MEDWIN PUBLISHERS
Committed to Create Value for Researchers

Petroleum & Petrochemical Engineering Journal
ISSN: 2578-4846

Machine Learning-Based Prediction and Optimization of Mechanical Specific Energy in Drilling Operations: A Gas Field Case Study in Romania

Dragomir V*, Mhanna A and Avram L#

Drilling Engineering, Petroleum, Gas University of Ploiesti, Romania

***Corresponding author:** Viorel Dragomir, Drilling Engineering, Petroleum, Gas University of Ploiesti, Romania.

#Equally contributed towards this article.

Research Article

Volume 10 Issue 4

Received Date: October 13, 2026

Published Date: November 12, 2026

DOI: 10.23880/ppej-16000421

Abstract

Drilling optimization in complex geological environments requires efficient control of operational parameters and energy consumption. This study presents a data-driven workflow for Mechanical Specific Energy (MSE) prediction and optimization using Artificial Neural Network (ANN), Random Forest (RF), and Support Vector Regression (SVR) models applied to drilling data from a gas field in southwestern Romania. Operational parameters, including weight on bit (WOB), rotary speed (RPM), flow rate, torque, and rate of penetration (ROP), were integrated to capture the nonlinear relationship between drilling conditions and MSE behavior. The results demonstrated that all developed models successfully predicted MSE under variable drilling conditions. Among the evaluated models, the Random Forest model achieved the highest prediction accuracy, while the ANN model showed strong generalization capability. Furthermore, an optimal operational window associated with reduced MSE, improved drilling efficiency, and stable drilling conditions was identified. The proposed workflow provides a practical framework for drilling optimization and supports the implementation of intelligent real-time drilling advisory systems for operational decision-making

Keywords: Drilling Optimization; Mechanical Specific Energy (MSE); Drilling Efficiency; Romania Gas Field; Data-Driven Optimization

Abbreviations

MSE: Mechanical Specific Energy; ANN: Artificial Neural Network; RF: Random Forest; SVR: Support Vector Regression; WOB: Weight on Bit; RPM: Rotary Speed; ROP: Rate of Penetration; RSS: Rotary Steerable System; FLW: Flow Rate; TORQ: Torque; SPP: Standpipe Pressure.

Introduction

Optimizing drilling has become an essential objective in modern petroleum engineering due to the growing technical and economic challenges of oil and gas well construction. In complex geological environments, drilling performance is strongly influenced by interactions among operational parameters, including Weight on Bit (WOB), rotational speed (RPM), flow rate, and drilling fluid properties. Improper parameter selection may lead to reduced drilling efficiency, excessive energy consumption, vibrations, poor hole cleaning, and stuck pipe incidents [1].

Mechanical Specific Energy (MSE) is widely used as an indicator of drilling efficiency because it reflects the amount of energy required to remove a unit volume of rock. Elevated

CC BY

MSE values are generally associated with inefficient drilling conditions and drilling dysfunctions. Consequently, MSE analysis has become a valuable tool for identifying non-productive drilling regimes and improving operational performance [2].

In recent years, Machine Learning techniques have been increasingly applied in drilling engineering due to their capability to model nonlinear relationships between drilling parameters and operational performance. Among the most commonly used methods are Artificial Neural Network (ANN), Random Forest (RF), and Support Vector Regression (SVR), which have demonstrated strong prediction capability in drilling optimization applications [3,4].

The gas field, located within the Getic Depression of Romania, represents a challenging environment for drilling operations due to lithological variability and tectonic complexity. Alternating marly-clayey and sandy formations significantly influence drilling stability and operational efficiency.

The main objective of this study is to develop a data-driven methodology for predicting and optimizing Mechanical Specific Energy (MSE) using Artificial Neural Network (ANN), Random Forest (RF), and Support Vector Regression (SVR) models applied to real drilling data from a southwestern gas field from Romania. The study aims to identify the optimal operational window for drilling parameters, including WOB, RPM, and flow rate, to improve drilling efficiency, reduce MSE, and minimize operational risks. Additionally, the developed workflow offers operational recommendations for maintaining stable, energy-efficient drilling conditions in complex geological environments.

Methodology

Field Data and Operational Dataset

The present study was conducted using operational drilling data collected from a gas well drilled in a southwestern gas field from Romania, located within the Getic Depression of Romania. The analyzed drilling interval is characterized by complex geological conditions, including alternating marly-clayey, sandy, and conglomeratic formations with significant lithological variability and tectonic compartmentalization. These geological conditions strongly influence drilling behavior, operational stability, and drilling efficiency, making the field suitable for advanced drilling optimization studies.

The analyzed drilling section was drilled using a Rotary Steerable System (RSS) push the bit combined with an underreamer assembly operating in continuous-rotation mode. The selected drilling configuration was designed to improve directional control, optimize hole enlargement performance, and maintain operational stability under complex drilling conditions.

The operational dataset was obtained from real-time drilling acquisition systems and included continuously recorded surface drilling parameters and calculated drilling indicators. The collected dataset consisted of Depth, Weight on Bit (WOB), Rotational Speed (RPM), Flow Rate (FLW), Torque (TORQ), Rate of Penetration (ROP), standpipe pressure (SPP), and Mechanical Specific Energy (MSE). These parameters were selected because they directly influence drilling efficiency, mechanical energy transfer, and operational stability.

Data Preprocessing and Filtering

The performance of the Artificial Neural Network (ANN) model depends strongly on the quality and consistency of the training dataset. Therefore, the raw drilling data collected from field operations were subjected to a rigorous preprocessing and filtering procedure before model development.

Initially, missing values (NaN), non-numerical records, and inconsistent operational points were removed from the dataset. In addition, drilling records with zero values for critical operational parameters such as Weight on Bit (WOB), Rotational Speed (RPM), and Flow Rate were excluded, as these conditions were not representative of actual drilling operations.

Subsequently, a Z-score analysis was applied to detect outliers. For each operational parameter, including ROP, WOB, RPM, Torque, Flow Rate, SPP, and Mechanical Specific Energy (MSE), the Z-score was calculated, and values exceeding the ±3 threshold were removed from the dataset. This filtering stage significantly reduced the influence of abnormal operational points and improved dataset consistency. Similar statistical filtering procedures are commonly used in drilling data analytics and machine learning applications to improve prediction reliability [4].

In the second filtering stage, the DBSCAN clustering algorithm was applied to identify isolated operational points that did not follow the dataset's dominant operational trend as illustrated in Figure 1. The clustering analysis was performed within the WOB–MSE operational space, and points classified as noise were removed from the dataset. Previous studies have shown that appropriate preprocessing and clustering techniques significantly improve the stability and predictive capability of data-driven drilling optimization models [5].

Copyright© Dragomir V, et al.

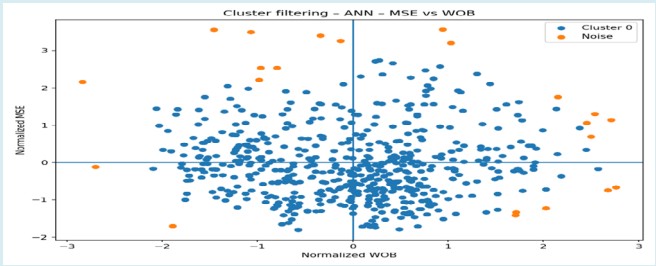


Figure 1: Cluster filtering (DBSCAN) for MSE function of WOB.

Following the preprocessing workflow, the total number of operational observations was reduced from 602 to 513, resulting in a more stable, coherent, and representative dataset suitable for ANN model training and MSE prediction.

Mechanical Specific Energy Calculation

Mechanical Specific Energy (MSE) was used in this study as the primary indicator for evaluating drilling efficiency and the transfer of mechanical energy during drilling operations. MSE represents the amount of mechanical energy required to remove a unit volume of rock and is widely applied in drilling engineering to identify efficient and inefficient drilling regimes. Elevated MSE values generally indicate inefficient drilling conditions, such as excessive friction, vibrations, poor hole cleaning, or drilling dysfunctions.

The conventional MSE relationship used in this study is expressed as:

$$MSE = \frac{WOB}{A} + \frac{120\pi \cdot T \cdot RPM}{A \cdot ROP} \quad (1)$$

where:
= Weight on Bit
= Torque
= Rotational Speed
= Rate of Penetration
= Bit Cross-Sectional Area

The first term represents the axial energy contribution, while the second term represents the rotational energy contribution during drilling operations.

Continuous MSE monitoring was performed throughout the analyzed drilling interval to identify operational regimes associated with efficient drilling performance and stable drilling conditions. Variations in MSE values were correlated with drilling parameters to identify operational windows characterized by reduced energy consumption and improved drilling efficiency. The variation of MSE with depth for the analyzed drilling section is presented in Figure 2, while Figure 3 illustrates the relationship between MSE and Rate of Penetration (ROP), providing additional insight into drilling efficiency and mechanical energy transfer during drilling operations.

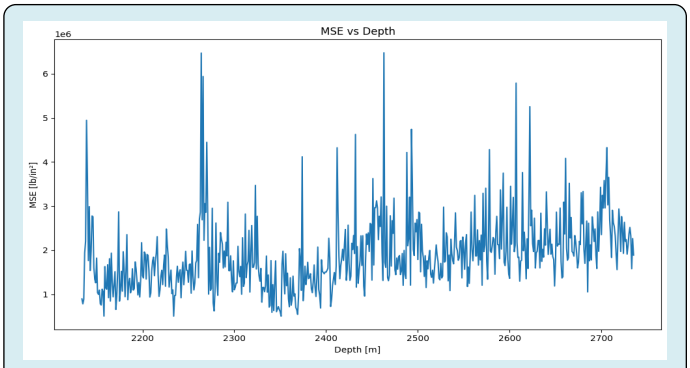


Figure 2: MSE variation versus drilling depth for the analyzed 6×7" drilling section.

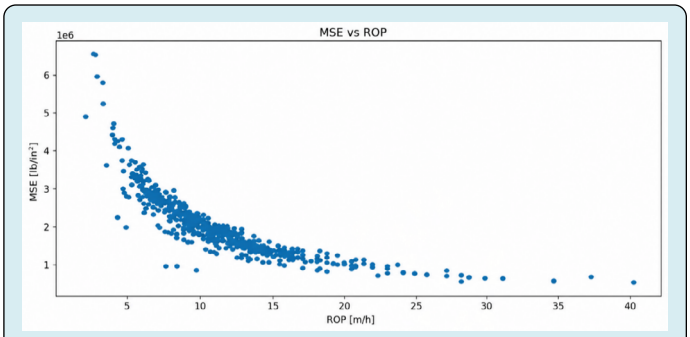


Figure 3: Relationship between MSE and Rate of Penetration (ROP).

Machine Learning Models

Three different Machine Learning models were developed and evaluated for MSE prediction and drilling optimization: Artificial Neural Network (ANN), Random Forest (RF), and Support Vector Regression (SVR). The selected input variables included Depth, WOB, RPM, Flow Rate, Torque, ROP, and SPP, while the target output variable for all developed models was MSE. The machine learning workflow was designed to evaluate the capability of different predictive models to reproduce the nonlinear relationship between drilling parameters and MSE behavior under complex drilling conditions. Similar machine learning workflows for drilling optimization and MSE prediction have been successfully applied in recent drilling engineering studies [2,3].

Model performance was evaluated using statistical indicators including correlation coefficient (R), coefficient of determination (R^2), Mean Absolute Percentage Error (MAPE), and Root Mean Square Error (RMSE). These statistical indicators are commonly used in machine learning-based drilling optimization and operational performance prediction studies [6].

Copyright© Dragomir V, et al.

Artificial Neural Network (ANN)

The ANN model was developed to capture the nonlinear interaction between operational drilling parameters and MSE behavior. The final ANN architecture consisted of seven input neurons, one hidden layer with 64 neurons, and one output neuron corresponding to the predicted MSE value. The hidden layer utilized the ReLU activation function, while a linear activation function was applied to the output layer for continuous prediction. The model was trained using the Adam optimization algorithm with an 80/20 training-testing split ratio. During training, the ANN model learned the complex operational relationships between drilling parameters and MSE by minimizing the prediction error between actual and predicted outputs.

The developed ANN model demonstrated strong predictive capability and good generalization performance under variable drilling conditions. Figure 4 presents the ANN architecture developed for MSE prediction using the selected drilling parameters.

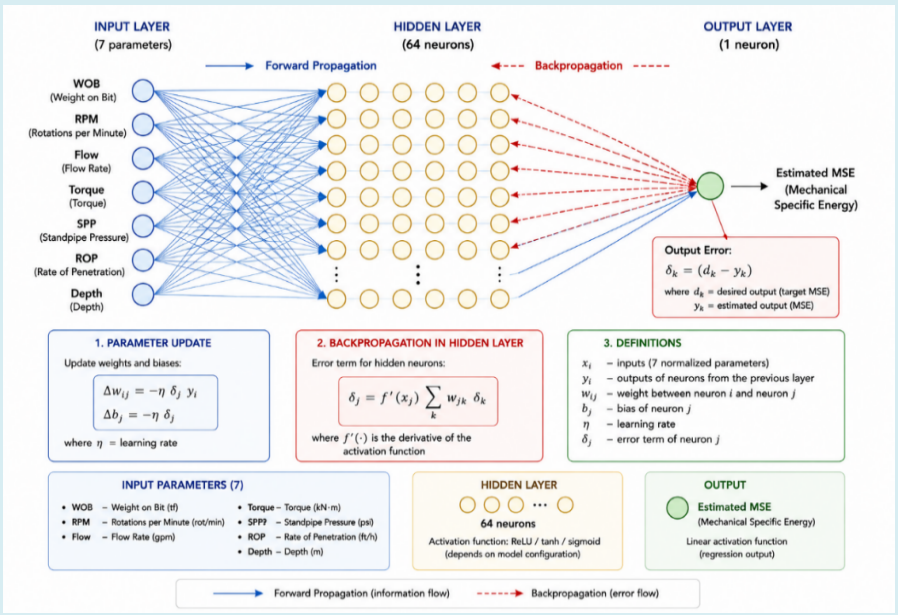


Figure 4: Forward and Backward Propagation Scheme of the ANN Model Used for MSE.

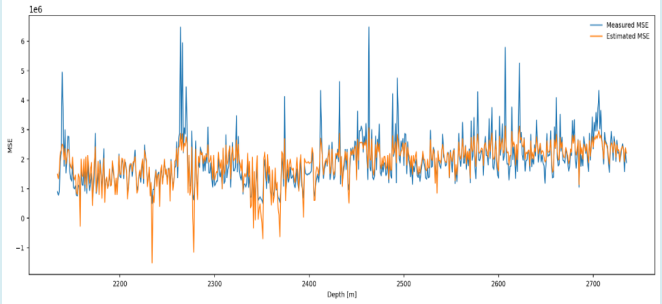


Figure 5: Variation of measured and estimated MSE versus depth using the ANN model for the 6×7" section.

Figure 5 presents the variation of measured and ANN-estimated MSE as a function of depth for the 6×7" drilling section. A good agreement can be observed between the estimated and measured values, indicating that the ANN model successfully reproduces the general MSE trend throughout the analyzed interval. Local differences occur mainly in zones characterized by abrupt MSE variations, where the model tends to smooth extreme values. Overall, the results confirm the capability of the ANN model to accurately estimate MSE and demonstrate its potential for real-time monitoring and drilling optimization applications.

Following the ANN prediction analysis, an empirical ANN-based relationship was established to describe the nonlinear dependency between operational drilling parameters and MSE behavior:

$$MSE = 5.24\times10^{4}\cdot WOB_n + 2.00\times10^{5}\cdot RPM_n - 1.50\times10^{4}\cdot Flow_n + 1.14\times10^{5}\cdot Torque_n - 1.59\times10^{4}\cdot SPP_n - 1.42\times10^{6}\cdot ROP_n + 1.28\times10^{5}\cdot Depth_n + 1.59\times10^{6} \quad (2)$$

where:
= Weight on Bit
= Rotational Speed
= Standpipe pressure
= Rate of Penetration

Copyright© Dragomir V, et al.

The developed ANN correlation provides a practical framework for future real-time drilling optimization and intelligent drilling advisory applications.

Random Forest (RF)

The Random Forest (RF) model was implemented as an ensemble-based regression technique for MSE prediction under complex drilling conditions. RF algorithms are widely used in drilling optimization studies due to their ability to capture nonlinear relationships among operational parameters while maintaining robustness to operational noise and overfitting [7].

The RF model operates by generating multiple decision trees using random subsets of the training dataset and combining their outputs to improve prediction stability and accuracy. In addition, RF models enable evaluation of feature importance, making them suitable for identifying the operational parameters that strongly influence MSE behavior.

The results demonstrated that the RF model achieved the highest prediction accuracy and the lowest statistical error among the evaluated machine learning models, indicating an excellent capability to reproduce complex operational drilling behavior. The comparison between measured and RF-predicted MSE values as a function of depth for the analyzed drilling interval is presented in Figure 6.

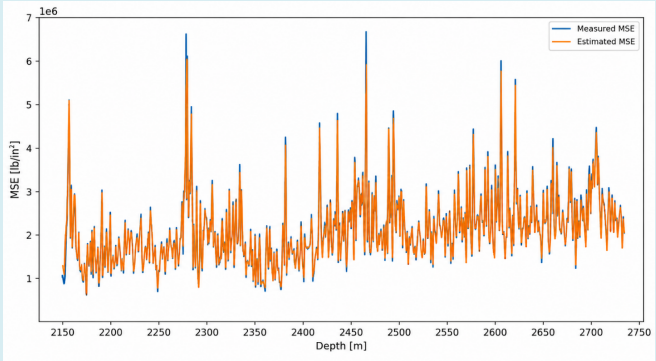


Figure 6: Variation of measured and estimated MSE versus depth using the Random Forest model for the 6×7" section.

Support Vector Regression (SVR)

The Support Vector Regression (SVR) model was implemented to evaluate nonlinear MSE prediction capability under variable drilling conditions. SVR models are particularly suitable for regression problems involving nonlinear operational trends because they utilize kernel-based transformations to model complex relationships between operational variables [8]. In this study, the SVR model demonstrated competitive prediction capability and satisfactory generalization performance for MSE estimation. The model successfully reproduced the dominant operational trends observed during drilling operations; however, it exhibited slightly higher prediction errors than the RF and ANN models, particularly during intervals characterized by increased operational instability. Despite these limitations, the SVR model confirmed the applicability of kernel-based machine learning approaches for drilling optimization and MSE prediction applications. The comparison between measured and SVR-predicted MSE values as a function of depth for the analyzed drilling interval is presented in Figure 7.

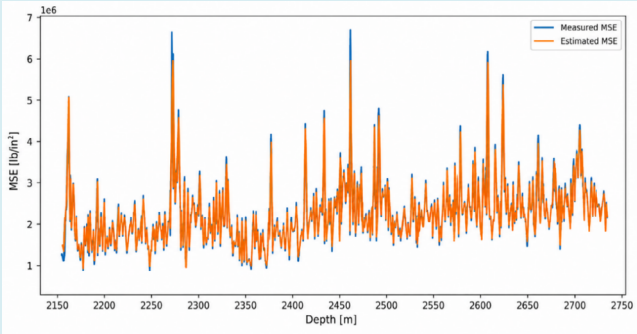


Figure 7: Variation of measured and estimated MSE versus depth using the Support Vector Regression model for the 6×7" section.

Results and Discussion

Machine Learning Model Performance

The developed machine learning models demonstrated strong capability for predicting MSE under variable drilling conditions. Model performance was evaluated using correlation coefficient (R), coefficient of determination (R^2), Mean Absolute Percentage Error (MAPE), and Root Mean Square Error (RMSE). The models were trained using filtered and normalized drilling data, including WOB, RPM, Flow Rate, Torque, SPP, ROP, and Depth as input variables. The objective of the developed models was to reproduce the variation of MSE and identify the most suitable approach for drilling optimization and real-time operational monitoring. The comparative plots presented in Figure 8 illustrate the variation of measured and estimated MSE with depth for the ANN, SVR, and Random Forest models. All three machine learning approaches successfully reproduced the general MSE trend throughout the analyzed drilling interval.

Among all the approaches investigated, the Random Forest model provided the best agreement between measured and estimated MSE values, showing excellent capability in reproducing both global trends and local variations. The

Copyright© Dragomir V, et al.

RF model demonstrated superior adaptability to nonlinear drilling behavior and better reproduction of extreme MSE fluctuations compared with the ANN and SVR models. Table 1 summarizes the performance indicators obtained for the developed machine learning models. The ANN model achieved overall values of approximately R≈0.98, RMSE ≈ 1.75×10^5lb/in^2, and AAPE ≈ 3.02%, confirming good predictive capability and stable model behavior.

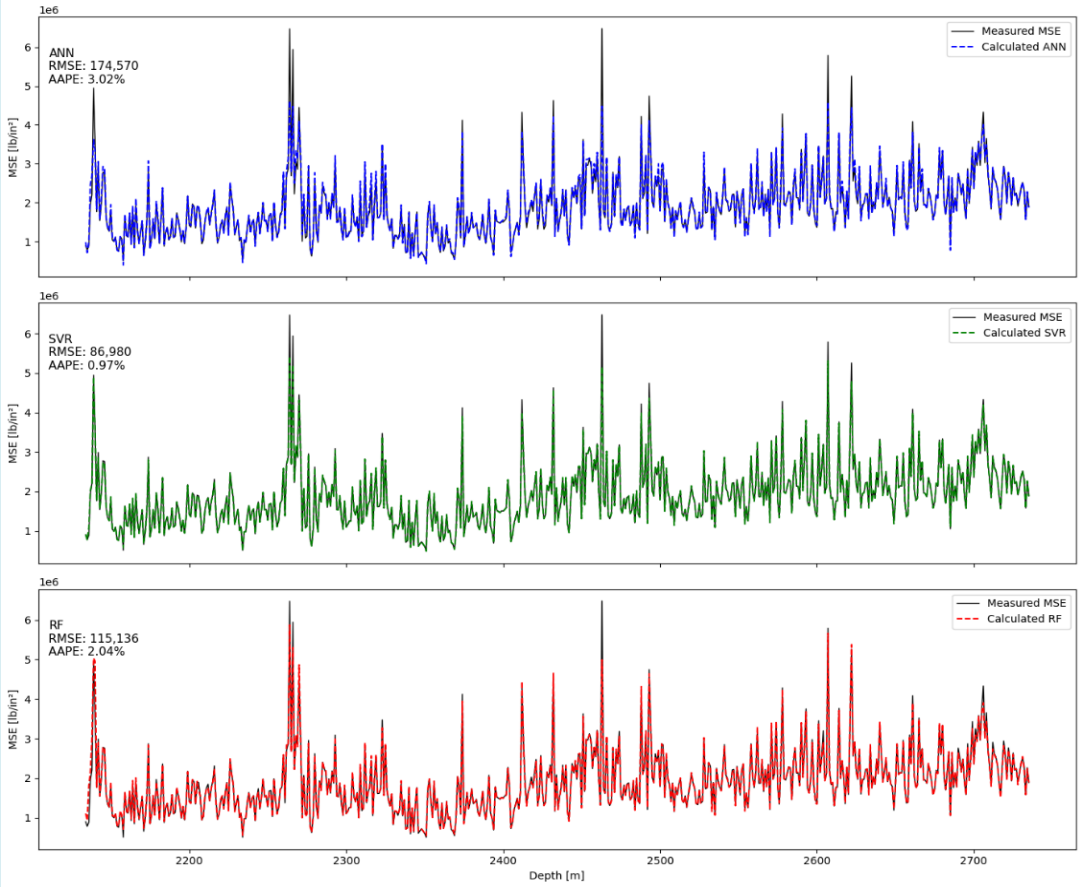


Figure 8: Comparative variation of measured and estimated MSE with depth using ANN, SVR, and RF models.

Model	Set	R	RMSE [lb/in^2]	AAPE (%)
ANN	Training	1	153114	2.92
	Testing	1	241710	3.41
	All data	1	174570	3.02
SVR	Training	1	107373	3.24
	Testing	0.9	224596	4.48
	All data	1	139106	3.49
Random Forest	Training	1	67000	1.42
	Testing	1	189000	4.5
	All data	1	115136	2.04

Table 1: Performance comparison of ANN, SVR, and RF models.

Copyright© Dragomir V, et al.

The SVR model produced slightly lower prediction errors, with RMSE ≈ lb/in^2, and AAPE ≈ 3.49%, while maintaining high correlation coefficients of approximately R≈0.97. The obtained results indicate strong prediction capability and good generalization performance for nonlinear MSE behavior. The obtained results confirm the applicability of machine learning techniques for real-time drilling optimization and operational parameter management under complex geological conditions [9,10]. The best overall performance was obtained using the Random Forest model, which achieved R≈0.99, RMSE ≈ 1.15×10^5lb/in^2, and AAPE ≈ 2.04% for the complete dataset. These results indicate that the RF model provided the highest prediction accuracy and the lowest overall prediction error among all investigated approaches.

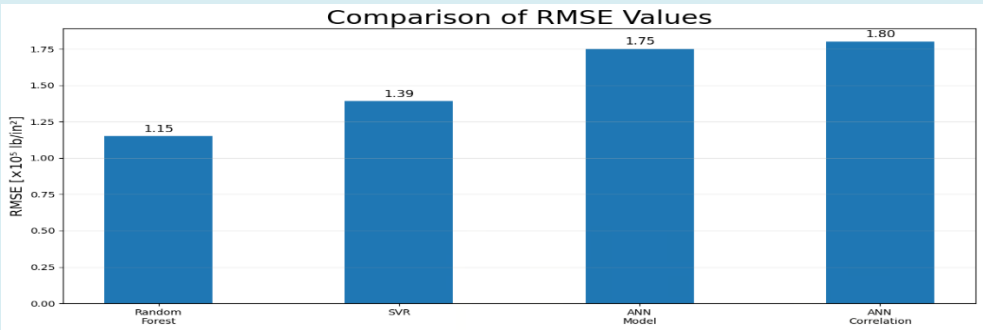


Figure 9: Comparison of RMSE Values for the Developed Machine Learning Models.

The RMSE comparison presented in Figure 9 indicates that the Random Forest model achieved the lowest prediction error, followed by the SVR and ANN models. The higher RMSE values obtained for the ANN-based approaches suggest reduced capability in reproducing extreme local MSE variations compared with the RF and SVR models.

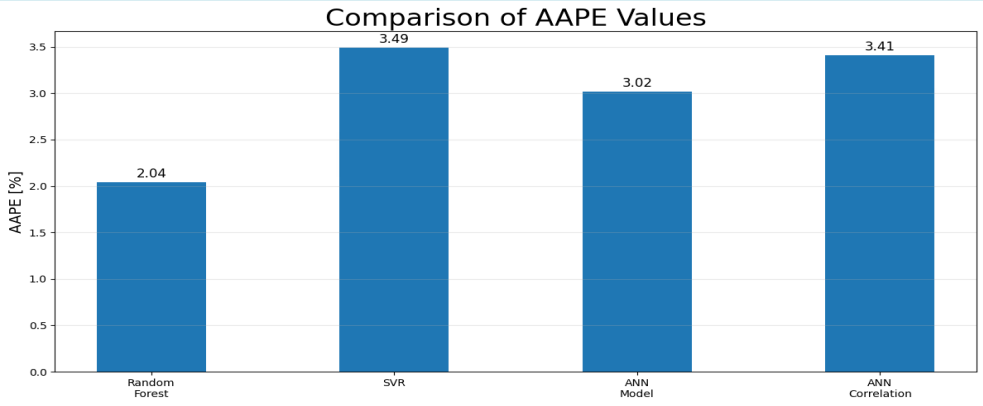


Figure 10: Comparison of AAPE Values for the Developed Machine Learning Models.

Similarly, the comparison of AAPE values confirms the superior performance of the Random Forest model, which achieved the lowest percentage prediction error. The SVR model showed intermediate accuracy, while the ANN model produced slightly higher errors. The comparison of AAPE values obtained for the developed machine learning models is presented in Figure 10. Nevertheless, all developed machine learning models demonstrated strong predictive capability and acceptable performance for drilling engineering applications.

MSE Behavior and Operational Window Analysis

The analysis of operational drilling parameters and MSE behavior allowed the identification of three distinct operational drilling regimes: conservative, optimal, and aggressive. The relationship between operational parameters and drilling performance is illustrated in Figure 11.

Copyright© Dragomir V, et al.

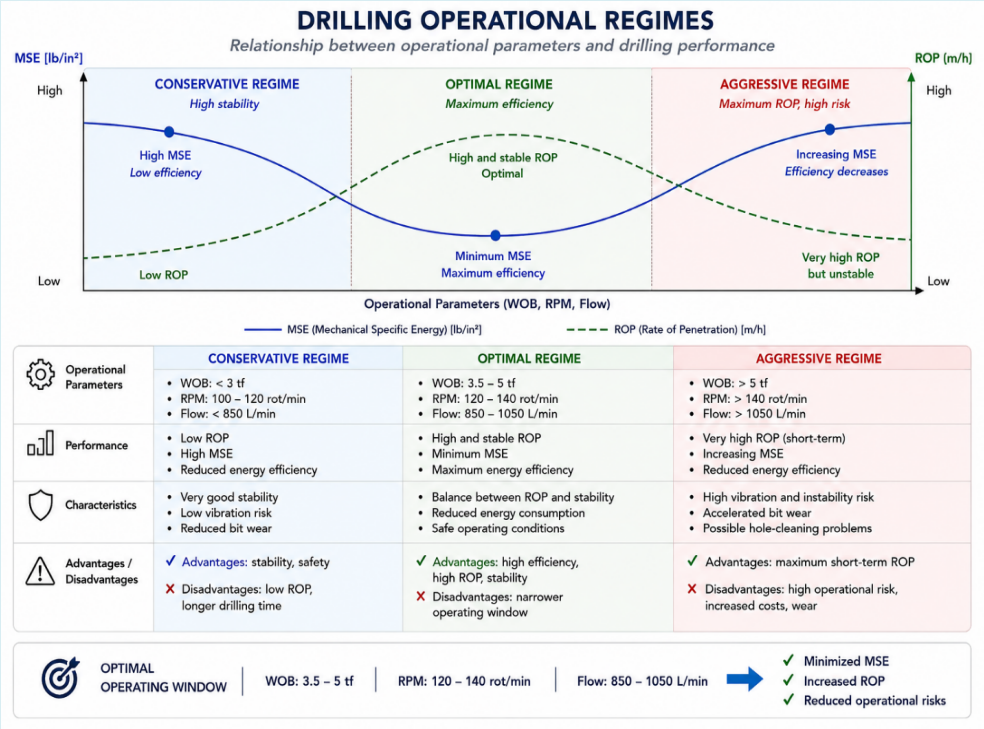


Figure 11: Operational drilling regimes and optimized operational window for the analyzed 6×7" drilling section.

The conservative operational regime was characterized by WOB values below 3 tf, rotational speeds below 120 rot/min, and flow rates lower than 850lpm. Under these conditions, the drilling process exhibited high operational stability and reduced vibration tendency; however, elevated MSE values and reduced ROP indicated inefficient energy transfer and lower drilling efficiency. The aggressive operational regime corresponded to WOB values exceeding 5 tf, rotational speeds above 140 rot/min, and flow rates higher than 1050 lpm. Under these operational conditions, short-term ROP values increased significantly; however, MSE values also increased, indicating reduced drilling efficiency and excessive mechanical loading. Furthermore, this regime was associated with increased vibration levels, operational instability, accelerated tool wear, and a higher probability of hole-cleaning problems. The optimal operational regime was identified within the interval of 3.5–5 tf, WOB, 120–140 rot/min RPM, and 850–1050 lpm flow rate. Within this operational window, the drilling process demonstrated minimum MSE values, stable and elevated ROP behavior, improved energy efficiency, and reduced operational instability. The analysis also demonstrated that maintaining drilling parameters within the identified operational window significantly reduces the probability of drilling dysfunctions such as stick-slip, excessive vibrations, poor hole cleaning, and unstable torque behavior. Consequently, the proposed operational window provides practical recommendations for improving drilling performance and maintaining stable drilling conditions under complex geological environments. The integration of MSE analysis with machine learning prediction models enabled the development of a data-driven optimization workflow capable of supporting future intelligent real-time drilling advisory systems and operational decision-making applications [11,12].

Conclusions

This study presented the development and evaluation of machine learning models for the prediction and optimization of Mechanical Specific Energy (MSE) using real drilling operational data from the analyzed 6×7" drilling section. Three machine learning approaches, namely Artificial Neural Network (ANN), Support Vector Regression (SVR), and Random Forest (RF), were developed using drilling parameters including WOB, RPM, Flow Rate, Torque, SPP, ROP, and Depth.

The results obtained demonstrated that all investigated models were capable of reproducing the nonlinear relationship between drilling parameters and MSE with good accuracy. Among the analyzed approaches, the Random Forest model provided the best overall prediction performance, achieving the lowest RMSE and AAPE values and the highest correlation coefficients.

Copyright© Dragomir V, et al.

The integration of MSE analysis with machine learning techniques enabled the identification of an optimal operational drilling window characterized by WOB values of approximately 3.5-5 tf, RPM values between 120–140 rot/min, and flow rates between 850–1050 lpm. Within this operating range, the drilling process exhibited reduced MSE, improved ROP behavior, higher energy efficiency, and increased operational stability.

The results obtained in this study confirm the potential of machine learning methods for real-time drilling optimization, operational monitoring, and intelligent drilling advisory applications. Future work may include the integration of additional drilling dysfunction indicators and real-time automated optimization systems for complex drilling environments.

References

1. Aadnøy BS, Looyeh R (2011) Petroleum Rock Mechanics: Drilling Operations and Well Design, Elsevier, Oxford, USA.

2. Bourgoyne AT, Millheim KK, Chenevert ME, Young FS (1986) Applied Drilling Engineering, SPE Textbook Series, Society of Petroleum Engineers, Richardson, Texas, USA.

3. Teale R (1965) The Concept of Specific Energy in Rock Drilling, International Journal of Rock Mechanics and Mining Sciences 2(1): 57-73.

4. Samuel R (2017) Drilling Engineering Optimization and Automation, Elsevier, Amsterdam, USA.

5. Mohaghegh S, Ameri S, Gaskari R (2017) Application of Artificial Intelligence and Data Mining in Drilling Optimization, Journal of Natural Gas Science and Engineering 45: 569-584.

6. Ahmed ST, Chen H, Wang Y (2021) Support Vector Machine Applications in Petroleum Engineering and Drilling Optimization, Journal of Petroleum Science and Engineering 196: 107640.

7. Hegde S, Gray D, Millwater M (2022) Application of Machine Learning Techniques for Real-Time Drilling Optimization, Journal of Petroleum Science and Engineering 208: 109276.

8. Al-Yami A, Schubert M, Al-Shehri M (2023) Machine Learning Applications for Drilling Optimization and Data Analytics, Energies 16(5).

9. Zhang H, Liu Y, Chen M (2023) Artificial Intelligence Applications for Real-Time Drilling Optimization and Decision Support Systems. Journal of Energy Resources Technology 145(4).

10. Chen X, Wang R, Yang J, Gao D, Li G, et al. (2025) A New Intelligent Optimization Method for Drilling Parameters of Extended Reach Wells Based on Mechanical Specific Energy and Machine Learning, Petroleum Research.

11. Barboza TR (2025) Drilling Optimization by Means of Decision Matrices and Mechanical Specific Energy Analysis, Applied Sciences.

12. Wang Y, Li H, Zhang X (2025) Data-Driven Drilling Optimization Using Machine Learning Techniques, Petroleum Exploration and Development 52(1).

Copyright© Dragomir V, et al.

YOUR RESEARCH DESERVES GLOBAL VISIBILITY

• Publish • Discover • Share • Connect

MEDICAL	CLINICAL	LIFESCIENCES	PHARMA
90+ OPEN ACCESS JOURNALS		400+ RESEARCH TOPICS	

SUBMIT YOUR RESEARCH
Submit your manuscript through our online submission system.

REPRINTS
If you require a reprint of any article, please contact us.
info@medwinpublishers.com

BECOME A REVIEWER
Share your expertise and contribute to scholarly publishing.

SCAN TO EXPLORE MEDWIN JOURNALS

• Journals • Submit Manuscript • Reviewer Program • Article Tracking

Our Journals indexed in

Embase® ELSEVIER | PubMed National Library of Medicine NLM | Crossref | Google Scholar | INDEX COPERNICUS INTERNATIONAL | ISSN INTERNATIONAL STANDARD SERIAL NUMBER INTERNATIONAL CENTRE | ROAD DIRECTORY OF OPEN ACCESS SCHOLARLY RESOURCES | EuroPub | Scilit | WorldCat® | Academic Resource Index ResearchBib | ADVANCED SCIENCE INDEX | INTERNATIONAL Scientific Indexing | ICMJE INTERNATIONAL COMMITTEE of MEDICAL JOURNAL EDITORS

This article published by **Medwin Publishers,** 28475 Witherspoon Dr, Novi, MI 48377, United States, +1-248-247-6042, WhatsApp +1 (248) 275-1133, info@medwinpublishers.com

© 2026, by the Authors. The articles published from this journal are distributed to the public under "**Creative Commons Attribution License**" (https://creativecommons.org/licenses/by/4.0/). Therefore, upon proper citation of the original work, all the articles can be used without any restriction or can be distributed in any medium in any form.