Optimization Of Rate Of Penetration With Real Time Measurements Using Machine Learning And Meta-Heuristic Algorithm

Sridharan Chandrasekaran, G. Suresh Kumar

Abstract: The energy industry has continuously strived to develop technology that maximizes drilling performance to reduce the cost per barrel of crude produced & at the same time, minimize the HSE (health, safety, and environment) risk. Rate of penetration (ROP) optimization is one of the primary factors to improve drilling efficiency and to minimize the operational cost of rigs, drilling operation and drilling tools. Traditional ROP models are empirical based which may be inconsistent in field environments and hence the predictive accuracy of such models are low and subjective. With immense drilling data, operational data, geological data collected over years, the drilling engineering started to shift from first principles modeling to data driven modelling which offers an easier way of extracting value in the data by intelligent algorithms. In this study, Artificial Neural Network (ANN) is developed to predict ROP by making use of the offset vertical wells' real-time surface parameters while drilling. In the ANN, the input-output mapping is designed with interconnected feed-forward back propagation neural network so that the ROP is efficiently predicted at the drilling bit. Data screening methods and feature engineering methods transform the raw data into a processed data so that the model learns effectively. The developed model is cross validated to generalize over a range of inputs and compared with field measurements. With the help of the developed ANN model, a meta-heuristic algorithm is incorporated to optimize ROP thereby reducing the overall cost per foot of the well. This is achieved by designing Particle Swarm Optimization (PSO) algorithm and allowing the PSO to find the best combination of drilling parameters namely weight on bit (WOB), revolutions per minute (RPM) of the drill bit, and flow in the pumps to maximize the ROP under field constraints. This study combines ANN with PSO to optimize ROP based on real time measurements which has immense potential for operating oil and gas companies to aid in well design or to add as an artificial intelligence component in drilling simulator or autonomous driller.

Index Terms :meta-heuristic, neural network,optimization, particle swarm.

1 INTRODUCTION

The drilling industry in the current times experience increasing complexity of the wells which escalate HSE (health, safety and environment) risk and the drilling cost. The focus of the drilling industry shifted towards finding unique ways to maximize the potential profits for drilling wells by improving the efficiency of the process. The inefficiency in drilling due to under optimized drilling process and poor rate of penetration(ROP) is estimated to cost 15% of the total well cost [1]. ROP is the related to the speed with which drilling is performed and maximizing it, is one of the critical factors affecting the commercial success of the drilling operation. Rate of penetration prediction with mathematical models and optimization of the drilling variables based on these models has been active area of research. ROP depends on many drilling variables and has complex relationships, like the operational parameters, formation properties, compressive strength, well hydraulics, borehole shape & size, mud properties, type of bit, hole-cleaning etc. For rotary drilling process, the ROP model proposed [2] is considered as a comprehensive model for the drilling optimization as it was based on physical behaviour and validated with field results. Warren [3] developed an imperfect cleaning model to predict the ROP in soft formation based on experimental data, dimensional analysis and generalized response curves which was later modified to include chip hold down effects [4].

- *Sridharan Chandrasekaran is currently pursuing PhD degree program in Department of Ocean Engineering, IIT-Madras, Chennai-600036, INDIA. Email: sridharavc@gmail.com. Phone: +91-9840261010*
- *Dr. G Suresh Kumar is a Professor in Petroleum Engineering Program, Department of Ocean Engineering, IIT-Madras, Chennai-600036, INDIA. Email: gskumar@iitm.ac.in. Phone: +91 - 44 - 2257 4814*

Such traditional empirical models contain a number of constants which needs to be evaluated based on field results or experimental methods and most optimization techniques would choose the constants that maximize ROP. This process is time consuming [5] and may be too specific for a formation. The abundance of real time drilling data together with increased big data infrastructure & computing resulted in the development of data driven models for predicting ROP and optimizing it. Neural network technique involving two parameters namely WOB and RPM [6] was used to optimize ROP. The data models suffer the drawback of understanding the physics behind the problem which was overcome by enabling coupling conditions between neural network and regression analysis [7]. Weak machine learning ROP predictors are coupled together (ensemble) so that an integrated model produces accurate predictions than individual predictors. The data modelling by such ensemble methods of machine learning [8] can infer parameters affecting ROP optimization rather than predicting them. There have been previous optimization efforts in the oil and gas industry. Heuristic approach by particle swarm optimization (PSO) was attempted to find the optimal well path [9] and well location [10] with the goal to minimize the total measured depth of the well. Jiang and Samuel [11] extended the ant colony optimization approach to estimate the best drilling parameters based on real time measurements. Self [12] presented work on optimization of drilling parameters with downhole parameters and for different combination of drill bits. In the current work, an artificial neural network model was developed to predict the ROP by performing data analytics from the offset well real time drilling data. In order to reduce the amount of uncertainties in the drilling scenario, only vertical wells drilled with similar BHAs from the same field is considered for the development of the model. In this work, a new approach to drilling optimization on real time surface parameters was introduced by incorporating PSO (particle swarm optimization) on the ANN model in order to optimize one section of the well. The overall objective was to use every available information in real-time to perform ROP optimization overcoming static pre-drill configuration and allowing the use of data driven analytics in the drilling industry.

2 ROP - ANN Modeling

This section describes the implementation of artificial neural network to model ROP from the surface drilling parameters. The ANN concept of data processing was inspired from biological neural system in human brains. A typical neuron consisting of dendrites that carry external impulses to the cell body. The cell body processes the information and provides an output to the axon which is interconnected to other neurons through synapses. Figure 1 show the mechanism of artificial neuron where input channel is analogous to the dendrites, cell body to the transfer/activation function and axons to the activation channel. A number of such artificial neurons form the neural network which learns by collecting data samples and understands the relationship between the output and input. The ANN continuously compares the output and the difference is fed back to the neuron to adjust the weights. The iterative process stops when the simulated output and the desired output agree within the threshold agreed by the user and at this point, the network has learned the relationship. The learning is stored in the form of model containing weights for certain type of activation function and biases. In this study, the output of the ANN is the instantaneous rate of penetration and the inputs to the network are WOB, RPM, Torque, DOC, Flow, Formation type, hole section, SPP. The problem is mathematically formulated as a classic supervised learning regression problem. One of the challenges while designing the neural network is to make a proper choice of the network architecture. In this work, a back propagation three layered neural network was constructed using scikit learn libraries of Python language[14] . The methodology of model creation is represented in figure 2.

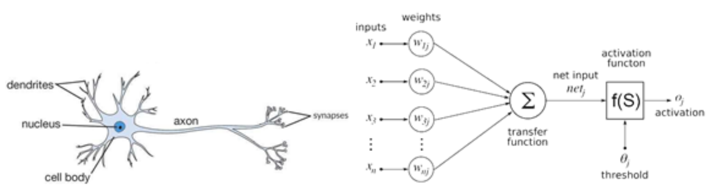

Figure 1: Mechanism of artificial neuron depicting biological neuron

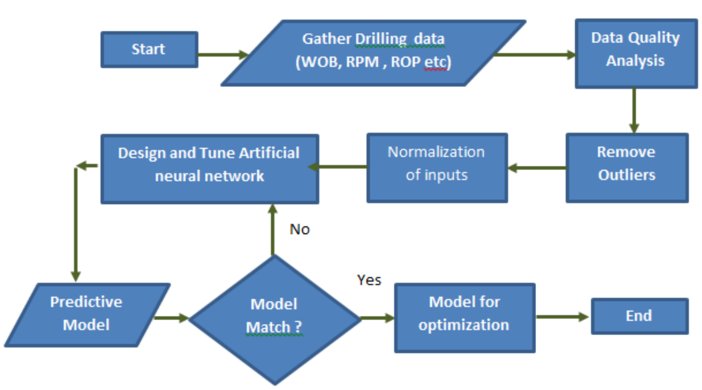

Figure 2: ROP Model development – Methodology

3 Particle Swarm Optimization

Particle swarm optimization (PSO) is a population based stochastic optimization technique inspired by social behavior of species in the nature like bird flocking or fish schooling. It is a meta-heuristic algorithm in which the group is referred as a swarm and each species in the group is referred as particle. The potential solutions (particles) swim through the problem space by following the current optimum particles when initialized with random positions. All particles have fitness values which are evaluated by the fitness function to be optimized, and have velocities which direct the flying of the particles. Each particle is governed by two equations namely the velocity equation and the position equation. The velocity equation for the particle is given by equation 1 and has three main components: the inertia component, the cognitive component, the social component given by, Inertial component : This tends to force the particle to move in the current trajectory Particle memory component : This tends to force the particle to move in the trajectory of its personal best Swarm (Social component): This tends to force the particle to move in the trajectory of the neighbourhood best performance. $v_{k+1}^i = w * v_k^i + c_1 * r_1 * (p_k^i - x_k^i) + + c_2 * r_2 * (p_k^g - x_k^i)$ (1) The updated particle position for (k+1) iteration is given in equation (2) and represented in figure 3, $x_{k+1}^i = x_k^i + v_{k+1}^i$ (2)

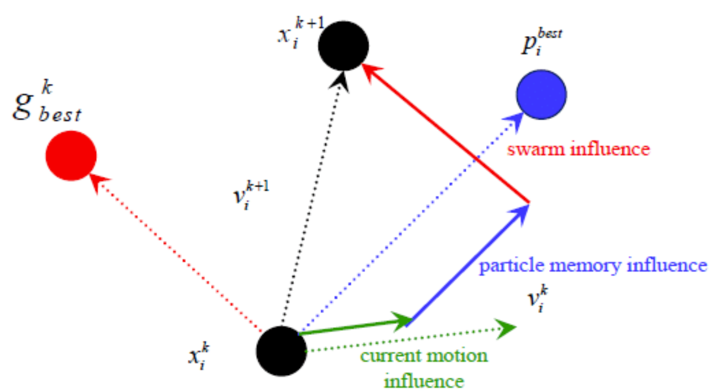

Figure 3: PSO Algorithm – Movement of the particle

In the above equation, w, c_1 and c_2 are the acceleration factors due to inertial effect, pbest and gbest respectively for the ith particle, r_1 and r_2 refers to the random numbers between 0 and 1. x and v refers to the position and the velocity of the particle. In this study the PSO algorithm is incorporated to find the best solution of three operational parameters which can be controlled by the drillers on the rig namely (WOB, RPM, FLOW). The other components are reactive in nature and are maintained at constant levels based on the well design conceived before the drilling process. In non-autonomous drilling operation, it may not be possible for the driller to change the operational parameters for every feet and hence, the optimization is performed over a 20 ft section interval. PSO Algorithm for Optimal ROP Set the parameters of the PSO algorithm (c1,c2,w), number of particles Set the number of dimensions (d = 3, WOB,RPM, FLOW) For every section, evaluate the non-varying parameters (SPP,Torque, Section diameter , formation type) For every particle and for every dimension, generate random numbers (rp and rg) Compute the ROP (fitness value) by passing the particle (WOB, RPM, FLOW) to the ANN model after scaling &

pre-processing If the current fitness value (ROP) is greater than the global best fitness value (gBest) in history, then update it. If the fitness value (ROP) is greater than the best personal fitness value (pBest) in history, then update it. Update the velocity and the positional parameter as per equation x Repeat steps 3 to steps 7 for every section until the position of the particle stabilizes to obtain the optimal parameters in that section

4. Data Pre-processing

4.1 Data Selection

In this study, data from vertical wells drilled in the same field in Middle East Asia is used for ROP modeling and optimization. The data in real time is collected from drilling data acquisition units on the surface of the rig. It includes WOB, RPM, torque, stand pipe pressure(SPP), flow in the pumps, Depth of cut (DOC). These data are depth indexed sampled at 0.5 ft interval and the three wells used in this study spans a total footage of 6750 ft. In order to ensure similar wells are used in the study, the datasets are selected such that static parameters in these wells namely the drill string configuration, bottom- hole assemblies, mud properties, the nature of drill bits, blade counts, drill – bit cutter configurations are similar between them. These datasets span five different types of formation (named 1-5) (rock types) and these formations were drilled with a 12.25 inch diameter bits as shown in Table 1. The instantaneous ROP measurement from Well A for different formations is shown in figure 4 which is measured as a ratio of the distance between two hole depth instances to the finite time for drilling this distance.

	Well A		Well B		Well C	
	Span (ft)	ROP ft/hr)	Span(ft)	ROP (ft/hr)	Span (ft)	ROP (ft/hr)
Formation 1	284.5	404.4	229.5	320.6	279.5	349.0
Formation 2	754.5	430.5	704.5	387.2	699.5	409.9
Formation 3	489.5	234.9	414.5	230.2	449.5	447.6
Formation 4	729.5	247.6	784.5	360.1	739.5	404.5
Formation 5	52.5	106.4	94.5	90.6	45.5	170.4

Table 1: Formation Span and average ROP of wells under consideration

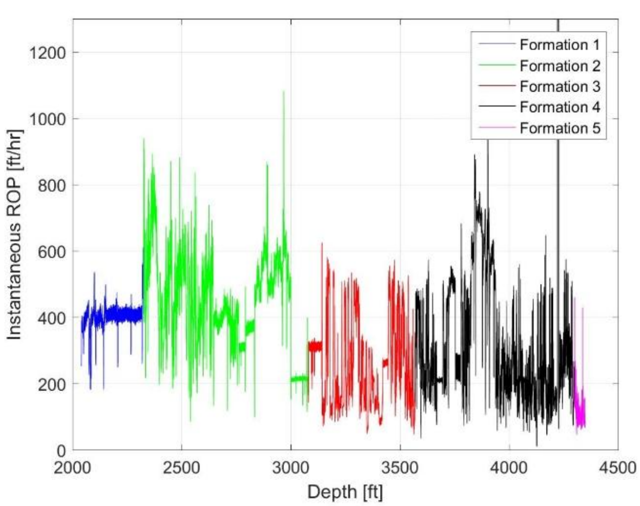

Figure 4: Well A – Instantaneous ROP measurement

4.2 Data Quality Analysis

Abnormal data can influence the model negatively and may restrict the model in its generalization. Suspect data is flagged in the range validation procedure which checks if the measurement can be within the limits of rig limitations and acceptable trends. For example, WOB or RPM can never be zero or negative during. Similarly, stand pipe pressure cannot be of a small magnitude during drilling activity since there should be adequate pressure in the drilling system to clean the hole and maintain formation integrity. The dataset is tested for anomaly by Z-score outlier detection algorithm where a threshold Z-Score of 3 is selected [13] and any data points above this threshold is marked as an outlier and excluded from the training data. The resulting data is normalized between (-1, 1) to remove geometrical biases towards some of the dimensions of the data vectors so that every bit of data gets treated in consistent manner.

5. Results and Discussion

5.1 ANN Training and Testing

The main concept of ANN is to set an algorithmic pattern for determining the output (ROP) of values to input data set (WOB, RPM, TORQUE, SPP, FLOW, formation type). It should be noted that formation type is a categorical variable which is converted to a numeric quantity using encoding methods to train the ANN. The dataset for Well A and Well C are combined to produce the training set and Well B is reserved as a blind dataset upon which the model will be evaluated. The combined dataset is utilized in such a way that 80% of the input data was taken for network training and 20% for validation in a completely randomized way. The accuracy of the data model is measured using R-metric (Coefficient of determination) which serves as a metric to score the predictions of the ANN model with the expected result. It is given by, The challenge in the design of neural network is to perfectly tune the hyper parameters and to make proper choice of

$$\text{R (Coefficient of Determination)} = 1 - \frac{\sqrt{\Sigma y^2_{\text{predicted}} - y^2_{\text{actual}}}}{n}$$

activation function so that the network generalizes well and do not perform over fitting. Tuning is usually done either by employing a grid search or manually changing the parameters by brute force. In this study, a loop of different values of hidden layers (neurons) and different values of activation function was run and then the best parameters with minimal error (R-metric closer to one) is chosen as the best model parameter. The available activation functions in scikit-learn in python module are shown in table 2.

Table 2: Activation function

Activation Function	Equation
Sigmoid	$y = \frac{1}{1+e^{-x}}$
Tanh	$y = \tanh(x)$
Rectified Linear Unit	$y = \max(0,x)$

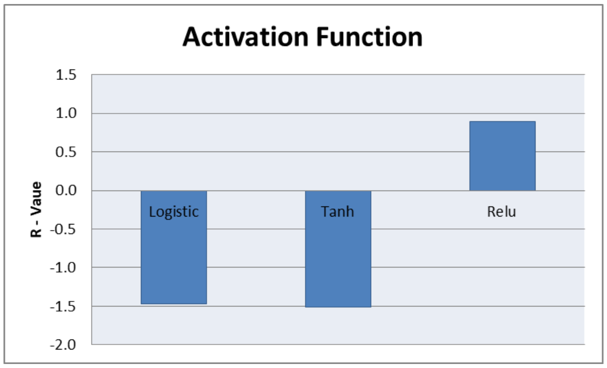

Figure 5: R value for different activation functions

A sensitivity analysis on the activation yielded Rectified Linear unit function to have the lowest training and testing error as shown in figure 5. After a number of simulations, the BP network utilized 60 neurons in the hidden layer to provide optimum accuracy in the training and validation dataset as shown in figure 6. The neural network gave satisfactory results with the Coefficient of determination around 0.99 and was not sensitive to the number of hidden layers beyond 40. Even though 40 hidden layers would suffice, the computational overhead at the rig will be adequate to accommodate more neurons and hence, 60 hidden layers were selected in the subsequent model application.

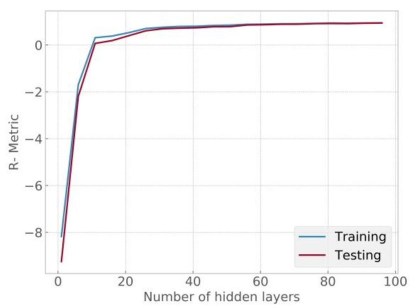

Figure 6: Optimum neurons in the hidden layer

5.2 Blind Well Testing

The ANN model is blind tested with Well B on its drilling parameters which were drilled with the same drill string configuration and with the same formation sequence to predict ROP. It must be noted that the same data quality engineering techniques applied on the Well A drilling parameters is also applied on Well B drilling parameters before the model is evaluated. This is performed to ensure that the model inputs between the training and the blind well testing are consistent. The plot of actual ROP and the predicted ROP for the formation 2 is evaluated as shown in figure 7 and the error metrics were also plotted in table 3. The R value between the predicted ROP and the actual ROP from the regression indicates that the predictions are not random and infers that the neural network model predictions of ROP are in close agreement with the actual ROP across all the formation. Formation 5 was not taken into consideration as its span is relatively very small and the data modeling will not be appropriate.

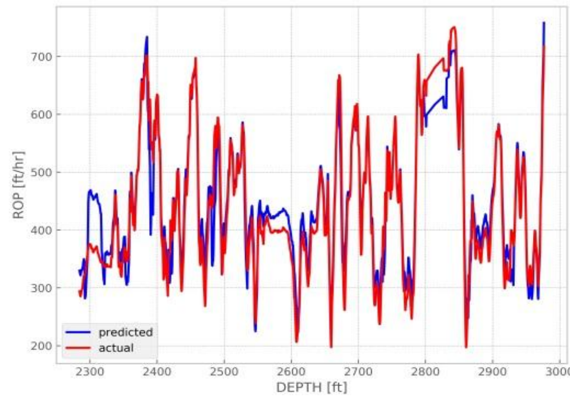

Figure 7: ROP prediction by ANN on formation 2 – Well C

Table 3: R – Value for all the formations in Well C

Formation	R - Value
1	0.70
2	0.73
3	0.89
4	0.65

5.3 BP-NN based PSO Optimization

In this section, the drilling parameters are optimized by the PSO algorithm to achieve the best ROP for a specific formation. In order to benchmark the resulting drilling parameters from this optimization algorithm, a brute force estimation of ROP is also developed. The brute force optimization generates every possible scenario of WOB, RPM, FLOW in a sequential basis and evaluates the ROP from the ANN model for every scenario. Theoretically, this could be considered to be the best ROP that can be achieved from the drilling parameters. Mathematically, maximum ROP can be achieved with maximum WOB, RPM or flow but due to the manufacturing, downhole and rig constraints, it may not be feasible In order to discuss the implementation of the algorithm, formation 2 is considered as a reference. Considering the practical nature by which the drilling parameters are changed in the drilling rig, the optimal parameters are evaluated for every 40 ft in formation 2. Hence, the dataset contains 18 depth intervals. The constraints on the drilling parameters are set based on the minimum and the maximum of WOB, RPM and FLOW on the previous section as the driller can change drilling parameters in smaller steps. The inertial, social and personal weights of the PSO are set to w=0.9, c_1=0.5, c_2=0.3. Calculations were performed for 100 iterations for 10 particles in three dimensions to determine the optimal parameters and compared with the brute force estimation and shown in table 4.

Table 4 : Results from PSO optimization

Depth [ft]	RPM		WOB (klbs)		ROP (ft/hr)	
	PSO	Actual	PSO	Actual	PSO	Actual
2330	83.2	83.2	4.0	2.4	458.2	459.7
2370	94.4	94.4	33.7	35.7	344.0	343.9
2410	111.6	111.6	19.9	19.8	453.0	453.8
2450	85.0	84.9	10.4	9.9	363.7	364.7
2490	92.6	92.4	24.0	24.3	513.9	514.8
2530	137.0	136.2	31.4	30.3	389.2	386.5
2570	104.0	103.9	29.6	31.3	275.6	276.8
2610	122.2	121.3	15.3	14.9	255.3	254.3
2650	120.7	119.5	8.2	8.1	348.7	347.2
2690	114.6	114.5	11.7	11.4	645.4	646.1
2730	123.7	122.7	10.0	9.3	541.1	542.5
2770	120.7	120.6	10.4	10.3	510.1	509.9
2810	131.3	131.0	5.9	5.8	818.7	818.3
2850	90.3	88.1	1.2	1.2	818.5	817.4
2890	126.7	126.3	3.2	2.9	535.3	534.3
2930	130.6	130.8	17.0	17.0	357.6	357.7
2970	126.0	121.7	13.8	20.0	402.5	394.1

The ANN-PSO algorithm showed promise in providing a better set of optimal drilling parameters at the same time achieving higher ROP. Moreover, the PSO algorithm was also able to produce the optimal ROP to the maximum achievable ROP in that section. The results from table 5 also indicate that percentage increase of RPM is much better than WOB or flow and hence controlling RPM has a better significance to produce optimal ROP for formation 2 which was under consideration. Another significant observation as shown in figure 8 is that the consistent way in which the PSO algorithm reduces the cost per iteration and achieved the optimal ROP. This showed that the running the PSO algorithm within 40 iteration can be practically achieved in the drilling rig where the computational resources will be limited. By providing such options with the help of simulations and optimization, the driller can choose the right drilling parameters depending on the practical feasibility of the rig, lithology type and the thickness of the formation in real time. This algorithm was also extended to other formation types in the available dataset and the table 6 provided a savings of 80 minutes in 6 hours of drilling time in identified formations with the ANN-PSO algorithm. During drilling there are many factors such as stick slip, vibrations which can additionally affect the drilling efficiency and the operational cost. Such problems can also be incorporated to find the optimal ROP under such circumstances by adding additional constraints to the PSO model. The solution search space of the PSO algorithm will now be reduced by the additional constraints. By this way, the optimal solutions will be realistic and representative of the field situation.

Table 5: Improvement in the drilling parameters (before and after) optimization

	Before Optimization	After Optimization
Average Flow (gal/m)	939	940
Average RPM	228	110
Average WOB (klbs)	21	14

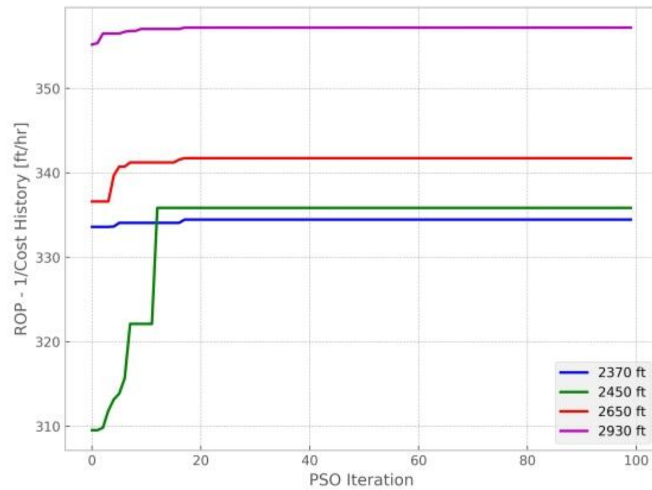

Figure 8: PSO Iterations

Table 6: Net savings in the time across formation due to PSO optimization

	Depth (ft)	Time to drill (hours) Before (Optimization)	After (Optimization)	Net Savings in Time (minutes)
Formation 2	704.5	1.82	1.48	20
Formation 3	414.5	1.80	1.44	21
Formation 4	784.5	2.18	1.54	39

6. Conclusion

This paper demonstrates the practical use of data model & learning methods in drilling engineering. Machine learning models can be effectively used to predict ROP along the entire well taking into account the surface operational parameters namely RPM, WOB, Flow, Torque and stand pipe pressure. The neural network type of algorithms enabled to capture the patterns of the drilling parameters that affect the ROP and predict the ROP in the blind well. This research also showed new ways to minimize drilling cost by the use of particle swarm optimization algorithm. When taking into account the operational parameters like WOB, RPM etc, there are infinite number of solutions depending on the rig limitation and algorithms like PSO can enable us to quickly determine the optimal parameters which could be incorporated in real-time. Since the implementation of such model is simple, this technique could be extended to real-time prediction and optimization at the surface of the rig without the need to rely on the downhole parameters and also be used in edge analytics.

7. Refrences

[1] Zakariya, Rafik, Alaa Zein, Emad Diab, Amr Lotfy, Chris Marland, Yaqoub Yousif, and Al Obaidli. "A Case Study of Real-Time Drilling Optimization to Improve Well Delivery through Enhancing Drilling Rates and Identifying Invisible Lost Time to Improve Performance." *SPE North Africa Technical Conference and Exhibition* ,Cairo, Egypt. Society of Petroleum Engineers. 14-16, September, 2015. https://doi.org/10.2118/175748-MS.

[2] Bourgoyne, A. T., and F.S. Young Jr. "A Multiple Regression Approach to Optimal Drilling and Abnormal Pressure Detection." *SPE Journal* 14 (4). pp. 371–84. 1974. https://doi.org/10.2118/4238-PA.

[3] Warren, T M, and W K Armagost. 1988. "Laboratory

Drilling Performance of PDC Bits." SPE Drilling Engineering. (SPE 15617). Society of Petroleum Engineers 3(02): pp.125–35. https://doi.org/10.2118/15617-PA

[4] Rampersad P. R., Hareland G., Boonyapaluk P. "Drilling optimization using drilling data and available technology," *SPE Latin America/Caribbean Petroleum Engineering Conference*, Buenos Aires, Argentina, 27-29 April, 1994. https://doi.org/10.2118/27034-MS

[5] Eskandarian, Sajad, Peyman Bahrami, and Pezhman Kazemi. "A Comprehensive Data Mining Approach to Estimate the Rate of Penetration: Application of Neural Network, Rule Based Models and Feature Ranking," *Journal of Petroleum Science and Engineering,*156(07),pp.605–15,2017. https://doi.org/10.1016/J.PETROL.2017.06.039.

[6] Dunlop, J, R Isangulov, W D Aldred, H A Sanchez, J L S Flores, J a Herdoiza, J Belaskie, and J C Luppens. "Increased Rate of Penetration Through Automation," *SPE/IADC Drilling Conference and Exhibition*. Society of Petroleum Engineers; 1-3 March, 2011. Amsterdam, The Netherlands. https://doi.org/10.2118/139897-MS.

[7] Mantha, B, and R Samuel. "ROP Optimization Using Artificial Intelligence Techniques with Statistical Regression Coupling," SPE Annual Technical Conference and Exhibition. Society of Petroleum Engineers. 2016.https://doi.org/10.2118/181382-MS.

[8] Hegde, Chiranth, and K. E. Gray. "Use of Machine Learning and Data Analytics to Increase Drilling Efficiency for Nearby Wells." *Journal of Natural Gas Science and Engineering*. Vol. 40. 2017.https://doi.org/10.1016/j.jngse.2017.02.019.

[9] Atashnezhad, A., Wood, D. A., Fereidounpour, A. "Designing and optimizing deviated wellbore trajectories using novel particle swarm algorithms. " *Journal of Natural Gas Science and Engineering*.21:pp.1184-1204.2014

[10] Onwunalu, J.E. and Durlofsky, 2010. "Application of a particle swarm optimization algorithm for determining optimum well location and type." *L.J. Computational Geosciences* 14: 183. 2010. https://doi.org/10.1007/s10596-009-9142-1

[11] Jiang, W., & Samuel, R. "Optimization of Rate of Penetration in a Convoluted Drilling Framework using Ant Colony Optimization." *Society of Petroleum Engineers. IADC/SPE Drilling Conference and Exhibition*, 1-3 March, 2016 Fort Worth, Texas, USA doi:10.2118/178847-MS

[12] Self, R., Atashnezhad, A., & Hareland, G. "Reducing Drilling Cost by finding Optimal Operational Parameters using Particle Swarm Algorithm." *Society of Petroleum Engineers. SPE Deepwater Drilling and Completions Conference*, 14-15 September, 2016, Galveston, Texas, USA . doi:10.2118/180280-MS

[13] Tripathy, Sushree Swarupa. " Testing Data on Analysis of Lead in Aqueous Solution." *American Journal of Theoretical and Applied Statistics* 2 (6):pp.233-240,2010. https://doi.org/10.11648/j.ajtas.20130206.21

[14] Pedregosa, Fabian, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss. "Scikit-Learn: Machine Learning in {Python}." *The Journal of Machine Learning Research* 12:pp.2825–2830.2011.