Mathematical Modeling to Predict the Rate of Penetration (ROP) Using Genetic Programming

- **Seyed Ali Seyedalangi[1], Mohammad Javad Nabavizadeh[2], Mastaneh Hajipour[2*]**

Department of Petroleum Engineering, Faculty of Petroleum and Chemical Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran.

Corresponding author Email address: m.hajipour@srbiau.ac.ir

Received: Feb 06, 2019 / Accepted: Mar 28, 2019

Abstract

Rate of penetration (ROP) model is a mathematical relation between bit penetration rate and properties of formation, drilling fluid and drilling operation conditions. Due to relatively high cost of drilling operations, it is essential to develop an accurate prediction of the ROP to estimate the drilling time and costs. In this paper, a new model has been developed for estimation of ROP in one of Iranian oil fields by implementation genetic programming. In the developed model, ROP has been correlated with 11 effective parameters reported in drilling master log and sonic log including weight on bit, bit rotational speed, total nozzle area size, mud weight, mud yield point, fluid loss and sonic time. For the evaluation of the proposed model, statistical parameters including root-mean-square deviation (RMSD), squared correlation coefficient ($R^{2)}$ and average absolute relative deviation (AARD) were calculated. Real data verification indicated that the developed model is accurate for estimating ROP and can provide useful information when drilling operation is running. The values of squared correlation coefficient and root-mean-square deviation show the reliability of the model.

Keywords: Rate of penetration, Genetic programming, Master log, Sonic log, Drilling operation.

1. Introduction

Among different energy sources, oil and gas play important roles. In petroleum industry, drilling operations always have high risks and costs. The cost of drilling operation depends on type of rig used, geographic location and target drilling depth. Cost of drilling is also influenced by drill bit rotating time, trip time, connection time, cost of drill bits and rig cost. The drill bit rotating time depends on several drilling variables e.g., wellbore stability, class of drill bit, weight on bit, rotary speed of drill bit, drilling mud properties, hydraulics of drilling mud, drill bit tooth wear and the cost of drilling increases with depth in a parabolic manner up to about 3,000 meters, and then exponentially increases beyond 4,000 meters (Masseron, 1990). Therefore, marginal improvement in drilling cost may reduce exploration and development costs to a significant extent (Guria et. al. 2014). In drilling operation, a large saving in time and money would be achieved by reducing the drilling time, since some of the costs are time-dependent. Drilling time could be minimized by raising the penetration rate (Bahari et. al. 2009).

The optimization of the rate of penetration (ROP) as a key parameter to understand and control the drilling process, is essential to reduce the drilling costs. Actually, having knowledge about the effects of different parameters on the ROP helps to adjust the variables and reach the maximum efficiency and minimum cost (Arabjamaloei and Shadizadeh, 2011). Different optimization studies have been conducted in this regard (Cheraghi Seifabad et. al. 2013; Zare et. al. 2014; Rahimzadeh et. al. 2011; Guria et. al. 2014; Kexiong et. al. 2007)

To predict the rate of penetration and the best operating drilling parameters, different models have been developed in terms of rock properties and drilling variables which have been discussed in the literature (Mitchell, 1992; Bourgoyne et al, 2003). It is very difficult to develop a mathematical model describing the details of drilling processes at the bottom of the well, completely. The difficulty mainly arises for two reasons. First, there are many parameters with complex relationships with the rate of penetration (Monazemi et al., 2012). These parameters consist of variables associated with drilling fluid, formation type, bit rotation speed and weight on bit. The identification of all these parameters is very difficult while some of them are controllable and some are not. Second problem in the mathematical modelling is the difficulty in involving all the mechanisms of drilling operation (grinding formation process by bit, transportation of cuttings to the surface, drilling fluid power effect on grinding and other physical effects) into mathematical equations. According to these complexities, different researchers mostly have used experimental data to produce empirical models for the prediction of the ROP.

So far, a number of ROP models have been presented in the literature. Gall and Woods (1963) investigated the effects of different drilling parameters on ROP using practical methods. In their experimental study, the following correlation between ROP and other parameters has been obtained.

$$\mathrm{ROP} = C_f \frac{N\,W^K}{a^b} \tag{1}$$

N, W, K, a and b are bit rotation speed, weight on bit (WOB), formation hardness, bit abrasion and bit dullness. C_f is a constant containing the effects of bit type, hydraulic, drilling fluid and formation.

Mechem and Fullerton (1965) presented another correlation based on six variables including hydraulic, mud pressure, well depth, bit rotation, WOB and formation drillability.

One of the most famous models was proposed by Bourgoyne and Young (1974). In this model, natural logarithm of ROP has been correlated with a multi-variable linear regression. In their correlation, eight variables including depth, formation compression, pressure difference at the bottom of well, bit diameter, WOB, bit rotation speed, bit wear and hydraulic have been considered. The general form of their model is as follows.

$$\mathrm{Ln\,ROP} = a_1 = a_2X_2 + a_3X_3 + a_4X_4 + a_5X_5 + a_6X_6 + a_7X_7 + a_8X_8 \tag{2}$$

Ziaja and Miska (1982) presented a mathematical model in which the effects of torque and WOB have been included. This mathematical relation has practical application for polycrystalline diamond compact (PDC) bit drilling processes.

Reza and Alcocer (1986) have developed a model for deep drilling operations. In their study, seven parameters including WOB (W), bit rotation speed (N), bit bearing diameter (d_b), drilling mud viscosity (μ), rock hardness (H), flow rate of drilling fluid (Q) and differential pressure (P_e) have been considered.

$$ROP = 0.33\, N\, d_b \left(\frac{N d_b^{\,2}}{\mathrm{i}}\right)^{0.43} \left(\frac{N d_b^{\,3}}{Q}\right)^{-0.68} \left(\frac{H d_b}{W}\right)^{-0.91} \left(\frac{P_e d_b}{W}\right)^{-0.15} \tag{3}$$

There are a number of statistical and mathematical approaches in order to develop empirical correlations. Artificial neural network (ANN), genetic programing (GP), particle swarm optimization (PSO), adaptive network-based fuzzy inference system (ANFIS), etc. are famous methods with diverse range of applications, especially for optimization and modelling purposes (Abooali and Khamehchi, 2016). Arabjamaloei and Shadizadeh (2011) have studied the effects of various parameters on ROP using artificial neural network (ANN) procedure. Monazemi et al. (2012) have used a three layer feed-forward neural network to estimate ROP.

The main target of this study was to propose a new model for rate of penetration with acceptable accuracy. In this regard, genetic programming (GP), as one of the most applicable methodologies, has been applied to obtain a mathematical model for the prediction of ROP. In this study, field data were collected from one of the Iranian southern oil fields and the considered variables were well depth, weight on bit, bit rotational speed, rate of drilling fluid, hole size, total nozzle area size, mud weight, plastic viscosity, mud yield point, fluid loss and sonic time. Using some new variables along with application of genetic programming for developing the ROP model is novel and innovative in this work.

2. Methodology

A dataset including 230 sets of experimental data was used in this study. Each set contains 11 parameters including depth, WOB, bit rotational speed, rate of drilling fluid, hole size, total nozzle area size, mud density (ρ_m), plastic viscosity (μ_p), mud yield point (YP), sonic time and fluid loss. All datasets were collected from three wells in one of the Iranian southern fields. At first, the dataset was divided into two subsets, randomly: training set (including 80% of data) and test set (including remaining data). Training data were used to construct the model and test set was applied for evaluating the estimation ability and accuracy of the developed model. In addition to 230 data using for developing the model, 76 data which is called "test #2 dataset" were collected from another well in the studied field and were used to verify the applicability of the new model.

2.1. Mathematical Optimization

Drilling cost optimization through mathematical techniques is based on the proposed models of the penetration rate, bit hydraulics, bit wear, etc., to predict and eventually optimize the rate of penetration and drilling cost (Kaiser M.J., 2007).

Genetic programming (GP) which has been developed in the early 1990s (Koza, 1992) is a powerful mathematical tool especially for optimization and modelling projects. In genetic approach, the algorithm randomly generates a population of computer programs in the form of tree structures (gene) and then, mutates and crosses over the best performing trees to generate a new population. This process is iterated until the last population containing the best programs solve the task well (Abooali and Khamehchi, 2014; Morrison et al, 2010).

After generating the first population (parents), the overall form of primary model is determined by weighted summation of all the genes with a bias term. A simple schematic of the gene (tree structure) has been shown in Figure 1.

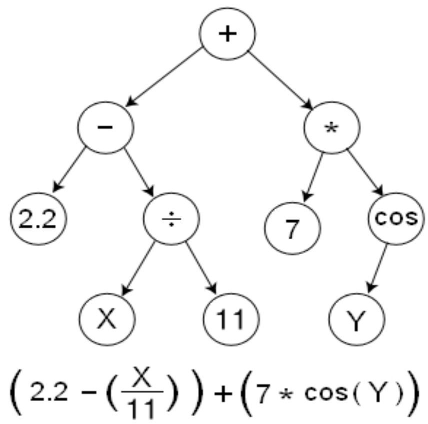

Fig 1. A simple gene (tree structure).

When the GP algorithm is applied for modelling purposes and specification of mathematical functions i.e., symbolic regression, the algorithm constructs the model form and then fits the model constants. If the algorithm creates several genes instead of one gene, it will be called "multi-gene symbolic regression". It is a more applicable technique developed in order to produce a population of mathematical relations. A multi-gene method consists of one or more genes that each one is individually a usual GP tree (Searson et al., 2010).

A free open source genetic programming toolbox prepared by Searson et. al. (2010) was used in the present study. It has been written for multi-gene symbolic applications. So, all the steps of genetic method are operated in this program. Before using the program, the basic parameters including the number of population, number of generation, maximum number of gene, maximum number of nodes in the genes, etc. should be determined.

3. Evaluation of the Model

For the evaluation of the proposed ROP model and optimization studies, usual statistical parameters including root-mean-square deviation (RMSD), squared correlation coefficient (R^2) and average absolute relative deviation percentage (AARD%) have been calculated. These parameters are defined as:

$$R^2=1-\frac{\sum_{i=1}^{n}(ROP_i^{exp}-ROP_i^{cal})^2}{\sum_{i=1}^{n}(ROP_i^{exp}-\overline{ROP^{exp}})^2} \quad (4)$$

$$RMSD=\sqrt{\left(\frac{1}{n}\right)\sum_{i=1}^{n}(ROP_i^{exp}-ROP_i^{cal})^2} \quad (5)$$

$$ARD(\%)=\left|\frac{ROP_i^{exp}-ROP_i^{cal}}{ROP_i^{exp}}\right|\times 100 \quad (6)$$

$$AARD(\%)=\left(\frac{1}{n}\right)\sum_{i=1}^{n}\left|\frac{(ROP_i^{exp}-ROP_i^{cal})}{ROP_i^{exp}}\right|\times 100 \quad (7)$$

ROP^{exp}, ROP^{cal}, $\overline{ROP}^{exp}$ and n stand for the real rate of penetration recorded in the master log, the value of the rate of penetration calculated by the model, average values of real rate of penetration and the number of data, respectively. The model efficiency is higher as the value of R^2 is closer to unity. Lower values of RMSD and AARD (i.e. closer values to 0) indicate higher accuracy of the developed model.

4. Results and Discussion

In this study, a mathematical model for the prediction of ROP has been obtained using the genetic programming. In the developed model, ROP has been correlated with a number of effective parameters reported in master log and sonic log i.e., well depth, weight on bit (WOB), bit rotational speed, rate of drilling fluid, hole size, total nozzle area size, mud weight, plastic viscosity, mud yield point, fluid loss and sonic time.

By application of genetic programming toolbox, an acceptable correlation for predicting the rate of penetration was obtained. The developed model is as follow:

$$ROP=6.978 * 10^{-4} * (X_1 - X_4+X_{11} - X_5) + X_{10} * (0.00676 * X_4 + 1.406 * X_6 - 1.391) - 3804 * X_6 + 0.04751 * X_8 - 0.02176 * \exp(6.456/X_6) + 1414 * \exp(X_6) - 0.04752 * \exp(-8.822 / X_4) * (X_9 - 3.059) + 0.04821 * X_6 * (X_{11} - 4.165) - (0.008708 * X_6 / (X_{11} - 3.904)) - 0.03541 * (X_{11} - X_4 * \exp(X_6) + X_4 * X_6 * (X_6 - 0.2606)) /(X_{10} - (X_8 / X_4)) + 13.67 * 10\text{-}^{4} / X_{10} + 0.6228 * X_7^{2} * Ln(Ln(X^7 * X^3)) / X_2 + 0.4003 * X_2^{2} * Ln(Ln(X_3 * X_2)) / X_7^{3} - 2.011 * 10^{-4} * X_7^{5} * X_3^{2} * Ln(X_7 * (X_7+X_2)) / X_2^{4} - 35.519 \quad (8)$$

in which the variables X_1 to X_{11} have been presented in Table 1.

Table 1 : The parameters of the developed model.

Variable	Quantity (unit)	Symbol
X_1	Depth (ft)	h
X_2	Sonic time (μs/ft)	-
X_3	Bit rotational speed (1/min)	RPM
X_4	Weight on bit (klb_f)	WOB
X_5	Mud injection rate (gal/min)	GPM
X_6	Total nozzle area size (in^2)	$A_{tot\text{-}nozzle}$
X_7	Hole size (in)	d
X_8	Mud density (pcf)	ρ_m
X_9	Plastic viscosity (cP)	μ_p
X_{10}	Yield point (lbf/100ft^2)	YP
X_{11}	Fluid loss	-

The statistical parameters of Eq. 8 have been shown in Table 2. As can be seen, the values of R^2, RMSD and AARD are appropriate and acceptable, therefore, the new developed model can be used for estimation of ROP with good accuracy.

Table 2 : Statistical parameters of the developed ROP model.

Parameter	Total data	Train data	Test data
n	230	185	45
RMSD	0.2901	0.2795	0.3299
R^2	0.9812	0.9824	0.9764
AARD	4.9682	4.4498	7.0994

For investigating the reliability and applicability of the developed model, sensitivity analysis on some of model parameters were carried out and the effects of several important drilling parameters on the variations of ROP were analyzed. The predicted results of the model for different values of RPM and hole size are shown in Figures 2a and b respectively. As expected, increasing the values of RPM enhances the penetration rate while increasing the hole size reduces the ROP values which have been predicted correctly using the new developed model.

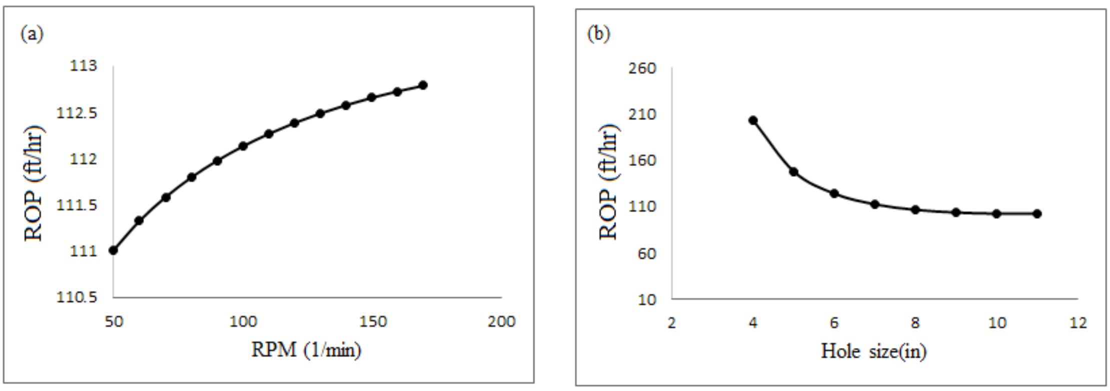

Fig 2. Variations of the predicted ROP with (a) RPM and (b) hole size.

Figures 3a and b show the effects of drilling fluid viscosity and sonic time on the values of ROP. It can be seen that the developed model predicts correctly the reduction of ROP as increasing the mud viscosity and sonic time. The estimated values of ROP have been compared with experimental data, in Figure 4. It is seen that the results of the model show an acceptable agreement with the experimental values.

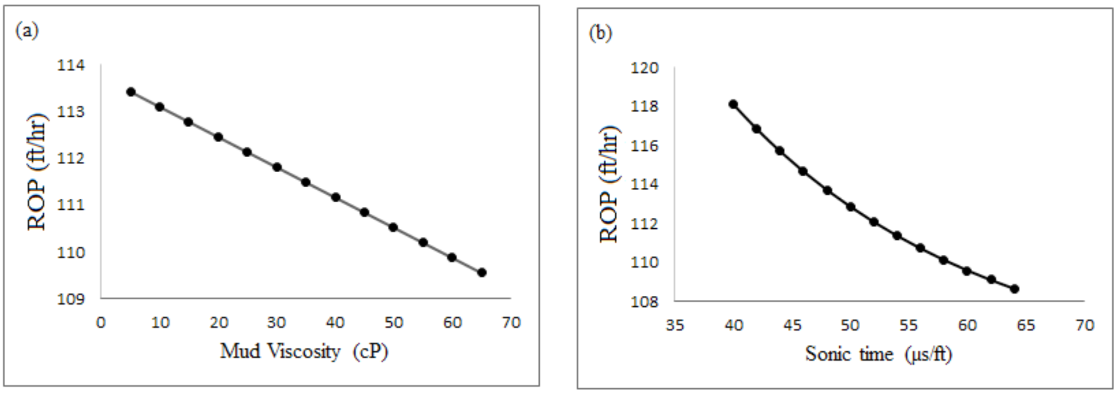

Fig 3. Variations of the predicted ROP with (a) mud viscosity and (b) sonic time.

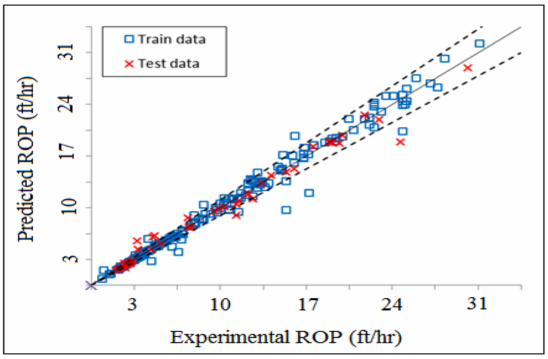

Fig 4. Predicted rate of penetration versus real data.

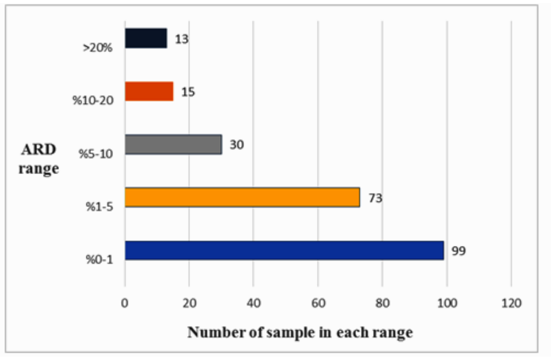

Fig 6. The absolute relative deviation (ARD) of the model over 230 data.

The absolute errors between estimated and experimental ROP values for all dataset have been calculated and shown in Figure 5. According to this figure, there are only 5 cases with absolute error of more than 1 among all 230 data.

The absolute relative deviation of all dataset have been also presented in Figure 6 versus the number of data samples in different range of relative error values. This figure shows that in the total dataset (230 data), there are 28 data with error exceeding 10%. In other words, 87.826% of all dataset samples have absolute relative deviation less than 10%. The average of errors lower than 10% is 2.146%.

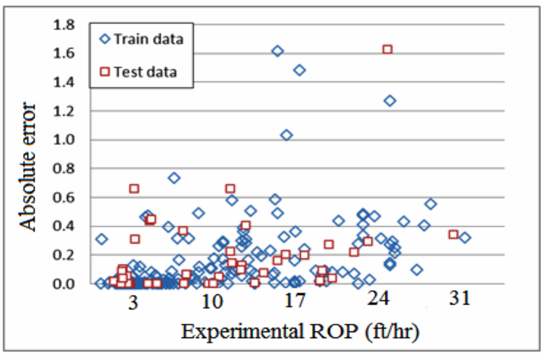

Fig 5. Absolute errors of all dataset versus experimental ROP.

To evaluate the model validity, a dataset including 76 experimental data collected from another well in the studied field was used as test #2 dataset. Figure 7 shows the estimated versus experimental ROP along with the statistical parameters of the new developed model over test #2 dataset. The absolute errors between the experimental and the predicted values for test #2 dataset have been presented in Figure 8. According to Figures 7 and 8, the prediction ability of the developed model is acceptable and the estimation accuracy is appropriate.

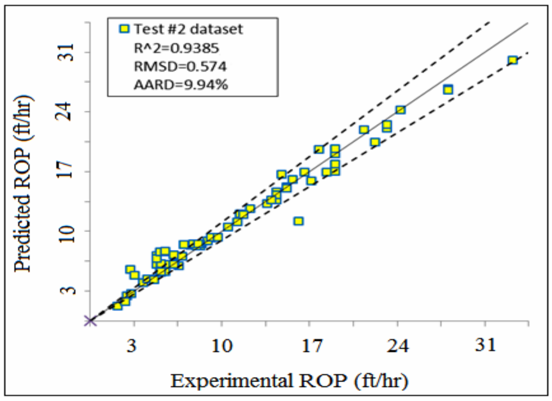

Fig 7. Validation of the predicted results of the model using test #2 dataset.

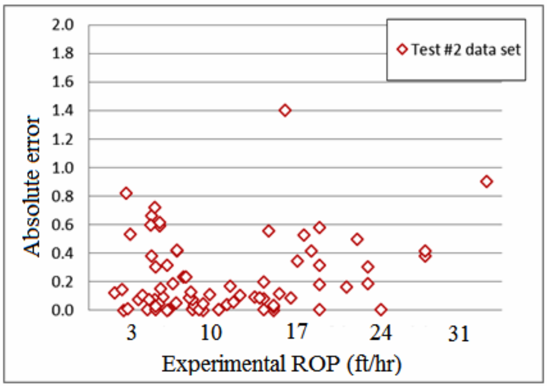

Fig 8. Absolute errors for test #2 dataset.

5. Conclusions

In this study, a new model has been developed for prediction of the rate of penetration (ROP) by application of genetic programming approach. Parameters of the new model are related to drilling operational condition, drilling fluid properties and sonic log which all of them are obtained from drilling master log and mud recap. Real data verification indicated that the developed model is accurate for estimating ROP and can provide useful information when drilling operation is running. As the ROP is a vital parameter affecting the drilling cost, it is important to find optimized condition of ROP and the new model can be applied in this area. This type of model can be directly used in drilling operations and also for drilling simulation processes.

References

1. Abooali, D., Khamehchi, E. 2014. Estimation of dynamic viscosity of natural gas based on genetic programming methodology. Journal of Natural Gas Science and Engineering 21: 1025-1031.

2. Abooali, D., Khamehchi, E. 2016. Toward predictive models for estimation of bubble-point pressure and formation volume factor of crude oil using an intelligent approach. Brazilian Journal of Chemical Engineering 33: 1083–1090.

3. Arabjamaloei, R., Shadizadeh, S. 2011. Modeling and Optimizing Rate of Penetration Using Intelligent Systems in an Iranian Southern Oil Field (Ahwaz OilField). Petroleum Science and Technology 29: 1637–1648.

4. Bahari A., Baradaran Seyed A. 2009. Drilling cost optimization in a hydrocarbon field by combination of comparative and mathematical method. Petroleum Science 6(4): 451–463.

5. Bourgoyne, A.T., Young, F.S. 1974. A Multiple Regression Approach to Optimal Drilling and Abnormal Pressure Detection. Society of Petroleum Engineers Journal 14: 371-384.

6. Bourgoyne A.T., Millheim K.K, Chenevert M.E., Young Jr F.S. 1991. Applied Drilling Engineering. Richardson, TX: SPE Textbook Series.

7. Cheraghi Seifabad M., Ehteshami P. 2013. Estimating the drilling rate in Ahvaz oil field. J Petrol Explor Prod Technol 3:169–173.

8. Galle, E. M., Woods, H.B. 1963. Best Constant Weight and Rotary Speed for Rotary Rock Bits. New York: American Petroleum Institute.

9. Guria C., Goli K.K., Pathak A.K. 2014. Multi-objective optimization of oil well drilling using elitist non-dominated sorting genetic algorithm. Petroleum Science 11(1): 97–110.

10. Kaiser M.J. 2007. A Survey of drilling cost and complexity estimation models. Int. J. Petroleum Science and Technology 1: 1-22.

11. Kexiong W., Laibin Z., Hongwei J. 2007. Relationship between formation water rate, equivalent penetration rate and volume flow rate of air in air drilling. Petroleum Science 4(4): 62-65.

Kia, S.M. 2012. Genetic Algorithms in MATLAB. Tehran: Kian publication.

Koza, J.R. (1992) Genetic Programming: on the Programming of Computers by Means of Natural Selection. MIT Press, Cambridge.

Masseron J. 1990. Petroleum Economics (4th ed.). Paris: Editions Technip.

12. Mechem, O.E., Fullerton, H.B.Jr. 1965. Computers invade the rig floor. Oil and Gas Journal.

13. Miska, S., Ziaja, M.B. 1982. Mathematical Model of the Diamond Bit Drilling Process and Its Practical Application. Annual Fall Meeting of the Society of Petroleum Engineers of AIME.

14. Mitchell B.J. 1992. Advanced Oil Well Drilling Engineering Handbook and Computer Programs (9th ed.). Dallas, TX: SPE Textbook Series.

15. 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. Electronic Scientific Journal, Oil and Gas Business 6: 21-31.

16. Moradi, H., Bahari, M.H., Naghibi Sistani, M.B., Bahari, A. 2010. Drilling rate prediction using an innovative soft computing approach, Scientific Research and Essays 5(13): 1583-1588.

17. Morrison, G.A., Searson, D.P. Willis, M.J., 2010. Using genetic programming to evolve a team of data classifiers. International Journal of Computer and Information Engineering 4(12): 1815-1818.

18. Rahimzadeh, H., Mostofi, M., and Hashemi, A. 2011. A New Method for Determining Bourgoyne and Young Penetration Rate Model Constants. Petroleum Science and Technology 29: 886–897.

19. Reza, M.R., Alcocer, C.F. A Unique Computer Simulation Model Well Drilling: Part I - The Reza Drilling Model., Paper SPE 15108 presented at the SPE 56th California Regional Meeting of SPE, 2-4 April 1986, Oakland, CA, USA.

20. Searson D.P., Leahy D.E., Willis M.J. 2011. Predicting the Toxicity of Chemical Compounds Using GPTIPS: A Free Genetic Programming Toolbox for MATLAB. In: Ao SI., Castillo O., Huang X., editors. Intelligent Control and Computer Engineering. Lecture Notes in Electrical Engineering, vol. 70. Dordrecht: Springer.

21. Searson, D.P., Leahy, D.E. Willis, M.J., 2010. GPTIPS: an open source genetic programming toolbox for multigene symbolic regression. IMECS, Hong Kong.

22. Zare, J., and Shadizadeh S.R. 2014. Managed Pressure Drilling to Increase Rate of Penetration and Reduce Formation Damage. Petroleum Science and Technology 32: 1833–1842.

مدلسازی ریاضی جهت پیش بینی نرخ نفوذ مته با روش برنامه‌ریزی ژنتیک

• سیدعلی سیدالنگی[۱]، محمدجواد نبوی زاده[۲]، مستانه حاجی پور[۲*]

۱. کارشناس ارشد مهندسی حفاری، گروه مهندسی نفت، دانشکده نفت و مهندسی شیمی، واحد علوم و تحقیقات، دانشگاه آزاد اسلامی، تهران، ایران.

۲. استادیار گروه مهندسی نفت، دانشکده نفت و مهندسی شیمی، واحد علوم و تحقیقات، دانشگاه آزاد اسلامی، تهران، ایران.

(ایمیل نویسنده مسئول: m.hajipour@srbiau.ac.ir)

چکیـــــده

مدل نرخ نفوذ مته، یک رابطه ریاضی بین سرعت نفوذ مته و ویژگی های سازند، سیال حفاری و شرایط عملیات حفاری است. به دلیل هزینه بالای عملیات حفاری، پیش بینی دقیق نرخ نفوذ مته جهت تخمین زمان و هزینه های حفاری ضروری است. در این مقاله، یک مدل جدید جهت پیش بینی نرخ نفوذ مته در یکی از میادین نفتی ایران با روش برنامه ریزی ژنتیک ارائه شده است. در مدل ارائه شده، نرخ نفوذ مته تابعی از ۱۱ پارامتر موثر گزارش شده در مستر لاگ حفاری و لاگ صوتی شامل وزن روی مته، سرعت چرخش مته، مساحت کل نازل ها، وزن گل، نقطه واروی گل، هرزروی سیال و زمان عبور صوت بدست آمد. برای ارزیابی مدل پیشنهادی، پارامترهای آماری شامل جذر میانگین مربعات خطا (RMSD)، مجذور ضریب همبستگی (R^2) و میانگین مطلق خطای نسبی (AARD) محاسبه شدند. اعتبار سنجی مدل با استفاده از داده های واقعی میدان نشان داد که مدل ارائه شده جهت پیش بینی نرخ نفوذ مته دقیق بوده و می تواند اطلاعات مفیدی حین عملیات حفاری در دسترس قرار دهد. مقادیر بدست آمده برای مجذور ضریب همبستگی و جذر میانگین مربعات خطا نشان دهنده قابل اطمینان بودن مدل هستند.

واژگان کلیدی: نرخ نفوذ مته، برنامه ریزی ژنتیک، مستر لاگ، لاگ صوتی، عملیات حفاری