

Article

Residual Decomposition for Lithotype-Aware Characterization of Rock Mechanical Parameters from Well Logs Under Lithological Heterogeneity

Xugang Liu ^{1,2}, Binghua Dang ^{1,2}, Lei Li ^{1,2}, Weixian Zhang ^{3,4,*}  and Wenzhe Zhou ^{3,4}

¹ Petroleum Engineering and Technology Research Institute, Sinopec North China Oil and Gas Company, Zhengzhou 450006, China; liuxugang.hbsj@sinopec.com (X.L.); dangbh.hbsj@sinopec.com (B.D.); lilei.hbsj@sinopec.com (L.L.)

² Sinopec Key Laboratory of Deep Coalbed Methane Exploration and Development, Zhengzhou 450006, China

³ College of Petroleum Engineering, China University of Petroleum (Beijing), Beijing 102249, China; zhouwenze@student.cup.edu.cn

⁴ College of Energy Innovation, China University of Petroleum (Beijing), Beijing 102249, China

* Correspondence: weixianzhang@student.cup.edu.cn; Tel.: +86-150-1023-3076

Featured Application

This study provides a lithotype-aware approach for reliable prediction of rock mechanical parameters from well logging data, supporting wellbore stability evaluation and drilling optimization in heterogeneous coal-bearing formations. The method is particularly effective in lithological transition zones where conventional approaches exhibit significant bias.

Abstract

Accurate characterization of rock mechanical parameters in heterogeneous geological formations remains challenging because lithological variations alter the relationship between logging signals and geomechanical responses. Existing approaches, including empirical formulas, pure machine learning models, and feature-augmented learning methods, often compress these variations into a single predictor, which can lead to biased estimates. To address this issue, this study proposes a heterogeneity-aware residual learning framework for rock mechanical parameter characterization from well logs. The method separates the prediction into a global component and a lithotype-conditioned correction, allowing lithological effects to be represented as structured residual behavior. This framework was developed and validated on deep coal-bearing formations in the Ordos Basin. By accounting for lithology-controlled response shifts, it produces predictions that better follow observed geological controls. Cross-well validation demonstrates reduced lithotype-induced bias and stable generalization within the studied formation. Further analysis shows that the performance gain is linked to the residual decomposition structure rather than to the addition of lithotype information alone. Compared with single-stage feature augmentation, the main advantage of the proposed framework is its ability to reduce systematic bias in lithological transition zones while preserving a transparent global–residual structure. Its demonstrated applicability is limited to wells within the studied coal-bearing formation, and broader transferability requires further validation.

Keywords: rock mechanical parameters; well logging; lithotype; residual learning; coal-bearing formations; Bayesian optimization



Academic Editor: Nikolaos Koukoulas

Received: 26 March 2026

Revised: 2 May 2026

Accepted: 7 May 2026

Published: 8 May 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

As energy exploration advances into deeper coal-bearing formations, ensuring wellbore stability and operational safety has become increasingly challenging [1–3]. In such environments, strong lithological heterogeneity, complex cleat systems, and frequent coal–rock interbedding lead to substantial variability in geomechanical properties. Reliable characterization of parameters such as uniaxial compressive strength (UCS), static Young’s modulus (E_s), and Poisson’s ratio (ν) is therefore essential for drilling design, wellbore stability evaluation, and hydraulic fracturing optimization. However, in practice, these parameters remain highly uncertain, significantly increasing operational risks and compromising engineering reliability [4–6].

The core difficulty of this problem is not merely nonlinearity. In heterogeneous formations, lithological composition and structure can change the mechanical meaning of otherwise similar well log signatures, making the inverse relationship ill-posed.

Existing approaches for rock mechanical parameter characterization can be broadly categorized into three groups. First, empirical methods establish predefined relationships between logging responses and mechanical properties [7–9], while computationally efficient, they rely on strong assumptions of formation homogeneity and often fail in heterogeneous intervals. Second, data-driven methods, including machine learning models such as random forests, gradient boosting, and neural networks, aim to capture complex nonlinear relationships [10–19]. Third, feature-augmented learning approaches incorporate additional geological information, such as lithology indicators, as input features to improve prediction accuracy.

Despite their differences, these approaches commonly treat the input–output relationship as globally consistent. This assumption is weak in coal-bearing intervals, where lithological regimes alter the local geomechanical response and can cause biased predictions, particularly near lithological transitions.

From a geological and engineering perspective, lithotype variations are a primary factor controlling mechanical behavior in coal-bearing formations [20–25]. Differences in maceral composition, fracture density, and structural integrity result in distinct strength and deformation characteristics across lithotypes. Ignoring these variations reduces predictive accuracy and weakens the geological interpretability of the model.

To address this limitation, we propose a heterogeneity-aware modeling framework based on lithotype-conditioned residual learning. Instead of treating lithotype as an auxiliary feature, the method separates the overall response into a global component and a lithological residual. This design allows the model to represent lithological effects as corrections to the baseline response rather than as another input in a single predictor.

It is developed for heterogeneous coal-bearing formations and evaluated under cross-well conditions.

Model effectiveness is evaluated through cross-well validation and comparative analysis. The results show lower prediction error across different wells and geological conditions. Residual analysis further indicates that the improvement is linked to the modeling structure rather than to additional inputs alone.

The main contributions of this study are summarized as follows: (1) A heterogeneity-aware residual learning paradigm is proposed for geomechanical characterization under lithological heterogeneity; (2) A structured decomposition formulation is introduced to separate global trends from lithological corrections; (3) The approach demonstrates improved cross-well generalization and lower prediction bias; (4) The geological consistency of the model is validated by linking residual behavior to lithological regimes.

The proposed approach also has drawbacks relative to traditional empirical methods. It introduces additional modeling complexity, depends on the reliability of HMLZ-based

lithotype identification, and requires laboratory control points for calibration, especially for lithotype-level strength parameters. These issues are addressed in this study through a unified hyperparameter-optimization protocol, explicit ablation against a single lithotype-feature model, independent-well testing, and a discussion of logging-resolution limits. Future work should further reduce these constraints by incorporating higher-resolution logs, additional wells, and rock-physics-based interface correction.

Overall, this study reframes rock mechanical parameter characterization as a structured learning problem under heterogeneity and provides a physically consistent modeling framework for the studied coal-bearing formation.

2. Materials and Methods

This section presents the study area, dataset construction, lithotype identification, and modeling workflow used in this study.

The overall methodology consists of three components. First, lithological heterogeneity is quantified through lithotype identification based on the HMLZ index, providing a structured representation of geological regimes along the wellbore. Second, a global predictor is trained from logging-depth features to geomechanical parameters. Third, a residual model conditioned on lithotype is constructed to adjust the baseline prediction in different lithological regimes. The final prediction is obtained by combining the two components.

2.1. Study Area Overview

The study area is located in the Ordos Basin, a large cratonic sedimentary basin in the western part of the North China Platform, with a sedimentary thickness of approximately 5000 m. The Upper Paleozoic strata are dominated by fluvial–deltaic clastic deposits and host extensive coal-bearing formations [26]. The investigated region lies within the Daniudi Gas Field, a key area for deep coalbed methane (CBM) exploration and development.

Coal seams in this region are typically buried at depths exceeding 2000 m and exhibit pronounced lithological heterogeneity characterized by frequent coal–rock interbedding and well-developed cleat systems. As drilling activities extend into deeper formations, engineering challenges such as wellbore instability and borehole enlargement have become increasingly severe. These challenges are closely associated with lithotype-controlled variations in geomechanical properties, which introduce significant uncertainty into stability evaluation and drilling design.

Accordingly, data from the Ordos Basin are used to validate the proposed heterogeneity-aware modeling framework under complex geological conditions.

2.2. Dataset Construction

A cross-well dataset was constructed to retain lithological variability across training, validation, and independent test wells.

Each depth sampling point is treated as a fundamental unit of analysis. For a given depth d , the corresponding logging responses and depth information are represented as the input vector $X(d)$, while the measured or experimentally derived geomechanical labels form the target vector $Y(d)$. In addition, lithotype information $L(d)$ derived from the HMLZ index is incorporated as a categorical condition. The resulting data structure is denoted as $(X(d), L(d), Y(d))$.

The complete input feature vector consists of AC, DEN, GR, CNL, LLD, LLS, SP, and depth z , as formally defined in Section 2.3.2. These measurements collectively characterize formation properties from complementary physical perspectives, including elastic response, density, shale or ash-related composition, porosity, resistivity, electrochemical response,

and burial-related trends. Lithotype is treated as a conditioning variable that governs regime-dependent shifts in the mapping from X to Y .

The definitions of input features and output targets are summarized in Table 1.

Table 1. Summary of input features and output targets.

Category	Feature Name	Symbol	Data Type
Input Features	Acoustic transit time	AC	Continuous
Input Features	Bulk density	DEN	Continuous
Input Features	Gamma ray	GR	Continuous
Input Features	Compensated neutron logging	CNL	Continuous
Input Features	Deep lateral resistivity	LLD	Continuous
Input Features	Shallow lateral resistivity	LLS	Continuous
Input Features	Spontaneous potential	SP	Continuous
Input Features	Depth	z	Continuous
Input Features	HMLZ-derived lithotype	Lithotype	Categorical
Output Targets	Static Young's modulus	E_s	Continuous
Output Targets	Static Poisson's ratio	ν	Continuous
Output Targets	UCS	UCS	Continuous
Output Targets	Cohesion	C	Lithotype-level estimate
Output Targets	Internal friction angle	ϕ	Lithotype-level estimate

The target vector is defined as the five-dimensional geomechanical parameter set in Equation (1):

$$Y(d) = [E_s(d), \nu(d), UCS(d), C(d), \phi(d)] \quad (1)$$

This multi-output formulation preserves intrinsic correlations among mechanical parameters and allows consistent characterization across depth. In this target set, UCS , E_s , and ν are point-scale laboratory measurements matched to core depths, whereas C and ϕ are lithotype-level estimates assigned from lithotype-specific triaxial Mohr–Coulomb regressions. Thus, C and ϕ should be interpreted as experimentally constrained mechanical baselines for different coal regimes rather than as independently measured continuous point-scale labels.

Data preprocessing ensures consistency and reliability of the dataset. All logging curves are aligned in the depth domain through datum unification and sampling interval standardization. Quality control procedures include the identification of missing intervals, suppression of spike noise, and removal of physically implausible values. Minor gaps are filled via interpolation to maintain continuity without altering global trends. Core measurements were matched to the nearest log depth sample on the unified 0.125 m depth grid, and no interpolation was applied between discrete core points. The continuous prediction curves presented in this study are generated by applying the trained model to every logging sample along the full well interval (point-wise inference on the logging grid), rather than by interpolating between discrete core measurements; core data are used only as supervised labels and evaluation references. After preprocessing, all features and targets are mapped onto a unified depth grid, forming a cross-well consistent data matrix.

A well-based partitioning strategy is adopted to ensure rigorous evaluation of generalization capability. By treating each well as an independent unit, the dataset avoids leakage caused by highly correlated adjacent depth samples. The training set consists of five wells, while separate wells are used for validation and testing (Table 2). Across the laboratory program, a total of 97 specimens supported model calibration, evaluation, and mechanism verification, covering all four coal lithotype classes (bright coal: 32; semi-bright coal: 36; semi-dull coal: 15; dull coal: 14). Among them, 40 specimens were used for triaxial compression testing, including 10 groups from Yangmei-1HF, 12 groups from D1-612, and 18 outcrop auxiliary groups used for mechanism verification. The remaining 57 specimens were used for uniaxial compression, CT scanning, and mineral-composition tests. This par-

tioning reflects the practical deployment scenario, where models are required to generalize to unseen wells with different lithological distributions.

Table 2. Dataset partitioning and corresponding well distribution used in this study.

Well ID	Dataset Role
Well-Train-1	Training
Well-Train-2	Training
Well-Train-3	Training
Well-Train-4	Training
Well-Train-5	Training
Well-Validate-1	Validation
Well-Test-1	Test (Independent/Blind)

A representative well profile is shown in Figure 1, illustrating the correspondence between logging responses and lithological variations.

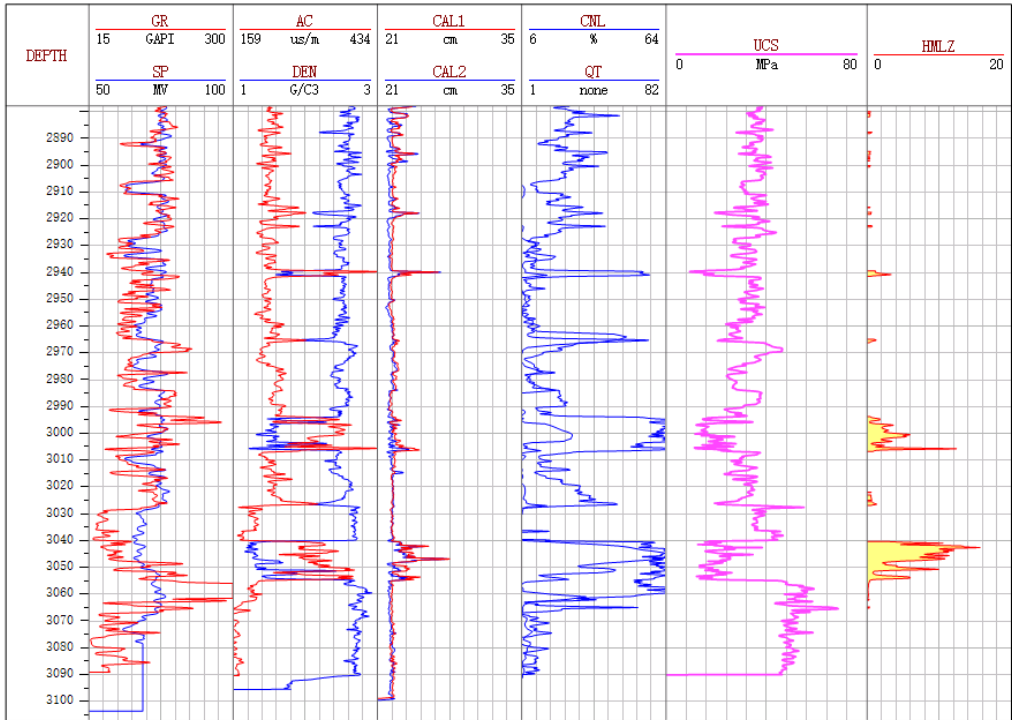


Figure 1. Comprehensive logging and training data profile of a representative well (Well-Train-5).

2.3. Lithotype Classification and Feature Engineering Based on the HMLZ Index

2.3.1. Experimental Determination of Rock Strength Parameters

To establish reference targets for model evaluation, laboratory geomechanical experiments were conducted on coal and parting (dirt band) samples collected from deep coal seams in the Ordos Basin.

Key mechanical parameters, including uniaxial compressive strength (*UCS*), static Young’s modulus (*E_s*), Poisson’s ratio (*ν*), cohesion (*C*), and internal friction angle (*φ*), were measured or derived from laboratory testing to provide benchmark labels for supervised learning and evaluation.

The experiments were performed using the MTS-816 electro-hydraulic servo-controlled rock testing system (Figure 2a), which provides stable closed-loop control for load application and displacement measurement. Uniaxial compression tests were conducted under

zero confining pressure until specimen failure. The UCS was calculated from the peak load using Equation (2):

$$UCS = \sigma_c = \frac{F_{max}}{A} \quad (2)$$

where A is the initial cross-sectional area. The static Young's modulus E_s was determined from the slope of the stress–strain curve within the linear elastic regime, while Poisson's ratio ν was computed as the ratio of lateral strain to axial strain in the same region.

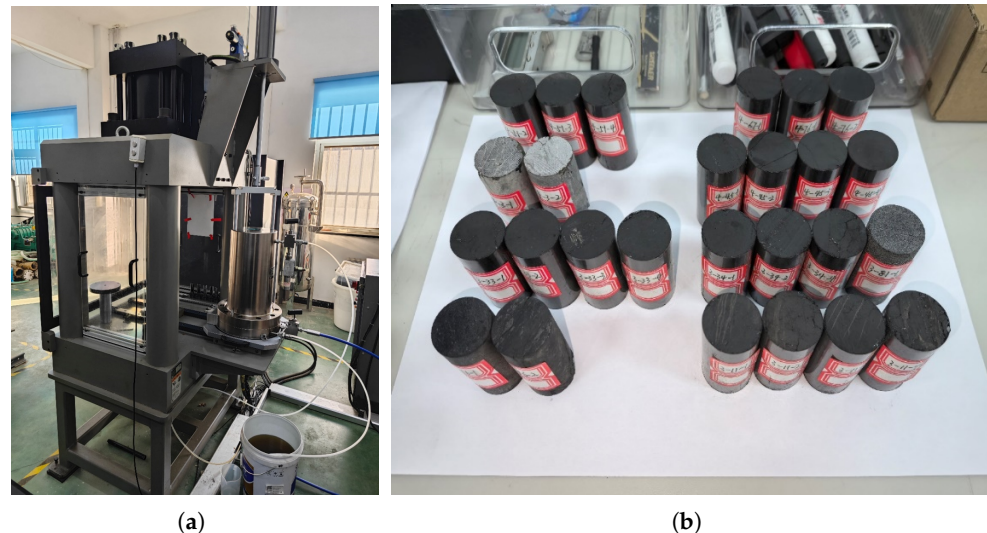


Figure 2. Experimental determination of rock mechanical parameters for coal specimens: (a) MTS-816 testing system; (b) standard specimens of various lithotypes and partings.

In addition to uniaxial tests, triaxial compression tests were conducted under multiple confining-pressure levels following the ISRM suggested method [27], providing the multi-stress-state data required for strength envelope construction. The Yangmei-1HF subset was tested at 5, 10, and 20 MPa, whereas the D1-612 subset was tested at 0, 10, and 20 MPa.

Cohesion C and internal friction angle ϕ were determined by fitting the peak axial stress σ_1 against confining pressure σ_3 using the Mohr–Coulomb criterion in Equation (3):

$$\sigma_1 = \frac{2C \cos \phi}{1 - \sin \phi} + \sigma_3 \frac{1 + \sin \phi}{1 - \sin \phi} \quad (3)$$

Linear regression over the three confining pressure levels yielded lithotype-specific values of C and ϕ for each specimen group.

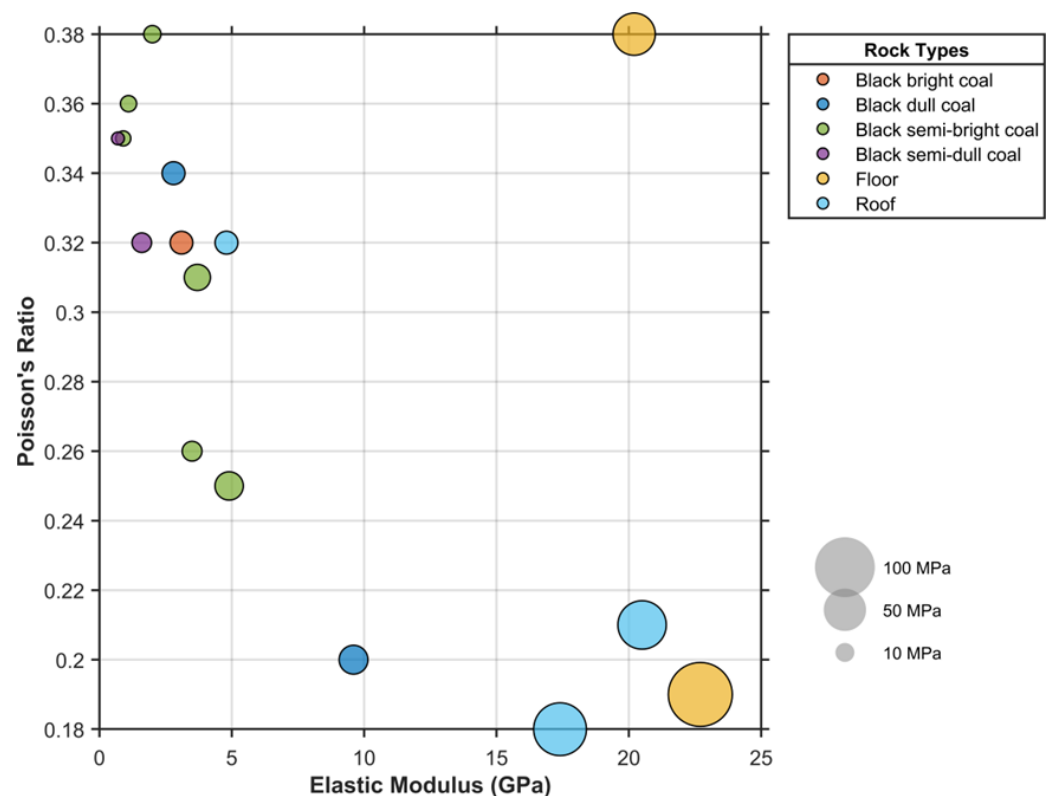
Because C and ϕ cannot be independently determined from UCS alone, these two parameters were obtained only from the triaxial strength envelopes. For each coal lithotype, the peak-strength data measured under different confining pressures were grouped and fitted separately, producing one representative pair of C_{litho} and ϕ_{litho} for bright coal, semi-bright coal, semi-dull coal, and dull coal. The continuous HMLZ response was first converted into the four discrete lithotype classes using the thresholds in Table 3. The corresponding lithotype-specific C_{litho} and ϕ_{litho} values were then assigned to each 0.125 m depth sample according to its HMLZ-derived lithotype, forming lithotype-specific supervised labels for cohesion and internal friction angle. In this way, the labels for C and ϕ retain a direct experimental basis while reflecting the structural mechanical baseline of different coal regimes.

During testing, failure modes were recorded to document structural differences among lithotypes. These observations provide qualitative evidence of mechanical variability across lithological regimes.

Table 3. Classification standards for coal lithotypes based on the HMLZ index.

Coal Lithotype	HMLZ Range	Lithological Characteristics	Expected Strength
Bright coal	$HMLZ > 15.2$	Extremely high vitrinite content (>75%); highly developed cleats; brittle texture; strong vitreous luster.	Minimum
Semi-bright coal	$5 < HMLZ \leq 15.2$	Dominant vitrinite with minor inertinite; developed fractures; banded structure.	Relatively low
Semi-dull coal	$1.9 < HMLZ \leq 5$	Increased inertinite and liptinite; tougher structure; fewer fractures.	Relatively high
Dull coal	$HMLZ \leq 1.9$	High inertinite and mineral content; dense structure; maximum toughness.	Maximum

Specimen preparation followed standard rock mechanics protocols to ensure comparability. Cylindrical samples were fabricated with strict control of geometry, and end faces were precision-ground to minimize loading artifacts. Defective specimens were excluded to maintain data integrity. Special care was taken during handling and mounting to avoid artificial damage, particularly given the fragile and heterogeneous nature of coal. The measured UCS , E_s , and ν for Well-Train-3 are shown in Figure 3.

**Figure 3.** Measured UCS , E_s , and ν for samples in Well-Train-3.

Experimental results reveal clear contrasts among lithotypes. Coal samples exhibit relatively low strength and stiffness, with UCS values ranging from 7.7 to 20.5 MPa and E_s between 4.1 and 7.8 GPa, while partings show significantly higher strength (up to 71.8 MPa). Poisson's ratio ν is generally higher in coal (0.23–0.34), reflecting its deformable structure. Substantial variability is also observed among the four coal lithotypes.

The corresponding measured mechanical properties for Well-Train-5 are shown in Figure 4.

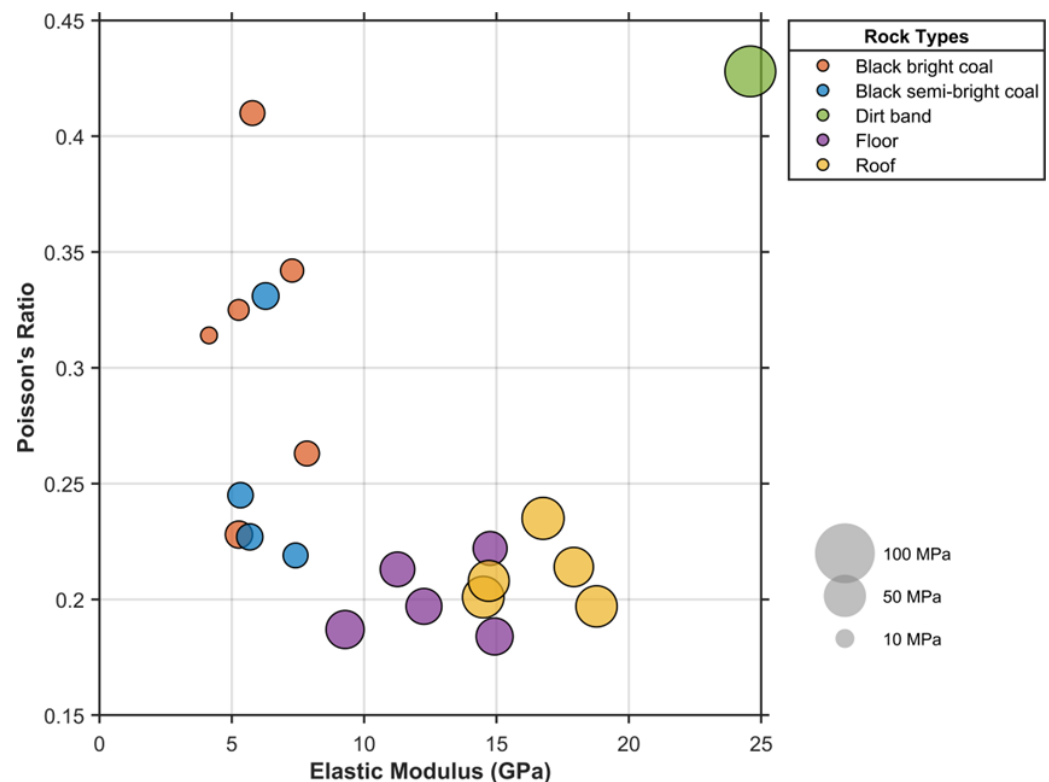


Figure 4. Measured mechanical properties for samples in Well-Train-5.

These observations support the use of lithotype information in the prediction framework.

The measured parameters are subsequently used as reference targets for model evaluation. Predictions at corresponding depths are extracted and compared against experimental values under cross-well conditions.

2.3.2. Feature Construction

Based on the experimentally measured rock mechanical parameters, the input representation is constructed for the residual learning framework.

The model input at depth d is defined as $(X(d), L(d))$, where the complete logging-depth feature vector is given by Equation (4):

$$X(d) = [AC(d), DEN(d), GR(d), CNL(d), LLD(d), LLS(d), SP(d), z(d)]. \quad (4)$$

Here, $X(d)$ contains the complete logging-depth feature vector, and $L(d)$ represents the HMLZ-derived lithotype condition.

The primary feature set therefore consists of acoustic transit time (AC), bulk density (DEN), gamma ray (GR), compensated neutron logging (CNL), deep and shallow lateral resistivity (LLD, LLS), spontaneous potential (SP), and depth (z). These features collectively describe formation properties from multiple physical perspectives, including elastic response, density distribution, mineral or ash-related composition, porosity, fluid-related conductivity, electrochemical response, and burial/stress-related trends.

Depth (z) is incorporated as a continuous feature to preserve vertical continuity and capture systematic variations associated with burial conditions and stress environments. This inclusion ensures that large-scale geological trends are retained in the global component of the model.

To account for lithological heterogeneity, lithotype information derived from the HMLZ index is introduced as the condition $L(d)$ in the residual component.

The Pearson correlation coefficients between the selected logging features and the target mechanical parameters are summarized in Figure 5.

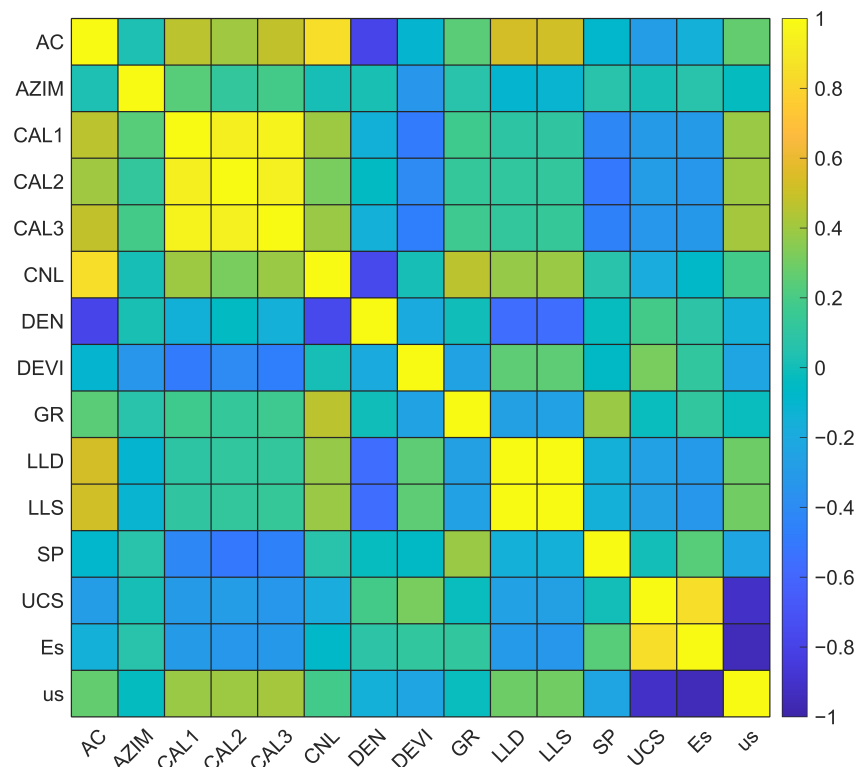


Figure 5. Pearson correlation coefficient matrix between logging features and rock mechanical parameters.

To assess the relevance of the selected features, Pearson correlation coefficients between logging variables and mechanical parameters are computed [28]. The results show that the selected logging features exhibit meaningful correlations with the target variables, indicating that they provide informative signals for geomechanical characterization. However, these correlations are primarily linear and therefore cannot fully describe the interactions present in the data.

Therefore, while the feature set provides a physically meaningful basis for modeling, the key gain comes from the residual decomposition formulation.

2.3.3. Coal Lithotype Identification (HMLZ)

Coal seams exhibit strong lithological heterogeneity governed by depositional environment, maceral composition, and structural development. Variations in fracture density, pore structure, and compositional characteristics lead to substantial differences in geomechanical behavior [24,25,29].

The HMLZ index is employed to identify lithological regimes along the wellbore. Rather than serving only as a descriptive classification, the HMLZ-derived lithotype sequence is used as the categorical condition in the residual model.

The HMLZ index is defined based on conventional well logging measurements, including lateral resistivity response (RD, represented by the resistivity logging suite), acoustic transit time (AC), bulk density (DEN), and gamma ray (GR), capturing key petrophysical characteristics associated with coal brittleness. It is calculated using Equation (5):

$$HMLZ = \frac{\lg(RD) \times AC}{DEN^2 \times GR} \quad (5)$$

Based on the computed HMLZ values, coal lithotypes are categorized into four classes—bright coal, semi-bright coal, semi-dull coal, and dull coal—according to pre-defined threshold intervals. These lithotypes reflect systematic differences in maceral composition, structural integrity, and fracture development, which are directly linked to mechanical properties. The resulting HMLZ-derived lithotype is treated as a categorical variable rather than as a continuous logging feature. The classification criteria and corresponding characteristics are summarized in Table 3.

The HMLZ index is computed continuously along depth, producing a lithotype sequence $L(z)$ that characterizes the spatial distribution of heterogeneity. Within this framework, $L(z)$ is used as a condition for the residual component of the model to learn regime-dependent corrections to the global geomechanical response.

The HMLZ-derived lithotype sequence provides a physically meaningful partition of the data space for the subsequent residual formulation.

2.4. Lithotype-Conditioned Residual Characterization Framework

Building on the above definitions, the geomechanical response is formulated as the sum of a global baseline and a residual term conditioned on lithotype.

Accordingly, the final prediction is expressed by the residual decomposition formulation in Equation (6):

$$\hat{Y}(d) = f(X(d)) + g(X(d), L(d)) \quad (6)$$

where $X(d)$ denotes the complete logging-depth feature vector defined in Section 2.3.2, $L(d)$ denotes the lithotype label derived from the HMLZ index, $f(\cdot)$ represents the baseline predictor from logging-depth responses to mechanical parameters, and $g(\cdot)$ represents the residual correction conditioned on lithotype. Equation (6) is not a simple model-stacking step; it separates the shared logging-to-mechanics trend from the lithotype-dependent deviation so that the correction term has a specific geological meaning.

In implementation, both $f(\cdot)$ and $g(\cdot)$ are constructed using CatBoost regressors [30–32]. The global model $f(X)$ is trained using only the logging-depth feature vector, while the residual model $g(X, L)$ takes the same feature vector together with lithotype as inputs. Lithotype is encoded as a categorical variable within CatBoost.

For multi-output prediction, separate models are trained for each target variable to ensure stable optimization. Hyperparameters are optimized using Bayesian Optimization based on validation error, and the same optimization protocol is applied to both stages to ensure fair comparison.

Residuals used to train $g(\cdot)$ are computed on the training set using predictions from $f(\cdot)$ without data leakage.

The function $f(X)$ is used to learn the dominant relationship between logging-depth responses and geomechanical parameters over the entire training dataset. It captures the general response trend shared across samples and reflects the lithology-independent component of the prediction. However, because the data are affected by lithological heterogeneity, this global model alone may leave systematic errors in intervals where different lithotypes exhibit different mechanical responses under similar logging signatures.

To characterize this effect explicitly, the residual is defined in Equation (7):

$$r(d) = Y(d) - f(X(d)) \quad (7)$$

where $Y(d)$ is the measured or lithotype-level target vector and $r(d)$ is the deviation between the observation and the global prediction. Equation (7) defines the quantity learned by the second-stage model: only the remaining lithotype-conditioned bias after the global prediction is modeled, rather than the full mechanical response. Here, this residual is

interpreted as a structured correction term associated with lithological heterogeneity rather than as purely random noise. The function $g(X, L)$ is then used to learn the relationship between the residual, the logging-depth features, and the lithotype condition.

The training procedure is implemented sequentially. First, the global model $f(X)$ is trained using the logging-depth feature vector to predict the target mechanical parameters. Second, the residuals are computed on the training set as the difference between the measured or lithotype-level target values and the predictions of the global model. Third, a residual model $g(X, L)$ is trained using the logging-depth feature vector together with the lithotype variable, with the residual term as the prediction target. During inference, the output of the global model and that of the residual model are added to obtain the final prediction.

In practical implementation, lithotype is introduced as a categorical condition in the residual stage. This design allows the residual model to learn regime-dependent corrections to the global trend.

This decomposition also improves the interpretability of the modeling framework. The global component describes the dominant mapping shared by the dataset, whereas the residual component accounts for lithotype-related deviations from that common trend. In this sense, the final prediction can be understood as a combination of baseline geomechanical response and lithology-dependent correction.

It should be noted that the present framework is developed and validated using data from coal-bearing formations in the Ordos Basin. The applicability of this approach to other basins, lithologies, and logging configurations requires further verification.

Hyperparameter Optimization and SHAP Analysis

To ensure stable model construction, hyperparameter optimization was performed using Bayesian Optimization (BO) [33]. In this study, BO was used to search for suitable parameter combinations for the predictive model by minimizing the Mean Squared Error on the validation set. The optimized parameters mainly include the number of iterations, learning rate, tree depth, and L2 regularization coefficient. BO treats the validation error as a black-box objective, updates a surrogate response surface after each trial, and uses an acquisition function to balance exploration of uncertain regions with exploitation of promising parameter combinations. This procedure improves the stability of model training and reduces the risk of overfitting caused by manual parameter selection.

The purpose of hyperparameter optimization in this study is to obtain a stable and reproducible model configuration for subsequent comparison and analysis. All model variants therefore follow the same optimization protocol.

In addition, SHAP analysis was employed to examine the contribution patterns of the input features in the trained model [34]. SHAP is, by definition, a post hoc interpretability method; it quantifies feature effects at both the global and sample levels and is reported as supplementary interpretation. For a given sample, the additive SHAP formulation can be written as Equation (8):

$$g(z') = \phi_0 + \sum_{j=1}^M \phi_j z'_j \quad (8)$$

where $z' = [z'_1, \dots, z'_M]$ is the simplified binary feature vector used in the additive SHAP explanation, z'_j indicates whether the j th feature is included in a feature coalition, M is the number of explained features, ϕ_0 is the baseline output, and ϕ_j represents the contribution of the j th feature to the local explanation. The subscripted attribution term ϕ_j should not be confused with the internal friction angle ϕ used as a mechanical target.

The SHAP value of each feature is computed as the weighted average marginal contribution over all possible feature subsets, as shown in Equation (9):

$$\psi_j = \sum_{S \subseteq \{x_1, \dots, x_p\} \setminus \{x_j\}} \frac{|S|!(p - |S| - 1)!}{p!} (f_x(S \cup \{x_j\}) - f_x(S)) \quad (9)$$

where ψ_j denotes the Shapley value for the original feature x_j , p is the total number of input features, S is a subset of features excluding x_j , and $f_x(S)$ is the model output evaluated using the feature subset S . The factorial term provides the standard Shapley weighting over all possible subset sizes.

In this work, TreeSHAP was used for efficient explanation of the tree-based model. Global SHAP importance identifies the dominant logging-depth variables, whereas local SHAP analysis inspects feature contributions across depth intervals and lithological settings. The main evidence for model effectiveness is derived from cross-well evaluation, case-study comparison, and ablation analysis; BO and SHAP serve supporting roles in training stability and interpretation, respectively.

3. Results and Discussion

3.1. Reliability-Oriented Hyperparameter Optimization and Model Performance

To obtain a stable training configuration under heterogeneous data conditions, hyperparameter optimization was conducted using Bayesian Optimization (BO) in combination with cross-validation (CV) [33,35]. The objective of this procedure is to identify a reliable set of hyperparameters that minimizes validation error while maintaining stable model behavior across different data splits.

BO iteratively refines the hyperparameter configuration by modeling the validation error as a black-box function and selecting candidate parameter sets through an acquisition strategy. Compared with manual tuning or grid-based search, this approach enables more efficient exploration of the parameter space and provides a more reproducible model-selection process. The search was stopped after the predefined optimization budget was completed or when no further meaningful decrease in validation error was observed.

As shown in Figure 6, the validation error exhibits clear trends with respect to the tested hyperparameters. Increasing the number of iterations reduces the validation error up to a certain point, after which the improvement becomes marginal, indicating convergence in model training. The learning rate also shows a limited effective range: excessively small values lead to underfitting, whereas overly large values result in unstable training and poorer validation performance.

The optimal tree depth is relatively low, suggesting that a simple model structure is sufficient to capture the dominant relationships in the data. Increasing the depth leads to higher validation error, indicating overfitting and increased sensitivity to noise. Similarly, larger values of the regularization coefficient (`l2_leaf_reg`) tend to increase the validation error, implying that excessive regularization may weaken the model's ability to fit the data adequately.

These observations indicate that the prediction task favors a configuration with moderate model complexity and stable training behavior. Overall, the selected hyperparameter region is consistent with the need to balance fitting capacity and regularization in a heterogeneous prediction setting.

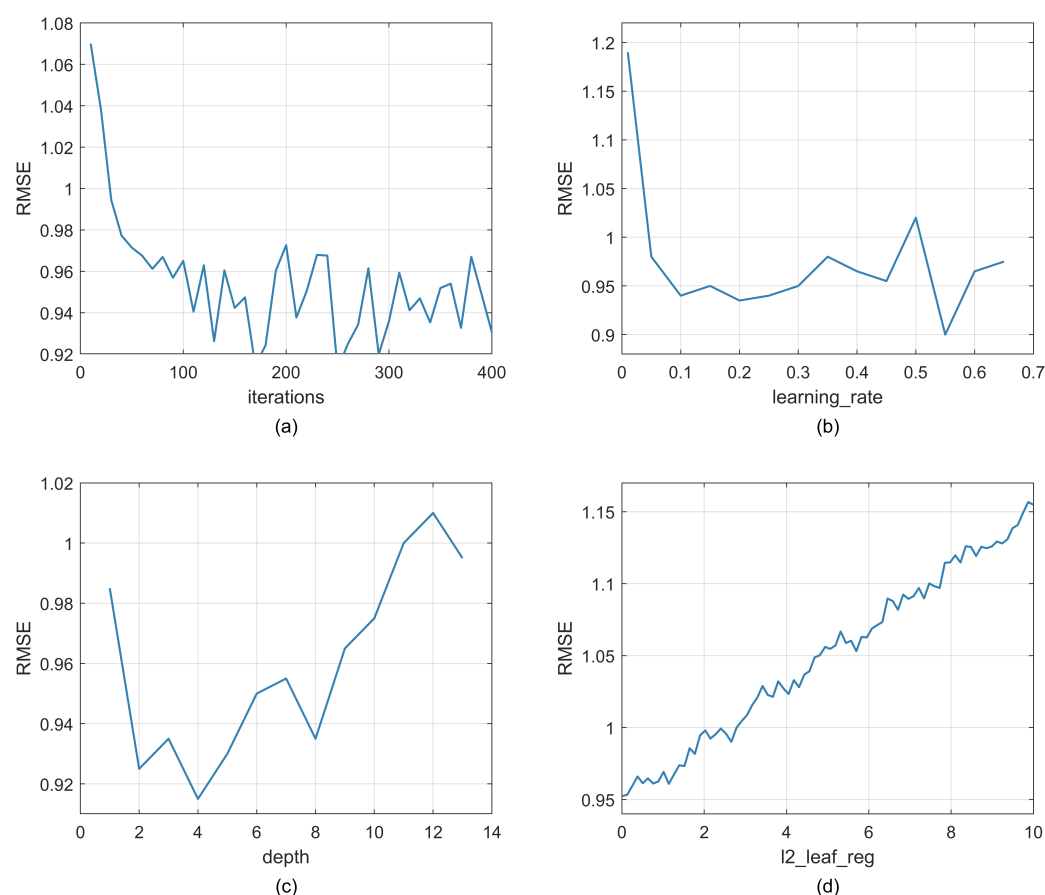


Figure 6. Search for optimal hyperparameter ranges via 5-fold cross-validation: (a) number of iterations, (b) learning rate, (c) tree depth, and (d) l2_leaf_reg.

The search ranges in Table 4 were selected to cover practical CatBoost configurations for a small, heterogeneous, cross-well dataset. The range of iterations [40, 200] allows sufficient boosting steps while limiting overfitting risk; the learning-rate interval [0.01, 0.5] covers both conservative and rapid update regimes; the tree-depth interval [2, 10] spans shallow trees for simple responses and deeper trees for nonlinear interactions; and the l2_leaf_reg interval [0.01, 1] provides a controlled range of regularization strengths. Figure 6d is shown over the broader sensitivity interval [0, 10] to indicate that excessive L2 regularization increases validation error, whereas the final BO search in Table 4 was restricted to [0.01, 1] after preliminary screening identified the effective low-regularization region. The lower bound avoids an unregularized leaf model, and the upper bound retains sufficient regularization without over-penalizing the small heterogeneous dataset.

Table 4. Optimal CatBoost hyperparameters obtained after tuning.

Parameter	Description	Search Range	Optimal Value
iterations	Number of boosting iterations	[40, 200]	150
learning rate	Step size controlling update magnitude	[0.01, 0.5]	0.38
depth	Maximum depth of decision trees	[2, 10]	3
l2_leaf_reg	L2 regularization coefficient	[0.01, 1]	0.04

Based on the optimization results, the final hyperparameter configuration is set as iterations = 150, learning rate = 0.38, depth = 3, and l2_leaf_reg = 0.04. This configuration provides a practical balance between model capacity and regularization, leading to stable training behavior in subsequent experiments.

The following experiments use this fixed configuration for all comparative evaluations.

3.2. Lithotype-Aware Characterization of Rock Mechanical Parameters

Following data preprocessing and model optimization, the heterogeneity-aware residual formulation was applied to the prediction wells for characterization of rock mechanical parameters. To evaluate the complete geomechanical response, predictive performance was assessed across all five output targets on the independent test well, including point-scale labels for UCS , E_s , and ν , as well as lithotype-level estimates for C and ϕ . Table 5 reports target-specific units and metrics, avoiding the unit inconsistency that would arise from aggregating these parameters into a single dimensional error value.

Table 5. Predictive performance of the residual model on the independent test set across five geomechanical targets. Cohesion and friction angle are evaluated as lithotype-level estimates derived from triaxial strength-envelope regressions; their metrics quantify recovery of lithotype-level mechanical baselines rather than independent point-scale continuous predictions.

Target Parameter	Unit	R^2	RMSE	MAE	MAPE (%)
UCS	MPa	0.928	0.875	0.641	5.32
E_s	GPa	0.915	0.322	0.215	4.85
ν	–	0.898	0.018	0.012	6.12
C	MPa	0.902	0.445	0.312	5.88
ϕ	°	0.912	1.352	0.985	4.25

The model maintains high predictive accuracy across both strength and elastic parameters, with R^2 values ranging from 0.898 to 0.928 and MAPE values below 6.2%. For C and ϕ , these metrics indicate that the model reconstructs the lithotype-specific mechanical baselines derived from triaxial strength envelopes. Accordingly, the high R^2 values of C and ϕ should not be interpreted as independent point-scale continuous predictions, but as successful recovery of lithotype-level baseline estimates. This consistency indicates that the residual formulation captures lithotype-controlled shifts in the broader mechanical response rather than improving UCS alone.

To further examine the effectiveness of the method under heterogeneous conditions, Well-Validate-1 and Well-Test-1 were selected as representative cases for comparison with conventional empirical approaches. Traditional methods apply a single parametric relationship across the entire well interval, which leads to systematic errors in lithological transition zones.

In contrast, residual decomposition separates the global response from lithological correction, reducing errors in heterogeneous intervals.

3.2.1. Case Study: Well-Validate-1

Well-Validate-1 was selected as a representative validation well. In the depth interval of 2700–2950 m, a relatively stable coal seam with a thickness of approximately 48 m is developed, interbedded with 4–6 parting layers totaling about 12 m. A total of 23 experimental measurements are available within this interval, covering all four lithotypes (bright, semi-bright, semi-dull, and dull coal). The predicted profiles, including HMLZ, lithotype classification, and mechanical parameters, are shown in Figure 7.

The predicted UCS profile shows clear alignment with lithotype variations along the depth axis. In the bright coal interval (2750–2800 m), where HMLZ values indicate low-strength lithotypes, the predicted UCS remains within a low range (18–28 MPa), consistent with measured values and yielding a Mean Absolute Error (MAE) of 2.3 MPa. Local fluctuations within this interval are also captured, reflecting sensitivity to small-scale structural variations.

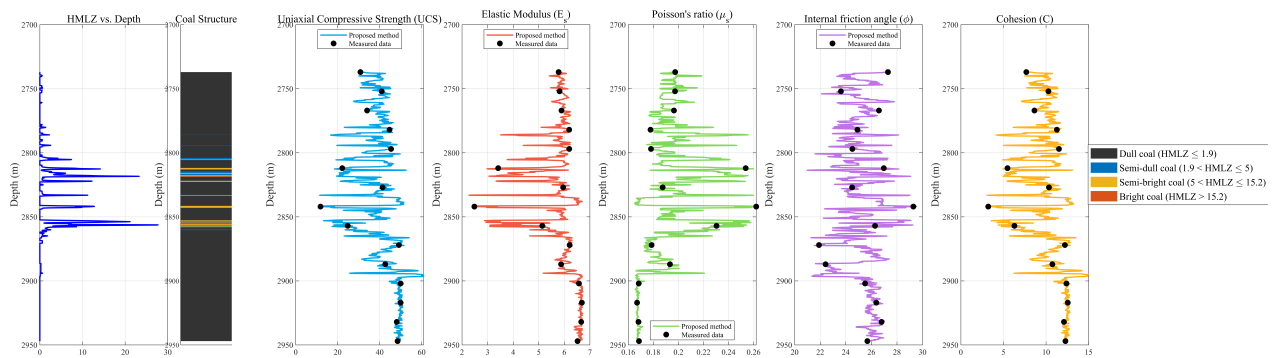


Figure 7. High-resolution predicted rock mechanical parameter profiles for Well-Validate-1.

At the lithological transition near 2800 m, an increase in HMLZ corresponds to a corresponding rise in predicted UCS over a short depth interval. This transition is closely aligned with measured data, indicating that the model responds to lithotype changes rather than producing a smoothed global trend. In the dull coal section around 2820 m, the predicted UCS stabilizes within the high-strength range and remains consistent with experimental observations.

In contrast, the traditional empirical approach exhibits systematic distortion across lithotypes. In bright coal, it overestimates UCS due to bias toward higher-strength samples, while in dull coal it underestimates strength. More critically, the traditional prediction curve lacks sensitivity to lithological transitions, resulting in a smooth, low-frequency trend that fails to capture abrupt changes in mechanical behavior.

A lithotype-wise comparison reveals that errors in the traditional method are concentrated in extreme lithotypes, whereas the residual model maintains consistently low errors across all categories. Across all 23 samples, MAE decreases from 9.2 MPa to 3.1 MPa and the coefficient of determination increases from 0.69 to 0.94.

The same tendency is observed for other mechanical parameters. E_s decreases in low-strength lithotypes and increases in high-strength lithotypes, while ν exhibits the opposite trend. These patterns are captured by the residual model but distorted in the traditional approach, indicating physically consistent relationships across multiple targets.

3.2.2. Case Study: Well-Test-1

Well-Test-1 serves as an independent test well that was not involved in either the training or validation stages. It is therefore used to evaluate the generalization capability of the proposed heterogeneity-aware formulation under more complex geological conditions. Compared with Well-Validate-1, this well exhibits significantly stronger lithological heterogeneity. Within the 2760–2920 m interval, the HMLZ curve identifies 37 lithotype transitions, corresponding to an average spacing of approximately 4.3 m, indicating a high-frequency heterogeneous system. In addition, multiple thin parting layers are present in the 2800–2850 m interval, forming frequent interbedded contacts with coal seams. This configuration imposes stringent requirements on the model's ability to resolve rapid lithological transitions and to maintain consistency across narrow depth intervals. A total of 18 experimental measurements are available, including nine points located within high-frequency transition zones, providing a demanding test scenario. The prediction results are shown in Figure 8.

In the high-frequency transition interval (2800–2850 m), the prediction curve responds clearly to lithotype variations. Multiple step-like changes in UCS are captured within short depth ranges, and the curve remains synchronized with lithotype transitions indicated by the HMLZ profile. At lithological interfaces, the model produces rapid and localized adjustments in predicted strength, consistent with measured data. Thin parting layers are

also correctly identified as high-strength zones, with predictions transitioning smoothly to adjacent coal intervals.

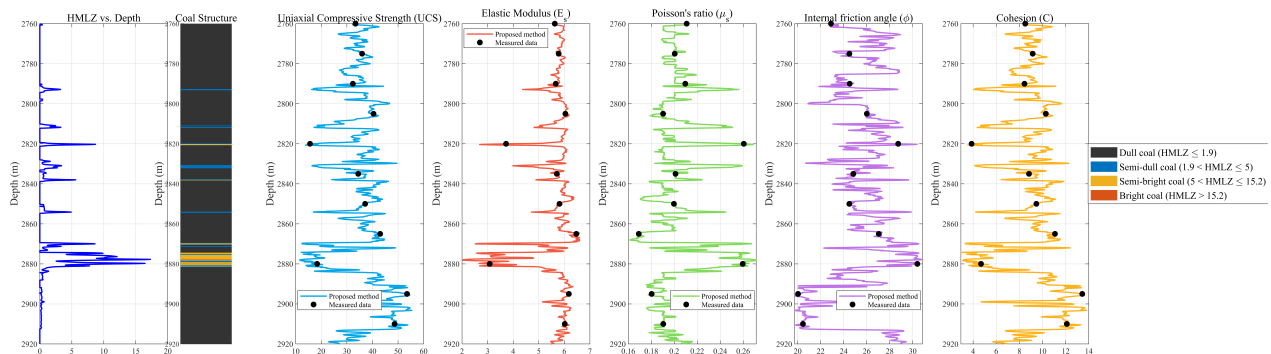


Figure 8. High-resolution predicted rock mechanical parameter profiles for Well-Test-1. The HMLZ category panel uses discrete color bands to distinguish lithotype classes; narrow non-gray bands mark high-frequency lithotype transitions.

In contrast, the traditional empirical approach shows a loss of stability under these conditions. The prediction curve exhibits irregular oscillations that are not aligned with lithotype changes, and abrupt variations appear even in relatively stable intervals.

Quantitative evaluation further highlights this difference. In the high-frequency transition zone, the residual model maintains low prediction error and a high proportion of samples within a narrow error band, whereas the traditional method shows significantly larger deviations and reduced consistency. Across the entire well, the residual model achieves substantially lower error and higher correlation with measured values, while maintaining similar performance levels to those observed in the validation well. The small difference in error between validation and test wells suggests stable predictive performance on unseen data.

The alignment between predicted step changes and lithotype transitions indicates that the model responds to regime changes rather than producing random fluctuations. Overall, the decomposition remains stable under both moderate and high-frequency heterogeneity.

3.3. Ablation Study Results and Quantitative Analysis

To quantify the role of the residual decomposition formulation, three modeling configurations were evaluated on the independent test well (Well-Test-1). The comparison focuses on uniaxial compressive strength (UCS) prediction and is summarized in Table 6.

Table 6. Quantitative performance comparison of different modeling strategies on the test set (UCS prediction).

Method	R^2	RMSE (MPa)	MAE (MPa)	MAPE (%)
Baseline ($f(X)$)	0.785	2.15	1.62	12.45
Lithotype as Feature ($f(X, L)$)	0.862	1.42	1.05	8.12
Residual ($f(X) + g(X, L)$)	0.928	0.875	0.641	5.32

The baseline model, which relies solely on logging-depth features, captures the overall variation trend but exhibits significant errors in intervals with frequent lithotype changes.

Introducing lithotype as an additional input improves performance, indicating that lithological information provides useful constraints. However, the improvement remains limited. It should be emphasized that the HMLZ-derived lithotype $L(d)$ is not independent external information; it is obtained from a nonlinear transformation and thresholding of logging responses that partly overlap with $X(d)$. Therefore, the comparison between

$f(X, L)$ and $f(X) + g(X, L)$ controls for the same lithotype condition, and the remaining improvement should be interpreted as an architectural effect of residual decomposition rather than as a gain from additional input information. The single-stage $f(X, L)$ model was optimized under the same BO–CV protocol, so the comparison is made under identical tuning conditions; nevertheless, more complex interaction-augmented single-stage designs may be explored in future work. Crucially, the $f(X, L)$ configuration—where lithotype is encoded as a categorical feature in a single CatBoost model—still retains the residual patterns discussed in the subsequent residual analysis despite CatBoost’s native categorical handling capability.

In contrast, the residual decomposition configuration models lithotype-induced deviations as a separate component. This leads to a noticeable performance increase, with R^2 increasing to 0.928 and MAE reduced by more than 60% compared to the baseline. More importantly, this change is achieved without introducing new information, but by restructuring the mapping itself.

The gap between $f(X, L)$ and $f(X) + g(X, L)$ indicates that lithotype-related errors are more effectively captured through an explicit residual correction than through a single feature-augmented mapping.

Overall, the ablation study provides quantitative evidence that the proposed approach differs from conventional feature augmentation.

3.4. Mechanism Analysis of Structured Residual Correction

To verify that the ablation benefit arises from the residual formulation rather than from a purely empirical two-stage refinement, a dedicated residual analysis was conducted. Here, “conditional bias” refers to systematic residual patterns that depend on lithological regime, rather than random noise. This analysis examines the error distribution of the baseline model $f(X)$ and evaluates the effect of the residual component $g(X, L)$.

3.4.1. Identification of Structured Bias in the Global Baseline

A fundamental assumption of this study is that the baseline mapping $f(X)$ can introduce biased errors in heterogeneous formations because it “averages” the mechanical responses of different lithotypes. To test this, the Signed Mean Residual (SMR) and Standard Deviation (SD) of the baseline model were calculated for each coal lithotype in the test set.

As shown in Table 7, the baseline residuals are not randomly distributed white noise; instead, they exhibit a clear polarity tied to the lithological regime. In low-strength bright coal intervals, the baseline model consistently overestimates UCS (SMR = +2.45 MPa), whereas in high-strength dull coal intervals, it tends to underestimate the strength (SMR = −3.12 MPa). This pattern shows that the baseline error is geologically structured.

Table 7. Signed mean residual (SMR) across lithotypes for baseline ($f(X)$), feature-augmented ($f(X, L)$), and proposed ($f(X) + g(X, L)$) models.

Coal Lithotype	SMR: M0 (MPa)	SMR: M1 (MPa)	SMR: M2 (MPa)	Std. Dev. (MPa)
Bright coal	+2.45	+2.15	+0.12	1.12
Semi-bright coal	+0.84	+0.65	+0.08	0.95
Semi-dull coal	−0.62	−0.58	−0.05	1.04
Dull coal	−3.12	−2.85	−0.18	1.48

3.4.2. Amplification of Bias in Transition Zones

The structured bias is further intensified in lithological transition zones. We defined “Transition Zones” as intervals within 0.5 m of a lithotype boundary identified by the HMLZ index. Figure 9 compares the residual density between stable intervals and transition zones.

The smooth probability–density curves were generated from the discrete residual samples at the 97 laboratory control points after assigning each sample to either a stable interval or a transition zone. A Gaussian kernel density estimator was used, and the bandwidth was selected automatically using Scott’s rule to avoid manual smoothing. The KDE curves are used only for visualizing the residual distribution; the SMR and SD statistics are calculated directly from the discrete residual values.

In stable lithological intervals, the baseline model exhibits a relatively narrow error distribution. However, in transition zones, the residual variance increases by approximately 140%, and the distribution becomes markedly bimodal. This phenomenon indicates that near lithological interfaces, the baseline model fails to track rapid shifts in geomechanical response, even when the logging signals (e.g., AC or DEN) show only subtle variations.

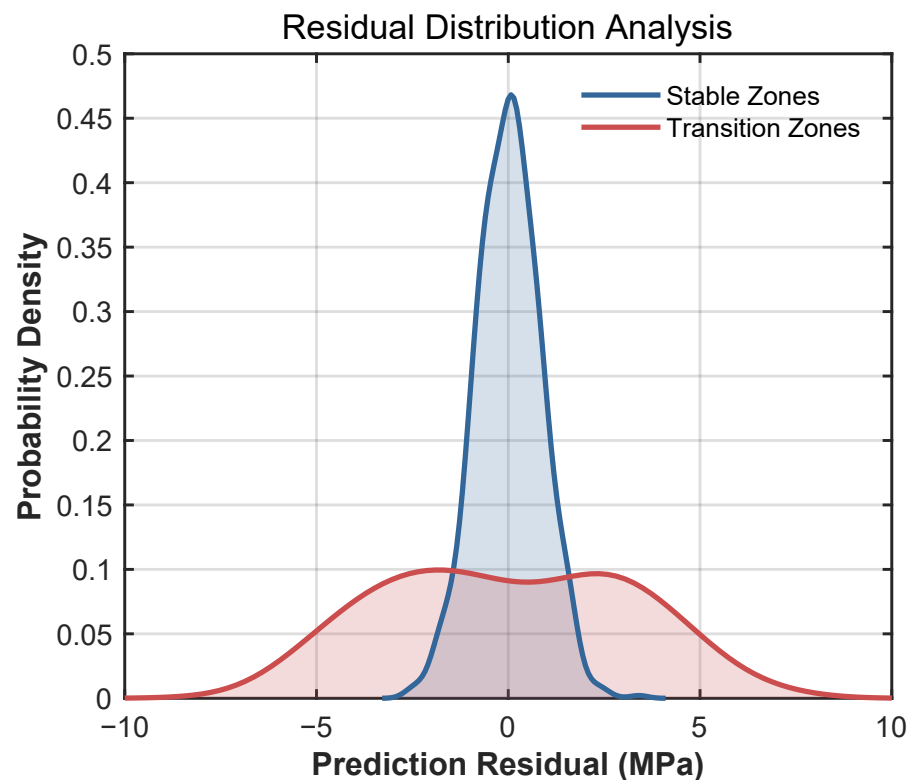


Figure 9. Comparison of baseline residual distributions in stable lithological zones versus high-frequency transition zones, illustrating the expansion of error variance under heterogeneity.

3.4.3. Effectiveness of the Lithotype-Conditioned Correction

The correction term $g(X, L)$ addresses these structured errors directly. Figure 10 illustrates the “flattening” effect of the residual correction. After incorporating this component, the SMR for all lithotypes converged toward zero (e.g., dull coal SMR improved from -3.12 MPa to -0.18 MPa).

Crucially, the standard deviation of the residuals also decreased across all regimes, indicating that $g(X, L)$ does not just shift the mean but also reduces the uncertainty within each lithotype. This transition from a “lithotype-biased” error to a “lithotype-neutral” error supports the residual correction design.

This observation indicates that the residual term is explicitly dependent on lithotype, supporting the use of lithological regimes as conditioning variables.

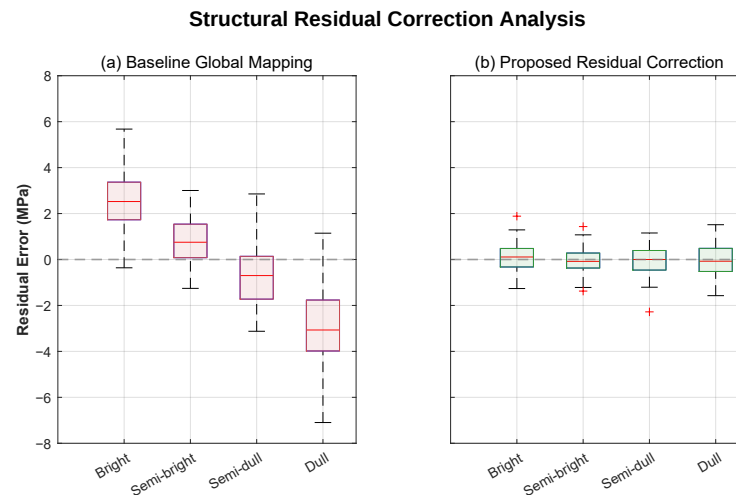


Figure 10. Box plots of prediction residuals for the four lithotypes: (a) baseline model, showing systematic bias; (b) residual model, showing centered and narrowed residual distributions.

The residual analysis demonstrates that baseline errors are geologically structured and that the proposed decomposition reduces this structure in the final predictions.

3.5. Heterogeneity-Focused Evaluation in Transition Zones

To further examine model behavior in transition zones, a focused evaluation was conducted on the 2800–2850 m interval of Well-Test-1, where bright coal and dull coal are frequently interbedded. The quantitative results are summarized in Table 8.

Table 8. Performance comparison in high-frequency lithological transition zones.

Method	R^2 (Transition)	RMSE (MPa)	Max Error (MPa)
Baseline ($f(X)$)	0.582	3.42	12.8
Lithotype as Feature ($f(X, L)$)	0.724	1.85	8.2
Residual ($f(X) + g(X, L)$)	0.894	0.92	3.1

The baseline model shows a pronounced degradation in performance within this interval, producing overly smoothed predictions that fail to capture sharp variations in mechanical properties across lithological boundaries.

Introducing lithotype as an additional feature improves sensitivity to lithological variation, but substantial errors remain.

In contrast, residual decomposition maintains stable predictive performance in the transition zone. The improvement is particularly evident in the reduction in maximum error, suggesting that abrupt changes in mechanical properties are better captured.

The error source in transition zones should be interpreted carefully. The residual component can correct lithotype-related systematic bias, as indicated by the reduction in maximum error from 8.2 MPa for $f(X, L)$ to 3.1 MPa for $f(X) + g(X, L)$ under identical logging inputs. However, part of the transition-zone error is controlled by measurement physics rather than by the mapping algorithm. Near lithological interfaces and thin interbeds, conventional wireline logs have limited vertical resolution (typically 0.3–0.6 m), and volume averaging together with shoulder-bed effects can produce mixed responses rather than a “pure” lithotype signal [36–38]. These sub-meter effects are intrinsic limitations of the logging tools and cannot be eliminated at the modeling level alone.

3.6. SHAP-Based Interpretability Analysis

SHAP (SHapley Additive Explanations) was employed to analyze the relationships between logging responses, lithological regimes, and predicted mechanical parameters. SHAP is used here as a post hoc interpretability tool to examine whether the learned feature-response relationships are consistent with known geomechanical principles.

At the global level, the mean absolute SHAP value was calculated for both the training and test sets using the complete model input vector. As shown in Table 9, the dominant contributions are associated with the HMLZ-derived lithotype, acoustic transit time (AC), gamma ray (GR), density (DEN), and resistivity-related features. CNL, SP, and depth were retained in the model input for physical completeness but had smaller mean SHAP values than the leading variables shown in the table and summary plot. The overall consistency of the dominant feature ranking between the training and test sets indicates that the model captures stable controlling factors across wells, rather than overfitting to specific local patterns.

Table 9. Dominant feature importance ranking based on SHAP values. The lithotype entry denotes the HMLZ-derived categorical lithotype; the complete model input vector is defined in Section 2.3.2.

Rank	Training Set	Mean SHAP	Test Set	Mean SHAP
1	HMLZ-derived lithotype	0.574	HMLZ-derived lithotype	0.558
2	AC	0.353	AC	0.364
3	GR	0.346	GR	0.353
4	DEN	0.332	DEN	0.320
5	LLD	0.248	LLD	0.286
6	LLS	0.216	LLS	0.188

This global ranking is physically interpretable. The dominant role of the HMLZ-derived lithotype indicates that categorical lithological regime controls a major part of the mechanical-response shift in heterogeneous coal seams. Among the raw logging-depth variables, AC is the most influential feature; it reflects the propagation behavior of acoustic waves and is closely associated with fracture development, pore structure, and structural integrity. DEN characterizes material compactness and bulk structural condition, while GR provides supplementary information related to compositional variability, ash content, and clay-related effects. LLD and LLS represent resistivity responses linked to pore-fluid, cleat, and compositional variations, and their secondary but stable ranking suggests that they complement AC, DEN, and GR rather than independently controlling the prediction. CNL, SP, and depth provide additional porosity, electrochemical, and burial-trend constraints even when their global SHAP rankings are secondary. Together, these features form a physically meaningful basis for geomechanical characterization.

To further examine the direction and distribution of dominant feature effects, a SHAP summary plot was generated, as shown in Figure 11. The summary plot presents the SHAP contribution of each sample together with the corresponding feature value for the leading variables, thereby revealing both the magnitude and polarity of feature influence. High AC values generally correspond to negative SHAP contributions, indicating a reduction in the predicted mechanical parameters, whereas low AC values tend to produce positive contributions. This agrees with rock physics expectations, since larger transit time is usually associated with poorer structural integrity and lower load-bearing capacity. In contrast, higher DEN values generally produce positive contributions, reflecting the higher strength expected in denser and more compact formations. The color separation and horizontal spread in Figure 11 further indicate that similar raw log values can produce different contributions when the lithotype condition changes, which supports the residual decomposition interpretation.

Unlike simple correlation analysis, SHAP can reveal conditional effects arising from nonlinear interactions. To assess whether lithological heterogeneity is reflected in the model behavior, the SHAP contributions associated with coal lithotypes were further examined. The analysis shows that even within similar AC or DEN intervals, SHAP values exhibit systematic offsets across lithotypes. Lower-strength lithotypes tend to produce more negative contributions, whereas higher-strength lithotypes more often produce positive contributions.

This result provides additional interpretive support. If lithotype-induced variation were merely random noise, samples with similar logging features would show similar SHAP contributions regardless of lithotype. Instead, the observed separation suggests that the model captures regime-dependent behavior.

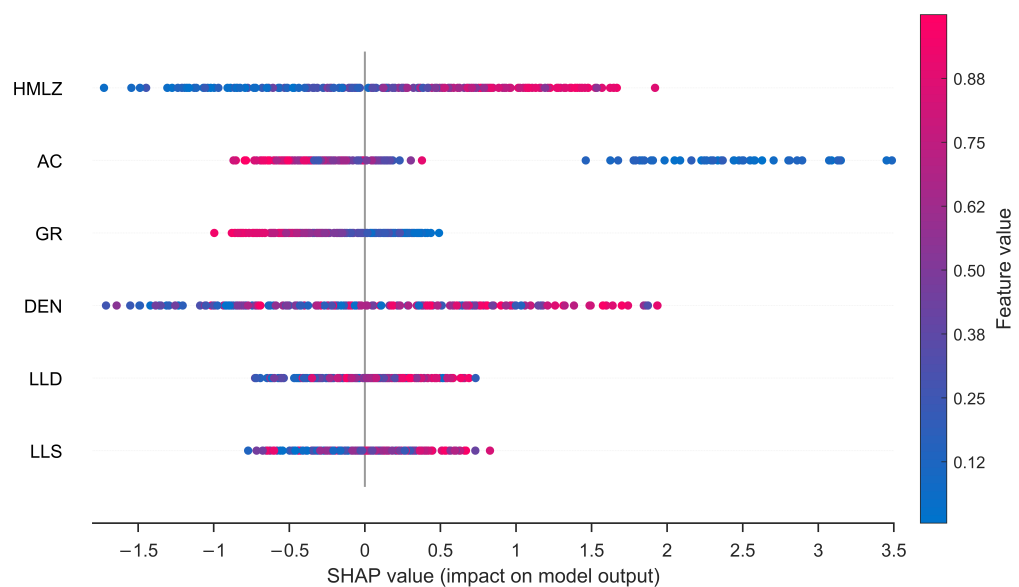


Figure 11. SHAP value distribution of the dominant features for model output. Feature colors represent min–max normalized feature values, with blue corresponding to values close to 0 and red corresponding to values close to 1.

To visualize this modulation more directly, local contribution analyses were performed for representative depth samples. Figure 12 shows how the final prediction is decomposed into additive feature contributions for specific examples. For low-strength samples located in bright coal intervals, the prediction below the baseline is typically driven by the joint effect of high AC, low DEN, and lithotype-associated negative contributions. For high-strength samples in dull or semi-dull coal intervals, the opposite pattern is observed, with low AC, high DEN, and lithotype-associated positive contributions driving the prediction above the baseline. This decomposition provides a traceable explanation for why the predicted mechanical parameter at a given depth is high or low.

Taken together, the SHAP results provide consistent evidence from multiple perspectives: global importance ranking, direction of feature influence, lithotype-related modulation, and local sample-wise decomposition. The interpretability analysis suggests physically meaningful and geologically consistent patterns. It should be acknowledged that the HMLZ-derived lithotype is obtained by thresholding an index calculated from a subset of the input logging features (AC, DEN, GR, and resistivity), while AC, DEN, GR, and resistivity-related curves are also included in the input feature set. This overlap introduces multicollinearity and can dilute or redistribute SHAP importance among correlated variables. Therefore, the absolute ranking and magnitude of SHAP values should be interpreted cautiously. This limitation affects feature-attribution interpretation, but it does

not invalidate the residual structure itself, because the main evidence for bias correction is provided by the SMR/SD residual analysis and the ablation comparison rather than by SHAP importance scores.

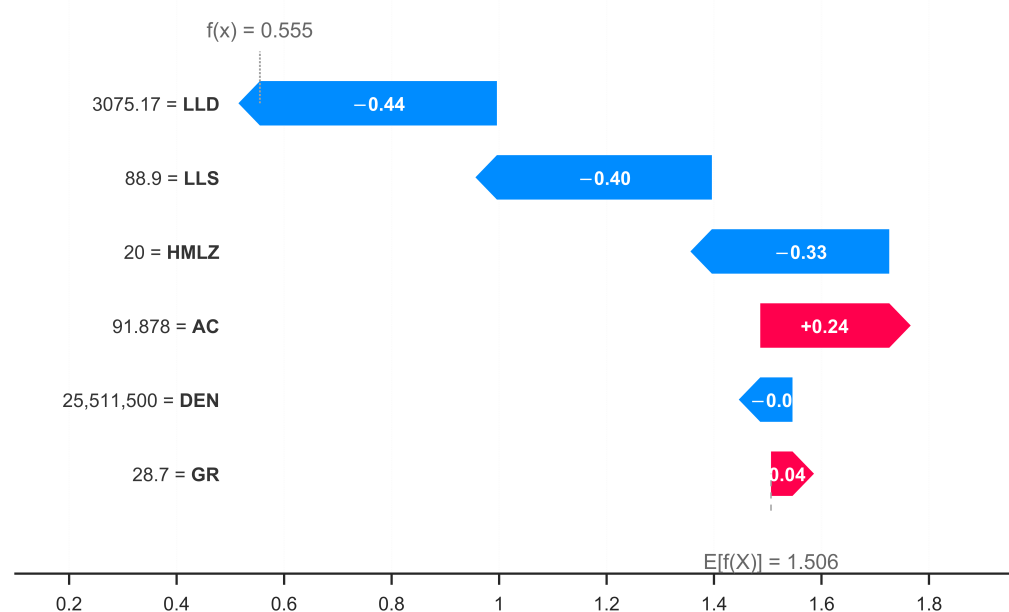


Figure 12. Analysis of feature contributions to the model output. Red arrows indicate positive SHAP contributions that increase the prediction, whereas blue arrows indicate negative SHAP contributions that decrease the prediction.

The SHAP-based analysis demonstrates geologically reasonable feature behavior and provides interpretability support for the residual formulation.

4. Conclusions

This study proposes a lithotype-conditioned residual decomposition framework for geomechanical characterization from well logs in heterogeneous coal-bearing formations. By separating a global baseline from a lithological correction, the approach reduces biased errors in heterogeneous intervals and improves the consistency between predicted mechanical properties and lithological regime.

Cross-well evaluation shows that the method maintains stable performance across training, validation, and independent test wells, suggesting that the learned relationships generalize across wells within the same formation. Case studies further demonstrate that the model responds effectively to lithological transitions and reduces prediction errors in extreme lithotypes, particularly in intervals where heterogeneity is pronounced.

Ablation experiments indicate that incorporating lithotype as a standard feature provides limited improvement, whereas residual decomposition introduces additional gains. Interpretability analysis suggests that feature contributions vary by lithotype, supporting the role of the residual component.

Despite these findings, several limitations should be noted. The dataset is derived from a specific coal-bearing formation within a single basin, and the number of wells is limited. Therefore, the strong performance observed in the independent test well should be interpreted as evidence of within-formation cross-well generalization, not as proof of basin-scale transferability. In addition, the reported metrics are point estimates from the available cross-well dataset; confidence intervals or bootstrap uncertainty estimates should be incorporated when larger multi-well datasets become available. Moreover, while the residual component is interpreted as representing structured bias, its generality across different datasets requires additional investigation. Furthermore, the lithotype classification

relies on the empirical HMLZ index rather than physics-based constraints, which limits the conclusions to local predictive performance within the specific coal-bearing formation studied; broader applicability requires validation under different basin conditions and logging configurations. Scale mismatch between core-based laboratory measurements and log-derived responses may also introduce calibration uncertainty [39]. Finally, the limited vertical resolution of conventional logging tools (volume averaging and shoulder-bed effects) constitutes an intrinsic constraint of the current approach; at lithological interfaces, this limitation cannot be eliminated by modeling alone, and future work may mitigate it by integrating higher-resolution imaging logs or rock-physics forward modeling for interface-response correction.

Future work will focus on multi-basin external validation, additional lithological systems, broader logging combinations, and rock-physics-informed constraints so that the observed within-formation performance can be tested under wider geological conditions.

Overall, this study provides a structured approach for incorporating lithological heterogeneity into geomechanical characterization and demonstrates its effectiveness in reducing bias and improving cross-well consistency under the conditions considered.

Author Contributions: Conceptualization: X.L. and W.Z. (Weixian Zhang); methodology, X.L.; software, X.L.; validation, B.D., L.L. and W.Z. (Wenze Zhou); formal analysis, W.Z. (Weixian Zhang); investigation, X.L.; resources, L.L.; data curation, X.L. and W.Z. (Weixian Zhang); writing—original draft preparation, X.L. and W.Z. (Weixian Zhang); writing—review and editing, B.D. and L.L.; visualization, X.L.; supervision, W.Z. (Weixian Zhang); project administration, W.Z. (Weixian Zhang); funding acquisition, W.Z. (Weixian Zhang). All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data and source code that support the findings of this study are available from the corresponding author upon reasonable request.

Acknowledgments: The authors wish to acknowledge the use of DeepSeek-V3.2 for English language polishing during the preparation of this manuscript.

Conflicts of Interest: Authors Xugang Liu, Binghua Dang, Lei Li were employed by the company Sinopec North China Oil and Gas Company. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Li, S.; Qin, Y.; Tang, D.; Shen, J.; Wang, J.; Chen, S. A comprehensive review of deep coalbed methane and recent developments in China. *Int. J. Coal Geol.* **2023**, *279*, 104369. [[CrossRef](#)]
2. Guo, Z.; Zhao, J.; You, Z.; Li, Y.; Zhang, S.; Chen, Y. Prediction of coalbed methane production based on deep learning. *Energy* **2021**, *230*, 120847. [[CrossRef](#)]
3. Zhu, Q.; Du, X.; Zhang, T.; Yu, H.; Liu, X. Investigation into the variation characteristics and influencing factors of coalbed methane gas content in deep coal seams. *Sci. Rep.* **2024**, *14*, 18813. [[CrossRef](#)]
4. Miao, Q.; Liu, H.; Wang, Y.; Wang, W.; Li, S.; Zhai, W.; Wei, K. Quantitative Mechanisms of Long-Term Drilling-Fluid-Coal Interaction and Strength Deterioration in Deep CBM Formations. *Processes* **2025**, *13*, 3183. [[CrossRef](#)]
5. Zhang, Y.; Zhou, J.; Li, J.; He, B.; Armaghani, D.J.; Huang, S. Advancing overbreak prediction in drilling and blasting tunnel using MVO, SSA and HHO-based SVM models with interpretability analysis. *Geomech. Geophys. Geo-Energy Geo-Resour.* **2025**, *11*, 53. [[CrossRef](#)]
6. Ding, Y.; Li, B.; Li, J.; Song, H.; Zeng, X. A mini-review on coal permeability under combined thermal and mechanical effects. *Energy Fuels* **2025**, *39*, 21659–21676. [[CrossRef](#)]

7. Gao, M.Z.; Gao, Z.; Yang, B.G.; Xie, J.; Wang, M.Y.; Hao, H.C.; Wu, Y.; Zhou, L.; Wang, J.Y. Macroscopic and microscopic mechanical behavior and seepage characteristics of coal under hydro-mechanical coupling. *J. Cent. South Univ.* **2024**, *31*, 2765–2779. [\[CrossRef\]](#)
8. Guo, H.; Sun, Z.; Ji, M.; Wu, Y.; Nian, L. An investigation on the impact of unloading rate on coal mechanical properties and energy evolution law. *Int. J. Environ. Res. Public Health* **2022**, *19*, 4546. [\[CrossRef\]](#) [\[PubMed\]](#)
9. Aziz, Q.A.A.; Awadh, S.M.; Al-Mimar, H.S. Estimation of rock mechanical properties of the Hartha Formation and their relationship to porosity using well-log data. *Iraqi Geol. J.* **2024**, *57*, 34–44. [\[CrossRef\]](#)
10. Xin, F.; Xu, H.; Tang, D.; Cao, C. Differences in accumulation patterns of low-rank coalbed methane in China under the control of the first coalification jump. *Fuel* **2022**, *324*, 124657. [\[CrossRef\]](#)
11. Zhang, Q.; Li, Y.; Li, Z.; Yao, Y.; Du, F.; Wang, Z.; Tang, Z.; Zhang, W.; Wang, S. Fracture Propagation Laws and Influencing Factors in Coal Reservoirs of the Baode Block, Ordos Basin. *Energies* **2024**, *17*, 6183. [\[CrossRef\]](#)
12. Li, C.; Zhang, W. Application of Deep Neural Networks in Spatial Estimation of Logging Parameters. In *Proceedings of the International Conference on Geology, Energy and Oil and Gas Exploration*; Springer: Cham, Switzerland, 2025; pp. 734–740.
13. Sanei, M.; Ramezanzadeh, A.; Delavar, M.R. Applied machine learning-based models for predicting the geomechanical parameters using logging data. *J. Pet. Explor. Prod. Technol.* **2023**, *13*, 2363–2385. [\[CrossRef\]](#)
14. Hiba, M.; Ibrahim, A.F.; Elkhatny, S.; Ali, A. Application of machine learning to predict the failure parameters from conventional well logs. *Arab. J. Sci. Eng.* **2022**, *47*, 11709–11719. [\[CrossRef\]](#)
15. Gabry, M.A.; Ali, A.G.; Elsayy, M.S. Application of Machine Learning Model for Estimating the Geomechanical Rock Properties Using Conventional Well Logging Data. In *Proceedings of the Offshore Technology Conference, OTC*, Houston, TX, USA, 1–4 May 2023; p. D021S028R004.
16. Rohit, Manda, S.R.; Raj, A.; Andraju, N. A machine learning approach to predict geomechanical properties of rocks from well logs. *Int. J. Data Sci. Anal.* **2025**, *20*, 653–670. [\[CrossRef\]](#)
17. Khetani, N.; Shah, V.; Gajera, D.; Pathak, O.; Ramalingam, V. Prediction of Geo-mechanical Parameter Logs from Petrophysical Well Logs using Machine Learning Approach. *J. Geol. Soc. India* **2024**, *100*, 1419–1432. [\[CrossRef\]](#)
18. Mollaei, F.; Moradzadeh, A.; Mohebian, R. Novel approaches in geomechanical parameter estimation using machine learning methods and conventional well logs. *Geosystem Eng.* **2024**, *27*, 252–277. [\[CrossRef\]](#)
19. Dong, M.; Zhu, T.; Wu, C.; Zhang, G.; Chen, H. Mechanical parameter prediction based on drilling engineering and well-logging information. *J. Geophys. Eng.* **2025**, *22*, 1333–1343. [\[CrossRef\]](#)
20. Pires, B.; Lima, V.; Silva, F.; Velloso, R. On the Role of Rock Lithotype, Porosity, and Permeability in Shear Bond Strength of Rock-Class G Cement Paste Interfaces. *SPE J.* **2025**, *30*, 3456–3475. [\[CrossRef\]](#)
21. Lobarinhas, R.; Dionísio, A.; Paneiro, G. High temperature effects on global heritage stone resources: A systematic review. *Heritage* **2024**, *7*, 6310–6342. [\[CrossRef\]](#)
22. Vigroux, M.; Eslami, J.; Beaucour, A.L.; Bourges, A.; Noumowé, A. High temperature behaviour of various natural building stones. *Constr. Build. Mater.* **2021**, *272*, 121629. [\[CrossRef\]](#)
23. 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. [\[CrossRef\]](#)
24. Shao, Y.; Wang, H.; Guo, Y.; Huang, X.; Wang, Y.; Zhao, S.; Zhu, Y.; Shen, L.; Huang, X.; Song, Y.; et al. Geological characteristics and gas-bearing evaluation of coal-measure gas reservoirs in the Huanghebei coalfield. *Front. Earth Sci.* **2023**, *11*, 1104418. [\[CrossRef\]](#)
25. Wang, Z.; Gong, K.; Chen, J.; Zhou, K.; Xu, S.; Lu, J.; Hu, J.; Cai, Y. The Influence of Coal Lithotype and Metamorphism on the Development Characteristics of Pore and Fracture in Coalbed Methane Reservoir in Ordos Basin and Qinshui Basin. *J. GeoEnergy* **2024**, *2024*, 5931614. [\[CrossRef\]](#)
26. Zhang, Y.; Chen, Y.; Zhang, S.; Feng, G.; Wang, Y.; Li, S.; Wang, Q.; Wang, B.; Zhao, L. Study on the Influence of Drilling Parameters on the Mechanical Properties and Pressure Relief Effect of Coal Rock. *Processes* **2025**, *13*, 993. [\[CrossRef\]](#)
27. Feng, X.T.; Haimson, B.; Li, X.; Chang, C.; Ma, X.; Zhang, X.; Ingraham, M.; Suzuki, K. ISRM suggested method: Determining deformation and failure characteristics of rocks subjected to true triaxial compression. *Rock Mech. Rock Eng.* **2019**, *52*, 2011–2020. [\[CrossRef\]](#)
28. Ibrahim, A.F.; Hiba, M.; Elkhatny, S.; Ali, A. Estimation of tensile and uniaxial compressive strength of carbonate rocks from well-logging data: Artificial intelligence approach. *J. Pet. Explor. Prod. Technol.* **2024**, *14*, 317–329. [\[CrossRef\]](#)
29. Su, J.; Zhang, J.; Wang, M.; Qin, Z.; Grebby, S. Vertical Distribution Heterogeneity of Pore Structure Collected from Deep, Thick Coal Seams. *Processes* **2026**, *14*, 240. [\[CrossRef\]](#)
30. Prokhorenkova, L.; Gusev, G.; Vorobev, A.; Dorogush, A.V.; Gulin, A. CatBoost: Unbiased boosting with categorical features. *Adv. Neural Inf. Process. Syst.* **2018**, *31*, 6638–6648.

31. Yi, Z.; Zou, Y. ADASYN-CatBoost method for lithology identification with imbalanced well-logging data: A case study of Zhaoxian Gold deposit in northwest Jiaodong peninsula, China. In *Proceedings of the Global Meeting Abstracts; Society of Exploration Geophysicists*: Tulsa, OK, USA, 2022; pp. 147–151. [\[CrossRef\]](#)
32. Hossain, T.M.; Hermans, M.; Olutoki, J.O. A novel stochastic catboost based shear wave velocity prediction and uncertainty analysis in sandstone reservoir using Multi-Seismic attributes. *IEEE Access* **2024**, *12*, 168160–168170. [\[CrossRef\]](#)
33. Wang, X.; Jin, Y.; Schmitt, S.; Olhofer, M. Recent advances in Bayesian optimization. *ACM Comput. Surv.* **2023**, *55*, 1–36. [\[CrossRef\]](#)
34. Antwarg, L.; Miller, R.M.; Shapira, B.; Rokach, L. Explaining anomalies detected by autoencoders using Shapley Additive Explanations. *Expert Syst. Appl.* **2021**, *186*, 115736. [\[CrossRef\]](#)
35. Widodo, S.; Brawijaya, H.; Samudi, S. Stratified K-fold cross validation optimization on machine learning for prediction. *Sink. J. Dan Penelit. Tek. Inform.* **2022**, *6*, 2407–2414. [\[CrossRef\]](#)
36. Molina Camargo, M.; Seyfang, B.; Basso, M.; Furlan Chinelatto, G.; Chandler, A.S.; Vidal, A.C. Challenges and suggestions for defining electrofacies models: The problem with well-log resolution and the resulting “shoulder bed effects” applied to a carbonate reservoir. *Interpretation* **2025**, *13*, T871–T886. [\[CrossRef\]](#)
37. Ma, J.; Huang, Y.; Kang, L.; Wang, Q.; Xia, L.; Hu, Y. Enhance the Vertical Resolution of Conventional Well Logs Using Auto-Encoder. In *Proceedings of the International Petroleum Technology Conference, IPTC*, Kuala Lumpur, Malaysia, 18–20 February 2025; p. D022S007R004. [\[CrossRef\]](#)
38. Xie, R.; Lin, X.; Jiang, S.; Wang, K.; Liu, J.; Lu, Y. Genetic Mechanism of Calcareous Interbeds in Shoreface Reservoirs and Implications for Hydrocarbon Accumulation: A Case Study of the Donghe Sandstone Reservoir in Hade Oilfield, Tarim Basin. *Minerals* **2026**, *16*, 259. [\[CrossRef\]](#)
39. El-Husseiny, A.; Al-Garadi, K.; Ali, A. Can laboratory ultrasonic measurements on core plugs reproduce results obtained from sonic well logging in carbonates? *J. Appl. Geophys.* **2022**, *204*, 104745. [\[CrossRef\]](#)

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.