



A new approach for real-time prediction of stick–slip vibrations enhancement using model agnostic and supervised machine learning: a case study of Norwegian continental shelf

Behzad Elahifar¹ · Erfan Hosseini²

Received: 5 April 2023 / Accepted: 14 August 2023 / Published online: 30 August 2023
© The Author(s) 2023

Abstract

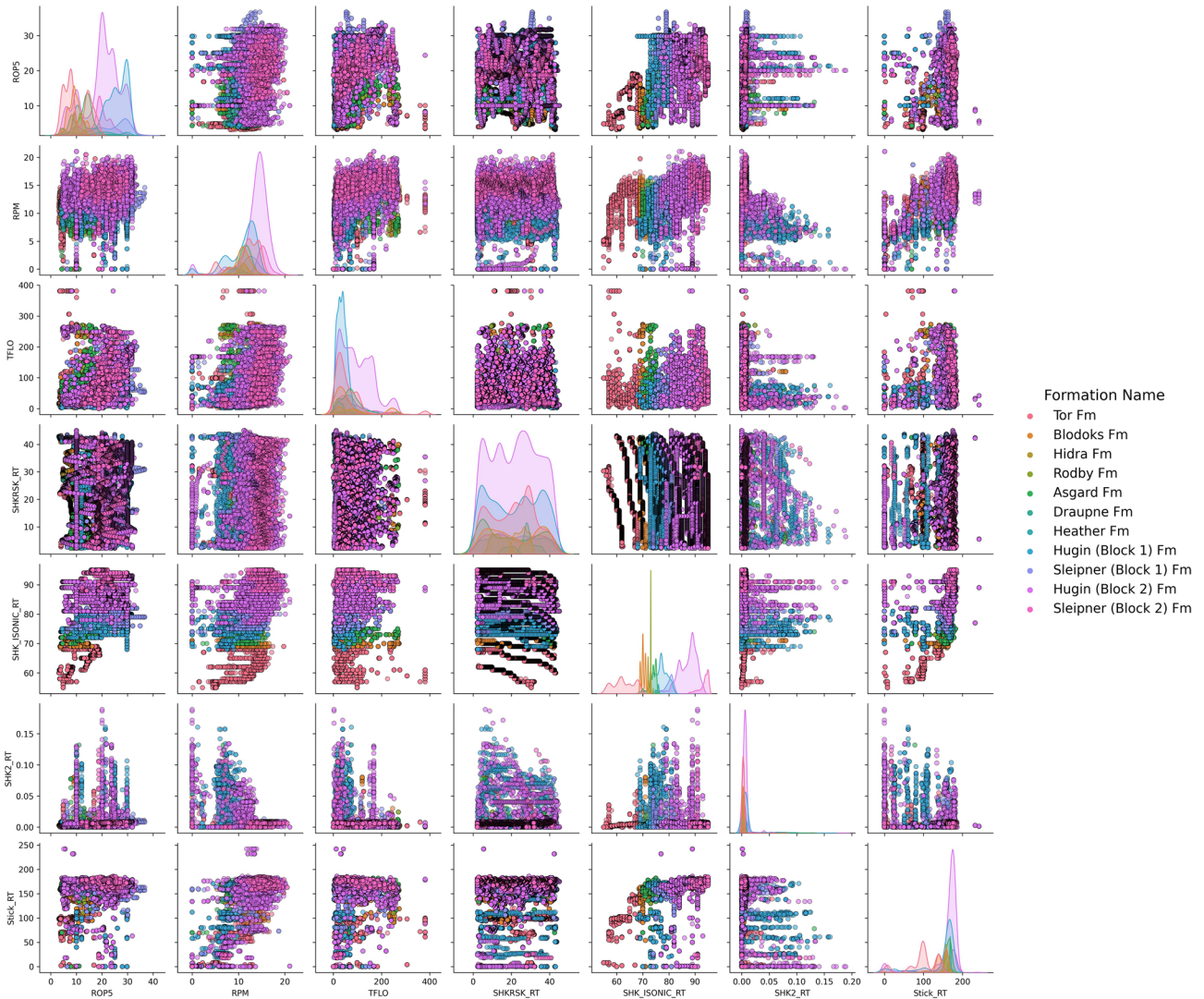
Efficient and safe drilling operations require real-time identification and mitigation of downhole vibrations like stick-slip, which can significantly diminish performance, reliability, and efficiency. This pioneering research introduces a robust machine learning approach combining model-agnostic regression techniques with Bayesian Optimized Extra Trees (BO_ET) to accurately predict stick-slip events in real-time using downhole sensor data. The model is rigorously tested and validated on a substantial offshore dataset comprising over 78,000 data points from a Norwegian continental shelf (NCS) oil field. The key input features encompassing real-time downhole and surface drilling parameters are carefully selected, including critical variables such as collar rotational speed, shock risks, annular pressure, torque, mud flow rate, drill string vibration severity, and other relevant measurements. These parameters offer significant insights into the occurrence of harmful stick-slip vibrations. Among several sophisticated machine learning models, the Extra Trees (ET) algorithm demonstrates superior performance with the lowest errors of 5.5056 revolutions per minute (r/min) Mean Absolute Error (MAE) and 9.9672 r/min Root Mean Square Error (RMSE) on out-of-sample test data. Further hyperparameter tuning of the ET algorithm via Bayesian Optimization dramatically reduces errors down to 0.002156 MAE and 0.024495 RMSE, underscoring the significant innovation and advantages of the proposed approach. By seamlessly incorporating real-time downhole sensor data and drill string mechanics, the model enables reliable identification of stick-slip events as they occur downhole. This grants opportunities to optimize critical drilling parameters including revolutions per minute (RPM), weight-on-bit (WOB), mud flow rates, and more to effectively mitigate stick-slip severity and improve the rate of penetration (ROP). Integrating the approach into automatic driller systems on offshore rigs offers immense benefits for drilling operations through substantially increased efficiency, fewer premature failures, lower costs, and significantly improved productivity and safety. Overall, this research strongly emphasizes the immense transformative potential of advanced data analytics and machine learning in enabling more efficient, economical, and sustainable drilling practices. The proposed model demonstrates clear superiority over existing methods and establishes a robust and reliable platform for real-time stick-slip prediction and mitigation, maximizing drilling performance.

✉ Erfan Hosseini
e.hosseini@niocexp.ir
Behzad Elahifar
behzad.elahifar@ntnu.no

¹ Department of Geoscience and Petroleum, Norwegian University of Science and Technology (NTNU), Trondheim, Norway

² Exploration Directorate, National Iranian Oil Company (NIOC), Tehran, Iran

Graphical abstract



Keywords Stick and slip vibrations · Drilling operations · Machine learning techniques · Optimization of drilling parameters · Automatic driller systems

List of symbols

Symbols

APRS_ARC Annular pressure: downhole annular pressure measurements (Bar)
 ARC_GR_UNC_RT Gamma-ray: a record of the emission of endogenous radioactive elements to the rock, Degrees API (API)
 ATMP Annular temperature: bottom annular space temperature measurement, Degrees Celsius (°C)
 BDTI Job on-bottom time, Hours (HR)

BPOS Traveling block position: height at which the traveling block is located on the mast or derrick, Meters (M)
 BVEL Traveling block velocity: velocity of movement of the block and the hoisting system, Meters per hour (M/HR)
 CRPM_RT Drill string revolutions or collar rotational speed: revolutions of the BHA at the level of the last reamer, Revolutions per minute (RPM)
 DEPT Depth: bit depth in meters or measured depth (MD), Meters (M)

ECD_ARC	Equivalent circulating density: density of the bottom-well fluid, Specific gravity (SG)	APRS_ARC	Annular Pressure
HKLD	Hook-load: measurement of the load on the hook by the working and drilling string, Kilokilogram-force (KKGf)	ATMP	Annular Temperature
MAE	Mean absolute error, Percent (%)	BAR	Barometric Pressure
PDSHKRSK	Power drive shock risk	BDTI	Job On-Bottom Time
RMSE	Root mean square error, Percent (%)	BHA	Bottom Hole Assembly
ROP5	Rate of penetration: drilling progress in distance and time, averaged every 5 feet, Meters per hour (M/HR)	BO	Bayesian Optimization
RPM	Revolution per minute: measurement of the revolutions of the turbine contained in the bottom hole assembly (BHA), Revolutions per minute (RPM)	BO_ET	Bayesian Optimization with Extra Tree model
SHK_ISONIC_RT	Isonic shock	BPOS	Traveling Block Position
SHK2_RT	ARC shock level	BVEL	Traveling Block Velocity
SHKPK_RT	Shock peak: magnitude of the biggest shock experienced by the Power Pulse as the Shock Peak, G (magnitude)	DDR _s	Daily Drilling Reports
SHKRSK_RT	Transverse shock risk	DEGC	Degree Celsius
SPPA	Pump pressure: friction losses in the hydraulic system (Bar)	ECD_ARC	Equivalent Circulating Density
Stick_RT	Stick and slip indicator: torsional vibration indicating stick and slip, Revolutions per minute (RPM)	ET	Extra Trees
SWOB	Surface weight on bit: measurement of the weight exerted by the string on the bit and the formation, Kilokilogram-force (KKGf)	Fm.	Formation
TFLO	Total pump flow: flow rate of drilling mud to the well, Liters per minute (LPM)	HKLD	Hook-load
TQA	Torque: torque exerted by the top drive system derived from string rotation, Kilometers decanewtons (KMN)	HR	Hour
TRPM_RT	Bottom turbine revolutions: measurement of the revolutions of the turbine contained in the BHA to energize downstream components, Revolutions per minute (RPM)	KKGf	Kilo Kilogram Force
		LPM	Liters per Minute
		LSSVM	Least Squares Support Vector Machine
		MLP-NN	Multilayer Perceptron Neural Network
		MNN	Modular Neural Networks
		MWD	Measurements While Drilling
		NPD	Norwegian Petroleum Directorate
		OOIP	Original Oil in Place
		PDSHKRSK	Power Drive Shock Risk
		PSO	Particle Swarm Optimization
		RF	Random Forest
		RMSE	Root Mean Square Error
		ROP	Rate of Penetration
		RPM	Revolutions per Minute
		SHK_ISONIC_RT	Isonic Shock
		SHK2_RT	ARC Shock Level
		SHKPK_RT	Shock Peak
		SHKRSK_RT	Transverse Shock Risk
		SPPA	Pump Pressure
		Stick_RT	Stick and Slip Indicator
		SVM	Support Vector Machine
		SWOB	Surface Weight on Bit
		TFLO	Total Pump Flow
		TQA	Torque
		TRPM_RT	Bottom Turbine Revolutions
		WOB	Weight on Bit
		XGBoost	Extreme Gradient Boosting

Abbreviations

AI	Artificial intelligence
AIC	Akaike information criterion
API	American Petroleum Institute

Introduction

An operation that aims to reach porous and permeable rocks in the basement, which may contain liquid or gaseous hydrocarbons, is known as drilling. Sedimentary basins

are studied geologically and geophysical to determine the location of oil drilling. These studies allow researchers to study subsurface compositions and potential deposits, but they cannot specify the existence or absence of hydrocarbon deposits (Hassan et al. 2022; Ayoub Mohammed et al. 2022; Alakbari et al. 2022). Drilling is the only method to demonstrate the hypotheses' validity and reveal the rocks' fluid content. The uncertainty as to the nature of the fluids trapped in the subsoil and the complexity of the sedimentary deposits explain the still high number of negative exploration boreholes and the significant share of boreholes in the cost of research (Chen and Guestrin 2016; Bailey et al. 2018; Zhong et al. 2022; Ayoub et al. 2022). The drilling objectives are always the same: To reach the depth desired by the geologists and to have the possibility of testing or putting the reservoir into production. Optimizing each operation's costs is necessary as the price of a drilling operation is high and negative explorations are unavoidable. Among the performance indicators of a borehole is its rate of penetration (ROP), which indicates the efficiency and speed of an operation more generally. Successfully optimizing the ROP will drastically reduce the costs of the drilling operation. Lateral, axial, or torsional vibrations directly impact this factor. These vibrations have a destructive effect on the drilling rig. In addition, they minimize the ROP and therefore increase the well's maintenance time and cost (Baarimah et al. 2022; Runia et al. 2013; Schwefe et al. 2014; Efteland et al. 2015).

To better understand the source of these vibrations, it must be understood that the drilling rig is, in principle, a motor on the surface. This motor rotates a drill string that can reach up to 5000 m in length, which gives these elements a behavior similar to that of a torsion spring. As the formations to be drilled differ, the torque applied by each of the formations is also different. This creates a discrepancy between the drill bit speed at the end of the string and the motor speed at the surface. This makes it Stick and Slip. Stick and Slip are the number one cause of drill string failure and wear. It is due to the aggressive torsional dynamics experienced by the drilling system, which behaves like a huge torsion spring (Nautiyal and Mishra 2023; Alakbari et al. 2016). These vibrations decrease efficiency, reliability, performance, and safety, which are significant aspects of deep well drilling. The inherent torsional dynamics of the drilling system primarily cause stick and slip vibrations in drilling engineering operations.

The drilling rig functions as a motor on the surface, rotating a drill string extending up to several kilometers in length. Due to variations in the drilled formations, each formation applies a different torque to the drill string, resulting in a mismatch between the desired angular velocity at the drill bit's end and the motor speed at the surface. This disparity between rotational speeds leads to stick and slip vibrations. Stick refers to the phase when the drill bit stops

completely, while slip refers to the stage when the bit rotates significantly faster than the desired angular velocity. These vibrations not only diminish drilling efficiency, reliability, and performance but also pose risks to safety. Stick and slip vibrations can cause premature fatigue and wear of the drill string, leading to failures and increased maintenance costs. Understanding the causes of stick and slip is crucial for effectively mitigating these vibrations and optimizing drilling performance. Drill string vibration can cost companies millions due to decreased performance and efficiency. Thus, vibration mitigation and improved drilling performance are of substantial economic interest to the petroleum industry. At the same time, safety remains the top priority. In other words, achieving performance improvement (e.g., by reducing vibrations) implies reduced costs (Elkatatny et al. 2019; Zakuan et al. 2011; Kyllingstad and Nessj en 2009; Craig et al. 2009). Stick and slip consist of phases in which a bit is completely stopped (sticking), followed by stages where a bit rotates several times faster than the desired angular velocity (slip). A simulation result of a drill pipe with a desired angular velocity of 50 RPM, which suffers from stick and slip vibrations. This phenomenon causes not only the failure of the bit due to premature fatigue of the drill string but also its premature wear due to its high speed. Stick and slip vibrations can be recognized in bottom and surface measurements. Modeling and analyzing these vibrations to attenuate them is an important aspect of drilling.

Approaches to vibration mitigation are classified into two major categories: passive and active. Passive techniques are grouped into three sub-categories: downhole assembly optimization (BHA), bit selection, and downhole equipment use. For example, the anti-stick and slip tools are mechanical tools intended to adjust the drilling torque automatically and reduce the oscillations to increase accuracy and efficiency while drilling (Gupta et al. 2019; Srivastava et al. 2022). Active methods are based on control techniques. The best-known approach is "Soft Torque". This technique's idea is to rotate the surface rotary system gently. For the synthesis of this regulator, the dynamics of this system are considered in the form of a torsion pendulum characterized by two degrees of freedom (Wu et al. 2012; Forster 2011; Bailey et al. 2008). A PI regulator based on surface data can dampen this torsion mode. But these regulators cannot always eliminate stick and slip vibrations, especially in deep wells and inclined wells. Another reason for these failures is uncertainty (non-linearity) in the interaction between the bit and the rock and a difference between the type of measurements used (downhole or surface measurements). Artificial intelligence (AI) techniques are increasingly being employed worldwide, including in the oil and gas sector, resulting in a substantial advancement in their use to increase accuracy (Otchere et al. 2021a, b; Shen et al. 2017; Vogel and Creegan 2016; Greenwood 2016). Several research papers

have utilized automated prediction models to forecast stick and slip vibrations, demonstrating their efficacy in creating patterns among complicated drilling parametric interactions (Zha and Pham 2018; Gupta et al. 2019; Saadeldin et al. 2022).

Srivastava (2022) proposed a hybrid machine learning model to enhance stick and slip prediction in drilling operations. They utilized two machine learning models: the least squares support vector machine (LSSVM) and the multilayer perceptron neural network (MLP-NN). Feature selection was applied to improve prediction accuracy by identifying relevant features, and a Savitzky–Golay (SG) filter was used to reduce noise in the data. The researchers employed the wrapper method in conjunction with evolutionary algorithms to enhance the performance of their models. The LSSVM-Cuckoo Optimization Algorithm achieved the highest prediction accuracy of 0.94 among the hybrid models developed. Additionally, Srivastava et al. (2022) employed an artificial neural network to predict stick and slip problems in offshore oil fields. They divided the datasets into training (60%), validation (20%), and testing (20%) sets and created databases with input features for bit bounce and bending vibrations for data normalization. The application of feature selection led to the passage of the Levenberg–Marquardt function for training the dataset, resulting in high accuracy. Their results showed accuracies of 0.96 and 0.93 for bit bounce and bending vibrations, respectively. Saadeldin et al. (2022) conducted a study to explore the potential of machine learning in predicting stick and slip occurrences during drilling operations. They focused on mining and analyzing drilling data to develop accurate prediction models. To reduce the dimensionality of the dataset, they employed a feature ranking method and selected 18 out of 23 studied parameters as input variables for stick and slip prediction. This approach improved computational processing time and enhanced the efficiency of the dataset. The results showed that the Gaussian kernel support vector machine (SVM) achieved a high level of accuracy with a maximum value of 0.92. The findings of this research paved the way for further investigations, including using different machine learning models that incorporate geological and operational parameters to improve stick and slip predictions.

In their study, Zhong et al. (2022) proposed a new technique for predicting stick and slip vibrations by combining modular neural networks (MNN) with particle swarm optimization (PSO). To improve the model's efficiency, they utilized data normalization to transform the data into a range of 0 to +1. The dataset was divided into 60% for training, 20% for validation, and 20% for testing. The results showed that incorporating PSO led to more accurate predictions by optimizing parameter variations. It was observed that the choice of machine

learning algorithm and input variables varied among different studies focusing on predicting drilling stick and slip vibrations. Supervised machine learning models are only as efficient as the information they are trained with; therefore, when irrelevant data is included as input, the model's performance suffers (Otchere et al. 2021a, b). As such, there is a need to develop a workflow capable of determining the most appropriate features, which involves engineering new features from existing input variables (Otchere et al. 2021a, b; Nautiyal and Kumar Mishra 2023). These new features should be capable of improving model performance compared to when they are inputted singularly. Again, the choice of machine learning models should not be a hindrance if a more robust model appending all robust models into one super machine learning model is used. This tends to mitigate the inherent problems of each model while complementing each other's strengths. This study uses data recorded from measurements while drilling (MWD) in offshore oil fields on the Norwegian continental shelf. The input features and their importance will be assessed by model-agnostic metrics for predicting stick and slip. Once a suitable dataset of relevant features is established, several machine learning models will be developed to predict drill string vibrations. The model with the most accurate results will be optimized using the Bayesian Optimization (BO) algorithm. Our primary objective in this study is to examine and analyze the impact and influence relating to individual features that will generate novel insight into their generalization of explaining the target. A major contribution of this research is the global explanation for variables in stick and slip prediction using explainable AI and developing a machine learning approach to enhance the prediction accuracy for drill string vibrations. This study aims to develop a new approach for real-time prediction of stick–slip vibrations in drilling operations, optimizing drilling parameters, improving drilling performance, and enhancing productivity and safety in the oil and gas industry. We achieve this by accurately identifying stick and slip occurrences using machine learning techniques and incorporating downhole sensor data and a drill string model. The proposed method holds significant potential for integration into automatic driller systems on offshore drilling rigs, leading to more efficient and sustainable drilling practices. This research uses machine learning models to introduce an innovative method for predicting stick and slip vibrations in drilling operations. The approach combines model-agnostic regression models and Bayesian Optimized Extra Trees, improving accuracy with the lowest error metrics. We can effectively reduce stick and slip severity and optimize penetration rate by applying this new model to automatic driller systems in offshore drilling rigs on the Norwegian continental shelf.

This innovation can revolutionize drilling practices and enhance the industry's efficiency and sustainability.

Strengths of the proposed method

In this article, the authors have successfully introduced a groundbreaking method for predicting stick and slip vibrations in drilling operations using machine learning models. The combination of model-agnostic regression models and Bayesian Optimized Extra Trees sets this research apart, bringing a novel and powerful approach to address this critical aspect of drilling operations. The strengths of this article lie in several key areas:

Firstly, the robustness of the dataset used in this research—comprising over 78,000 data points from one of the Norwegian continental shelf oil fields—provides a strong foundation for validating the effectiveness of the proposed method. The sheer volume of data ensures that the model is well-trained and tested, increasing its reliability and applicability in real-world scenarios. Secondly, the impressive accuracy of the Extra Trees (ET) model, demonstrated by its low error metrics (r/min MAE and r/min RMSE), highlights the potential of this method to outperform existing techniques in predicting stick and slip vibrations. This improved accuracy is crucial in enabling more efficient and cost-effective drilling operations, ultimately benefiting the oil and gas industry as a whole.

Additionally, the practical implications of the proposed method cannot be overstated. By integrating this model into automatic driller systems in offshore drilling rigs, operators can optimize the penetration rate, reduce stick and slip severity, and increase overall drilling efficiency. These benefits will contribute to reduced non-productive time, lower operational costs, and a decreased environmental impact, showcasing the immense value of this research. Lastly, the article presents a clear and concise explanation of the methodology, making it accessible to a wide range of readers, from industry professionals to academic researchers. As a result, it establishes a solid foundation for further exploration and development of this innovative approach to predicting stick and slip vibrations in drilling operations.

In conclusion, the proposed method's strengths are its accuracy, applicability, and potential impact on drilling optimization and automation. This research marks a significant milestone in enhancing the understanding and management of stick and slip vibrations, paving the way for continued advancements in offshore drilling technology.

Background of machine learning regression models

This research stems from earlier studies where several workflows and techniques were used to predict stick and slip using available input variables. As academics increasingly move away from empirical correlations, machine learning has become entrenched. Individual drilling parameters offer critical information concerning drill string vibrations. However, when some subsets are used, they can predict stick and slip more accurately. Researchers sometimes employ the Pearson correlation coefficient to identify significant variables, while others rely on wrapper approaches, intricacy techniques, and algorithms that use metaheuristics to identify powerful features. This indicates that no definitive collection of features exists for making this prediction. However, a robust workflow is needed to utilize model-agnostic techniques to solve this prediction problem. This study investigates the correlation among the most crucial input-relevant variables in predicting vibrations caused by sticks and slips in oil well drill strings. The relevant features selected by the model-agnostic techniques will then be input features. Based on their learning theory and ability to work with high-dimensional and complex data, several machine learning models will be used to predict stick and slip. The reviewed models perform differently depending on learning theory, dimensionality, small or large data, and different data distributions. Identifying a suitable model that can handle all these problems has become necessary since drilling data are commonly highly dimensional, small or large, with different data distributions.

Based on these assertions, this study establishes relevant features that can lead to lower prediction errors. Improved accuracy in estimating stick and slip vibration will have a massive influence on field operations and the overall integrity of the well. Hence, although dependent on data type, machine learning models can be applied to drilling operations with similar input features. The most common issue with machine learning models and their varying performance is centered on data. Having the capability of explaining the input variables and determining causation has been the pinnacle of this study as more innovative approaches are created to solve this issue. The slightest gain in accuracy is critical to improving the decision-making process within the petroleum business, thus indicating the importance of this field for advancing research. The summary analysis of the reviewed models used in this study is presented in Table 1, which can be found as follows.

Table 1 An overview of the algorithms described in this research and the corresponding authors

Model	Description	Developed or implemented by
Ridge Regression	A robust technique against multicollinearity minimizes standard errors by applying some bias to the model estimates, resulting in a more reliable prediction. The Ridge regression model works by calculating differences between standardized dependent and non-dependent variables and dividing them by their standard deviations	A study by Hoerl and Kennard published in (1970)
Least Absolute Shrinkage and Selection Operator (LASSO)	Robust against multicollinearity performs better with lower-dimensional data, prevents overfitting, and reduces standard errors	Tibshirani (1996)
Support Vector Machine (SVM)	Structural risk minimization induction produces improved generalized robust, sparse, and unique results based on a geometrically straightforward approach	Vapnik in collaboration with Lerner (1963)
Decision Tree (DT)	It uses induction and pruning techniques to build hierarchical decision boundaries and remove unnecessary structures from the decision tree to battle overfitting	Li et al. (1984)
The Extreme Gradient Boosting method (XGBoost)	Gradient-boosted decision trees are implemented of the first- and second-order maximize loss functions, adding regularization terms to avoid overfitting by adjusting the final weights	Chen and Guestrin (2016)
Extremely Randomized Trees (Extra Tree or ET)	Potentially achieves better performance than the random forest, a simpler algorithm is employed to construct the decision trees used as an ensemble member	Geurts et al. (2006)
Random Forest (RF)	An advanced decision tree that is robust against overfitting and offers easy interpretability	Breiman (2001)

Data collection and data description

This research's main objective focuses on improving the prediction of stick and slip using data from the offshore oilfield on the Norwegian continental shelf during drilling operations. Fields located in the North Sea, near the southern end of the Norwegian sector, are described as fault block structures with approximately 173 million barrels of original oil in place (OOIP). Reservoirs are small dome-shaped structures formed when adjacent salt deposits collapsed over the Middle of the Jurassic (Geurts et al. 2006). The data consists of 78,000 data points with 23 features compiled from resources such as measurement while drilling (MWD), daily drilling reports (DDRs), final drilling well reports, and literature. However, after removing missing data, 49,000 points remained. Table 2 shows the description of real-time MWD measurement records in this study. Also, Table 3 shows the descriptive statistics of some selected features. The main input features for this research are;

a. *Real-time downhole drilling data* Stick and slip indicator, collar rotational speed, turbine revolution per minute, transverse shock risk, ARC shock level, Isonic

shock, power drive shock risk, equivalent circulating density, annular pressure, and annular temperature.

- b. *Real-time surface drilling data* Surface weight on bit, torque, the total flow rate of all active pumps, and pump pressure.
- c. *Formation characteristics* Gamma-ray, formation type, bit depth, and penetration rate.

One of the main advantages of applying data analytics and machine learning to data is finding patterns and hidden information in high-dimensional data. As such, some input variables will be deemed irrelevant in predicting stick and slip vibrations. Table 4 presents the typical lithology observed in the wells of the oil fields located on the Norwegian continental shelf. For this study, the lithology was encoded into numeric variables ranging from 1 to 19, illustrated in Table 5. The coordinates for each well were also removed.

Table 2 Description of real-time MWD measurement records used in this study

Name	Mnemonic	Unit	Description
Depth	DEPT	M	Bit depth in meters or measured depth (MD)
Rate of penetration	ROP5	M/HR	Drilling progress in distance and time-averaged every 5 (ft.)
Gamma-ray	ARC_GR_UNC_RT	° API	The record of the emission of endogenous radioactive elements to the rock
Traveling block position	BPOS	M	The height at which the traveling block is located on the mast or derrick
Traveling block velocity	BVEL	M/HR	The velocity of movement of the block and the hoisting system
Surface weight on bit	SWOB	KKGF	Measurement of the weight exerted by the string on the bit and, therefore, on the formation
Hook-load	HKLD	KKGF	Measurement of the load on the hook by the working and drilling string
Torque	TQA	KMN	The torque exerted by TDS derived from string rotation in units of kilometers decanewtons
Revolution per minute	RPM	RPM	Measurement of the revolutions of the turbine contained in the BHA to energize downstream components
Drill string revolutions or collar rotational speed	CRPM_RT	RPM	Revolutions of the BHA at the level of lastraberrenas
Total pump flow	TFLO	LPM	The flow rate of drilling mud to well
Bottom turbine revolutions	TRPM_RT	RPM	Measurement of the revolutions of the turbine contained in the BHA to energize downstream components
Pump pressure	SPPA	BAR	Friction losses in the hydraulic system
Transverse shock risk	SHKRSK_RT	--	--
ARC shock level	SHK2_RT	--	--
Isonic shock	SHK_ISONIC_RT	--	--
Power drive shock risk	PDSHKRSK	--	--
Shock peak	SHKPK_RT	G	The magnitude of the biggest shock experienced by the Power Pulse is the Shock Peak
Equivalent circulating density	ECD_ARC	SG	The density of the bottom-well fluid
Annular pressure	APRS_ARC	BAR	Downhole annular pressure measurements
Annular temperature	ATMP	DEGC	Bottom annular space temperature measurement
Job on-bottom time	BDTI	HR	--
Stick and slip indicator	Stick_RT	RPM	Torsional vibration

Methodology

Data analysis and visualization

To aid in the understanding of how the input features relate to the output, data analysis and visualization were used. The degree of relationship between input and output variables was quantified using the Spearman rho matrix of covariance. To study linear and nonlinear relationships, we used Pearson's coefficients, Kendall correlations, as well as Spearman's rank coefficients based on Eqs. (1), (2), and (3).

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad (1)$$

$$\tau = \frac{n_c - n_d}{\frac{1}{2}n(n-1)} \quad (2)$$

$$r_s = 1 - \frac{6 \sum_{i=1}^n D^2}{n(n^2 - 1)} \quad (3)$$

In Eq. (1) x and y are our samples, and the variables \bar{y} and \bar{x} are an average between samples, and n is a sample size. n_c and n_d indicate the number of concordant and discordant pairs, as given by Eq. (2). Furthermore, Eq. (3) specifies D as the difference between the ranks of the two groups and n as their volume. Pearson, Kendall, and Spearman coefficients are correlated more strongly if they are close to +1 or -1. In contrast, the input parameters are independent if these coefficients are close to zero. Due to this independence of input parameters, more information can be incorporated into the models as a result of selecting features that integrate more information. Knowing the difference between these two is important, and that correlation does not imply causation. This Heatmap of Fig. 1 shows the features that have a moderate correlation to the target. Figure 2 shows stacked horizontal bar charts for different correlation coefficients.

Table 3 Summary statistics of input and target characteristics

	Count	Mean	std	min	25%	50%	75%	max
DEPT	49,099.00	3207.16	295.60	2780.83	2927.10	3173.92	3483.09	3750.05
ROP5	49,099.00	17.66	7.88	3.20	10.10	18.35	24.24	36.73
ARC_GR_UNC_RT	49,099.00	44.29	39.09	0.00	24.52	37.36	49.63	257.22
BPOS	49,099.00	137.10	4.81	62.96	133.01	136.54	141.00	150.25
BVEL	49,099.00	174.06	12.56	8.95	172.40	176.27	180.60	193.21
SWOB	49,099.00	60.50	12.55	0.00	50.03	61.38	70.71	152.01
HKLD	49,099.00	60.21	12.44	39.29	49.80	61.24	70.83	109.87
TQA	49,099.00	6.40	2.96	0.00	4.44	6.41	8.39	55.25
RPM	49,099.00	12.16	3.31	0.00	10.76	12.67	14.44	21.04
CRPM_RT	49,099.00	145.38	44.43	0.00	139.00	163.00	172.00	242.00
TFLO	49,099.00	148.59	44.86	0.00	140.00	169.00	179.00	182.00
TRPM_RT	49,099.00	73.67	62.58	0.00	27.00	54.00	99.00	381.00
SPPA	49,099.00	0.00	0.02	0.00	0.00	0.00	0.00	2.00
SHKRSK_RT	49,099.00	0.48	4.07	0.00	0.00	0.00	0.00	344.00
SHK2_RT	49,099.00	21.62	12.08	2.22	10.97	22.19	31.59	45.07
SHK_ISONIC_RT	49,099.00	0.01	0.01	0.00	0.00	0.01	0.01	0.19
PDSHKRSK	49,099.00	78.80	9.62	55.00	73.00	79.00	88.00	95.00
SHKPK_RT	49,099.00	1964.06	97.29	0.00	1962.83	1962.83	2000.25	2067.15
ECD_ARC	49,099.00	1.46	0.03	1.37	1.45	1.47	1.48	1.50
APRS_ARC	49,099.00	416.78	23.04	365.84	402.27	421.36	433.40	460.85
ATMP	49,099.00	2797.04	141.66	2031.25	2773.44	2812.50	2851.56	4960.94
BDTI	49,099.00	0.00	0.15	0.00	0.00	0.00	0.00	10.00
Stick_RT	49,099.00	44.33	3.26	0.00	44.00	44.00	45.00	48.00

Some features correlate to torsional vibration, including shock peak, equivalent circulating density, and Isonic shock. In addition, there is a slight correlation between total pump flow, annular temperature, and bottom turbine revolutions related to stick and slip. The pair plot also depicts the non-linear distribution between some of the input and the target. Figure 3 shows that none of the features correlated linearly with the target based on the nonlinear distribution. Therefore, nonlinear models were deemed appropriate for this research.

Application of machine learning models

The cross-validation approach involves resampling input data into multiple groups to evaluate its predictive performance and minimize prediction errors on unknown data (Otchere et al. 2021a, b). This study employed the hold-out cross-validation approach, where randomly partitioned data was divided into 85% for training and 15% for testing, with 7,350 data points dedicated to testing. This approach ensures that separate data points, not seen by the trained model, are used to validate the model, preventing overfitting and selection bias. This research aimed to identify a model capable of predicting stick and slip vibrations with minimal errors. Consistency was maintained by using the

same out-of-sample data throughout the study, ensuring that similar data points were used across all models.

The hyper-parameter tuning process was carried out using the Bayesian Optimization (BO) algorithm to improve the performance of the final model. BO is a model-based sequential optimization technique that utilizes the Bayes Theorem to explore the extrema of objective functions efficiently. It involves iteratively training a probabilistic Bayesian approximation of the target model based on prior outcomes approximation. The probabilistic model, which relates hyper-parameters to the objective function's probability score, is then evaluated using an acquisition function to select the best measurements for optimizing the accurate parameter. Hyper-parameter tuning aims to update the model parameters that control accuracy and performance during the training process, thereby optimizing the model's performance by minimizing the cost function (Otchere et al. 2022a, b, c).

Explainable AI

This research will employ model-agnostic methods to assess the significance of input features and provide explanations independent of the machine learning model used. The objective is to identify features that contribute

Table 4 General lithology in oil field wells of the Norwegian continental shelf (Frankiewicz 2019)

Lithology	Description
Seabed	Consists of dense to very dense sands overlaying stiff clay
QUARTERNARY	Clay with thin stringers of sand. Coarser material up to boulder size may occur
NORDLAND GP.—Pliocene and Pleistocene	Gray claystone with thin stringers of sand and siltstone
Utsira formation	Fine to medium-grained, moderately well and well-sorted sandstone with minor silt and limestone stringers
HORDALAND GP.—Eocene to Miocene	It is dominated by claystone and minor limestone/dolomite stringers, except for the sandy Skade and Grid formations
Skade Fm	Medium-grained and moderately sorted sandstone, occasionally calcareous cemented
Grid Fm	Very fine to fine-grained sandstone
ROGALAND GP.— <i>L. Paleocene to L. Eocene</i>	
Balder Fm	Vari-colored claystone, partly tuffaceous with some limestone stringers
Sele Fm	Claystone and minor limestone stringers
Lista Fm	Non-calcareous claystone with minor limestone stringers
Ty Fm	Very fine to medium-grained sandstone, moderately to poor sorted, with some interbedded claystone, siltstone, and a few limestone stringers
SHETLAND GP.— <i>U. Cretaceous</i>	
Ekofisk Fm	Chalky off-white to light gray limestone, moderately hard with traces of claystone and sandstone
Tor Fm	White limestone, moderately hard, becoming pale red-brown and very hard with depth, traces of claystone
Hod Fm	Off-white to white limestone, moderately hard, chalky, grading to marl with depth, glauconite
Blodøks Fm	Medium to dark gray marl, argillaceous laminations, glauconitic in parts
Hidra Fm	Off-white, firm limestone
CROMER KNOLL GP.— <i>U. Cretaceous to L. Cretaceous</i>	
Rødby Fm	Marl with argillaceous laminations
Sola Fm	Marl and claystone
Åsgard Fm	Interbedded limestone and marl with some minor layers of claystone and siltstone
VIKING GP.— <i>U. Jurassic</i>	
Draupne Fm	Very organic-rich claystone, micaceous, carbonaceous, and traces of pyrite
Heather Fm	Claystone with limestone stringers and interbedded claystone, kaolin, sandstone, and limestone in the lowermost part
VESTLAND GP.— <i>M. Jurassic</i>	
Hugin Fm	Sandstone, very fine to very coarse-grained, moderately to well sorted. Rare claystone stringers
Sleipner Fm	Sandstone, very fine to medium-grained, moderately to well sorted, gray claystone, and layers of coal
HEGRE GP.— <i>U. Triassic</i>	
Skagerrak Fm	Fine-grained sandstone with some interbedded silty sections
Smith Bank Fm	Reddish brown claystone with occasionally sandstone stringers

meaningfully to the model's predictions, regardless of their type or nature. Two techniques, namely permutation feature importance (PFI) and Shapley values, will be utilized to analyze the results generated by the models. These techniques require the model, feature vectors, target variable, and error metrics as inputs. By applying PFI and Shapley values, the features that fail to adequately explain the target variable will be identified and removed from the input feature vector.

Permutation feature importance

The Permutation Feature Importance (PFI) approach, as described by Otchere et al. (2022a, b, c), is a model-agnostic method used to assess the significance of features. It works by permuting the association between a variable and the model result, assigning random values to each feature. The importance of a component is determined by measuring the increase in estimation error after permuting the feature multiple times. If shuffling the parameters leads to higher

Table 5 Numeric codes are assigned to formations in the intended oil field on the Norwegian continental shelf

Formation type	Numeric code
Utsira Fm	1
Skade Fm	2
Grid Fm	3
Balder Fm	4
Sele Fm	5
Lista Fm	6
Ty Fm	7
Ekofisk Fm	8
Tor Fm	9
Blodøks Fm	10
Hidra Fm	11
Rødby Fm	12
Åsgard Fm	13
Draupne Fm	14
Heather Fm	15
Hugin (Block 1) Fm	16
Sleipner (Block 1) Fm	17
Hugin (Block 2) Fm	18
Sleipner (Block 2) Fm	19

errors, the feature is considered important, whereas if the error remains constant, the feature is deemed unimportant. The PFI method offers advantages such as straightforward interpretation, providing a global overview of a predictive model, and not requiring retraining. However, whether training or test data should be used to measure PFI is unclear. The permutation importance function in Scikit-Learn (Pedregosa et al. 2011) can be used for PFI analysis, calculating the feature importance scores based on specified criteria, with higher values indicating better predictive power.

Shapley values

A Shapley value, introduced by Shapley (1997) from the theory of cooperative games, serves as a critical means of measuring the importance of each feature and its contribution to a model's performance. Shapley values work by allocating among participants the entire surplus value created through the coalition of all participants in each cooperative game. According to this theory, four properties must be met: efficiency, symmetry, dummy, and additive properties. The players represent the input features used in machine learning and interpretability, and the Shapley value determines how each component influences the outcome. Knowing how the features contribute to the model output helps understand the essential factors in generating the model's results. The Shapley value is not dependent upon any particular model kind; hence, it applies to any model type and structure.

In contrast to other methods of interpreting model results, Shapley's values are based on a rigorous conceptual framework. While intuitive reasoning is essential for interpretability, other methods lack the same rigid conceptual framework. Shapley values have the advantage of being evenly distributed among an instance's feature values. It has been suggested that the Shapley decision may provide the only way to offer detailed and complete clarifications in cases where the law mandates it. This is the right to explanations. Shapley, like any other approach, has inherent drawbacks. In most real-world scenarios, it is only possible to calculate an approximate solution because it is computationally expensive. It is also easy to misunderstand. Shapley's value does not reflect the difference between a model's expected and actual values when the feature is removed from the training process. How much does an input value contribute to the real and average prediction gap?

The open-source SHAP library works well with SHAP values and other metrics. SHAP establishes a link between efficient credit allocation and local explanations, making it model-agnostic. Doing so brings together optimal credit allocation and reasons at the local level. The use of game theory's traditional sharp values and their associated expansions has been the focus of numerous recent works. Recent academics have expanded on this notion by developing approaches incorporating SHAP values into deep learning models and gradient explainers. These approaches can estimate SHAP values for any model using specifically weighted local linear regression. It also provides various charts to help view the data and comprehend the model. This study does not include the mathematical foundations of the models utilized because there is a vast collection of published material detailing them. An overview of the methodology and workflow applied to this research is shown in Fig. 4.

Results and discussion

Model performance evaluation

These models employed in this study were evaluated using an Akaike Information Criterion (AIC) to establish which model provided the most optimal fit concerning the data analysis. The assessment of the results shows that lower AIC values have a higher probability of relevant data than larger AIC. The following factors were considered when comparing and selecting the most suitable model. A difference in AIC findings across models is considered insignificant if it is less than 10, moderate between 10 and 50, significant between 50 and 100, and strong if above 100. The results show that the extra tree algorithm was superior and most accurate for stick and slip vibration predictions. There was a difference of ~932 AIC when the extra tree model was

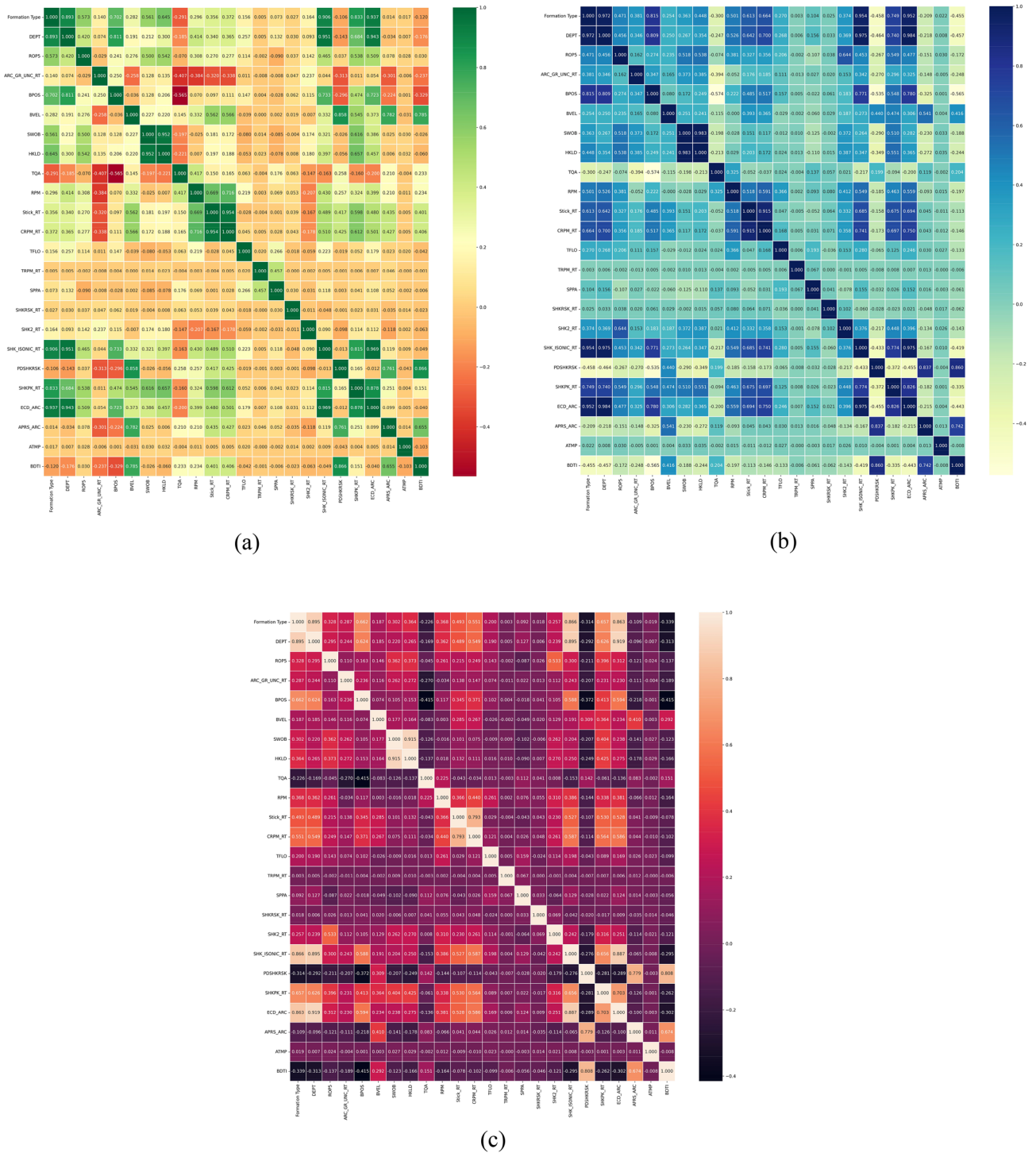


Fig. 1 Heatmap to assess the relationships between the input variables and targets; **a** Pearson correlation, **b** Spearman correlation, **c**

compared to the random forest model representing the two models with the lowest AIC (Fig. 5). Hence, the extra tree model showed a strong disparity and improved likelihood of fit compared to the other models. Even though AIC can calculate the relative quality of a model in terms of its

convergence with the actual data, the model's bias for over-fitting or under-fitting remains.

The models' performance was assessed based on the out-of-sample data results. Since the models have been trained with the training data, any attempt to reproduce the values

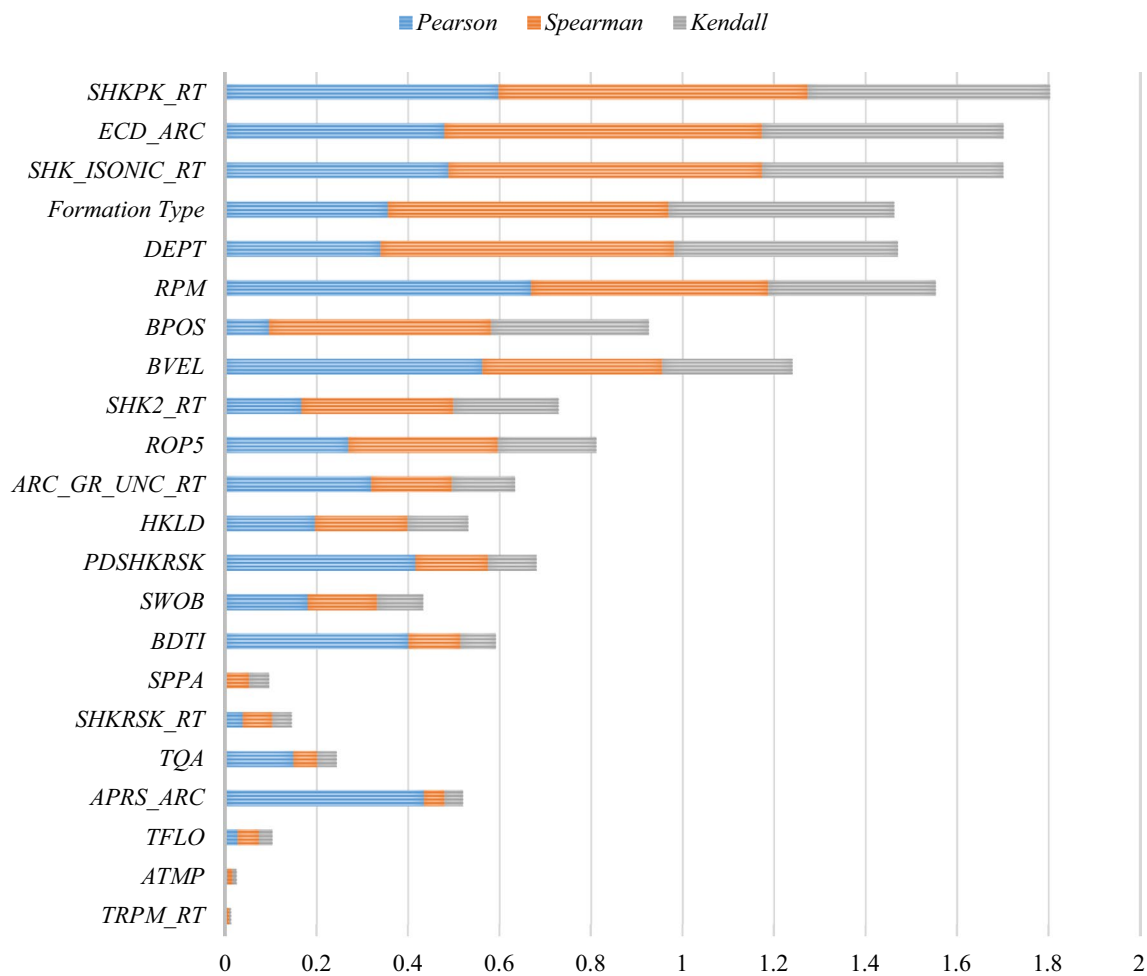


Fig. 2 Stacked bar chart for the different correlation coefficients

will likely result in high accuracy. High training accuracy, however, is not always desirable since it may overfit the data, collecting intrinsic noise and yielding a non-generalized model. Test accuracy based on non-sample data refers to the possibility of accurate predictions made by algorithms based on variables that have not yet been seen or observed. Since the model aims to improve unseen data, models must have high test accuracy. Correlation coefficient (R^2) values greater than 0.85 in training and accuracy test scores were deemed an acceptable trade-off over biases and variance for this work. Table 6 shows that the Extra Tree model had the highest R^2 on the test data for predicting stick and slip. Figure 6 compares all of the model estimates to actual results.

Any meaningful gain in accuracy, and hence a reduction in inaccuracy, achieved by machine learning models have a major influence on making decisions. A result of the models employed in the study produced mixed performance results which were heavily influenced by both their theoretical and statistical underpinnings. Figure 7 depicts the models' performance as a cross-validation evaluation to

analyze their exactness, efficiency, and reliability. Based on each model's root mean square error (RMSE) findings, it can be concluded that the Extra Tree model evaluation is the most reliable compared to the real stick and slip values. The Mean Absolute Error (MAE) also suggests that the Extra Tree model is the most accurate for predicting stick and slip vibrations. In selecting the most suitable model, the ranking features used in this study are MAE, RMSE, AIC, and R^2 . The performance of the extra tree model is attributed to the bias-variance concept used to build this model. When the cut-point and features are explicitly randomized, coupled with ensemble averaging, the algorithm can significantly reduce variance compared to weaker randomization techniques employed by several algorithms. Aside from having similar advantages as Decision Trees based on consistency for universal generalization and approximation, Extra Trees also provide resilience regarding gross model errors. This is because outliers affect its predictions slightly and locally. When the input features are significantly more than the random splits, Extra Trees potentially outperform Decision

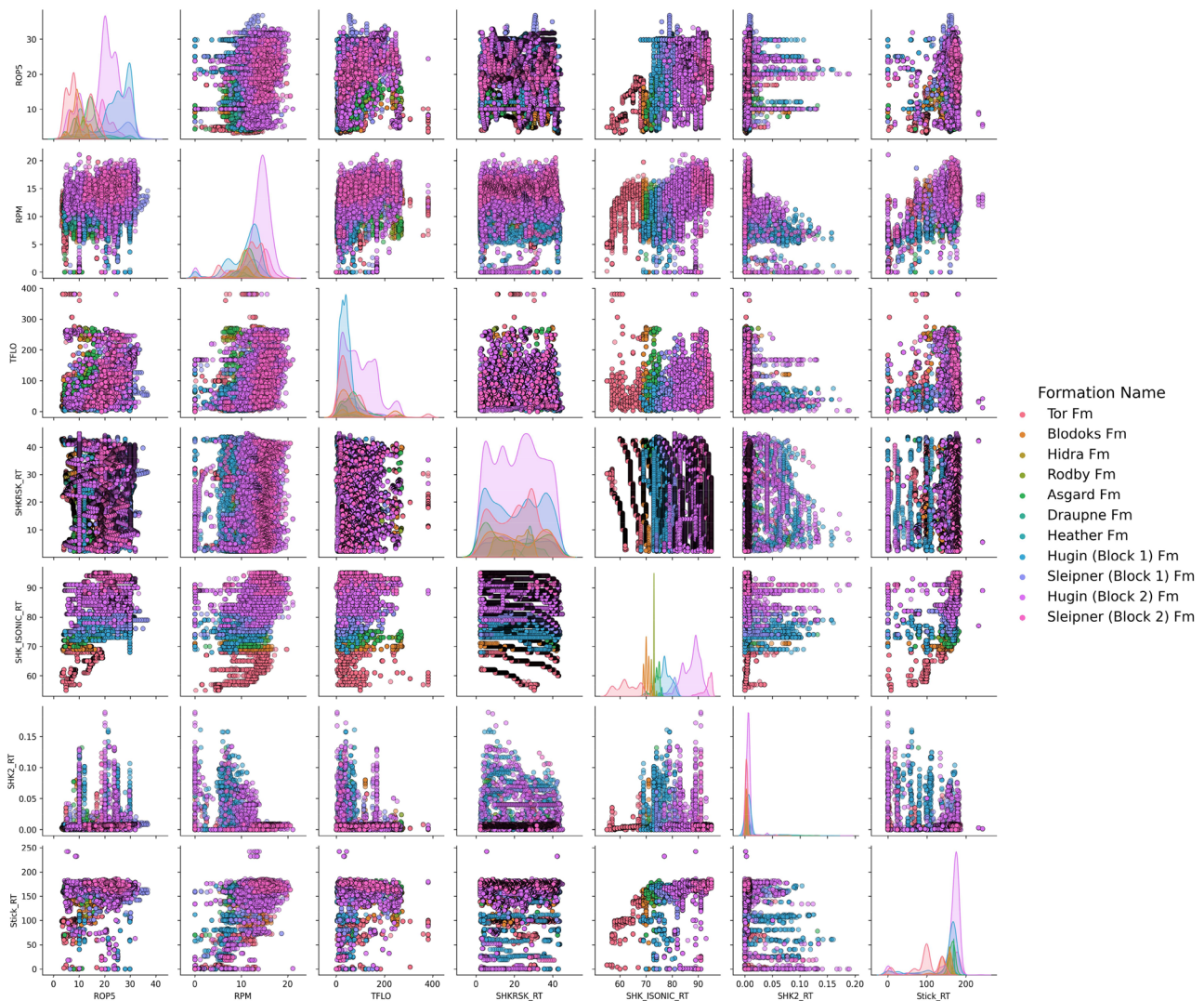


Fig. 3 Pair plot distribution of some variables color-coded against the formation type

Trees in terms of computational efficiency while resisting irrelevant input features.

Model agnostic results

Permutation importance based on XGBoost has been computed to estimate the variables' significance and analyze the influence of multicollinearity among input parameters. Accordingly, the results in Fig. 8 show that the collar rotational speed, annular temperature, and gamma-ray highly correlate to predict stick and slip variations. The results indicate that permuting an input feature reduces the model's accuracy by, at most, 0.125, suggesting that some of the inputs are relevant. However, Isonic shock, transverse shock risk, ARC shock level, rate of penetration, and pump pressure do not offer any improvements to the model prediction. A comparison is made between this result and the result

of the XGBoost algorithm presented in Fig. 9. A model-agnostic approach can be used to infer causality, despite the discrepancy in the result. This is necessary because correlation does not mean causation; hence, feature selection techniques based on correlation may not entirely improve model performance.

Analysis of features using agnostic model metrics

An illustration of the PFI results for the input features in the entire data set is shown in Fig. 10. The final model prediction was made using the same train/test split as in the previous study to maintain consistency. As a result of the initial testing, R^2 using this Random Forest (RF) model reached 0.94. The importance score is calculated so that higher values indicate higher predictive power. A significant part of the accuracy score 0.94 is attributed

Fig. 4 A flow chart summarizes the procedure and methodology used in the research

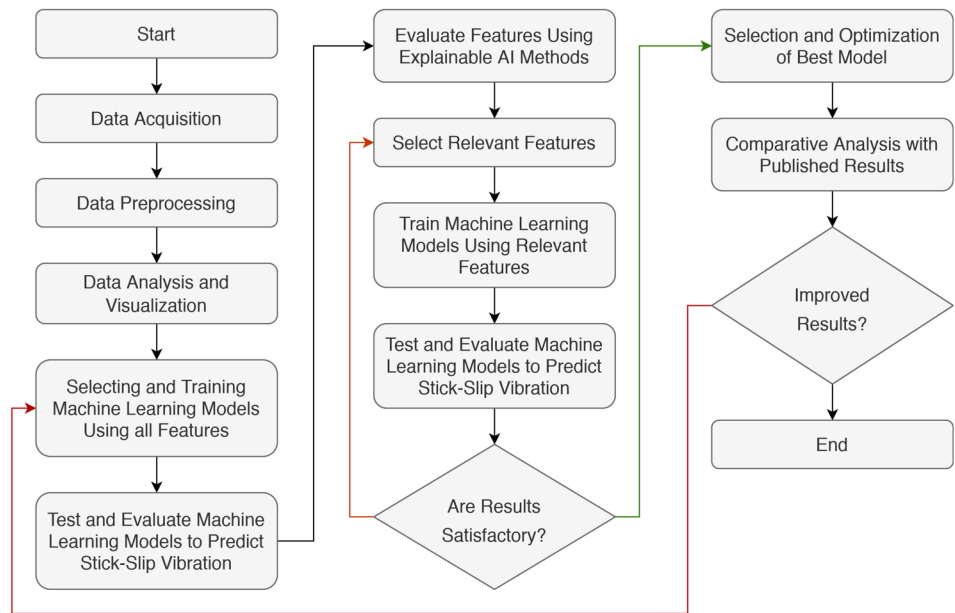


Fig. 5 AIC results comparing all the models on the test data

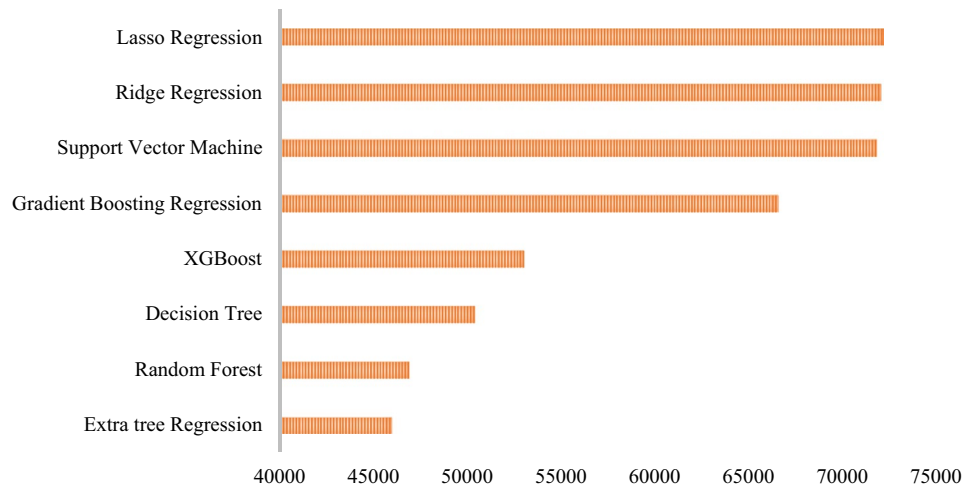


Table 6 Training and testing the coefficient of correlation score for each model used in this study

Models	R^2 score for training	R^2 score for testing
Ridge regression	0.818433313	0.776743
Lasso regression	0.814570912	0.773679
Support vector machine	0.999954119	0.757682
Decision tree	1	0.920748
Random forest	0.999274577	0.938261
Extra tree regression	1	0.944944
Gradient boosting regression	0.913981113	0.867297
XGBoost	0.99113411	0.936788

to the value of the most relevant features. According to the results, some input variables are significantly underweighted. According to these outcomes, we can conclude that a very limited number of variables have significant predictive power. This result indicates that some of the characteristics are more relevant than others. Based on the high test accuracy computation, this outcome follows the results.

The initial permutation importance calculated on the training data shows the model dependency on each feature during training. However, it is vital to continue to mention that this analysis is done for both the training and testing datasets. This is to help account for features that may help with the generalization power of the model. If a feature is considered important for the train set but not for the test

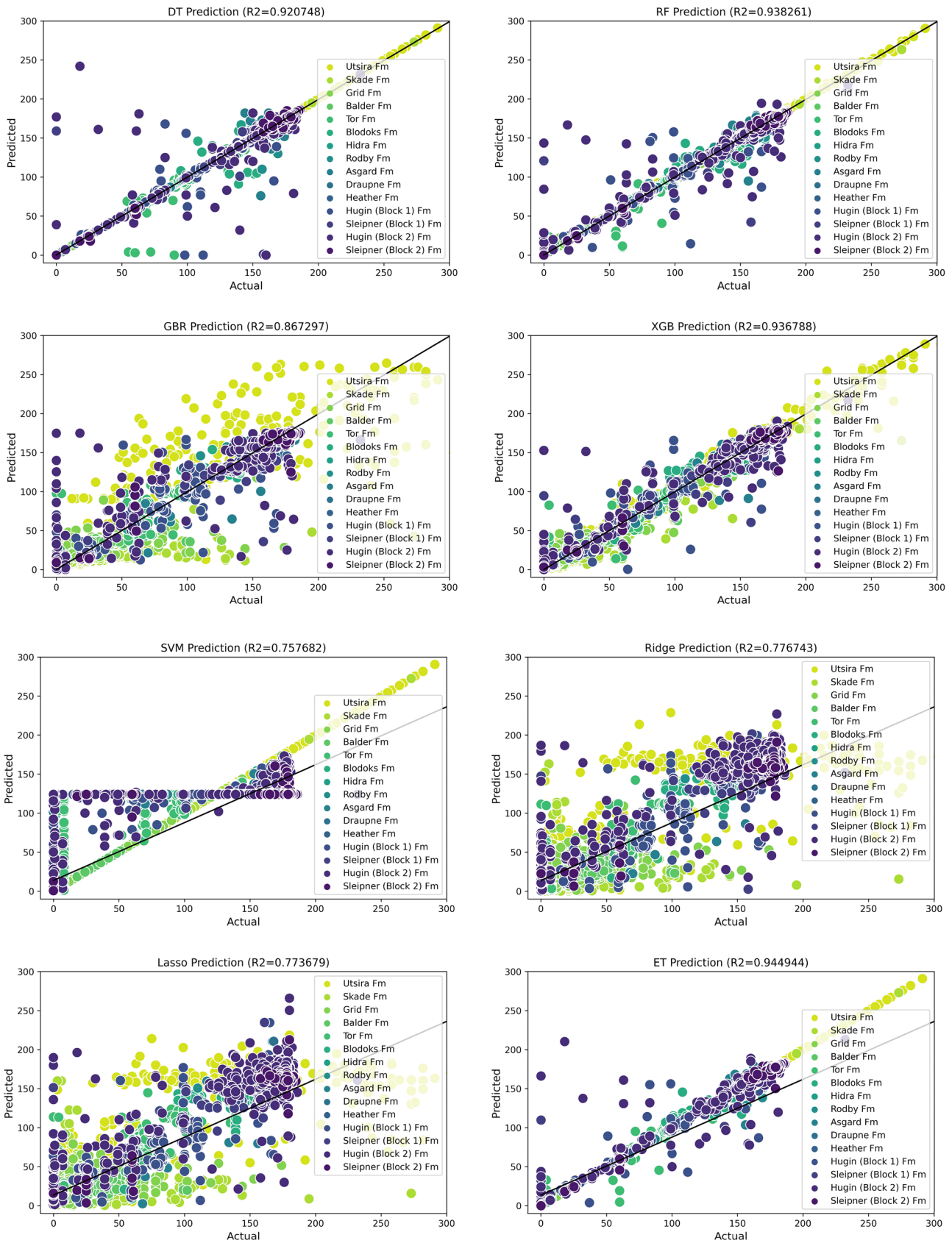


Fig. 6 Cross-plot of predicted vs. actual stick and slip for all models based on test data

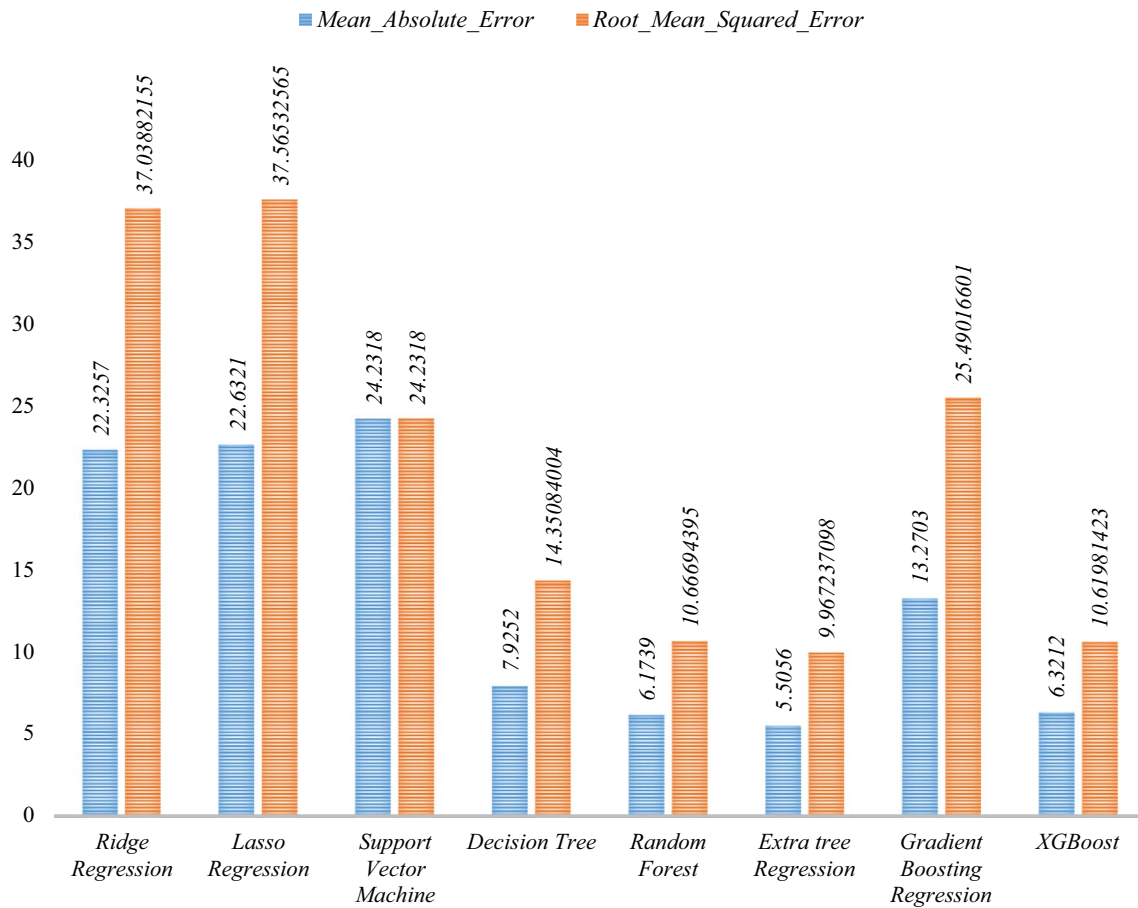


Fig. 7 Comparison of test data RMSE and MAE prediction errors for all models

set, the model will likely overfit. Sensitivity analysis was performed to select the relevant features for further prediction. The RF model was used to predict stick and slip vibrations using the top 24–18 features, and the result is illustrated in Table 7. The results indicate that the optimal number of features to use in predicting stick and slip is 22.

Analysis of features using shapley values model agnostic metrics

The extra tree model was used to predict drill string vibrations using all the features. The same training and testing datasets used in the PFI analysis were used. Since machine learning models are supposed to be interpretable to help understand how a model made a particular prediction, model-agnostic metrics are helpful. Shapley's values originated from game theory, and it is necessary to clarify its application to supervised machine learning interpretability. A game represents the prediction assignment for a specific event in the dataset. The gain is calculated by comparing the actual forecast for a particular instance to the average estimates applied to all the cases. A given instance's feature

values collaborate to achieve the gain in the stick and slip forecast. Figure 11 visualizes the shapely values as absolute values in a feature importance plot. In shapely feature importance, the most important features are those with high absolute values. The average mean of the Shapley values for each feature across the data is computed to evaluate the global importance. The collar rotational speed, annular temperature, and gamma-ray feature variables were highly correlated, with gamma-ray playing the most significant role and affecting the target by an average of 34.06 r/min on the x-axis. The plot also indicated that the correlation of formation type, bottom turbine revolutions, and ARC shock level to the target was low, with the inclusion of ARC shock level, disagreeing with PFI results.

Based on this result, further investigation was necessary to ascertain which features were irrelevant to the target. The importance of Shapley values is an alternative to the significance of permutation features. Both importance measures have significant differences. The decline in model performance determines the permutation feature's relevance, whereas Shapley values are primarily determined by the magnitude of the features' provenances. In terms of feature

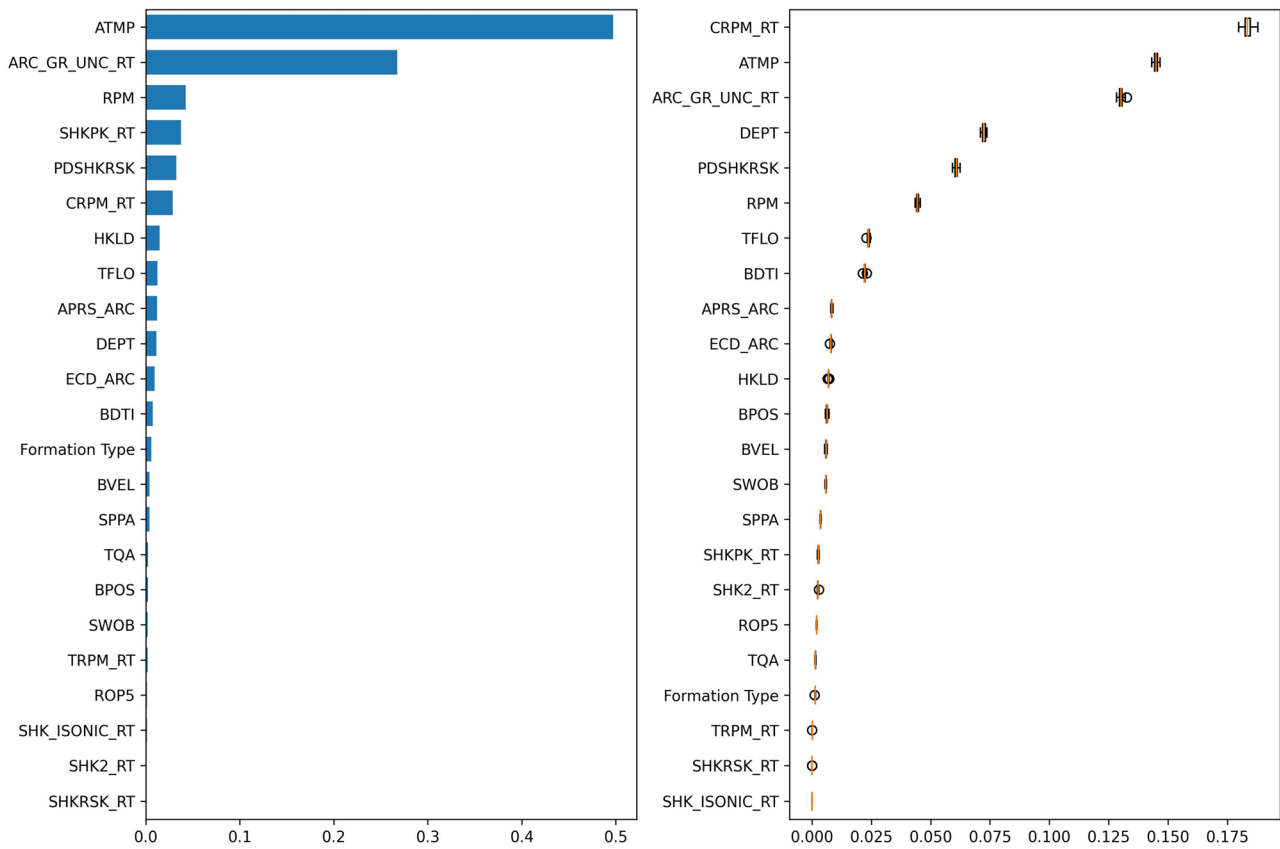


Fig. 8 Permutation importance plot

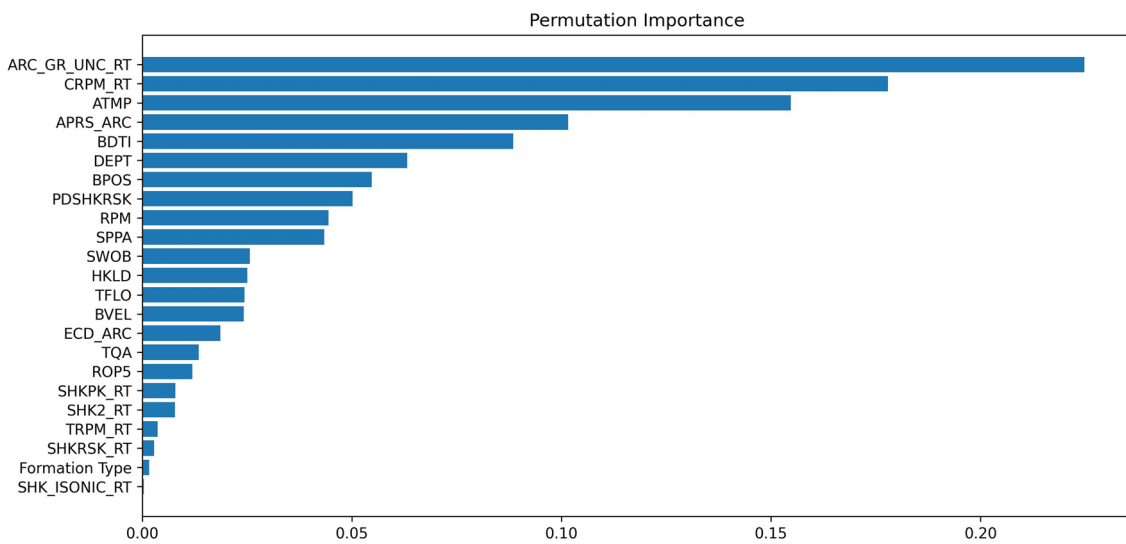


Fig. 9 XGboost feature importance to the target features

importance, the plot is informative; however, it provides few other details outside of this significance. The bee swarm technique, a more informative plot, is used for analysis.

Figures 12 and 13 visualize the shapely values as absolute values in a bee swarm summary plot for train and test data, respectively. The y-axis determines a feature, and the x-axis defines a Shapley value. The feature importance with

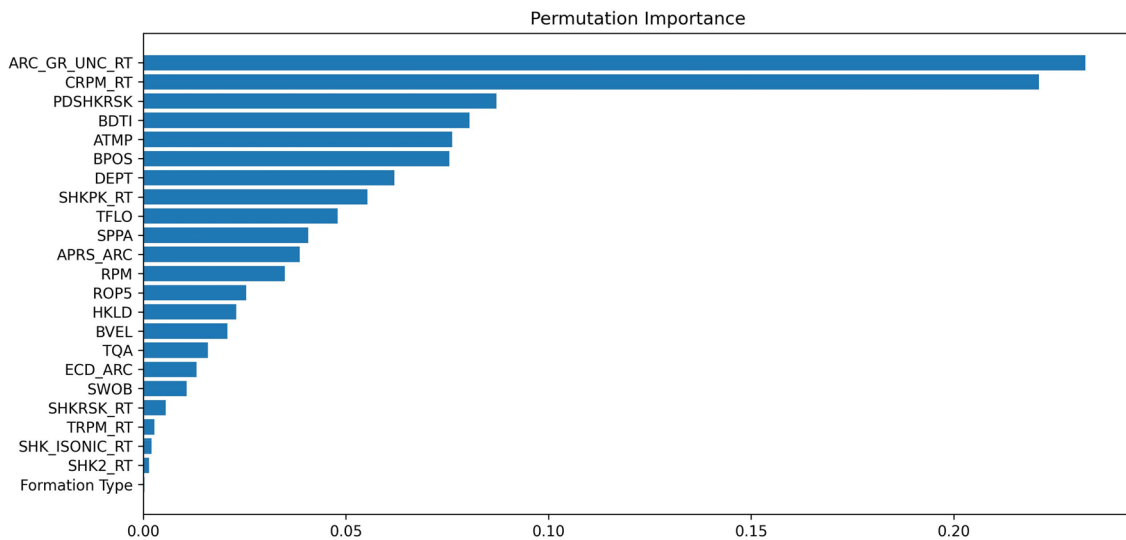


Fig. 10 Permutation features the importance of the input variables

Table 7 Selection of input features based on accuracy performance

Input features	Features dropped	Accuracy
24	None	0.938261
23	Formation type	0.944203
22	Formation type, bottom turbine revolutions	0.954385
21	Formation type, bottom turbine revolutions, ARC shock level	0.954103
20	Formation type, bottom turbine revolutions, ARC shock level, transverse shock risk	0.954188
19	Formation type, bottom turbine revolutions, ARC shock Level, transverse shock risk, rate of penetration	0.954245
18	Formation type, bottom turbine revolutions, ARC shock level, transverse shock risk, rate of penetration, torque	0.954261

associated effects is combined into global explanations whereby a feature matrix of Shapley values is achieved for every instance. According to the summary plot, the first indications of a positive and negative correlation between the value of a feature and its impact on a target are identified. All the data points, made up of all the training data points, overlap and are scattered on the y-axis. This type of visualization depicts how Shapley values are distributed across features. Feature values are arranged in descending order of significance. From the plots, low feature values are represented in blue, while high feature values are denoted in red. Analyzing the influence of gamma ray and annular temperature, it is observed that low values predict high stick and slip values, whereas high values predict low stick and slip values. The distribution of the gamma-ray data points also indicates a global explanation that explains the entire model's behavior. This analysis confirms that this correlation alone cannot be termed causality. As a result, to fully understand the influence of other variables on stick and slip

prediction, a model-agnostic approach is needed. From the bee swarm plot, it can be observed that about 22 features can globally explain how the predictions were made.

The results of the bee swarm plot confirmed that features that have high importance in both the train and test results exhibit their significance in the global explanation of the target. From the demonstrated results, the following analysis was derived:

1. *Feature importance:* The features are ranked in descending order, and torque and bit depth are ranked first from the train and test data plot. The formation type feature has close to zero importance because it does not have any causal effect on predicting stick and slip.
2. *Impact:* There is a negative correlation between torque and bit depth, with a high prediction effect in general, as shown by the horizontal location of the data points.

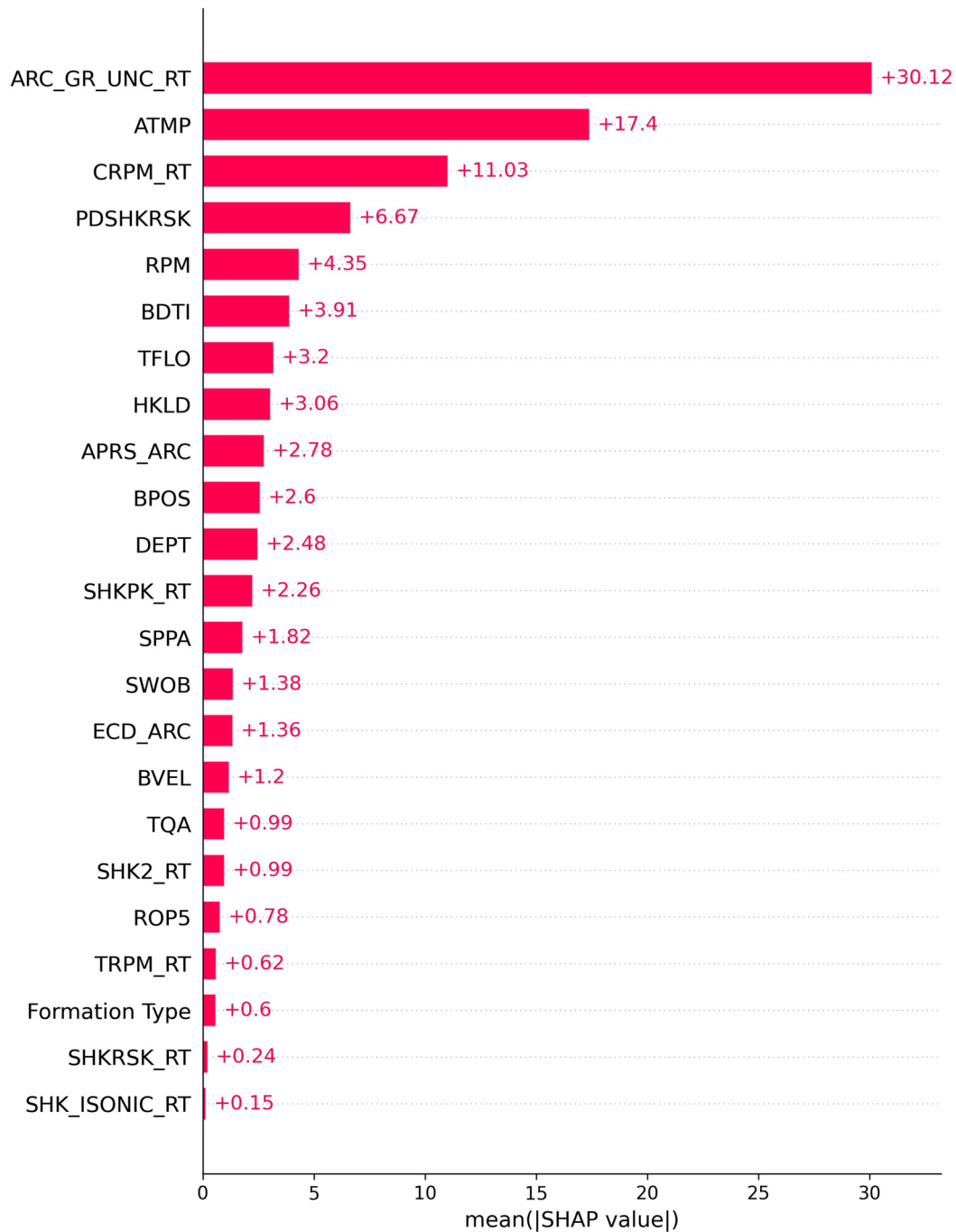


Fig. 11 A measure of the importance of a SHAP feature is the mean absolute Shapley value

Feature values based on Shapley values are based on a solid theoretical basis in game theory and compute feature values reasonably distributed to the target. This method also contributes to the unification of linguistic learning. Fast computations allow multiple Shapley values to be calculated for global model interpretations. Because Shapley values represent a small element of global variations, those interpretations are compatible with local explanations.

Evaluation of top features

Further analysis was performed using the top 5 performing models for the top 22 features based on PFI and Shapley values. This was done to see if they would outperform the initial model predictions. PFI resulted in the removal of formation type and Isonic shock. While Shapley values resulted in the reduction of formation type and bottom turbine

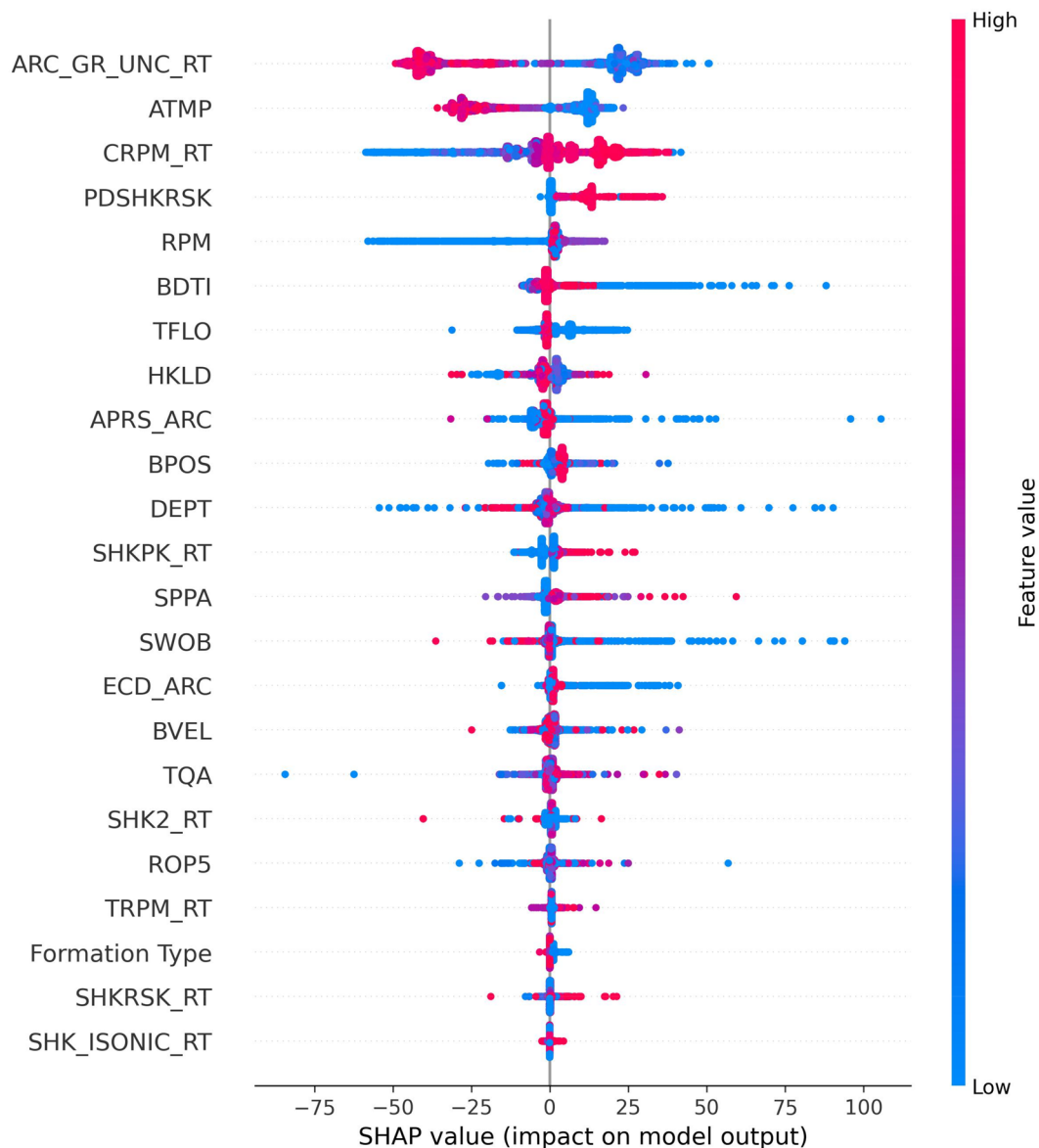


Fig. 12 A bee swarm summary plot of the effect of SHAP values on a model target using the train data

revolutions. Table 8 provides a summary of the results of the research. All three top-performing approaches improved when new features based on PFI and Shapley values were used. Although Shapley values selected the most relevant features, most of the enormous improvement in model accuracy can be seen.

Model optimization

This study employed the Bayesian optimization (BO) technique to tune specific model parameters to improve model performance. In developing our Bayesian optimized extra trees (ET) model, a range of hyper-parameters were carefully

selected and tuned to optimize the algorithm's performance. A comprehensive table detailing the chosen hyper-parameters, their descriptions, and their respective values can be found in Table 9. This table provides a clear understanding of the algorithm's configuration. It allows for easy replication of our methodology. By referring to Table 9, readers can gain insights into the intricacies of our model and understand how the selected hyper-parameters contribute to the prediction of stick and slip vibrations in drilling operations.

The results in Fig. 14 show the model performance of the extra tree model using the top 22 features identified by Shapley values and the BO_ET model. The results were compared to those of Srivastava et al. (Srivastava et al.

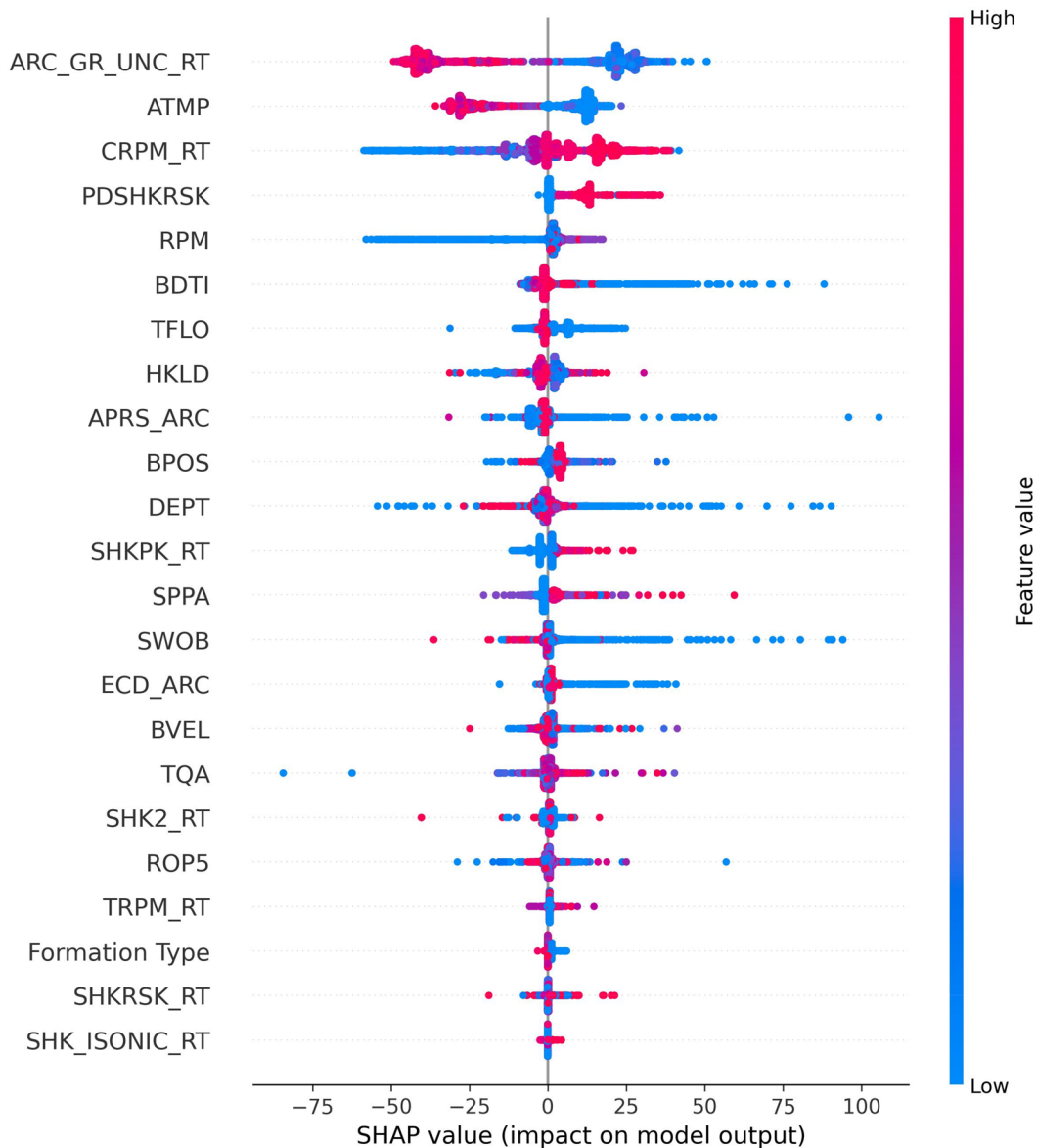


Fig. 13 A bee swarm summary plot of the effect of SHAP values on a model target using the train data

2022). The authors used the same data as this study but employed different techniques and models for predicting stick and slip vibrations. Notably, although machine learning models depend, most data are expected to be identical. Consequently, this comparison focuses on the input features and how they explain the target rather than downgrading the other researchers' results.

From the error analysis performed, it was observed that the features selected based on the Shapley values greatly reduced the model's error. The initial extra tree model recorded an MAE of 5.50 r/min, while the MAE from the Shapley selected features was 1.E-06 r/min. This result represents about a 95% reduction in MAE. Similarly, the RMSE for the extra tree model using all 24 features was

9.9672 r/min. The RMSE from the Shapley selected features was recorded as 4.4616 r/min. Again, this represents about a 99% reduction in this error metric. The extra tree was further improved by hyper-parameter tuning where BO_ET recorded MAE and RMSE values of 0.002156 as well as 0.024495 r/min, accordingly. Moreover, the superiority of the extra tree model can largely be attributed to the bias-variance concept used to build this model, which makes it resilient to outliers. Based on all evaluation criteria, it can be concluded that Shapley selected parameters are highly relevant, offers a global generalization of the target, and improve model efficiency.

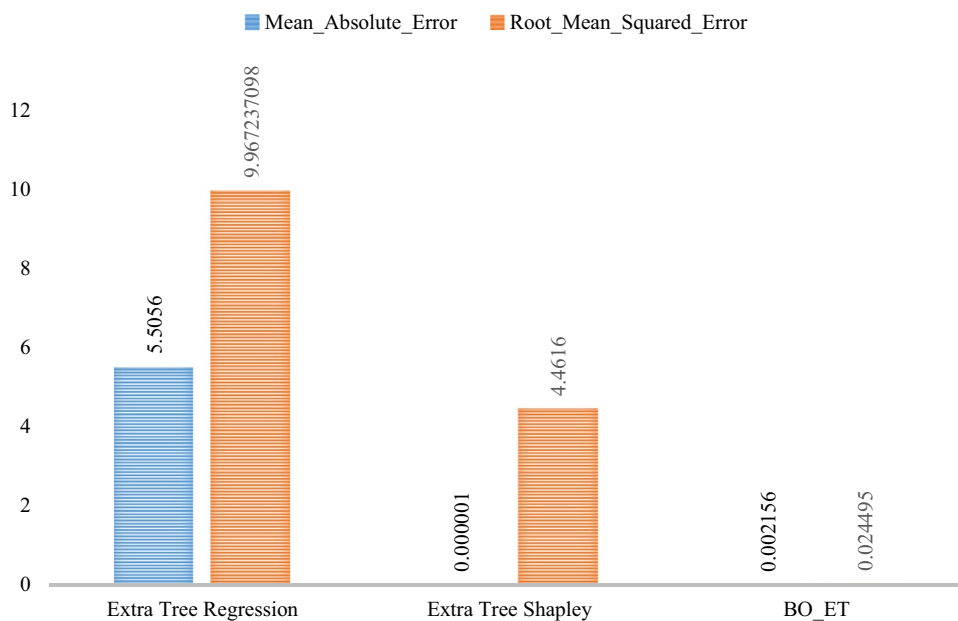
Table 8 Computed the accuracy of the test data using the top 5 algorithms

Various models	Achieving accuracy based on all features	Accuracy based on PFI's top 22 features	Accuracy based on Shapley values top 22 features
Gradient Boosting	0.867297	0.872997	0.995597
Decision Tree	0.920748	0.934348	0.993548
Random Forest	0.938261	0.943361	0.991361
Extra Tree	0.944944	0.956144	0.999888
XGBoost	0.936788	0.941888	0.991288

Table 9 Hyper-parameters of Bayesian Optimized Extra Trees (ET) Algorithm

Hyper-parameter	Description	Value
n_estimators	Number of trees in the forest	100
max_depth	Maximum depth of the tree	10
min_samples_split	Minimum number of samples required to split an internal node	2
min_samples_leaf	Minimum number of samples required to be at a leaf node	1
max_features	The number of features to consider when looking for the best split	sqrt
bootstrap	Whether bootstrap samples are used when building trees	False

Fig. 14 Comparing the estimation errors of top-performing models



Sensitivity analysis

In Fig. 15, we demonstrate the kernel density estimates for predicted and actual stick and slip results. A blue line indicates predicted values, while a purple line indicates actual values. As a result of observing the testing data, the BO_ET has a much higher degree of agreement with the real data. Specifically, this investigation indicates that the BO_ET can capture a wider range of values than the other algorithms used in this study. Consequently, this observation makes it suitable for evaluating stick and slip vibrations in other wells using the relevant drilling parameters. As a result, the

sensitivity analysis verifies the assessment metrics outlined above.

In this study, explainable AI was demonstrated using model-agnostic metrics. Since correlation does not mean causation, simple statistical techniques may not be able to indicate the causal effects of each input on the target. The findings show that gradually increasing the input variables above 22 causes a degradation in the model's accuracy. Based on Shapley values, formation type and bottom turbine revolutions are among the variables that have no direct influence on stick and slip vibrations. Hence, it should not

Fig. 15 Kernel density estimation for BO_ET model stick and slip prediction demonstrating the closeness of projected values to real values

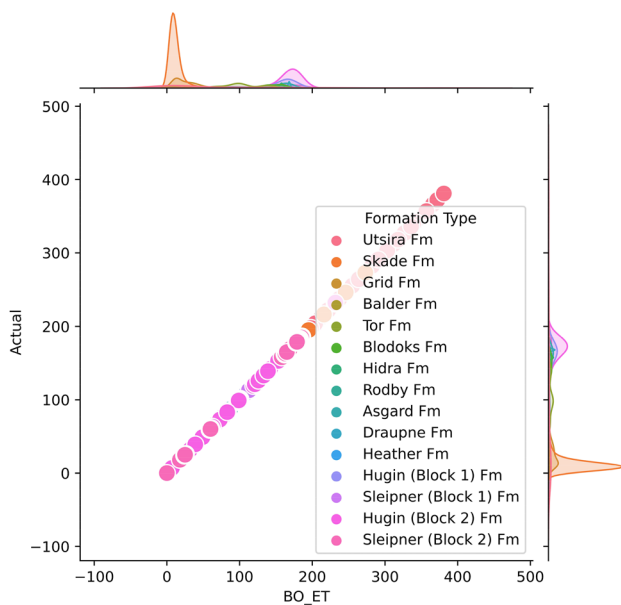
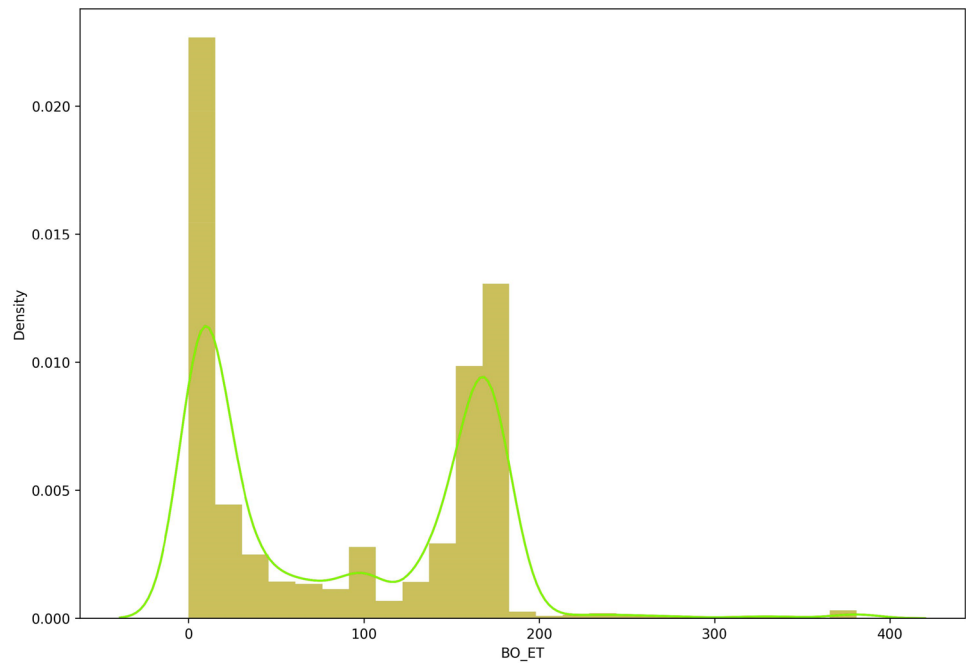


Fig. 16 BO_ET versus actual stick and slip per formation type ($R^2=0.9999$)

be used to train the model. Using relevant causal features, model-agnostic approaches have provided some trustworthiness and useful solutions for drilling-string vibration prediction. The ability to adequately predict stick and slip before drilling gives vital information to help choose optimal RPM and WOB for suppressing such a phenomenon. As a result, forecasting stick and slip can reduce costly downtime, including downhole motor failures, twisting off, and damage

to drill bits (Fig. 16). Predicting stick and slip vibrations can give valuable insight into the type of interventions required to be implemented.

Assessing results validity in stick–slip prediction models for oil fields

Regarding the validity of the results, we point out that the test set was kept separate from both the training and validation sets throughout the entire procedure. It was essential to determine the models' performance on unseen data and compare them to one another. This is discussed in Sect. 4.2. However, it is critical to note that test accuracy varies significantly within the individual wells of the test set. This is despite our efforts to test the model as accurately as possible. In this regard, the findings might have differed had the data been split differently. Additionally, the data itself had a significant impact on this paper. The data sets for the desired oil field were carefully chosen. This is discussed in section "Data analysis and visualization". The data sets for the desired oil field were carefully selected for accurate generalization.

Because the model was not trained on the test set, it could not predict torsional vibrations. This is because there was a stick or slip in the test set. As stated in the paper, we used data from the Norwegian Petroleum Directorate (NPD), which can be regarded as a valid source of information. There are various sources and periods from which the data were reported to NPD, which may result in the differing quality of the reported data. Since the equipment used for data collection has evolved and improved over the years,

the time aspect of data collection is relevant. In addition, differences in data collection processes between companies may also have affected the results. This paper presents artificial intelligence (AI) learning algorithms that can only be applied to predict stick and slip based on the data given to them. Humans can determine the right stick and slip formations using their experience and possibly other measurements. A machine learning algorithm will detect false patterns if fed erroneous data.

Recognizing and overcoming limitations

Although our research demonstrates promising results regarding predicting stick–slip vibrations in drilling operations using machine learning, it is important to acknowledge some potential limitations of this study. By addressing these limitations, we can improve the validity of our findings and identify areas for future research.

1. *Limited data:* The model's performance depends on the quality and quantity of downhole sensor data and drill string model information. The availability of comprehensive and accurate data is essential for the model to make accurate predictions. However, if the data is insufficient or of low quality, it may hinder the model's performance.
2. *Generalizability:* Our model may be limited to the specific drilling operations or conditions it was trained on, reducing its applicability across different geological formations or drilling operations. Further testing and adaptation may be necessary to ensure the model's effectiveness in diverse scenarios.
3. *Model complexity:* Our study's Bayesian Optimized Extra Trees approach may be computationally expensive. This could limit the model's applicability in real-time drilling operations or where computational resources are constrained. Future research could explore optimizing the computational efficiency of the model.
4. *Comparison with other models:* While our model outperformed several reported models, a more extensive comparison with a wider range of existing models would provide greater insight into its true value and relative performance. This could also highlight potential areas of improvement for our model.
5. *Implementation challenges:* Integrating our model into automatic driller systems could encounter hardware limitations, integration difficulties, and resistance from industry practitioners reluctant to adopt new technologies. Future work should address these challenges and facilitate the model's adoption in practical settings.

In conclusion, recognizing and addressing these limitations can further refine our model and enhance its applicability in predicting stick–slip vibrations in drilling operations, ultimately contributing to more efficient and cost-effective drilling processes.

Conclusions

1. This research introduces a novel method for predicting stick and slip vibrations in drilling operations using machine learning techniques, specifically model-agnostic regression models combined with Bayesian Optimized Extra Trees. The approach demonstrates significant innovation in addressing downhole vibrations and holds the potential to enhance drilling efficiency and safety in the oil and gas industry.
2. Various data analysis and visualization techniques were employed to quantify the relationships between input and output variables, revealing both linear and nonlinear correlations. The selection of input features based on metrics such as MAE, RMSE, AIC, and R2 aided in identifying the most relevant variables for accurate stick and slip predictions.
3. The Extra Tree model outperformed other models, achieving superior accuracy in stick and slip vibration predictions as indicated by lower AIC values and higher R2 scores. The model's performance was validated using out-of-sample data, emphasizing its reliability and ability to generalize to unseen data.
4. Integration of downhole sensor data and a drill string model enabled the model to accurately identify stick and slip occurrences in real-time, providing opportunities for optimizing drilling parameters such as revolutions per minute (RPM) and weight on the bit (WOB). By mitigating stick and slip severity, the proposed method improves the rate of penetration (ROP) and contributes to enhanced productivity and safety in the oil and gas industry.
5. The research highlights the significance of feature selection and the importance of relevant input features. Collar rotational speed, annular temperature, and gamma-ray exhibit high correlations with stick and slip vibrations prediction. On the other hand, variables such as Isonic shock, transverse shock risk, ARC shock level, rate of penetration, and pump pressure have limited impact on the model's prediction.
6. The methodology employed in this research, including permutation feature importance and Shapley values, provides insights into the relevance and impact of input variables on stick and slip predictions. The use of Bayesian Optimization for hyper-parameter tuning

enhances the performance of the Extra Tree model, resulting in even lower MAE and RMSE values.

7. The proposed method demonstrates its superiority over existing models reported in the literature through rigorous testing on data from a Norwegian oil field. The potential integration of this method into automatic driller systems on offshore drilling rigs offers real-time identification and management of stick and slip vibrations, contributing to improved drilling efficiency and safety in the oil and gas industry.
8. The adoption of machine learning techniques, as demonstrated in this research, presents the potential for solving complex problems in the oil and gas sector. The reduction in the need for human operators to identify and manage stick and slip vibrations saves time and resources while minimizing drill string failures and associated costs.
9. By embracing data-driven technologies and integrating this approach into existing drilling systems, the oil and gas industry can transform its practices, leading to improved performance, cost savings, and positive environmental impacts. The proposed method's ability to optimize the penetration rate contributes to more sustainable drilling operations by reducing energy requirements.
10. This research provides a comprehensive analysis of the proposed method, its advantages over existing models, and the potential practical implications for the oil and gas industry. The findings emphasize the innovative nature of the approach and offer valuable insights for decision-making and future advancements in machine learning in the oil and gas sector.

Funding The research, authorship, and publication of this article were not supported by external funding.

Declarations

Conflict of interest The authors have not declared a conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Alakbari FS, Elkatatny S, Baarimah SO (2016). Prediction of bubble point pressure using artificial intelligence AI techniques. In: SPE Middle East artificial lift conference and exhibition, Manama, Kingdom of Bahrain. SPE-184208-MS. <https://doi.org/10.2118/184208-MS>
- Alakbari FS, Mohyaldinn ME, Ayoub MA, Muhsan AS, Abdulkadir SJ, Hussein IA, Salih AA (2022) Prediction of critical total drawdown in sand production from gas wells: machine learning approach. *Can J Chem Eng* 101(5):2493. <https://doi.org/10.1002/cjce.24640>
- Ayoub Mohammed MA, Alakbari FS, Nathan CP, Mohyaldinn ME (2022) Determination of the gas-oil ratio below the bubble point pressure using the adaptive neuro-fuzzy inference system (ANFIS). *ACS Omega* 7(23):19735–19742. <https://doi.org/10.1021/acsomega.2c01496>
- Ayoub MA, Elhadi A, Fatherlhman D, Saleh MO, Alakbari FA, Mohyaldinn ME (2022) A new correlation for accurate prediction of oil formation volume factor at the bubble point pressure using Group Method of Data Handling approach. *J Petrol Sci Eng* 208(B):109410. <https://doi.org/10.1016/j.petrol.2021.109410>
- Baarimah SO, Al-Gathe AA, Baarimah AO (2022) Modeling yemeni crude oil reservoir fluid properties using different fuzzy methods. In: International conference on data analytics for business and industry (ICDABI), Sakhir, Bahrain, <https://doi.org/10.1109/ICDABI56818.2022.10041519>
- Bailey JR, Biediger EAO, Gupta V, Ertas D, Elks WC, Dupriest FE (2008) Drilling vibrations modeling and field validation. In: IADC/SPE drilling conference, Orlando, Florida, USA. SPE-112650-MS. <https://doi.org/10.2118/112650-MS>
- Bailey JR, Payette GS, Wang L (2018) Improved methods to understand and mitigate stick-slip torsional vibrations. In: IADC/SPE drilling conference and exhibition, Fort Worth, Texas, USA. SPE-189673-MS. <https://doi.org/10.2118/189673-MS>
- Breiman L (2001) Random forests. *Machine learning*, 45:5-32. <https://doi.org/10.1023/A:1010933404324>
- Chen T, Guestrin C (2016) Xgboost: a scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining, New York, USA, pp 785–794. <https://doi.org/10.1145/2939672.2939785>
- Craig AD, Hanley CM, McFarland B, Shearer D, King P (2009) A proven approach to mitigating drilling vibration problems in offshore Western Australia. In: International petroleum technology conference, Doha, Qatar, IPTC, 151. <https://doi.org/10.3997/2214-4609-pdb.151.iptc13399>
- Efteland F, Creegan A, Jordan L, Caraway C (2015) The significance of pro-active online monitoring with stick-slip mitigation. In: Abu Dhabi international petroleum exhibition and conference, Abu Dhabi, UAE. SPE-177950-MS. <https://doi.org/10.2118/177950-MS>
- Elkatatny S, Shahzad Kamal M, Alakbari F, Mahmoud M (2019) Optimizing the rheological properties of water-based drilling fluid using clays and nanoparticles for drilling horizontal and multi-lateral wells. *Appl Rheol* 28(4). <https://doi.org/10.3933/applrheol-28-43606>
- Forster I (2011) Axial excitation as a means of stick slip mitigation-small scale rig testing and full-scale field testing. In: SPE/IADC drilling conference and exhibition, Amsterdam, The Netherlands. SPE-139830-MS. <https://doi.org/10.2118/139830-MS>
- Frankiewicz J (2019) The application of data analytics and machine learning for formation classification and bit dull grading prediction. Dissertation, University of Stavanger, Norway. <http://hdl.handle.net/11250/2633873>
- Geurts P, Ernst D, Wehenkel L (2006) Extremely randomized trees. *Mach Learn* 63:3-42. <https://doi.org/10.1007/s10994-006-6226-1>

- Greenwood JA (2016) Vibration monitoring and mitigation-an integrated measurement system. In: IADC/SPE drilling conference and exhibition, Fort Worth, Texas, USA. SPE-178773-MS. <https://doi.org/10.2118/178773-MS>
- Gupta S, Chatar C, Celaya JR (2019) Machine learning lessons learnt in stick-slip prediction. In: Abu Dhabi international petroleum exhibition and conference, Abu Dhabi, UAE, SPE-197584-MS. <https://doi.org/10.2118/197584-MS>
- Hassan AM, Ayoub MA, Mohyadinn ME, Al-Shalabi EW, Alakbari FS (2022) A new insight into smart water assisted foam SWAF technology in carbonate rocks using artificial neural networks ANNs. In: Offshore Technology Conference Asia, Virtual and Kuala Lumpur, Malaysia, OTC-31663-MS. <https://doi.org/10.4043/31663-MS>
- Hoerl AE, Kennard RW (1970) Ridge regression: biased estimation for nonorthogonal problems. *Technometrics* 12(1):55–67. <https://doi.org/10.1080/00401706.1970.10488634>
- Kyllingstad Å, NessjØen PJ (2009) A new stick-slip prevention system. In: SPE/IADC drilling conference and exhibition, Amsterdam, the Netherlands. SPE-119660-MS. <https://doi.org/10.2118/119660-MS>
- Li B, Friedman J, Olshen R, Stone C (1984) Classification and regression trees (CART). *Biometrics* 40(3):358–361. <https://doi.org/10.2307/2530946>
- Nautiyal A, Mishra AK (2023) Machine learning application in enhancing drilling performance. *Procedia Comput Sci* 218:877–886. <https://doi.org/10.1016/j.procs.2023.01.068>
- Otchere DA, Ganat TOA, Gholami R, Lawal M (2021a) A novel custom ensemble learning model for an improved reservoir permeability and water saturation prediction. *J Nat Gas Sci Eng* 91:103962. <https://doi.org/10.1016/j.jngse.2021.103962>
- Otchere DA, Ganat TOA, Gholami R, Ridha S (2021b) Application of supervised machine learning paradigms in the prediction of petroleum reservoir properties: comparative analysis of ANN and SVM models. *J Petrol Sci Eng* 200:108182. <https://doi.org/10.1016/j.petrol.2020.108182>
- Otchere DA, Abdalla Ayoub Mohammed M, Ganat TOA, Gholami R, Aljunid Merican ZM (2022a) A novel empirical and deep ensemble super learning approach in predicting reservoir wettability via well logs. *Appl Sci* 12(6):2942. <https://doi.org/10.3390/app12062942>
- Otchere DA, Ganat TOA, Nta V, Brantson ET, Sharma T (2022b) Data analytics and Bayesian optimised extreme gradient boosting approach to estimate cut-offs from wireline logs for net reservoir and pay classification. *Appl Soft Comput* 120:108680. <https://doi.org/10.1016/j.asoc.2022.108680>
- Otchere DA, Ganat TOA, Ojero JO, Tackie-Otoo BN, Taki MY (2022c) Application of gradient boosting regression model for the evaluation of feature selection techniques in improving reservoir characterisation predictions. *J Petrol Sci Eng* 208:109244. <https://doi.org/10.1016/j.petrol.2021.109244>
- Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Duchesnay E (2011) Scikit-learn: machine learning in python. *J Mach Learn Res* 12:2825–2830
- Runia DJ, Dwars S, Stulemeijer IP (2013) A brief history of the Shell soft torque rotary system and some recent case studies. In: SPE/IADC drilling conference, Amsterdam, the Netherlands. SPE-163548-MS. <https://doi.org/10.2118/163548-MS>
- Saadeldin R, Gamal H, Elkhatmy S, Abdulraheem A (2022) Intelligent model for predicting downhole vibrations using surface drilling data during horizontal drilling. *J Energy Resour Technol*. <https://doi.org/10.1115/1.4052794>
- Schwefe T, Ledgerwood LW, Jain JR, Fuselier DM, Oueslati H, Endres L (2014) Development and testing of stick/slip-resistant PDC bits. In: IADC/SPE drilling conference and exhibition, Fort Worth, Texas, USA. SPE-168026-MS. <https://doi.org/10.2118/168026-MS>
- Shapley LS (1997) A value for n-person games. *Classics in game theory*, 69. <https://doi.org/10.1515/9781400881970-018>
- Shen Y, Zhang Z, Zhao J, Chen W, Hamzah M, Harmer R, Downton G (2017) The origin and mechanism of severe stick-slip. In: SPE annual technical conference and exhibition, San Antonio, Texas, USA. SPE-187457-MS. <https://doi.org/10.2118/187457-MS>
- Srivastava S (2022) An experimental, modeling and machine learning based investigation of stick-slip vibrations. PhD dissertation. University of Oklahoma
- Srivastava S, Shah RN, Teodoriu C, Sharma A (2022) Impact of data quality on supervised machine learning: case study on drilling vibrations. *J Petrol Sci Eng* 219:111058. <https://doi.org/10.1016/j.petrol.2022.111058>
- Tibshirani R (1996) Regression shrinkage and selection via the lasso. *J R Stat Soc: Ser B (methodol)* 58(1):267–288. <https://doi.org/10.1111/j.2517-6161.1996.tb02080.x>
- Vapnik V (1963) Pattern recognition using generalized portrait method. *Autom Remote Control* 24:774–780
- Vogel SK, Creegan AP (2016) Case study for real time stick/slip mitigation to improve drilling performance. In: SPE/IADC middle east drilling technology conference and exhibition, Abu Dhabi, UAE. SPE-178176-MS. <https://doi.org/10.2118/178176-MS>
- Wu X, Karupiah V, Nagaraj M, Partin U, Machado M, Franco M, Duvvuru H (2012) Identifying the root cause of drilling vibration and stick-slip enables fit-for-purpose solutions. In: IADC/SPE drilling conference and exhibition, San Diego, California, USA. SPE-151347-MS. <https://doi.org/10.2118/151347-MS>
- Zakuan A, Junaida A, Subroto B, Hermawan H, Fatakh A, Halim A (2011) Stick slip mitigation plan to improve drilling. In: SPE Asia Pacific oil and gas conference and exhibition, Jakarta, Indonesia. SPE-141988-MS. <https://doi.org/10.2118/141988-MS>
- Zha Y, Pham S (2018) Monitoring downhole drilling vibrations using surface data through deep learning. In: SEG technical program expanded abstracts, 2101–2105. Society of Exploration Geophysicists. <https://doi.org/10.1190/segam2018-2964198.1>
- Zhong R, Salehi C, Johnson Jr, R (2022) Machine learning for drilling applications: a review. *J Nat Gas Sci Eng* 104807. <https://doi.org/10.1016/j.jngse.2022.104807>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.