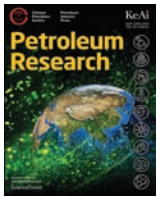
Petroleum Research 11 (2026) 895–911

Contents lists available at ScienceDirect

Petroleum Research

journal homepage: http://www.keaipublishing.com/en/journals/petroleum-research/



Full Length Article

Data-driven real-time prediction of cutting's transport index in directional wellbores



Mohammed Al-Shargabi [a], Shadfar Davoodi [a,b,*], David A. Wood [c,**], Mohammad Mehrad [a], Mohammed Al-Rubaii [d]

[a] School of Earth Sciences & Engineering, Tomsk Polytechnic University, Lenin Avenue, Tomsk, Russia
[b] Artificial Intelligence Center, Skolkovo Institute of Science and Technology, Bolshoy Boulevard 30, bld. 1, 121205, Moscow, Russia
[c] DWA Energy Limited, Lincoln, United Kingdom
[d] Saudi Aramco, Dhahran, Saudi Arabia

ARTICLE INFO

Article history:
Received 19 April 2025
Received in revised form
24 November 2025
Accepted 25 November 2025
Available online 28 November 2025

Keywords:
Drilling cutting's transport index
Hybrid machine learning
Radial basis function neural networks
Real-time monitoring
Deviated/horizontal hole cleaning
Two-stage optimization

ABSTRACT

The drilling of non-vertical wellbores requires effective real-time monitoring techniques to improve drilling efficiency. This study innovatively develops nine robust hybrid machine learning (HML) models to develop a real-time surrogate model for the drilling cuttings transport index (DCTI). The DCTI, a comprehensive indicator of hole-cleaning efficiency, is typically calculated with a 2–4 h delay due to its dependency on laboratory-measured fluid properties. Our hybrid ML models overcome this limitation by accurately emulating the DCTI value using only real-time, surface-recorded drilling data. A 6818-record dataset was collected from three oil and gas fields (Eastern Province of Saudi Arabia), comprising three offshore and two onshore wells. Following a meticulous dataset preprocessing, a non-dominated sorting genetic algorithm (NSGA-II) was applied with a multi-layer perceptron neural network (MLPNN), identifying seven key features influencing DCTI – inclination angle (ANG), rate of penetration, flow in, depth, torque, stand-pipe pressure (SPP), and revolutions per minute. Extreme learning machine (ELM), MLPNN, and radial basis function neural networks (RBFNN) were hybridized with particle swarm optimization (PSO) and a growth optimizer in two stages: (1) to determine the optimal model structure, and (2) to fine-tune the model hyperparameters to precisely estimate DCTI based on the selected features. The HML models achieved better estimation performance and generalizability than the standalone models. Achieving the lowest root mean square error (RMSE = 0.00344), the RBFNN-PSO model exhibited the best DCTI estimation performance. Analytical scoring, robustness, and Monte Carlo simulation confirmed this model's superior performance, robustness, and high generalization. Shapley additive explanations (SHAP) feature influence analysis revealed that ANG exerted the most influence on the RBFNN-PSO estimations, while SPP had the least impact. During drilling operations, the proposed data-driven approach practically enables drilling crews to diagnose and monitor hole-cleaning performance in real time and improve the rate of penetration. Furthermore, SHAP analysis transforms the model into a prescriptive tool, identifying the primary real-time parameters affecting hole cleaning. This enables drillers to make targeted operational decisions, such as adjusting the rate of penetration or flow rate, to mitigate developing issues.

© 2025 Chinese Petroleum Society. Publishing services by Elsevier B.V. on behalf of KeAi Communications Co. Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

* Corresponding author. School of Earth Sciences & Engineering, Tomsk Polytechnic University, Lenin Avenue, Tomsk, Russia.
** Corresponding author.
E-mail addresses: shadfardav@gmail.com, davoodis@hw.tpu.ru (S. Davoodi), dw@dwasolutions.com (D.A. Wood).
Peer review under the responsibility of Editorial Office of Petroleum Research.

https://doi.org/10.1016/j.ptlrs.2025.11.010
2096-2495/© 2025 Chinese Petroleum Society. Publishing services by Elsevier B.V. on behalf of KeAi Communications Co. Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

Growing global energy demand, along with declining conventional resources (Wang et al., 2024), has necessitated the use of highly deviated and horizontal wells for hydrocarbon production, especially in unconventional reservoirs (Al-Rubaii et al., 2023b; Al-Shargabi et al., 2024). These advanced drilling approaches extend contact with the formation and improve productivity (Al-Shargabi et al., 2024). The market for such technologies was estimated at USD 3.5 billion in 2024 and is expected to grow to USD 5.8 billion by 2033, with a compound annual growth rate of 6.1% from 2026 (Deloitte, 2025; VMP, 2025). However, as wellbore complexity has increased, hole cleaning (HC) has emerged as one of the most pressing operational challenges (Al-Rubaii et al., 2023b; Elsayed Salem and Nasr El-Din, 2006; Khosravanian and Aadnøy, 2021).

Accumulated cuttings in highly deviated and horizontal wellbores sections often form beds at the annular bottom, impeding circulation, causing pump blockages while circulating and stuck pipe events. Such problems elevate non-productive time (NPT), compromise safety, and inflate costs (Al-Rubaii et al., 2023b; Costa et al., 2023). Consequently, HC efficiency is fundamental to overall drilling performance, as effective cuttings removal is essential for safety, operational continuity, and economic viability (Yeu et al., 2019). This complexity arises from the dynamic interplay of three primary factors: drilling fluid (DF) properties, operational and mechanical parameters, and wellbore geometry (Al-Rubaii et al., 2023d; Nazari et al., 2010).

DF is the primary medium for cuttings transport; adequate density prevents particle settling and fluid loss (Yu et al., 2007), while optimized rheology ensures effective suspension and conveyance (Adari et al., 2000). Operational and mechanical parameters further complicate HC optimization (Al-Shargabi et al., 2024). Weight on bit (WOB) has a dual effect: higher loading produces finer cuttings but increases cuttings bed risk (Al-Rubaii et al., 2023b; Elsayed Salem and Nasr El-Din, 2006; Khosravanian and Aadnøy, 2021). Flow rate (GPM) is critical—insufficient flow impairs evacuation, excessive flow risks wellbore-wall erosion or turbulence (Khosravanian and Aadnøy, 2021), while drill-string rotation (RPM) and rate of penetration (ROP) influence cuttings distribution (Al-Shargabi et al., 2024; Zhu et al., 2023). Wellbore inclination also affects transport mechanisms in an inclination-dependent manner (Ahmed, 2020; Zhu et al., 2023). In wellbores with angles ranging from 0 to 45°, laminar flow dominates, and HC efficiency is governed by fluid rheology and stress distribution (Mahmoud et al., 2020). In 45–65°, cuttings sliding disrupts flow predictability and reduces efficiency (Iqbal et al., 2023), with the 30–70° range having the highest risk of stuck pipe, necessitating flow rates 2–3 times higher than vertical wells (Piroozian et al., 2012; Yeo et al., 2021). In 65–90° sections, turbulent flow enhances transport but remains sensitive to rheological properties affecting lift capacity (Piroozian et al., 2012; Yeo et al., 2021).

The multifaceted challenges mentioned highlight the need for robust HC monitoring and management in highly deviated and horizontal wells (Al-Rubaii et al., 2023b). Real-time evaluation of HC performance using actual drilling parameters, coupled with dynamic optimization of cuttings transport, is essential to mitigate downhole incidents, reduce NPT, and lower operational costs (Al-Rubaii et al., 2023b; Al-Shargabi et al., 2024). Currently, three primary approaches have been used to estimate HC performance in non-vertical wellbores: empirical indexing methods, machine learning (ML), and deep learning (DL) models (Al-Shargabi et al., 2024; Davoodi et al., 2025).

Since 1961, several empirical models have been utilized with the initial development of cutting concentration in annulus (CCA_1) model by Newitt et al. (1961), which estimated the accumulation of drilling cuttings in the annulus. Mitchell (1992) later expanded this model to provide empirical evidence for a concentric concentration model, accounting for pre- and post-connection circulation. Sifferman et al. (1974) developed a cutting carrying index (CCI) based on transport ratio (TR) and GPM to evaluate the fluid velocity needed to lift cuttings with the actual annular velocity (V_{ann}) in the well. Subsequently, several models have been developed to provide additional insights for HC performance. The most widely used of these are transport index (TI) (Luo et al., 1992, 1994), HC ratio (HCR) (Rasi and Member, 1994), and carrying capacity ratio (CCR) (Robinson and Morgan, 2004). For instance, CCR assesses cuttings lifting in vertical wellbores using V_{ann}, DF density, and a power law constant (k) (Robinson and Morgan, 2004).

Martins et al. (2011) enhanced wellbore monitoring by integrating cuttings flow meters with hydraulic models, improving HC assessment. Jimmy et al. (2022) developed a lifting coefficient (β^1) to assess cutting's lifting ability in wellbores. Despite these efforts, these models have primarily been tested in laboratory settings, limiting their applicability and reliability and necessitating further validation under diverse field conditions, especially in highly deviated and horizontal wellbore sections. Additionally, current models rely on non-real-time drilling parameters, such as DF properties including plastic viscosity (PV), yield point (YP), k, and flow behavior index (n), restricting their use in real-time drilling operations, which depend on time-consuming measurements that take 2–4 h per day and are limited to specific well sections. Al-Rubai et al. (2023a, 2023c) modified three empirical models (CCR, HCR, and TI) to generalize and apply them across various complex profiles, including horizontal wells, using only field data records. However, despite model validations from field datasets, these modified models rely on non-real-time drilling parameters, which limit their applications in real-time HC performance.

In the realm of ML and DL, few studies have been utilized to estimate HC related issues (Al-Shargabi et al., 2024; Davoodi et al., 2025). Ozbayoglu et al. (2002) pioneered the use of artificial neural networks (ANN) to forecast cuttings-bed height (CBH), incorporating input variables such as feed cuttings concentration, fluid densities, ROP, and wellbore geometry, achieving an average absolute percentage error (AAPE) below 10%. Rooki et al. (2014) expanded and applied ANN and multiple linear regression (MLR) to predict cuttings concentration in foam DF based on experimental foam properties, with ANN yielding the best performance (mean average percentage error (MAPE) of 3.2%). Li et al. (2014) used a radial basis function (RBF) network to estimate particle settling behavior based on fluid and particle properties, reporting a root mean square error (RMSE) of 0.0272. Ulker and Sorgun (2016) evaluated various ML models, including k-nearest neighbor (k-NN), support vector regression (SVR), linear regression, and ANN, to estimate the dimensionless cuttings-bed area ratio from DF velocity, ROP, wellbore inclination, and RPM, showing that ANN achieved the best results with RMSE of 0.2592. Al-Azani et al. (2019) and Alkinani et al. (2019) further validated ANN's reliability using extensive experimental datasets and highlighted how DF properties influence HC efficiency.

Han et al. (2022) applied ML (ANN, SVR) and DL models (recurrent neural networks (RNN) and long short-term memory (LSTM)) to predict CBH using the inputs GPM, cuttings diameter, ROP, eccentricity, and wellbore diameter. Their ANN model performed the best, achieving an RMSE of 1.069%. Chowdhury and Hovda (2023, 2022) utilized fuzzy logic (FL) to estimate cuttings concentration using 509 experimental data points and multiple drilling parameters. Their FL model achieved an RMSE of 3.41%, which was improved upon by a hybrid FL and genetic algorithm (GA) applied to 702 data points, reducing RMSE to 1.5722%. However, current ML and DL studies only rely on measurements or

laboratory-based analysis that are unavailable in real-time drilling operations. They are specifically tailored toward predicting intermediate parameters like cuttings concentration or CBH rather than providing holistic evaluations of HC conditions or the likelihood of operational problems such as stuck pipe.

The literature review indicates that there has been minimal research focused on developing ML models that can estimate HC performance in complex well profiles. So far, only one study has suggested an approach for estimating HC performance. Wood (2023) interpolated data points along the curved trends experimentally in horizontal sections to generate additional points for ML models. Subsequently, seven empirical H% relationships were developed, using various influencing parameters on HC such as RPM, DF density, and V_{ann}. Among the ML models for predicting these empirical H% relationships, XGBoost demonstrated the best performance, achieving an RMSE of 0.76%. However, the study relies on a small dataset derived from experimentally-interpolated information, and only two real-time inputs were employed in each of the developed empirical relationships. This limitation makes them unsuitable for broader application across diverse, complex wells and restricts their generalizability in real-time scenarios.

To the best of the authors' knowledge, no data-driven models have been proposed to date that can provide a real-time estimation of HC performance using field data from complex well profiles. Hence, there is a pressing need for a novel and robust real-time approach to estimate HC performance in highly deviated and horizontal wellbores. This advancement would ultimately aid a drilling team's ability to manage and diagnose HC more efficiently. All empirical models developed so far have been based on experimental data or only field data, without HC experimental verification (Al-Rubaii et al., 2023a, 2023b, 2023c; Luo et al., 1992, 1994; Martins et al., 2011; Mitchell, 1992; Newitt et al., 1961; Rasi and Member, 1994; Robinson and Morgan, 2004; Sifferman et al., 1974). Likewise, ML and DL models have been used to predict HC issues without providing insights into HC performance itself. They have relied solely on experimental data records without any field validation (Al-Azani et al., 2019; Alkinani et al., 2019; Chowdhury and Hovda, 2022, 2023; Han et al., 2022; Li et al., 2014; Ozbayoglu et al., 2002; Rooki et al., 2014; Ulker and Sorgun, 2016; Wood, 2023).

To bridge this gap, this study presents a novel, data-driven framework that enables real-time, accurate, and robust estimation of HC efficiency in non-vertical wellbores using only readily available field parameters, eliminating reliance on time-consuming laboratory analyses. The primary innovation of this study is that, for the first time, it develops a high-fidelity, real-time surrogate model for the empirical DCTI calculation. This approach explicitly tackles the time-lag problem inherent in the original model rather than attempting to validate the DCTI concept itself. The objective is to create a diagnostic tool that can provide an instantaneous estimate of the DCTI, effectively bypassing its dependency on lagged laboratory data. The standalone models evaluated are radial basis function neural network (RBFNN), multilayer perceptron neural network (MLPNN), and extreme learning machine (ELM). The hybrid ML (HML) models enhance each of the standalone models by integrating them with two powerful metaheuristic optimization algorithms: particle swarm optimization (PSO) and growth optimizer (GO). These HML models are designed to automatically tune critical hyperparameters, thereby improving prediction accuracy, convergence speed, and generalization under diverse drilling conditions.

A central estimated output of the developed models is the drilling cuttings transport index (DCTI (Al-Rubaii et al., 2023b);), a metric that evaluates HC efficiency in drilling. Although the empirical DCTI model offers reliable HC condition estimates, certain inputs, such as PV, YP, and viscometer readings (R_3, R_6, R_{100}, R_{200}, R_{300}, R_{600}), require laboratory measurement, causing a 2–4 h delay daily. This time lag restricts real-time monitoring and impairs timely decision-making, which is essential for sustaining optimal drilling performance. Therefore, the study develops standalone ML and HML models using drilling variables recorded in real time, including RPM, inclination angle, bit aggressiveness (BA), ROP, and GPM.

The developed models are evaluated using real field data from five non-vertical wells, including three offshore and two onshore wells drilled in Saudi Arabia. These wells represent different drilling conditions: (1) smooth ROP resulting in efficient HC conditions; (2) low ROP due to cuttings accumulation during drilling; and (3) instances of stuck drilling strings due to insufficient cuttings transport. As a result, the complete dataset encompasses a wide range of HC conditions. The proposed ML and HML models are deployed to generate real-time predictions of HC conditions by estimating DCTI values without requiring laboratory measurements. Consequently, such DCTI estimates can be used in real time to make timely decisions that improve wellbore HC conditions. Furthermore, by employing the Shapley Additive Explanations (SHAP) feature-influence, interpretability technique, the model transitions from a black-box predictor to a diagnostic tool, revealing the primary drilling parameters influencing HC and providing actionable guidance for operators.

2. Data collection and description

2.1. The drilling cutting transport index (DCTI)

Al-Rubaii et al. (2023b) modified the DCTI based on TI developed by Luo et al. (Luo et al., 1992, 1994) to generate the highest level of semi-automations in real-time HC estimation. Hence, this work addresses the primary deficiency in the semi-automated DCTI model by assessing the HC circumstances and devising a real-time ML and HML model to overcome the limitations of the modified DCTI. The modified DCTI is expressed as Eq. (1) (Al-Rubaii et al., 2023b):

$$DCTI = MW \cdot (RF_m + LC_m) \cdot AF_m \quad (1)$$

Where MW denotes the DF density (lb/cf), AF_m denotes the modified angle factor, LC_m denotes the modified original lifting capacity, and RF_m denotes the average rheology factor.

The modified DCTI model is a significant advancement in drilling operations, as it incorporates the variables hydraulic velocities, rheological characteristics, flow regime, cuttings properties, equivalent circulation density, lifting capacity, and angle factors. It provides a comprehensive assessment of downhole conditions and serves as a reliable indicator of HC conditions. A DCTI value over 1 indicates satisfactory cleaning, while a DCTI below 1 indicates insufficient cleaning (Al-Rubaii et al., 2023b; Luo et al., 1992, 1994; Unegbu, 2010). For more information about the DCTI refer to (Al-Rubaii et al., 2023b). The next subsection describes how the dataset was compiled to determine the output DCTI.

It is crucial to clarify that the DCTI values serving as the target variable for this study were not directly measured but were calculated retrospectively using Eq. (1). The requisite rheological properties (e.g., PV, YP, R3-R600) were obtained from routine mud checks performed at the rig site at approximately 2-4-h intervals. These lab measurements were meticulously time-synchronized with the real-time drilling data records to construct the comprehensive dataset. Therefore, the objective of the ML models developed herein is to function as a high-fidelity, real-time surrogate model for this empirical DCTI calculation, effectively bypassing its

dependency on lagged laboratory data.

2.2. Standalone ML and HML input parameters description

Building on the results of previous studies (Al-Rubaii et al., 2023d; Al-Shargabi et al., 2024), this study employs seven real-time parameters to forecast the DCTI, enabling the assessment and monitoring of HC conditions using RBFNN, MLPNN, and ELM optimized with PSO and GO. GO and PSO are applied to the standalone ML models to optimize the hyperparameters of the HML models. These optimizers avoid local optima, enhancing generalization, improving accuracy, and providing a robust solution for forecasting DCTI, thereby reducing training time and minimizing errors (Davoodi et al., 2023). The input parameters and factor include depth (ft); rate of penetration (ROP, ft/h); drilling fluid flow out (FLo, %); drilling fluid flow in (FLin, gal/min); hook load (HL, klbf); revolution per minute (RPM, rpm) for rotary-steerable system (RSS); stand pipe pressure (SPP, psi); torque (Trq, kft.lbf); weight on bit (WOB, klbf); angle inclinations (ANG, deg); and bit aggressiveness factor (BA). These parameters can be used to evaluate and estimate the HC conditions in real-time by monitoring the drilling operation and adjusting the HC process accordingly. For example, higher FLin can help remove cuttings more efficiently, but can also increase mud pressure and lead to lost circulation. Monitoring FLin in real time can help optimize HC (Al-Azani et al., 2019). Similarly, higher ROP can indicate better HC, but it can also lead to instability and lost circulation. Monitoring ROP in real time can help identify optimum drilling conditions (Brenjkar and Biniaz Delijani, 2022). Table 1 displays a statistical summary of the dataset variables, including the predicted target (dependent) variable DCTI.

All data variables were collected through drilling rig sensors in real time, installed on the offshore and onshore drilling rigs involved in the wells drilled in three oil and gas fields located in the Eastern Province of Saudi Arabia. The recorded data from five well sites was transmitted to a remote operations center for monitoring and collation. The recorded data is divided into three hole-cleaning classes: (1) effective HC due to the high ROP; (2) poor HC due to cuttings accumulations during drilling, causing a decrease in ROP; and (3) stuck drilling string due to insufficient cutting transport.

The five wells studied all involve non-vertical wellbore sections varying from low to high deviations, including horizontal components. All drilling mechanical inputs used were recorded and filtered every 5 s and stored in a central system that can be accessed through structured query language (SQL) queries. Additionally, the data for these drilling sections were selected based on matching static characteristics among the wells and field experience related to the identification of HC conditions. These wells possess the following uniform features: borehole diameter (12.25 inches and 8.5 inches), DF type (Innovert DF and oil-based DF), type of drill bit (PDC), strength of drill bit (M3), drilling conducted with a motorized RSS, and main reservoir targets (Hanifa and Khafji Formation). Moreover, the wells were all drilled for production purposes and followed the same drilling plan.

The compiled input variables were used to determine DCTI (target variable) for promptly estimating the HC conditions in near-real time. The data from an offshore well, which was used to develop the DCTI model (Al-Rubaii et al., 2023b), was employed for comparison with the results of the five wells evaluated in this study.

3. Methodology

Fig. 1 shows a schematic diagram depicting the sequential procedure for forecasting DCTI with the standalone ML and HML models. The process involves compiling historical field data to measure the DCTI and influencing independent variables. The data are initially analyzed to identify and remove outliers by evaluating the relationship between mechanical specific energy (MSE) and drilling ROP. The separation ratios of the filtered data records are examined for dividing the dataset into training and testing categories. A genetic algorithm (GA) with non-dominated sorting (NSGA-II), combined with MLPN, is employed to select a subset of the available input features that have the most significant impact on DCTI. Appropriate settings are applied to predictive ML and optimization algorithms to develop both standalone and hybrid forms on training data and DCTI predictive models. These trained models are evaluated on testing data to assess their generalizability. Scoring and robustness techniques are applied to select the best model. The best model is then further assessed with a Monte Carlo simulation to determine the level of uncertainty. The best model is also analyzed with Shapley additive explanations (SHAP) to determine the relative importance of the input features.

3.1. Predictive algorithms

Each ML algorithm is associated with its own strengths and weaknesses, which influence its prediction performance when applied to a specific dataset. In this regard, the ML algorithms considered, ELM, MLPNN, and RBFNN, are compared based on various criteria in Table 2.

3.1.1. The extreme learning machine (ELM)

Huang developed the ELM in 2004 to overcome conventional algorithms' limitations like low learning rates and slow training (Huang et al., 2004; Liang et al., 2006). The model employs input, hidden, and output layers, assigning weights randomly across them. Output weights are computed in a single step using the Moore-Penrose generalized inverse operation. ELM is a user-friendly algorithm that efficiently analyzes network parameters, reduces human involvement, and offers superior performance and compatibility with nonlinear activation and kernel functions. It eliminates iterative backpropagation training in multi-layer perceptron networks but tends to lack sufficient precision and accuracy (Rashidi et al., 2021; Yeom and Kwak, 2017). Furthermore, ELMs, unlike networks with multiple hidden layers, have a single level of abstraction, making them more flexible through optimized algorithms for better performance (Zeng et al., 2017).

Table 1
Statistical summary of the variable distributions, each involving 6818 data records compiled from the five non-vertical offshore and onshore wells used to predict DCTI.

Data variables	Maximum	Minimum	Mean	Skewness	Kurtosis
Depth, ft	12820.00	5950.00	9397.46	−0.01	−1.20
ROP, ft/h	267.86	0.55	93.38	0.74	0.34
FLo, %	43.40	22.96	34.21	−0.01	−0.04
FLin, gal/min	787.20	272.50	565.32	−5.24	55.87
HL, klbf	204.62	154.55	174.50	0.57	−0.61
RPM, rpm	103.40	42.30	81.50	−0.26	−0.57
SPP, psi	4418.56	1059.33	3933.96	−1.60	4.86
Trq, kft.lbf	15.62	0.01	8.77	−0.41	−0.68
WOB, klbf	38.6	0.00	26.7	−0.91	−0.23
ANG, DEG	88.70	22.90	75.50	−1.40	0.50
BA	1.84	0.001	1.03	−0.41	−0.68
DCTI	1.90	0.26	0.57	1.61	1.28

Note: the abbreviations used in this table are ROP = rate of penetration, Flo = drilling fluid flow out, FLin = drilling fluid flow in, HL = hook load, RPM = revolution per minute, SPP = stand-pipe pressure, Trq = torque, WOB = weight on bit, ANG = angle inclination, BA = bit aggressiveness factor, and DCTI = drilling cutting's transport index.

3.1.2. The multilayer perceptron neural network (MLPNN)

Artificial neural networks (ANNs) are configured in analogous ways to neurons in the human brain and are designed to uncover complex nonlinear relationships between variables (Ali, 1994). Key parameters for building ANN models include data preparation, input variable selection, choice of neural network type and structure, transfer function, and training algorithm (Maimon and Rokach, 2005). The MLPNN is a popular neural network structure due to its ability to estimate functions with reasonable accuracy, provided there are enough hidden units. It is ideal for modeling relationships between multiple input variables and output variables (Bishop and Nasrabadi, 2006). An MLPNN model was used to estimate ROP in a study well (Maimon and Rokach, 2005). MLPNNs' layered architecture enables them to model complex data relationships and patterns, making them adaptable to nonlinear and non-parametric problems, as nodes are connected through adjustable weights (Awad and Khanna, 2015). It starts with an input layer where each neuron represents an input-data feature. The data then moves through hidden layers, where each neuron computes the weighted sum of its inputs and applies an activation function to introduce nonlinearity (Hornik et al., 1989). The output layer generates the final predictions (LeCun et al., 2015).

3.1.3. The radial basis function neural network (RBFNN)

Moody and Darken (1989) developed the RBFNN as an ANN alternative to address scattered data point approximation problems. The RBFNN excels in tackling non-linear problems, demonstrating enhanced training efficiency, robust generalization abilities, and a streamlined design methodology (Chandra et al., 2020; Yu et al., 2011). RBFNNs and k-NN models share a similar framework but differ in their implementation. RBFNNs calculate Euclidean distance between test cases and neuron centers, assign weights, multiply output, and pass the modified data through output layers to generate predictions (Rezaei et al., 2022). Nevertheless, the application of RBFNNs may be influenced by specific data-point values, potentially reducing their prediction efficiency compared to MLP (Xie et al., 2011).

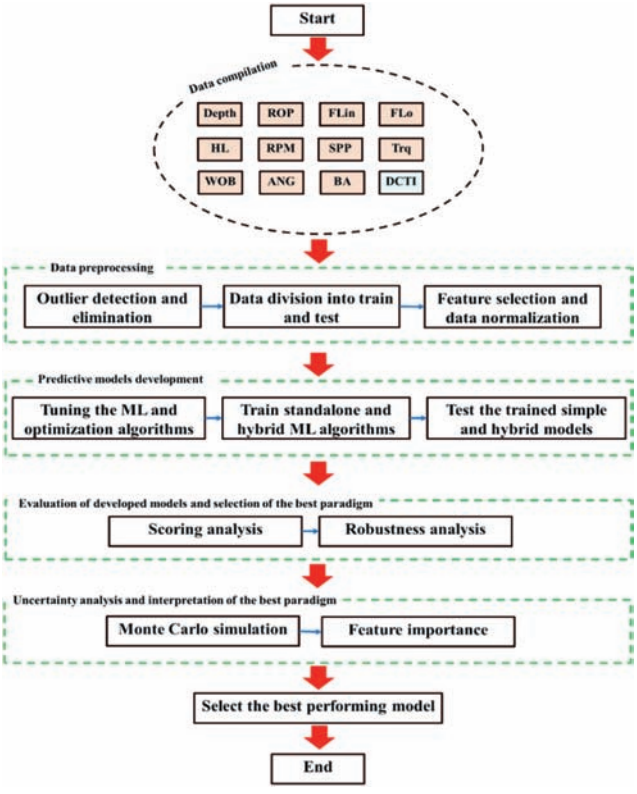


Fig. 1. A schematic diagram depicting the sequential procedure for forecasting the standalone ML and HML of DCTI.

3.2. Optimization algorithms

3.2.1. Particle swarm optimization (PSO)

PSO is a stochastic optimization technique inspired by animal swarming behaviors (Eberhart and Kennedy, 1995), which generates a group of possible solutions based on collections of "particles" that progress towards the most optimal solution (Onwunalu et al., 2009). The position vector $x_i(t)$ and velocity vector $v_i(t)$ are determined by individual experience and collective knowledge of the swarm's accurate positions (Sharma and Onwubolu, 2009). The effectiveness and simplicity of PSO implementation (Eq. (2)) depend on the selected algorithm control parameter values, which are influenced by different neighborhood topologies (Lin and Hong, 2007; Onwunalu et al., 2009).

$$v_{id}(t+1) = wv_{id}(t) + c_1 r_1 \left(P_{best,id}(t) - x_{iid}(t)\right) + c_2 r_2 \left(G_{best,id}(t) - x_{id}(t)\right), d = 1 \quad (2)$$

The PSO algorithm employs three distinct components, namely inertia (wv_{id}), cognitive (c_1), and social (c_2) constants, to update the velocity and position of particles (Onwunalu et al., 2009). The velocity update equation effectively balances exploration, determined by a weighting factor w, and exploitation, influenced by cognitive and social factors (Onwunalu et al., 2009). Cognitive terms are used to denote the individual's best position ($P_{best,id}$) of a particle, whereas social terms are employed to represent its social interaction (t), which guides it towards the collective best-known (global best) position ($G_{best,id}$) of the swarm. Random coefficients r and c are used to vary weights applied to these terms, resulting in a stochastic behavior. The subsequent particle position (Eq. (3)) is determined by the revised velocity (Lin and Hong, 2007; Pedersen and Chipperfield, 2010).

$$x_{id}(t+1) = x_{id}(t) + v_{id}(t+1), d = 1, 2, \ldots, D \quad (3)$$

3.2.2. Growth optimizer (GO)

The growth optimization (GO) (Fatani et al., 2023; Zhang et al., 2023) is an optimization technique that uses learning and introspection to enhance problem-solving abilities. The technique generates an initial list of candidate solutions, X, for optimization issues, similar to how individuals acquire knowledge by monitoring their environment. The algorithm then assesses its search method, identifying areas for improvement, and is used to address complex challenges in various engineering approaches (Zhang et al., 2023). The first stage involves generating a population of candidate solutions, which are assessed and improved throughout the optimization process (Zhang et al., 2023). The generation of novel solutions uses indexes to determine factors within each solution, with index "*i*" representing the specific solution within the population and dimension index "*j*" representing attribute values. The index's range is limited by upper and lower bounds (μb and lb) specific to the optimization problem's search space. The *rand(0,1)* function generates uniform random integers within the range of 0–1 (Fatani et al., 2023).

The GO algorithm is used in three phases: *population* phase segmentation, *learning* stage, and reflection stage (Gao et al., 2024;

Table 2
Performance comparisons of ELM, MLPNN, and RBFNN algorithms using different criteria.

Criteria	ELM	MLPNN	RBFNN
Training speed	Fast training due to single layer feed forward network and analytical determination of weights (Huang et al., 2006)	Moderate training speed; can be optimized with various algorithms like backpropagation (Rumelhart et al., 1986)	Generally slower than ELM but can be faster than MLPNN if the number of RBF units is optimized (Broomhead and Lowe, 1988)
Generalization performance	Good generalization with a large number of hidden neurons (Huang et al., 2011)	Capable of capturing complex patterns and relationships due to multiple hidden layers (Bishop, 2007)	Good generalization ability with proper selection of centers and width of RBF units (Moody and Darken, 1989)
Implementation simplicity	Easy to implement due to its simple architecture and training process (Huang et al., 2006)	More complex due to multiple layers and iterative training (Bishop, 2007)	Simpler than MLPNN but more complex than ELM due to the need for optimizing RBF units (Park and Sandberg, 1991)
Handling nonlinearity	Efficiently handles nonlinearity with enough hidden neurons (Huang et al., 2011)	Excellent at modeling highly nonlinear relationships (Rumelhart et al., 1986)	Effectively handles nonlinearity with properly tuned RBF units (Moody and Darken, 1989)
Overfitting	High risk of overfitting with too many hidden neurons (Huang et al., 2011)	Prone to overfitting if not properly regularized (Bishop, 2007)	Can overfit if the number of RBF units is too high or parameters are not properly tuned (Goodfellow et al., 2014; Park and Sandberg, 1991)
Parameter sensitivity	Sensitive to the number of hidden neurons and input weights initialization (Huang et al., 2011)	Highly sensitive to network architecture, learning rate, and other hyperparameters (Bishop, 2007)	Sensitive to the selection of RBF centers and the spread parameter (Moody and Darken, 1989)
Computational complexity	Low computational complexity in training phase (Huang et al., 2006)	Higher computational complexity due to backpropagation and multiple layers (Rumelhart et al., 1986)	Moderate computational complexity, can be high if many RBF units are used (Broomhead and Lowe, 1988)

Zhang et al., 2023). In *population phase segmentation*, the algorithm organizes the initial set of possible solutions (X) in a hierarchical structure based on a parameter scale of 5 classes. The upper section contains the current optimal solution, the "leader," and the two subsequent "elite" candidates. The algorithm uses leadership characteristics in human society to direct the optimization process, arranging solutions in ascending order from P1+1 to N–P1, and listing them in ascending order (Zhang et al., 2023). The GO algorithm's *learning stage* focuses on examining individual variations and their underlying causes, replicating four fundamental variations. It aims to understand factors that differentiate individuals and the underlying causes of these disparities (Zhang et al., 2023). *In reflection stage,* the reflection and learning components of the GO algorithm collaborate, necessitating the reevaluation of early insights and the resumption of systematic learning when an issue remains unresolved (Gao et al., 2024; Zhang et al., 2023).

3.3. The evaluation of forecasting errors and inaccuracies

Several widely used prediction-error statistical metrics are calculated to determine the DCTI prediction performance of each ML and HML model evaluated (Davoodi et al., 2023; Hosseini et al., 2023). These errors are residual error (Err_i, Eq. (4)), average relative error (ARE, Eq. (5)), average absolute relative error (AARE, Eq. (6)), root-mean-square error (RMSE, Eq. (7)), coefficient of determination (R^2, Eq. (8)), and standard deviation of error (SD, Eq. (9)).

$$Err_i = DCTI_{m_i} - DCTI_{p_i} \tag{4}$$

$$ARE = \frac{\sum_{i=1}^{n} RE_i}{n} \tag{5}$$

$$AARE = \frac{1}{l}\sum_{i=1}^{1}\left|\frac{Err_i}{DCTI_{m_i}}\right| \times 100 \tag{6}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Err_i)^2} \tag{7}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(Err_i)^2}{\sum_{i=1}^{n}\left(DCTI_{p_i} - \frac{\sum_{i=1}^{n} DCTI_{m_i}}{n}\right)^2} \tag{8}$$

$$SD = \sqrt{\frac{\sum_{i=1}^{l}\left(Err_i - \frac{\sum_{i=1}^{l} Err_i}{l}\right)}{l}} \tag{9}$$

Where $DCTI_{m_i}$ and $DCTI_{p_i}$ represent the actual measured and predicted values of the dependent variable i, respectively, and l represents the total number of data points in the dataset.

3.4. Development of HML models to predict DCTI

To achieve the best possible prediction performance with the ML algorithms, it is essential to optimize two categories of ML-control parameters. The first category pertains to the structural parameters of the algorithms, which include specifying the number of hidden layers and/or the number of neurons within these layers. The second category involves hyperparameters, including weights and biases. Consequently, the integration of the ML algorithms with the optimizers to configure the HML models involves two stages of optimization. In the first step, the structural parameters are optimized. In the second step, the optimization algorithms are applied to fine-tune the weights and biases, thereby enhancing the overall performance of the models.

For ELM, the first HML step is to determine the optimal number of nodes in the single hidden layer. To do this, the ELM algorithm is

applied as a loop to the normalized training set with hidden-layer nodes, ranging from 1 to 50, recording RMSE for each number of nodes. The number of hidden-layer nodes with the lowest RMSE is selected. The HML optimizer then focuses on the weights and biases. Each node has weights that are influenced by the number of input variables, and each node also requires a bias. The HML optimizer fine-tunes the ELM weights and biases to once again minimize the DCTI prediction error (RMSE) for the training data. This two-step optimization process is depicted in Fig. 2a.

The first HML optimization stage for MLPNN determines the number of hidden layers (between one and three). The PSO algorithm determines the optimal number of neurons for each number of possible hidden layers, recording the DCTI prediction error (RMSE) for the training data. The model structure with the optimal number of hidden layers and neurons is selected. The optimizer then focuses on fine-tuning the weights and biases, as shown in Fig. 2b. The optimized MLPNN model is then trained with the training set, and the trained model is applied to predict the testing set.

Fig. 3 illustrates the workflow used for optimizing RBFNN's structure and hyperparameter values. The first optimization step determines the number of neurons in the hidden layer. The k-means clustering algorithm is used for this step, and the optimum number of clusters it identifies with the training set is selected as the hidden-layer number of neurons. In the second optimization step, the centers of the optimum clusters and the weights applied to the output layer are optimized. The optimized RBFNN model is then trained with the training set, and the trained model is applied to predict the testing set.

4. Results

4.1. Data exploration

Evaluation of the dataset's Pearson correlation coefficient matrix (Fig. 4) is an essential step before creating ML predictive models, as it helps uncover the relationships between various features and the target variable. This examination offers deeper insights into the data, revealing underlying patterns and connections that guide feature engineering and preprocessing. Additionally, it assists in identifying anomalies or unexpected relationships in the data, leading to further investigation or necessary data cleaning.

Fig. 4 reveals that most of the variables exhibit an inverse relationship with DCTI, except for HL and ROP, which have a direct linear effect. The ANG feature has the highest correlation coefficient with DCTI, whereas the FLo feature displays the lowest correlation. While the correlation matrix reveals the linear relationships within the dataset, several trends warrant further discussion as they diverge from expectations of efficient drilling operations. These divergences are not due to data artefacts but are characteristic of the poor hole-cleaning conditions that this dataset comprehensively captures. For instance, the negative correlations between ROP and both WOB (−0.45) and RPM (−0.49) are counter-intuitive for optimal drilling but are a classic signature of dysfunction. During periods of cuttings accumulation, ROP decreases due to bit-balling and increased bottom-hole pressure. In response, drillers often instinctively increase WOB and RPM as an attempt to penetrate through the obstruction. This common operational response results in a dataset where high WOB and RPM are associated with low ROP. Similarly, the specific positive and negative correlations between other parameters (e.g., HL vs. WOB) are influenced by the wellbore geometry, friction, and the downhole conditions being experienced. Therefore, the correlation matrix reflects the complex, real-world interplay of parameters under a wide range of conditions, including severe dysfunction, which is precisely what enables the ML model to learn to identify and predict these inefficient states.

4.2. Dataset preprocessing

Outliers significantly impact the performance and reliability of ML models by altering model parameters, leading to skewed outcomes that don't accurately represent the data distribution. This decreases the model's overall prediction performance and generalization capabilities. Additionally, outliers can prolong training time as algorithms may require more iterations to converge (Davoodi et al., 2023). The relationship between drilling mechanical specific energy (MSE) and ROP (Fig. 5) was analyzed to identify data records relating to potential drilling problems. Generally, a low ROP combined with high MSE is an indication of potential issues in the drilling operation. In the raw dataset (Fig. 5a), several data points exhibit low ROP coupled with high MSE. Considering that the maximum confined compressive strength for the collected wells was approximately 25,000 psi, 52 outlying data points were identified and removed from the dataset (Fig. 5b).

To test the developed models on unseen data, a portion of the collected data must be set aside as a testing data subset. This is achieved by separating the filtered dataset (with outliers removed) into training and testing sets, using a specific partition ratio. Initially, 100 data records with a uniform distribution were excluded from the collected data. Then, the remaining data were divided into training and testing groups with five distinct partition ratios (70:30, 75:25, 80:20, 85:15, and 90:10). The RBFNN algorithm was used to predict DCTI with the different partition ratios, testing each trained model with the 100 isolated data records. Fig. 6 displays the RMSE for these DCTI predictions. As the quantity of training data increases, the prediction errors generated by the training set increase slightly, whereas the prediction error generated by the trained models applied to the unseen testing subset decreases from partitions of 70:30 to 75:25 and then remains approximately constant. Based on these results, a dataset partition ratio of 75% training: 25% testing was selected for detailed analysis of all ML and HML models considered.

Data normalization was applied to ensure all input features have equal influence in the analysis and modeling processes. When data variables are expressed on varying scales, the variables with larger scales can disproportionately affect the prediction results, leading to biased conclusions. In this study, data normalization was performed by specifying the minimum (X_{min}) and maximum (X_{max}) values measured for each feature (Davoodi et al., 2023). The values of each feature were then scaled between zero and one using Eq. (10).

$$X_{norm} = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \tag{10}$$

The NSGA-2 optimizer was combined with the MLPNN model to select the features with the most significant impact on DCTI. The MLPN structure used in this process consists of a three-layer hidden structure with 4 neurons in each layer. Fig. 7 illustrates the trend of RMSE and R^2 changes with the best set of selected features for a specific number of inputs. It is apparent that an increase in the number of input features up to seven reduces the model error. The rate of error reduction is initially high as features are added, but decreases as the number of inputs increases, becoming imperceptible beyond 7 included variables. Therefore, the DCTI predictive ML/HML models are developed with just seven input variables. Table 3 presents the best set of selected features for each specified number of inputs. Based on this analysis of the

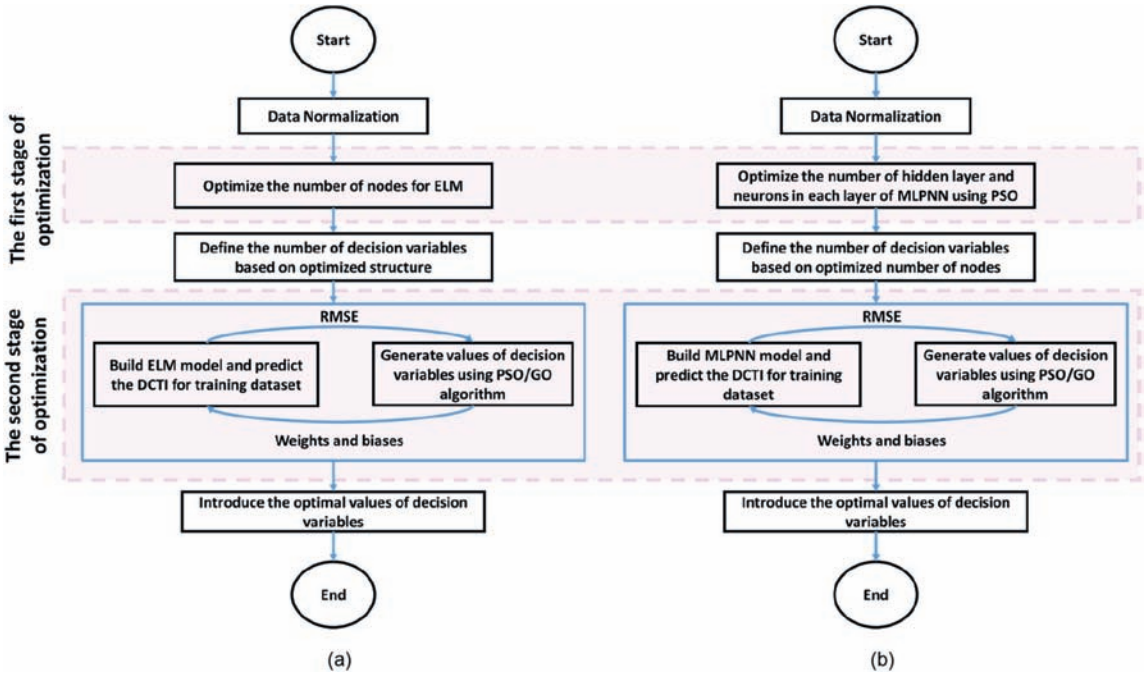


Fig. 2. Flowcharts for the HML optimization sequence applied to: (a) ELM; and (b) MLPNN.

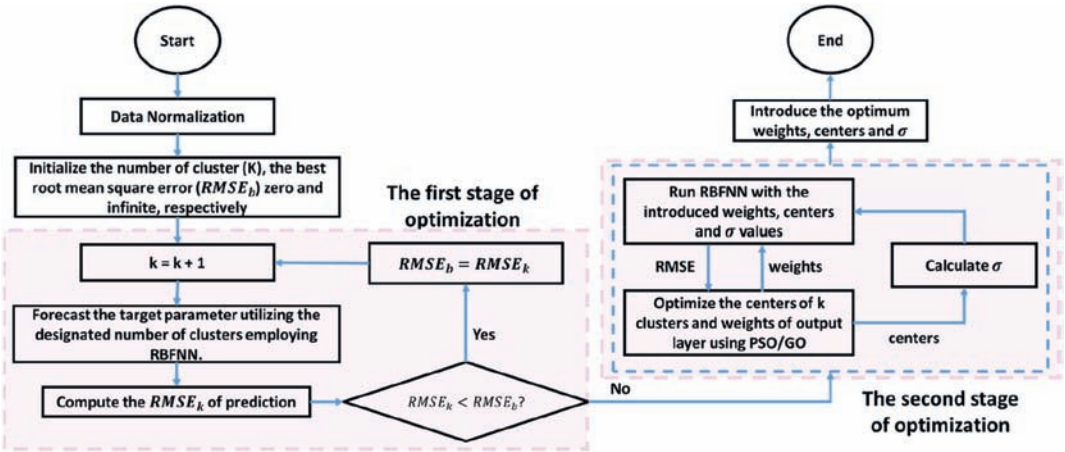


Fig. 3. Flowchart for the HML optimization sequence applied to RBFNN.

error reduction, the input features ANG, ROP, FLin, Depth, Trq, SPP, and RPM were selected for detailed DCTI modeling.

4.3. Tuning the standalone machine learning algorithms

The performance of ML and optimization algorithms is highly dependent on the values assigned to their control parameters (Davoodi et al., 2023). Therefore, to achieve a model with high accuracy and generalization ability, it is necessary for the control parameters of these algorithms to be adjusted according to the problem conditions before they are used. This tuning process was rigorously conducted on the designated training set (75% of the total data). To mitigate the variability inherent in model initialization and to prevent overfitting, a 10-fold cross-validation strategy was employed for robust evaluation and selection of the optimal model structure. In this technique, the training data is divided into ten equal subsets. The ML algorithm is executed ten times, with each iteration using one of the ten subsets for testing and the remaining nine subsets for training. This ensures that each subset is used for testing at least once, providing a robust evaluation of the model's performance. The error value in this technique is determined by averaging the results of the three runs with the highest accuracy on the testing fold. In this subsection, the structures of the ELM, MLPNN, and RBFN algorithms will be discussed. The PSO algorithm was used to adjust the structure of these algorithms. Since the final prediction performance of models obtained from ML algorithms depends on the initialization of their hyperparameters during the training process. Hence, the prediction performance of the resulting model can vary with each run.

Due to the presence of a single hidden layer in the ELM algorithm, only the number of neurons in this hidden layer needs to be determined. Consequently, only one decision variable is defined in the PSO algorithm to specify the number of neurons in the single hidden layer. The range for this decision variable was set between

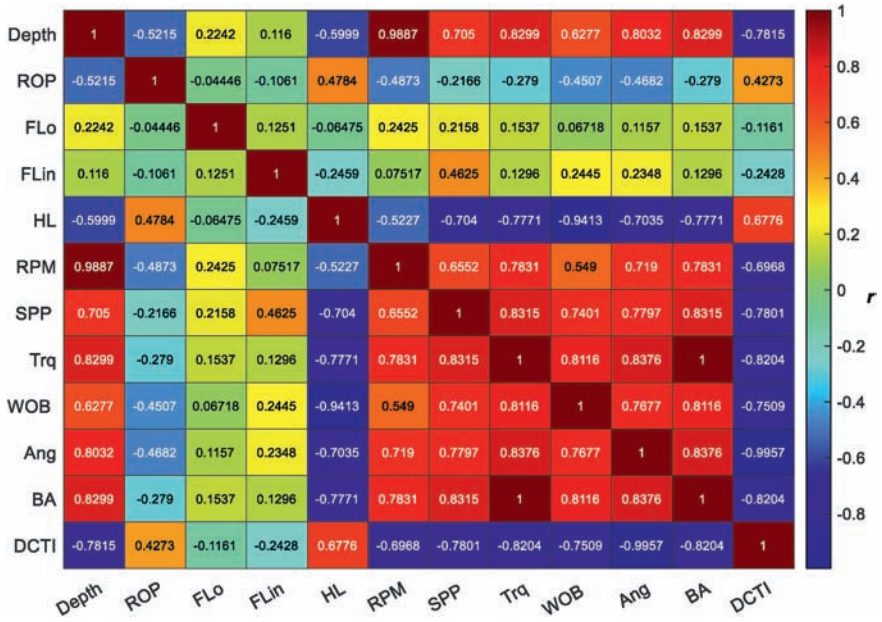


Fig. 4. Correlation heatmap matrix for recorded features exerting influence on the dependent variable, DCTI.

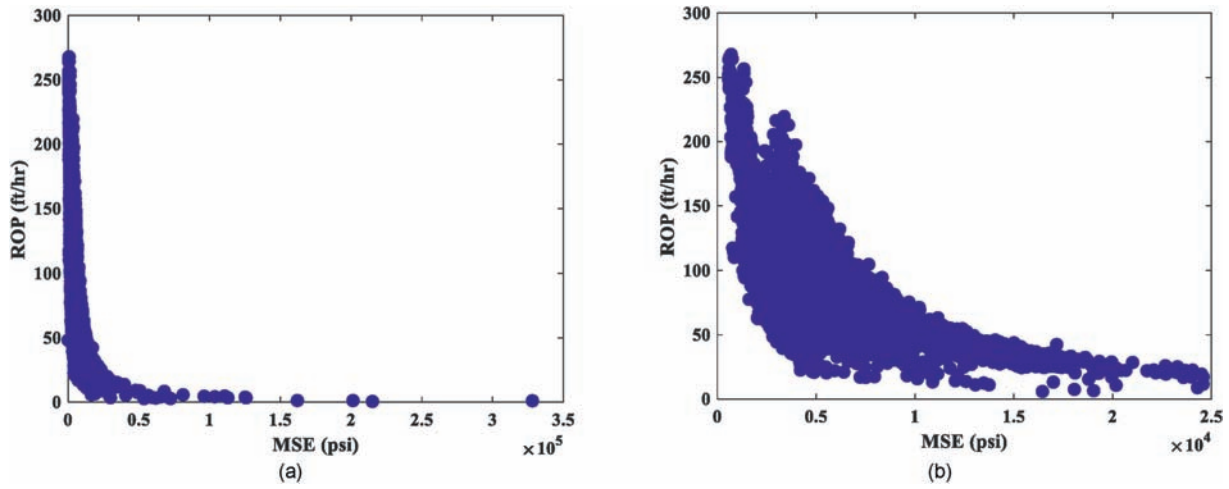


Fig. 5. Mechanical specific energy (MSE) versus rate of penetration (ROP) cross plots: (a) raw dataset and (b) dataset with outliers removed. Note the order of magnitude horizontal scale difference between these two graphs.

1 and 50. Applying the PSO algorithm to optimize the number of neurons in the ELM revealed that using 11 neurons in the single hidden layer leads to high accuracy (RMSE on training folds = 0.17925) with the training data. The activation function for the hidden layer was selected as the sigmoid function after empirical testing against alternatives (ReLU, sine), as it provided the most stable convergence for this dataset. The output layer employs a linear activation function. The model weights were determined analytically using the Moore-Penrose generalized inverse.

In the MLPN algorithm, it is necessary to determine both the number of neurons in each hidden layer and the number of hidden layers. Since a maximum of three hidden layers can effectively address very complex and non-linear problems, a separate PSO algorithm was employed to determine the optimal number of neurons for models with one, two, and three hidden layers. This approach allows for the comparison of error values across the best models obtained for each of these three configurations to select the most appropriate structure. Fig. 8a illustrates the error reduction over a series of iterations by the PSO algorithm using training set. The MLPNN with two hidden layers, comprising 3 neurons in the first layer and 5 neurons in the second layer, yields the lowest RMSE compared to models with one and three hidden layers. The hyperbolic tangent (tanh) activation function was used for all hidden layers, and the output layer utilized a linear activation function. The network was trained using the Levenberg-Marquardt algorithm, chosen for its rapid convergence on medium-sized datasets. To improve generalization and prevent

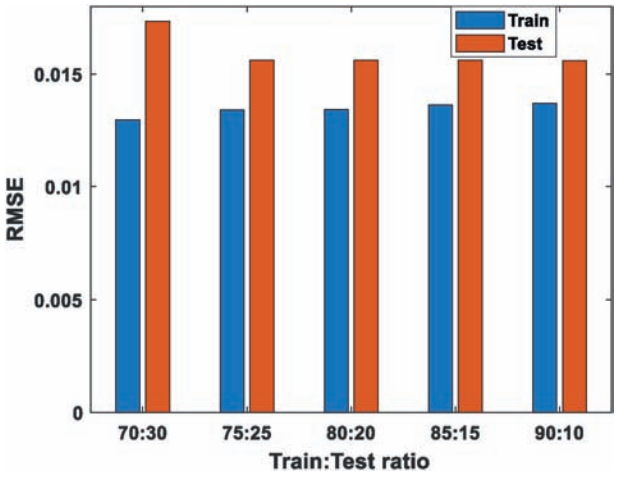


Fig. 6. RMSE of DCTI for the trained RBFNN models with different partition ratios applied to the 100 isolated (unseen) data records. Note in the studied dataset, the scale range of DCTI is from ~0.26 to ~1.9.

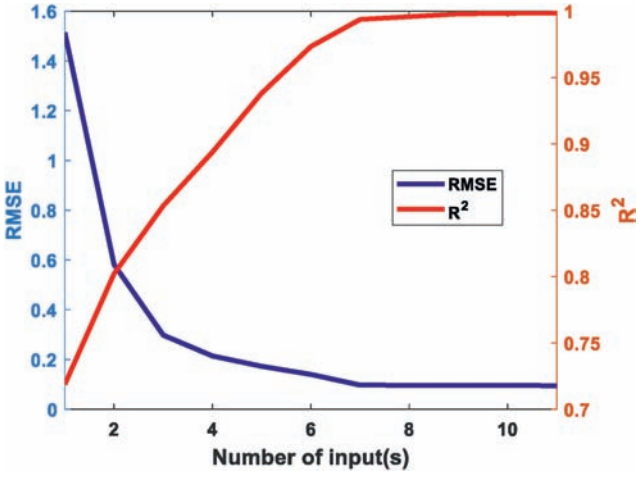


Fig. 7. Comparison of RMSE and R^2 for the selected optimal input feature combinations determined using the NSGA–II–MLPNN model.

overfitting, L2 regularization (weight decay) with a lambda value of 0.0001 was incorporated into the training process. Furthermore, early stopping was implemented with a patience of 10 epochs, monitoring the validation loss from the 10-fold cross-validation.

The number of neurons in the RBFNN algorithm layer was determined using an iteration loop set to 25. This analysis identified eight neurons as the optimal configuration, which resulted in the lowest DCTI prediction error (RMSE = 0.02529) for the standalone model (Fig. 8b). The centers of the radial basis functions were initialized using the k-means clustering algorithm applied to the training data. The spread (width) parameter of the Gaussian basis functions was set proportionally to the average distance between the centers within each cluster. The weights connecting the hidden layer to the linear output neuron were calculated using a least-squares method.

4.4. Developing and evaluating the HML predictive models

By configuring the predictive algorithms with suitable parameters, they were subsequently deployed on the training data, incorporating the seven selected input features. The evolution of RMSE reduction through various iterations of the optimization algorithms (Fig. 9) illustrates the fine-tuning of hyperparameters during the training phase. Notably, convergence to an optimal solution is evident across all algorithms by iteration 20. Nevertheless, except for the RBFNN-PSO algorithm, the remaining algorithms become ensnared in local optima.

Fig. 10 presents a cross-plot contrasting estimated and predicted values of DCTI applying the ML and HML algorithms to the training data. The ELM model predictions exhibit greater dispersion compared to the ideal Y = X line, indicating subpar performance. Conversely, HML models and the standalone RBFNN predictions align closely with the Y = X line, suggesting lower error rates relative to other models. The MLPNN model displays a distinct trend, with poor alignment to the Y = X line, indicating a tendency to underestimate higher DCTI values and slightly overestimate the lower end of the DCTI range below values of ~0.5. Table 4 presents an evaluation of both standalone and hybrid models using various error criteria and the R^2. HML models generate significantly lower error values compared to those from standalone algorithms, underscoring the efficacy of employing optimization algorithms to refine hyperparameter values. Among the array of developed models, it's evident that the ELM algorithm yields the highest error, while the model derived from the RBFNN-PSO algorithm achieves the lowest error, indicative of superior training performance in minimizing discrepancies between predicted and actual values.

Fig. 11 illustrates the comparison between the estimated and predicted DCTI values of each trained model applied to the testing set. Notably, the dispersion of data points in the ELM model is conspicuous, deviating noticeably from the ideal Y = X line. This dispersion phenomenon, observed during the training phase, persists in the test data, aligning with prior expectations. Of the standalone ML models, the RBFNN model exhibits superior performance. However, the HML models substantially outperform the standalone ML models. Notably, the MLPNN model displays

Table 3
RMSE and R^2 for DCTI predictions based on input feature combinations determined by the NSGA–II–MLPNN model. The best feature combination selected is displayed in bold type.

No	Selected feature(s)	RMSE	R^2
1	ANG	1.5135	0.7188
2	ANG, ROP	0.5820	0.8021
3	ANG, ROP, FLin	0.2978	0.8533
4	ANG, ROP, FLin, Depth	0.2141	0.8938
5	ANG, ROP, FLin, Depth, Trq	0.1725	0.9381
6	ANG, ROP, FLin, Depth, Trq, SPP	0.1400	0.9733
7	**ANG, ROP, FLin, Depth, Trq, SPP, RPM**	**0.0971**	**0.9937**
8	ANG, ROP, FLin, Depth, Trq, SPP, RPM, HL	0.0964	0.9958
9	ANG, ROP, FLin, Depth, Trq, SPP, RPM, HL, BA	0.0956	0.9979
10	ANG, ROP, FLin, Depth, Trq, SPP, RPM, HL, BA, FLo	0.0954	0.9983
11	ANG, ROP, FLin, Depth, Trq, SPP, RPM, HL, BA, FLo, WOB	0.0950	0.9986

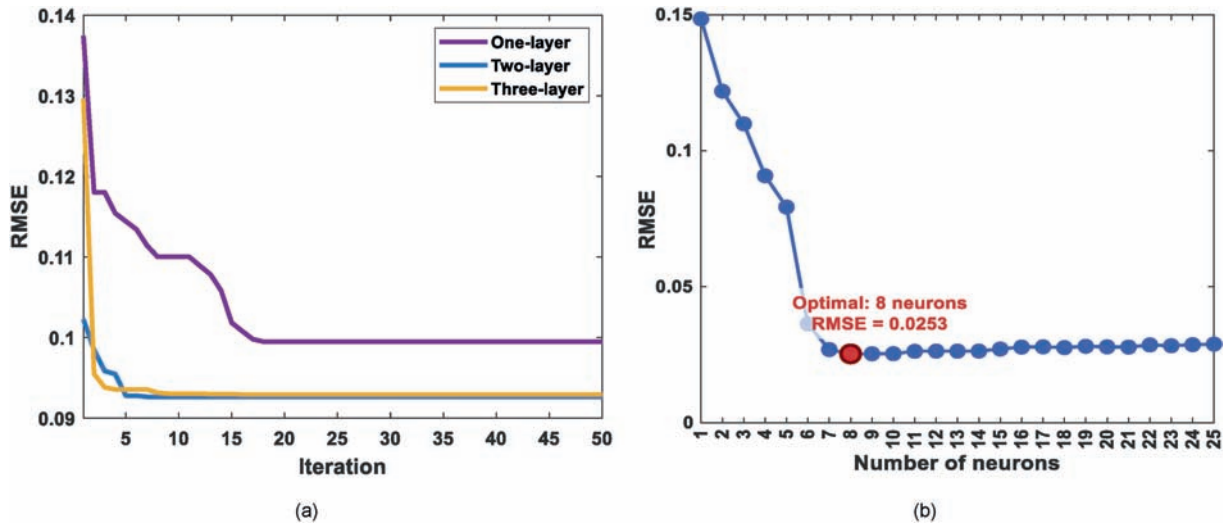


Fig. 8. RMSE reduction in the optimization of the structure of (a) MLPNN with one to three hidden layers applying the PSO algorithm using training data and (b) RBFNN neurons in the hidden layer.

substantial underestimation of higher DCTI values. Hence, caution is advised when employing this model for predictions involving elevated DCTI values. Table 5 compares ML and HML trained models applied to the testing set in terms of five error criteria. The superior prediction performance of the HML models is apparent. ELM generates the poorest DCTI prediction performance, and RBFNN-PSO generates the best DCTI prediction performance.

5. Discussion

Further insight into the prediction performances of the ML and HML models can be gained from additional analysis.

5.1. Score analysis of the models

ML/HML prediction models can excel based on one performance measure but underperform when assessed by others. To address this inconsistency, a scoring system that integrates multiple performance metrics is used. For each metric, the model with the best performance among all compared models receives the highest score, while the model with the weakest performance is assigned the lowest score. The maximum score is based on the total number of models evaluated; therefore, the scores in this study range from 1 to 9. These scores are calculated separately for the predictions generated by the training and testing sets. An overall score, representing the combined total of training and testing scores, is provided to give a more comprehensive view of each model's performance. Table 6 displays the ML and HML models ranked (highest to lowest) according to their DCTI prediction scores: RBFNN-PSO (highest performance) > RBFNN-GO > MLPNN-PSO > ELM-PSO > MLPNN-GO > ELM-GO > RBFNN > MLPNN > ELM (lowest performance). This ranking is displayed as a radar diagram (Fig. 12).

5.2. Robustness analysis of the models

Robustness is a crucial aspect of ML models that influences their reliability and applicability in real-world situations. According to Zhang et al. (2021), a robust model should exhibit consistent performance across diverse datasets, enabling reliable generalization beyond the training data. This can be a challenge in noisy datasets such as those generated during the drilling of wellbores. A robust model is resistant to noise and outlying data points (2014). By accurately distinguishing relevant patterns from irrelevant noise, robust models enhance prediction performance and stability, which is essential for making informed decisions in dynamic environments.

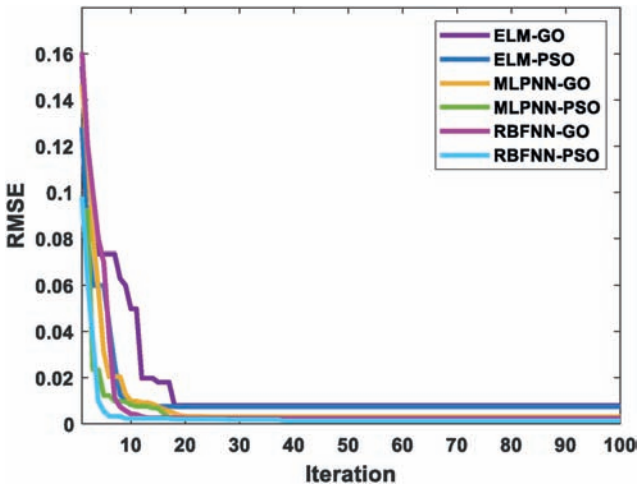


Fig. 9. Error reduction in each iteration of the optimization algorithm in the training phase of HML models configured to predict DCTI.

The reliability of the developed models was assessed by introducing different levels of white noise to the input variable distributions to gauge robustness in terms of R^2 values. Varying levels of Gaussian white noise were generated by adjusting the variance parameter "s" while maintaining a mean (μ) of zero in Eq. (11) governing the probability density function, $P_x(x)$.

$$P_x = \frac{1}{s\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2s}} \tag{11}$$

Fig. 13 illustrates the robustness test results. As expected, a decreasing trend of R^2 values is generated as noise levels increase. However, the severity of that trend varies among the models. Notably, the model developed with ELM exhibits a more pronounced decline in R^2 compared to other models, indicating its lower robustness. Conversely, the RBFNN-GO and RBFNN-PSO

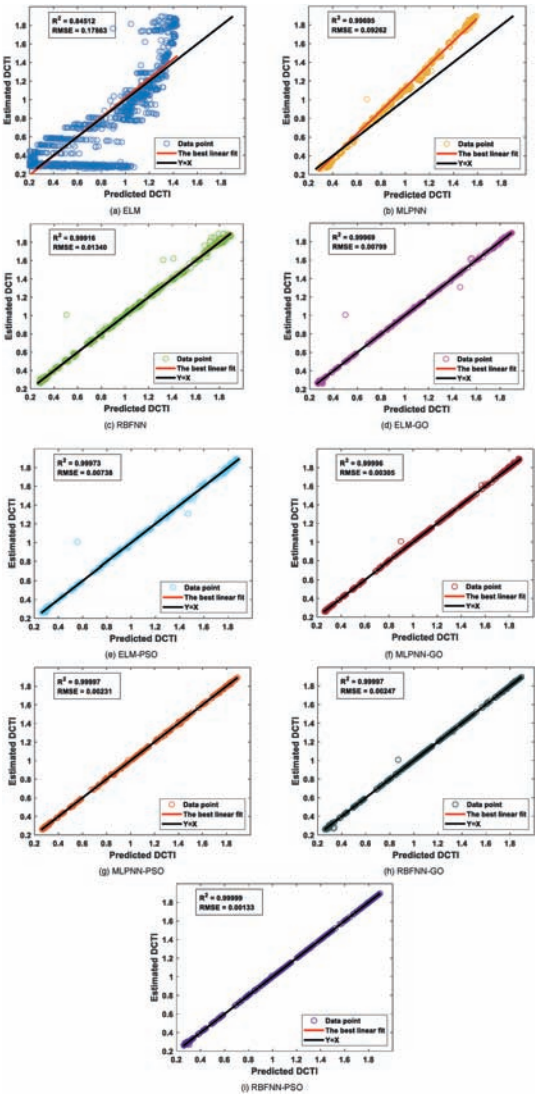


Fig. 10. Cross plots illustrating the comparison between estimated and predicted DCTI using (a) ELM, (b) MLPNN, (c) RBFNN, (d) ELM-GO, (e) ELM-PSO, (f) MLPNN-GO, (g) MLPNN-PSO, (h) RBFNN-GO, and (i) RBFNN-PSO models applied to the training data subset.

Table 4
Comparison of prediction performance for standalone ML and HML models applied to the DCTI training data subset.

Type	Models	ARE	AARE	RMSE	R^2	SD
Standalone	ELM	0.06972	0.24585	0.17863	0.84512	0.17864
	MLPNN	0.03355	0.09156	0.09262	0.99695	0.09046
	RBFNN	0.00828	0.01797	0.01340	0.99916	0.01326
Hybrid	ELM-GO	0.00041	0.00481	0.00799	0.99969	0.00799
	ELM-PSO	0.00009	0.00229	0.00738	0.99973	0.00738
	MLPNN-GO	−0.00242	0.00484	0.00305	0.99996	0.00288
	MLPNN-PSO	0.00009	0.00255	0.00231	0.99997	0.00231
	RBFNN-GO	0.00009	0.00128	0.00247	0.99997	0.00247
	RBFNN-PSO	0.00002	0.00226	0.00133	0.99999	0.00133

models display a more gradual decrease in R^2, indicating that they are more robust. Up to a variance level of 0.35, these two models exhibit a gentler decline in R^2 compared to higher variance levels, suggesting a noise tolerance threshold of ~0.35.

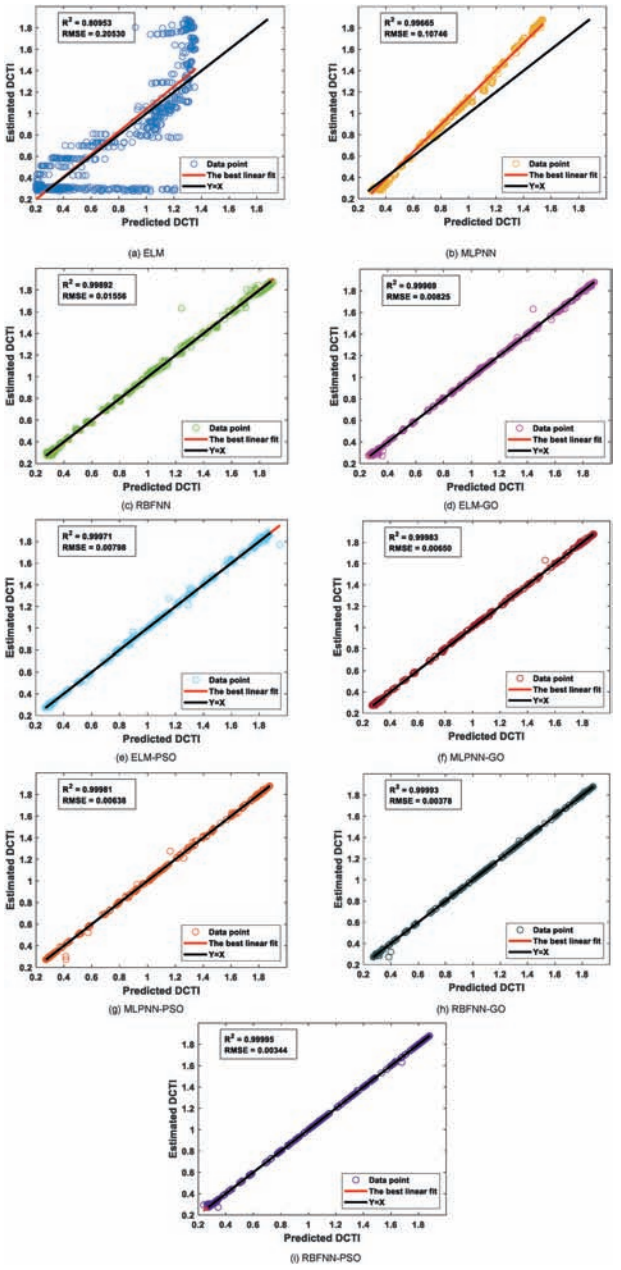


Fig. 11. Cross plots illustrating the comparison between estimated and predicted DCTI using (a) ELM, (b) MLPNN, (c) RBFNN, (d) ELM-GO, (e) ELM-PSO, (f) MLPNN-GO, (g) MLPNN-PSO, (h) RBFNN-GO, and (i) RBFNN-PSO trained models applied to the testing data subset.

5.3. Monte Carlo simulation

To comprehensively evaluate the efficacy of the RBFNN-PSO model, especially in the absence of extensive unseen data, a Monte Carlo simulation was employed to generate a synthetic dataset. The actual data distribution of each feature was accounted for in the simulation process, using the minimum and maximum values from the original training data. In generating the synthetic dataset, not only were the individual distributions of each feature considered, but also their joint distributions, including covariances between input variables. This approach ensures that the relationships and dependencies observed in the original dataset are preserved in the simulated data, which is crucial for accurately modeling non-linear distributions. Specifically, the covariance

Table 5
Comparison of prediction performance of trained standalone and hybrid ML and HML models applied to the DCTI testing data subset.

Type	Models	ARE	AARE	RMSE	R^2	SD
Standalone	ELM	0.05112	0.26442	0.20530	0.80953	0.20471
	MLPNN	0.02957	0.10099	0.10746	0.99665	0.10380
	RBFNN	0.00876	0.02123	0.01556	0.99892	0.01542
Hybrid	ELM-GO	0.00104	0.01236	0.00825	0.99969	0.00825
	ELM-PSO	−0.00002	0.00381	0.00798	0.99971	0.00798
	MLPNN-GO	−0.00608	0.01126	0.00650	0.99983	0.00602
	MLPNN-PSO	0.00042	0.00351	0.00638	0.99981	0.00638
	RBFNN-GO	0.00033	0.00203	0.00378	0.99993	0.00378
	RBFNN-PSO	−0.00033	0.00307	0.00344	0.99995	0.00344

matrix of the training data was calculated to capture the linear relationships between pairs of input features. This matrix was then used to generate multivariate normal distributions for synthetic data. By doing so, it was ensured that the synthetic dataset not only mimicked the marginal distributions of each feature but also their covariances, thus preserving the statistical properties of the original dataset. A full dataset with input features that followed these analyzed distributions was generated by the Monte Carlo simulation, maintaining both individual feature distributions and their covariances. These simulated input records were then fed into the best-performing RBFNN-PSO model to predict the DCTI values. Fig. 14 illustrates the comparison between the distributions of the estimated and predicted DCTI values generated by the RBFNN-PSO model using input data distributions derived from the Monte Carlo simulation. The histograms display a close resemblance between the two distributions, which indicates that the model's predictions align closely with the statistical characteristics of the synthetic data. This resemblance affirms the robustness and generalization capabilities of the RBFNN-PSO model, as it successfully captures the underlying data patterns observed in the collected dataset.

5.4. Feature importance analysis

Examining the influence of each input feature on the performance of the highest-scoring prediction model (RBFNN-PSO) can establish its relative importance in determining DCTI predictions. This is achieved by conducting a Shapley additive explanations (SHAP) influence analysis of the RBFNN-PSO model's predictions.

Table 6
Model scores for specific criteria and total scores for training, testing, and overall performance in predicting DCTI.

Models	Subset	ARE	AARE	RMSE	R^2	SD	Score	Total score
ELM	Train	3	1	1	2	1	8	14
	Test	2	1	1	1	1	6	
MLPNN	Train	4	2	2	3	2	13	24
	Test	3	2	2	2	2	11	
RBFNN	Train	5	3	3	4	3	18	34
	Test	4	3	3	3	3	16	
ELM-GO	Train	7	5	4	5	4	25	47
	Test	6	4	4	4	4	22	
ELM-PSO	Train	8	7	5	6	5	31	61
	Test	9	6	5	5	5	30	
MLPNN-GO	Train	6	4	6	7	6	29	59
	Test	5	5	6	7	7	30	
MLPNN-PSO	Train	8	6	8	8	8	38	71
	Test	7	7	7	6	6	33	
RBFNN-GO	Train	8	9	7	8	7	39	80
	Test	8	9	8	8	8	41	
RBFNN-PSO[a]	Train	9	8	9	9	9	44	87
	Test	8	8	9	9	9	43	

[a] The best performing model.

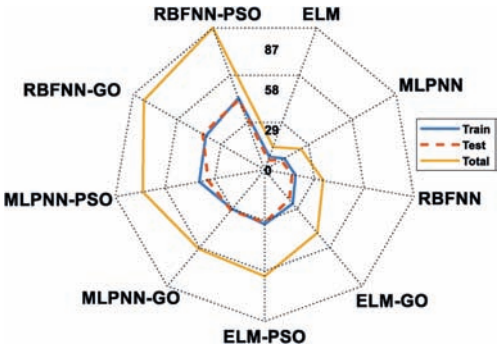


Fig. 12. A radar diagram comparing the prediction performance scores for DCTI prediction, showing the difference between the training and testing sets and the total score achieved by both ML and HML models.

SHAP determines the impact of each data record on the overall prediction performance of the model, quantifying those impacts in positive and negative SHAP values (Fig. 15). The individual data variable values distributions are depicted by a color scale (red = high variable value; blue = low variable value) with the positive and negative SHAP value magnitudes expressed on the horizontal axis (Fig. 15a) The SHAP summary plot (Fig. 15b) displays the mean absolute SHAP values of each input data variable, ranked in descending order, and identifies the overall relative importance of each input feature. When the ANG variable has low values (near vertical wellbore), its SHAP value is positive, whereas at high ANG values (near horizontal wellbore), its SHAP value is substantially negative (Fig. 15a). This contrasts with other input features for the RBFNN-PSO model, where the SHAP values deviate only slightly from either side of zero. The SHAP summary values indicate that the ANG input variable has the most significant impact on the RBFNN-PSO model's DCTI predictions, while the SPP feature has the least influence (Fig. 15b).

The power of the SHAP analysis extends beyond static feature importance to the real-time diagnosis of dysfunctional drilling states. By examining the combination and magnitude of SHAP values for a specific low DCTI prediction, the model can immediately flag the root cause of the inefficiency. For instance, a scenario where a high ROP exerts a strong negative influence (large negative SHAP value) while the FLin provides only a minimal positive contribution would be a clear, quantifiable indicator of a dysfunctional state where cuttings generation is drastically outpacing the

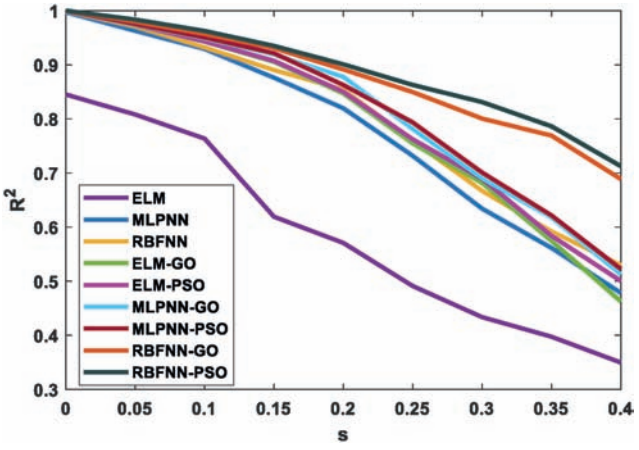


Fig. 13. R^2 trends associated with the introduction of various noise levels to the input variable distributions used to predict DCTI.

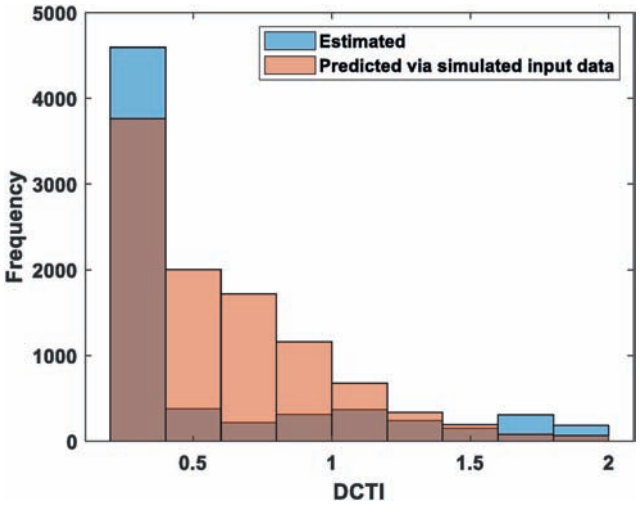


Fig. 14. Histograms comparing the distributions of estimated and predicted DCTI values derived by RBFNN-PSO from the data generated via Monte Carlo simulations. The dark shading represents values where the two distributions overlap.

system's hole-cleaning capacity. Conversely, if the ANG is the dominant negative contributor in a horizontal section, it highlights the inherent geometric challenge, prompting proactive measures rather than reactive parameter changes. This ability to deconstruct a poor hole-cleaning condition into its constituent, weighted parameter influences in real time is what transforms the model from a simple alarm into a diagnostic decision-support system.

The SHAP analysis provides more than just model interpretability; it unlocks the primary practical value of this work by transforming the predictive surrogate model into a prescriptive decision-support system. While the real-time prediction of a declining DCTI value serves as a valuable alarm, the SHAP analysis provides the diagnostic reason for the alarm, which is necessary for implementing the most appropriate corrective action. By examining the contribution of each feature for a specific low prediction, a drilling operator can identify the root cause and execute a targeted mitigation strategy in real time .

- Scenario 1: High ROP is the dominant negative contributor. This indicates that cuttings are being generated faster than the hydraulic system can transport them. The prescribed action is to reduce the ROP to re-establish equilibrium between cuttings generation and evacuation.
- Scenario 2: Low FLin is the dominant negative contributor. This indicates insufficient hydraulic energy for cuttings evacuation. The action is to increase the pump FLin, provided the specified operational limits (ECD, SPP) are not exceeded.
- Scenario 3: High ANG is a strong contributor. This is a non-controllable geometric factor. Its high influence serves as a constant reminder of the inherent cleaning challenge in deviated sections, prompting proactive measures like pumping high-viscosity sweeps or increasing RPM to agitate cuttings beds preemptively.

This level of interpretability demonstrates that the model has learned physically meaningful relationships between controllable drilling parameters and hole-cleaning efficiency. It provides a data-driven rationale for interventions, moving beyond a simple numerical prediction to offer specific, actionable guidance. The SHAP analysis confirms that the model's value lies not merely in replicating an empirical equation but in capturing the underlying physics of the drilling process in a directly useable and operationally transformative way.

6. Conclusions

This study successfully developed robust hybridized machine learning (HML) models to serve as a real-time surrogate for the empirical drilling cuttings transport index (DCTI). By leveraging only real-time, surface-recorded drilling data, the models overcome the fundamental time-delay limitation associated with the DCTI method's calculation, which requires laboratory-measured fluid properties. The models effectively provide a real-time diagnostic estimate of the DCTI, enabling immediate assessment of hole-cleaning conditions. Optimizing extreme learning machine (ELM), multi-layer perceptron neural network (MLPNN), and radial basis function neural networks (RBFNN) with either a particle swarm optimizer (PSO) or growth optimizer (GO) leads to effective HML models. Pre-processing identified and removed outlier data from a 6818-data-point dataset (from three oil and gas fields located in the Eastern Province of Saudi Arabia, comprising five wells) compiled with real-time recorded variables during drilling operations drilling wells. The filtered data points were divided into training and testing sets using a range of separation ratios.

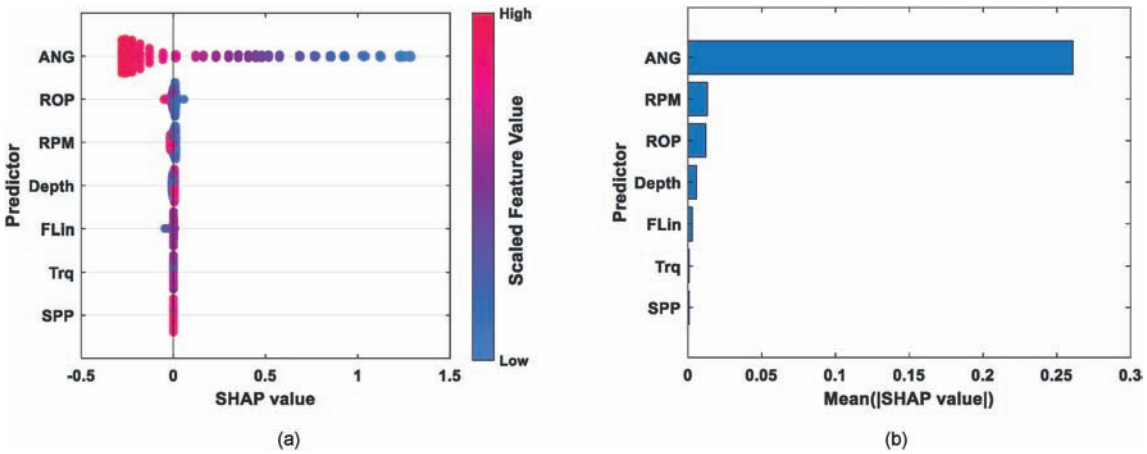


Fig. 15. SHAP analysis results for the RBFNN-PSO model's DCI predictions: (a) SHAP detailed input variable impact plot, and (b) SHAP summary plot of feature importance.

Feature selection identified a subset of the available recorded features based on their relative impacts on DCTI. Standalone machine learning (ML) and optimized HML algorithms were applied to the training data. The trained models were then evaluated with the testing set. Various prediction performance scoring metrics were calculated for each model, applied to both the training and testing data subsets. The best-performing model was then evaluated with Monte Carlo simulation and Shapley additive explanations (SHAP) influence analysis to assess its generalizability, the relative importance of each input characteristic on the model's output, and the marginal effect of each input feature. Crucially, the SHAP analysis transforms the model from a predictive tool into a prescriptive one, identifying the key drilling parameters influencing HC in real time and enabling data-driven operational decisions to prevent non-productive time. The results lead to the following conclusions.

- Examining the relationship between mechanical specific energy (MSE) and rate of penetration (ROP) identified 52 outlier data points, which were removed to generate a filtered input dataset.
- A 75% training and 25% testing data-partition split led to the best DCTI prediction performance and strong model generalization.
- Applying the second version of the non-dominated sorting genetic algorithm (NSGA-II) with MLPNN identified the most-effective subset of input features (inclination angle (ANG), ROP, flow rate (FLin), depth, torque (Trq), standpipe pressure (SPP), and drill-string rotation (RPM)) in terms of DCTI influence.
- ELM sensitivity analysis identified 11 neurons in its hidden layer to be most effective.
- PSO to determine the optimal structure of MLPNN to consist of two hidden layers, consisting of 3 neurons in the first layer and five neurons in the second layer.
- RBFNN Sensitivity analysis identified 25 neurons in its hidden layer that generated the lowest error value (root mean square error (RMSE) = 0.02529).
- The HML models outperformed the standalone ML algorithms, as confirmed by “Score” analysis. This confirmed the value of incorporating metaheuristic optimization algorithms.
- ELM generated the poorest prediction performance training (RMSE training = 0.17863; RMSE testing = 0.20530), whereas RBFNN-PSO achieved the best prediction performance (RMSE training = 0.00133; RMSE testing = 0.00344).
- Noise-robustness analysis confirmed that RBFNN-PSO was most resilient to noisy data inputs.
- RBFNN-PSO was further evaluated with the dataset generated by Monte Carlo simulation, demonstrating the ability of the trained model to perform well with unseen data.
- SHAP analysis applied to RBFNN-PSO revealed ANG as the most influential feature and SPP as the least influential feature on this high-performing model. This transforms the model from a simple numerical surrogate into a diagnostic and prescriptive tool, providing actionable insights for optimizing drilling parameters in real time.

CRediT authorship contribution statement

Mohammed Al-Shargabi: Writing – review & editing, Writing – original draft, Visualization, Software, Methodology, Investigation, Formal analysis, Data curation. **Shadfar Davoodi:** Writing – review & editing, Writing – original draft, Validation, Supervision, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **David A. Wood:** Writing – review & editing, Visualization, Validation, Methodology, Formal analysis. **Mohammad Mehrad:** Writing – original draft, Validation, Software, Investigation, Formal analysis, Data curation. **Mohammed Al-Rubaii:** Writing – review & editing, Validation, Methodology, Formal analysis, Data curation, Conceptualization.

Data availability

Data used in the present study are confidential and the authors do not have permission to share the dataset publicly or privately.

Funding

The work was supported by the grant for research centers in the field of AI provided by the Ministry of Economic Development of the Russian Federation in accordance with the agreement 000000C313925P4F0002 and the agreement with Skoltech N°139-10-2025-033.

Declaration of competing interest

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.

Abbreviations

AAPE	Average Absolute Percentage Error
AF_m	Modified Angle Factor
ANG	Angle Inclinations
ANN	Artificial Neural Networks
API	American Petroleum Institute
BA	Bit Aggressiveness Factor
CBH	Cuttings-Bed Height
CCA	Cutting Concentration in the Wellbore Annulus
CCA_v	Volumetric Cutting Concentration in the Wellbore Annulus
CCI	Cutting Carrying Index
DCTI	The Drilling Cutting's Transport Index
DF	Drilling Fluid
DL	Deep Learning
ELM	Extreme Learning Machine
FL	Fuzzy Logic
FLin	(Drilling Fluid) Flow In
FLo	(Drilling Fluid) Flow out
GA	Genetic Algorithm
GO	Growth Optimizer
GPM	Flow Rate
HC	Hole Cleaning
HL	Hook load
HML	Hybrid Machine Learning
k-NN	k-Nearest Neighbor
LC_m	Modified Original Lifting Capacity
LSTM	Long Short-Term Memory
MAPE	Mean Average Percentage Error
ML	Machine Learning
MLPNN	Multi-Layer Perceptron Neural Network
MLR	Multiple Linear Regression
MSE	Mechanical Specific Energy
NSGA-II	Non-Dominated Sorting Genetic Algorithm
PDC	Polycrystalline Diamond Compact Drill Bit
PSO	Particle Swarm Optimization
PV	Plastic Viscosity
R^2	Coefficient of Determination
RBF	Radial Basis Function
RF_m	Modified Average Rheology Factor
RMSE	Root Mean Square Error
RNN	Recurrent Neural Network

ROP	Rate of Penetration
RPM	Revolutions Per Minute
RSS	Rotary-Steerable System
SD	Standard Deviation of Error
SHAP	Shapley Additive Explanations
SPP	Stand Pipe Pressure
SQL	Structured Query Language
SVR	Support Vector Regression
TI	Transport Index
TR	Transport Ratio
Trq	Torque
WOB	Weight on Bit
YP	Yield Point
β^1	Lifting Coefficient

References

Adari, R.B., Miska, S., Kuru, E., Bern, P., Saasen, A., 2000. Selecting drilling fluid properties and flow rates for effective hole cleaning in high-angle and horizontal Wells. In: SPE Reservoir Engineering. Society of Petroleum Engineers), OnePetro, pp. 273–282. https://doi.org/10.2118/63050-MS.

Ahmed, A.S.D., 2020. Hole Widening Drilling Penetration Mechanism and Performance Evaluation in Hard Rock Utilizing Cutting Analysis. Memorial University of Newfoundland. https://doi.org/10.48336/cf2h-dz84.

Al-Azani, K., Elkatatny, S., Ali, A., Ramadan, E., Abdulraheem, A., 2019. Cutting concentration prediction in horizontal and deviated wells using artificial intelligence techniques. J. Pet. Explor. Prod. Technol. 9, 2769–2779. https://doi.org/10.1007/s13202-019-0672-3.

Ali, J.K., 1994. Neural networks: a new tool for the petroleum industry?. In: Society of Petroleum Engineers - European Petroleum Computer Conference. SPE, pp. 233–242. https://doi.org/10.2523/27561-ms, 1994, EPCC 1994.

Alkinani, H.H., Al-Hameedi, A.T.T., Dunn-Norman, S., Al-Alwani, M.A., Missouri, D.L., Al-Bazzaz, W.H., 2019. Application of descriptive data analytics: how to properly select the best ranges of viscosity and flow rate for optimal hole cleaning?. In: SPE Eastern Regional Meeting https://doi.org/10.2118/196597-MS.

Al-Rubaii, Al-Shargabi, M., Al-Shehri, D., 2023a. A novel model for the real-time evaluation of hole-cleaning conditions with case studies. Energies (Basel) 16, 4934. https://doi.org/10.3390/en16134934.

Al-Rubaii, Al-Shargabi, M., Al-Shehri, D., Alyami, A., Minaev, K.M., 2023b. A novel efficient borehole cleaning model for optimizing drilling performance in real time. Appl. Sci. 13, 7751. https://doi.org/10.3390/app13137751.

Al-Rubaii, Al-Shargabi, M., Al-Shehri, D.A., 2023c. A novel automated model for evaluation of the efficiency of hole cleaning conditions during drilling operations. Appl. Sci. 13, 6464. https://doi.org/10.3390/APP13116464.

Al-Rubaii, M., Al-Shargabi, M., Al-Shehri, D., 2023d. Hole cleaning during drilling oil and gas Wells: a review for hole-cleaning chemistry and engineering parameters. Adv. Mater. Sci. Eng. 2023, 1–33. https://doi.org/10.1155/2023/6688500.

Al-Shargabi, M., Davoodi, S., Wood, D.A., Al-Rubaii, M., Minaev, K.M., Rukavishnikov, V.S., 2024. Hole-cleaning performance in non-vertical wellbores: a review of influences, models, drilling fluid types, and real-time applications. Geoenergy Sci. Eng. 233, 212551. https://doi.org/10.1016/J.GEOEN.2023.212551.

Awad, M., Khanna, R., 2015. Efficient Learning Machines: Theories, Concepts, and Applications for Engineers and System Designers. Springer nature.

Bishop, C.M., 2007. Pattern Recognition and Machine Learning. Choice Reviews Online. https://doi.org/10.5860/choice.44-5091.

Bishop, C.M., Nasrabadi, N.M., 2006. Pattern Recognition and Machine Learning. Springer.

Brenjkar, E., Biniaz Delijani, E., 2022. Computational prediction of the drilling rate of penetration (ROP): a comparison of various machine learning approaches and traditional models. J. Pet. Sci. Eng. 210, 110033. https://doi.org/10.1016/J.PETROL.2021.110033.

Broomhead, D.S., Lowe, D., 1988. Multivariable functional interpolation and adaptive networks. Complex Syst. 2, 321–355.

Chandra, S., Gaur, P., Pathak, D., 2020. Radial basis function neural network based maximum power point tracking for photovoltaic brushless DC motor connected water pumping system. Comput. Electr. Eng. 86, 106730. https://doi.org/10.1016/J.COMPELECENG.2020.106730.

Chowdhury, D., Hovda, S., 2023. A hybrid fuzzy logic/genetic algorithm model based on experimental data for estimation of cuttings concentration during drilling. Geoenergy Sci. Eng. 231, 212387. https://doi.org/10.1016/J.GEOEN.2023.212387.

Chowdhury, D., Hovda, S., 2022. Estimation of downhole cuttings concentration from experimental data – Comparison of empirical and fuzzy logic models. J. Pet. Sci. Eng. 209, 109910. https://doi.org/10.1016/J.PETROL.2021.109910.

Costa, L.C., Carvalho, C.F., Soares, A.S.F., Souza, A.C.P., Bastos, E.F.T., Guimarães, E.C.B.T., Santos, J.C., Carvalho, T., Calderari, V.H., Marinho, L.S., Marques, M.R.C., 2023. Physical and chemical characterization of drill cuttings: a review. Mar. Pollut. Bull. https://doi.org/10.1016/j.marpolbul.2023.115342.

Davoodi, S., Al-Shargabi, M., Wood, D.A., Mehrad, M., 2025. Advancement of artificial intelligence applications in hydrocarbon well drilling technology: a review. Appl. Soft Comput. 176, 113129. https://doi.org/10.1016/j.asoc.2025.113129.

Davoodi, S., Mehrad, M., Wood, D.A., Ghorbani, H., Rukavishnikov, V.S., 2023. Hybridized machine-learning for prompt prediction of rheology and filtration properties of water-based drilling fluids. Eng. Appl. Artif. Intell. 123, 106459. https://doi.org/10.1016/j.engappai.2023.106459.

Deloitte, 2025. 2025 Oil and Gas Industry Outlook [WWW Document]. URL. https://www.deloitte.com/us/en/insights/industry/oil-and-gas/oil-and-gas-industry-outlook.html.

Eberhart, R., Kennedy, J., 1995. New optimizer using particle swarm theory. In: Proceedings of the International Symposium on Micro Machine and Human Science, pp. 39–43. https://doi.org/10.1109/MHS.1995.494215.

Elsayed Salem, S.K., Nasr El-Din, M.A., 2006. Drillpipe eccentricity prediction during drilling directional Wells. In: Canadian International Petroleum Conference. https://doi.org/10.2118/2006-047.

Fatani, A., Dahou, A., Abd Elaziz, M., Al-qaness, M.A.A., Lu, S., Alfadhli, S.A., Alresheedi, S.S., 2023. Enhancing intrusion detection systems for IoT and cloud environments using a growth optimizer algorithm and conventional neural networks. Sensors 23, 4430. https://doi.org/10.3390/S23094430.

Gao, H., Zhang, Q., Bu, X., Zhang, H., 2024. Quadruple parameter adaptation growth optimizer with integrated distribution, confrontation, and balance features for optimization. Expert Syst. Appl. 235, 121218. https://doi.org/10.1016/J.ESWA.2023.121218.

Goodfellow, I.J., Shlens, J., Szegedy, C,. 2014. Explaining and harnessing adversarial examples. In: Proceedings of the International Conference on Learning Representations, San Diego, CA, USA. https://doi.org/10.48550/arXiv.1412.6572.

Han, Y., Zhang, X., Xu, Z., Song, X., Zhao, W., Zhang, Q., 2022. Cuttings bed height prediction in microhole horizontal Wells with artificial intelligence models. Energies 15, 8389. https://doi.org/10.3390/EN15228389.

Hornik, K., Stinchcombe, M., White, H., 1989. Multilayer feedforward networks are universal approximators. Neural Netw. 2, 359–366.

Hosseini, S., Khatti, J., Taiwo, B.O., Fissha, Y., Grover, K.S., Ikeda, H., Pushkarna, M., Berhanu, M., Ali, M., 2023. Assessment of the ground vibration during blasting in mining projects using different computational approaches. Sci. Rep. 13, 1–29. https://doi.org/10.1038/s41598-023-46064-5.

Huang, G., Zhu, Q., Siew, C., 2004. Extreme learning machine: a new learning scheme of feedforward neural networks. In: IEEE International Conference on Neural Networks - Conference Proceedings, vol. 2, pp. 985–990. https://doi.org/10.1109/IJCNN.2004.1380068.

Huang, G.-B., Zhou, H., Ding, X., Zhang, R., 2011. Extreme learning machine for regression and multiclass classification. IEEE Trans. Syst. Man Cybern., Part B (Cybernetics) 42, 513–529.

Huang, G.-B., Zhu, Q.-Y., Siew, C.-K., 2006. Extreme learning machine: theory and applications. Neurocomputing 70, 489–501.

Iqbal, S.M., Hussain, A., Ali, N., Hussain, W., Hussain, H., Hussain, S., Shah, S.Y.A., Nyakilla, E.E., 2023. Experimental evaluation of different influencing parameters on cutting transport performance (CTP) in deviated wells. Geosystems and Geoenvironment 2, 100110. https://doi.org/10.1016/j.geogeo.2022.100110.

Jimmy, D., Wami, E., Ogba, M.I., 2022. Cuttings lifting coefficient model: a criteria for cuttings lifting and hole cleaning quality of mud in drilling optimization. In: Society of Petroleum Engineers - SPE Nigeria Annual International Conference and Exhibition, NAIC 2022. OnePetro. https://doi.org/10.2118/212004-MS.

Khosravanian, R., Aadnøy, B.S., 2021. Wellbore hydraulics and hole cleaning: optimization and digitalization. In: Khosravanian, R., Aadnøy, B.S. (Eds.), Methods for Petroleum Well Optimization: Automation and Data Solutions. Elsevier, pp. 1–538. https://doi.org/10.1016/B978-0-323-90231-1.00007-8.

LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. Nature 521, 436–444. https://doi.org/10.1038/nature14539.

Li, M., Zhang, G., Xue, J., Li, Y., Tang, S., 2014. Prediction of the wall factor of arbitrary particle settling through various fluid media in a cylindrical tube using artificial intelligence. Sci. World J. 2014. https://doi.org/10.1155/2014/438782.

Liang, N., Huang, G., Member, S., Saratchandran, P., Member, S., Sundararajan, N., 2006. A fast and accurate online sequential learning algorithm for feedforward networks, 17, 1411–1423.

Lin, C.J., Hong, S.J., 2007. The design of neuro-fuzzy networks using particle swarm optimization and recursive singular value decomposition. Neurocomputing 71, 297–310. https://doi.org/10.1016/J.NEUCOM.2006.12.016.

Luo, Y., Bern, P.A., Chambers, B.D., 1992. Flow-rate predictions for cleaning deviated Wells. In: Drilling Conference - Proceedings, pp. 367–376. https://doi.org/10.2118/23884-MS.

Luo, Y., Bern, P.A., Chambers, B.D., Kellingray, D.S., 1994. Simple charts to determine hole cleaning requirements in deviated Wells. In: Drilling Conference - Proceedings of the IADC/SPE Drilling Conference, pp. 499–505. https://doi.org/10.2118/27486-MS.

Mahmoud, A.A., Elzenary, M., Elkatatny, S., 2020. New hybrid hole cleaning model for vertical and deviated wells. Journal of Energy Resources Technology, Transactions of the ASME 142. https://doi.org/10.1115/1.4045169/1065573.

Maimon, O., Rokach, L., 2005. Data mining and knowledge discovery handbook, second. In: Data Mining and Knowledge Discovery Handbook. Springer. https://doi.org/10.1007/0-387-25465-x_2.

Martins, A.L., Aragão, A.F.L., Aranha, P.E., Folsta, M.G., Waldmann, A.T.A., 2011. Well construction hydraulics in challenging environments. In: SPE/IADC Drilling

Conference 2, pp. 782–793. https://doi.org/10.2118/140145-ms.
Mitchell, B., 1992. Advanced Oilwell Drilling Engineering Handbook, a Textbook. Society of Petroleum Engineers of the AIME, Dallas, TX, USA.
Moody, J., Darken, C.J., 1989. Fast learning in networks of locally-tuned processing units. Neural Comput. 1, 281–294. https://doi.org/10.1162/NECO.1989.1.2.281.
Nazari, T., Hareland, G., Azar, J.J., 2010. Review of cuttings transport in directional well drilling: systematic approach. In: Society of Petroleum Engineers Western North American Regional Meeting 2010 - in Collaboration with the Joint Meetings of the Pacific Section AAPG and Cordilleran Section GSA, vol. 1, pp. 108–122. https://doi.org/10.2118/132372-MS.
Newitt, D.M., Richardson, J.F., Gliddon, B.J., 1961. Hydraulic conveying of solids in vertical pipes. Trans. Inst. Chem. Eng. 39, 93–100.
Onwunalu, Jérôme E., Durlofsky, Louis J., Onwunalu, J.E., Durlofsky, L.J., 2009. Application of a particle swarm optimization algorithm for determining optimum well location and type. Comput. Geosci. 14, 183–198. https://doi.org/10.1007/S10596-009-9142-1.
Ozbayoglu, E.M., Miska, S.Z., Reed, T., Takach, N., 2002. Analysis of bed height in horizontal and highly-inclined wellbores by using artificial neural networks. In: SPE International Thermal Operations and Heavy Oil Symposium and International Horizontal Well Technology Conference. OnePetro, Alberta, Canada. https://doi.org/10.2118/78939-MS.
Park, J., Sandberg, I.W., 1991. Universal approximation using radial-basis-function networks. Neural Comput. 3, 246–257.
Pedersen, M.E.H., Chipperfield, A.J., 2010. Simplifying particle swarm optimization. Appl. Soft Comput. 10, 618–628. https://doi.org/10.1016/J.ASOC.2009.08.029.
Piroozian, A., Ismail, I., Yaacob, Z., Babakhani, P., Ismail, A.S.I., 2012. Impact of drilling fluid viscosity, velocity and hole inclination on cuttings transport in horizontal and highly deviated wells. J. Pet. Explor. Prod. Technol. 2, 149–156. https://doi.org/10.1007/S13202-012-0031-0/FIGURES/6.
Rashidi, S., Mehrad, M., Ghorbani, H., Wood, D.A., Mohamadian, N., Moghadasi, J., Davoodi, S., 2021. Determination of bubble point pressure & oil formation volume factor of crude oils applying multiple hidden layers extreme learning machine algorithms. J. Pet. Sci. Eng. 202. https://doi.org/10.1016/j.petrol.2021.108425.
Rasi, M., Member, S., 1994. Hole cleaning in large, high-angle wellbores. In: Drilling Conference - Proceedings, pp. 299–310. https://doi.org/10.2118/27464-MS.
Rezaei, F., Jafari, S., Hemmati-Sarapardeh, A., Mohammadi, A.H., 2022. Modeling of gas viscosity at high pressure-high temperature conditions: integrating radial basis function neural network with evolutionary algorithms. J. Pet. Sci. Eng. 208, 109328. https://doi.org/10.1016/J.PETROL.2021.109328.
Robinson, L., Morgan, M., 2004. Effect of hole cleaning on drilling rate and performance. In: AADE 2004 Drilling Fluids Conference, pp. 6–7.
Rooki, R., Doulati Ardejani, F., Moradzadeh, A., 2014. Hole cleaning prediction in foam drilling using artificial neural network and multiple linear regression. Geomaterials 04 47–53. https://doi.org/10.4236/gm.2014.41005.
Rumelhart, D.E., Hinton, G.E., Williams, R.J., 1986. Learning representations by back-propagating errors. Nature 323, 533–536.
Sharma, A., Onwubolu, G., 2009. Hybrid particle swarm optimization and GMDH system. Studies in Computational Intelligence 211, 193–231. https://doi.org/10.1007/978-3-642-01530-4_5/CO VER.
Sifferman, T.R., Myers, G.M., Haden, E.L., Wahl, H.A., 1974. Drill cutting transport in full scale vertical annuli. J. Petrol. Technol. 26, 1295–1302. https://doi.org/10.2118/4514-PA.
Ulker, E., Sorgun, M., 2016. Comparison of computational intelligence models for cuttings transport in horizontal and deviated wells. J. Pet. Sci. Eng. 146, 832–837. https://doi.org/10.1016/J.PETROL.2016.07.022.
Unegbu, C.T., 2010. Hole Cleaning and Hydraulics. University of Stavanger, Norway.
VMP, 2025. Horizontal Completions Market Insights [WWW Document]. URL. https://www.verifiedmarketreports.com/product/horizontal-completions-market/ (accessed 9.8.25).
Wang, F., Xu, H., Wang, S., Deng, J., Wang, Y., 2024. Fluid flow and efficient development technologies in unconventional reservoirs: state-of-the-art methods and future perspectives. Advances in Geo-Energy Research 12, 237–240. https://doi.org/10.46690/AGER.2024.06.07.
Wood, D.A., 2023. Variable interaction empirical relationships and machine learning provide complementary insight to experimental horizontal wellbore cleaning results. Advances in Geo-Energy Research 9, 172–184. https://doi.org/10.46690/ager.2023.09.05.
Xie, T., Yu, H., Wilamowski, B., 2011. Comparison between traditional neural networks and radial basis function networks. In: Proceedings - ISIE 2011: 2011 IEEE International Symposium on Industrial Electronics, pp. 1194–1199. https://doi.org/10.1109/ISIE.2011.5984328.
Yeo, L., Feng, Y., Seibi, A., Temani, A., Liu, N., 2021. Optimization of hole cleaning in horizontal and inclined wellbores: a study with computational fluid dynamics. J. Pet. Sci. Eng. 205, 108993. https://doi.org/10.1016/J.PETROL.2021.108993.
Yeom, C.-U., Kwak, K.-C., 2017. Short-term electricity-load forecasting using a TSK-based extreme learning machine with knowledge representation. Energies (Basel) 10, 1613.
Yeu, W.J., Katende, A., Sagala, F., Ismail, I., 2019. Improving hole cleaning using low density polyethylene beads at different mud circulation rates in different hole angles. J. Nat. Gas Sci. Eng. 61, 333–343. https://doi.org/10.1016/J.JNGSE.2018.11.012.
Yu, H., Xie, T., Paszczyński, S., Wilamowski, B.M., 2011. Advantages of radial basis function networks for dynamic system design. IEEE Trans. Ind. Electron. 58, 5438–5450. https://doi.org/10.1109/TIE.2011.2164773.
Yu, M., Takach, N.E., Nakamura, D.R., Shariff, M.M., 2007. An experimental study of hole cleaning under simulated downhole conditions. In: Proceedings - SPE Annual Technical Conference and Exhibition, vol. 2, pp. 1186–1199. https://doi.org/10.2118/109840-MS.
Zeng, N., Zhang, H., Liu, W., Liang, J., Alsaadi, F.E., 2017. A switching delayed PSO optimized extreme learning machine for short-term load forecasting. Neurocomputing 240, 175–182. https://doi.org/10.1016/j.neucom.2017.01.090.
Zhang, C., Bengio, S., Hardt, M., Recht, B., Vinyals, O., 2021. Understanding deep learning (still) requires rethinking generalization. Commun. ACM 64, 107–115.
Zhang, Q., Gao, H., Zhan, Z.H., Li, J., Zhang, H., 2023. Growth optimizer: a powerful metaheuristic algorithm for solving continuous and discrete global optimization problems. Knowl. Base Syst. 261, 110206. https://doi.org/10.1016/J.KNOSYS.2022.110206.
Zhu, N., Ding, S.-D., Shi, X.-L., Huang, W.-J., Gao, D.-L., 2023. Cuttings transport: back reaming analysis based on a coupled layering-sliding mesh method via CFD. Pet. Sci. https://doi.org/10.1016/J.PETSCI.2023.06.009.