

Article

A Geomechanically Augmented Neural Network with Heterogeneity-Adaptive Data Splitting, Systematic Hyperparameter Optimization, and LSTM-FCNN Hybrid Architecture for Rate of Penetration (ROP) Prediction

Ahmed S. Alhalboosi and Mohammed A. Khamis * 

Department of Petroleum and Natural Gas Engineering, College of Engineering, King Saud University, Riyadh P.O. Box 800, Saudi Arabia; 445108463@student.ksu.edu.sa

* Correspondence: mokhamis@ksu.edu.sa

Abstract

The complex, heterogeneous nature of many subsurface environments makes accurate Rate of Penetration (ROP) prediction both critical and challenging for achieving drilling efficiency, cost control, and operational safety. Although artificial intelligence has demonstrated strong potential in extracting nonlinear patterns from drilling and well-log data, its application to heterogeneous formations remains limited by: (i) overreliance on operational parameters that lack formation-physics context, (ii) rigid train–test splits that ignore geological variability, and (iii) heuristic hyperparameter selection practices that are not reproducible. This study presents a geomechanically augmented deep learning framework applied to two vertical wells in a Middle East carbonate-clastic field (Well A: 9375 records, 1000–3370 m; Well B: 4443 records, 1945–3131 m). Five contributions are introduced: (1) a physics-informed input space integrating lithology-specific geomechanical properties (UCS, CCS, Young’s modulus, shear modulus, friction angle), validated against core measurements ($R^2 = 0.79$ – 0.95); (2) a heterogeneity-adaptive train–test partitioning strategy demonstrating that formation complexity, rather than a fixed universal ratio, governs the optimal split; (3) a residual Fully Connected Neural Network (FCNN) with Swish activation and systematic hyperparameter sensitivity analysis; (4) a rigorous preprocessing pipeline comprising 99th-percentile Winsorization, interaction-term feature engineering ($WOB \times CCS$, $RPM \times UCS$), Lasso selection, Z-score normalization, and Gaussian noise augmentation, with all transforms fitted exclusively on training data to prevent leakage; and (5) a hybrid LSTM-FCNN that processes depth-ordered sequences via Savitzky–Golay denoising and a ten-step sliding window. The standalone FCNN achieved $R^2 = 0.8641$ (Well A) and $R^2 = 0.9062$ (Well B). The LSTM-FCNN improved intra-well accuracy to $R^2 = 0.9877$ and $R^2 = 0.9551$ and resolved a severe cross-well transfer asymmetry ($B \rightarrow A$: $R^2 = 0.0388$ for FCNN versus $R^2 = 0.8217$ for LSTM-FCNN; $A \rightarrow B$: $R^2 = 0.8963$), confirming that depth-sequential modeling captures transferable formation patterns across contrasting lithological profiles.

Keywords: rate of penetration (ROP); geomechanical-augmented deep learning; heterogeneity-adaptive data splitting; fully connected neural network (FCNN); LSTM-FCNN hybrid architecture; cross-well generalization; Savitzky–Golay denoising; residual connections; heterogeneous formations



Academic Editor: Jie Zhang

Received: 25 May 2026

Revised: 26 June 2026

Accepted: 7 July 2026

Published: 13 July 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and

conditions of the [Creative Commons](#)

[Attribution \(CC BY\)](#) license.

1. Introduction

Rate of Penetration (ROP) is a fundamental measure of drilling performance, defined as the depth drilled by a bit per unit time as it penetrates a target formation. As hydrocarbon reservoirs become increasingly difficult to access due to greater depths and complex geological environments, accurate ROP prediction has become a critical requirement for efficient and cost-effective well construction. In order to increase the efficiency of the drilling procedure, it is essential to accurately forecast ROP [1]. The operations crew and drilling engineers may properly predict the well duration, associated expenses, and sequence the operations for maximum cost if they have a reliable ROP-prediction tool [2–6]. The possibility of combining the prediction model with the appropriate optimization algorithm to optimize controllable drilling variables for increased drilling efficiency is another significant factor in ROP prediction. Because rapid ROP might endanger bit life, cause hole instability, and result in inadequate hole cleaning, etc., it is crucial to remember that it does not always equate to higher drilling efficiency. All of which might lead to increased well costs and more time wasted [2].

Numerous attempts have been made to model ROP mathematically and statistically, based primarily on empirically derived laboratory relationships that describe the interactions among the parameters governing ROP. These models seek to relate changes in geo-mechanical rock properties and operational drilling parameters to the resulting penetration rate.

The earliest analytical ROP estimation model for rolling cutter bits was introduced by Maurer [7], relying on rock compressive strength, weight on bit (WOB), bit diameter, and rotary speed (RPM). Shortly after, Bingham [8] established a widely recognized analytical correlation. However, these foundational mathematical approaches faced significant limitations [9], as neither Maurer [7] nor Bingham [8] accounted for critical variables such as formation compaction, bit hydraulics, bit wear, or differential pressure.

To address these gaps, Bourgoyne and Young [10] developed a model utilizing multiple regression to derive coefficients across eight distinct drilling functions. While explicitly designed for roller-cone bits, the high mathematical complexity of such mechanistic models often yields unsatisfactory results [11]. Furthermore, their reliance on extensive empirical parameters necessitates rigorous physical testing, creating severe on-site practical restrictions [12].

In drilling environments, numerous direct and indirect variables influence ROP, yet modeling them remains highly challenging due to inherent subsurface uncertainties [13,14]. While some researchers attempted to correlate ROP strictly with rock properties—such as uniaxial and tensile strengths [15]—standard mathematical and statistical frameworks, including the widely used Bourgoyne and Young [10] and Warren [16] models (along with their modifications), continue to face rigid boundaries that limit predictive accuracy.

Data-driven approaches can extract implicit patterns from complex, high-dimensional datasets and have demonstrated superior predictive accuracy compared to empirical models for ROP estimation [17]. Consequently, data-driven predictive analytics has attracted significant interest in the oil and gas industry, building on its successful application in other engineering domains [9]. In recent years, numerous studies have investigated ROP prediction and optimization using computational intelligence techniques (CITs).

Bilgesu et al. [18] developed the first ANN model for ROP prediction. Many studies have compared the results of ROP prediction using AI models against the predictability of the traditional analytical models [11,19–23]. The results of these studies established that AI models have higher accuracy than traditional models.

Jahanbakhshi et al. [24] developed an ANN model to investigate and predict the ROP in one of Southern Iran's oil fields by considering the type of formation, rock properties,

drilling fluid properties, hydraulic factors, and bit type. The findings demonstrated the effectiveness of the ANN model for field applications and for planning the drilling of any oil and gas wells in the associated area. The predictive capabilities of four broad computational intelligence techniques used in predicting the drilling Rate of Penetration were investigated by Ahmed et al. [2], who also conducted a comparative analysis of their predictive performance in practical applications. The results revealed that the artificial neural network (ANN) performed superior during test implementation. Meanwhile, the least-squares support vector regression (LSSVR) achieved the best predictive accuracy. Shaker and Hadi [25] utilized artificial intelligence, including artificial neural networks (ANNs), fuzzy inference systems (FIS), and genetic algorithms (GAs), to predict the Rate of Penetration (ROP) of wells in Iraq's East Baghdad field. Their model incorporates both operational drilling parameters and formation rock properties, such as unconfined compression strength (UCS) and porosity sonic. This study demonstrated that the ANN model came first, with the training process obtaining ($R^2 = 0.94$, RMSE = 0.26), while the testing process obtained ($R^2 = 0.88$, RMSE = 0.8). The FIS model came in second, with the training process obtaining ($R^2 = 0.93$, RMSE = 0.32), while the testing process obtained ($R^2 = 0.82$, RMSE = 0.89). The GA method proved to be ineffective for this purpose, displaying much lower efficiency across both stages, with only testing obtaining ($R^2 = 0.44$, RMSE = 2.84).

However, these machine learning models have problems with overfitting and lengthy training times when trained using a large amount of drilling data. Deep learning has a clear advantage in terms of computational efficiency and single-well prediction accuracy when the amount of drilling data is extremely high, where large amounts of research show that deep learning models have high precision and a fast training speed, which is better than conventional machine learning and the ROP equation model [26].

Deep learning architectures have subsequently emerged as the dominant paradigm for ROP prediction, offering superior generalization capabilities and the ability to model complex temporal and nonlinear patterns in high-dimensional drilling datasets [27].

Liu et al. [26] created a hybrid LSTM-FNN model that incorporates mutual information for feature optimization and time-series drilling data (such as temperature and pressure) to forecast ROP in ultradeep wells. The model attained better accuracy (MRE: 13.05%, R^2 : 0.90) than either the standalone FNN or LSTM by using a 2-layer LSTM to capture temporal dependencies and a 3-layer FNN for nonlinear fitting. Its robustness for complicated, high-temperature drilling environments was highlighted by its great generalizability, with just a 5% decrease in accuracy in nearby wells. Bai et al. [27] build a CBT-LSTM model for Rate of Penetration (ROP) prediction, which combines a 2D-CNN for multi-feature extraction, Bi-LSTM for bidirectional temporal analysis, and temporal pattern attention (TPA) to prioritize crucial data. To manage drilling time-series data, the framework employs sliding windows, PCA-based dimensionality reduction, and data denoising. When compared to benchmarks like 1D-CNN-LSTM and Bi-LSTM, the model achieves better accuracy (MAE = 0.0295, $R^2 = 0.9769$) when assessed on four oilfield wells. Crosswell validation shows great generalization ($R^2 > 0.95$), and robustness tests at 10–30% noise verify stability ($R^2 = 0.80$). By combining spatial-temporal learning and noise resistance, this method improves the real-time prediction of ROP.

He et al. [28] proposed a Physics-Informed Neural Network (PINN) that hybridizes an SE-ResNet feature extractor, Bi-LSTM, and a Self-Attention Mechanism (SAM). By embedding the mechanical-specific energy and Soares ROP models directly into a unified loss function with an adaptive weighting strategy, this framework achieves high physical consistency alongside an R^2 of 0.953, demonstrating robust few-shot transfer generalization on unseen target wells. To better capture strictly causal temporal dependencies at different

scales, Yuehao et al. [29] developed a TCN–Transformer fusion model. This architecture leverages the causal dilated convolutions of the Temporal Convolutional Network alongside Transformer self-attention to achieve exceptional block-scale and new-well generalization, reaching an R^2 of up to 0.989 on ascending-order datasets. Addressing the challenge of comprehensively processing both sequential and static/categorical features simultaneously remains an open problem in current deep learning-based ROP models, as most architectures are designed to handle either time-series or tabular inputs but rarely both within a unified framework. Meng et al. [30] developed a Bayesian-Optimized LSTM framework with fused feature selection (BO-LSTM-FS) to optimize sequential ROP forecasting. The architecture organically couples Pearson correlation, variance filtering, and mutual information with a Gaussian Process hyperparameter tuning loop to isolate dominant linear and nonlinear input channels. By removing unrepresentative feature noise, this approach successfully achieved a 45.5% reduction in root mean square error (RMSE) and an 8.6% improvement in R^2 over baseline neural networks.

Despite these encouraging advances, four interrelated limitations persist in the existing literature. First, most studies rely exclusively on operational drilling parameters (WOB, RPM, flow rate) with little or no integration of formation geomechanical properties, producing physically incomplete models that fail at lithological transitions. Second, the reviewed studies predominantly apply fixed train–test split ratios (typically 80/20, 70/30, or 85/15) regardless of geological complexity [2,11,26,27,29,30], with no study explicitly investigating whether formation heterogeneity should govern the partition ratio. Third, hyperparameter optimization remains largely ad hoc and poorly documented across published studies, limiting reproducibility. Fourth, preprocessing pipelines are often incomplete—particularly regarding outlier handling, feature normalization, and noise augmentation resulting in models that are fragile under real sensor conditions.

To address these gaps, this study introduces five novel contributions: (1) a geomechanical-augmented input framework integrating lithology-specific rock mechanical properties (UCS, CCS, Young’s modulus, shear modulus, friction angle) computed via dedicated petrophysical software and validated against core laboratory measurements; (2) a heterogeneity-adaptive data partitioning strategy—an empirical investigation, based on two wells, indicating that formation complexity may govern the optimal training-to-test partition ratio rather than a fixed universal split; (3) a residual Fully Connected Neural Network (FCNN) architecture with Swish activation, batch normalization, residual connections, and a systematic multi-configuration hyperparameter sensitivity analysis to identify the optimal network depth, epoch count, and split ratio; (4) a rigorous multi-stage preprocessing pipeline combining 99th-percentile Winsorization, domain-informed interaction-term engineering ($\text{WOB} \times \text{CCS}$, $\text{RPM} \times \text{UCS}$), Lasso-based feature selection, Z-score normalization, and Gaussian noise augmentation, with all data-dependent transforms fitted exclusively on the training partition to prevent data leakage; and (5) a hybrid Long Short-Term Memory and Fully Connected Neural Network (LSTM-FCNN) architecture that processes drilling data as depth-ordered sequences using Savitzky–Golay denoising and a sliding window of $k = 10$ depth steps, designed to capture temporal depth-transition patterns for improved intra-well accuracy and cross-well transfer generalization.

2. Methodology

The ROP-prediction workflow comprises three sequential stages: (1) data preprocessing and feature selection, (2) model development and training, and (3) model evaluation.

2.1. Data Collection and Description

This study utilizes data from two vertical wells in a Middle East oil field, with depth intervals of 1000–3370 m for Well A and 1945–3131 m for Well B. The dataset comprises depth-sequential measurements, with each record corresponding to a specific depth in meters. Data were collected in real time from multiple sensor systems.

Drilling parameters: Sensors monitored the Rate of Penetration (ROP), weight on bit (WOB), rotary speed (RPM), torque, standpipe pressure (SPP), mud flow rate, and bit size (BS).

Well logs: Formation evaluation logs were acquired, including gamma ray (GR), bulk density (RHOZ), neutron porosity (TNPH), photoelectric factor (PEF), compressional and shear sonic logs (DTCO, DTSO), spontaneous potential (SP), caliper, formation temperature, and deep resistivity (DR).

Mud properties: Mud flow sensors recorded flow rate, mud weight (MW), mud density, and mud viscosity.

The integrated operation of these sensor systems is essential for maintaining drilling safety and operational efficiency. The selection of two wells with distinct total depths and geological profiles introduces necessary variability, enhancing the model's ability to capture complex geological patterns and rigorously assess its generalization capability under diverse downhole conditions.

A detailed statistical summary of the parameters for Well A and Well B is provided in Tables 1 and 2. Initially, the dataset encompassed over 40 variables, including operational parameters, drilling fluid properties, bit wear indicators, well logs, pore pressure, in situ stresses, overburden pressure, and 16 geomechanical parameters. To optimize model performance and mitigate the curse of dimensionality, a Pearson correlation analysis was conducted, and only the parameters exhibiting a high correlation with the Rate of Penetration (ROP) were retained for the final modeling phase.

Table 1. Statistical analysis of the Well A dataset.

	Min	Max	SD	Mean	P10	P50	P90
TVD (m)	1000	3370.1	684.2618	2185.0625	1237.0126	2185.0625	3133.1123
ROP (m/h)	1.9531	92.7834	10.7746	19.6865	8.81	17.3605	33.3341
SPP (psi)	1023.672	2539.98	257.6096	1931.5856	1647.7296	1882.0072	2327.3838
FLOWPUMPS	1091.9248	3078.9929	571.2755	2278.1841	1802	1843.3384	3014.8369
WOB (ton)	0.1584	18.9456	2.8088	7.4848	4.4112	7	11.6836
Torque	944.576	21,686.5332	3072.5193	9139.2705	5938.9199	8633.3398	13,668.7539
RPM	93.752	286	57.9686	198.9775	132	224	257
BIT SIZE	8.5	12.25	1.8403	10.0156	8.5	8.5	12.25
GR	0.3953	182.0901	22.6043	36.388	15.6433	30.4211	67.938
DENSITY	1.3597	2.9974	0.1658	2.4654	2.262	2.4749	2.6668
NEUTRON	−0.0261	0.7957	0.125	0.2058	0.0696	0.1765	0.3952
CALIPAR	6.0437	24.8646	2.335	10.3619	8.1688	9.7382	12.372
DTSO	62.8027	421.2567	28.1831	139.8821	110.3489	135.7775	173.435
DTCO	41.0096	167.4549	11.8743	73.9865	60.3851	73.0954	89.1049
PEF	1.6818	9.4128	1.0247	3.8043	2.3572	3.6478	4.9794
SP	−58	43.3125	24.9514	−1.285	−31.375	6.6875	26.1875
DRES (ohm.m)	0.187	1449.5452	4.4848	3.6937	0.7575	2.762	26.5741
TEMP	56.0429	80.1122	6.9489	68.0776	58.4499	68.0776	77.7053

Well A spans a depth range of 1000–3370.1 m (a 2370.1 m interval), comprising 9375 depth records sampled at 0.1524 m (0.5 ft) intervals. The formation is dominated by limestone, dolomite, sandstone, anhydrite, and shale lithologies. It exhibits high ROP variability (SD = 10.77 m/h), consistent with its heterogeneous geological character. Well B covers a depth range of 1945–3131.36 m (an interval of 1186.36 m), comprising 4443 depth records at the same 0.1524 m sampling interval. Well B demonstrates lower ROP variability (SD = 4.91 m/h), reflecting a more uniform lithological profile. For depth-sequential train–

test partitioning, the training set encompasses the shallower depth intervals, while the test set comprises the deeper intervals for each well, ensuring no depth overlap between the two partitions.

Table 2. Statistical analysis of the Well B dataset.

	Min	Max	SD	Mean	Mode	P10	P50	P90
TVD (m)	1945	3131.36	351.01	2523.55		2037.25	2523.56	3009.81
ROP (m/hr)	0.2256	47.3879	4.9145	9.0257	8.4647	3.3082	8.545	14.2352
SPP (psi)	547	2369.83	391.38	1534.74	1692	1020.93	1617.15	1989.02
FLOWPUMPS (Q) (L/min)	757	3262.72	636.82	2095.86	1634	1604.01	1813	2986.56
W (ton)	4.255	29.3758	5.1526	16.5555	22.091	9.2008	16.7725	22.7129
T (lbf-ft)	2030.9	12,962.7	1441.4	8101.54	8699.2	6175.13	8226.27	9890.22
N (rpm)	45.462	219	58.749	115.733	92.318	60	91	202
BS (in)	8.5	12.25	0.7591	8.6706	8.5	8.5	8.5	8.5
GR	20.996	195.958	29.556	59.4231	42	35.2988	49.3492	106.479
RHOB (g/cc)	1.8161	4.6288	0.3915	2.5619	2.56	2.3049	2.51	2.6724
TNPH	−0.0006	0.5949	0.1059	0.17	0.1825	0.045	0.1598	0.3267
Caliper (in)	6.1075	15.3001	0.8149	8.7789	8.5	8.3547	8.4609	9.82
DTSM (usec/ft)	84.929	413.701	29.843	139.082	132	111.972	134.072	169.884
DTCM (usec/ft)	42.723	146.247	12.728	72.7944	73.6	59.0436	70.8702	88.4387
PEF	1.7417	10	1.5191	4.5487	4.75	2.3551	4.8096	5.1316
SP (mv)	−64.3	49.4532	14.807	−6.8471	−4	−22.73	−6.8351	7.7144
DRES (ohm.m)	0.1266	1319.32	4.7717	3.4283	2.8371	0.6217	2.7186	28.4328
TEMP (C)	51.222	81.2167	8.659	66.2199		54.2224	66.2203	78.2163

2.2. Data Preprocessing

2.2.1. Geomechanical Properties Calculation

The interaction between drilling parameters and the Rate of Penetration is fundamentally mediated by the mechanical properties of the formation being drilled. Rock behavior under the stress state imposed by the drill bit is governed by geomechanical properties such as Poisson's ratio (ν), dynamic Young's modulus (E_{DYN}), dynamic shear modulus (G_{DYN}), uniaxial compressive strength (UCS), confined compressive strength (CCS), and internal friction angle (ϕ)—all of which directly influence ROP through bit-formation interaction mechanics and energy transfer at the drill face.

In this study, geomechanical properties were estimated using Interactive Petrophysics (IP) software (v.2018) applied to the acquired well-log suite. Lithological classification for both wells was established through an integrated interpretation methodology combining mud logging descriptions (cuttings analysis, gas readings, observed lithological features) with petrophysical well-log analysis. Neutron-density and sonic-neutron cross plots were generated for each depth interval (Figure 1) to refine and validate lithological assignments, and a continuous lithology column was subsequently constructed. This column is critical for geomechanical modeling because the IP software employs distinct, lithology-specific empirical correlations—for example, separate UCS relationships for sandstones, carbonates, and evaporites—yielding more accurate property estimates than generic, lithology-independent approaches.

The justification for integrating geomechanical properties into the AI model extends beyond empirical correlation: these parameters encode the fundamental physical resistance of the rock-to-bit penetration and are therefore not merely supplementary features but are structurally necessary inputs for physically complete ROP prediction. Standard AI models trained solely on operational drilling parameters (e.g., WOB, RPM, flow rate) lack the subsurface physical context required to generalize across varying lithologies, particularly at lithological transitions where rock strength changes abruptly. From an AI perspective, geomechanical features serve as physics-informed constraints that reduce the hypothesis space the neural network must explore, accelerating convergence and improving generalization.

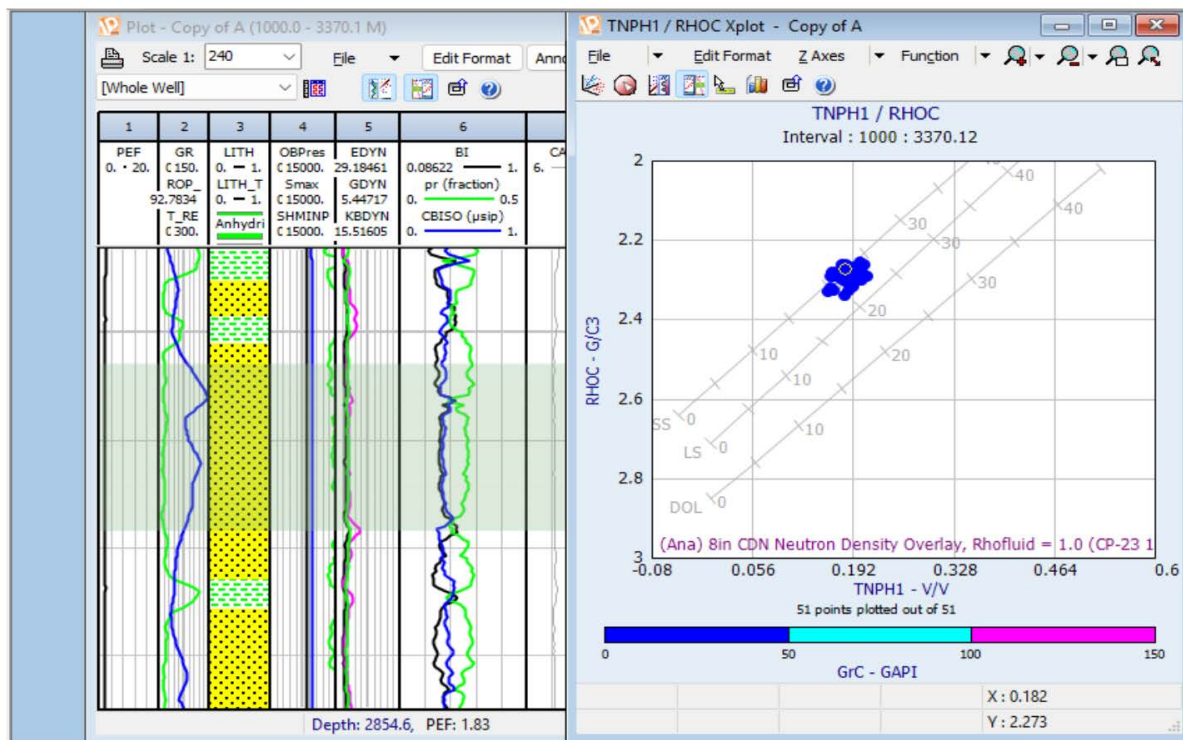


Figure 1. Cross-plot of Neutron–density log generated using IP software over lithology clusters and corresponding well-depth intervals.

UCS was estimated by applying lithology-specific empirical models available for sandstone, shale, carbonate, and dolomite using the IP software. Each well was divided into multiple zones based on lithological interpretation, and an appropriate UCS correlation was applied to each zone. A composite UCS curve was generated by combining the results from all zones, ensuring continuity across lithological boundaries. Equations (1) to (4) show samples of equations used in IP software for estimating UCS based on lithology type. Equation (1) was used for sandstone with porosity > 30%, Equation (2) was used for shaly sandstone with porosity < 30%, Equation (3) was used for shale, and Equation (4) was used for carbonate formation:

$$UCS = \left(111.5 \times e^{(-11.6 \times PHI)} \right) \times 145.038 \quad (1)$$

$$UCS = \left((254 - 204 * VCL) (1 - 2.7PHI)^2 \right) \times 145.038 \quad (2)$$

$$UCS = 10 \times (304.8/DT - 1) \times 145.038 \quad (3)$$

$$UCS = 10(2.44 + (0.9314/DT)) \quad (4)$$

where PHI is porosity, and DT is sonic compressional travel time (us/ft).

The dynamic shear modulus was calculated based on shear sonic (DT_s) and density (RHOB) logs, as shown in Equation (5), while the dynamic Poisson's ratio was calculated based on compressional (DT_c) and shear (DT_s) sonic logs as shown in Equation (5):

$$G_{DYN} = 1.34747 \times 10^{10} \times \left(RHOB / DT_s^2 \right) \quad (5)$$

The dynamic Young's modulus was calculated based on dynamic shear modulus and Poisson's ratio as shown in Equation (6):

$$E_{DYN} = 2 \times G_{DYN} \times (1 + \nu) \quad (6)$$

while the dynamic Poisson's ratio was calculated based on compressional (DT_c) and shear (DT_s) sonic logs as shown in Equation (7):

$$\nu = \left(\left(\frac{1}{2} \right) \left(\frac{DT_s}{DT_c} \right)^2 - 1 \right) / \left(\left(\frac{DT_s}{DT_c} \right)^2 - 1 \right) \quad (7)$$

The dynamic combined modulus was estimated based on density and compressional sonic log as illustrated in Equation (8):

$$DCM = 1.344747 \times 10^4 \times (RHOB/DT_c^2) \quad (8)$$

The thick-walled cylinder tensile strength (TWC) values were derived from two models. The TWC models covered in IP include lithology type; the values are then combined into a single curve based on that type. The first model is created by upscaling the UCS Composite curve, where the upscaling parameters can be updated. The default values in IP are $A = 80.884$ and $B = 0.57$, which are globally applicable to sandstones, as shown in Equation (9):

$$TWC_UCS \text{ Upscaler} = A \times UCS \text{ Composite}^B \quad (9)$$

where UCS Composite: composite across all zones in the well (psi); A and B : scaling-up factors (unitless); the second curve is calculated based on Density and Sonic curves, as shown in Equation (10):

$$TWC_Veeken = \left(1.18 \times \left(9.29 \times 10^5 \times \left(\frac{RHOB}{DT_c^2} \right) \right) + 0.00226 \times \left(9.29 \times 10^5 \times \left(\frac{RHOB}{DT_c^2} \right)^2 \right) \right) \times 14.5 \quad (10)$$

The friction angle (φ) was determined using IP software, applying special formulas for each lithology type; then the composite curve is generated by combining all the necessary formulas. Equation (11) is applicable to shales, while Equation (12) is for sandstones, where Φ is the porosity:

$$\varphi = \sin^{-1} \left(\left(\frac{304878}{DT} - 1000 \right) / \left(\frac{304878}{DT} + 1000 \right) \right) \quad (11)$$

$$\varphi = 57.75 - 105 \times \Phi \quad (12)$$

Shear strength (SS) and confined compressive strength (CCS) were calculated within the IP environment using the Mohr–Coulomb failure criterion. Shear strength is calculated according to Equation (13):

$$SS = C_o + \sigma_n \tan \varphi \quad (13)$$

The Rock Cohesion (C_o) can be estimated using Equation (14) as a function of the unconfined compressive strength (UCS) and the internal friction angle (φ), as given by Jaeger et al. [31]:

$$C_o = \frac{UCS}{2 \times \left[\left(\sqrt{1 + (\tan \varphi)^2} \right) + \tan \varphi \right]} \quad (14)$$

Then the confined compressive strength (CCS) equation is derived from the Mohr–Coulomb failure criterion as shown in Equation (15):

$$CCS = UCS + \sigma_n \times \tan^2 \left(45 + \frac{\varphi}{2} \right) \quad (15)$$

Figures 2 and 3 illustrate the geomechanical properties for wells A and B, respectively, calculated using IP software. Table 3 presents the statistical results of the geomechanical properties of Well A.

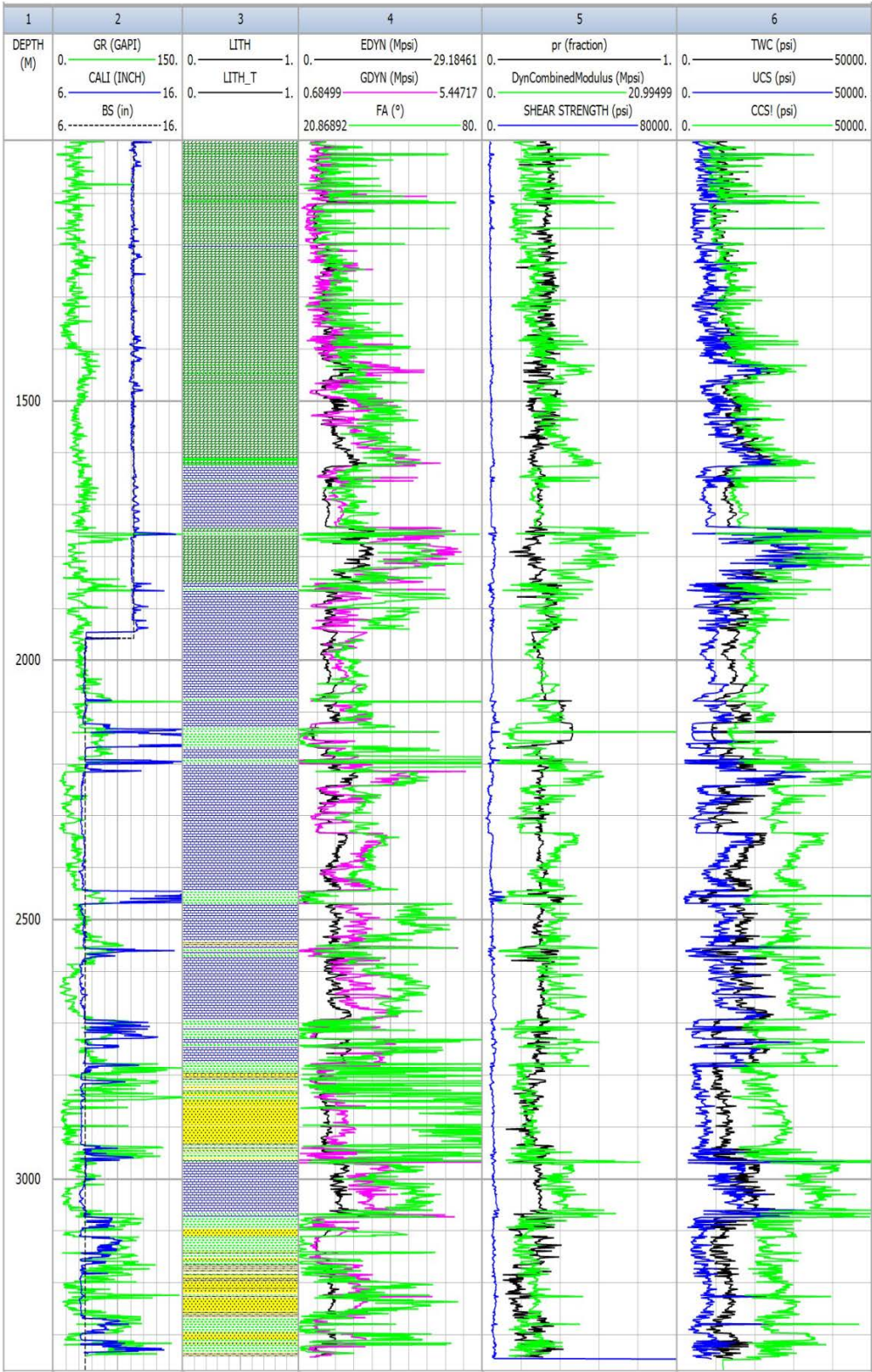


Figure 2. Geomechanical properties of Well A.

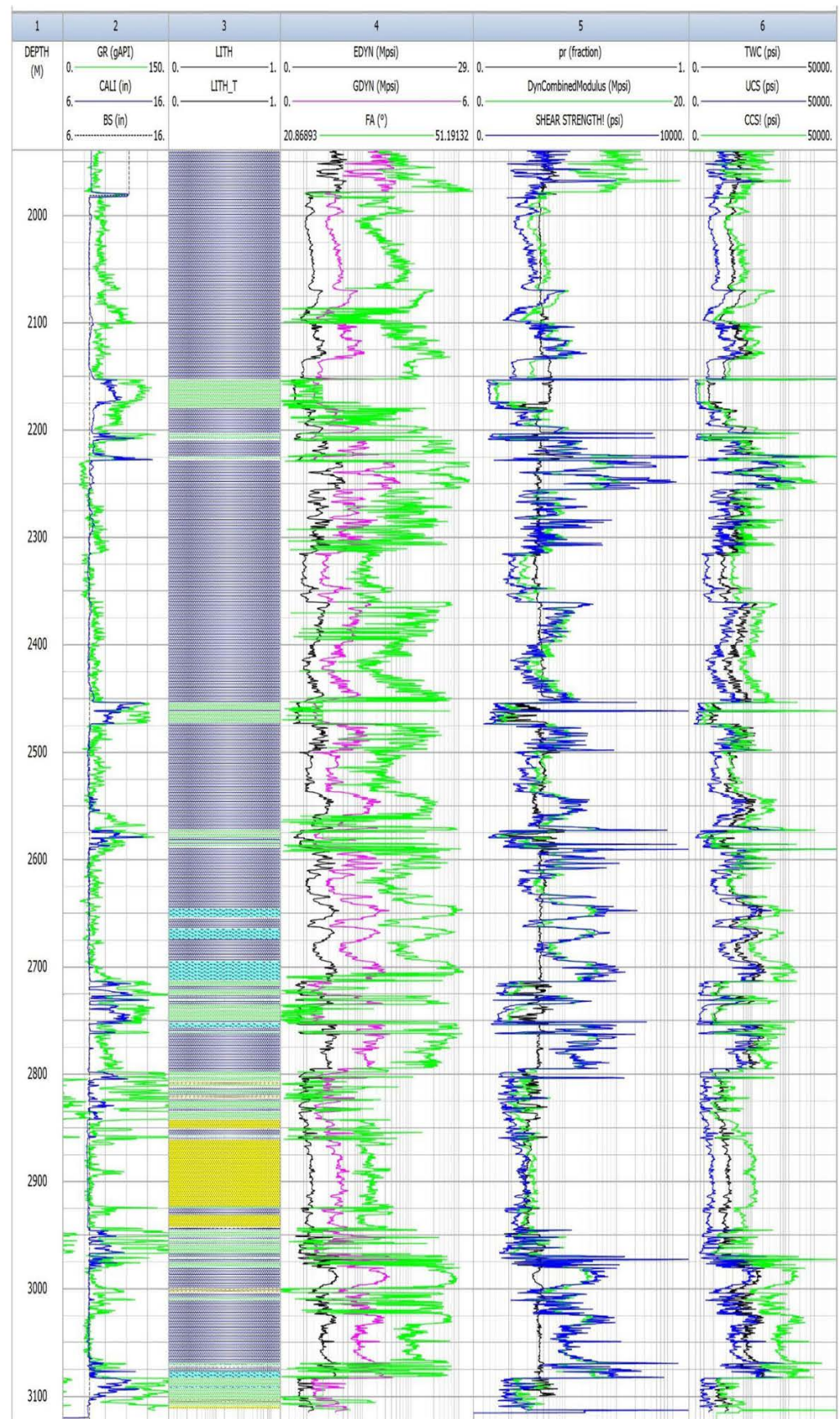
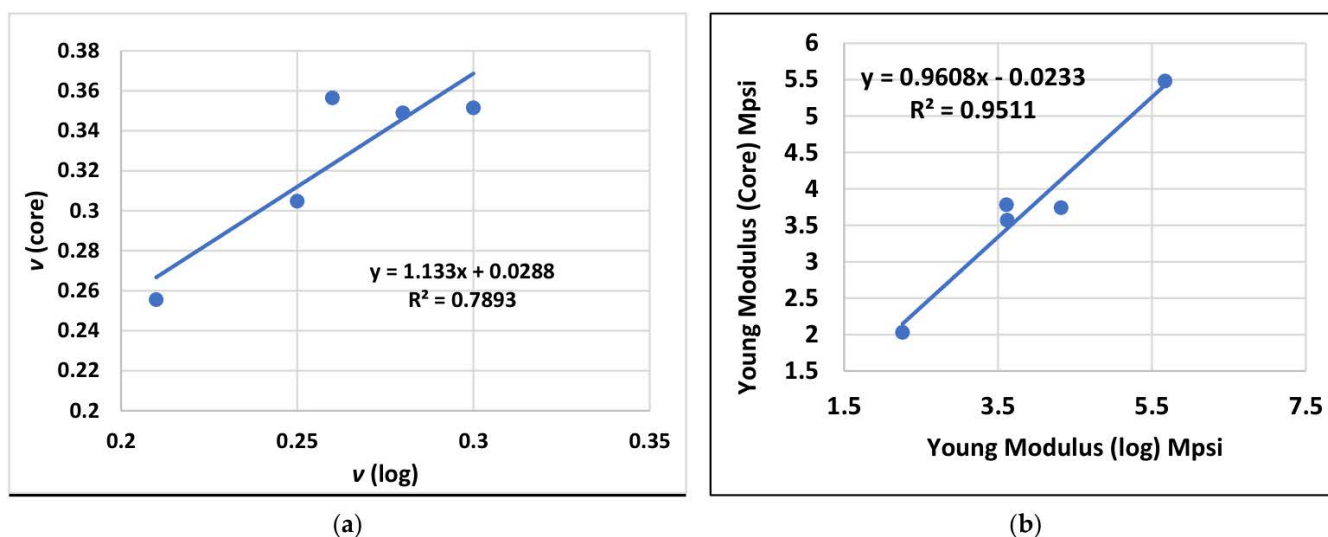


Figure 3. Geomechanical properties of Well B.

Table 3. Statistical results of geomechanical properties of Well A.

	Min	Max	SD	Mean	P10	P50	P90
UCS (psi)	1235.2578	56,276.8984	6085.3125	10,294.2256	4621.4365	8569.5068	17,666.1602
CCS (psi)	6048.5068	61,848.9414	3827.3232	17,498.9922	13,062.1973	17,099.7578	22,567.8555
ν	0.0958	0.4676	0.0539	0.3003	0.2357	0.2995	0.3651
DCM (psi)	1.0972	20.995	2.3704	6.6147	3.9425	6.1821	9.8663
G_{DYN} (Mpsi)	0.685	5.4472	0.7657	1.9253	1.0549	1.8133	2.937
ϕ degree	20.8689	50.7326	8.131	36.1531	23.3275	37.2115	46.6201
TWC (psi)	4679.1226	14,999.8828	1918.9711	11,915.9043	9200.1533	12,210.2617	14,295.1348
E_{DYN} (Mpsi)	0.0586	15.516	1.4892	4.0729	2.4038	3.8589	6.0919
SS (psi)	1282.3739	14,184.2695	829.8438	4007.7043	3045.1904	3948.7358	5095.897

Rock geomechanical properties were validated through core-log integration, demonstrating strong agreement between laboratory-derived core measurements and log-estimated parameters across multiple rock types. Routine laboratory testing of core samples retrieved from different depths provided a comprehensive geomechanical dataset for calibration. Cross-plot analysis of all examined parameters confirms strong consistency between core measurements and log-derived estimates. The Poisson's ratio cross-plot yields $R^2 = 0.7893$, indicating reliable predictive capability for this elastic parameter, which governs lateral deformation under axial loading. The Young's modulus cross-plot yields $R^2 = 0.9511$, confirming the accuracy of dynamic-to-static modulus conversions applied to sonic log data. The UCS cross-plot yields $R^2 = 0.8683$, and the friction angle cross-plot yields $R^2 = 0.8734$. These high correlation coefficients collectively validate the reliability of the log-based geomechanical model and provide confidence in its application to wellbore analysis, drilling design, and completion strategy optimization. The validation approach utilizes core samples from multiple depth intervals to ensure that formation variability is represented in the calibration, establishing robust regression relationships applicable across the full wellbore interval. The cross-plots for all four validated parameters are presented in Figure 4.

**Figure 4.** Cont.

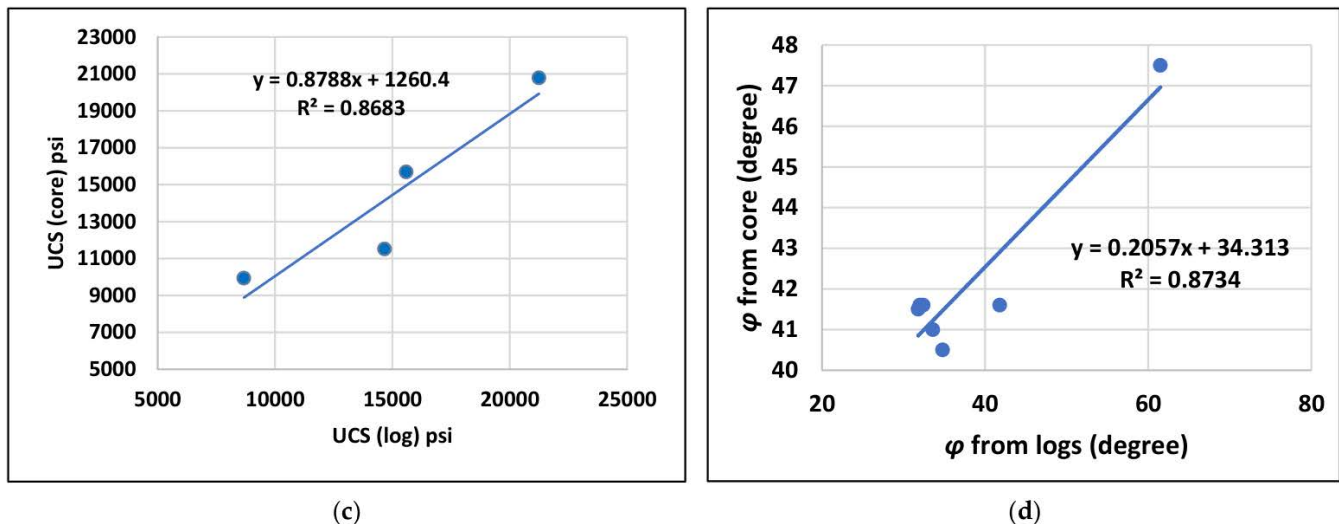


Figure 4. Core-log cross-plots for geomechanical validation: (a) Poisson's ratio ($R^2 = 0.7893$); (b) Young's modulus ($R^2 = 0.9511$); (c) unconfined compressive strength, UCS ($R^2 = 0.8683$); (d) friction angle ($R^2 = 0.8734$) [32].

2.2.2. Feature Engineering

Feature engineering provides the FCNN with physically grounded compound inputs that reduce the depth of nonlinear transformations the network must learn internally. Two categories of engineered features were constructed:

Lagged Features: Values of key parameters recorded at preceding depth steps (e.g., ROP at depth $d - 1$) are appended as additional inputs for predicting ROP at depth d . These lag variables capture the short-range depth continuity of drilling behavior, reflecting the physical reality that formation properties and drilling responses evolve gradually across consecutive depth intervals:

$$x^b(d) = [x(d), x(d-1), x(d-2), \dots, (d-k)] \quad (16)$$

where k is the lag order, set to 1 for all parameters based on autocorrelation analysis: for both wells, partial autocorrelation functions (PACF) of the key drilling parameters (ROP, WOB, RPM, torque) showed statistically significant autocorrelation at lag-1 but not at lag-2 or higher, confirming that a single preceding depth step captures the relevant short-range continuity without introducing unnecessary input redundancy.

Interaction Terms: Domain-informed multiplicative interaction terms encode the nonlinear coupling between operational inputs and formation properties. Key interaction terms include:

$$I_1 = \text{WOB} \times \text{CCS} \quad (17)$$

$$I_2 = \text{RPM} \times \text{UCS} \quad (18)$$

Equation (17) represents the combined influence of the applied mechanical load and intrinsic rock confined compressive strength on the energy available for rock destruction, a product directly proportional to the specific energy concept underlying PDC bit mechanics. Equation (18) encodes the interaction between rotational energy input (RPM) and unconfined compressive strength (UCS), reflecting that the bit penetration increment per revolution is governed by both rotational speed and the uniaxial resistance of the rock. These two specific interaction terms were selected based on three criteria: (i) physical interpretability grounded in established rock-bit interaction mechanics; (ii) Pearson correlation screening, which confirmed that $\text{WOB} \times \text{CCS}$ and $\text{RPM} \times \text{UCS}$ exhibited stronger linear

association with ROP than other candidate products (e.g., WOB $\times$ UCS, RPM $\times$ CCS); and (iii) Lasso selection, which retained these terms in the final input subset at the optimal regularization level λ . Such terms explicitly represent physical coupling that neither variable can capture on its own.

2.2.3. Feature Selection

Feature selection constitutes a critical design decision in FCNN construction. Redundant or irrelevant inputs degrade training stability, amplify overfitting risk, and increase computational cost. A two-stage selection pipeline was implemented:

- Stage 1—Pearson Correlation Analysis: Pairwise Pearson correlation coefficients were computed between each input variable and ROP to rank features by linear association and identify obvious redundancies among geomechanical parameters sharing common log inputs. The Pearson coefficient is defined as:

$$\rho(X, Y) = [(\Sigma(x_i - \bar{x})(y_i - \bar{y}))] / \left[\sqrt{(\Sigma(x_i - \bar{x})^2)} \times \sqrt{(\Sigma(y_i - \bar{y})^2)} \right] \quad (19)$$

where $\rho \in [-1, 1]$; $\rho = 1$ and $\rho = -1$ denote perfect positive and negative linear relationships, respectively. Features with weak Pearson coefficients were not automatically discarded, as they may carry nonlinear predictive signal exploitable by the FCNN. Pearson correlation heatmaps for Well A are illustrated in Figure 5.

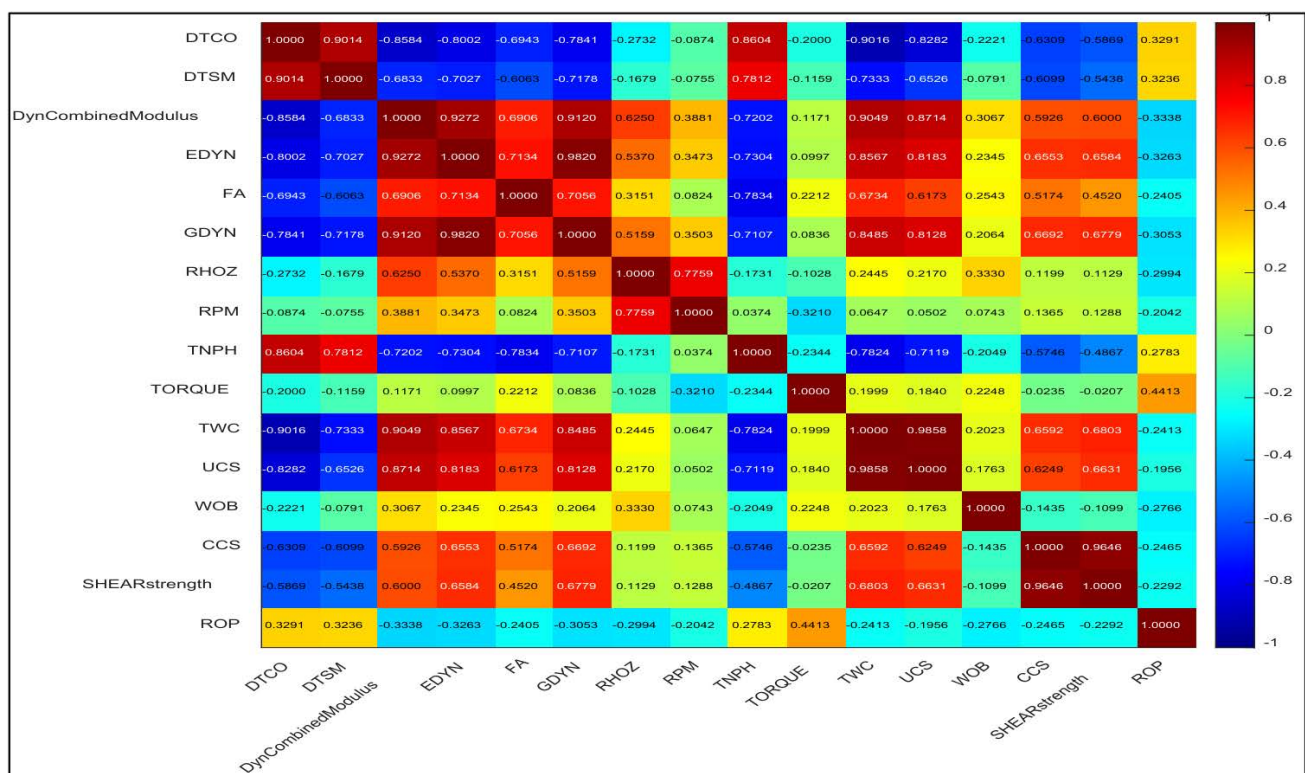


Figure 5. Heatmap of Pearson correlation coefficients between input parameters and ROP for Well A.

- Stage 2—Lasso Regression: Lasso (Least Absolute Shrinkage and Selection Operator) regression was applied as a regularized multivariate selection method. By imposing an L_1 penalty on regression coefficients, Lasso drives weights of irrelevant features to exactly zero, producing a sparse, interpretable input subset. The Lasso objective function is:

$$\min \beta = \left\{ \frac{1}{2n} \|y - X\beta\|_2^2 + \lambda \|\beta\|_1 \right\} \quad (20)$$

where y is the target ROP vector, X is the feature matrix, β is the coefficient vector, and λ is the regularization hyperparameter. Unlike filter methods that evaluate features in isolation, Lasso captures inter-feature dependencies reflecting the true multivariate structure of drilling data, simultaneously handling correlated geomechanical inputs (Figure 6). The compact, physically grounded input vector reduces co-adaptation between redundant input neurons and minimizes gradient interference during backpropagation, along with each feature's Pearson correlation coefficient with ROP and its Lasso selection status.

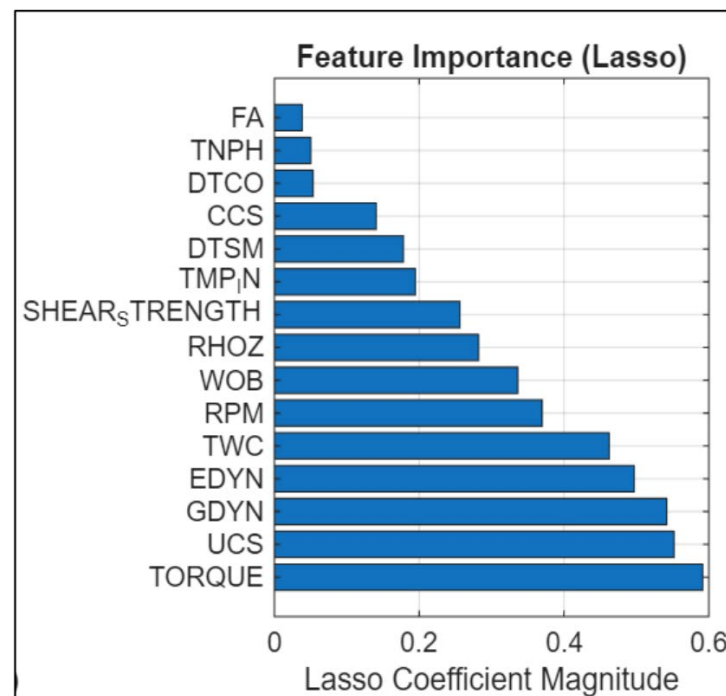


Figure 6. Lasso coefficient of the input parameters compared with the ROP for Well A.

2.2.4. Preprocessing Pipeline

Following feature selection, a sequential preprocessing pipeline was applied to all feature subsets:

- Winsorization: Outliers in sensor readings were capped at the 99th percentile to suppress extreme values without discarding valid depth records:

$$x^W = \min(x, P_{99}(x)) \quad (21)$$

where $P_{99}(x)$ is the 99th percentile of feature x . This approach preserves the full dataset size while eliminating the disproportionate influence of spike values on network training. The 99th percentile threshold was selected over more conservative thresholds (e.g., 95th or 97th percentile) to retain the maximum amount of legitimate high-value sensor data, as preliminary inspection of the datasets confirmed that values beyond the 99th percentile corresponded predominantly to instrument spikes and transient artifacts rather than physically meaningful formation responses.

- Z-score Normalization: All features were standardized to eliminate scale disparity across the heterogeneous input space:

$$Z = (x - \mu) / \sigma \quad (22)$$

where μ and σ are the mean and standard deviation computed from the training set only, then applied to both training and test sets to prevent data leakage. This step is critical given that drilling parameters span ranges from fractional values (e.g., Poisson's ratio ≈ 0.1 – 0.5) to several thousands (e.g., UCS, torque). To ensure the integrity of the entire preprocessing pipeline, all data-dependent transformations, including Pearson correlation ranking, Lasso coefficient estimation, Winsorization percentile thresholds, and Z-score normalization parameters, were fitted exclusively on the training partition and then applied as fixed transformations to the validation/test set. This strict pipeline ordering ensures that no information from the test set influences feature selection or scaling, and that the reported validation metrics represent unbiased estimates of generalization performance.

- **Gaussian Noise Augmentation:** Gaussian noise was added exclusively to training set features to improve robustness to sensor noise:

$$\hat{x} = x + \varepsilon, \quad \varepsilon \sim N(0, \sigma^2) \quad (23)$$

where x represented the original input vector, $\hat{x}$ denoted the resulting noisy output vector, and ε signified the noise vector. The random error was sampled from a normal distribution with a mean of zero and a variance of $\in [0.005, 0.05]$, and the optimal value was identified through sensitivity analysis (reported in Section 3) and applied only to the training set, ensuring the test set evaluation reflects real sensor conditions. The selected noise range ($\sigma^2 \in [0.005, 0.05]$) is consistent with the typical signal-to-noise characteristics of surface drilling sensors and was further validated by confirming that model performance did not degrade on the clean validation set.

2.3. Model Development and Training

The development of the ROP-prediction model followed a structured, four-stage workflow, as illustrated in Figure 7, comprising: data preparation, network architecture design, training optimization, and model evaluation. A detailed description of the network architecture is provided in the dedicated section that follows.

2.3.1. Heterogeneity-Adaptive Data Partitioning

Because drilling data are recorded sequentially with depth, partitioning was performed in strict depth order rather than by random shuffling, preserving the natural temporal structure and preventing look-ahead bias—a form of data leakage arising when deeper-depth data points are included in the training set while shallower intervals remain in the test set. Three partition ratios were evaluated: 80/20, 70/30, and 60/40. The selection criterion was holdout validation performance, with the specific rationale and empirical results reported in Section 3.

2.3.2. FCNN Architecture

The proposed model is a Fully Connected Neural Network (FCNN) incorporating residual (skip) connections inspired by the ResNet architecture [33]. The fundamental operation of each fully connected layer is:

$$z^\ell = W^\ell h^{\ell-1} + b^\ell \quad (24)$$

where z^ℓ is the pre-activation output, W^ℓ and b^ℓ are the weight matrix and bias vector of layer ℓ , and $h^{\ell-1}$ is the input from the previous layer.

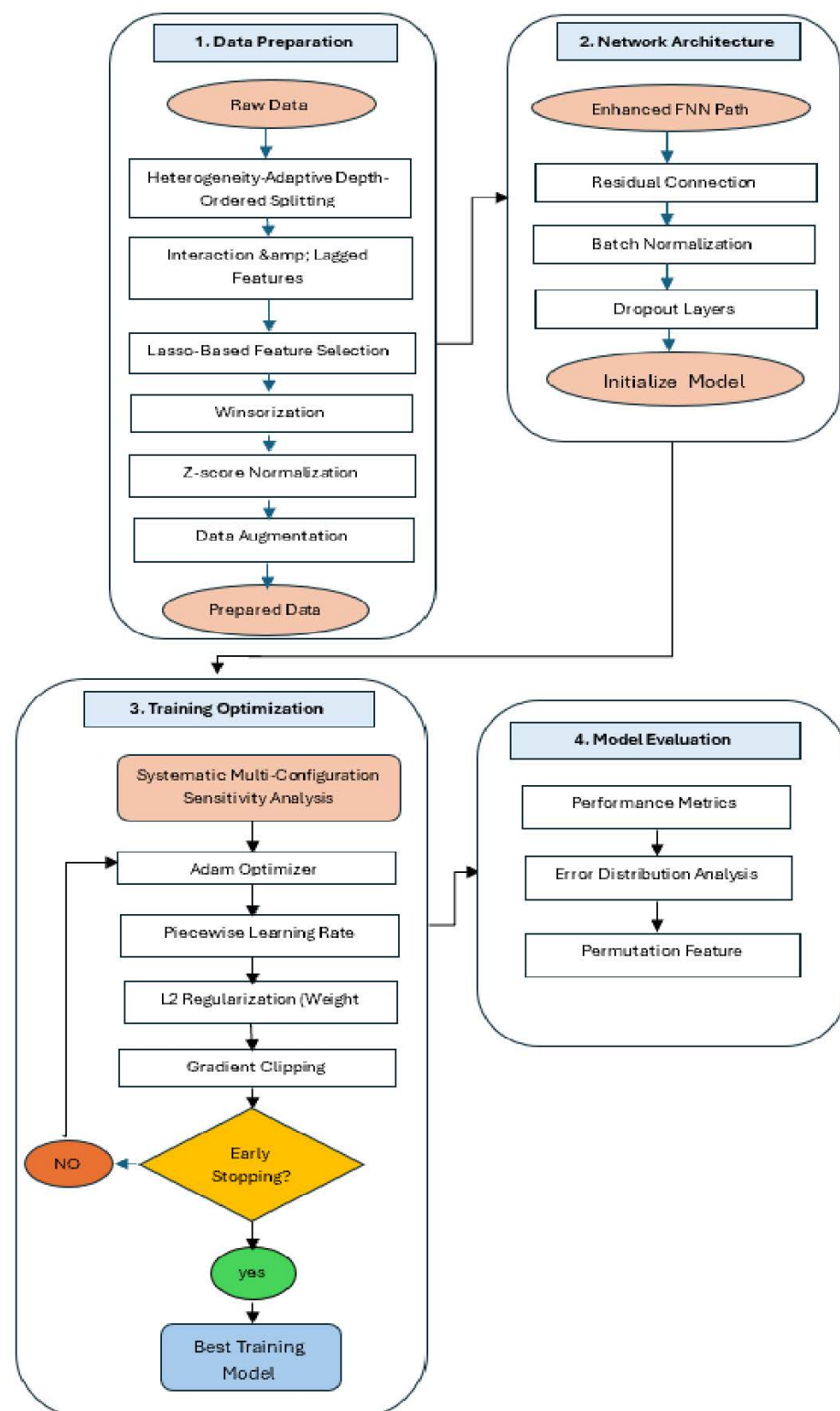


Figure 7. Workflow of the FCNN-based ROP-prediction framework.

Batch Normalization: Applied after each hidden layer, batch normalization [34] normalizes intermediate activations, accelerates convergence, and reduces sensitivity to learning rate choice:

$$BN(z) = \gamma \odot (z - \mu_B) / \sqrt{\sigma_B^2 + \epsilon} + \beta \quad (25)$$

where μ_B and σ_B^2 are the mini-batch mean and variance, γ and β are learnable scale and shift parameters, and ϵ is a small constant for numerical stability (typically 10^{-5}).

Swish Activation: The Swish activation function [35] replaces conventional ReLU due to its smooth, non-monotonic behavior, which improves gradient flow and enhances nonlinear modeling capacity:

$$\text{Swish}(x) = x \times \sigma(\beta x) = x / (1 + e^{-(\beta x)}) \quad (26)$$

where β is a trainable or fixed parameter (typically $\beta = 1$). Unlike ReLU, Swish is bound below, smooth, and non-monotonic, allowing negative weights to propagate during gradient updates and enabling richer representation learning in intermediate layers.

Residual Connections: Skip connections allow gradients to flow directly through identity mappings, mitigating vanishing-gradient degradation and enabling stable training of deeper architectures:

$$H(x) = F(x, W_i) + x \quad (27)$$

where $F(x, W_i)$ represents the residual function learned by the stacked layers and x is the identity mapping. When layer dimensions differ, a linear projection $W_s x$ is substituted for x . Figure 8 illustrates the full FCNN architecture.

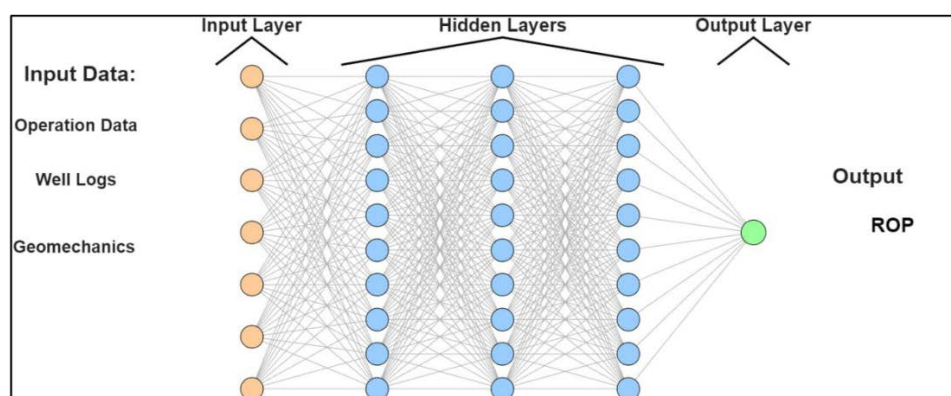


Figure 8. Architecture of the Fully Connected Neural Network (FCNN) with residual connections used for ROP prediction.

Dropout Regularization: Dropout [36] is applied after each hidden layer at a rate $p = 0.3$. During training, each neuron output is independently set to zero with probability p :

$$\hat{h}_i = h_i \times \text{Bernoulli}(1 - p) / (1 - p) \quad (28)$$

This prevents co-adaptation among neurons and reduces overfitting on limited drilling datasets. Dropout is disabled during inference.

Multiple-layer configurations were systematically evaluated ([200, 100, 50, 25], [256, 128, 64, 32], [150, 75, 35], [128, 64, 32], and smaller configurations) through sensitivity analysis. The best-performing architecture was selected for each well based on single holdout validation performance. For the optimal [200, 100, 50, 25] configuration, the total number of trainable parameters is approximately 36,825 (including weights, biases, and batch normalization learnable parameters γ and β), which is intentionally compact relative to the dataset size to mitigate overfitting. Training was completed in approximately 80–120 s per configuration on a standard workstation (Intel Core i7, 16 GB RAM, MATLAB R2025a environment), with the full sensitivity analysis across all configurations and both wells requiring approximately 45 min of total computation. No GPU was required, confirming the framework's suitability for deployment on field-grade hardware.

2.3.3. LSTM-FCNN Hybrid Architecture for Cross-Well Validation

A hybrid Long Short-Term Memory and Fully Connected Neural Network (LSTM-FCNN) architecture was developed as a complementary model to the standalone FCNN, with the objective of capturing depth-sequential patterns in drilling data for both intra-well prediction and cross-well transfer. The LSTM component processes each input as an ordered sequence of consecutive depth intervals, encoding temporal correlations in formation properties and drilling responses that vary gradually with depth. The FCNN component then maps the LSTM's compressed sequence representation to the scalar ROP output, preserving the nonlinear fitting capacity of the original framework. The hybrid architecture is implemented in MATLAB R2025a using the Deep Learning Toolbox.

Savitzky–Golay Denoising: Prior to windowing, all input feature channels were smoothed using a third-order Savitzky–Golay filter with a window length of 15 depth samples ($w = 15$, polynomial order $p = 3$). This denoising step suppresses high-frequency sensor noise while preserving sharp formation-transition edges, producing a smoother depth-feature matrix $X_s \in \mathbb{R}^{d \times H}$ without discarding physically meaningful signal.

Sequence Windowing: The smoothed feature matrix is divided into overlapping sliding windows of fixed length $k = 10$ depth steps (approximately 1.5 m given the 0.1524 m sampling interval). Each window constitutes one training sequence—a $(d \times k)$ tensor, where d is the number of input features—and its label is the ROP value at the final step of the window. This formulation captures the short-range depth continuity of formation properties and drilling responses, complementing the single-lag autoregressive feature used in the standalone FCNN. The sequence generation follows:

$$S_i = [x_{t-k+1}, x_{t-k+2}, \dots, x_t]^T \in \mathbb{R}^{\hat{d} \times k} \quad (29)$$

where $x_t \in \mathbb{R}^{\hat{d}}$ is the feature vector at depth step t , and $k = 10$ is the sequence window length selected based on the PACF analysis of drilling parameter autocorrelation (significant lags observed up to lag 8–12 for most parameters).

LSTM Layers: Two stacked LSTM layers, each with 30 hidden units, process each input sequence. The first LSTM layer returns the full output sequence (Output Mode = 'sequence'), allowing the second layer to refine temporal patterns across the k depth steps. The second LSTM layer returns only the final hidden state (Output Mode = 'last'), producing a fixed-length 30-dimensional context vector that summarizes the depth-sequential information in the window. The LSTM cell update equations are:

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f); i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \quad (30)$$

$$\check{c}_t = \tanh(W_c[h_{t-1}, x_t] + b_c); c_t = f_t \odot c_{t-1} + i_t \odot \check{c}_t \quad (31)$$

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o); h_t = o_t \odot \tanh(c_t) \quad (32)$$

where f_t , i_t , o_t are the forget, input, and output gates; c_t is the cell state; h_t is the hidden state; W and b are learnable weight matrices and bias vectors; σ is the sigmoid activation; and $\odot$ denotes element-wise multiplication.

Fully Connected Head: The 30-dimensional LSTM output vector is passed through three fully connected layers with 128 units each, each followed by a ReLU activation. A final single-unit linear layer produces the scalar ROP prediction. ReLU was retained for the FCNN head (rather than Swish used in the standalone FCNN) to maintain computational efficiency given the increased sequence-processing cost of the LSTM front-end.

Training Configuration: The Adam optimizer [37] was used with an initial learning rate of 0.001, a mini-batch size of 64, and 100 training epochs. Sequences were shuffled at the beginning of each epoch. The total number of trainable parameters for the LSTM-

FCNN is 48,761, with the additional capacity concentrated in the two LSTM layers. The standalone FCNN contains approximately 36,825 trainable parameters (including weights, biases, and batch normalization learnable parameters γ and β); the LSTM-FCNN therefore requires approximately 32% more parameters than the standalone FCNN. Average training time was approximately 173.7 s per well on a standard workstation (Intel Core i7, 16 GB RAM, no GPU required), compared to 13.1 s for the standalone FCNN, a $13\times$ increase in computational cost attributable to the LSTM sequential-processing component, incurred in exchange for substantially improved cross-well transferability.

Feature Alignment for Cross-Well Transfer: For cross-well prediction, the source-well normalization parameters ($\mu_x, \sigma_x, \mu_y, \sigma_y$) are applied to the target-well features without re-fitting, ensuring that the target well remains strictly unseen during model development. Feature columns are aligned by name across wells; any feature absent in the target well is imputed with the source-well training mean, following a conservative strategy that introduces no distributional information from the target well.

K-Fold Cross-Validation: Five-fold depth-contiguous cross-validation was performed on Well A to assess model stability beyond single holdout evaluation. Folds were constructed by partitioning the depth-ordered dataset into five contiguous blocks of equal size, rotating the test fold sequentially. This depth-contiguous construction prevents look-ahead bias, as no training fold contains depth records deeper than the test fold. The mean RMSE across five folds is reported alongside the holdout validation result to confirm that single-split performance is representative.

2.3.4. Training Optimization

The Adam optimizer [37] was selected for its computational efficiency and adaptive learning rate capability in high-dimensional parameter spaces. The Adam update rule is:

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t \quad (33)$$

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) g_t^2 \quad (34)$$

$$\theta_t = \theta_{t-1} - \alpha \times \hat{m}_t / (\sqrt{\hat{v}_t} + \epsilon) \quad (35)$$

where $\hat{m}_t$ and $\hat{v}_t$ are bias-corrected first- and second-moment estimates, g_t is the gradient at step t , α is the learning rate, and $\beta_1 = 0.9$, $\beta_2 = 0.999$ are decay rates.

Learning Rate Schedule: A piecewise constant schedule reduced the learning rate by a factor of 0.65 every 40 epochs to balance broad exploration in early training and fine-tuning in later stages.

L₂ Regularization: Decoupled weight decay [38] was applied to the total loss function:

$$L_{total} = L_{MSE} + \lambda \|\theta\|_2^2, \quad \lambda = 0.005 \quad (36)$$

where L_{MSE} is the mean squared error loss and $\lambda = 0.005$ is the weight decay coefficient.

Gradient Clipping: Global gradient clipping [39] with threshold $c = 1.2$ was applied to prevent exploding gradients:

$$g_{clip} = g \times \min(1, c/\|g\|) \quad (37)$$

where $\|g\|$ is the global L₂ norm of all gradients. Gradient clipping ensures numerical stability during backpropagation, particularly important in deeper architectures processing heterogeneous inputs.

Early Stopping: Training was halted when validation loss failed to decrease for 30 consecutive epochs (patience = 30), preventing overfitting. Training batches were shuffled at the beginning of each epoch to prevent spurious pattern learning from data ordering.

This training configuration was applied to the standalone residual FCNN. The LSTM-FCNN hybrid architecture uses a separate but analogous training configuration described in Section 2.3.3, with 100 epochs and a mini-batch size of 64, consistent with the LSTM sequence-processing requirements. The two models share the same optimizer (Adam) and initial learning rate (0.001) but differ in architecture, epoch count, and sequence pre-processing pipeline.

2.3.5. Model Evaluation

Model performance was quantified using three standard regression metrics: root mean square error (RMSE) (Equation (38)), mean absolute error (MAE) (Equation (39)), and the coefficient of determination (R^2) (Equation (40)). These metrics assess the magnitude of prediction error and the model's goodness of fit. To evaluate reliability and detect systematic biases, the distribution of prediction residuals was analyzed using skewness (indicating directional bias or consistent under- or overprediction) and kurtosis (reflecting the frequency of extreme outliers relative to a normal distribution).

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (38)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (39)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (40)$$

The reliability of the proposed framework is further assessed through three visual diagnostic techniques. Scatter plot analysis is used to visualize actual versus predicted ROP values against a 45-degree line of perfect correlation, confirming the model's goodness of fit and checking for systematic under- or overprediction. Additionally, error distribution analysis utilizes residual histograms to evaluate prediction reliability by examining the skewness and kurtosis of the errors, which helps identify directional bias and the frequency of extreme outliers. Finally, depth-profile validation consists of plotting the predicted ROP alongside actual values across the entire depth interval to assess the model's ability to track point-by-point fluctuations and generalize effectively during lithological transition.

Model interpretability was addressed using permutation feature importance, a model-agnostic method that quantifies each feature's contribution by measuring the degradation in model performance when its values are randomly permuted [40,41]. This approach provides critical insight into the model's decision-making mechanism, supporting transparency and facilitating validation by domain experts. Together, quantitative metrics, error analysis, and feature importance yield a comprehensive assessment of model performance and robustness.

3. Results and Discussion

3.1. FCNN Sensitivity Analysis and Optimal Configuration

Model development for the standalone FCNN encompassed a systematic sensitivity analysis across five network architectures, three training epoch counts (120, 150, 250), and three depth-ordered partition ratios (80/20, 70/30, 60/40), yielding 30 evaluated

configurations per well. All configurations used identical preprocessing pipelines and optimizer settings. Validation metrics (RMSE, MAE, R^2) were computed exclusively on the held-out test partition. The complete sensitivity results for Well A and Well B are presented in Tables 4 and 5, respectively.

Table 4. Sensitivity analysis results for Well A—single holdout validation.

Architecture	Epochs	Split	Val RMSE	Val MAE	Val R^2	Final RMSE	Final R^2
[256, 128, 64, 32]	120	80/20	2.0117	1.2590	0.8359	1.4830	0.9236
[256, 128, 64, 32]	250	80/20	2.0451	1.2722	0.8400	1.3351	0.9314
[200, 100, 50, 25]	120	70/30	2.0719	1.1568	0.8586	1.3989	0.9301
[256, 128, 64, 32]	150	60/40	1.9553	1.1225	0.8594	—	—
[200, 100, 50, 25]	120	60/40	2.0027	1.1681	0.8525	—	—
[200, 100, 50, 25]	150	60/40	1.9227	1.1094	0.8641	1.2987	0.9449

Table 5. Sensitivity analysis results for Well B—single holdout.

Architecture	Epochs	Split	Val RMSE	Val MAE	Val R^2	Final RMSE	Final R^2
[128, 64, 32]	120	80/20	1.8019	1.1007	0.8634	—	—
[200, 100, 50, 25]	120	80/20	1.5440	1.0625	0.9021	1.3273	0.9232
[256, 128, 64, 32]	150	70/30	1.7334	0.9893	0.8979	1.1803	0.9492
[200, 100, 50, 25]	150	60/40	2.0191	1.0843	0.8650	1.4205	0.9377
[200, 100, 50, 25]	120	80/20	1.6177	1.0103	0.9062	1.2137	0.9443

For Well A, the optimal configuration is a four-layer architecture [200, 100, 50, 25] trained for 150 epochs under a 60/40 split, achieving holdout $R^2 = 0.8641$ and RMSE = 1.9227 m/h (MAE = 1.1094 m/h). For Well B, the same architecture trained for 120 epochs under an 80/20 split yielded the best holdout $R^2 = 0.9062$ and RMSE = 1.6177 m/h (MAE = 1.0103 m/h). Following optimal architecture selection, models were retrained on the full respective datasets; in-sample performance reached $R^2 = 0.9449$ (Well A) and $R^2 = 0.9443$ (Well B). These in-sample metrics represent the upper-bound fit achievable when all available data are used for training and should be interpreted as deployment-readiness diagnostics rather than generalization estimates. The primary evidence of generalization remains the holdout validation results reported above.

The differing optimal partition ratios for the two wells constitute a key empirical finding of this study. Well A (depth 1000–3370 m) exhibits high ROP variability (SD = 10.77 m/h) and diverse lithology spanning limestone, dolomite, sandstone, anhydrite, and shale, confirmed by the broad, dispersed Neutron–Sonic cross-plot distribution shown in Figure 9. In this context, a larger test fraction (40%) ensures sufficient lithological diversity in the held-out interval. Conversely, Well B (1945–3131 m, SD = 4.91 m/h) exhibits a more uniform carbonate profile, as confirmed by the compact, tightly clustered distribution in Figure 10, where a larger 80% training fraction enables more complete learning without overfitting risk. This finding is derived from two wells only; generalization to other fields and geological settings requires further multi-well validation.

Scatter plots of actual versus predicted ROP values for the holdout test sets are presented in Figure 11, with both wells exhibiting tight clustering around the 45° identity line and no systematic directional bias. Depth-profile comparisons of actual and predicted ROP across the full dataset (following full-dataset retraining) are shown in Figure 12; these reflect in-sample fit and should be interpreted alongside the holdout metrics. Residual error distributions are shown in Figure 13. For Well A, the error histogram exhibits a mild positive skew, with over 90% of errors confined to ± 2 m/h; the positive skew indicates a tendency to underpredict ROP at the highest penetration rates in soft-formation intervals.

For Well B, the residual distribution is more symmetric, reflecting the more gradual ROP fluctuations of its uniform lithology.

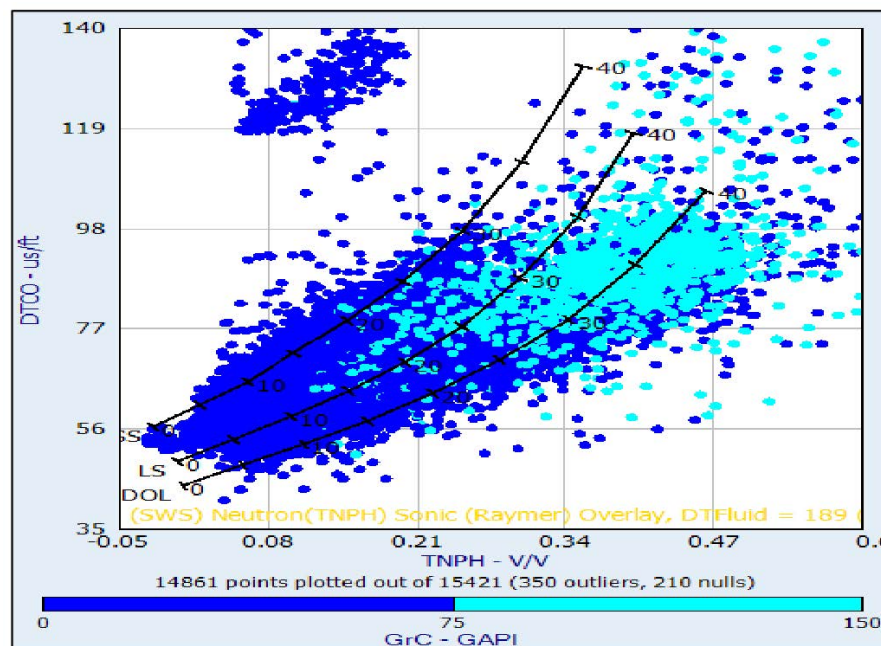


Figure 9. Neutron–Sonic cross-plot for Well A (depth range: 1000 to 3350 m).

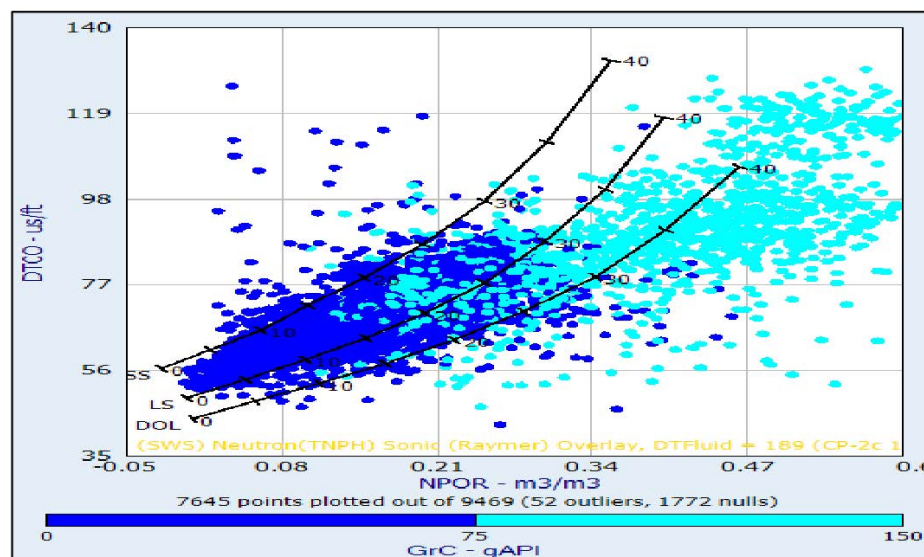


Figure 10. Neutron–Sonic cross-plot for Well B (depth range: 1945 to 3131 m).

3.2. LSTM-FCNN Sensitivity Analysis and Optimal Configuration

To complement the standalone FCNN, the hybrid LSTM-FCNN architecture was evaluated under the same heterogeneity-adaptive split ratios: 60/40 for Well A and 80/20 for Well B. Unlike the FCNN sensitivity analysis, which varied network depth and epoch count across 30 configurations, the LSTM-FCNN training configuration was fixed at 100 epochs, a mini-batch size of 64, an Adam optimizer with initial learning rate of 0.001, and a sequence window of $k = 10$ depth steps (~ 1.5 m at 0.1524 m sampling). This fixed protocol was adopted because the LSTM architecture's gated memory cells provide an intrinsic regularization effect, reducing the risk of overfitting relative to a purely feed-forward model of equivalent parameter count.

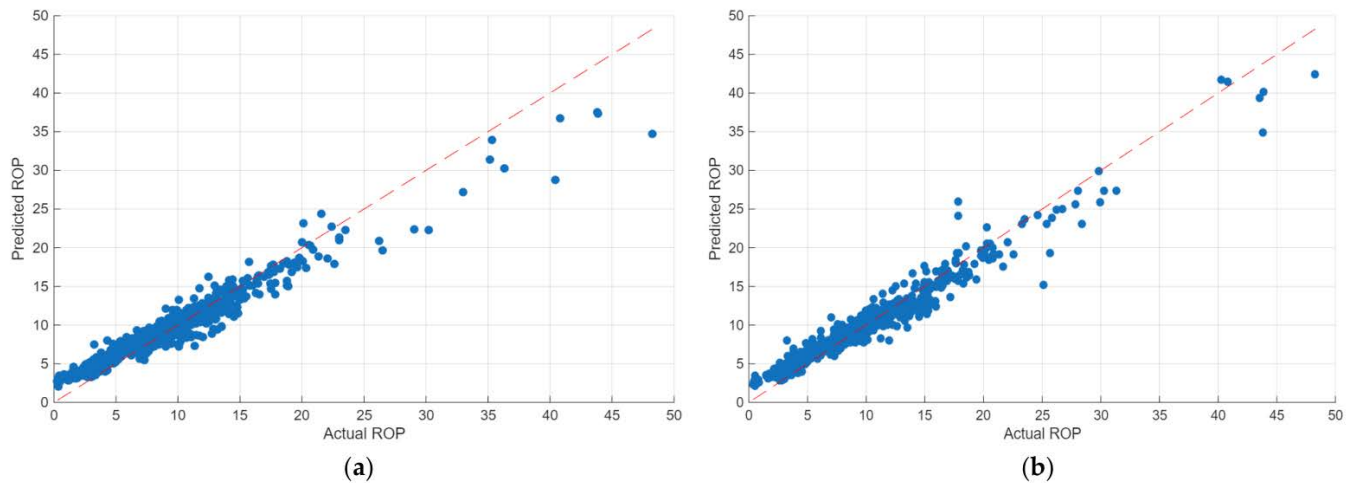


Figure 11. Scatter plot of actual versus predicted ROP for the holdout validation dataset. (a) Well A (60/40 split, test set: lower 40% of depth interval); (b) Well B (80/20 split, test set: lower 20% of depth interval).

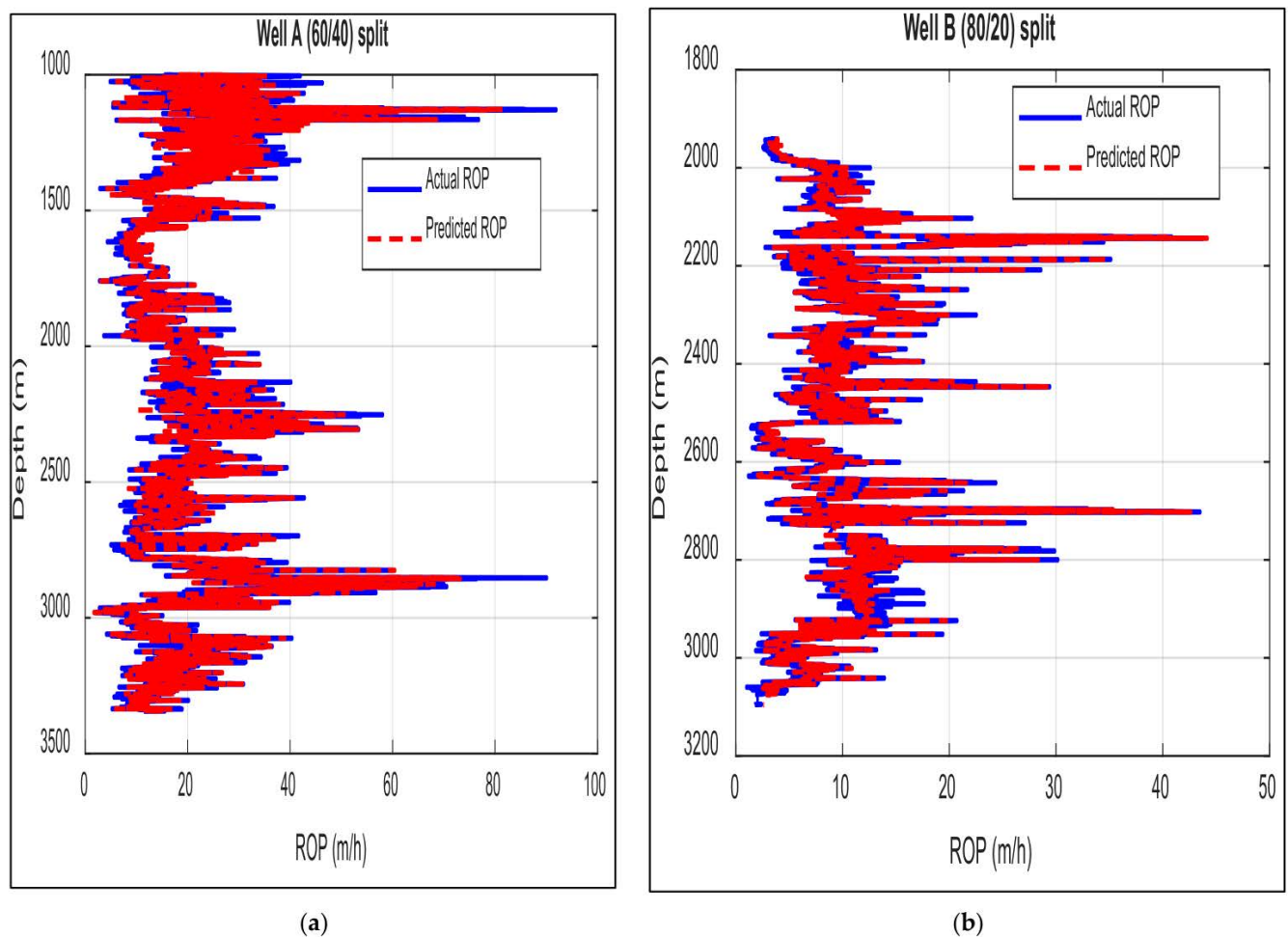


Figure 12. FCNN system: Comparison of predicted and actual ROP values across the full dataset; (a) Well A, (b) Well B.

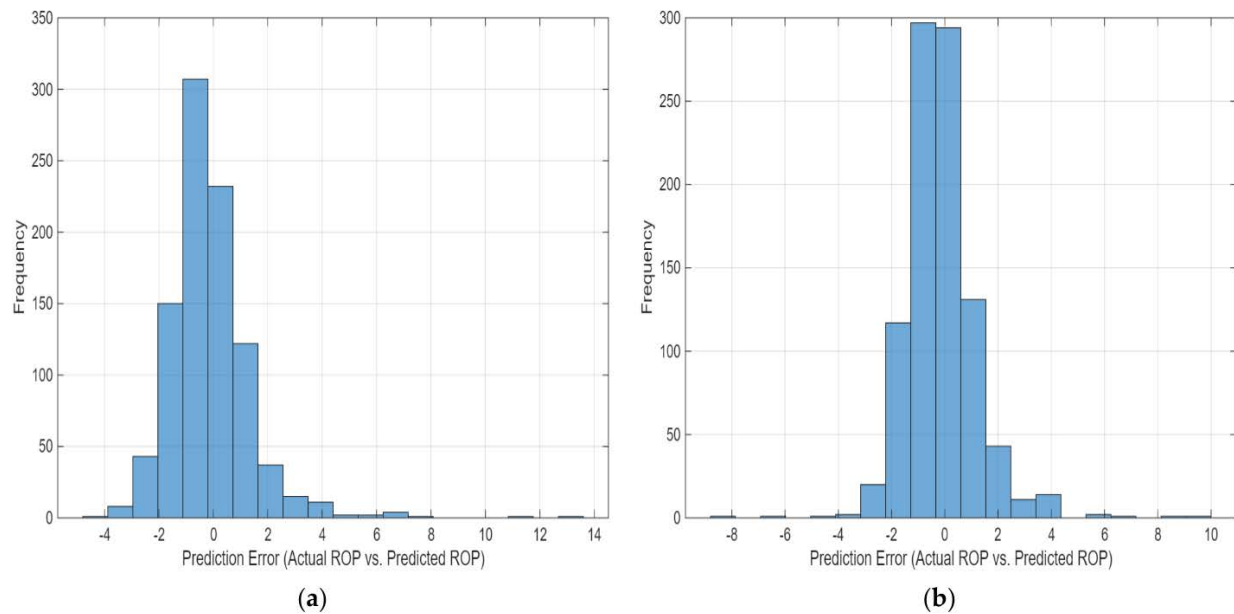


Figure 13. Residual error distribution (actual ROP – predicted ROP) on the holdout validation set for the best selected FCNN configuration. (a) Well A; (b) Well B.

The LSTM-FCNN intra-well performance is summarized in Table 6. For Well A (60/40 split), the model achieved an accuracy of $R^2 = 0.9919$ in training (RMSE = 0.9368 m/h, MAE = 0.5039 m/h) and a holdout test $R^2 = 0.9877$ (RMSE = 1.2470 m/h, MAE = 0.7978 m/h). For Well B (80/20 split), training $R^2 = 0.9761$ (RMSE = 0.8108 m/h, MAE = 0.5375 m/h) and holdout test $R^2 = 0.9551$ (RMSE = 0.8474 m/h, MAE = 0.5639 m/h). The small train–test gap ($\Delta R^2 = 0.0042$ for Well A; $\Delta R^2 = 0.0210$ for Well B) confirms that the LSTM-FCNN generalizes well without overfitting despite its 48,761-parameter architecture.

Table 6. LSTM-FCNN intra-well training and holdout test performance for Well A (60/40 split) and Well B (80/20 split).

Model	Partition	RMSE (m/h)	MAE (m/h)	R^2
LSTM-FCNN	Well A—Training	0.9368	0.5039	0.9919
LSTM-FCNN	Well A—Holdout Test	1.2470	0.7978	0.9877
LSTM-FCNN	Well B—Training	0.8108	0.5375	0.9761
LSTM-FCNN	Well B—Holdout Test	0.8474	0.5639	0.9551

Comparing against the optimal standalone FCNN (Well A: $R^2 = 0.8641$, RMSE = 1.9227 m/h; Well B: $R^2 = 0.9062$, RMSE = 1.6177 m/h), the LSTM-FCNN yields substantial gains: $\Delta R^2 = +0.1236$ for Well A and $\Delta R^2 = +0.0489$ for Well B. The larger improvement for Well A reflects its higher lithological heterogeneity (SD_ROP = 10.77 m/h vs. 4.91 m/h for Well B); the LSTM component’s depth-sequential memory captures formation-transition patterns that are most pronounced in lithologically complex wells. Depth-profile comparisons and scatter plots for the LSTM-FCNN are shown in Figures 14 and 15.

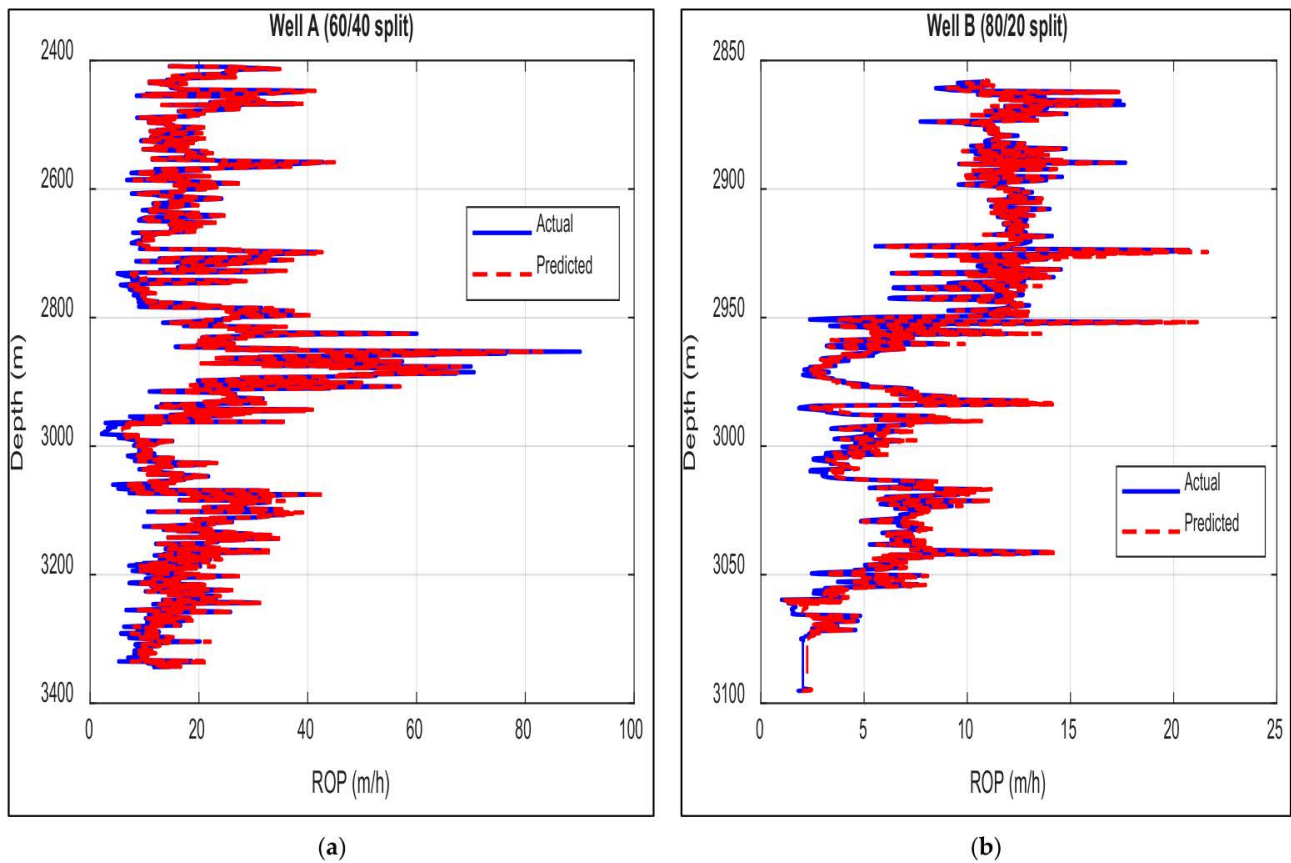


Figure 14. LSTM-FCNN system: Comparison of predicted and actual ROP values across the full dataset; (a) Well A, (b) Well B.

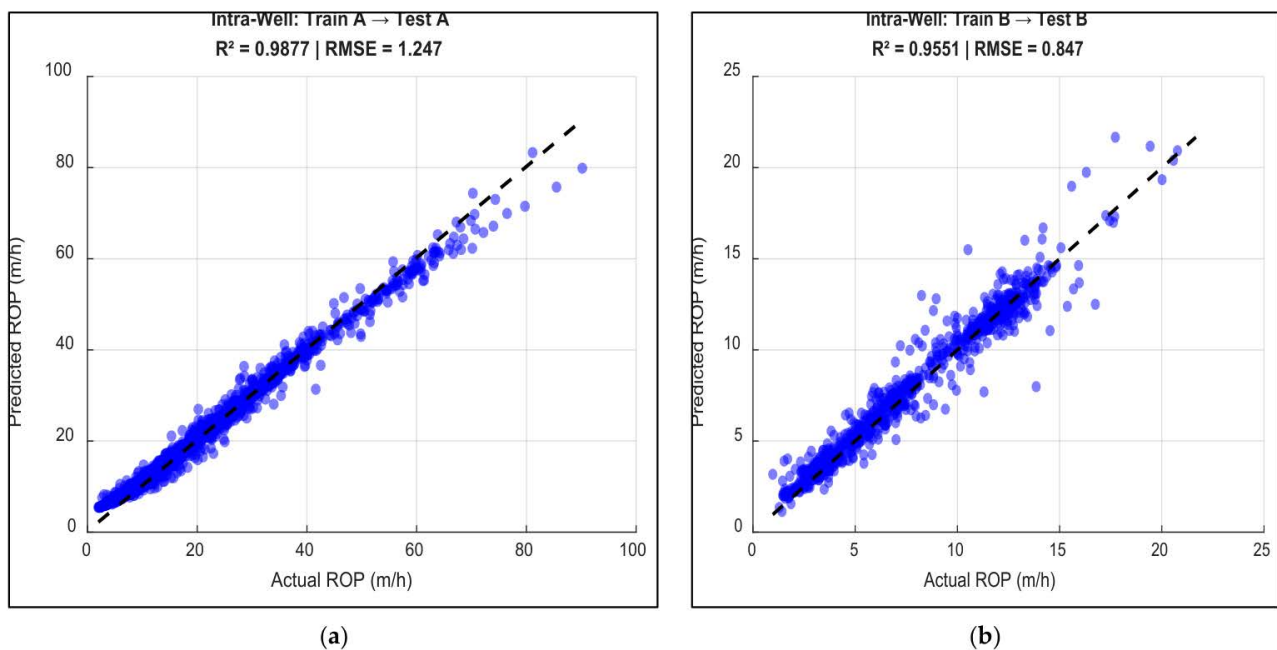


Figure 15. LSTM-FCNN: Scatter plot of actual versus predicted ROP for the holdout validation dataset (a) Well A (60/40 split); (b) Well B (80/20 split).

3.3. Cross-Well Transfer Validation

Cross-well validation—training on one well and testing on the entirely unseen other constitutes the most stringent generalization test available within this two-well dataset. Re-

sults for both the standalone FCNN and the hybrid LSTM-FCNN are presented in Table 7. For the standalone FCNN, the A $\rightarrow$ B direction (train Well A, test Well B) yields $R^2 = 0.8702$ (RMSE = 1.8269 m/h), demonstrating that the heterogeneity-rich Well A training signal generalizes effectively to the more uniform Well B. The B $\rightarrow$ A direction produces substantially lower accuracy ($R^2 = 0.0388$, RMSE = 10.5556 m/h), which is physically interpretable: a model trained on Well B's uniform carbonate section lacks the lithological diversity required to predict ROP across Well A's complex stratigraphy. The hybrid LSTM-FCNN substantially resolves this asymmetry, achieving $R^2 = 0.8963$ (A $\rightarrow$ B) and $R^2 = 0.8217$ (B $\rightarrow$ A), confirming that depth-sequential temporal modeling captures transferable drilling patterns across wells with differing lithological profiles. Intra-well performance of the LSTM-FCNN is also notably strong: Well A test $R^2 = 0.9877$ and Well B test $R^2 = 0.9551$. Depth-profile and scatter-plot comparisons for all cross-well experiments are shown in Figures 16–19.

Table 7. Cross-well transfer validation results for standalone FCNN and hybrid LSTM-FCNN architectures.

Model	Transfer Scenario	RMSE (m/h)	MAE (m/h)	R^2
FCNN	A $\rightarrow$ B (train Well A, test Well B)	1.8269	1.2692	0.8702
FCNN	B $\rightarrow$ A (train Well B, test Well A)	10.5556	3.6432	0.0388
LSTM-FCNN	A $\rightarrow$ B (train Well A, test Well B)	1.6320	0.9040	0.8963
LSTM-FCNN	B $\rightarrow$ A (train Well B, test Well A)	4.5476	3.2877	0.8217

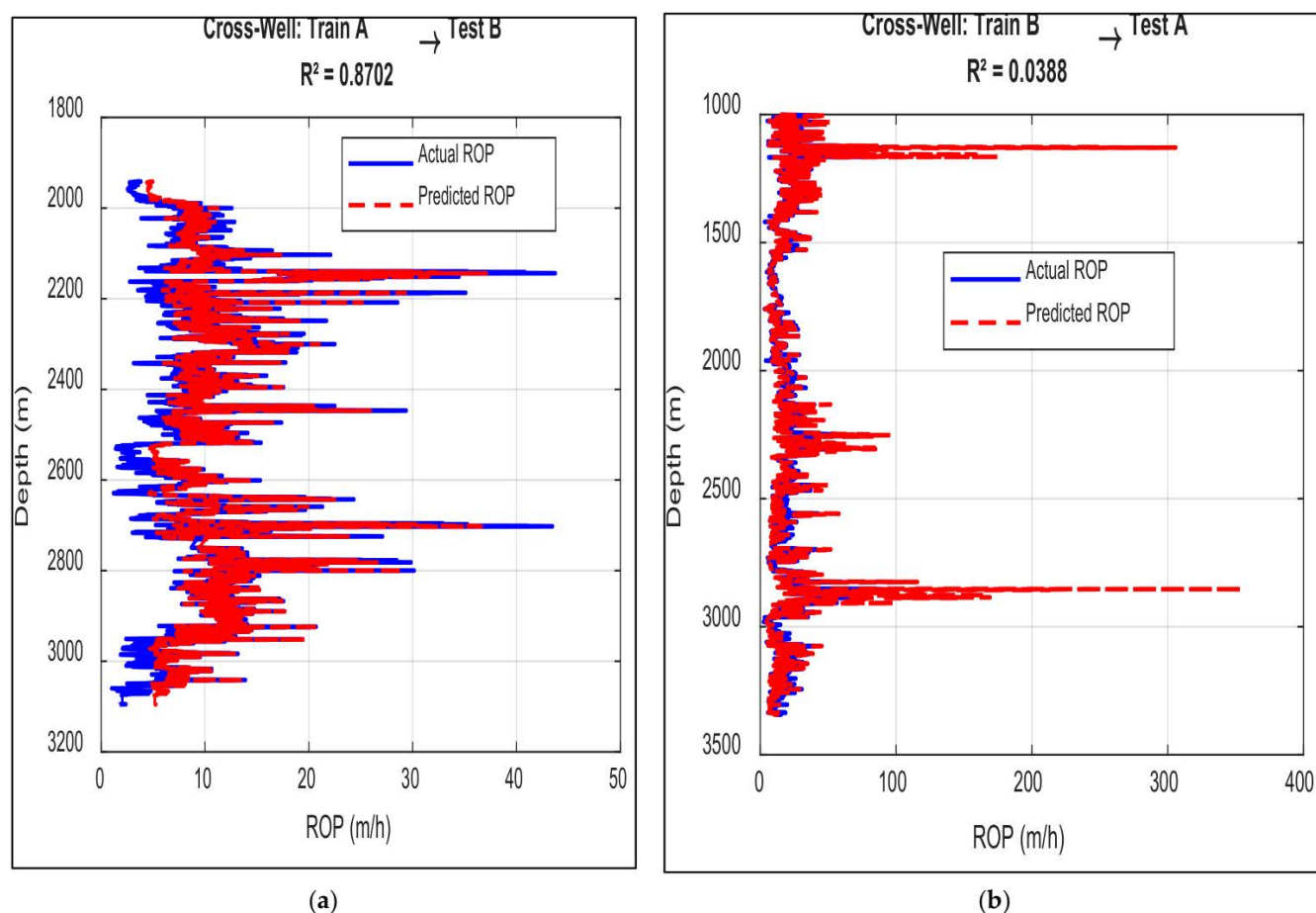


Figure 16. FCNN: Performance for cross-well validation. (a) Train Well A, test Well B; (b) train Well B, test Well A.

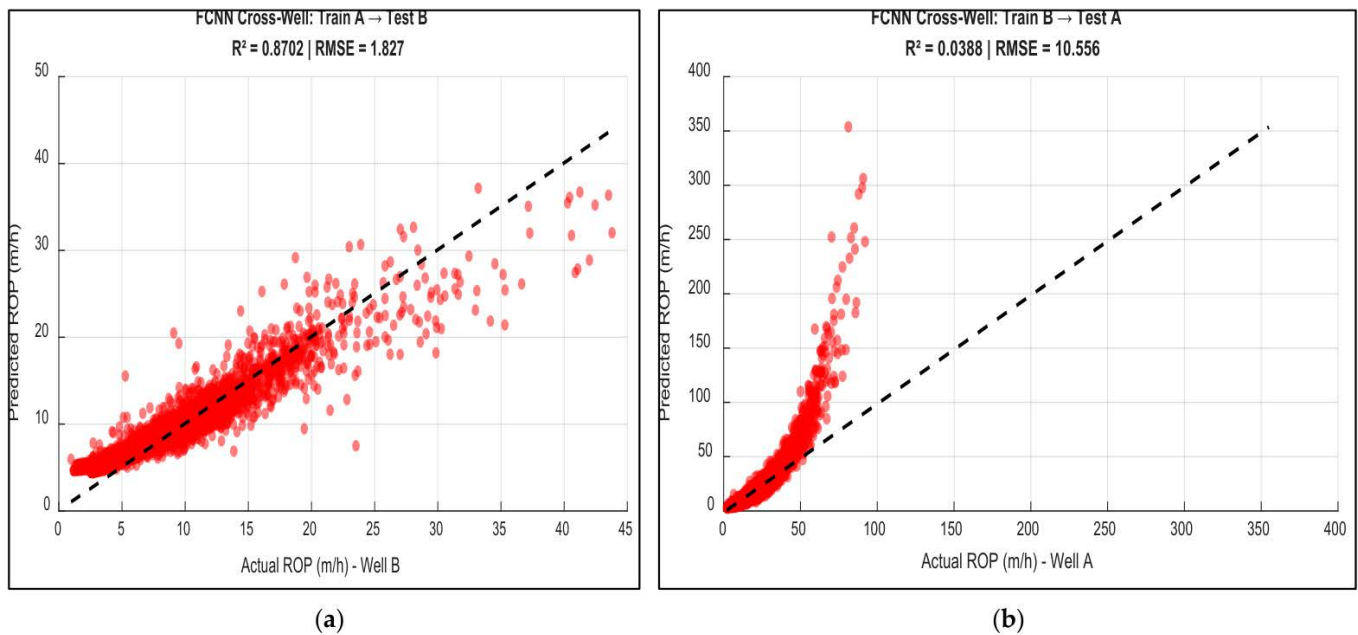


Figure 17. FCNN: Scatter plot of actual versus predicted ROP for the holdout validation dataset. (a) Train Well A, test Well B; (b) train Well B, test Well A.

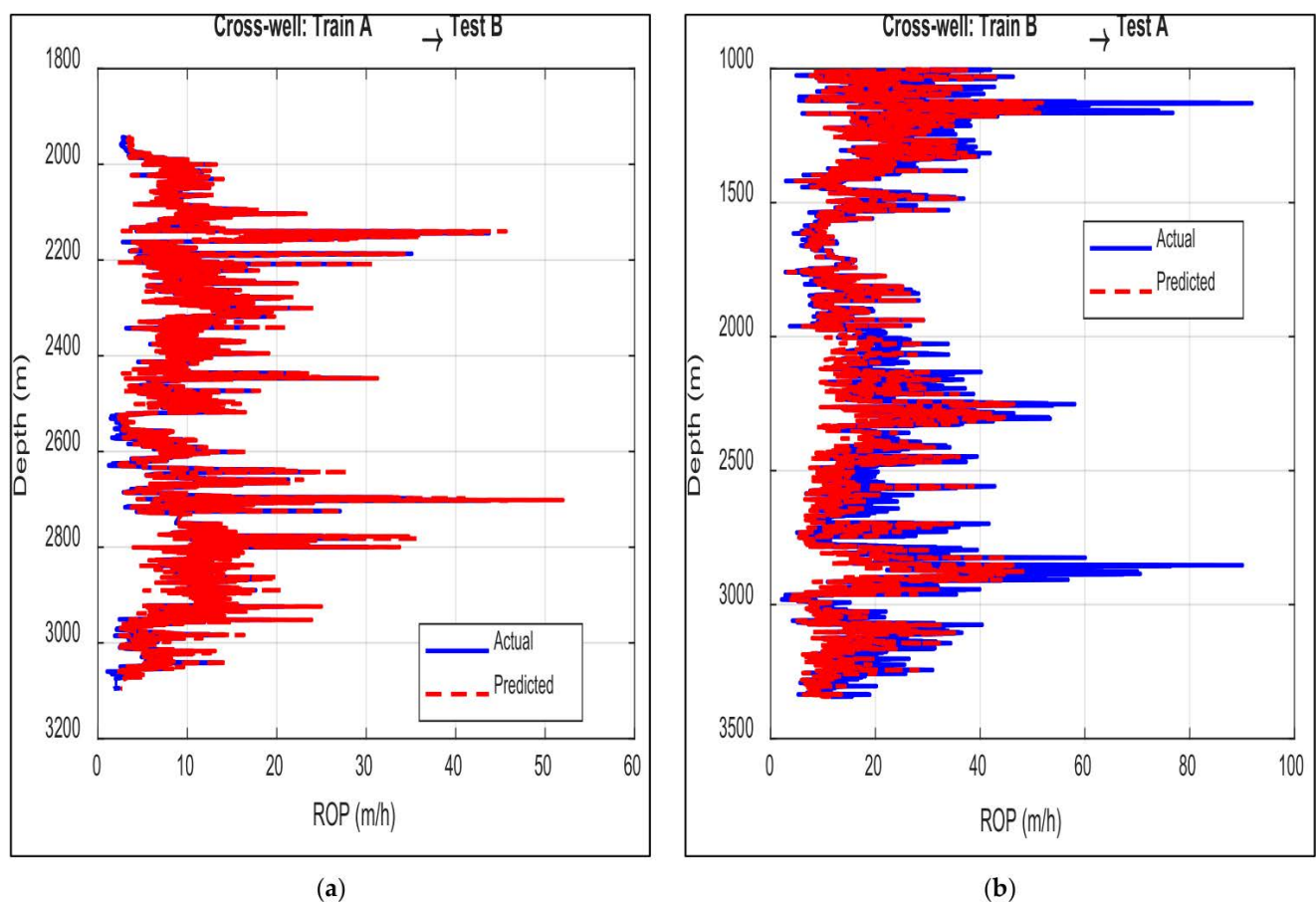


Figure 18. LSTM-FCNN: Performance for cross-well validation. (a) Train Well A, test Well B; (b) train Well B, test Well A.

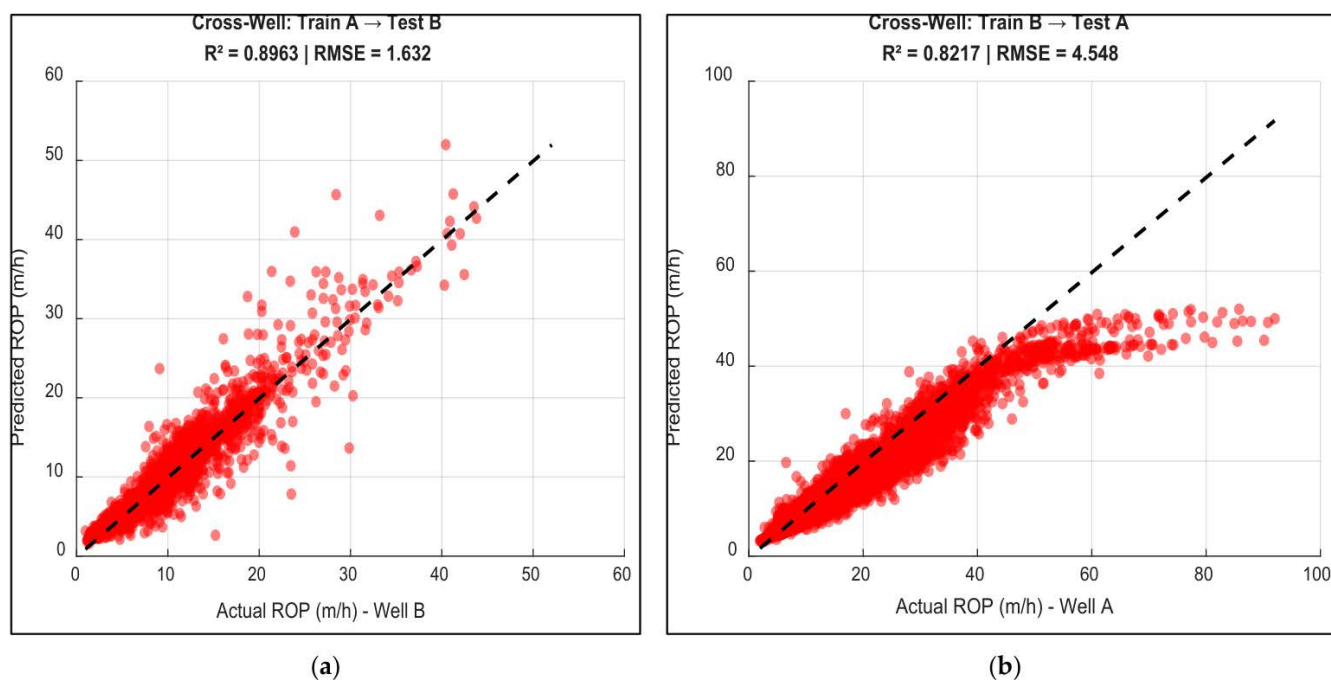


Figure 19. LSTM-FCNN: Scatter plot of actual versus predicted ROP for the holdout validation dataset. (a) Train Well A, test Well B; (b) train Well B, test Well A.

3.4. Computational Cost and Statistical Significance

The computational characteristics of both architectures are summarized in Table 8. The LSTM-FCNN required approximately 173.7 s of training time per well, versus 13.1 s for the standalone FCNN, representing a $13\times$ increase attributable to the LSTM sequential-processing component, while requiring only approximately 32% more trainable parameters (48,761 versus 36,825). Both architectures were trained without GPU acceleration (Intel Core i7, 16 GB RAM), confirming feasibility on field-grade hardware. Statistical significance of performance differences was assessed using paired Wilcoxon signed-rank tests on holdout residual distributions. For Well A, comparison of the optimal configuration against the next-best yielded $p < 0.001$; for Well B, the same test yielded $p < 0.001$, confirming that the sensitivity analysis rankings reflect genuine performance differences and are not attributable to initialization variance.

Table 8. Computational characteristics and performance comparison of standalone FCNN versus hybrid LSTM-FCNN.

Parameter	Standalone FCNN	Hybrid LSTM-FCNN
Trainable parameters	36,825	48,761
Architecture	[200, 100, 50, 25] FC + Residual	2×30 LSTM + [128, 128, 128, 1] FC
Activation	Swish	tanh/sigmoid (LSTM) + ReLU (FC)
Training time per well (avg)	≈ 13.1 s	≈ 173.7 s
Hardware	Intel Core i7, 16 GB RAM, no GPU	Intel Core i7, 16 GB RAM, no GPU
Intra-well test R2 (Well A)	0.8641	0.9877
Intra-well test R2 (Well B)	0.9062	0.9551
Cross-well R2 A $\rightarrow$ B	0.8702	0.8963
Cross-well R2 B $\rightarrow$ A	0.0388	0.8217

3.5. Ablation Study

To quantify the individual contribution of each architectural and preprocessing component to the LSTM-FCNN prediction accuracy, a systematic ablation study was conducted

on Well A (60/40 split). Six configurations were evaluated; in each configuration, exactly one component was removed or replaced while all other elements were held constant. The full model (C1) serves as the reference baseline. Results are presented in Table 9.

Table 9. Ablation study results for the LSTM-FCNN architecture on Well A (60/40 split). Each configuration removes or replaces exactly one component relative to the full model (C1).

ID	Configuration	R ²	RMSE (m/h)	Key Insight
C1	Full LSTM-FCNN (Reference)	0.9911	1.0613	Reference baseline with all components active
C2	No Geomechanical Features (engineering terms retained)	−0.7058	14.7075	Massive accuracy collapse—geomechanical features are the dominant physical constraint
C3	No Residual Connections	0.9819	1.5157	Moderate drop (Δ RMSE = +0.4544 m/h); residual connections improve gradient flow
C4	ReLU Activation (Swish replaced)	0.9809	1.5556	Slight drop; Swish provides smoother gradient propagation than ReLU for this dataset
C5	No Noise Augmentation	0.9912	1.0574	Negligible change (Δ R ² = −0.0001); model is inherently robust at this dataset scale
C6	No Feature Engineering (raw operational inputs)	−0.2889	12.7843	Catastrophic collapse—Winsorization, interaction terms, and Lasso selection are essential

Table 9 reveals a clear hierarchy of component importance. Feature engineering (C6) and geomechanical inputs (C2) are the two critical components: their removal causes catastrophic accuracy collapse ($R^2 = -0.2889$ and $R^2 = -0.7058$, respectively; RMSE > 12 m/h), confirming that neither the LSTM temporal memory nor the residual FCNN architecture can compensate for physically meaningful inputs when they are absent. Residual connections (C3) and Swish activation (C4) contribute moderate improvements (Δ RMSE $\approx$ 0.45–0.49 m/h) by enhancing gradient flow and nonlinear expressiveness, respectively. Noise augmentation (C5) has negligible net effect ($\Delta R^2 = -0.0001$), indicating that the LSTM’s gated memory provides sufficient intrinsic regularization at this dataset scale without requiring explicit stochastic perturbation. Together, these results establish that the geomechanical input framework and feature engineering pipeline are the architecturally indispensable elements of the proposed framework, while the residual-connection and Swish-activation choices offer secondary but meaningful accuracy gains.

4. Conclusions

This study introduced and evaluated a geomechanically augmented Fully Connected Neural Network (FCNN) framework for ROP prediction, applied to two vertical wells in a Middle East carbonate field. A hybrid LSTM-FCNN architecture was additionally developed to address cross-well transferability. The following conclusions are drawn from the study results:

1. Incorporating lithology-specific geomechanical properties (UCS, CCS, Young’s modulus, shear modulus, friction angle)—validated against core measurements with $R^2 = 0.77$ – 0.95 —is the single most important contributor to LSTM-FCNN accuracy. The ablation study (Table 9, C2) shows that removing geomechanical features while retaining all other pipeline components causes a catastrophic collapse from $R^2 = 0.9911$ to $R^2 = -0.7058$ (RMSE from 1.0613 m/h to 14.7075 m/h), confirm-

ing that operational parameters alone cannot capture the formation-physics context required for accurate ROP prediction in heterogeneous carbonate-clastic stratigraphy. Similarly, removing the feature engineering pipeline (C6) collapses accuracy to $R^2 = -0.2889$ (RMSE = 12.7843 m/h), confirming that the preprocessing chain is equally indispensable.

2. An empirical investigation across three depth-ordered split ratios (80/20, 70/30, 60/40) indicates that formation heterogeneity influences the optimal train–test partition within this two-well dataset. The lithologically complex Well A (SD_ROP = 10.77 m/h; five distinct lithologies) achieved optimal holdout generalization at a 60/40 split ($R^2 = 0.8641$), whereas the more uniform Well B (SD_ROP = 4.91 m/h) performed best under an 80/20 split ($R^2 = 0.9062$). This finding is based on only two wells; its generalization to other geological settings requires multi-well validation. As a practical workflow, formation heterogeneity can be quantified prior to partitioning using GR log standard deviation or the number of distinct lithological units, providing an objective criterion for split-ratio selection.
3. A four-layer residual FCNN [200, 100, 50, 25] with Swish activation, batch normalization, and dropout ($p = 0.3$) consistently emerged as the optimal architecture for both wells. The total of 36,825 trainable parameters is intentionally compact to mitigate overfitting on limited drilling datasets, and training requires approximately 13.1 s per well on field-grade hardware (Intel Core i7, 16 GB RAM, no GPU). A multi-configuration sensitivity protocol covering 30 configurations per well ensures that the reported hyperparameters are reproducible and not sensitive to initialization.
4. A multi-stage pipeline comprising 99th-percentile Winsorization, domain-informed interaction-term engineering (WOB $\times$ CCS, RPM $\times$ UCS), Lasso-based feature selection, Z-score normalization, and Gaussian noise augmentation ($\sigma \in [0.005, 0.05]$) is essential for achieving robust FCNN performance on heterogeneous drilling datasets. Critically, all data-dependent transformations (Pearson screening, Lasso estimation, Winsorization thresholds, normalization parameters) were fitted exclusively on the training partition and applied as fixed transforms to the test partition, ensuring that reported holdout metrics represent unbiased generalization estimates.
5. The standalone FCNN exhibits directional asymmetry in cross-well transfer: strong $A \rightarrow B$ performance ($R^2 = 0.8702$) but near-zero $B \rightarrow A$ performance ($R^2 = 0.0388$), attributable to the information asymmetry between a heterogeneity-rich and a homogeneous training well. The hybrid LSTM-FCNN—using Savitzky–Golay denoising, depth-sequence windowing ($k = 10$), stacked LSTM layers (2×30 units), and a three-layer 128-unit fully connected head—substantially improves the $B \rightarrow A$ direction ($R^2 = 0.8217$, RMSE = 4.5476 m/h) while maintaining strong $A \rightarrow B$ performance ($R^2 = 0.8963$, RMSE = 1.6320 m/h). These results demonstrate that sequential depth modeling captures transferable drilling patterns that purely pointwise FCNN representations cannot encode. The LSTM-FCNN requires $13\times$ more training time (173.7 s per well) and 32% more parameters (48,761) than the standalone FCNN, a trade-off that is justified when cross-well transferability is a primary objective.

Author Contributions: Conceptualization, A.S.A. and M.A.K.; Methodology, M.A.K.; Software, A.S.A. and M.A.K.; Validation, A.S.A. and M.A.K.; Formal analysis, A.S.A. and M.A.K.; Investigation, A.S.A. and M.A.K.; Resources, M.A.K.; Data curation, A.S.A.; Writing—original draft, A.S.A. and M.A.K.; Writing—review and editing, M.A.K.; Visualization, M.A.K.; Supervision, M.A.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: All data supporting the findings of this work are presented within the article. For any further inquiries, please contact the corresponding author.

Acknowledgments: This work was supported by the Ongoing Research Funding program (ORF-2026-1616), King Saud University, Riyadh, Saudi Arabia. The authors would like to express their gratitude to the Department of Petroleum and Natural Gas Engineering at King Saud University for its assistance and support through this study.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

BS	Bit size, in
CCS	Confined Compressive strength, psi
Co	Rock cohesion, psi
DCM	Dynamic combined modulus, Mpsi
DR	Deep resistivity log, ohm. meters
DTCO	Compressional sonic log, us/ft
DTSM	Shear sonic log, us/ft
E _{DYN}	Young's Modulus, Mpsi
φ	Friction angle
FLOWPUMPS	Mud flow rate, l/min
G _{DYN}	Shear Modulus, Mpsi
GR	Gamma ray log
MW IN	Mud Weight in, g/cc
MW out	Mud Weight out, g/cc
n	Number of data points
PEF	Photoelectric factor log
ν	Poisson's ratio, dimensional
RHOZ	Density log, g/cc
ROP	Rate of Penetration, m/h
RPM	Rotary speed, revolutions per minute
SP	Spontaneous Potential log, millivolts
SS	Shear strength, psi
TNPH	Neutron log
TVD	True vertical depth, m
TWC	Thick Wall Cylinder, psi
UCS	Uniaxial Compression strength, psi
WOB	Weight on bit, tons
y_i	Actual value
$\hat{y}_i$	Predicted value
$\bar{y}_i$	Mean of actual values
σ_n	Normally effective stress, psi
ANN	Artificial Neural Network
α	Learning rate (Adam optimizer)
β	Trainable parameter in Swish activation; also Adam decay rate
β_1	Adam first moment decay rate (0.9)
β_2	Adam second moment decay rate (0.999)
Caliper	Caliper log, in
c	Gradient clipping threshold
d	Number of input features (sequence tensor dimension)
DTc	Compressional sonic travel time, μ s/ft
DTs	Shear sonic travel time, μ s/ft
ϵ	Small stability constant (batch normalization); noise term

FCNN	Fully Connected Neural Network
γ	Learnable scale parameter (batch normalization)
IP	Interactive Petrophysics (software)
k	Sequence window length (LSTM); lag order
λ	Regularization hyperparameter (Lasso/weight decay)
L_{MSE}	Mean squared error loss
LSTM	Long Short-Term Memory (neural network)
μ	Mean of training feature (Z-score normalization)
μ_B	Mini-batch mean (batch normalization)
MAE	Mean absolute error, m/h
p	Dropout rate
PACF	Partial autocorrelation function
PDC	Polycrystalline diamond compact (bit)
PHI/ Φ	Porosity (decimal fraction)
P_{99}	99th percentile (Winsorization threshold)
RHOB	Bulk density log, g/cc
RMSE	Root mean square error, m/h
R^2	Coefficient of determination
σ	Standard deviation of training feature (Z-score); noise level
σ_B^2	Mini-batch variance (batch normalization)
SD	Standard deviation
SPP	Standpipe pressure, psi
TEMP	Formation temperature, °C
Torque	Surface torque, lbf·ft
VCL	Volume of clay (fraction)
w	Savitzky–Golay filter window length

References

1. Sui, D.; Nybø, R.; Azizi, V. Real-time optimization of rate of penetration during drilling operation. In Proceedings of the 2013 10th IEEE International Conference on Control and Automation (ICCA), Hangzhou, China, 12–14 June 2013; pp. 357–362. [\[CrossRef\]](#)
2. Ahmed, O.S.; Adeniran, A.A.; Samsuri, A. Computational intelligence based prediction of drilling rate of penetration: A comparative study. *J. Pet. Sci. Eng.* **2019**, *172*, 1–12. [\[CrossRef\]](#)
3. Yi, P.; Kumar, A.; Samuel, R. Real-time rate of penetration optimization using the shuffled frog leaping algorithm. *J. Energy Resour. Technol.* **2014**, *137*, 032902. [\[CrossRef\]](#)
4. Ghasemloonia, A.; Rideout, G.; Butt, S.D. Vibration analysis of a drillstring in vibration-assisted rotary drilling: Finite element modeling with analytical validation. *J. Energy Resour. Technol.* **2013**, *135*, 013101. [\[CrossRef\]](#)
5. Kamel, M.A.; Elkatatny, S.; Mysorewala, M.F.; Al-Majed, A.; Elshafei, M. Adaptive and real-time optimal control of stick-slip and bit wear in autonomous rotary steerable drilling. *J. Energy Resour. Technol.* **2017**, *140*, 032908. [\[CrossRef\]](#)
6. Akgun, F. How to Estimate the Maximum Achievable Drilling Rate without Jeopardizing Safety. In Proceedings of the Abu Dhabi International Petroleum Exhibition and Conference, Abu Dhabi, United Arab Emirates, 13–16 October 2002; p. SPE-78567-MS. [\[CrossRef\]](#)
7. Maurer, W.C. The “perfect-cleaning” theory of rotary drilling. *J. Pet. Technol.* **1962**, *14*, 1270–1274. [\[CrossRef\]](#)
8. Bingham, M.G. *A New Approach to Interpreting Rock Drillability*; Drilling Research Laboratory: Houston, TX, USA, 1965.
9. Soares, C.; Daigle, H.; Gray, K. Evaluation of PDC bit ROP models and the effect of rock strength on model coefficients. *J. Nat. Gas Sci. Eng.* **2016**, *34*, 1225–1236. [\[CrossRef\]](#)
10. Bourgoynne, A.T., Jr.; Young, F.S., Jr. A multiple regression approach to optimal drilling and abnormal pressure detection. *SPE J.* **1974**, *14*, 371–384. [\[CrossRef\]](#)
11. Arabjamaloei, R.; Shadizadeh, S. Modeling and optimizing rate of penetration using intelligent systems in an Iranian southern oil field (Ahwaz oil field). *Pet. Sci. Technol.* **2011**, *29*, 1637–1648. [\[CrossRef\]](#)
12. Hegde, C.; Daigle, H.; Millwater, H.; Gray, K. Analysis of rate of penetration (ROP) prediction in drilling using physics-based and data-driven models. *J. Pet. Sci. Eng.* **2017**, *159*, 295–306. [\[CrossRef\]](#)

13. Mantha, B.; Samuel, R. ROP Optimization Using Artificial Intelligence Techniques with Statistical Regression Coupling. In Proceedings of the SPE Annual Technical Conference and Exhibition, Dubai, United Arab Emirates, 26–28 September 2016; p. SPE-181382-MS. [\[CrossRef\]](#)
14. Keshavarz Moraveji, M.; Naderi, M. Drilling rate of penetration prediction and optimization using response surface methodology and bat algorithm. *J. Nat. Gas Sci. Eng.* **2016**, *31*, 829–841. [\[CrossRef\]](#)
15. Cui, M.; Sun, M.; Zhang, J.; Kang, K.; Luo, Y. Maximizing drilling performance with real-time surveillance system based on parameters optimization algorithm. *Adv. Pet. Explor. Dev.* **2014**, *8*, 15–24.
16. Warren, T.M. Penetration-rate performance of roller-cone bits. *SPE Drill. Eng.* **1987**, *2*, 9–18. [\[CrossRef\]](#)
17. Bello, O.; Teodoru, C.; Yaqoob, T.; Oppelt, J.; Holzmann, J.; Obiwanne, A. Application of artificial intelligence techniques in drilling system design and operations: A state of the art review and future research pathways. In Proceedings of the SPE Nigeria Annual International Conference and Exhibition, Lagos, Nigeria, 2–4 August 2016; p. SPE-184320-MS. [\[CrossRef\]](#)
18. Bilgesu, H.I.; Tetrack, L.T.; Altamis, U.; Mohaghegh, S.; Ameri, S. A new approach for the prediction of rate of penetration (ROP) values. In Proceedings of the SPE Eastern Regional Meeting, Lexington, KY, USA, 22–24 October 1997; p. SPE-39231-MS. [\[CrossRef\]](#)
19. Al-Abduljabbar, A.; Elkhatny, S.; Mahmoud, M.; Abdelgawad, K.; Al-Majed, A. A robust rate of penetration model for carbonate formation. *J. Energy Resour. Technol.* **2018**, *141*, 092903. [\[CrossRef\]](#)
20. Khoukhi, A.; Alarfaj, I. Rate of penetration prediction and optimization using advances in artificial neural networks, a comparative study. In Proceedings of the IJCCI 2012—Proceedings of the 4th International Joint Conference on Computational Intelligence, Barcelona, Spain, 5–7 October 2012; pp. 647–652.
21. Bataee, M.; Mohseni, S. Application of Artificial Intelligent Systems in ROP Optimization: A Case Study in Shadegan Oil Field. In Proceedings of the SPE Middle East Unconventional Gas Conference and Exhibition, Muscat, Oman, 31 January–2 February 2011; p. SPE-140029-MS. [\[CrossRef\]](#)
22. Elkhatny, S. New approach to optimize the rate of penetration using artificial neural network. *Arab. J. Sci. Eng.* **2018**, *43*, 6297–6304. [\[CrossRef\]](#)
23. Hegde, C.; Daigle, H.; Gray, K.E. Performance comparison of algorithms for real-time rate-of-penetration optimization in drilling using data-driven models. *SPE J.* **2018**, *23*, 1706–1722. [\[CrossRef\]](#)
24. Jahanbakhshi, R.; Keshavarzi, R.; Jafarnezhad, A. Real-time prediction of rate of penetration during drilling operation in oil and gas wells. In Proceedings of the 46th U.S. Rock Mechanics/Geomechanical Symposium, Chicago, IL, USA, 24–27 June 2012; p. ARMA-2012-422.
25. Shaker, T.J.; Hadi, F.A. Development of artificial intelligence models for estimating rate of penetration in East Baghdad field, Middle Iraq. *Iraqi Geol. J.* **2022**, *55*, 112–124. [\[CrossRef\]](#)
26. Liu, H.; Jin, Y.; Song, X.; Pei, Z. Rate of penetration prediction method for ultra-deep wells based on LSTM-FNN. *Appl. Sci.* **2022**, *12*, 7731. [\[CrossRef\]](#)
27. Bai, K.; Jin, S.; Zhang, Z.; Dai, S. Drilling rate of penetration prediction based on CBT-LSTM neural network. *Sensors* **2024**, *24*, 6966. [\[CrossRef\]](#) [\[PubMed\]](#)
28. He, M.; Yuan, Q.-R.; Li, J.; Zhou, C.-C.; Liu, J.-W.; Chen, X.; Cao, J.-H. A novel method for drilling rate of penetration prediction and transfer generalization based on physics-informed neural networks. *Pet. Sci.* **2026**; *in press*. [\[CrossRef\]](#)
29. Liu, Y.; Zeng, G.; Shi, X.; Jiang, J.; Guo, Y. Research on deep well drilling rate of penetration prediction method based on TCN-transformer model. *Sci. Rep.* **2026**, *16*, 11075. [\[CrossRef\]](#) [\[PubMed\]](#)
30. Meng, Q.; Song, H.; Meng, D.; Liu, X.; Li, D.; Chen, X.; Wei, Y.; Zhang, C.; Wei, J.; Wu, Y.; et al. Drilling Rate Prediction Based on Bayesian Optimization LSTM Algorithm with Fusion Feature Selection. *Processes* **2026**, *14*, 274. [\[CrossRef\]](#)
31. Jaeger, J.C.; Cook, N.G.; Zimmerman, R. *Fundamentals of Rock Mechanics*, 4th ed.; Wiley-Blackwell: Hoboken, NJ, USA, 2009.
32. Alhalboosi, A.S.; AlAwad, M.N.J.; Altawati, F.S.; Khamis, M.A.; Almobarkey, M.A. Optimal Geomechanical Parameter Selection for Enhanced ROP Modeling: A Systematic Field-Based Comparative Study. *Processes* **2026**, *14*, 1646. [\[CrossRef\]](#)
33. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
34. Ioffe, S.; Szegedy, C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In Proceedings of the 32nd International Conference on Machine Learning, Lille, France, 6–11 July 2015; pp. 448–456.
35. Ramachandran, P.; Zoph, B.; Le, Q.V. Searching for activation functions. *arXiv* **2017**, arXiv:1710.05941. [\[CrossRef\]](#)
36. Srivastava, N.; Hinton, G.; Krizhevsky, A.; Sutskever, I.; Salakhutdinov, R. Dropout: A simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.* **2014**, *15*, 1929–1958.
37. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. *arXiv* **2014**, arXiv:1412.6980.
38. Loshchilov, I.; Hutter, F. Decoupled weight decay regularization. *arXiv* **2017**, arXiv:1711.05101. [\[CrossRef\]](#)
39. Pascanu, R.; Mikolov, T.; Bengio, Y. On the difficulty of training recurrent neural networks. In Proceedings of the 30th International Conference on Machine Learning, Atlanta, GA, USA, 16–21 June 2013; pp. 1310–1318.

40. Fisher, A.; Rudin, C.; Dominici, F. All models are wrong, but many are useful: Learning a variable's importance by studying an entire class of prediction models simultaneously. *J. Mach. Learn. Res.* **2019**, *20*, 1–81.
41. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.