

SPE/IADC-233045-MS

Physics-Based AI/ML Approach for Assessing and Predicting Hydraulics and Hole Cleaning in Deviated Boreholes—A High-Temperature High-Pressure Case Study

Rahman Ashena and Vassilios C. Kelessidis, Smart Energy Data Analytics-Group, Houston

Copyright 2026, SPE/IADC Asia Pacific Drilling Technology Conference and Exhibition DOI [10.2118/233045-MS](https://doi.org/10.2118/233045-MS)

This paper was prepared for presentation at the SPE/IADC Asia Pacific Drilling Technology Conference and Exhibition held in Bali, Indonesia, 5 – 6 August 2026.

This paper was selected for presentation by an SPE/IADC program committee following review of information contained in an abstract submitted by the author(s). Contents of the paper have not been reviewed by the Society of Petroleum Engineers or the International Association of Drilling Contractors and are subject to correction by the author(s). The material does not necessarily reflect any position of the Society of Petroleum Engineers or the International Association of Drilling Contractors, its officers, or members. Electronic reproduction, distribution, or storage of any part of this paper without the written consent of the Society of Petroleum Engineers or the International Association of Drilling Contractors is prohibited. Permission to reproduce in print is restricted to an abstract of not more than 300 words; illustrations may not be copied. The abstract must contain conspicuous acknowledgment of SPE/IADC copyright.

Abstract

Accurate hydraulics and hole-cleaning assessment in high-pressure, high-temperature (HPHT) drilling environments remains challenging due to the combined effects of temperature, pressure, non-Newtonian fluid behavior, and drillstring geometry. This study presents an innovative physics-based AI/ML methodology for modeling drilling hydraulics and hole cleaning using a Herschel–Bulkley (HB) fluid model while accounting for pressure- and temperature-dependent fluid properties and tool-joint effects.

The data-driven methodology was applied to a real deep well exceeding 12,000 ft in depth deviated up to 31° and with static formation temperatures of approximately 220°F. Hydraulics and hole-cleaning performance were evaluated using 17 key performance indicators (KPIs) or criteria. Nine (9) machine-learning algorithms were investigated, with two of them, Extra Trees and Random Forest, providing the strongest prediction performance. These models were integrated with physics-based hydraulics calculations to improve assessment capability while maintaining consistency with established drilling physics.

The results demonstrate successful assessment of hydraulics and hole-cleaning condition behavior, achieving high validation accuracy. The results also indicate that incorporating additional high-quality field and operational data can further improve model accuracy and robustness. The study demonstrates the potential of integrating physics-based hydraulics modeling with AI/ML techniques to provide reliable and practical evaluations for challenging HPHT drilling operations.

Introduction

High-pressure, high-temperature (HPHT) and geothermal drilling operations present significant challenges for maintaining effective hydraulics and hole cleaning, particularly in deep and deviated wellbores. Increasing temperature and pressure can substantially alter drilling-fluid properties, such as density, and rheological behavior, which directly affect pressure losses, equivalent circulating density (ECD), cuttings transport efficiency, and overall wellbore hydraulics. These effects become particularly important in deviated sections, where gravitational settling of cuttings and reduced transport efficiency can lead to cuttings accumulation and poor hole cleaning.

The present study was conducted on a real deep well with a measured depth exceeding 12,000 ft (TVD >10,760 ft), a maximum inclination of approximately 31°, and downhole temperatures reaching approximately 220°F. The combination of significant depth, elevated temperature, and wellbore inclination provides a challenging environment for evaluating hydraulics and hole-cleaning behavior. Poor hole cleaning can contribute to excessive torque and drag, pack-off, increased ECD, and stuck-pipe events, making accurate prediction an important component of safe and efficient drilling operations.

A physics-based hydraulics and hole-cleaning model was developed using the Herschel–Bulkley (HB) fluid model. The selection and implementation of the HB model are supported by previous work by Ashena, including the evaluation of HB pressure-estimation methods for annular flow and the detailed derivation and comparison of HB laminar-flow equations for annuli (Ashena, 2024, 2025) or Kelessidis et al. (2006, 2011). These studies provide a foundation for applying the HB model to drilling hydraulics and for improving pressure-loss prediction under realistic annular-flow conditions.

In the present work, the physics-based model considers the effects of temperature and pressure on fluid viscosity, density, and rheological properties, together with the effects of tool-joint geometry on annular hydraulics. The model also incorporates the influence of cuttings concentration, rate of penetration (ROP), borehole inclination, and other drilling parameters on cuttings transport and hole-cleaning performance. This integrated approach allows the coupled effects of HPHT conditions and drilling parameters to be represented more realistically than approaches based only on surface fluid properties or simplified annular geometries.

The physics-based model was subsequently integrated with nine (9) AI/ML algorithms, including artificial neural-network-based modeling, to develop predictive models for hydraulics and hole-cleaning performance. The use of neural networks and other AI/ML techniques in petroleum engineering has been previously demonstrated by Ashena and Thonhauser, including their 2024 work on artificial neural networks for geoscience and petroleum applications. Similar AI/ML approaches have also been applied to well-control problems, including data-driven early kick prediction, as documented in Chapter 13, “Early Kick Prediction by Data-Driven AI/ML Techniques,” of *The Practical Handbook of Well Control* (Ashena et al., 2025); or for drilling optimization (Bataee et al., 2010; Ashena et al., 2010;; Ashena and Thonhauser, 2015; Ashena et al., 2021).

The AI/ML models in this study were evaluated against the physics-based hydraulics including the hole-cleaning model and available well data. The results demonstrate convincing predictive performance while maintaining consistency with the underlying physical behavior of the drilling system. The integration of physics-based modeling with AI/ML therefore provides a practical framework for predicting hydraulics and hole-cleaning performance under challenging HPHT conditions, while providing a pathway for further improvement as additional field and operational data become available.

Theory

In this work, nine (9) advanced machine learning algorithms were evaluated for evaluating hole-cleaning performance. The algorithms used consist of: Extra Trees, Random Forest (RF), Support Vector Machine Regression (SVM), Extreme Gradient Boosting (XGBoost), HistGradientBoosting (HGB), K-Nearest Neighbors Regression (KNN), Logistic Regression (LR), Gaussian Naive Bayes (GNB), and Artificial Neural Networks (ANN). They were selected because of their ability to model complex nonlinear relationships among drilling parameters while maintaining strong predictive performance and computational efficiency.

Extra Trees

Extra Trees (Extremely Randomized Trees) is an ensemble machine-learning method that combines multiple randomized decision trees to improve prediction accuracy and reduce overfitting. Unlike conventional

decision trees, it introduces additional randomness by selecting random features and split points when building each tree. For regression, the final prediction is generally the average of all tree predictions. Extra Trees is particularly effective for nonlinear, complex engineering problems involving multiple input variables and can be applied to drilling, well-control, and petroleum-engineering prediction tasks.

For a regression problem, the predictions from all trees are typically averaged, while for classification, the trees' predictions are combined through voting. The additional randomization reduces the correlation between individual trees and can improve generalization and computational efficiency. A simplified regression formulation is:

$$\hat{y} = \frac{1}{N} \sum_{i=1}^N T_i(x) \quad \text{Eq. 1}$$

where N is the number of decision trees, $T_i(x)$ is the prediction from the i -th tree for input vector x , and $\hat{y}$ is the final Extra Trees prediction.

Random Forest Regression (RF)

Random Forest is an ensemble learning technique that combines a large number of decision trees. Each tree is trained using a bootstrap sample of the original dataset, while a random subset of features is considered at each split. Random Forest generally improves prediction accuracy and robustness compared to a single decision tree by reducing model variance and sensitivity to noise. The final prediction is obtained by averaging the outputs from all trees.

Support Vector Regression (SVM)

Support Vector Regression extends the principles of Support Vector Machines to regression problems. The objective is to determine a function that approximates the target values while maintaining maximum margin around the prediction function. The optimization problem can be expressed as:

$$f(x) = \sum_{i=1}^n (\alpha_i - \alpha_i^*) K(x_i, x) + b \quad \text{Eq. 2}$$

For kernel:

$$K(x_i, x) = \exp(-\gamma \|x_i - x\|^2) \quad \text{Eq. 3}$$

where α_i, α_i^* are SVR coefficients, K is the kernel function, b is the bias term, and γ controls the influence of the training data points.

Note that Kernel functions allow SVR to capture nonlinear relationships between drilling parameters and standpipe pressure. However, computational requirements increase substantially as dataset size grows.

Extreme Gradient Boosting (XGBoost)

Extreme Gradient Boosting (XGBoost) is an enhanced implementation of gradient boosting that incorporates regularization, efficient tree construction, and advanced optimization techniques. Regularization improves generalization performance and reduces the risk of overfitting. Due to its computational efficiency and predictive capability, XGBoost has become one of the most widely used algorithms for structured engineering datasets. The objective function consists of a training loss term and a regularization term:

$$Obj = L(\theta) + \Omega(\theta) \quad \text{Eq. 4}$$

where $(L(\theta))$ represents the prediction loss and $(\Omega(\theta))$ penalizes model complexity.

HistGradientBoosting (HGB)

HGB is a gradient-boosting machine-learning algorithm that builds a sequence of decision trees, where each new tree is trained to reduce the errors of the previous trees. It uses histogram-based binning of continuous features, which makes the training process faster and more memory-efficient than conventional gradient boosting. HGB can effectively capture nonlinear relationships and interactions between variables and is suitable for complex engineering prediction problems.

K-Nearest Neighbors Regression (KNN)

K-Nearest Neighbors is an instance-based learning algorithm that predicts the target value using the average response of the (k) nearest observations in feature space. KNN is simple to implement and can capture nonlinear behavior; however, prediction accuracy is highly dependent on the choice of distance metric and the number of neighbors selected.

Logistic Regression (LR)

Logistic Regression is a supervised machine-learning method used primarily for binary classification. It estimates the probability of an outcome using a logistic (sigmoid) function and classifies observations based on a selected probability threshold. It is widely used for predicting events such as engineering incidents, equipment failures, or other yes/no engineering outcomes.

This model is a simple classification model that always predicts the most frequently occurring class in the training dataset. It requires no feature learning or complex calculations and serves as a basic reference model for comparing the performance of more advanced machine-learning algorithms.

Gaussian Naive Bayes (GNB)

GNB is a probabilistic classification algorithm based on Bayes' theorem. It assumes that the input features are conditionally independent and follow a Gaussian (normal) distribution within each class. GNB estimates the probability of each class from the training data and assigns a new observation to the class with the highest probability. It is computationally efficient and suitable for fast classification of engineering datasets. The basic Bayes equation is:

$$P(C|X) = \frac{P(X|C)P(C)}{P(X)} \quad \text{Eq. 5}$$

where $P(C|X)$ is the probability of class C given the input data X .

Artificial Neural Networks (ANN)

An Artificial Neural Network (ANN) is a machine-learning model designed to solve complex nonlinear engineering problems by processing information through interconnected artificial neurons. The most commonly used structure in petroleum engineering is the Feed-Forward Artificial Neural Network (FF-ANN), consisting of an input layer, one or more hidden layers, and an output layer. Information propagates forward through the network, with the neurons applying weighted connections and activation functions to learn relationships between input variables and the desired output (see Fig. 1).

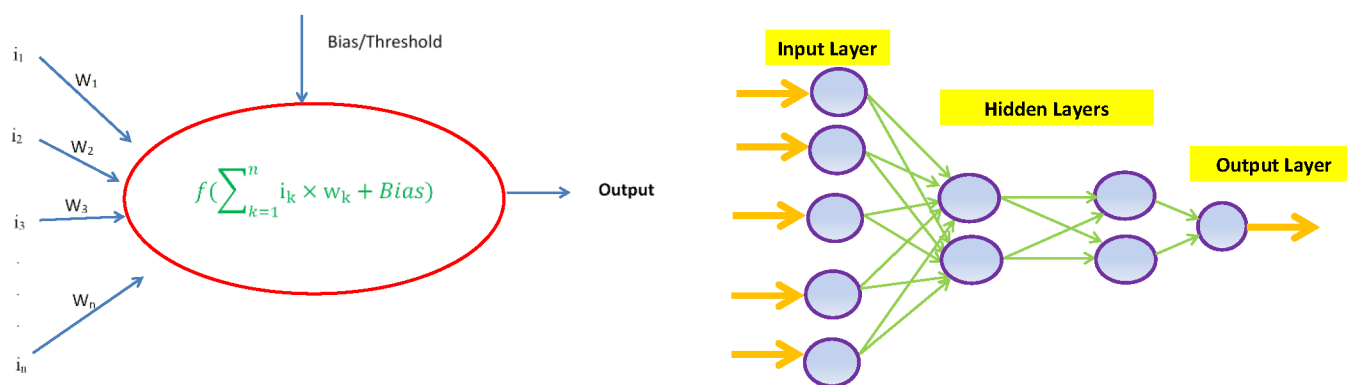


Figure 1—Typical Feed Forward Artificial Neural Networks (FF-ANN)

Well Data and Data Processing

The study was conducted on an actual HPHT well with a measured depth exceeding 12,000 ft, a maximum inclination of approximately 31° from vertical, and a static downhole temperature of approximately 220 °F, with the temperature profile shown in Fig. 2. Because static and circulating temperatures can differ significantly during drilling, the dynamic downhole temperature was calculated using the dynamic temperature-modeling equations implemented in the proposed methodology. This provided a more representative temperature profile for the real-time hydraulics and hole-cleaning calculations.

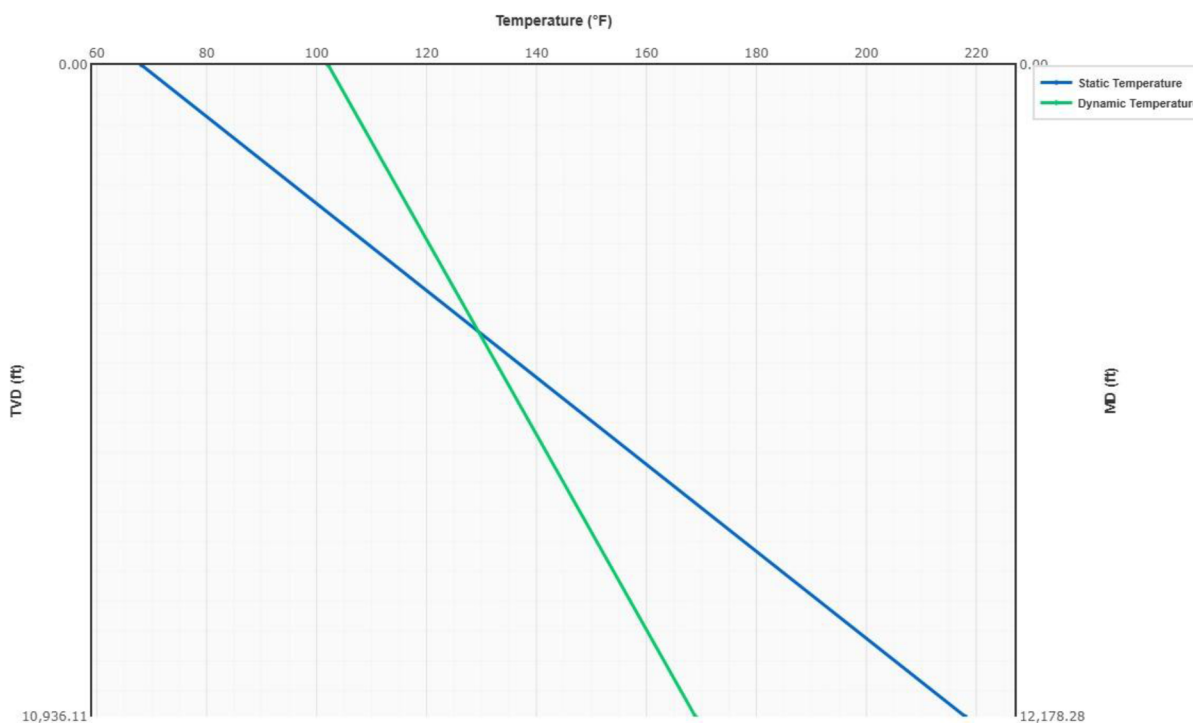


Figure 2—Well temperature profiles versus depth and xOptim™DrL -simulated dynamic temperatures

The well was drilled with a BHA incorporating a mud motor and MWD sensors under actual HPHT operating conditions. The formation consisted primarily of sandstone and siltstone, with intervals containing clay, creating challenging conditions for cuttings transport and hole cleaning. The well also experienced hole-cleaning difficulties and stuck-pipe events, providing a practical basis for evaluating the proposed methodology.

Real-time drilling and downhole data were acquired at a 5-second sampling frequency. An automated data-processing and quality-control (QC) procedure was applied to identify and remove or correct data outliers and anomalous measurements. The processed dataset passed the QC criteria established by the innovative data-processing system, providing a consistent dataset for subsequent modeling. Drilling activity recognition was performed automatically at 1-minute intervals to identify and classify the operational state of the well (Fig. 3).

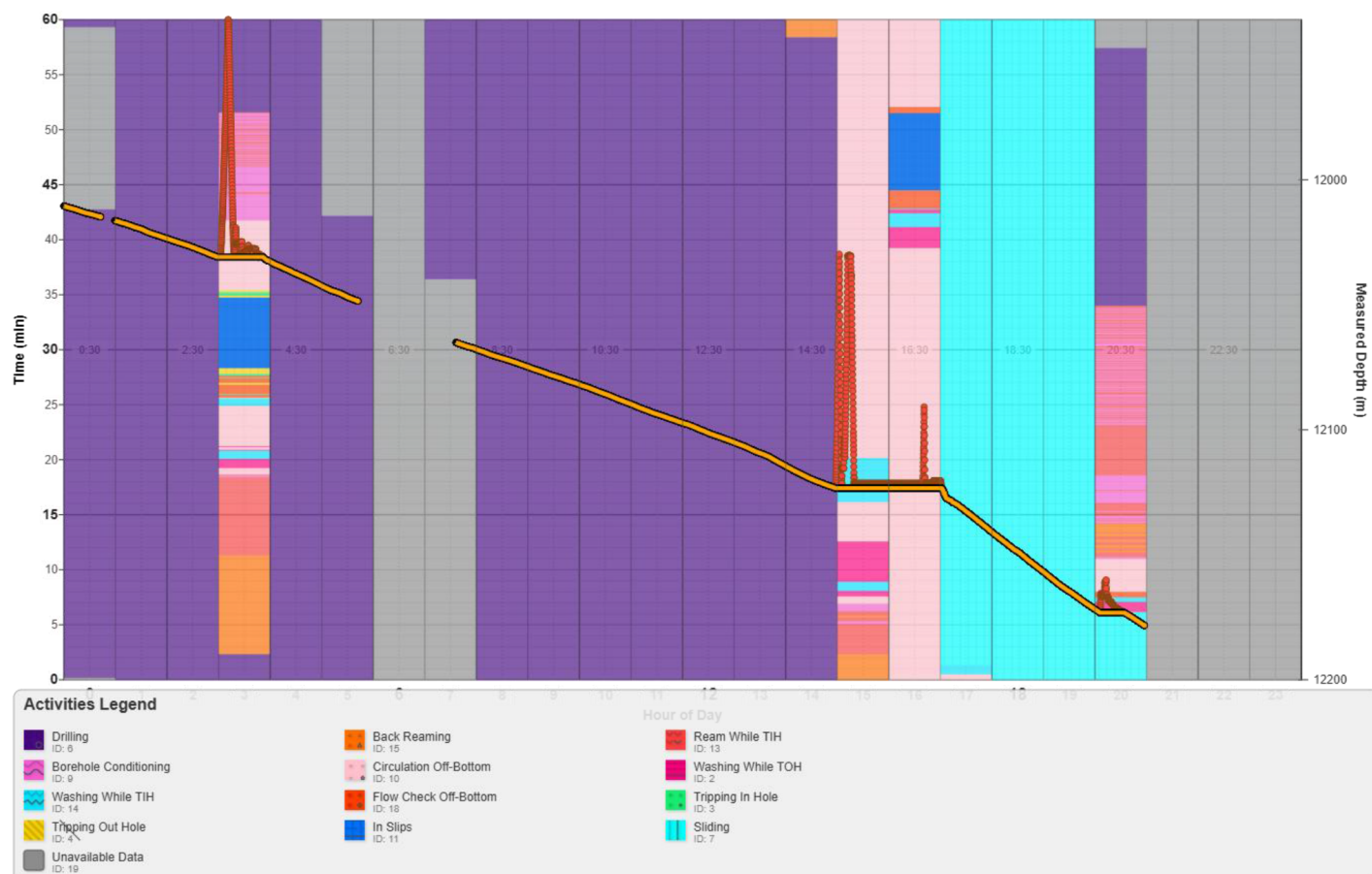


Figure 3—Hourly-based activity recognition using xOptim™DrL simulation

For each 5-second data point, real-time tool xOptim™DrL's Hydraulics (Hydro) and Torque and Drag (TnD) analyses were performed. The dynamic temperature model was integrated into these calculations to account for the evolving downhole thermal conditions and their effects on fluid properties and hydraulics. The resulting high-frequency physics-based outputs, together with the QC-processed drilling data, were subsequently used for AI/ML modeling and evaluation of hydraulics and hole-cleaning performance.

Results and Discussion

The 17 hole-cleaning criteria comprising the HC-Index were estimated using the developed physics-based models. These criteria are of flow regimes in the two conduits of pipe and annulus; percentage of the bit pressure loss vs total pressure loss; HSI values, in-situ cuttings concentration in the annulus, annular velocity and rate, low bed thickness, torque and drag values, and lifting ratio, etc. In parallel, nine AI/ML models were applied to independently estimate the same criteria and cross-check the physics-based results. The performance of the different models was evaluated using selected key performance indicators (KPIs), providing a quantitative comparison of their accuracy, consistency, and predictive capability.

Fig. 4 Fig. 5 to Fig. 7 illustrate four (4) of the 17 KPIs for this case study well: flow regime in the two conduits of pipe and annulus; percentage of the bit pressure loss vs total pressure loss; HSI values, and in-situ cuttings concentration in the annulus.,

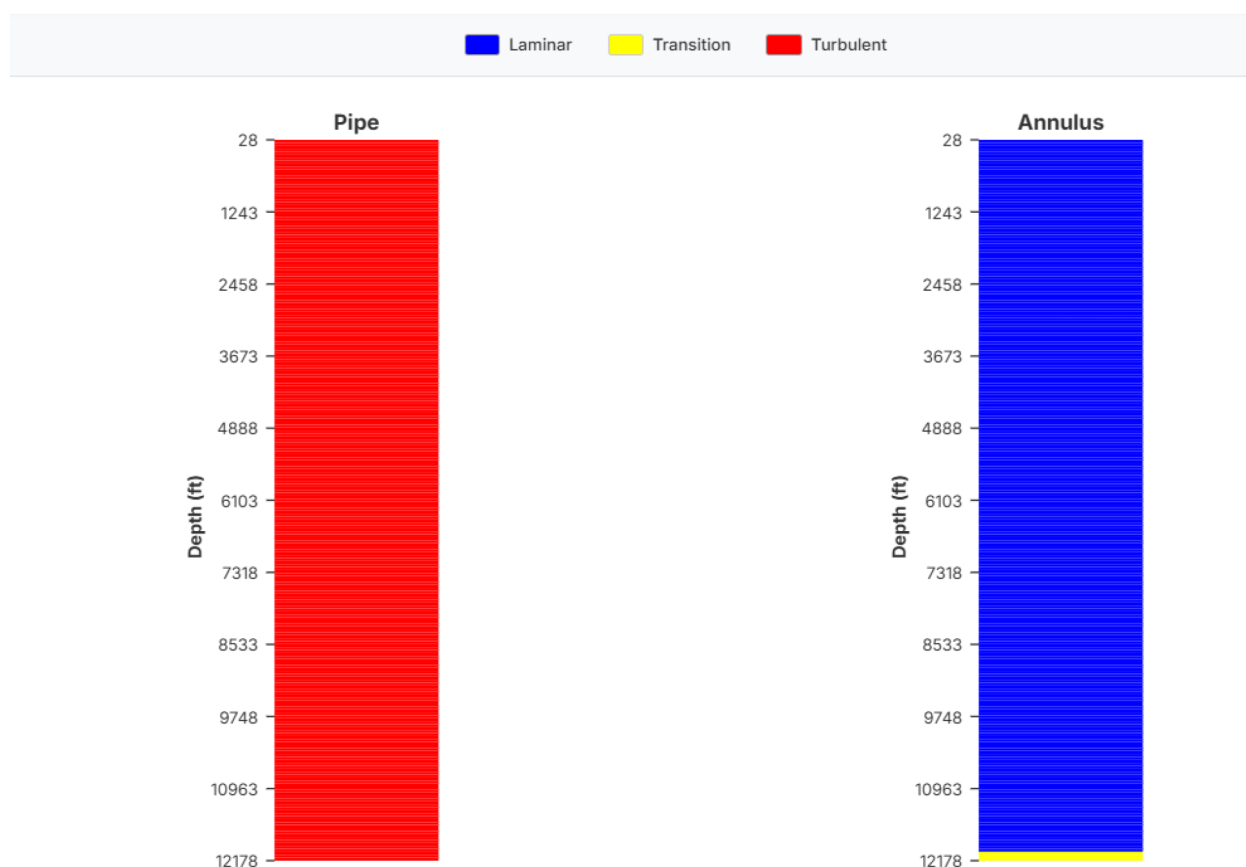


Figure 4—Flow regime recognition in two conduits of drill-string and annulus, a KPI

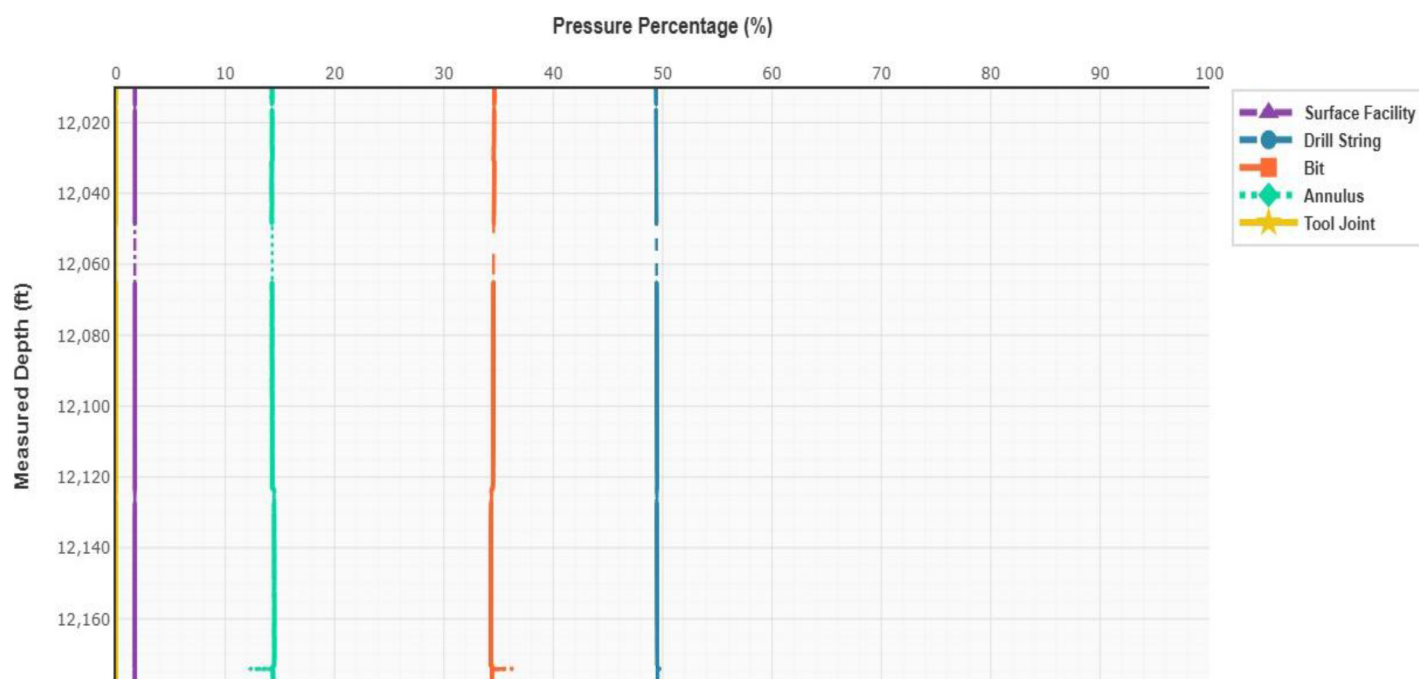


Figure 5—The pressure distribution percentage of the total pressure drop (TDP) indirectly used as one of the 17 KPIs

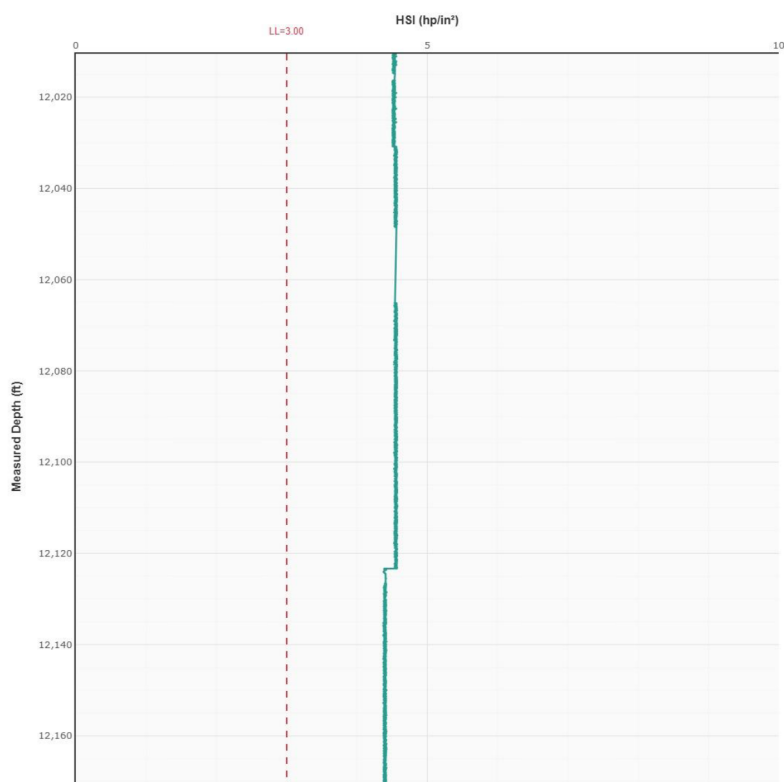


Figure 6—One of the 17 KPIs for HC-index

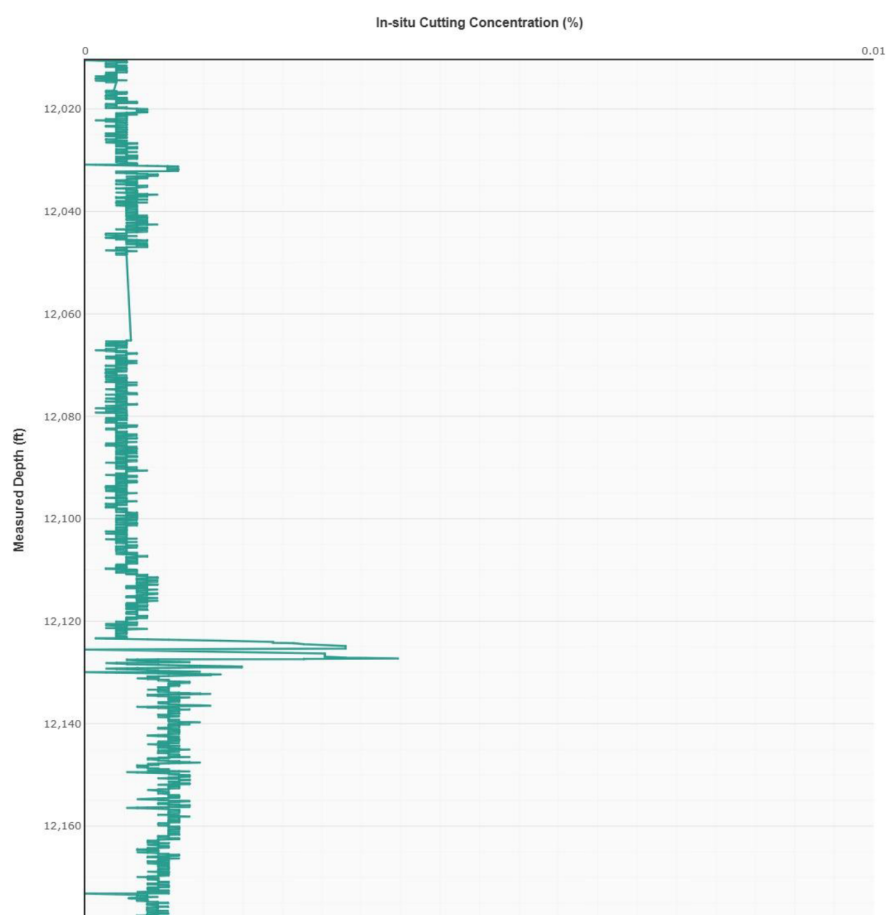


Figure 7—In-situ cuttings concentration as one of Hydro's KPIs

For the classification models, ROC (Receiver Operating Characteristic) curves and corresponding AUC (Area Under the ROC Curve) values were also used to further evaluate and compare their classification performance (Fig. 8 and Fig. 9). ROC curves are used to evaluate and compare classification machine-learning models. The curve plots True Positive Rate (TPR) or sensitivity on the y-axis, and False Positive Rate (FPR) on the x-axis. It is generated by varying the model's classification threshold. The AUC summarizes the model's ability to distinguish between two classes, with a value of 1.0 indicating perfect discrimination and 0.5 representing performance equivalent to random classification. Note that "Chance" is a statistical baseline used for comparison. It represents the performance expected from random guessing, typically with an ROC-AUC of 0.50 for a balanced binary classification problem.

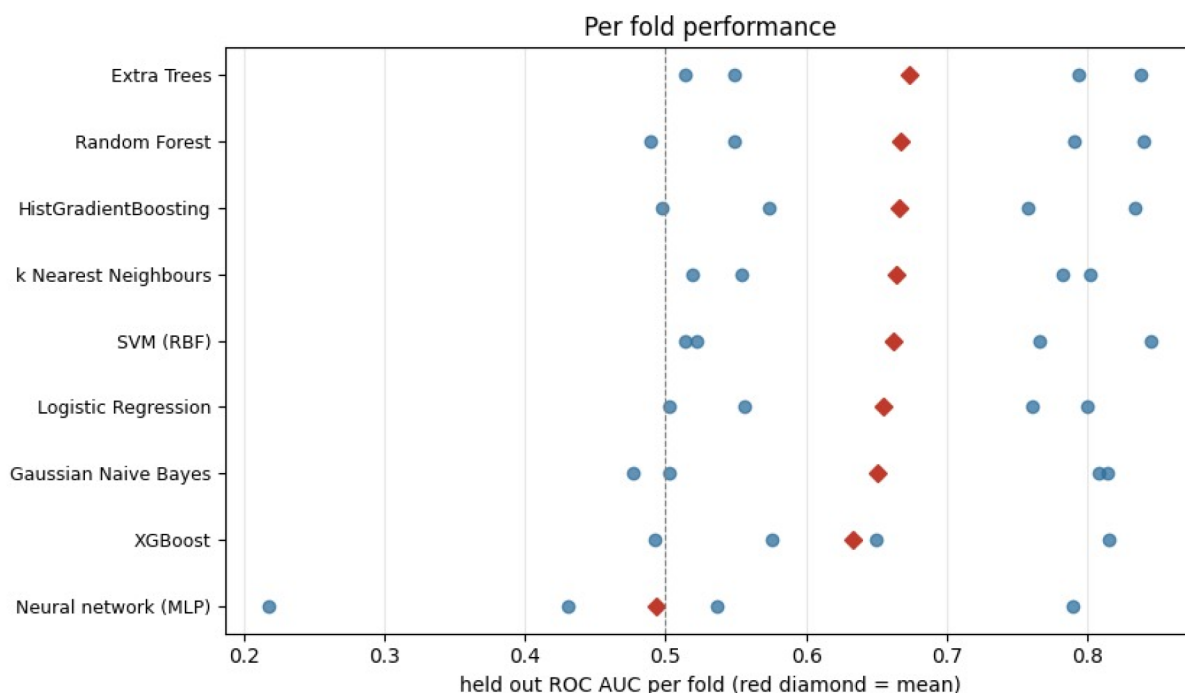


Figure 8—Comparison of different ML algorithms used for estimating hole-cleaning performance (HC-index)

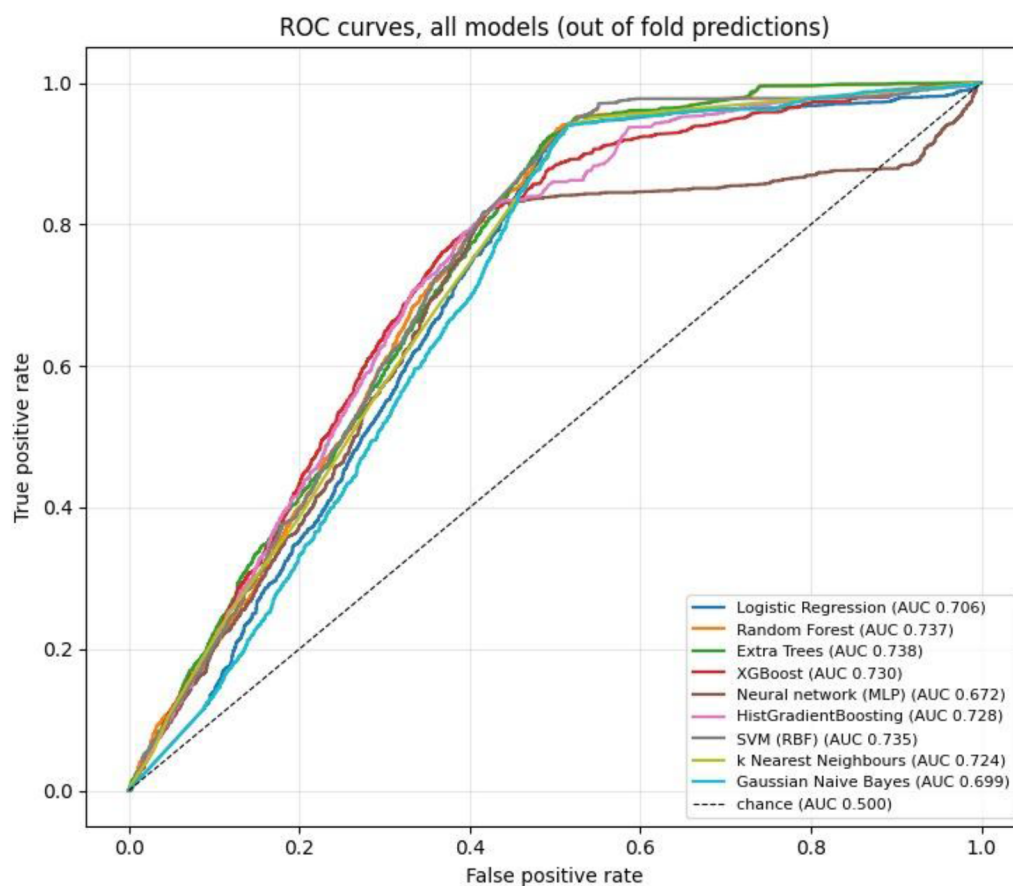


Figure 9—Another comparison of different ML algorithms using receiver operator characteristic (ROC) curves

Based on the combined results, the final HC-Index was calculated and is presented as an integrated measure of hole-cleaning performance. The results indicate that hole cleaning in the studied well was

modelled well by the AI/ML method. The cleaning was generally good to excellent (Fig. 10), with the estimated HC-Index criteria showing satisfactory hole-cleaning conditions over the evaluated well interval. The agreement between the physics-based and AI/ML results provides an additional cross-check of the assessment and supports the reliability of the final HC-Index determination.

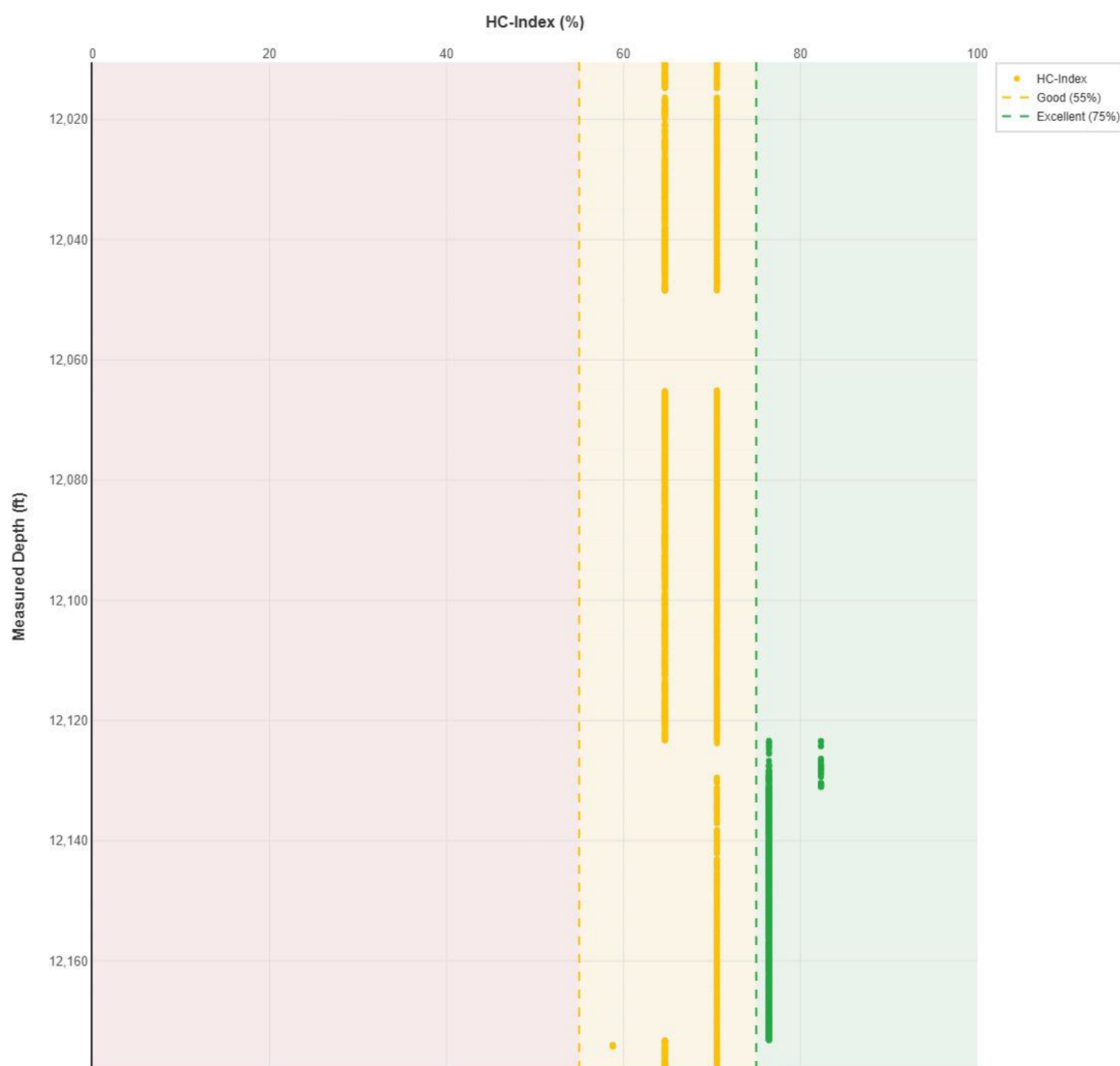


Figure 10—Hole-cleaning index or performance (HC-index) estimation as the result of 17 Hydro's KPIs

Summary and Conclusions

This study presented an integrated physics-based and AI/ML approach for evaluating hole-cleaning performance and estimating the HC-Index of an HPHT well. A total of 17 hole-cleaning criteria were evaluated using the developed physics-based models, while nine AI/ML models were applied independently to cross-check and validate the results. Model performance was assessed using selected key performance indicators (KPIs), while ROC curves and AUC values were used for classification-based models.

1. The comparative results showed that Extra Trees (ET) and Random Forest (RF) provided the strongest overall performance among the evaluated AI/ML approaches for the studied HPHT well, providing reliable predictions of the hole-cleaning criteria. The agreement between the physics-based and AI/ML results provided an additional cross-check of the assessment. The final HC-Index indicated that

hole-cleaning performance in the studied well was generally good to excellent over the evaluated interval.

2. Overall, the results demonstrate the value of combining physics-based modeling with AI/ML techniques, particularly ET and RF, to provide a comprehensive and quantitative assessment of hole-cleaning performance in complex HPHT drilling conditions.

Reference

- Ashena, R. (2024). Recognizing the Most Accurate Herschel-Bulkley Fluid Pressure Estimation Method for Annulus: An Update to API RP 13D, 7th Edition. *SPE J.* (2024;): doi: <https://doi.org/10.2118/223089-PA>
- Ashena, R. (2025) Stepwise Derivation and Validation of the Herschel–Bulkley Laminar Fluid Flow Equations – in Annulus. *J Petrol Explor Prod Technol* **15**, 100. <https://doi.org/10.1007/s13202-024-01919-z>
- Ashena R. (2025) *The Practical Handbook of Well Control*, Published by Taylor and Francis, ORC Publishing, USA. https://www.routledge.com/The-Practical-Handbook-of-Well-Control/Ashena/p/book/9781032753850?srsId=AfmBOopo88n5b-SbosYm1800b_eDmC5kQn2YokeWC8VAZI_Bfzmt-ses
- Ashena, R., Rabiei, M., Rasouli, V., Mohammadi, A. H., and Mishani, S. (2021). "Drilling Parameters Optimization Using an Innovative Artificial Intelligence Model." *ASME. J. Energy Resour. Technol.* May 2021; **143**(5): 052110. <https://doi.org/10.1115/1.4050050>.
- Ashena, R. and Thonhauser, G. (2015). Application of artificial neural networks in geoscience and petroleum industry. *Chapter 4 of the book "Artificial intelligent approaches in petroleum geosciences"*, Published by Springer International Publishing, Edited by C. Cranganu, H. Luchian and M. E. Breaban, <https://www.springerprofessional.de/en/application-of-artificial-neural-networks-in-geoscience-and-petr/2397732>.
- Ashena, R., Moghadasi, J., Ghalambor, A. et al. (2010). Neural Networks in BHCP Prediction Performed Much Better Than Mechanistic Models. Paper presented at the International Oil and Gas Conference and Exhibition in China, Beijing, China, June 2010. doi: <https://doi.org/10.2118/130095-MS>.
- Bataee, M., Edalatkhah, S., and R. Ashena (2010). Comparison between Bit Optimization Using Artificial Neural Network and Other Methods Base on Log Analysis Applied in Shadegan Oil Field. Paper presented at the International Oil and Gas Conference and Exhibition in China, Beijing, China. doi: <https://doi.org/10.2118/132222-MS>
- Kelessidis, V.C., Maglione, R., Tsamantaki, C., Aspirtakis, Y. (2006) Optimal determination of rheological parameters for Herschel–Bulkley drilling fluids and impact on pressure drop, velocity profiles and penetration rates during drilling, *Journal of Petroleum Science and Engineering*, Volume **53**, Issues 3–4, <https://doi.org/10.1016/j.petrol.2006.06.004>.
- Kelessidis, V. C., Dalamarinis, P., Maglione, R. (2011) Experimental study and predictions of pressure losses of fluids modeled as Herschel–Bulkley in concentric and eccentric annuli in laminar, transitional and turbulent flows, *Journal of Petroleum Science and Engineering*, Volume **77**, Issues 3–4, <https://doi.org/10.1016/j.petrol.2011.04.004>.