A New Hybrid Analytical-Machine Learning Method for Real-Time ROP Modeling

Ali Reza Moazzeni [1] , Ehsan Khamehchi [1,*]

[1]   Department of Petroleum Engineering, Amirkabir University of Technology (Tehran Polytechnic), Tehran, Iran
*   Correspondence: khamehchi@aut.ac.ir;

Abstract: Real-time drilling optimization refers to operations and equipment that could minimize total drilling costs. Drilling speed that is called the rate of penetration (ROP) in the drilling industry can be used as a good indicator for the performance evaluation of the drilling operation. Real-time control for drilling ROP is limited to just a few controllable parameters during drilling operations, that is, WOB, RPM, and hydraulics. These parameters can be controlled from the surface by the driller in real-time. In the traditional methods of ROP modeling, an inflexible equation could be developed between some important effective drilling parameters such as weight on the bit or bit rotational speed and drilling rate of penetration. These models had a low degree of accuracy, and they were not applicable in the newly drilled wells even in the same field with an acceptable degree of accuracy. In this study, a new real-time continues-learning method for ROP modeling was developed. In this method, as the drilling operation gets starts and the drilling data reaches the surface, ROP modeling starts, and as the drilling continues, the model accuracy increases. For the method evaluation, 5 famous existing analytical drilling model was selected. Also, a new ROP model was developed in this work. All of these 6 models contain some constant coefficients that were obtained using a new machine learning method named Rain Optimization Algorithm. In the end, the accuracy of the models was compared. Results show that the presented method for ROP modeling is a very flexible method with a high degree of accuracy that can be easily used in any formation. Also, the newly presented model could increase the accuracy of ROP prediction from 75% to 81%.

Keywords: Rate of penetration; Real-Time; Continues learning; ROP modeling; Rain Optimization Algorithm.

1. Introduction

Drilling optimization refers to operations and equipment that could minimize total drilling costs. Drilling speed that is called the rate of penetration (ROP) in the drilling industry can be used as a good indicator for the performance evaluation of the drilling operation. Optimizing ROP to reduce drilling costs in oil and gas development wells is the permanent objective of drilling researchers [1-6]. Exact Prediction of ROP can lead to the improvement of drilling operation efficiency, causing drilling time and cost reduction [7]. Also, having a correct forecast of drilling rate and an exact ROP model can help to use a precise optimization algorithm leading to drilling parameter optimization and increasing drilling operation efficiency. An important point that should be noticed is that drilling speed increase might not cause drilling efficiency increase in every situation. However, it might cause some drilling

problems such as hole cleaning problems, wellbore instability or bit failure, causing drilling time and drilling cost to growth [8].

There are lots of direct and indirect equations for drilling rate prediction that relate some of the drilling parameters to the rate of penetration. Some of these equations use parameters that are not easy to obtain without performing complicated experiments or using uncertain empirical equations [9]. Some of these equations and methods of utilizing them would be presented in the next section.

Many parameters affect the rate of penetrations, some of them are controllable during drilling operations, and some of them are uncontrollable. Sikes (1936) determined that exerted weight on the bit (WOB), the rotational speed of bit (RPM), the diameter of the hole that is being drilled (d_h), hole cleaning and hydraulics, bit type, and formation properties are the most important parameters affecting drilling ROP. Some of these parameters are a function of some other parameters [10]. For example, when a PDC bit is used for drilling a hole, PDC cutters number, side rake angle of the cutters, back rake angle of the cutters, size of cutters and arrangement of the cutters on the bit can influence on the drilling ROP strongly [11]. In the same manner, drilling fluid density and mud rheological properties are some of the important properties of drilling fluids' effect on ROP when talking about hydraulics [12,13]. Some of these parameters, such as bit type, drill string design, or rheological properties of drilling fluid, are determined before starting to drill a hole and are constant while drilling. These parameters are described as uncontrollable drilling parameters. On the other hand, some of the other drilling parameters such as WOB, RPM, and pump flow rate can be changed in real-time while drilling in the surface by the driller, which are called controllable parameters [14].

Therefor real-time control for drilling ROP is limited to just a few controllable parameters during drilling operations, that is, WOB, RPM, and hydraulics (usually flow rate of the pump). These parameters can be controlled from the surface by the driller in real-time.

Several investigators worked on the effect of WOB, N, and q on the drilling rate of penetration [15-17]. Dupriest *et al*. (2005) stated that the effect of WOB on the ROP could be considered in three regions [17]. At low WOBs, the bit is not engaged in the formation properly. In this region increasing, WOB has a minimal effect on ROP since there is no enough pressure on the bit for drilling to start. When WOB gets to a certain value, the intermediate region starts. In this section, ROP increases linearly with respect to WOB, and this trend continues until getting to the maximum point, which is called the founder point. This point is the optimum WOB for the drilling. After this point, increasing weight on the bit has a minimal effect on ROP increase. Even ROP can starts to decrease because of hole cleaning problems, drill pipe vibration or bit balling as it was considered by Gandelman (2012) [18].

Bourgoyne *et al*. (1986), stated that ROP has a linear relationship with the bit rotational speed (RPM) [16]. At the high RPMs, this relation deviates from the linearity because of drill pipe vibration or hole cleaning problems, as was stated by Gandelman (2012) [18].

Effect of pump flow rate (q) on the ROP can be considered from two aspects [19]: Pump flow rate increase can elevate friction pressure drop in the annulus causing drilling fluid pressure increase in the bottom of the hole, especially on the cutting at the bottom of the hole. This issue can hold the cutting on the bottom of the hole, causing chips to hold down phenomena and ROP decrease. On the other hand, the pump flow rate increase can raise the pressure drop across the bit, causing better bottom hole cleaning. From this point of view, increasing the pump flow rate would increase ROP.

It should be noticed that modeling drilling speed is a difficult problem since lots of parameters affect ROP, and some of these parameters are not easy to obtain. Many attempts had been made to develop an ROP model from available data of previously drilled wells. These models tried to be used in the new adjacent wells, but results were not satisfactory. In this work, an ROP model would be obtained from real-time drilling data, and this model would be used for optimizing the ROP in the remained parts of the same well. While the drilling is in progress, more data would be obtained, and the model would be updated. Once the drilled segment is longer the predicted model would be more accurate.

2. Materials and Methods

2.1. ROP models.

In the 1950s and 1960s, many attempts had done to develop a relationship between WOB, RPM, and ROP. These equations that are called R-W-N equations (representing ROP, WOB, and RPM) was developed regardless of the hole cleaning conditions, vibration, and depth of the hole. In fact, the researchers were seeking a universal equation for every drilling conditions. Some of these equations are given in Table 1 where W determines weight on bit, N represents bit rotational speed, D is the bit diameter, and R shows the rate of penetration.

Table 1. Some of the important R-W-N equations developed in the 1940 to 1960 decades [12].

Developer	R-W-N equation
van Lingen	$R = W^{1.25}N^{0.8}$
Gatlin	$R = a + bWN^c \qquad c < 1$
Eckel , Bielstein and Cannon	$R = aWN^{0.5}$
Wardroup and Cannon	$R = aWN^{0.4}$
Moore	$R = aWN$
Cunningham	$R = \dfrac{aWN^b}{D} \qquad b < 1$
Brantley	$R = a + bWN$

Bingham (1964) significantly changed the path of R-W-N equations by adding an imperial exponent to the WOB [20]. He also could increase the applicability of his equation by confining that to a special formation and adding coefficient 'a' as formation drillability to it.

$$R = a\left(\frac{W}{D}\right)^b N \qquad (1)$$

That a and b in this equation are dimensionless coefficients determined separately for each formation, W determines weight on the bit (lb), N represents bit rotational speed in RPM, D is the bit diameter (in), and R shows the rate of penetration. This model would be called BM (an abbreviation of the Bingham model) in this study.

The effect of hydraulics on the drilling rate of penetration was first entered into the R-W-N equations by Eckel (1967) [13]. He noticed that the nozzle velocity and flow rate had an important effect on ROP as follow:

$$R = a_1 w^{a_2} N^{a_3}\left(\frac{a_4 q\rho}{d_n \mu}\right)^{0.5} \qquad 2 < \frac{a_4 q\rho}{d_n \mu} < 100 \qquad (2)$$

Where in this equation a_1 to a_5 are dimensionless coefficients determined separately for each formation, q is the flow rate in GPM, ρ is the specific gravity of drilling fluid, d_n is nozzle diameter in in., and μ is fluid viscosity in cp. This model would be called EM (an abbreviation of the Eckel model) in this study.

Bourgoyne and Young (1986) developed the most comprehensive ROP equation until that time [16]. They stated that 8 factors are effective in drilling speed as follow:

$$ROP = f_1 f_2 f_3 f_4 f_5 f_6 f_7 f_8 \qquad (3)$$

Drilling parameter items considered to have an effect on the rate of penetration are as summarized below:

$$f_1 = e^{2.303 a_1} = K \qquad (4)$$

$$f_2 = e^{2.303 a_2 (10000 - TVD)} \qquad (5)$$

$$f_3 = e^{2.303 a_3 TVD^{0.69}(g_p - 9.0)} \qquad (6)$$

$$f_4 = e^{2.303 a_4 TVD(g_p - \rho_c)} \qquad (7)$$

$$f_5 = \left[\frac{\left(\frac{w}{d_b}\right) - \left(\frac{w}{d_b}\right)_t}{4 - \left(\frac{w}{d_b}\right)_t} \right] \qquad (8)$$

$$f_6 = \left(\frac{N}{60}\right)^{a_6} \qquad (9)$$

$$f_7 = e^{-a_7 h} \qquad (10)$$

$$f_8 = \left(\frac{F_j}{1000}\right)^{a_8} \qquad (11)$$

In these equations,

TVD = true vertical depth, ft

g_p = pore pressure gradient, lbm/gal

ρ_c = equivalent circulating density, ppg

$\left(\frac{w}{d_b}\right)_t$ = threshold bit weight per inch of bit, 1000lbf/in

h = fractional tooth dullness

F_j = hydraulic impact force beneath the bit, lbf

$a_1 \ to \ a_8$ = constants that must be chosen based on local drilling conditions.

In Equation 3, f_1 is representative of formation strength; f_2 shows the effect of depth and it is normalized for the depth of 10000 ft; f_3 determines the effect of pore pressure normalized to pore pressure gradient of 9 ppg/ft; f_4 displays the effect of bottom hole differential pressure; f_5 demonstrates the effect of weight on bit, normalized to 4000 lbs; f_6 illustrates the effect of bit rotary speed normalized to 60 RPM; f_7 shows the effect of bit tooth wear and finally f_8 determines the effect of hydraulics using impact force effect. This model would be called BYM (an abbreviation of the Bourgoyne and Young model) in this study.

In equation 11, the impact force can be calculated as follow:

$$F_j = 0.000518 \rho q v_n \qquad (12)$$

$$v_n = 0.3208 \frac{q}{TFA} \qquad (13)$$

By substituting equation 13 to equation 12 it can be concluded that $F_j \propto q^2$. Also, Bourgoyne and Young (1986) stated that it is possible to use hydraulic horsepower (HHP) instead of impact force to consider the effect of hydraulics in equation 3 as follow [16]:

$$f_8 = (HSI)^{a_8} \qquad (14)$$

$$HSI = \frac{HHP}{A_b} = \frac{\frac{\Delta p_b q}{1714}}{\frac{\pi d_b^2}{4}} \qquad (15)$$

$$\Delta p_b = \frac{\rho q^2}{10861 TFA^2} \qquad (16)$$

Where in the Equations 14 to 16 HSI represents of hydraulic horsepower drop related to every square inch of bit, Δp_b is the pressure drop in the bit in psi, d_b is bit diameter (in), and TFA is the total flow area of the bit nozzles (in^2). By substituting equation 15 to equation 16, it can be concluded that $HSI \propto q^3$.

Some of the parameters in equation 3, when this equation is used for a special hole with a specific formation, are constant [21]. Some of these parameters are bit diameter, nozzle diameter, apparent viscosity, pore pressure gradient, and mud weight. Therefore when this equation is used as a real-time optimization equation for drilling a specific formation, factors f_1, f_3, f_4 and f_8 are nearly constant, and they can be considered as the constant multipliers in Bourgoyne and Young model (BYM).

$$a = f_1 f_3 f_4 f_8 \qquad (17)$$

Substituting Equation 17 to equation 3, and supposing threshold WOB to be zero yields:

$$R = a \times e^{2.303 a_2 (10000 - TVD)} \times \left(\frac{w}{4d_b}\right)^{a_5} \times \left(\frac{N}{60}\right)^{a_6} \times \left(\frac{F_j}{1000}\right)^{a_8} \qquad (18)$$

Gray (2018) more simplified Equation 18 and presented the following equation as modified Bourgoyne and Young model (MBYM):

$$R = a_1 TVD^{a_2} W^{a_5} N^{a_6} q^{a_8} \qquad (19)$$

2.2. Developing a new ROP model using the concept of MSE.

Mechanical specific energy (MSE) is the amount of work done for the excavating unit volume of rock [22]. In rotary table drilling, work is done both by the piercing the bit, WOB (lb), and the exerted force while rotation of bit or torque, T (lb-ft). Teale (1965) calculated the MSE as follow [22]:

$$Work = W \times R + 60 \times 2\pi N \times T \qquad (20)$$

Where:

N: rotation speed in rev/min

W: the weight on the bit in lb

T: the torque in lb-ft

R: penetration rate in ft/hr

Work: the work is done for removing the rock in lb-ft/hr

Also, the volume of excavated rock in one hour is:

$$V = A \times R \qquad (21)$$

Where A is the cross-section area of the hole in in^2, and V is the volume of removed rock in one hour. So, the mechanical specific energy (in lb/in^2) can be computed by dividing work by volume:

$$MSE = \frac{W}{A} + \frac{60 \times 2\pi N \times T}{A \times R} \qquad (22)$$

Cherif (2012) stated that each bit has a mechanical efficiency related to its cutter size and structure [22]. Including mechanical efficiency (E_m) to the Teal's equation, the mechanical specific energy can be expressed as:

$$MSE = \frac{E_m \cdot Work}{V} = \frac{E_m}{A}\left(W + \frac{60 \times 2\pi \times N \times T}{R}\right) \qquad (23)$$

Chen *et al*. (2014) stated that provided mechanical energy on the surface has a great difference with the mechanical energy received by the bit due to friction between pipes and

borehole, especially in directional and horizontal drilling [24]. So he formulated a relation between measured surface weight on bit and bottom hole weight on the bit:

$$W_b = W.e^{-\mu_s.\gamma_b} \qquad (24)$$

$$T = \frac{\mu_b.D_b.W.e^{-\mu_s.\gamma_b}}{36} \qquad (25)$$

Therefore when torque data is available in vertical wells:

$$\text{MSE} = \frac{E_m}{A}.\left(W.e^{-\mu_s.\gamma_b} + \frac{10.47\text{N}.\mu_b.D_b.W.e^{-\mu_s.\gamma_b}}{R}\right) \qquad (26)$$

Where in this equation:

Em: Mechanical efficiency of the bit

γ_b: Bit sliding coefficient (between 0.3 and 0.85)

μ_s: Drill string sliding coefficient (between 0.25 and 0.4)

Letting

$a_1 = \mu_s.\gamma_b$

$a_2 = 10.47.\mu_b.D_b$

$a_3 = a \; constant \; for \; unit \; changes$

Equation 26 can be summarized as follow:

$$a_3.MSE = \frac{E_m.W.e^{-a_1}}{A}\left(1 + \frac{a_2.N}{R}\right) \qquad (27)$$

Rearranging Equation 27 for ROP, yields:

$$R = \frac{a_2.N.W.e^{-a_1}}{\frac{a_3.A.MSE}{E_m} - W.e^{-a_1}} \qquad (28)$$

Dupriest *et al*. (2005) stated that Bits tend to transfer 30% to 40% of their input energy to the rock, even when operating at peak performance [17].

Pessier (1992) mentioned that in the best situation, the minimum amount of MSE has a value of the order of confining compressive strength (CCS) of the rock. Therefore, in the most efficient situation, the minimum amount of MSE is equal to CCS, and just about 35% of the energy is transferred by the bit. This means that in Equation 28, CCS can be used instead of MSE [25]. Also, we can suppose the mechanical efficiency of the bit to be 0.35. The amount of CCS can be found from adjacent wells or can be calculated from the log or empirical equations. Therefore, Equation 28 can be rewritten as follow:

$$R = \frac{a_2 N.T}{\frac{a_3.A.CCS}{0.35} - W.e^{-a_1}} \qquad (29)$$

This model that is discussed by Mozzeni and Khamehchi (2019) would be call MKM1 in this study [23].

2.3. New ROP model by including torque.

In neither of the previous models for ROP, the parameter of torque is not available. Torque is a function of weight on bit, bit properties, and formation properties, and by entering it to the ROP model, all of these parameters would be included in the model. It is possible to use from the available torque data for developing or fitting a model, but the main problem is while using this model where the torque is not a controllable parameter. Torque is a function of WOB, bit properties, and formation properties. So it is not possible to adjust the torque exactly on an arbitrary value while drilling, but it is possible to adjust the WOB. On the other hand, the drilling rate has a high dependency on the torque, and it is hard to ignore such important data.

The proposed method for this problem is to predict the torque as a function of WOB using a search algorithm from the real-time drilling data. As it was proposed by Chen *et al.* (2014) torque is a function of WOB as follow [24]:

$$T = \frac{\mu_b . D_b . W . e^{-\mu_s . \gamma_b}}{36} \tag{30}$$

Letting $b_1 = \frac{\mu_b . D_b}{36}$ and $b_2 = -\mu_s . \gamma_b$ this equation would be more simplified as:

$$T = b_1 . W . e^{-b_2} \tag{31}$$

In this equation b_1 and b_2 are dimensionless constants that can be tuned from real-time drilling data. So, having torque, the following equation can be proposed for predicting ROP as a new model (MKM2) according to previous discussions on effecting parameters on ROP:

$$R = \frac{a_1 . W^{a_2} N^{a_3} T^{a_4} HSI^{a_5}}{A} \tag{32}$$

After model fitting and obtaining a_1 to a_5 and b_1 to b_2 from drilling data of a special section, by substituting Equation 31 to Equation 32 the following model would be applicable for ROP prediction or ROP optimization in the next drilling sections of a hole:

$$R = \frac{a_1 . W^{a_2} N^{a_3} (b_1 . W . e^{-b_2})^{a_4} HSI^{a_5}}{A} \tag{33}$$

2.4. Model performance evaluation.

In all of six previously discussed ROP models (BM, EM, BYM, MBYM, MKM1, and MKM2), there are empirical coefficients that should be calculated in real-time. For this purpose, the constants would be guessed using a metaheuristic search algorithm, and the amount of ROP for each model would be calculated. The calculated ROP would be compared with the field ROP. The most common method for calculating cost function is the sum of the squares of difference:

$$C = \sum_{i=1}^{M} \left(R_{field,i} - R_{model,i} \right)^2 \tag{34}$$

Which in this equation, C is the cost function value, M is the number of training data, R_{field} is the value of obtained ROP in the field and R_{model} is the value of calculated ROP using the model.

The main objective of model fitting is minimizing the cost function. Therefore, after finding the best constants for each model, these models should be compared to each other. For this purpose, the value of error would be calculated using another part of the data:

$$RMSE = \sqrt{\frac{1}{K} \sum_{i=1}^{K} \left(R_{field,i} - R_{model,i} \right)^2} \tag{35}$$

In this equation, RMSE is the root mean square error, and k represents the number of test data. An important point about RMSE is that the importance of the value of error is determined related to the ROP value interval. For resolving this problem, RMSE can be normalized to the mean of drilling ROP as follow:

$$Normalized\ RMSE = \frac{RMSE}{\frac{1}{k} \sum_{i=1}^{k} R_{field,i}} \tag{36}$$

In this study, the cost function value for each model would be calculated using Equation 34 and the error value for each model would be compared using Equations 35 and 36.

2.5. Model fitting method.

In this study, it will be used for the Rain Optimization Algorithm (ROA), one of the newest metaheuristic search algorithm introduced by Moazzeni and Khamehchi (2019) [26].

In this algorithm, each solution of the problem can be modeled by a raindrop. Depending on the problem, some points in the answer space can be selected randomly as the raindrops fall in the ground randomly. The main property of each drop of rain is its radius.

The radius of every raindrop can be reduced as time goes by, and it can be increased as a raindrop is connected to other drops. When the initial population of answers is produced, the radius of each droplet can be assigned randomly in an appropriate range. In each iteration, every droplet checks its neighborhood dependent on its size. Single droplets that are not still connected to any other droplet, just check for the end limit of the place that it has covered.
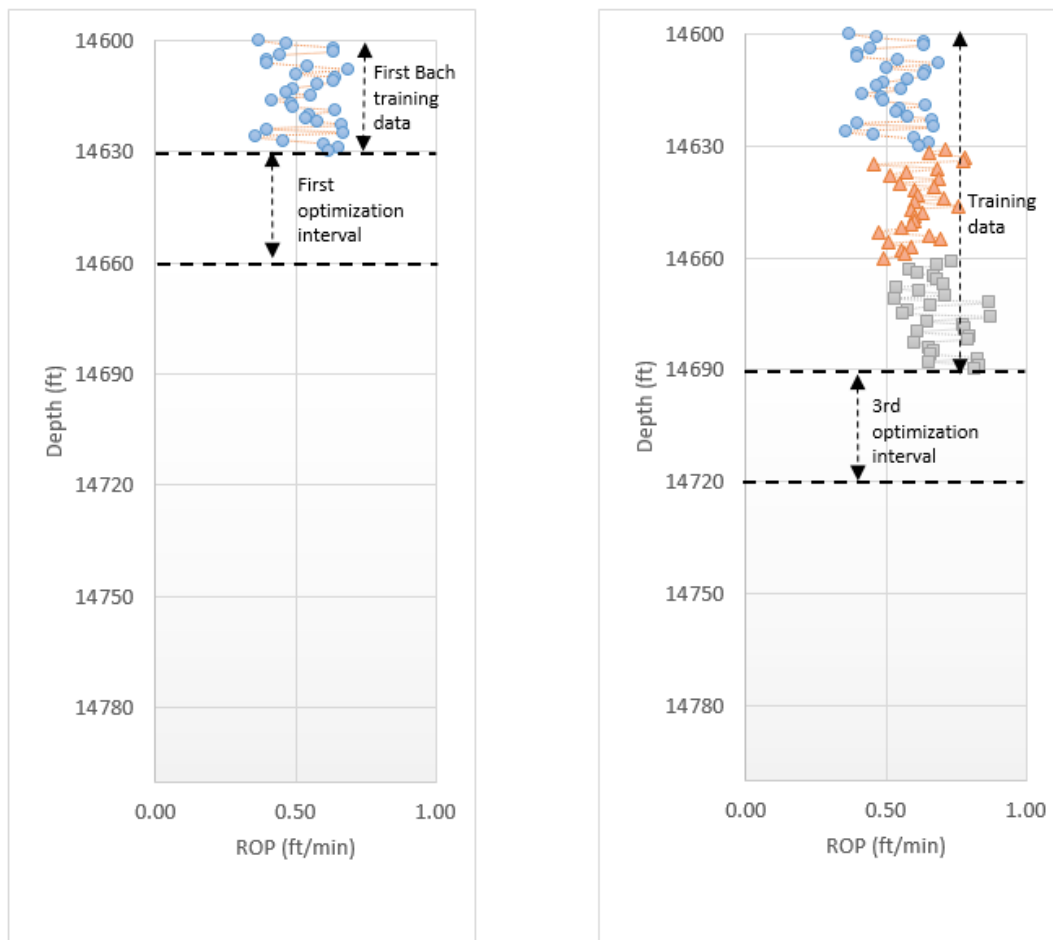

Figure 1. 30ft drilling optimization intervals in a 200ft long formation.

When we are solving a problem in n-dimensional space, every droplet consists of n variable. So at the first step, the lower and upper limit of variable one will be checked as these limits would be determined by the radius of the droplet. At the next step, two endpoints of variable two would be tested, and this is continued until the last variable. In this stage, the cost of the first droplet would be updated by moving it downward. This is not the end action for this droplet, and while cost function is reducing, it will move downward in the same direction. This action will be performed for all droplets, then the cost and position of all droplets will be assigned.

Therefore this algorithm would find the constants in the selected ROP models by simulating them to rain droplets. By starting drilling operations, the first batch of drilling data would be produced. In each step, the newly obtained data, in conjunction with the previously obtained data, would be entered into ROA, and the best value of the constants of the model

would be found. This method of solving a problem that is called the real-time continues-learning method is shown in Figure 1.

As it can be seen in Figure 1, a formation with a 200-foot hight is divided into 30-foot sections. When the first batch data related to the first 30-foot was obtained, the constants of each ROP model would be calculated separately by ROA, and the value of error for each model would be calculated. Then the best model would be selected and using that model, optimum drilling parameters for the next 30-foot would be calculated. Using these drilling parameters, drilling proceeded until the second batch of data related to the second 30-foot segment obtained. Again using the first and second batch of data, drilling optimization would be performed, and this process would be continued to the end of the formation. The important point is that the length of the optimization segment can be reduced or extended arbitrarily, but it should be noticed that it is not common to change drilling parameters in very short periods.

3. Results and Discussion

Model fitting was done on six analytical ROP models (BM, EM, BYM, MBYM, MKM1, and MKM2) using ROA, and the error of each model was calculating using Equations 35 and 36. Used data for this work is from K4 formation in the south Pars gas field in Iran from the depth of 14608 to 14800 ft. Table 2 shows the first batch of the used data. Available data has been normalized as follow before entering to ROA:

$$W_n = \frac{W}{4000} \tag{37}$$

$$N_n = \frac{N}{60} \tag{38}$$

$$T_n = \frac{T}{10000} \tag{39}$$

$$q_n = \frac{q}{400} \tag{40}$$

$$TVD_n = \frac{TVD}{10000} \tag{41}$$

$$F_{j_n} = \frac{F_j}{1000} \tag{42}$$

$$HSI_n = \frac{HSI}{1000} \tag{43}$$

Table 2. A sample of used drilling data for solving the ROP models.

data number	Depth (ft)	wob (lb)	Torque (lb*ft)	RPM (rpm)	Spp (psi)	q (gpm)	ROP (ft/min)
1	14608	7093.98	18542	121	1932	490	0.533413
2	14609.66	7066.8	18777	121	1932	490	0.556653
3	14611.32	7057.74	18807	121	1932	490	0.546693
4	14612.98	7030.56	18789	121	1930	490	0.573253
5	14614.64	7066.8	18759	121	1923	490	0.540607
6	14616.3	7062.27	19086	122	1922	490	0.5561
7	14617.96	6998.85	19439	119	1916	490	0.62416
8	14619.62	7098.51	18708	122	1909	490	0.390653
9	14621.28	6545.85	18212	122	1867	491	0.93624
10	14622.94	6754.23	18459	123	1865	491	0.604793
11	14624.6	6491.49	19022	121	1865	491	0.785733
12	14626.26	6215.16	19658	119	1862	491	1.00264
13	14627.92	6505.08	19075	121	1859	491	0.765813
14	14629.58	6713.46	18384	123	1851	491	0.553333
15	14631.24	6654.57	18544	122	1851	490	0.639653
16	14632.9	6872.01	18542	122	1852	490	0.680047
17	14634.56	6872.01	18558	122	1851	490	0.6889
18	14636.22	6804.06	19284	122	1847	490	0.850473
19	14637.88	6930.9	18508	123	1849	490	0.702733

data number	Depth (ft)	wob (lb)	Torque (lb*ft)	RPM (rpm)	Spp (psi)	q (gpm)	ROP (ft/min)
20	14639.54	6890.13	18747	123	1846	490	0.720993
21	14641.2	6939.96	17983	124	1849	490	0.495787
22	14642.86	6935.43	17883	124	1845	490	0.455393
23	14644.52	6813.12	18589	124	1852	490	0.618073
24	14646.18	6749.7	18745	122	1854	490	0.747
25	14647.84	6795	18636	123	1854	490	0.67728
26	14649.5	6967.14	18011	124	1862	490	0.508513
27	14651.16	6962.61	18046	124	1863	490	0.545587
28	14652.82	6980.73	17731	125	1871	490	0.50962
29	14654.48	7003.38	17471	125	1872	490	0.45318
30	14656.14	6980.73	17505	124	1879	490	0.3984

In the proposed procedure in this study, after drilling a particular section of a hole (every 50 centimeters), a new data series will be added to the previous data set. In this case, all previous data will be used for solving the model. So by increasing the depth, available drilling data will be more complete, and the model will be more precise. Figure 2 shows the ROP and some of the drilling parameter values in the optimization interval with respect to the depth.
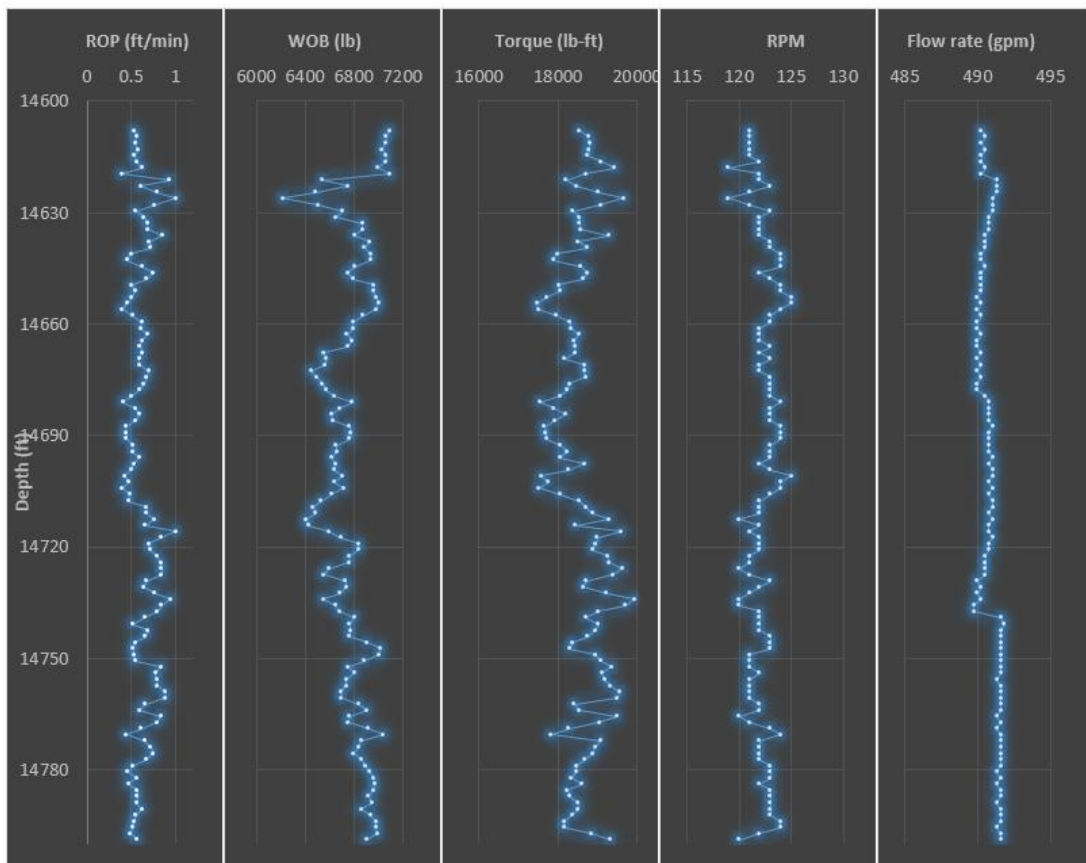

Figure 2. ROP, WOB,Torque, RPM, and Flow rate values vs. depth in the optimization intervals.

At the first trial, ROA will guess the amount of the a-constants in all six models separately. At the next iterations, this algorithm tries to change a-constants to reduce the cost function of each model, and this process would be continued for 100 iterations.

For solving this problem using ROA, the initial population was 100, the minimum amount of each variable was zero. Also, the initial radius of each raindrop was set to be 0.02, and rain speed was equal to rain adsorption that was 10 in each iteration, and 100 iterations for each run was considered.

Table 3. Obtained a-constants, RMSE, NORM RMSE, and Cost Function Value for each model, obtained using the first 30 data series input.

| Model Name	Model	a-constants	RMSE	NORM RMSE	Cost Function Value
BM	$R = a_1 \left(\dfrac{W}{D}\right)^{a_2} N$	$a_1 = 0.30$ $a_2 = 1.30\text{E-}10$	0.149	0.239	0.664
EM	$R = a_1 w^{a_2} N^{a_3} \left(\dfrac{a_4 q \rho}{d_n \mu}\right)^{0.5}$	$a_1 = 0.28$ $a_2 = 0.06$ $a_3 = 0.37$ $a_4 = 17.17$	0.147	0.236	0.649
BYM	$R = a_1 \times e^{2.303 a_2 (10000 - TVD)}$ $\times \left(\dfrac{w}{4 d_b}\right)^{a_3}$ $\times \left(\dfrac{N}{60}\right)^{a_4}$ $\times \left(\dfrac{F_j}{1000}\right)^{a_5}$	$a_1 = 2.052$ $a_2 = 2.27E - 05$ $a_3 = 0.0006$ $a_4 = 0.861$ $a_5 = 2.054$	0.147	0.236	0.647
MBYM	$R = a_1 TVD^{a_2} W^{a_3} N^{a_4} q^{a_5}$	$a_1 = 0.56$ $a_2 = 1.40$ $a_3 = 5.1E - 05$ $a_4 = 0.02$ $a_5 = 0.90$	0.145	0.233	0.629
MKM1	$R = \dfrac{a_2 N.T}{\dfrac{a_3.A.CCS}{0.35} - W.e^{-a_1}}$	$a_1 = 9.04E - 10$ $a_2 = 0.184$ $a_3 = 3$	0.161	0.259	0.780
MKM2	$R = \dfrac{a_1 . W^{a_2} N^{a_3} T^{a_4} HSI^{a_5}}{A}$	$a_1 = 0.010$ $a_2 = 0.010$ $a_3 = 0.723$ $a_4 = 2.714$ $a_5 = 2.557$	0.122	0.195	0.445
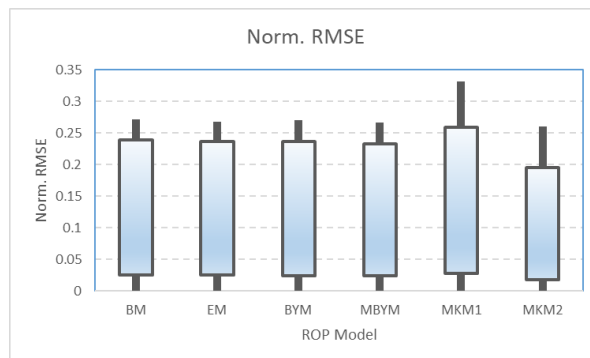
Figure 3. Trend of changing error while a-constants are tuning for all six models.

In the first attempt, just the first data set was used for finding a-constants of each model. Then, when second data sets arrived, first and second data series used for finding a-constants, and this process continued for 30 data series. Table 3 shows the obtained a-constants, RMSE, Normalized RMSE, and Cost Function Value for each model, obtained using the first 30 data series input.

As can be seen in Table 3 MKM2 model has the least amount of Normalized RMSE of 19.5 percent among the 6 considered models in this study. On the other hand, the MKM1 model has the maximum amount of RMSE (26 percent) between the considered models. The trend of changing error while a-constants are tuning for all six models can be seen in Figure 3.

In Figure 3, the performance of each model is shown using candlelight. Each candle has a body, a downward shadow, and an upward shadow. The minimum showed value with the body (opening point) shows the error value for the first data series. The maximum value of the body (close point) determines the last obtained error using all data series. The minimum amount

of error in the tuning process is shown by the downward shadow (min), and the maximum error is shown by the upward shadow (max). Figure 3 shows that the MKM2 model has the best performance in every four subjects compared with other models.

Table 4. Normalized data used for ROP model tuning.

Data point	depth	wob n	Torqe n	RPM n	Rop (ft/min)
1	4400	1.773495	1.8542	2.016667	0.533413
2	4400.5	1.7667	1.8777	2.016667	0.556653
3	4401	1.764435	1.8807	2.016667	0.546693
4	4401.5	1.75764	1.8789	2.016667	0.573253
5	4402	1.7667	1.8759	2.016667	0.540607
6	4402.5	1.765568	1.9086	2.033333	0.5561
7	4403	1.749713	1.9439	1.983333	0.62416
8	4403.5	1.774628	1.8708	2.033333	0.390653
9	4404	1.636463	1.8212	2.033333	0.93624
10	4404.5	1.688558	1.8459	2.05	0.604793
11	4405	1.622873	1.9022	2.016667	0.785733
12	4405.5	1.55379	1.9658	1.983333	1.00264
13	4406	1.62627	1.9075	2.016667	0.765813
14	4406.5	1.678365	1.8384	2.05	0.553333
15	4407	1.663643	1.8544	2.033333	0.639653
16	4407.5	1.718003	1.8542	2.033333	0.680047
17	4408	1.718003	1.8558	2.033333	0.6889
18	4408.5	1.701015	1.9284	2.033333	0.850473
19	4409	1.732725	1.8508	2.05	0.702733
20	4409.5	1.722533	1.8747	2.05	0.720993
21	4410	1.73499	1.7983	2.066667	0.495787
22	4410.5	1.733858	1.7883	2.066667	0.455393
23	4411	1.70328	1.8589	2.066667	0.618073
24	4411.5	1.687425	1.8745	2.033333	0.747
25	4412	1.69875	1.8636	2.05	0.67728
26	4412.5	1.741785	1.8011	2.066667	0.508513
27	4413	1.740653	1.8046	2.066667	0.545587
28	4413.5	1.745183	1.7731	2.083333	0.50962
29	4414	1.750845	1.7471	2.083333	0.45318
30	4414.5	1.745183	1.7505	2.066667	0.3984

So the MKM2 would be selected as the best model in this formation. The ROP model for this formation until this point can be considered as follow:

$$R = \frac{0.01 W^{0.01} N^{0.723} T^{2.714} HSI^{2.557}}{33.16} \tag{44}$$

This equation can more be simplified as follow:

$$R = 0.0003 W^{0.01} N^{0.723} T^{2.714} HSI^{2.557} \tag{45}$$

It is possible to use Equation 44 for predicting ROP in every depth. Having $WOB_n = 1.688$, $N_n = 2.050$, $T_n = 1.845$ and $HSI_n = 2.166$ in the data point 10 from Table 4, the amount of calculated ROP in this point using Equation 44 is:

$$ROP = 0.0003 * 1.688^{0.01} * 2.050^{0.723} * 1.845^{2.714} * 2.166^{2.557} = 0.596 \text{ ft/min}$$

Real ROP in this point is 0.60, and the calculated ROP is 0.604 which is very close to the reality. Changing the used data series for model tuning changes the model accuracy. For model tuning in each point, it is possible to use some arbitrary previous data series. Increasing the number of used data series decreases the model accuracy, but the model would be more universal. For considering the effect of used data series number (DSN) in the model tuning, it is used of 5 DSN, 10 DSN, and 30 DSN for the MKM2 model. Figure 4 shows the predicted ROP vs. real ROP in these three situations.

It should be considered that choosing the DSN for model tuning is arbitrary, but selecting less DSN caused the model to be more local that means the model is valid just for a few next adjacent points, but the error of model fitting would be reduced. Figure 4(a) shows

the predicted ROP vs. real ROP using MKM2 and DSN of 30 with the normalized RMSE equal to 0.166. On the other hand, Figure 4(b) shows the predicted ROP vs. real ROP using MKM2 and DSN of 10 with the normalized RMSE equal to 0.128, and Figure 4(c) shows the same thing with DSN of 5 and normalized RMSE of 0.113. Noticing that the number of a-constants in the MKM2 is 5, it is unreasonable to reduce the DSN below 5.
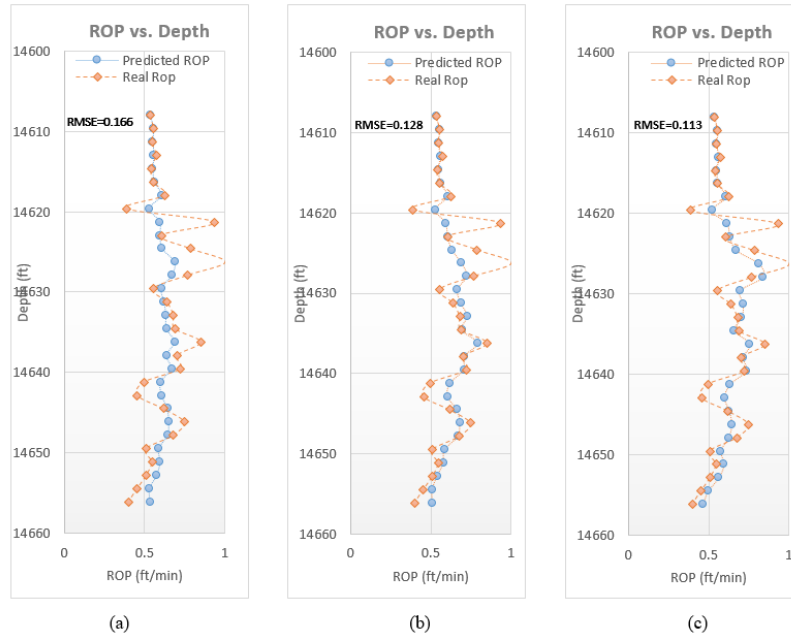

Figure 4. (a) the predicted ROP vs. real ROP using MKM2 and DSN of 30; (b) the predicted ROP vs. real ROP using MKM2 and DSN of 10; (c) the predicted ROP vs. real ROP using MKM2 and show the same thing with DSN of 5

Also, when drilling is progressing from a data point, it is possible to predict ROP in the next data point that is not still drilled when WOB, RPM, Torque, and HSI at this point are available. For this aim, MKM2 model with DSN of 10 would be used to predict ROP in the next point that is not steel drilled. It is clear that before doing this simulation, the amount of torque in the intended point should be calculated. Chen *et al.* (2014) proposed the following equation for predicting torque on the bit [24]:

$$T = \frac{\mu_b . D_b . WOB . e^{-\mu_s . \gamma_b}}{36} \tag{46}$$

This equation can be considered as

$$T = b_1 WOB . e^{-b_2} \tag{47}$$

Where b_1 and b_2 are some empirical constants relating to the bit-type, formation, bit life, and drilling parameters. b_1 and b_2 can be obtained with the proposed method in this study using ROA and DSN of 10. The obtained b_1 and b_2 coefficient and RMSE, Normalized RMSE, and cost function value for predicting torque is given in Table 5. Figure 5 compares the predicted torque in each point before drilling that point and the real torque at this point after drilling. Also, a comparison of predicted ROP using MKM2 and DSN of 10 before drilling and real value of ROP after drilling can be seen in Figure 6.

Table 5. Obtain a and b coefficients and RMSE, Normalized RMSE, and cost function value for predicting torque.

depth	b_1	b_2	RMSE	norm_RMSE	cost
14608	2.365217	0.816368	0	0	0
14609.66	2.502496	0.864569	0.015331	0.008216431	0.00047

| depth	b_1	b_2	RMSE	norm_RMSE	cost
14611.32	2.337684	0.792745	0.015893	0.008494748	0.000758
14612.98	2.478146	0.848533	0.016097	0.008595026	0.001037
14614.64	2.463726	0.842498	0.014417	0.007695517	0.001039
14616.3	2.335388	0.785857	0.018635	0.009915658	0.002084
14617.96	1.192181	0.107314	0.033415	0.017693575	0.007816
14619.62	2.29899	0.76597	0.032751	0.017362334	0.008581
14621.28	2.353505	0.785407	0.038427	0.020449731	0.01329
14622.94	1.329826	0.209645	0.035436	0.018867102	0.011301
14624.6	2.690463	0.904146	0.058187	0.030935555	0.030471
14626.26	1.264865	0.132063	0.102157	0.054040993	0.093924
14627.92	1.58874	0.349814	0.103718	0.054775107	0.096818
14629.58	1.432528	0.242374	0.09861	0.052192144	0.087516
14631.24	2.593527	0.832112	0.095153	0.050523223	0.081486
14632.9	2.661415	0.861217	0.098525	0.052592161	0.087365
14634.56	1.301895	0.14292	0.092612	0.049479755	0.077193
14636.22	1.316788	0.152277	0.092187	0.048941502	0.076486
14637.88	2.829824	0.920152	0.096595	0.051266879	0.083976
14639.54	2.596971	0.842475	0.094964	0.050483142	0.081164
14641.2	2.484546	0.819096	0.062748	0.033690273	0.035436
14642.86	1.235216	0.134108	0.053494	0.028927112	0.025754
14644.52	2.836251	0.965705	0.053192	0.028728833	0.025465
14646.18	2.355415	0.78022	0.052542	0.028343574	0.024846
14647.84	2.817027	0.957392	0.05337	0.028773635	0.025635
14649.5	2.72475	0.928979	0.05935	0.032103214	0.031702
14651.16	2.537801	0.867905	0.051398	0.028010165	0.023776
14652.82	2.47327	0.847737	0.057983	0.031748378	0.030259
14654.48	2.508115	0.871433	0.063799	0.03520582	0.036633
14656.14	2.472308	0.86069	0.068524	0.037924234	0.042259
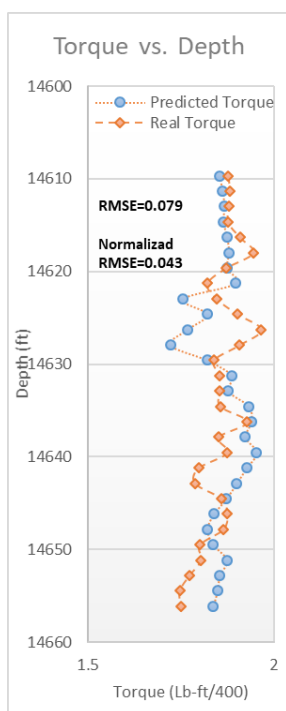
Figure 5. Comparison of the predicted torque in each point before drilling that point and the real torque at this point after drilling.
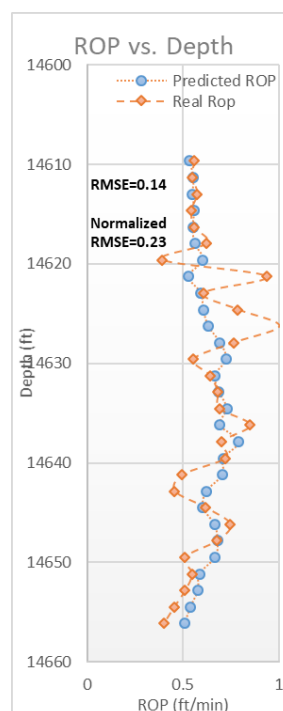

Figure 6. Comparison of predicted ROP using MKM2 and DSN of 10 before drilling and real value of ROP after drilling.

As it can be seen in Figure 5, the real-time continuous learning method with DSN of 10 could predict Torque value before drilling with normalized RMSE of about 4%, which means the accuracy of the method is about 96%. Also, the proposed method in this study could predict the ROP before drilling with Normalized RMSE of 23% when DSN was set to 10, as can be seen in Figure 6.

4. Conclusions

In this study, a new hybrid method for ROP modeling was presented. This model can predict the drilling rate using real-time drilling data. Also, continuous learning property of the proposed method causes the accuracy of the model to increase as the drilled section length increases. Moreover, a new empirical ROP method named MKM2 was presented. In this model, torque on the bit was used as a new input parameter, and a method for predicting torque at the next undrilled segment was proposed that the model could predict the torque value with 96% accuracy. Also, five famous models named BM, EM, BYM, MBYM, MKM1 were used for predicting ROP that could predict the drilling rate with about 75 percent degree of accuracy. The new proposed model (MKM2) could increase the degree of accuracy to 81%, which could yield 6% improvement in the ROP prediction accuracy. In addition, the MKM2 could predict the ROP before drilling with Normalized RMSE of 23% when DSN was set to 10.

Funding

Acknowledgments

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Chen, X.; Gao, D.; Guo, B. A Method for Optimizing Jet-Mill-Bit Hydraulics in Horizontal Drilling. *SPE-187225-PA* **2016**, *21*, 416-422, https://doi.org/10.2118/178436-PA.
2. Chen, X.; Gao, D.; Guo, B. Optimal design of jet mill bit for jet comminuting cuttings in horizontal gas drilling hard formations. *Journal of Natural Gas Science and Engineering* **2016**, *28*, 587-593,https://doi.org/10.1016/j.jngse.2015.12.033.
3. Li, B.; Guo, B.; Li, H.; Shi, Y. An analytical solution to simulate the effect of cement/formation stiffness on well integrity evaluation in carbon sequestration projects. *Journal of Natural Gas Science and Engineering* **2015**, *27*, 1092-1099,https://doi.org/10.1016/j.jngse.2015.09.058.
4. Bahari, A.; Khamehchi, E. Interfacial Phenomena Effect on Sand Production Due to Optimized Smart Water with/without the Presence of Nanoparticles. *Journal of Biointerface Research in Applied Chemistry*, **2020**, *Volume 10, Issue 6*, https://doi.org/ 10.33263/BRIAC106.66526668
5. Ezzati, S.; Khamehchi, E. Sandstone Reservoir Wettability Alteration Due to Water Softening: Impact of Silica Nanoparticles on Sand Production Mechanism. *Journal of Biointerface Research in Applied Chemistry*, **2020**, *Volume 10, Issue 5*, https://doi.org/10.33263/BRIAC105.63286342
6. Ejtemaee, P.; Khamehchi, E. Experimental investigation of rheological properties and formation damage of water-based drilling fluids in the presence of Al2O3, Fe3O4, and TiO2 nanoparticles. *Journal of Biointerface Research in Applied Chemistry*, **2020**, *Volume 10, Issue 4*, https://doi.org/10.33263/BRIAC104.886894.
7. Sui, D.; Nybo, R.; Azizi, V. Real-time optimization of rate of penetration during drilling operation. *IEEE Int. Conf. Control Autom. ICCA*, **2013**, pp. 357–362. https://doi.org/10.1109/ICCA.2013.6564893.
8. Ahmed, O.S.; Adeniran, A.A.; Samsuri, A. Computational intelligence based prediction of drilling rate of penetration: A comparative study, *Journal of Petroleum Science and Engineering***2018**, *172*, 1-12,https://doi.org/10.1016/j.petrol.2018.09.027.
9. Hegde, C.; Daigle, H.; Millwater, H.; Gray, K. Analysis of rate of penetration (ROP) prediction in drilling using physics-based and data-driven models. *Journal of Petroleum Science and Engineering* **2017**, *159*, 295-306, https://doi.org/10.1016/j.petrol.2017.09.020.
10. Barbosa, L.F.F.M.; Nascimento, A.; Mathias, M.H.; de Carvalho, J.A. Machine learning methods applied to drilling rate of penetration prediction and optimization - A review. *Journal of Petroleum Science and Engineering* **2019**, *183*,https://doi.org/10.1016/j.petrol.2019.106332.

11. Sinor, L.A.; Powers, J.R.; Warren, T.M. The Effect of PDC Cutter Density, Back Rake, Size, and Speed on Performance. In:*Proceedings of IADC/SPE Drilling Conference.* Dallas, Texas,1981, https://doi.org/10.2118/39306-MS.

12. Zhang, S.; Wang, H.; Qiu, Z.; Cao, W.; Huang, H.; Chen, Z. Calculation of safe drilling mud density window for shale formation by considering chemo-poro-mechanical coupling effect. *Petroleum Exploration and Development* **2019**, *46*, 1271-1280, https://doi.org/10.1016/S1876-3804(19)60280-6.

13. Eckel, J. R.; Microbit Studies of the Effect of Fluid Properties and Hydraulics on Drilling Rate. *Journal of Petroleum Technology,***1967***, 19, Issue 4*, pp. 541-546. https://doi.org/10.2118/1520-PA.

14. Moazzeni, A.; Khamehchi, E.; Drilling Rate Optimization by Automatic Lithology Prediction Using Hybrid Machine Learning. *Journal of Petroleum Science and Technology***2019,** *9*, 77-88.

15. Maurer, W.C. The "Perfect - Cleaning" Theory of Rotary Drilling. *SPE-13932-PA* **1962**, *14*, 1270-1274, https://doi.org/10.2118/408-PA.

16. Bourgoyne, A. T.; Millheim, K.K.; Chenevert, M.E.; Young, F.S. Applied Drilling Engineering. *Society of Petroleum Engineers***1986**.

17. Dupriest, F.E.; Koederitz, W.L. Maximizing Drill Rates with Real-Time Surveillance of Mechanical Specific Energy. SPE/IADC Drilling Conference. Amsterdam, The Netherlands, February 23-25. 2005, https://doi.org/10.2118/92194-MS.

18. Gandelman, R.A. Prediction of ROP and Real Time Optimization of Operational Parameters while Drilling Offshore Oil-Wells. M.S. Thesis. Universidade Federal do Riode Janeiro, 2012.

19. Gray, K. Real-time predictive capabilities of analytical and machine learning rate of penetration (ROP) models, *Journal of Petroleum Science and Engineering* **2018**, *172*, 934-959,https://doi.org/10.1016/j.petrol.2018.08.083.

20. Bingham, M.G. How Rock Properties Are Related to Drilling. *The Oil and Gas Journal***1964**, *62,* 94-101.

21. Nayak, R., Nanda, B., A review on formulation, design of nanostructured material through oil-in-water microemulsion, *Journal of Letters in Applied NanoBioScience*, **2020**, Volume 9, 2; https://doi.org/10.33263/LIANBS92.945951;

22. Teale, R. The concept of specific energy in rock drilling. *International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts* **1965**, *2*, 57-73,https://doi.org/10.1016/0148-9062(65)90022-7.

23. Cherif, H.; Bits. S. FEA modelled MSE/UCS values optimise PDC design for entire hole section. In: A*t the North Africa Technical Conference and Exhibition held in Cairo, Egypt*2012;pp. 20-22.

24. Chen, X.; Fan, H.; Guo, B.; Gao, D.; Wei, H.; Ye, Z. Real-Time Prediction and Optimization of Drilling Performance Based on a New Mechanical Specific Energy Model. *Arabian Journal for Science and Engineering* **2014**, *39*, 8221-8231, http://dx.doi.org/10.1007/s13369-014-1376-0.

25. Pessier, R.C.; Fear, M.J. Quantitative common drilling problems with mechanical specific energy and bit-specific coefficient of sliding friction. *Paper SPE 24584 presented at the 67th Annual Technical Conference and Exhibition*, *Washington, DC* **1992**, https://doi.org/10.2118/24584-MS.

26. Moazzeni, A. Real-Time Optimization of Drilling Parameters During Drilling Operation, PhD thesis, Amirkabir University of Technology, Department of petroleum engineering, Iran, 2020.