rate of penetration prediction and optimization using advances in artificial neural networks, a comparative study. In: Proceedings of the 4th International Joint Conference on Computational Intelligence, pp. 647–652. https://doi.org/10.5220/0004172506470652.

Amer, M.M., Dahab, A.S., El-Sayed, A.A.H., 2017. An ROP predictive model in nile delta area using artificial neural networks. In: SPE Kingdom of Saudi Arabia Annual Technical Symposium and Exhibition, 24-27 April, Dammam, Saudi Arabia. Society of Petroleum Engineers. https://doi.org/10.2118/187969-MS.

Anemangely, M., Ramezanzadeh, A., Tokhmechi, B., Molaghab, A., Mohammadian, A., 2018. Drilling rate prediction from petrophysical logs and mud logging data using an optimized multilayer perceptron neural network. J. Geophys. Eng. 15, 1146–1159. https://doi.org/10.1088/1742-2140/aaac5d.

Ansari, H.R., Hosseini, M.J.S., Amirpour, M., 2017. Drilling rate of penetration prediction through committee support vector regression based on imperialist competitive algorithm. Carbonates Evaporites 32, 205–213. https://doi.org/10.1007/s13146-016-0291-8.

Antonio, L.M., Coello, C.A.C., 2017. Coevolutionary multi-objective evolutionary algorithms: a survey of the state-of-the-art. IEEE Trans. Evol. Comput. 1–16. https://doi.org/10.1109/TEVC.2017.2767023. (Preprint).

Arabjamaloei, R., Karimi Dehkordi, B., 2012. Investigation of the most efficient approach of the prediction of the rate of penetration. Energy Sources, Part A Recovery, Util. Environ. Eff. 34, 581–590. https://doi.org/10.1080/15567036.2010.493925.

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.

Arehart, R., 1990. Drill-bit diagnosis with neural networks. SPE Comput. Appl. 2, 24–28. https://doi.org/10.2118/19558-PA.

Arnaout, A., Zoellner, P., Johnstone, N., Thonhauser, G., 2013. Intelligent data quality control of real-time rig data. In: SPE Middle East Intelligent Energy Conference and Exhibition, 28-30 October. Society of Petroleum Engineers, Manama, Bahrain. https://doi.org/10.2118/167437-MS.

Ashrafi, S.B., Anemangely, M., Sabah, M., Ameri, M.J., 2019. Application of hybrid artificial neural networks for predicting rate of penetration (ROP): a case study from Marun oil field. J. Pet. Sci. Eng. 175, 604–623. https://doi.org/10.1016/j.petrol.2018.12.013.

Awotunde, A.A., Mutasiem, M.A., 2014. Efficient drilling time optimization with differential evolution. In: SPE Nigeria Annual International Conference and Exhibition. Society of Petroleum Engineers, Lagos, Nigeria. https://doi.org/10.2118/172419-MS. sPE-172419-MS.

Ayoub, M., Shien, G., Diab, D., Ahmed, Q., 2017. Modeling of drilling rate of penetration using adaptive neuro-fuzzy inference system. Int. J. Appl. Eng. Res. 12, 12880–12891.

Badgwell, T.A., Lee, J.H., Liu, K.H., 2018. Reinforcement learning – overview of recent progress and implications for process control. In: Eden, M.R., Ierapetritou, M.G., Towler, G.P. (Eds.), 13th International Symposium on Process Systems Engineering (PSE 2018). Elsevier. volume 44 of Computer Aided Chemical Engineering, pp. 71–85. https://doi.org/10.1016/B978-0-444-64241-7.50008-2.

Bahari, A., Baradaran, A., 2007. Trust-region approach to find constants of Bourgoyne and Young penetration rate model in Khangiran Iranian gas field. In: Latin American & Caribbean Petroleum Engineering Conference, 15-18 April, Buenos Aires, Argentina. Society of Petroleum Engineers. https://doi.org/10.2118/107520-MS.

Bahari, A., Seyed, A.B., 2009. Drilling cost optimization in a hydrocarbon field by combination of comparative and mathematical methods. Pet. Sci. 6, 451–463. https://doi.org/10.1007/s12182-009-0069-x.

Bahari, M., Bahari, A., Moharrami, F.N., Sistani, M.N., 2008. Determining Bourgoyne and Young model coefficients using genetic algorithm to predict drilling rate. J. Appl. Sci. 8, 3050–3054. https://doi.org/10.3923/jas.2008.3050.3054.

Balaji, K., Rabiei, M., Suicmez, V., Canbaz, C.H., Agharzeyva, Z., Tek, S., Bulut, U., Temizel, C., 2018. Status of data-driven methods and their applications in oil and gas industry. In: SPE Europec Featured at 80th EAGE Conference and Exhibition. Society of Petroleum Engineers, Copenhagen, Denmark. https://doi.org/10.2118/190812-MS.

Barbosa, L.F.F.M., Nascimento, A., Carvalho, J.A., Mathias, M.H., 2018. Machine learning methods applied to rate of penetration prediction - a technical review. In: Rio Oil & Gas Expo and Conference 2018 Proceedings, Brazilian Petroleum. Gas and Biofuels Institute - IBP, Rio de Janeiro, Brazil IBP1871_18.

Basarir, H., Tutluoglu, L., Karpuz, C., 2014. Penetration rate prediction for diamond bit drilling by adaptive neuro-fuzzy inference system and multiple regressions. Eng. Geol. 173, 1–9. https://doi.org/10.1016/j.enggeo.2014.02.006.

Bataee, M., Irawan, S., Kamyab, M., 2014. Artificial neural network model for prediction of drilling rate of penetration and optimization of parameters. J. Jpn. Pet. Inst. 57, 65–70. https://doi.org/10.1627/jpi.57.65.

Bataee, M., Mohseni, S., 2011. Application of artificial intelligent systems in ROP optimization : a case study in Shadegan oil field. In: SPE Middle East Unconventional Gas Conference and Exhibition, 31 January-2 February, Muscat, Oman. Society of Petroleum Engineers. https://doi.org/10.2118/140029-MS.

Bello, O., Teodoriu, C., Yaqoob, T., Oppelt, J., Holzmann, J., Obiwanne, A., 2016. Application of artificial intelligence techniques in drilling system design and operations: a state of the art review and future research pathways. In: SPE Nigeria Annual International Conference and Exhibition. Society of Petroleum Engineers. https://doi.org/10.2118/184320-MS.

Berrar, D., 2019. Cross-validation. In: Ranganathan, S., Gribskov, M., Nakai, K., Schönbach, C. (Eds.), Encyclopedia of Bioinformatics and Computational Biology. Academic Press, Oxford, pp. 542–545. https://doi.org/10.1016/B978-0-12-809633-8.20349-X.

Bezminabadi, S.N., Ramezanzadeh, A., Jalali, S.m.E., Tokhmechi, B., Roustaei, A., 2017. Effect of rock properties on ROP modeling using statistical and intelligent methods: a case study of an oil well in Southwest of Iran. Arch. Min. Sci. 62, 131–144. https://doi.org/10.1515/amsc-2017-0010.

Bilgesu, H., Tetrick, L., Altmis, U., Mohaghegh, S., Ameri, S., 1997. A new approach for the prediction of rate of penetration (ROP) values. In: SPE Eastern Regional Meeting, , 22-24 October. Society of Petroleum Engineers, Lexington, Kentucky. https://doi.org/10.2118/39231-MS.

Bingham, M.G., 1965. A New Approach to Interpreting Rock Drillability. Petroleum Publishing Company.

Bishop, C.M., 1995. Neural Networks for Pattern Recognition. Oxford University Press, Inc., New York, NY, USA.

Bishop, C.M., 2006. Pattern Recognition and Machine Learning. Springer-Verlag, Berlin, Heidelberg.

Bishop, C.M., 2013. Model-based machine learning. Phil. Trans. R. Soc. Lond. A Math. Phys. Eng. Sci. 371. https://doi.org/10.1098/rsta.2012.0222.

Bizanti, M., 1989. Drilling rate model for Louisiana formations. J. Pet. Sci. Eng. 2, 261–267. https://doi.org/10.1016/0920-4105(89)90003-X.

Bodaghi, A., Ansari, H.R., Gholami, M., 2015. Optimized support vector regression for drilling rate of penetration estimation. Open Geosci. 7, 870–879. https://doi.org/10.1515/geo-2015-0054.
Bourgoyne, A., Millheim, K., Chenevert, M., Young, F., 1986. Applied Drilling Engineering. Society of Petroleum Engineers.
Bourgoyne, A., Young, F.S., 1974. A multiple regression approach to optimal drilling and abnormal pressure detection. Soc. Pet. Eng. J. 14, 371–384. https://doi.org/10.2118/4238-PA.
Breiman, L., 1996. Bagging predictors. Mach. Learn. 24, 123–140. https://doi.org/10.1007/BF00058655.
Breiman, L., 2001a. Random forests. Mach. Learn. 45, 5–32. https://doi.org/10.1023/A:1010933404324.
Breiman, L., 2001b. Statistical modeling: the two cultures (with comments and a rejoinder by the author). Stat. Sci. 16, 199–231. https://doi.org/10.1214/ss/1009213726.
Broomhead, D.S., Lowe, D., 1988. Multivariable functional interpolation and adaptive networks. Complex Syst. 2, 321–355. https://www.complex-systems.com/abstracts/v02_i03_a05/.
Browne, M.W., 2000. Cross-validation methods. J. Math. Psychol. 44, 108–132. https://doi.org/10.1006/jmps.1999.1279.
Buntine, W.L., Weigend, A.S., 1991. Bayesian back-propagation. Complex Syst. 5, 603–643. https://www.complex-systems.com/abstracts/v05_i06_a04/.
Cao, W., Wang, X., Ming, Z., Gao, J., 2018. A review on neural networks with random weights. Neurocomputing 275, 278–287. https://doi.org/10.1016/j.neucom.2017.08.040.
Chiandussi, G., Codegone, M., Ferrero, S., Varesio, F., 2012. Comparison of multi-objective optimization methodologies for engineering applications. Comput. Math. Appl. 63, 912–942. https://doi.org/10.1016/j.camwa.2011.11.057.
Cortes, C., Vapnik, V., 1995. Support-vector networks. Mach. Learn. 20, 273–297. https://doi.org/10.1007/BF00994018.
Cui, Y., Geng, Z., Zhu, Q., Han, Y., 2017. Review: multi-objective optimization methods and application in energy saving. Energy 125, 681–704. https://doi.org/10.1016/j.energy.2017.02.174.
Dashevskiy, D., Dubinsky, V., Macpherson, J., 1999. Application of neural networks for predictive control in drilling dynamics. In: SPE Annual Technical Conference and Exhibition. Society of Petroleum Engineers. https://doi.org/10.2118/56442-MS.
De Togni, R., 2018. Predictive model for drilling phase duration of oil & gas wells. In: Master's thesis in computer science and engineering. Politecnico di Milano, . http://hdl.handle.net/10589/141802.
Deng, H., 2018. Interpreting tree ensembles with intrees. Int. J. Data Sci. Anal. https://doi.org/10.1007/s41060-018-0144-8.
Deng, N., Tian, Y., Zhang, C., 2012. Support Vector Machines: Optimization Based Theory, Algorithms, and Extensions, first ed. Chapman and Hall/CRC, New York. https://doi.org/10.1201/b14297.
Deng, Y., Chen, M., Jin, Y., Zhang, Y., Zou, D., Lu, Y., 2016. Theoretical and experimental study on the penetration rate for roller cone bits based on the rock dynamic strength and drilling parameters. J. Nat. Gas Sci. Eng. 36, 117–123. https://doi.org/10.1016/j.jngse.2016.10.019.
Detournay, E., Defourny, P., 1992. A phenomenological model for the drilling action of drag bits. Int. J. Rock Mech. Min. Sci. Geomech. Abstr. 29, 13–23. https://doi.org/10.1016/0148-9062(92)91041-3.
Diaz, M.B., Kim, K.Y., Kang, T.H., Shin, H.S., 2018. Drilling data from an enhanced geothermal project and its pre-processing for ROP forecasting improvement. Geothermics 72, 348–357. https://doi.org/10.1016/j.geothermics.2017.12.007.
Diaz, M.B., Kim, K.Y., Shin, H.S., 2017a. On-line prediction model for rate of penetration (ROP) with cumulating field data in real time. In: 4th ISRM Young Scholars Symposium on Rock Mechanics, 10-13 May, Jeju, Korea. International Society for Rock Mechanics and Rock Engineering. https://www.onepetro.org/conference-paper/ISRM-YSS-2017-071.
Diaz, M.B., Kim, K.Y., Shin, H.s., 2017b. Rate of penetration (ROP) forecast based on artificial neural network with online learning. In: The 2017 World Congress on Advances in Structural Engineering and Mechanics (ASEM17), Ilsan(Seoul), Korea, . http://www.i-asem.org/publication_conf/asem17/7.TS/W4G.6.TS1404_3747F1.pdf.
Donne, C.P.J., 2017. Parameter detection in real time drilling data. Master's thesis (petroleum engineering master). Faculty of Engineering Department of Geoscience and Petroleum, Norwegian University of Science and Technology, Trondheim, Norway. http://hdl.handle.net/11250/2450329.
Drucker, H., Burges, C.J.C., Kaufman, L., Smola, A.J., Vapnik, V., 1997. Support vector regression machines. In: In: Mozer, M.C., Jordan, M.I., Petsche, T. (Eds.), Advances in Neural Information Processing Systems, vol. 9. MIT Press, pp. 155–161. http://papers.nips.cc/paper/1238-support-vector-regression-machines.pdf.
Duan, J., Zhao, J., Xiao, L., Yang, C., Chen, H., 2014. A ROP prediction approach based on improved BP neural network. In: 2014 IEEE 3rd International Conference on Cloud Computing and Intelligence Systems. IEEE, pp. 668–671. https://doi.org/10.1109/CCIS.2014.7175818.
Duan, J., Zhao, J., Xiao, L., Yang, C., Li, C., 2015. A crop optimization approach based on improved bp neural network pso. In: Advances in Swarm and Computational Intelligence. Springer International Publishing, pp. 11–18. https://doi.org/10.1007/978-3-319-20469-7_2.
Dupriest, F.E., Koederitz, W.L., 2005. Maximizing drill rates with real-time surveillance of mechanical specific energy. In: SPE/IADC Drilling Conference, Amsterdam, The Netherlands, 23-25 February. Society of Petroleum Engineers, Amsterdam, The Netherlands. https://doi.org/10.2118/92194-MS.
Elkatatny, S., 2018. New approach to optimize the rate of penetration using artificial neural network. Arabian J. Sci. Eng. 43, 6297–6304. https://doi.org/10.1007/s13369-017-3022-0.
Eren, T., 2010. Real-time-optimization of Drilling Parameters during Drilling Operations. Phd thresis. Middle East Technical University.
Eren, T., Ozbayoglu, M.E., 2010. Real time optimization of drilling parameters during drilling operations. In: SPE Oil and Gas India Conference and Exhibition. Society of Petroleum Engineers. https://doi.org/10.2118/129126-MS.
Eskandarian, S., Bahrami, P., Kazemi, P., 2017. A comprehensive data mining approach to estimate the rate of penetration: application of neural network, rule based models and feature ranking. J. Pet. Sci. Eng. 156, 605–615. https://doi.org/10.1016/j.petrol.2017.06.039.
Esmaeili, A., Elahifar, B., Fruhwirth, R.K., Thonhauser, G., 2012. ROP modeling using neural network and drill string vibration data. In: SPE Kuwait International Petroleum Conference and Exhibition, 10-12 December, Kuwait City, Kuwait. Society of Petroleum Engineers. https://doi.org/10.2118/163330-MS.
Fonseca, T.C., Mendes, J.R.P., Serapião, A.B.S., Guilherme, I.R., 2006. A genetic neuro-model reference adaptive controller for petroleum wells drilling operations. In: 2006 International Conference on Computational Intelligence for Modelling Control and Automation and International Conference on Intelligent Agents Web Technologies and International Commerce (CIMCA'06). IEEE. https://doi.org/10.1109/CIMCA.2006.8.
Formighieri, S., Filho, P.J.d.F., 2015. Estimation of Bourgoyne and Young model coefficients using Markov chain Monte Carlo simulation. In: 2015 Winter Simulation Conference (WSC). IEEE, pp. 1172–1183. https://doi.org/10.1109/WSC.2015.7408243.
Friedman, J.H., 2001. Greedy function approximation: a gradient boosting machine. Ann. Stat. 29, 1189–1232. https://doi.org/10.1214/aos/1013203451.
Friedman, J.H., 2006. Recent advances in predictive (machine) learning. J. Classif. 23, 175–197. https://doi.org/10.1007/s00357-006-0012-4.
Fruhwirth, R.K., Thonhauser, G., Mathis, W., 2006. Hybrid simulation using neural networks to predict drilling hydraulics in real time. In: SPE Annual Technical Conference and Exhibition. Society of Petroleum Engineers, San Antonio, Texas. https://doi.org/10.2118/103217-MS.
Gandelman, R.A., 2012. Predição da ROP e Otimização em Tempo Real de Parâmetros Operacionais na Perfuração de Poços de Petróleo Offshore (Portuguese). Dissertação (metrado em tecnologia de processos químicos e bioquímicos). Escola de Química, Universidade Federal do Rio de Janeiro. Rio de Janeiro, RJ - Brasil. http://epqb.eq.ufrj.br/download/predicao-da-rop-e-otimizacao-em-tempo-real.pdf.
Garavand, A.A., Esmaeilian, G., 2015. Development of an imperialist competitive algorithm (ICA)-based committee machine to predict bit penetration rate in oil wells of Iran. Cienc. Nat. 37, 173–182. https://doi.org/10.5902/2179460X20844.
Garćia, L.P.F., de Carvalho, A.C.P.L.F., Lorena, A.C., 2013. Noisy data set identification. In: Hybrid Artificial Intelligent Systems. Springer Berlin Heidelberg, pp. 629–638. https://doi.org/10.1007/978-3-642-40846-5_63.
Gelenbe, E., 1989. Random neural networks with negative and positive signals and product form solution. Neural Comput. 1, 502–510. https://doi.org/10.1162/neco.1989.1.4.502.
Gidh, Y., Ibrahim, H., Purwanto, A., Bits, S., Company, A.S., 2011. Real-time drilling parameter optimization system increases ROP by predicting/managing bit wear. In: SPE Digital Energy Conference and Exhibition, 19-21 April, the Woodlands. Society of Petroleum Engineers, Texas, USA. https://doi.org/10.2118/142880-MS.
Gidh, Y.K., Purwanto, A., Ibrahim, H., 2012. Artificial neural network drilling parameter optimization system improves ROP by predicting/managing bit wear. In: SPE Intelligent Energy International. Society of Petroleum Engineers. https://doi.org/10.2118/149801-MS.
Graham, J., Muench, N., 1959. Analytical determination of optimum bit weight and rotary speed combinations. In: Fall Meeting of the Society of Petroleum Engineers of AIME. Society of Petroleum Engineers. https://doi.org/10.2118/1349-G.
Guria, C., Goli, K.K., Pathak, A.K., 2014. Multi-objective optimization of oil well drilling using elitist non-dominated sorting genetic algorithm. Pet. Sci. 11, 97–110. https://doi.org/10.1007/s12182-014-0321-x.
Guyon, I., Elisseeff, A., 2003. An introduction to variable and feature selection. J. Mach. Learn. Res. 3, 1157–1182. https://doi.org/10.1016/j.aca.2011.07.027.
Hagan, M.T., Menhaj, M.B., 1994. Training feedforward networks with the marquardt algorithm. IEEE Trans. Neural Netw. 5, 989–993. https://doi.org/10.1109/72.329697.
Hajizadeh, Y., 2019. Machine learning in oil and gas; a swot analysis approach. J. Pet. Sci. Eng. 176, 661–663. https://doi.org/10.1016/j.petrol.2019.01.113.
Hankins, D., Salehi, S., Saleh, F.K., 2015. An integrated approach for drilling optimization using advanced drilling optimizer. J. Pet. Eng. 2015, 12. https://doi.org/10.1155/2015/281276.
Hareland, G., Hoberock, L., 1993. Use of drilling parameters to predict in-Situ stress bounds. In: SPE/IADC Drilling Conference, 22-25 February. Society of Petroleum Engineers, Amsterdam. https://doi.org/10.2118/25727-MS.
Hareland, G., Rampersad, P., 1994. Drag - bit model including wear. In: SPE Latin America/Caribbean Petroleum Engineering Conference. Society of Petroleum Engineers, Buenos Aires, Argentina. https://doi.org/10.2118/26957-MS.
Hasan, B.M., Aboozar, B., Hamidreza, M., 2011. Intelligent drilling rate predictor. Int. J. Innov. Comput. Inf. Contr. 7, 1511–1519.
Hastie, T., Tibshirani, R., Friedman, J., 2009. The Elements of Statistical Learning. Springer Series in Statistics, 12 ed. Springer New York, New York, NY. https://doi.org/10.1007/978-0-387-84858-7. https://web.stanford.edu/~hastie/ElemStatLearn/.
Haykin, S., 1994. Neural Networks - a Comprehensive Foundation. Prentice-Hall Inc., Upper Saddle Rive, New Jersey.
Haykin, S.S., 2009. Neural Networks and Learning Machines, third ed. Pearson Education, Upper Saddle River, NJ.
Hedayatzadeh, M., Shahriar, K., Hamidi, J.K., 2010. An Artificial Neural Network Model to Predict the Performance of Hard Rock TBM. International Society for Rock

Mechanics and Rock Engineering, New Delhi, India. https://www.onepetro.org/conference-paper/ISRM-ARMS6-2010-107.
Hegde, C., Daigle, H., Gray, K.E., 2018a. Performance comparison of algorithms for real-time rate-of-penetration optimization in drilling using data-driven models. SPE J. 23, 1706–1722. https://doi.org/10.2118/191141-PA.
Hegde, C., Daigle, H., Millwater, H., Gray, K., 2017. Analysis of rate of penetration (ROP) prediction in drilling using physics-based and data-driven models. J. Pet. Sci. Eng. 159, 295–306. https://doi.org/10.1016/j.petrol.2017.09.020.
Hegde, C., Gray, K., 2018. Evaluation of coupled machine learning models for drilling optimization. J. Nat. Gas Sci. Eng. 56, 397–407. https://doi.org/10.1016/j.jngse.2018.06.006.
Hegde, C., Gray, K.E., 2017. Use of machine learning and data analytics to increase drilling efficiency for nearby wells. J. Nat. Gas Sci. Eng. 40, 327–335. https://doi.org/10.1016/j.jngse.2017.02.019.
Hegde, C., Soares, C., Gray, K., 2018b. Rate of penetration (ROP) modeling using hybrid models: deterministic and machine learning. In: SPE/AAPG/SEG Unconventional Resources Technology Conference, 23-25 July. Unconventional Resources Technology Conference, Houston, Texas, USA. https://doi.org/10.15530/URTEC-2018-2896522.
Hegde, C., Wallace, S., Gray, K., 2015a. Using trees , bagging , and random forests to predict rate of penetration. In: SPE Middle East Intelligent Oil and Gas Conference and Exhibition, 15-16 September, Abu Dhabi. UAE, Society of Petroleum Engineers This. https://doi.org/10.2118/176792-MS.
Hegde, C.M., Wallace, S.P., Gray, K.E., 2015b. Use of regression and bootstrapping in drilling inference and prediction. In: SPE Middle East Intelligent Oil and Gas Conference and Exhibition, 15-16 September, Abu Dhabi. UAE, Society of Petroleum Engineers. https://doi.org/10.2118/176791-MS.
Hornik, K., Stinchcombe, M., White, H., 1989. Multilayer feedforward networks are universal approximators. Neural Netw. 2, 359–366. https://doi.org/10.1016/0893-6080(89)90020-8.
Howarth, D., Adamson, W.R., Berndt, J.R., 1986. Technical note correlation of model tunnel boring and drilling machine performances with rock properties. Int. J. Rock Mech. Min. Sci. Geomech. Abstr. 23, 171–175. https://doi.org/10.1016/0148-9062(86)90344-X.
Huang, G.B., Chen, L., Siew, C.K., 2006. Universal approximation using incremental constructive feedforward networks with random hidden nodes. IEEE Trans. Neural Netw. 17, 879–892.
Huang, G.B., Zhu, Q.Y., Siew, C.K., 2004. Extreme learning machine: a new learning scheme of feedforward neural networks. In: 2004 IEEE International Joint Conference on Neural Networks (IEEE Cat. No.04CH37541), vol. 2. pp. 985–990. https://doi.org/10.1109/IJCNN.2004.1380068.
Jacinto, C.M.C., Freitas Filho, P.J., Nassar, S.M., Roisenberg, M., Rodrigues, D.G., Lima, M.D.C., 2013. Optimization models and prediction of drilling rate (ROP) for the Brazilian pre-salt layer. Chem. Eng. Trans. 33, 823–828. https://doi.org/10.3303/CET1333138.
Jahanbakhshi, R., Keshavarzi, R., Jafarnezhad, A., 2012. Real-time prediction of rate of penetration during drilling operation in oil and gas wells. In: 46th U.S. Rock Mechanics/Geomechanics Symposium. American Rock Mechanics Association, Chicago, Illinois. https://www.onepetro.org/conference-paper/ARMA-2012-244.
Jang, J.R., 1993. ANFIS: adaptive-network-based fuzzy inference system. IEEE Trans. Syst. Man Cybern. 23, 665–685. https://doi.org/10.1109/21.256541.
Jiang, W., Samuel, R., 2016. Optimization of rate of penetration in a convoluted drilling framework using Ant colony optimization. In: IADC/SPE Drilling Conference and Exhibition. Society of Petroleum Engineers, Fort Worth, Texas. https://doi.org/10.2118/178847-MS.
Kaelbling, L.P., Littman, M.L., Moore, A.W., 1996. Reinforcement learning: a survey. J. Artif. Intell. Res. 4, 237–285. https://doi.org/10.1613/jair.301.
Kahraman, S., 2002. Correlation of TBM and drilling machine performances with rock brittleness. Eng. Geol. 65, 269–283. https://doi.org/10.1016/S0013-7952(01)00137-5.
Kahraman, S., 2016. Estimating the penetration rate in diamond drilling in laboratory works using the regression and artificial neural network analysis. Neural Process. Lett. 43, 523–535. https://doi.org/10.1007/s11063-015-9424-7.
Kasabov, N.K., Song, Q., 2002. DENFIS: dynamic evolving neural-fuzzy inference system and its application for time-series prediction. IEEE Trans. Fuzzy Syst. 10, 144–154. https://doi.org/10.1109/91.995117.
Khosravanian, R., Choodar, B., Wood, D.A., Aadnoy, B.S., 2016a. ROP fuzzy-logic model proposed for intelligent drilling in Iran, Malaysia. Oil Gas J. 114.
Khosravanian, R., Sabah, M., Wood, D.A., Shahryari, A., 2016b. Weight on drill bit prediction models: Sugeno-type and Mamdani-type fuzzy inference systems compared. J. Nat. Gas Sci. Eng. 36, 280–297. https://doi.org/10.1016/j.jngse.2016.10.046.
Khosravanian, R., Wood, D.A., Choodar, B., Aadnoy, B.S., 2016c. Prediction Rate of Penetration (ROP) Using Fuzzy Logic in Intelligent Drilling (Technical Report November). .
Kutas, D.T., Nascimento, A., Elmgerbi, A.M., Roohi, A., Prohaska, M., Thonhauser, G., Mathias, M.H., 2015. A study of the applicability of Bourgoyne & Young ROP model and fitting reliability through regression. In: International Petroleum Technology Conference, International Petroleum Technology Conference, https://doi.org/10.2523/IPTC-18521-MS.
LeCun, Y., 1988. A theoretical framework for back-propagation. In: Touretzky, D., Hinton, G., Sejnowski, T. (Eds.), Proceedings of the 1988 Connectionist Models Summer School. Morgan Kaufmann, Pittsburgh, Pa, pp. 21–28.
LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. Nature 521, 436–444. https://doi.org/10.1038/nature14539.
Lee, Y., Oh, S., Kim, M.W., 1991. The effect of initial weights on premature saturation in back-propagation learning. In: IJCNN-91-Seattle International Joint Conference on Neural Networks, vol. 1. pp. 765–770. https://doi.org/10.1109/IJCNN.1991.155275.
Little, R.J.A., Rubin, D.B., 2002. Statistical Analysis with Missing Data, second ed. John Wiley and Sons, Inc., Hoboken, NJ - EUA.
Lyons, W.C., Plisga, G.J., 2004. Standard Handbook of Petroleum and Natural Gas Engineering, second ed. Gulf Professional Publishing.
MacKay, D.J., 1991. Bayesian interpolation. Neural Comput. 4, 415–447. https://doi.org/10.1162/neco.1992.4.3.415.
Mahdevari, S., Shahriar, K., Yagiz, S., Akbarpour, M., 2014. A support vector regression model for predicting tunnel boring machine penetration rates. Int. J. Rock Mech. Min. Sci. 72, 214–229. https://doi.org/10.1016/j.ijrmms.2014.09.012.
Mamdani, E., Assilian, S., 1975. An experiment in linguistic synthesis with a fuzzy logic controller. Int. J. Man Mach. Stud. 7, 1–13. https://doi.org/10.1016/S0020-7373(75)80002-2.
Mantha, B., Samuel, R., 2016. ROP optimization using artificial intelligence techniques with statistical regression coupling. In: SPE Annual Technical Conference and Exhibition, 26-28 September, Dubai. UAE, Society of Petroleum Engineers. https://doi.org/10.2118/181382-MS.
Mathis, W., Thonhauser, G., Wallnoefer, G., Ettl, J., 2007. Use of real-time rig-sensor data to improve daily drilling reporting, benchmarking, and planning. SPE Drill. Complet. 22, 217–226. https://doi.org/10.2118/99880-PA.
Mathur, N., Glesk, I., Buis, A., 2016. Comparison of adaptive neuro-fuzzy inference system (ANFIS) and Gaussian processes for machine learning (GPML) algorithms for the prediction of skin temperature in lower limb prostheses. Med. Eng. Phys. 38, 1083–1089. https://doi.org/10.1016/j.medengphy.2016.07.003.
Maurer, W., 1962. The "perfect - cleaning" theory of rotary drilling. J. Pet. Technol. 14, 1270–1274. https://doi.org/10.2118/408-PA.
Mendel, J.M., 1995. Fuzzy logic systems for engineering: a tutorial. Proc. IEEE 83, 345–377. https://doi.org/10.1109/5.364485.
Mendes, J.R.P., Fonseca, T.C., Serapião, A.B.S., 2007. Applying a genetic neuro-model reference adaptive controller in drilling optimization. World Oil 1–5.
Mendes-Moreira, J.a., Soares, C., Jorge, A.M., Sousa, J.F.D., 2012. Ensemble approaches for regression: a survey. ACM Comput. Surv. 45, 10:1–10:40. https://doi.org/10.1145/2379776.2379786.
Mitchell, R.F., Miska, S.Z., 2011. Fundamentals of Drilling Engineering. Society of Petroleum Engineers, Richardson, TX.
Momeni, M., Hosseini, S.J., Ridha, S., Laruccia, M.B., Liu, X., 2018. An optimum drill bit selection technique using artificial neural networks and genetic algorithms to increase the rate of penetration. J. Eng. Sci. Technol. 13, 361–372.
Monazami, M., Hashemi, A., Shahbazian, M., 2012. Drilling rate of penetration prediction using artificial neural network: a case study of one of Iranian southern oil fields. Oil Gas Bus. 6, 21–31.
Moradi, H., Bahari, M.H., Bagher, M., Sistani, N., 2010. Drilling rate prediction using an innovative soft computing approach. Sci. Res. Essays 5, 1583–1588.
Moran, D.P., Ibrahim, H.F., Purwanto, A., Osmond, J., 2010. Sophisticated ROP prediction technology based on neural network delivers accurate results sophisticated ROP prediction technology based on neural network delivers accurate results. In: IADC/SPE Asia Pacific Drilling Technology Conference and Exhibition, 1-3 November. Society of Petroleum Engineers, Ho Chi Minh City, Vietnam. https://doi.org/10.2118/132010-MS.
Moraveji, M.K., Naderi, M., 2016. Drilling rate of penetration prediction and optimization using response surface methodology and bat algorithm. J. Nat. Gas Sci. Eng. 31, 829–841. https://doi.org/10.1016/j.jngse.2016.03.057.
Motahhari, H., Hareland, G., James, J., 2010. Improved drilling efficiency technique using integrated PDM and PDC bit parameters. J. Can. Pet. Technol. 49, 45–52. https://doi.org/10.2118/141651-PA.
Nascimento, A., 2016. Mathematical modeling for drilling optimization in pre-salt sections: a focus on South Atlantic ocean operations. Doctoral thesis. São Paulo State University (UNESP), School of Engineering, Guaratinguetá. http://hdl.handle.net/11449/136182.
Nascimento, A., Elmgerbi, A., Roohi, A., Prohaska, M., Thonhauser, G., Gonçalves, J.L., Mathias, M.H., 2016. Reverse engineering: a new well monitoring and analysis methodology approaching playing-back drill-rate tests in real-time for drilling optimization. J. Energy Resour. Technol. 139https://doi.org/10.1155/2015/631290. 12902–12902–5.
Nascimento, A., Tamas Kutas, D., Elmgerbi, A., Thonhauser, G., Hugo Mathias, M., 2015. Mathematical modeling applied to drilling engineering: an application of Bourgoyne and Young ROP model to a presalt case study. Math. Probl. Eng. 2015, 1–9. https://doi.org/10.1155/2015/631290.
Ning, J., Honghai, F., Yinghu, Z., Tianyu, L., 2013. A new model of ROP prediction for drilling engineering with data source data output desired output. Adv. Inf. Sci. Serv. Sci. (AISS) 5, 597–604. https://doi.org/10.4156/AISS.vol5.issue7.70.
Noshi, C.I., Schubert, J.J., 2018. The role of machine learning in drilling operations; a review. In: SPE/AAPG Eastern Regional Meeting. Society of Petroleum Engineers, Pittsburgh, Pennsylvania, USA. https://doi.org/10.2118/191823-18ERM-MS.
Otalvora, W.C., AlKhudiri, M., Alsanie, F., Mathew, B., 2016. A comprehensive approach to measure the realtime data quality using key performance indicators. In: SPE Annual Technical Conference and Exhibition, 26-28 September. UAE, Society of Petroleum Engineers, Dubai. https://doi.org/10.2118/181315-MS.
Pao, Y., Takefuji, Y., 1992. Functional-link net computing: theory, system architecture, and functionalities. Computer 25, 76–79. https://doi.org/10.1109/2.144401.
Pao, Y.H., Park, G.H., Sobajic, D.J., 1994. Learning and generalization characteristics of the random vector functional-link net. Neurocomputing 6, 163–180. https://doi.org/10.1016/0925-2312(94)90053-1.
Payette, G.S., Spivey, B.J., Wang, L., Bailey, J.R., Sanderson, D., Kong, R., Pawson, M., Eddy, A., 2017. Real-time well-site based surveillance and optimization platform for

drilling: technology, basic workflows and field results. In: SPE/IADC Drilling Conference and Exhibition. Society of Petroleum Engineers, Hague, The Netherlands. https://doi.org/10.2118/184615-MS.
Pollock, J., Stoecker-Sylvia, Z., Veedu, V., Panchal, N., Elshahawi, H., 2018. Machine learning for improved directional drilling. In: Offshore Technology Conference, 30 April - 3 May. Offshore Technology Conference, Houston, Texas, USA. https://doi.org/10.4043/28633-MS.
Powell, M.J.D., 1987. Algorithms for Approximation. Clarendon Press, New York, NY, USA, pp. 143–167 chapter Radial Basis Functions for Multivariable Interpolation: A Review.
Quinlan, J.R., 1986. The effect of noise on concept learning. In: Machine Learning: Artificial Intelligence Approach ume 2. Morgan Kaufmann Publishers Inc., pp. 149–166 (chapter 6).
Rahimzadeh, H., Mostofi, M., Hashemi, A., 2011. A new method for determining Bourgoyne and Young penetration rate model constants a new method for determining Bourgoyne and Young penetration rate model constants. Pet. Sci. Technol. 29, 886–897. https://doi.org/10.1080/10916460903452009.
Rahimzadeh, H., Mostofi, M., Hashemi, A., Salahshoor, K., 2010. Comparison of the penetration rate models using field data for one of the gas fields in Persian Gulf area. In: International Oil and Gas Conference and Exhibition in China, 8-10 June, Beijing, China. Society of Petroleum Engineers, Beijing, China. https://doi.org/10.2118/131253-MS.
Rahmanifard, H., Plaksina, T., 2018. Application of artificial intelligence techniques in the petroleum industry: a review. Artif. Intell. Rev. 1. https://doi.org/10.1007/s10462-018-9612-8.
Ren, Y., Zhang, L., Suganthan, P., 2016. Ensemble classification and regression-recent developments, applications and future directions [review article]. IEEE Comput. Intell. Mag. 11, 41–53. https://doi.org/10.1109/MCI.2015.2471235.
Ross, T.J., 2010. Fuzzy Logic with Engineering Applications. John Wiley, Chichester.
Rumelhart, D.E., Hinton, G.E., Williams, R.J., 1986. Learning internal representations by error propagation. In: Parallel Distributed Processing: Explorations in the Microstructure of Cognition, vol. 1. MIT Press, Cambridge, MA, USA, pp. 318–362. (chapter 8). https://web.stanford.edu/class/psych209a/ReadingsByDate/02_06/PDPVolIChapter8.pdf.
Sabah, M., Talebkeikhah, M., Wood, D.A., Khosravanian, R., Anemangely, M., Younesi, A., 2019. A machine learning approach to predict drilling rate using petrophysical and mud logging data. Earth Sci. Inform. https://doi.org/10.1007/s12145-019-00381-4.
Sagi, O., Rokach, L., 2018. Ensemble learning: a survey. Wiley Interdiscip. Rev.: Data Min. Knowl. Discov. 8, e1249. https://doi.org/10.1002/widm.1249.
Salgado, C.M., Azevedo, C., Proença, H., Vieira, S.M., 2016. Noise versus Outliers. Springer International Publishing, Cham, pp. 163–183. https://doi.org/10.1007/978-3-319-43742-2_14.
Schmidhuber, J., 2015. Deep learning in neural networks: an overview. Neural Netw. 61, 85–117. https://doi.org/10.1016/j.neunet.2014.09.003.
Seifabad, M.C., Ehteshami, P., 2013. Estimating the drilling rate in Ahvaz oil field. J. Pet. Explor. Prod. Technol. 3, 169–173. https://doi.org/10.1007/s13202-013-0060-3.
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. Probl. Eng. 2016, 13. https://doi.org/10.1155/2016/3575380. article ID 3575380.
Shokouhi, S.V., Skalle, P., Aamodt, A., 2014. An overview of case-based reasoning applications in drilling engineering. Artif. Intell. Rev. 41, 317–329. https://doi.org/10.1007/s10462-011-9310-2.
Skalle, P., Aamodt, A., ErikGundersen, O., 2013. Experience transfer for process improvement. Eng. Appl. Artif. Intell. 26, 2206–2214. https://doi.org/10.1016/j.engappai.2013.06.012.
Skjerpen, T., Storrøsten, H.B., Rosendahl, K.E., Osmundsen, P., 2018. Modelling and forecasting rig rates on the Norwegian continental shelf. Resour. Energy Econ. 53, 220–239. https://doi.org/10.1016/j.reseneeco.2018.05.004.
Smola, A.J., Schölkopf, B., 2004. A tutorial on support vector regression. Stat. Comput. 14, 199–222. https://doi.org/10.1023/B:STCO.0000035301.49549.88.
Soares, C., Daigle, H., Gray, K., 2016. Evaluation of PDC bit ROP models and the effect of rock strength on model coefficients. J. Nat. Gas Sci. Eng. 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. 172, 934–959. https://doi.org/10.1016/j.petrol.2018.08.083.
Sugeno, M., 1985. Industrial Applications of Fuzzy Control. Elsevier Science Inc., New York, NY, USA.
Sui, D., Nybø, R., Azizi, V., 2013. Real-time optimization of rate of penetration during drilling operation. In: 2013 10th IEEE International Conference on Control and Automation. (ICCA), pp. 357–362. https://doi.org/10.1109/ICCA.2013.6564893.
Suparta, W., Alhasa, K., 2016. Modeling of tropospheric delays using ANFIS. In: Adaptive Neuro-Fuzzy Interference System. Springer International Publishing, Cham, pp. 5–18. https://doi.org/10.1007/978-3-319-28437-8_2. (chapter 2).
Suykens, J., Gestel, T.V., Brabanter, J.D., Moor, B.D., Vandewalle, J., 2002. Least Squares Support Vector Machines. World Scientific, Singapore. https://doi.org/10.1142/5089.
Swearingen, T., Drevo, W., Cyphers, B., Cuesta-Infante, A., Ross, A., Veeramachaneni, K., 2017. Atm: a distributed, collaborative, scalable system for automated machine learning. In: 2017 IEEE International Conference on Big Data. (Big Data), pp. 151–162. https://doi.org/10.1109/BigData.2017.8257923.
Szlek, J., Mendyk, A., 2018. Cran-r project. fscaret: automated feature selection from 'caret'. URL: https://CRAN.R-project.org/package=fscaret.
Takagi, T., Sugeno, M., 1985. Fuzzy identification of systems and its applications to modeling and control. IEEE Trans. Syst. Man Cybern. SMC 15, 116–132. https://doi.org/10.1109/TSMC.1985.6313399.
Tansev, E., 1975. A heuristic approach to drilling optimization. In: Fall Meeting of the Society of Petroleum Engineers of AIME. Society of Petroleum Engineers, Dallas, Texas, pp. 18. https://doi.org/10.2118/5546-MS.
Tavares, R.M., 2006. Interpretação e Análise de Dados de Perfuração em Poços de Petróleo (Portuguese). Dissertação de mestrado. Universidade Estadual de Campinas. http://repositorio.unicamp.br/bitstream/REPOSIP/263676/1/Tavares_RogerioMartins_M.pdf.
Teale, R., 1965. The concept of specific energy in rock drilling. Int. J. Rock Mech. Min. Sci. Geomech. Abstr. 2, 57–73. https://doi.org/10.1016/0148-9062(65)90022-7.
Tewari, S., Dwivedi, U.D., 2017. A novel neural network framework for the prediction of drilling rate of penetration. In: APCEC17 International Conference on Advances in Petroleum Chemical & Energy Challenges.
Thonhauser, G., 2004. Using real-time data for automated drilling performance analysis. Oil Gas Eur. Mag. 4, 170–173.
Tsukamoto, Y., 1979. An approach to fuzzy reasoning method. In: Grupta, M., Ragade, R., Yager, R. (Eds.), Advances in Fuzzy Set Theory and Applications. Elsevier, Amsterdam, pp. 137–149.
Valisevich, A., Ruzhnikov, A., Bebeshko, I., Moreno, R., Zhentichka, M., Bits, S., 2015. Drillbit optimization system: real-time approach to enhance rate of penetration and bit wear monitoring. In: SPE Russian Petroleum Technology Conference, 26-28 October. Society of Petroleum Engineers, Moscow, Russia. https://doi.org/10.2118/176517-MS.
Vapnik, V.N., 1998. Statistical Learning Theory. Wiley, New York, NY.
Vapnik, V.N., 2000. The Nature of Statistical Learning Theory, second ed. Springer, New York, NY. https://doi.org/10.1007/978-1-4757-3264-1.
Wallace, S.P., Hegde, C.M., Gray, K.E., 2015. A system for real-time drilling performance optimization and automation. In: SPE Middle East Intelligent Oil and Gas Conference and Exhibition, 15-16 September, Abu Dhabi. UAE, Society of Petroleum Engineers. https://doi.org/10.2118/176804-MS.
Wang, Y., Salehi, S., 2015. Application of real-time field data to optimize drilling hydraulics using neural network approach. J. Energy Resour. Technol. 137, 062903. https://doi.org/10.1115/1.4030847.
Warren, T., 1987. Penetration rate performance of roller cone bits. SPE Drill. Eng. 2, 9–18. https://doi.org/10.2118/13259-PA.
Werbos, P.J., 1982. Applications of advances in nonlinear sensitivity analysis. In: Drenick, R.F., Kozin, F. (Eds.), System Modeling and Optimization. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 762–770.
Wong, T.T., 2015. Performance evaluation of classification algorithms by k-fold and leave-one-out cross validation. Pattern Recognit. 48, 2839–2846. https://doi.org/10.1016/j.patcog.2015.03.009.
Xue, B., Zhang, M., Browne, W.N., Yao, X., 2016. A survey on evolutionary computation approaches to feature selection. IEEE Trans. Evol. Comput. 20, 606–626. https://doi.org/10.1109/TEVC.2015.2504420.
Yavari, H., Sabah, M., Khosravanian, R., Wood, D.A., 2018. Application of an adaptive neuro-fuzzy inference system and mathematical rate of penetration models to predicting drilling rate. Iran. J. Oil Gas Sci. Technol. 7, 73–100. https://doi.org/10.22050/IJOGST.2018.83374.1391.
Yılmaz, S., Demircioglu, C., Akin, S., 2002. Application of artificial neural networks to optimum bit selection. Comput. Geosci. 28, 261–269. https://doi.org/10.1016/S0098-3004(01)00071-1.
Young, F., 1969. Computerized drilling control. J. Pet. Technol. 21, 483–496. https://doi.org/10.2118/2241-PA.
Yuan, P., Yan, T., Feng, J., Chang, L., 2009. Application of case-based reasoning method on drilling parameter optimization. In: 2009 International Conference on Computational Intelligence and Software Engineering, Wuhan, China, pp. 1–5. https://doi.org/10.1109/CISE.2009.5364105.
Zadeh, L., 1965. Fuzzy sets. Inf. Control 8, 338–353. https://doi.org/10.1016/S0019-9958(65)90241-X.
Zare, J., Shadizadeh, S.R., 2014. Managed pressure drilling to increase rate of penetration and reduce formation damage and reduce formation damage. Pet. Sci. Technol. 32, 1833–1842. https://doi.org/10.1080/10916466.2010.540618.
Zhou, A., Qu, B.Y., Li, H., Zhao, S.Z., Suganthan, P.N., Zhang, Q., 2011. Multiobjective evolutionary algorithms: a survey of the state of the art. Swarm Evol. Comput. 1, 32–49. https://doi.org/10.1016/j.swevo.2011.03.001.