Energy Proceedings
Vol 41, 2024
ISSN 2004-2965

Real-time prediction of oil and gas drilling rate based on physics-based model and particle filter method

Chengkai Zhang [1], Xianzhi Song [1*], Yinao Su [1, 2]

1 College of Petroleum Engineering, China University of Petroleum (Beijing), Beijing, 102249, China

2 CNPC Engineering Technology R&D Company Limited, Beijing, 102206, China

(*Corresponding Author: songxz@cup.edu.cn)

ABSTRACT

Oil and gas drilling, essential for exploring and exploiting petroleum resources, involves significant time, labor, and costs, often exceeding $300,000 daily. Predicting the drilling rate (Rate of Penetration, ROP) accurately and promptly is crucial for improving efficiency and reducing expenses. In drilling, physics-based and machine learning models are typically used for ROP forecasting. Physics-based models, while intuitive, often lack precision in complex conditions. Machine learning models, though precise, face challenges with data availability and training costs in real-time settings. This paper introduces a novel approach combining a physics-based model with a particle filter algorithm for real-time ROP prediction. It adapts the Bourgoyne-Young ROP model and Markov assumptions into a state space model, using the particle filter to estimate elusive coefficients through probability theory. This enables real-time data updates for more accurate ROP predictions. The proposed framework is evaluated against traditional models using open-source and field drilling datasets in post-drilling and real-time scenarios. Results show conventional physics-based models fall short in both scenarios, while machine learning and the new particle filter model show significant improvements. In post-drilling analysis, these models achieve under 5% mean relative error. For real-time predictions, machine learning models have over 20% error, but the particle filter model reduces this to approximately 15%. This highlights the particle filter model's superiority in accuracy and cost-effectiveness under dynamic and uncertain drilling conditions. This paper presents a robust, efficient solution for ROP prediction and optimization, marking a significant advancement in the drilling field.

Keywords: rate of penetration, particle filter, real-time prediction, physics-based model, machine learning

NONMENCLATURE

Abbreviations	
ROP	Rate of Penetration
D	Depth
WOB	Weight on Bit
N	Rotary Speed
Q	Flow Rate
RMSE	Root Mean Squared Error
MRE	Mean Relative Error
RF	Random Forest
SVM	Support Vector Machine
Symbols	
a_1 to a_5	Equation coefficients
$\boldsymbol{x}_k$	The state variables at time k
$\boldsymbol{y}_k$	The observed variable at time k
$\boldsymbol{u}_k$	The state noise
$\boldsymbol{v}_k$	The observation noise
$\boldsymbol{z}_k$	The random walk value at time k
$w_{k,j}^{i}$	The weight of the j-th state variable of the i-th particle at time k

1. INTRODUCTION

Drilling represents a costly, high-risk, and formidable undertaking, where the cornerstone of efficiency hinges upon the intricate process of wellbore extension by penetrating formations. This drilling process is intricately woven into a tapestry of influences, ranging from geological formations, drilling equipment, and downhole fluids, to an array of engineering parameters such as weight on bit, rotary speed, and flow rate. These factors, in conjunction with geological parameters, collectively shape the dynamics of the drilling process. Therefore,

This is a paper for 15th International Conference on Applied Energy (ICAE2023), Dec. 3-7, 2023, Doha, Qatar.

optimizing the engineering parameters of the drilling process stands as a potent strategy for elevating ROP, leading to improved drilling efficiency, shorter drilling durations, and reduced overall drilling costs. In essence, a profound understanding of how various factors impact drilling rates, and the establishment of precise drilling rate prediction models, form the bedrock upon which parameter optimization is built [1].

Currently, research on drilling rate prediction models mainly focuses on two aspects: physics-based drilling rate models and intelligent drilling rate models [2]. Physics-based models are the most used drilling rate prediction models. Previous researchers have analyzed the mechanisms by which various factors influence drilling rate and have constructed various forms of drilling rate equations, such as the Maurer equation [3], Warren equation [4], Bourgoyne-Young equation [5], Hareland equation [6], Detournay equation [7], Motahhari equation [8], and the modified Yang equation [9]. Researchers [10-12] have determined the coefficients of drilling rate equations using field drilling data through multivariate regression to predict and optimize drilling rates. Physics-based drilling rate models align with drilling principles, exhibit good comprehensibility, have a simple form, and offer easy parameter determination. As a result, they have found wide-ranging applications. However, the drilling process is exceedingly complex, and the strong nonlinear relationships between various factors and drilling rates are often challenging to accurately describe. Moreover, dynamic changes in downhole conditions and the presence of uncertainty factors can disrupt the accuracy of physics-based models.

In recent years, data-driven models, with machine learning models as representatives, have witnessed rapid development in the field of drilling rate prediction and optimization [13-16]. Data-driven models, primarily based on extensive drilling data, can identify nonlinear relationships between drilling rates and various influencing factors, often without the need for extensive physics-based analysis. This allows them to accurately predict drilling rates. Hegde and Soares [17-20] have explored the effectiveness of both physics-based and intelligent models in predicting drilling rates in different scenarios and with different models. Hassan et al. [21] employed artificial neural networks for drilling rate prediction and coupled the predicted rates with the mechanical-specific energy for comprehensive drilling efficiency optimization. Negara and Bilal [22] established machine-learning models based on input parameters from physics-based models to predict drilling rates. Ahmed et al. [23] compared the performance of four commonly used prediction models, including neural networks. Elkatatny [24] constructed neural networks based on an adaptive differential evolution algorithm for prediction. Li et al. [25] investigated the issue of training data volume for intelligent models. Li [26] applied algorithms such as Bagging, Random Forest, and Gradient Boosting Trees to train drilling rate regression models. Previous research has shown that the accuracy of intelligent models far surpasses that of physics-based models. However, they heavily rely on data, necessitate a wide range of input parameters for model training, demand a large quantity of high-quality data, and suffer from poor interpretability and transferability.

In summary, both physics-based models and data-driven models have their respective strengths and weaknesses. The former are simple to use, and have clear underlying principles, but exhibit lower accuracy. The latter yield superior prediction results but struggle with transferability and require high-quality data. In addition to accuracy and transferability, practical drilling rate models should possess dynamics, uncertainty, and real-time capability in the actual drilling process. Firstly, drilling rates are influenced by a multitude of factors, and the underground environment is complex. Therefore, the parameters of drilling rate models should be dynamic as there are inevitably unaccounted factors during the modeling process. Secondly, due to the complexity of the underground environment and noise introduced by measurement instruments, on-site data may carry some uncertainty. Consequently, the parameters of drilling rate models should be estimated, and they come with associated uncertainties and confidence intervals. Finally, the ever-changing underground conditions necessitate drilling rate models to have the ability to adapt dynamically, meaning the established models should be able to update in real time with incoming data. However, the current literature on drilling rate prediction primarily relies on post-drilling analysis scenarios, where researchers use historical drilling data for modeling and prediction, making full use of all drilling data. Only a limited number of researchers [10, 12, 18, 20, 27] have worked on establishing drilling rate models based on real-time drilling scenarios, utilizing partial, real-time transmitted data. Few have studied the dynamic, uncertain, and real-time aspects of the drilling process. Particle filtering is a widely applied algorithm for estimating the state of nonlinear systems based on observed data [28, 29]. It employs probabilistic methods to estimate the current system state, making it an effective tool to handle the dynamics, uncertainties, and real-time aspects of the drilling process [30, 31].

This paper aims to investigate the applicability of the particle filtering algorithm in real-time estimation of drilling rate equations and real-time prediction of drilling rates. Building upon the modified Bourgoyne-Young physics-based drilling rate equation [18], this study utilizes the particle filtering method to estimate the coefficients of the drilling rate equation in real time. The study compares the drilling rate prediction performance of physics-based drilling models, intelligent drilling models, and particle-filter-based physics-based models in two scenarios: post-drilling analysis and real-time prediction. It reveals the advantages of the particle filtering model in handling the dynamic, uncertain, and real-time characteristics of the drilling process. The findings of this research can provide valuable insights for real-time drilling rate prediction, optimization, and decision-making in drilling operations.

2. PARTICLE FILTERING-BASED DRILLING RATE EQUATION ESTIMATION MODEL

2.1 Physics-based ROP model

Physics-based drilling rate models have gained widespread acceptance due to their simplicity, interpretability, and alignment with drilling principles. The Bourgoyne-Young drilling rate equation [5] stands as the most used physics-based drilling rate model. This model encompasses various factors, such as formation strength, drilling pressure, rotary speed, flow rate, tooth wear, hydraulics, and pressure differentials, that influence drilling rates. Despite its comprehensive consideration of factors affecting drilling rates, many parameters within the Bourgoyne-Young model are challenging to obtain in practical applications, such as formation compactness and tooth wear. Furthermore, the model incorporates several standardized empirical parameters, like normalized depth and normalized drilling pressure, which need re-estimation when different drill bit equipment is used. This significantly limits the model's applicability. Soares and colleagues [18], based on theoretical analysis, have introduced a modified Bourgoyne-Young model that eliminates standardized empirical parameters and focuses solely on key influencing factors.

$$ROP = a_1 D^{a_2} WOB^{a_3} N^{a_4} Q^{a_5} \quad (1)$$

In the equation, where a_1 to a_5 are the equation coefficients, ROP represents the drilling rate (m/h), D represents the well depth (m), WOB represents the weight on bit (kN), N stands for rotary speed in RPM, and Q represents the flow rate in L/min.

Soares and his team compared the performance of the modified model and the original model on drilling data from 18 different formations. They found that the modified Bourgoyne-Young model had the lowest absolute error in 14 formations and the lowest root mean square error in all 16 formations, indicating that the modified model simplified parameter acquisition and calculation processes while maintaining computational accuracy. Therefore, this paper employs the modified Bourgoyne-Young model to describe the relationship between drilling rates and other parameters in the drilling process.

2.2 State space model of the drilling process

Based on a first-order Markov assumption, the drilling process can be described using the state-space model shown in Fig. 1 [32, 33]. This model consists of the state equation and the observation equation in Eq. (2), where the state equation characterizes the evolution of state variables, and the observation equation describes the relationship between system variables and observed values. Due to the complexity of the subsurface environment, the coefficients of the drilling rate model should be dynamic during the drilling process, while drilling rate, well depth, and drilling engineering parameters are measurable. Eq. (1) links the drilling rate with well depth and engineering parameters through the coefficients a_1 to a_5. Therefore, for the drilling process described by Eq. (1), this paper regards the equation coefficients a_1 to a_5 as the state variables x of the system and drilling rate (ROP), well depth (D), weight on bit (WOB), rotary speed (N), and flow rate (Q) as the observable variables y. Using filtering methods based on observed data, Eq. (2) estimates the current state variables, enabling the determination of the current drilling rate model. This forms the foundation for subsequent drilling rate prediction and optimization.

$$\begin{cases} \boldsymbol{x}_k = f(\boldsymbol{x}_{k-1}) + \boldsymbol{u}_k \\ \boldsymbol{y}_k = h(\boldsymbol{x}_k) + \boldsymbol{v}_k \end{cases} \quad (2)$$

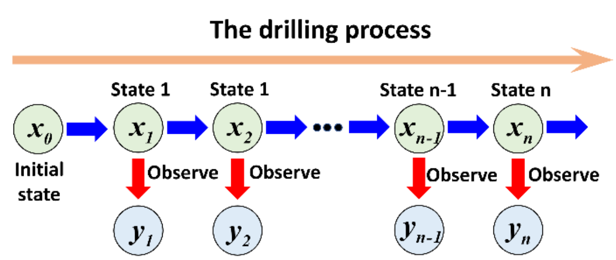

Fig.1. State space model of the drilling process

In the equation, $\boldsymbol{x}_k$ and $\boldsymbol{x}_{k-1}$ represent the state variables at time k and time $(k-1)$, respectively. $\boldsymbol{y}_k$ denotes the observed variable at time k, while $\boldsymbol{u}_k$ and

$\boldsymbol{v}_k$ represent the noise in the state transition process and the observation process, respectively. It is assumed that the system's observed variables $\boldsymbol{y}$ and the state noise $\boldsymbol{u}_k$ and observation noise $\boldsymbol{v}_k$ are independent, and both the state noise and observation noise follow Gaussian distributions.

Based on Eq. (1), we can derive the observation equation as follows:

$$\begin{cases} y_{k,1} = x_{k,1} + x_{k,2}y_{k,2} + x_{k,3}y_{k,3} + x_{k,4}y_{k,4} + x_{k,5}y_{k,5} + v_k \\ y_{k,2} = (y_{k,1} - x_{k,1} - x_{k,3}y_{k,3} - x_{k,4}y_{k,4} - x_{k,5}y_{k,5})/x_{k,2} + v_k \\ y_{k,3} = (y_{k,1} - x_{k,1} - x_{k,2}y_{k,2} - x_{k,4}y_{k,4} - x_{k,5}y_{k,5})/x_{k,3} + v_k \\ y_{k,4} = (y_{k,1} - x_{k,1} - x_{k,2}y_{k,2} - x_{k,3}y_{k,3} - x_{k,5}y_{k,5})/x_{k,4} + v_k \\ y_{k,5} = (y_{k,1} - x_{k,1} - x_{k,2}y_{k,2} - x_{k,3}y_{k,3} - x_{k,4}y_{k,4})/x_{k,5} + v_k \end{cases} \tag{3}$$

In the equation, $\boldsymbol{y}_k = [y_{k,1}, y_{k,2}, y_{k,3}, y_{k,4}, y_{k,5}]^T = [\ln ROP_k, \ln D_k, \ln WOB_k, \ln N_k, \ln Q_k]^T$ represents the natural logarithm of the observed values of drilling rate, well depth, weight on bit, rotary speed, and flow rate at time k. $\boldsymbol{x}_k = [x_{k,1}, x_{k,2}, x_{k,3}, x_{k,4}, x_{k,5}]^T = [\ln a_{k,1}, a_{k,2}, a_{k,3}, a_{k,4}, a_{k,5}]^T$ represents the values of the state variables at time k.

The state equation reflects the relationship between state variables at consecutive time steps; however, the coefficients in Eq. (1) are unknown. According to the literature [34], when the state equation is unknown, it can be approximated using a random walk, as follows:

$$\boldsymbol{x}_k = \boldsymbol{x}_{k-1} + \boldsymbol{z}_k \tag{4}$$

In the equation, $\boldsymbol{z}_k$ represents the random walk value at time k, which is drawn from a Gaussian distribution.

With this formulation, the paper has established a nonlinear state-space model for the drilling process based on Eq. (1):

$$\begin{cases} \boldsymbol{x}_k = \boldsymbol{x}_{k-1} + \boldsymbol{z}_k + \boldsymbol{u}_k \\ \boldsymbol{y}_k = h(\boldsymbol{x}_k) + \boldsymbol{v}_k \end{cases} \tag{5}$$

2.3 Particle filter algorithm

Particle filtering, a filtering method suitable for nonlinear system state estimation, has gained popularity in recent years. It has found widespread applications in various fields such as target tracking, fault diagnosis, communications, satellite navigation, and sonar localization [28, 29]. Particle filtering employs Monte Carlo methods to estimate the Bayesian optimal estimate. Its principle involves representing the posterior probability of a random event through multiple random samples (referred to as particles) with associated weights, allowing for the estimation of the system's state from noisy observed sequential data [33]. This paper utilizes the Sequential Importance Resampling (SIR) particle filtering algorithm to estimate the state of the drilling process. The value of the j-th state variable at the k-th time step (or sequence step based on depth, hereafter referred to as a time step) is approximated using particles $x_{k,j}^i$, where $i = 1,2,\dots,N$ represents the particle number, N is the total number of particles, k denotes the time step, $j = 1,2,3,4,5$ indicates the state variable. The workflow of the SIR particle filtering is as follows:

(1) Assuming the probability density function of the initial state, $p_{0,j}$, is known (in this paper, it's assumed to be a Gaussian distribution), random samples are drawn from $p_{0,j}$ to initialize each particle, $x_{0,j}^i$, and each particle is assigned an initial weight $w_{0,j}^i = 1/N$.

(2) $k = 1,2,\dots$, perform the following steps:

Predict the state values for each particle at time k using the state equation based on the state values at the time $(k-1)$:

$$x_{k,j}^i = x_{k-1,j}^i + z_{k,j} + u_{k,j} \tag{6}$$

Update the particle weights based on the new observation $y_{k,j}^i$ at time k:

$$w_{k,j}^i = g\left(y_{k,j}^i - h\left(x_{k,j}^i\right)\right) w_{k-1,j}^i \tag{7}$$

where $g()$ represents the probability density function of the observation noise $\boldsymbol{v}_k$.

Normalize the weights of all particles:

$$w_{k,j}^i = \frac{w_{k,j}^i}{\sum_{i=1}^N w_{k,j}^i} \tag{8}$$

Resample N particles individually: For the i-th particle, first generate a random number r from a uniform distribution in [0,1]. Then, accumulate the weights of the particles until their sum is greater than r. That is, find m such that $\sum_{i=1}^{m-1} w_{k,j}^i < r$ and $\sum_{i=1}^{m} w_{k,j}^i > r$. Set the new particle's weight as $w_{k,j}^i = w_{k,j}^m$.

Calculate estimates of statistical quantities like the mean, variance, and confidence intervals of the j-th state variable at time k:

$$x_{k,j} = \sum_{i=1}^N w_{k,j}^i x_{k,j}^i \tag{9}$$

Reset all particle weights for the next filtering step:

$$w_{k,j}^i = 1/N \tag{10}$$

Following this process allows the calculation of estimates for the system's various state variables at each time step.

3. EXPERIMENT CONFIGURATION

3.1 Dataset

This study is based on publicly available drilling data from the 58-32 geothermal well in the FORGE geothermal project in Utah, USA [35, 36], and actual drilling data from the K5 well in a domestic oilfield. The drilling rate data for both wells is shown in Fig. 2 and 3. The 58-32 geothermal well has a depth of 2297 meters, and this study analyzes 400 data points in the interval from 1425 to 1596 meters. This section used a drill bit with a diameter of 222.25mm and consists of a poor quartz sand granite formation. The K5 well has a depth of 6538 meters, and this study analyzes 1000 data points in the interval from 3931 to 4930 meters. This section used a drill bit with a diameter of 333.3mm.

As observed in Fig. 2 - 3, the drilling rate trends for these two wells vary significantly. The 58-32 geothermal well exhibits frequent fluctuations in drilling rate with relatively good overall regularity. In contrast, the K5 well's drilling rate in the first half displays irregular "zigzag" variations. This difference is attributed to the high-frequency adjustments of engineering parameters like drilling pressure, rotary speed, and flow rate in the 58-32 geothermal well, while the engineering parameters for the K5 well often remain constant for several meters. Comparing these two datasets with vastly different trends will help reflect the applicability of the various models.

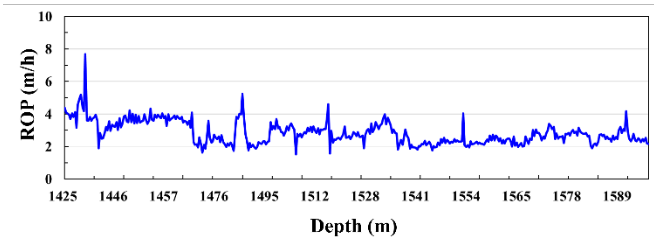

Fig. 2. Rate of penetration versus depth of geothermal well 58-32

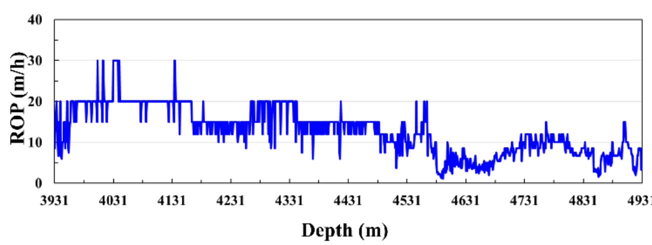

Fig. 3. Rate of penetration versus depth of well K5

3.2 Evaluation Metrics

In this paper, we first compare the performance of physics-based models, filter-based physics-based models, and commonly used intelligent models on two datasets. To do this, it is necessary to select evaluation metrics that measure the performance of each model. This paper selects Root Mean Squared Error (RMSE) and Mean Relative Error (MRE) as performance evaluation metrics. RMSE measures the deviation between estimated values and actual values, and its calculation formula is as follows:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2} \tag{11}$$

MRE represents the average relative error between estimated values and actual values, and its calculation formula is as follows:

$$MRE = \frac{1}{n}\sum_{i=1}^{n}\frac{|\hat{y}_i - y_i|}{y_i} \tag{12}$$

Here, n represents the sample size, $\hat{y}$ represents the estimated sample values, and y represents the actual sample values.

3.3 Model comparison

The two most common types of models in drilling speed prediction are physics-based regression models based on physics-based equations and intelligent models based on artificial intelligence algorithms. In this paper, we compared the performance of physics-based regression models, intelligent models, and filter-based models in two scenarios: post-drilling data analysis and real-time drilling speed prediction. The settings for each model are as follows:

(1) Physics-based Regression Model

The modified Bourgoyne-Young drilling speed equation is selected as the physics-based model, and a multiple regression method is used to obtain the values of the drilling speed equation coefficients. In the post-drilling data analysis scenario, the drilling speed equation is fitted using all n data points, and then all the data is input into the fitted equation to compare the RMSE and MRE of the regression drilling speed with the actual drilling speed, as shown in the process in Fig. 4. In the real-time drilling speed prediction scenario, the process is carried out as shown in Fig. 4. First, the drilling speed equation is fitted using the first k data points, and then the engineering parameters of the $(k+1)$-th data point are input to predict the drilling speed of the $(k+1)$-th data point, which is then compared with the actual value to calculate its MRE. After that, the drilling speed equation is fitted using the first $(k+1)$ data points to predict and compare the drilling speed of the $(k+2)$-nd point, and this process is repeated.

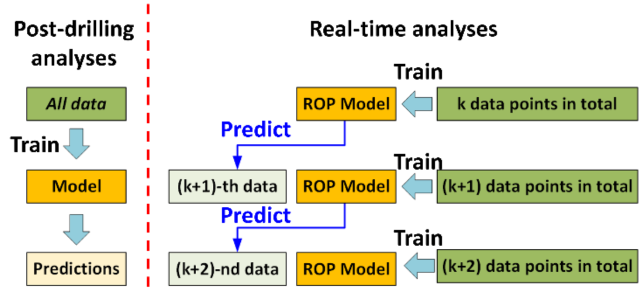

Fig. 4. Schematic of post-drilling data analysis and ROP real-time prediction

(2) Machine learning Models

In drilling speed prediction research, common intelligent models include Artificial Neural Networks (ANN), Support Vector Machines (SVM), Random Forest (RF), and Extreme Gradient Boosting (XGBoost), among others. Hegde et al. [20] and Echeverri Duque [37] compared the performance of these commonly used intelligent models in drilling speed prediction and found that SVM and RF had higher accuracy in drilling speed prediction. Therefore, this paper selects the SVM model and RF model for comparison, with the same operational procedures in both scenarios as the physics-based regression model. To optimize the performance of SVM and RF, this paper tunes their hyperparameters (Table 1 and Table 2) using the grid search method and uses the best-performing set for comparison with other models.

Table 1 Hyperparameters of Random Forest

Hyperparameters	Candidates
Tree number	{10, 50, 100, 200, 500}
Minimum sample	{2, 3, 4}
Maximum features	{2, 3, 4}

Table 2 Hyperparameters of Support Vector Machine

Hyperparameters	Candidates
Kernel function	{'linear', 'rbf', 'sigmoid'}
Regularization parameter	{0.1, 1, 10, 100}
Gamma	{0.0001, 0.001, 0.01, 0.1}
Epsilon	{0.01, 0.1, 1, 10}

(3) Particle Filter Models

The estimation model based on the SIR particle filtering algorithm and the modified Bourgoyne-Young drilling speed equation is described in Section 1, and the operational procedures for both post-drilling data analysis and real-time drilling speed prediction scenarios are shown in Fig. 4. In the particle filtering algorithm, theoretically, the more particles used, the more accurate the results, but an excessive number of particles can significantly increase computational costs. In this paper, using the K5 well dataset as an example, calculations were performed using an Intel(R) Core (TM) i7-8700 @3.20GHz processor, and comparisons of MRE and runtimes were made for different numbers of particles. The results are shown in Fig. 5. As seen from the figure, when the number of particles exceeds 2000, the error is already small enough, and the runtime increases significantly. Therefore, in the case analysis, this paper sets the number of particles to 2000.

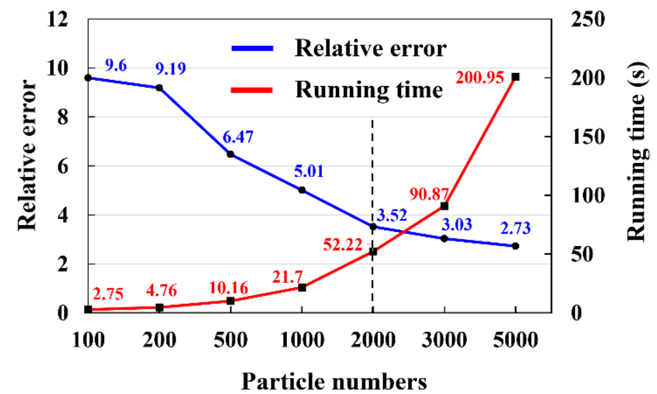

Fig. 5. Mean relative error and run time using different numbers of particles

4. RESULTS AND DISCUSSION

4.1 Model comparison

Fig. 6 to 13 respectively illustrate the performance of the physics-based model, SVM model, RF model, and particle filtering model on the 58-32 geothermal well dataset and the K5 well dataset. This includes regression drilling speed in the post-drilling data analysis scenario, predicted drilling speed in the real-time drilling speed prediction scenario, and comparisons with the actual drilling speed.

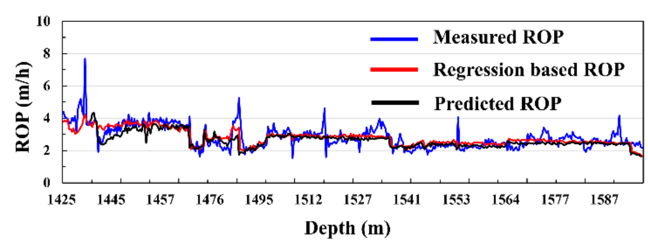

Fig. 6. Performance of physics-based model on well 58-32

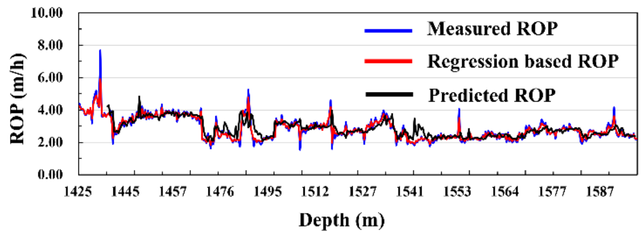

Fig. 7. Performance of RF model on well 58-32

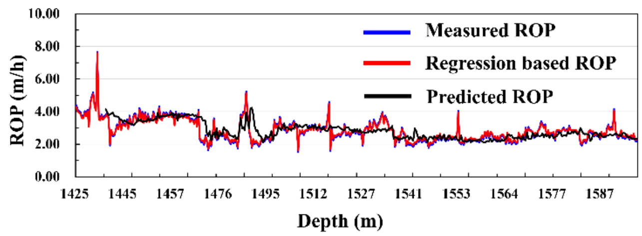

Fig. 8. Performance of SVM model on well 58-32

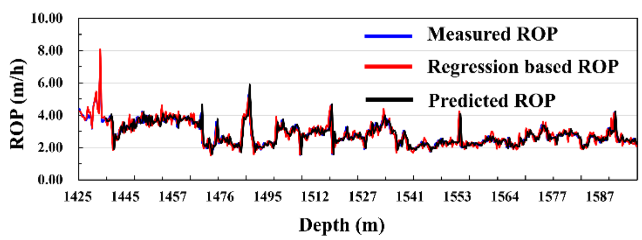

Fig. 9. Performance of particle filter model on well 58-32

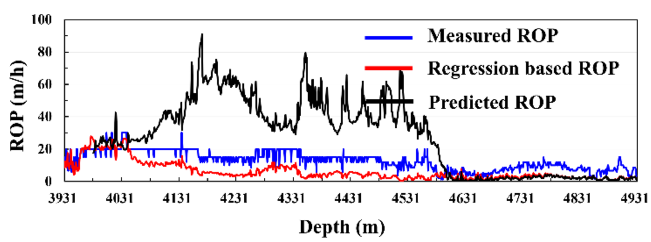

Fig. 10. Performance of physics-based model on well K5

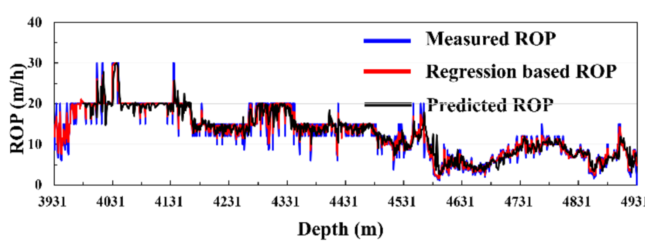

Fig. 11. Performance of RF model on well K5

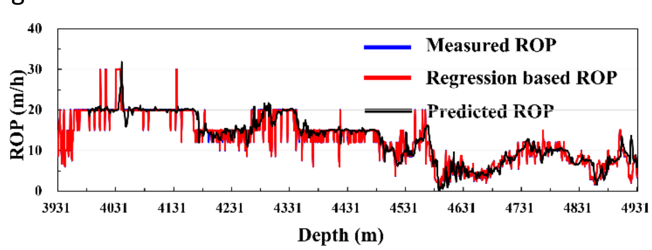

Fig. 12. Performance of SVM model on well K5

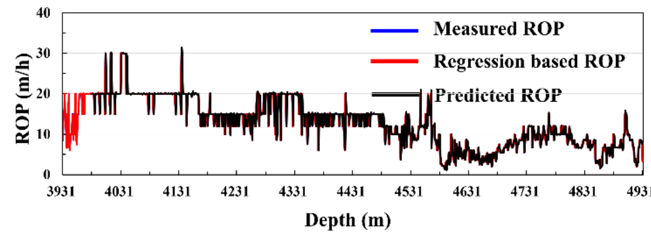

Fig. 13. Performance of particle filter model on well K5

Combining Fig. 6 to 13, it is evident that for the 58-32 geothermal well, in the scenario of using all data for post-drilling analysis, all models perform well, with the RF model, SVM model, and particle filtering model having similar performance, resulting in an average relative error of less than 5%. However, when using data from the already drilled sections for real-time prediction, the prediction errors for all four models increase to around 12%. The particle filtering model and the RF model demonstrate the best performance.

For the K5 well, which has a less regular drilling speed trend, the physics-based model performs poorly in both post-drilling analysis and real-time prediction scenarios, with errors reaching 58% and 158%, respectively. In contrast, the RF model, SVM model, and particle filtering model all show errors of less than 5% in the post-drilling analysis scenario. Among these, the SVM and particle filtering models perform the best. When performing real-time prediction, the RF model and SVM model have prediction errors of around 20%, while the particle filtering model's error is 15.58%, outperforming the other models.

It can be concluded that the particle filtering model, whether used for post-drilling analysis with all data or for real-time prediction based on the already drilled data, outperforms the traditional physics-based model and commonly used RF and SVM models. In post-drilling analysis, the RF and SVM models perform similarly to the particle filtering model, but to achieve optimal results, the RF and SVM models require hyperparameter tuning, resulting in higher computational and time costs. In real-time prediction, the RF and SVM models perform well on the 58-32 geothermal well data, which has a more regular speed trend. However, when faced with less regular data, such as that from the K5 well, their performance is less satisfactory. This is primarily due to differences in the drilling speed data:

Firstly, in the real-time prediction scenario, every prediction is based on models trained using all the previous data. To achieve the best predictive performance, an ideal situation would involve conducting hyperparameter tuning for every prediction.

This process incurs high computational and time costs, particularly when using the "real-time prediction" scenario, where such costs are considerable [20] [37]. For data with good regularity, global hyperparameters can describe the overall nonlinear relationships well, leading to satisfactory results in real-time prediction when dealing with local data. However, for data with less regularity, local data may differ significantly from global data, making it difficult for global hyperparameters to accurately describe the relationships between local data, resulting in larger prediction errors. Additionally, in practical applications, RF and SVM models may not even have access to all-well data for hyperparameter tuning, and they may only use neighboring well data. This raises concerns about model transferability and further increases errors.

Secondly, even with real-time hyperparameter tuning, when local data exhibits significant variations, the prediction errors are still considerable. Taking the RF model as an example, for data with good regularity, such as the 58-32 geothermal well, the error with global hyperparameter tuning is relatively small and comparable to the particle filtering algorithm. However, for data with less regularity, such as the K5 well, the maximum relative error reaches 6.7 times and 9.2 times that of the global hyperparameter tuning, which does not meet the requirements for precision. Based on the above analysis, this is because local data with less regularity exhibit significantly different trends, and even with real-time hyperparameter tuning, the errors in predictions are still substantial.

In contrast, regardless of the data's regularity, particle filtering can consistently produce good results. This is because the particle filtering algorithm treats the previously estimated state at time t-1 as the current state for predicting the drilling speed at time t. In this scenario, if the environmental conditions at time t have not changed significantly from those at time t-1, it is assumed that the state at time t is very similar to that at time t-1, and therefore the predicted drilling speed is accurate. Once the observed values for time t (real drilling speed, drilling pressure, rotation speed, etc.) are available, particle filtering can correct the estimated state using these observations, yielding an accurate current state. This process continues recursively. Therefore, particle filtering can accurately estimate the current state and can predict the next time's drilling speed accurately when the environment remains stable.

4.2 The Performance Evaluation of the Particle Filtering Model

The accuracy of the particle filtering model is not only assessed based on the error in estimating drilling speed data but also by evaluating the estimated depth, drilling pressure, rotary speed, and flow rate, which are the four observed data. Table 3 presents the errors in estimating depth, drilling speed, weight on bit, rotary speed, and flow rate on the K5 well dataset.

Table 3 MRE of the observed data of particle filter model.

	D	ROP	WOB	N	Q
MRE	1.71%	1.30%	1.45%	1.49%	1.53%
Max	24.07%	10.30%	15.46%	16.98%	18.71%
<10%	98.50%	99.90%	99.60%	99.60%	99.50%

According to Table 3, the errors between the estimated data and the real data using the particle filtering model are very small, with estimation errors less than 10% for approximately 99% of the data.

The particle filtering model not only provides real-time estimates for the drilling speed equation but also offers the uncertainty of the estimated parameters. Fig. 14 illustrates the coefficient of rotary speed and its 90% confidence interval in the drilling speed equation estimated using the particle filtering algorithm on the K5 well data. A 90% confidence interval means that there is a 90% chance that the coefficient falls within this interval, indicating the uncertainty in the estimation of the drilling speed equation. From the graph, it can be observed that the coefficient of rotary speed changes continuously with depth, reflecting the dynamic nature of the drilling process. Similar confidence intervals can be obtained for the other four coefficients in the drilling speed equation. By calculating the mean values, you can estimate the coefficients and use these estimates in the drilling speed equation. Furthermore, the uncertainty of the coefficient terms can be used to calculate the uncertainty of the observed data. Fig. 15 displays the actual rotary speed, the estimated rotary speed, and their 90% confidence intervals. It is evident that the estimated rotary speed closely matches the actual rotary speed, and the actual rotary speed falls within the confidence interval. This indirectly validates the accuracy of the estimates obtained using the particle filtering algorithm.

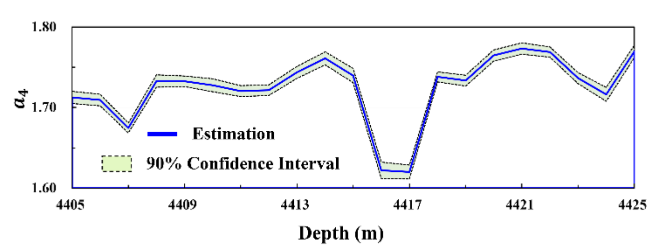

Fig. 14. Estimated coefficient and its 90% confidence interval of rotary speed in the drilling model using data from well K5

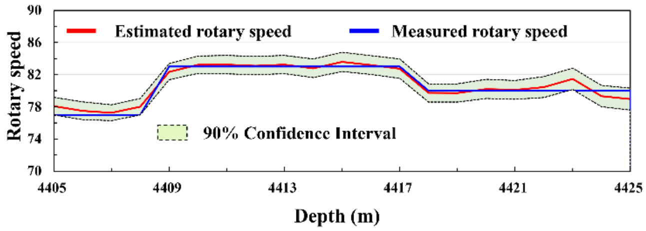

Fig. 15. Estimated rotary speed and its 90% confidence interval using data from well K5

5. CONCLUSIONS

This paper delves into the real-time prediction of drilling speed by combining the particle filtering algorithm with a physics-based drilling speed equation. Using the modified Bourgoyne-Young physics-based drilling speed equation as a foundation, the particle filtering method is introduced to estimate the coefficients of the drilling speed equation in real time. The performance of physics-based drilling speed models, intelligent drilling speed models, and particle filtering models in both post-drilling data analysis and real-time drilling speed prediction scenarios is compared. The paper provides the following conclusions:

(1) Physical models based on multivariate regression offer simple calculations and good interpretability but suffer from significant prediction errors.

(2) RF and SVM models are highly accurate when analyzing post-drilling data, with an average relative error of less than 5%. However, they exhibit errors exceeding 20% when predicting drilling speed in real time. Additionally, they face issues related to hyperparameter tuning and model transferability, making them unsuitable for real-time ROP prediction.

(3) The particle filtering drilling speed model exhibits an average relative error of less than 5% in the post-drilling data analysis scenario and a 15.58% error in real-time drilling speed prediction. The prediction results are accurate, making it suitable for both post-drilling analysis and real-time drilling speed prediction.

(4) The drilling speed equation based on the particle filtering algorithm combines accuracy, simplicity, good interpretability, and real-time updating. It can dynamically estimate the coefficients of the physics-based drilling speed equation and their uncertainties based on observable data such as drilling pressure and rotary speed. This model aligns with the dynamic, uncertain, and real-time nature of the drilling process and serves as a foundational tool for optimizing engineering parameters during drilling processes.

ACKNOWLEDGEMENT

The authors would like to thank the National Natural Science Foundation of China (No. 52125401) for the financial supports.

DECLARATION OF INTEREST STATEMENT

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. All authors read and approved the final manuscript.

REFERENCE

[1] LI, G., SONG, X., & TIAN, S. Intelligent Drilling Technology Research Status and Development Trends [J]. Petroleum Drilling Techniques, 2020, 48(01):1-8.

[2] Zhang, C., Song, X., Su, Y., & Li, G. (2022). Real-time prediction of rate of penetration by combining attention-based gated recurrent unit network and fully connected neural networks. Journal of Petroleum Science and Engineering, 213, 110396.

[3] Maurer, W. C. (1962). The" perfect-cleaning" theory of rotary drilling. Journal of Petroleum Technology, 14(11), 1270-1274.

[4] Warren, T. M. (1987). Penetration-rate performance of roller-cone bits. SPE Drilling Engineering, 2(01), 9-18.

[5] Bourgoyne Jr, A. T., & Young Jr, F. S. (1974). A multiple regression approach to optimal drilling and abnormal pressure detection. Society of Petroleum Engineers Journal, 14(04), 371-384.

[6] Hareland, G., & Rampersad, P. R. (1994, April). Drag-bit model including wear. In SPE Latin America and Caribbean Petroleum Engineering Conference (pp. SPE-26957). SPE.

[7] Detournay, E., Richard, T., & Shepherd, M. (2008). Drilling response of drag bits: Theory and experiment. International Journal of Rock Mechanics and Mining Sciences, 45(8), 1347-1360.

[8] Motahhari, H. R., Hareland, G., & James, J. A. (2010). Improved drilling efficiency technique using integrated PDM and PDC bit parameters. Journal of Canadian Petroleum Technology, 49(10), 45-52.

[9] Guan, Z., Chen, T., & Liao, H. (2021). Theory and technology of drilling engineering. Springer.

[10] Eren, T., & Ozbayoglu, M. E. (2010, January). Real time optimization of drilling parameters during drilling operations. In SPE Oil and Gas India Conference and Exhibition? (pp. SPE-129126). SPE.

[11] Hadi, F., Altaie, H., & AlKamil, E. (2019, November). Modeling rate of penetration using artificial intelligent system and multiple regression analysis. In Abu Dhabi International Petroleum Exhibition and Conference (p. D021S032R001). SPE.
[12] Zhang, C., Song, X., Liu, Z., Ma, B., Lv, Z., Su, Y., ... & Zhu, Z. (2023). Real-time and multi-objective optimization of rate-of-penetration using machine learning methods. Geoenergy Science and Engineering, 223, 211568.
[13] Barbosa, L. F. F., Nascimento, A., Mathias, M. H., & de Carvalho Jr, J. A. (2019). Machine learning methods applied to drilling rate of penetration prediction and optimization-A review. Journal of Petroleum Science and Engineering, 183, 106332.
[14] Moazzeni, A. R., & Khamehchi, E. (2020). Rain optimization algorithm (ROA): A new metaheuristic method for drilling optimization solutions. Journal of Petroleum Science and Engineering, 195, 107512.
[15] Ashrafi, S. B., Anemangely, M., Sabah, M., & Ameri, M. J. (2019). Application of hybrid artificial neural networks for predicting rate of penetration (ROP): A case study from Marun oil field. Journal of petroleum science and engineering, 175, 604-623.
[16] Zhao, Y., Noorbakhsh, A., Koopialipoor, M., Azizi, A., & Tahir, M. M. (2020). A new methodology for optimization and prediction of rate of penetration during drilling operations. Engineering with Computers, 36, 587-595.
[17] Hegde, C., Daigle, H., Millwater, H., & Gray, K. (2017). Analysis of rate of penetration (ROP) prediction in drilling using physics-based and data-driven models. Journal of petroleum science and Engineering, 159, 295-306.
[18] Soares, C., & Gray, K. (2019). Real-time predictive capabilities of analytical and machine learning rate of penetration (ROP) models. Journal of Petroleum Science and Engineering, 172, 934-959.
[19] Hegde, C., Soares, C., & Gray, K. (2018, September). Rate of penetration (ROP) modeling using hybrid models: deterministic and machine learning. In Unconventional Resources Technology Conference, Houston, Texas, 23-25 July 2018 (pp. 3220-3238). Society of Exploration Geophysicists, American Association of Petroleum Geologists, Society of Petroleum Engineers.
[20] Hegde, C., Daigle, H., & Gray, K. E. (2018). Performance comparison of algorithms for real-time rate-of-penetration optimization in drilling using data-driven models. Spe Journal, 23(05), 1706-1722.
[21] Hassan, A., Elkatatny, S., & Al-Majed, A. (2020). Coupling rate of penetration and mechanical specific energy to Improve the efficiency of drilling gas wells. Journal of Natural Gas Science and Engineering, 83, 103558.
[22] Negara, A., & Saad, B. (2020, January). Combining Insight from Physics-Based Models into Data-Driven Model for Predicting Drilling Rate of Penetration. In International Petroleum Technology Conference (p. D032S227R001). IPTC.
[23] Ahmed, O. S., Adeniran, A. A., & Samsuri, A. (2019). Computational intelligence based prediction of drilling rate of penetration: A comparative study. Journal of Petroleum Science and Engineering, 172, 1-12.
[24] Elkatatny, S. (2018, April). Rate of penetration prediction using self-adaptive differential evolution-artificial neural network. In SPE Kingdom of Saudi Arabia Annual Technical Symposium and Exhibition (pp. SPE-192186). SPE.
[25] Li, Q., Cao, Y. W., & Zhu, H. Y. (2021). Discussion on the lower limit of data validity for ROP prediction based on artificial intelligence. Drilling Engineering, 48(3), 21-30.
[26] LI, Y. (2020) Research on Parameters Optimization Using Machine Learning Algorithm in Deep Sea Oil Drilling Field (Doctoral dissertation, Beijing University of Posts and Telecommunications, China).
[27] Wallace, S. P., Hegde, C. M., & Gray, K. E. (2015, September). A system for real-time drilling performance optimization and automation based on statistical learning methods. In SPE middle East intelligent oil and gas conference and exhibition. OnePetro.
[28] HU, S., JING Z. Overview of particle filter algorithm [J]. Control and Decision, 2005, 20(004): 361-5,71
[29] Li, T. C., Fan, H. Q., & Sun, S. D. (2015). Particle filtering: Theory, approach, and application for multitarget tracking. Acta Automatica Sinica, 41(12), 1981-2002.
[30] Ambrus, A., Daireaux, B., Carlsen, L. A., Mihai, R. G., Karimi Balov, M., & Bergerud, R. (2020, August). Statistical determination of bit-rock interaction and drill string mechanics for automatic drilling optimization. In International Conference on Offshore Mechanics and Arctic Engineering (Vol. 84430, p. V011T11A027). American Society of Mechanical Engineers.
[31] Daireaux, B., Ambrus, A., Carlsen, L. A., Mihai, R., Gjerstad, K., & Balov, M. (2021, March). Development, Testing and Validation of an Adaptive Drilling Optimization System. In SPE/IADC Drilling Conference and Exhibition (p. D051S022R003). SPE.
[32] Thrun, S. (2002). Probabilistic robotics. Communications of the ACM, 45(3), 52-57.

[33] Doucet, A., & Johansen, A. M. (2009). A tutorial on particle filtering and smoothing: Fifteen years later. Handbook of nonlinear filtering, 12(656-704), 3.
[34] DING, J. (2012) Study on Performance Parameters Estimation of High Speed Train (Doctoral dissertation, Southwest Jiaotong University, China).
[35] Podgorney, R., McLennan, J., & Moore, J. (2018). Utah FORGE: Drilling Data for Student Competition (No. 1113). USDOE Geothermal Data Repository (United States); Idaho National Lab.(INL), Idaho Falls, ID (United States).
[36] Balamir, O., Rivas, E., Rickard, W. M., McLennan, J., Mann, M., & Moore, J. (2018, February). Utah FORGE reservoir: Drilling results of deep characterization and monitoring well 58-32. In Proc. 43rd Workshop on Geothermal Reservoir Engineering, Stanford University, Stanford, California.
[37] Echeverri Duque, S. (2019). Evaluation of machine learning for optimization and anomaly detection in offshore drilling operations (Master's thesis, University of Stavanger, Norway).