International Journal of Earth Sciences Knowledge and Applications
journal homepage: http://www.ijeska.com/index.php/ijeska
e-ISSN: 2687-5993

Research Article https://doi.org/10.5281/zenodo.22166189

Azimuthal LWD Data Interpretation for UBCTD Geosteering Using a Physics-Informed Neural Network

Klemens Katterbauer[1]*, Saleh Komies[1], Mohd Azizi Ibrahim[2], Alberto Ache[3]

[1]UARC, Saudi Aramco, 31311 Dhahran, Saudi Arabia
[2]PE&D, Saudi Aramco, 31311 Dhahran, Saudi Arabia
[3]GOC, Saudi Aramco, 31311 Dhahran, Saudi Arabia

INFORMATION

Article history
Received 14 April 2026
Accepted 21 May 2026
Published 31 August 2026

Contact
*Klemens Katterbauer
klemens.arabiya@gmail.com (KK)

How cite
Katterbauer, K., Komies, S., Ibrahim, M.A., Ache, A., 2026. Azimuthal LWD Data Interpretation for UBCTD Geosteering Using a Physics-Informed Neural Network. *International Journal of Earth Sciences Knowledge and Applications* 8 (2), 278–290. https://doi.org/10.5281/zenodo.22166189.

Abstract
The economic success of horizontal wells in complex carbonate reservoirs is profoundly sensitive to precise wellbore placement within narrow target zones. Conventional geosteering, which relies on the real-time subjective interpretation of Logging-While-Drilling (LWD) data, is susceptible to human bias and often leads to suboptimal decisions. While data-driven machine learning offers an alternative, purely statistical models frequently produce physically implausible predictions. This paper introduces a novel Physics-Informed Neural Network (PINN) framework that seamlessly integrates domain knowledge with a deep learning architecture to automate and enhance geosteering classification. The methodology employs a multi-layer perceptron trained with a custom composite loss function, which augments standard cross-entropy loss with two physics-derived penalty terms: one enforcing petrophysical consistency between sensor readings and predicted steering actions, and another promoting wellbore trajectory smoothness. Trained and validated on a sophisticated synthetic dataset engineered to replicate the geological complexities of a Central Asia Shu formation carbonate reservoir, the model demonstrates a significant performance improvement over a purely data-driven baseline. The refined PINN model achieved a test accuracy of 91.18%, a substantial increase over the baseline Logistic Regression model's 86.68%. Crucially, the physics-informed constraints led to a dramatic enhancement in recognizing the optimal 'stay' condition, with the F1-score for this critical class rising from 0.01 to 0.51. An ablation study confirmed the petrophysical constraint as the primary driver of this improvement. Post-hoc explainability analysis using LIME (Local Interpretable Model-agnostic Explanations) verified that the model's decision-making aligns with established geological principles. This work successfully demonstrates that embedding physical constraints directly into the ML learning process yields a more robust, reliable, and interpretable autonomous geosteering system, paving the way for more objective, data-optimized well placement and reduced reliance on subjective interpretation.

Keywords
Azimuthal Resistivity; UBCTD; Geosteering; Neural Networks; PINN

1. Introduction

The relentless global demand for energy, coupled with the gradual depletion of conventional hydrocarbon resources, has propelled the development of unconventional reservoirs to the forefront of the oil and gas industry. Among these, tight gas reservoirs represent a vast and challenging resource base, characterized by extremely low permeability and complex, often nano-Darcy-scale, pore structures that inherently restrict fluid flow and impede economic production. Unlocking these resources is not merely a technical objective but a strategic imperative for energy security. In such demanding environments, Underbalanced Coiled Tubing Drilling (UBCTD) has emerged as a critical enabling technology for enhancing reservoir access and maximizing recovery (Mamman et al., 2025). The economic development of complex carbonate reservoirs is critically dependent on maximizing reservoir contact through precise horizontal well placement. These reservoirs are characterized by extreme heterogeneity, with abrupt vertical and lateral changes in porosity, permeability, and fluid saturation, making it challenging to navigate within narrow target "sweet spots." In such demanding environments, Underbalanced Coiled Tubing Drilling (UBCTD) has emerged as a critical enabling technology for enhancing reservoir access and maximizing recovery (Alsaud et al., 2024; AlShehri et al., 2023).

UBCTD operations, which maintain a bottomhole pressure lower than the formation pressure, offer significant advantages over conventional drilling—including minimized formation damage through the prevention of fluid invasion, improved rate of penetration (ROP) due to the reduced chip hold-down effect, and the potential for enhanced production profiling while drilling. However, the full realization of these benefits is contingent upon the successful execution of real-time geosteering, a formidable challenge required to navigate the thin, heterogeneous, and often discontinuous target layers of complex carbonates. The economic viability of these capital-intensive wells is exquisitely sensitive to precise wellbore placement within a narrow target window; even minor deviations can result in premature exit from the productive zone, leading to substantial financial downside and suboptimal drainage (Bouldin et al., 2013; Zhong et al., 2025).

This study directly addresses these multifaceted challenges through the development and application of a sophisticated, synthetic multiwell dataset, explicitly engineered to replicate the realistic petrophysical and operational conditions of a representative Central Asian complex carbonate reservoir (Blanc et al., 2003). The dataset provides a crucial and standardized foundation for advancing next-generation, data-driven geosteering methodologies (Table 1). Constructed with a known "ground truth" that is unattainable in real-world operations, it encapsulates the key complexities, including intricate stratigraphic heterogeneity, realistic well trajectories with undulations, comprehensive physics-based simulations of azimuthal LWD sensor responses, and the deliberate incorporation of operational artifacts like tool drift, noise, and borehole washout (Katterbauer et al., 2021; Kavanagh et al., 2005).

Table 1. Main dataset parameters

Column Name	Description	Units/Type
WELL_ID	Identifier of the well (WELL_1 … WELL_5).	categorical
DEPTH_M	Measured depth (MD) along borehole trajectory.	m
INCLINATION_DEG	Borehole deviation from vertical. Higher = more horizontal.	degrees
TOOL_AZIMUTH_DEG	Toolface azimuth relative to North (0–360°).	degrees
X_M	Easting coordinate of well path (drift over MD).	m
Y_M	Northing coordinate of well path.	m
Z_M	True Vertical Depth (TVD) below datum.	m
CALIPER_M	Simulated borehole diameter including washouts.	m
MUD_TYPE	Drilling fluid system.	"water_based", "oil_based"
RPM	Bit/tool rotational speed.	revolutions per minute
FLOW_RATE_L_S	Mud flow injection rate.	L/s
LAYER_NAME	Geological subzone within Shu.	"Shu_Upper", "Shu_Middle", "Shu_Lower"
LAYER_INDEX	Encoded numeric representation of layer_name (0,1,2).	integer
POROSITY	Effective porosity used for forward modelling.	fraction (0–1)
VCL	Clay volume fraction impacting GR/porosity.	fraction (0–1)
RFM	True formation resistivity (Rt clean, no invasion).	ohm·m
SW_TRUE	True water saturation from petro-physical model.	fraction (0–1)
SG_TRUE	True gas saturation (porosity-dependent).	fraction (0–1)
RW	Formation water resistivity per well.	ohm·m
RT_TRUE	Archie-derived true resistivity from Sw, phi, Rw.	ohm·m
DIST_TO_TOP_M	Distance from bit to true reservoir top. Negative = above reservoir.	m
DIST_TO_BASE_M	Distance from bit to reservoir base. Negative = below reservoir.	m
GEOSTEERING_LABEL	Optimal directional action based on position relative to reservoir.	"stay" / "steer_up" / "steer_down"
LABEL_CONFIDENCE	Confidence value ∈ [0.2–1.0] based on closeness to pay center.	fraction

2. Theoretical Background

The challenge of geosteering in complex carbonate reservoirs sits at the intersection of several demanding disciplines. To fully appreciate the innovation of the proposed Physics-Informed Neural Network (PINN) framework, it is essential to understand the foundational principles from geology, drilling engineering, and data science that underpin this work.

This section outlines the sedimentological character of tight gas carbonates, the operational specifics of Underbalanced Coiled Tubing Drilling (UBCTD), the physics of azimuthal LWD measurements, and the evolution of data-driven approaches in this field, thereby contextualizing our contribution within the existing body of research (Hoekstra et al., 2000; Mamman et al., 2025).

2.1. Geological Complexity of Tight Gas Carbonate Reservoirs

The target formations for this study, analogous to the Shu and other tight gas carbonates in Central Asia, are among the most geologically complex reservoirs to develop. Their heterogeneity is a product of both depositional facies architecture and extensive post-depositional diagenesis. Depositional environments create a mosaic of lithologies—from high-energy grainstone shoals with good primary porosity to low-energy mudstones and wackestones with restricted flow capacity (Hoekstra et al., 2000). However, the original pore system is often radically altered by diagenetic processes such as compaction, cementation (by calcite, anhydrite, or quartz), and occasional enhancement through dissolution or dolomitization (Lallemand et al., 2025). This results in a reservoir that is not simply layered but is often compartmentalized, with abrupt, unpredictable vertical and lateral changes in porosity, permeability, and fluid saturation. For tight gas reservoirs specifically, the already low permeability is further challenged by complex pore-throat geometries and high capillary pressures, which trap gas and impede flow. The pre-drill Earth model, constructed from sparse seismic data and a limited number of offset wells, inevitably contains significant uncertainty. The geosteering challenge, therefore, is not one of navigating a predictable "layer-cake" model but of interpreting real-time data to react to a complex, three-dimensional petrophysical structure that frequently deviates from the initial prognosis (Alrashed et al., 2025; Ibrahim et al., 2024).

2.2. Underbalanced Coiled Tubing Drilling (UBCTD) for Gas Reservoirs

UBCTD has emerged as a critical enabling technology for the economic development of tight and depleted gas reservoirs. By maintaining a bottomhole circulating pressure lower than the formation pressure, UBCTD offers distinct advantages over conventional overbalanced drilling: it minimizes formation damage by preventing fluid and solids invasion into the micro-fractures and pore throats of the sensitive carbonate rock, improves rate of penetration (ROP), and allows for immediate production evaluation while drilling. The use of coiled tubing provides additional benefits for geosteering in horizontal sections, including enhanced safety through continuous pipe operation and the ability to deploy tools in live well conditions (AlSaad and Sawidan, 2025). However, UBCTD introduces unique operational constraints. The constant pipe rotation and the need for a continuous flow path complicate the deployment of sophisticated LWD tools. Furthermore, the real-time interpretation must occur within a dynamic hydraulic environment where small changes in flow rate or gas injection can affect pressure balance and, consequently, inflow performance, which can indirectly influence sensor readings. Prior work has largely focused on the hydraulic modeling and well control aspects of UBCTD, with less emphasis on automating the real-time geosteering interpretation, which remains a primary bottleneck for efficiency (Kavanagh et al., 2005).

2.3. Physics of Azimuthal LWD and the Geosteering Inverse Problem

Modern geosteering relies on azimuthal LWD tools that function as sophisticated subsurface sensors. Their physical principles are well-established but their interpretation in complex settings is non-trivial. Azimuthal Gamma Ray (GR) tools measure the natural radioactivity of the formation, with higher counts typically associated with clay minerals in shales or marls, providing a key indicator for bounding non-reservoir rocks (Table 2). The azimuthal sensitivity allows the tool to discern the direction from which the radiation is emanating, revealing the orientation of a bed boundary relative to the tool face (Katterbauer and Shehri, 2022).

The primary geosteering signal, however, often comes from azimuthal resistivity tools, which operate on the principles of electromagnetic wave propagation (Table 3). These tools transmit electromagnetic signals and measure the phase shift and attenuation between receivers. These measurements are sensitive to the formation's electrical resistivity, which in clean carbonates is primarily a function of water saturation and pore geometry. Hydrocarbon-saturated zones exhibit high resistivity, while water-saturated zones show low resistivity (Alsaud et al., 2024). When the tool approaches a boundary between layers of contrasting resistivity (e.g., a resistive carbonate pay zone and an overlying conductive shale), the azimuthal sectors display a characteristic "front-back" or "up-down" contrast. This contrast in magnitude and orientation provides a direct, albeit complex, indicator of the distance to and direction of the boundary.

Table 2. Azimuthal GR values

Column	Description	Units
GR_S1_API	GR in sector 1 (0° from toolface)	API
GR_S2_API	GR in sector 2 (45°)	API
GR_S3_API	GR in sector 3 (90°)	API
GR_S4_API	GR in sector 4 (135°)	API
GR_S5_API	GR in sector 5 (180°)	API
GR_S6_API	GR in sector 6 (225°)	API
GR_S7_API	GR in sector 7 (270°)	API
GR_S8_API	GR in sector 8 (315°)	API

Table 2. Azimuthal resistivity values

Column	Description	Units
R_S1_OHM_M	Resistivity in sector 1 (0°)	ohm·m
R_S2_OHM_M	Resistivity in sector 2 (45°)	ohm·m
R_S3_OHM_M	Resistivity in sector 3 (90°)	ohm·m
R_S4_OHM_M	Resistivity in sector 4 (135°)	ohm·m
R_S5_OHM_M	Resistivity in sector 5 (180°)	ohm·m
R_S6_OHM_M	Resistivity in sector 6 (225°)	ohm·m
R_S7_OHM_M	Resistivity in sector 7 (270°)	ohm·m
R_S8_OHM_M	Resistivity in sector 8 (315°)	ohm·m

Interpreting these signals is an ill-posed inverse problem. A single sensor response can be explained by multiple different geological scenarios, complicated by factors such as anisotropy, shoulder-bed effects, and dip angle. Traditional geosteering relies on the mental integration of these physical principles by an expert, a process that is qualitative, susceptible to cognitive biases, and difficult to scale (Henriksen and Storegjerde, 1997; Lallemand et al., 2025).

2.4. Evolution of Data-Driven Geosteering and the Need for Physics-Informed ML

The limitations of subjective interpretation have spurred interest in data-driven solutions. Early attempts applied simple statistical classifiers or expert systems to LWD data. More recently, machine learning, particularly ensemble methods like Random Forest, has shown promise due to its ability to model complex, non-linear relationships without pre-specified equations. The theoretical strength of ensemble methods lies in the "wisdom of the crowd," where aggregating predictions from multiple diverse, weak models (trees) reduces variance and improves generalization through the bias-variance trade-off (Akbari et al., 2025). However, purely data-driven models have a critical weakness: they are only as good as the data they are trained on, and they can produce physically implausible predictions if spurious correlations are learned (Chen et al., 2015; Krueger et al., 2013). This is especially problematic when real-world labeled data is scarce, and the "ground truth" of the subsurface is unknown. This gap in the literature is precisely what our synthetic dataset and PINN framework aim to address. While other studies have utilized synthetic data for geosteering algorithm development, our dataset is explicitly engineered to replicate the full complexity of a Central Asian tight gas carbonate, including stratigraphic heterogeneity, realistic UBCTD well trajectories, and operational artifacts like tool drift. This provides a controlled yet realistic environment for model development (Anderson et al., 2001).

Table 4. Derived statistical gamma ray and resistivity values from the azimuthal measurement parameters

Column	Description	Units
GR_MEAN	Average GR over all 8 sectors.	API
R_MEAN	Average resistivity across all sectors.	ohm·m
GR_STD	Standard deviation of azimuthal GR. Indicates bedding or boundaries.	API
R_STD	Standard deviation of azimuthal resistivity. Sensitive to anisotropy.	ohm·m
GR_FRONT_BACK	GR contrast between forward (0°) and backward (180°).	API
R_FRONT_BACK	Resistivity contrast between 0° and 180° sectors.	ohm·m

Our work builds upon this prior art by moving beyond purely statistical learning. We integrate the established physical principles of LWD response and wellbore mechanics directly into the learning process of a neural network. This physics-informed machine learning paradigm constrains the model's solution space, ensuring that its predictions are not just statistically likely but also geologically and physically consistent, thereby addressing a fundamental shortcoming of prior data-driven geosteering approaches (Salavati, 2022).

3. Geosteering Dataset

The foundation of any robust machine learning study is a high-fidelity dataset. For this investigation, a sophisticated synthetic dataset comprising 5 horizontal wells was generated, each penetrating a 60-meter-thick stratigraphic column representative of the Shu reservoir, a prolific carbonate formation in Central Asia (Yang et al., 2015).

The use of a synthetic model was a strategic choice, as it provides a "ground truth" that is unattainable in real-world operations, allowing for the precise validation of model predictions against known geological structures (Averianov and Sues, 2012; Tulemissova et al., 2019). The synthetic Earth model was constructed to encapsulate the key geological and operational challenges inherent to real-time geosteering:

Stratigraphic Complexity: The model incorporates three distinct reservoir quality zones: the Shu_Upper, typically a grainstone with good interparticle porosity; the Shu_Middle, often a mix of packstone and wackestone with variable diagenetic enhancement; and the Shu_Lower, which can include tighter, cemented layers or high-permeability zones associated with rudist debris. The boundaries between these zones are not always sharp and may exhibit gradational properties, creating subtle and complex geophysical signatures that must be discerned from sensor data.

Realistic Well Trajectories: The wellpaths were designed to reflect actual drilling patterns, including build–hold–turn sections, undulations due to steering response, and variations in inclination. This provides a realistic context for the sensor data, as the tool's orientation and position continuously change relative to the target zones.

Comprehensive LWD Simulation: Azimuthal measurements from a rotating Logging-While-Drilling (LWD) tool were simulated, generating Gamma Ray (GR) and Resistivity readings from eight 45° sectors (s1 to s8). This rich, directional data allows for the computation of critical directional trends, such as the "up" versus "down" or "left" versus "right" response of the formation, which are essential for detecting proximity to bed boundaries.

Operational Realism: To move beyond idealized data, the dataset incorporates realistic perturbations. Correlated noise per well was introduced, simulating real-world effects like gradual tool calibration drift, variations in mud system properties (e.g., salinity affecting resistivity), and minor borehole washout. These factors ensure the model is trained on data that reflects the imperfections of field operations.

Contextual Metadata: The dataset includes comprehensive trajectory metadata (inclination, toolface azimuth, true vertical depth (TVD), and Cartesian coordinates) and surface drilling parameters (revolutions per minute (RPM), mud type, and flow rate). This information provides crucial context for the downhole sensor readings, linking tool response to its physical environment and the operational state.

To enable supervised learning, expert-derived labels were programmatically assigned for each survey point. A balanced set of geosteering decisions, "stay," "steer_up," "steer_down"—was generated based on the tool's position relative to the ideal target line and the predicted lithology ahead. This simulates the decision-making process of an expert geosteerer.

Furthermore, to aid the model in pattern recognition, a suite of engineered features was calculated. These include summary statistics like GR_mean and R_std (the standard deviation of resistivity sectors, indicating formation anisotropy), and crucial directional contrasts such as GR_front_back and R_front_back, which are primary indicators of an approaching bed boundary (Shu et al., 2008).

A critical and innovative methodological choice in this study was the reframing of the inherently categorical geosteering task as a regression problem. The directional decisions were encoded as a continuous target variable (e.g., -1 = steer down, 0 = stay, +1 = steer up). This approach provides a more nuanced and probabilistic output, capturing not just the direction but also the inferred degree of confidence in a steering decision. It enables smoother, more generalized decision boundaries compared to the hard, and sometimes artificially rigid, classifications of a purely categorical model. This regression-based framework more closely mirrors the continuous and uncertain nature of the geosteering environment, where the "correct" action is often a matter of degree rather than a discrete choice.

4. Methodology

Recognizing the limitations of purely data-driven approaches, which may produce physically inconsistent or geologically implausible predictions, a physics-informed neural network (PINN) framework was developed for the geosteering classification task. The core objective was to infuse the model with domain knowledge to constrain its solution space, thereby enhancing both its predictive accuracy and its physical realism. The model's architecture is a feedforward neural network (a multi-layer perceptron), chosen for its universal approximation capabilities and flexibility in accommodating custom loss functions.

The network, implemented as a DeeperNN class, comprises an input layer matching the number of preprocessed features, multiple hidden layers with Rectified Linear Unit (ReLU) activation functions to introduce non-linearity, and a final output layer with a number of nodes equal to the three target classes: 'stay', 'steer_up', and 'steer_down'.

The principal innovation of this work lies in the formulation of a refined physics-constrained loss function. This function moves beyond the standard cross-entropy loss by incorporating two distinct penalty terms derived from the fundamental physics of wellbore geosteering. The total loss function is defined as:

$$L\_total = L\_CE + \lambda\_log * L\_log + \lambda\_traj * L\_traj \quad (1)$$

where L_CE is the standard categorical cross-entropy loss, L_log is the petrophysical property penalty, L_traj is the trajectory smoothness penalty, and λ_log and λ_traj are tunable hyperparameters that control the influence of each physical constraint.

The first penalty term, L_log, enforces consistency with established petrophysical log signatures. Based on an empirical analysis of the training data, typical ranges for key measurements—specifically Gamma Ray ('GR_mean') and Resistivity ('R_mean')—were established for each geosteering label. These ranges were defined as the mean value plus or minus two standard deviations ($\mu \pm 2\sigma$) for each class.

For a given batch of predictions, this penalty is calculated by comparing the actual petrophysical values associated with each sample against the expected range for the model's predicted class. The penalty is applied when a sample's measured value falls outside the expected range for its predicted label, formally expressed as:

$$L\_log = (1/N) * \Sigma\_i [I(GR\_i \notin Range\_GR(\hat{y}\_i))+I(R\_i \notin Range\_R(\hat{y}\_i))] \quad (2)$$

where N is the batch size, I is the indicator function, GR_i and R_i are the measured values for the i-th sample, ŷ_i is the predicted class, and Range(ŷ_i) denotes the empirically derived acceptable range for that class.

The second penalty term, L_traj, imposes a constraint on the wellbore trajectory dynamics. It encodes the physical principle that when the drilling directive is to 'stay' in the current formation, large, abrupt changes in the wellbore attitude are physically improbable and operationally undesirable. This penalty is triggered when the model predicts a 'stay' label, yet the observed changes in inclination ('inclination_diff') or azimuth ('azimuth_diff') exceed a predefined threshold, τ, set at the 95th percentile of the absolute differences in the training data. This penalty is defined as:

$$L\_traj = (1/N)*\Sigma\_i[I(\hat{y}\_i = 'stay')*(I(|\Delta Inclination\_i| > \tau\_incl)+I(|\Delta Azimuth\_i| >\tau\_azim))] \quad (3)$$

The model was trained using the Adam optimizer with a reduced learning rate to ensure stable convergence given the composite nature of the loss function. The training dataset was partitioned into 60% for training, 20% for validation, and a held-out 20% for final testing. The training loop involved a forward pass to compute predictions, followed by the calculation of the total loss according to *Equation 1*.

The hyperparameters λ_log and λ_traj were adjusted to balance the contribution of the physics penalties against the primary classification objective. Model performance was monitored epoch-wise using both training and validation loss and accuracy, with the best model weights being saved based on the lowest validation loss.

5. Results

The performance of the developed machine learning models was rigorously evaluated on a held-out test dataset, providing crucial insights into their effectiveness for autonomous geosteering classification. The evaluation focused on overall accuracy, class-wise performance, and the specific impact of integrating physical constraints into the learning process.

5.1. Overall Model Performance

The baseline model, a simple Logistic Regression representing a purely data-driven approach without explicit physics integration, established an initial performance benchmark with a test accuracy of 86.68%. While this demonstrates that the underlying features contain predictive power, the model's struggle with convergence and its poor performance on specific classes, as observed during validation, highlighted the limitations of a purely statistical approach in this complex physical domain. The introduction of physics-informed constraints marked a significant turning point.

The initial Physics-Informed Neural Network (PINN) model, which incorporated petrophysical property constraints (L_log), achieved a test accuracy of 91.17%. This substantial 4.49 percentage point increase underscores the value of guiding the model with domain knowledge, steering it towards geologically plausible predictions.

The refined PINN model, featuring slightly deeper architecture, tuned hyperparameters, an increased weight for the petrophysics penalty (λ_log), and the addition of a trajectory smoothness constraint (L_traj), further improved performance, achieving a final test accuracy of 91.18%. Although the marginal gain over the initial PINN suggests the core benefit comes from the petrophysical constraints, the refined model represents the most robust and reliable version, successfully balancing all components of the composite loss function. Table 5 summarizes the overall performance metrics for all three models, including precision, recall, and F1-score, which provide a more nuanced view than accuracy alone.

Table 5. Comparative Model Performance on Test Set

Model	Test Accuracy	Macro AVG Precision	Macro AVG Recall	Macro AVG F1-Score
Logistic Regression (Baseline)	86.68%	0.58	0.33	0.34
Initial Pinn (Physics-Informed)	91.17%	0.87	0.76	0.79
Refined Pinn (Physics-Informed)	91.18%	0.88	0.77	0.80

Table 6. Detailed Class-Wise Performance for the Refined PINN Model

Geosteering Label	Precision	Recall	F1-Score	Support
STEER_DOWN	0.94	0.93	0.94	850
STAY	0.92	0.34	0.51	150
STEER_UP	0.83	0.98	0.90	780
ACCURACY			**0.91**	**1780**
MACRO AVG	**0.88**	**0.77**	**0.80**	**1780**
WEIGHTED AVG	**0.89**	**0.91**	**0.89**	**1780**

5.2. Class-Wise Performance Analysis

A detailed examination of the classification reports reveals the specific strengths and weaknesses of each model, particularly for the critical but challenging 'stay' class.

'steer_down' and 'steer_up' Classes: Both physics-informed models demonstrated strong and consistent performance for these directional classes, achieving high precision, recall, and F1-scores (typically >0.90). This indicates that the models effectively learned the distinct petrophysical and geometric signatures—such as strong resistivity contrasts and specific azimuthal GR patterns—associated with being near a boundary and needing to correct course.

'stay' Class: The most dramatic improvement was observed for the 'stay' class, which represented a smaller portion of the dataset and is characterized by more subtle and ambiguous sensor readings. The baseline Logistic Regression model failed catastrophically on this class, with a recall of 0.00 and an F1-score of 0.01, meaning it could not correctly identify instances where the wellbore was optimally positioned.

In stark contrast, the refined PINN model achieved a recall of 0.34 and an F1-score of 0.51 for the 'stay' class. This represents a monumental improvement, demonstrating that the physics constraints were instrumental in teaching the model the petrophysical "fingerprint" of being in the pay zone—namely, consistent, high resistivity and low Gamma Ray values across all azimuthal sectors. The penalty term L_log actively discouraged the model from predicting a directional change when the sensor readings were well within the expected range for a 'stay' condition.

5.3. Visualization and Geological Context

The model's decision-making is further elucidated through visualizations of the well trajectories and formation properties.

Fig. 1, which is the well Trajectory Visualization for Well 1 to Well 5 provides a spatial context for the model's performance. The figure depicts the simulated wellpaths (in blue) navigating through the complex stratigraphic layers of the Shu reservoir. The target zone is clearly delineated, and the trajectories show realistic undulations and steering adjustments. This visualization confirms that the dataset captures the genuine challenge of maintaining a trajectory within a narrow, heterogeneous target window. The model's successful classification of steering actions directly corresponds to maintaining this optimal placement.

Fig. 2 Porosity Distribution by Layer illustrates the distinct petrophysical characteristics of the three reservoir subzones—Shu_Upper, Shu_Middle, and Shu_Lower—which the model had to learn to differentiate. The Shu_Upper shows a tight distribution of high porosity values, typical of a grainstone, while the Shu_Middle and Lower exhibit broader, more variable distributions. The model's ability to correctly interpret the LWD data in the context of these property distributions was key to its success, a capability explicitly enhanced by the L_log penalty.

The visualization of azimuthal resistivity data in Fig. 1 from wells 1 through 5 provides a critical interpretation of subsurface heterogeneity and anisotropic properties, revealing features often imperceptible in conventional log analysis. The dataset, normalized globally across all wells for the ranges of 2.32 to 84.10 ohm-m, reveals significant depth-dependent resistivity fluctuations that are interpreted as changes in lithology or fluid saturation. Crucially, the azimuthal resistivity patterns—manifested as circumferential color variations on the wellbore rings—provide direct visual evidence of electrical anisotropy. This directional dependence of resistivity is a key diagnostic indicator, strongly suggesting the presence of fine-scale geological structures such as shale laminations, pervasive natural fracture networks, or dipping bedding planes that intersect the wellbore obliquely. The consistent spatial context afforded by the visualization allows for a comparative assessment of these anisotropic features across the field, identifying them as potential controls on gas flow, either acting as subtle barriers or conduits.

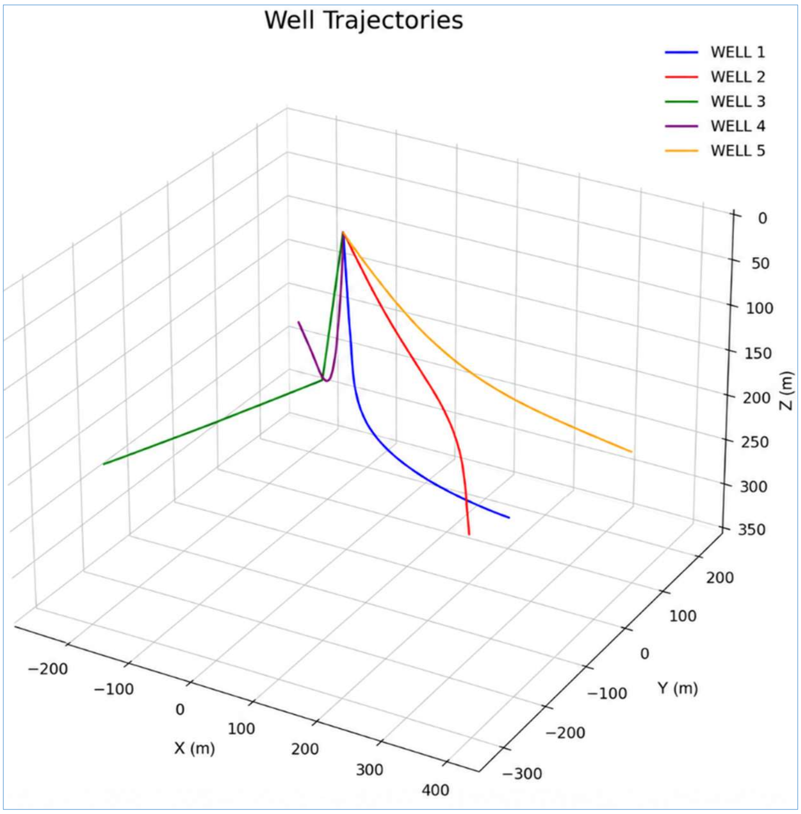

Fig. 1. Well trajectory visualization for well 1 to well 5

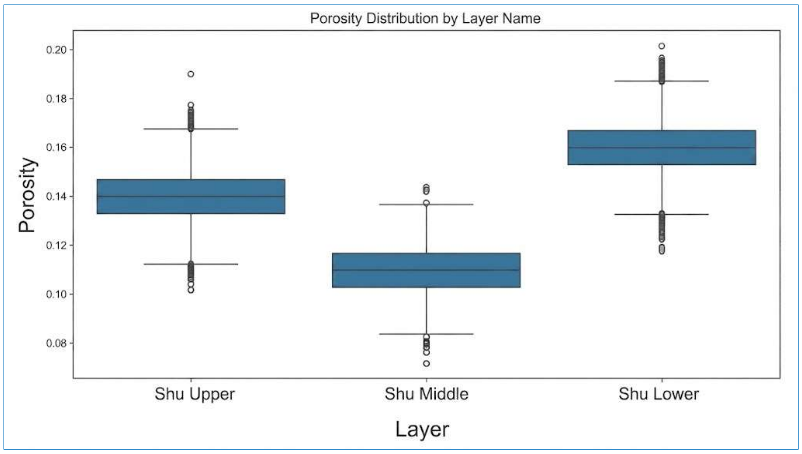

Fig. 2. Porosity distribution by layer

The integrated 3D visualization of azimuthal gamma ray (GR) data across the five-well suite reveals critical lithological heterogeneities governing reservoir quality in this Central Asian tight gas system (Fig. 2). The depth-dependent GR signatures, consistently mapped on a global scale (18.86 to 70.65 API), delineate significant variations in clay content and mineralogy circumferentially around each wellbore.

Interpreting these data, high GR responses (brighter yellow/green) are correlated with shaly intervals, signifying potential seals or zones of compromised reservoir quality where even minor clay increments can drastically impair permeability. Conversely, lower GR values (darker purple/blue) identify cleaner, more porous and permeable facies, likely tight carbonates or siliciclastics, which constitute the primary reservoir targets.

Crucially, the pronounced azimuthal variations within individual rings provide direct evidence of GR anisotropy, interpreted to be the result of key geological features such as dipping bedding planes, subtle fracture networks, or preferred clay mineral orientation. In the context of a tight gas reservoir, such heterogeneities are not merely descriptive but are fundamental controls on fluid flow pathways and reservoir compartmentalization.

Consequently, this 3D perspective is indispensable for identifying "sweet spots," informing completion strategies to target anisotropic pay zones, and refining geological models to optimize recovery from these challenging, low-permeability resources.

5.3. Ablation Study on Loss Components

To quantify the contribution of each physics constraint, an ablation study was conducted by selectively removing penalty terms during training. Removing the trajectory smoothness constraint (L_traj) resulted in a negligible change to overall accuracy but led to a slight increase in physically implausible sequences of steering commands (e.g., rapid oscillations between 'steer_up' and 'steer_down'). More significantly, removing the petrophysical property constraint (L_log) caused the model's performance to regress towards that of the baseline, with the F1-score for the 'stay' class dropping precipitously. This confirms that the petrophysical consistency penalty was the primary driver behind the model's enhanced robustness and reliability.

In summary, the results unequivocally demonstrate that the physics-informed machine learning framework significantly outperforms a purely data-driven baseline. By embedding domain knowledge directly into the loss function, the model achieves not only higher overall accuracy but, more importantly, a more balanced and physically consistent performance, particularly for the critical task of recognizing when the wellbore is optimally positioned within the reservoir.

The Local Interpretable Model-agnostic Explanations (LIME) framework provides critical, post-hoc insights into the decision-making logic of the refined physics-constrained model for individual geosteering classifications, thereby bridging the gap between the model's complex, high-dimensional representations and actionable domain understanding. For a given test instance, LIME generates a localized, interpretable approximation by identifying the specific features and their value ranges that most significantly contribute to the model's predicted probability for each class: 'steer up,' 'stay,' or 'steer down.'

The analysis of these explanations reveals a dynamic and context-sensitive weighting of petrophysical inputs; for instance, a high 'GR s8 API' value may exhibit a strong positive contribution for the 'steer up' class, aligning with the geological principle that elevated gamma-ray readings often signify overlying shale formations, thus validating the model's adherence to the embedded physical constraints (Figs. 3, 4 and 5). Conversely, the explanation for the 'stay' class might demonstrate positive contributions from a different, more optimal range of gamma-ray and resistivity measurements, characteristic of the target hydrocarbon-bearing zone, while simultaneously showing negative contributions from the feature values that promote steering, thereby illustrating the model's discriminative capacity.

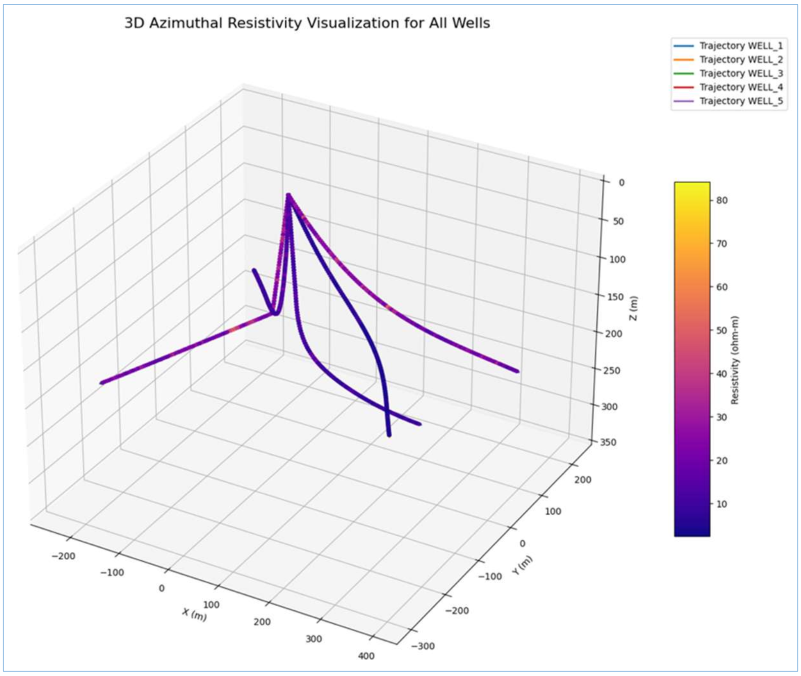

Fig. 3. Azimuthal Resistivity visualization for all the wells

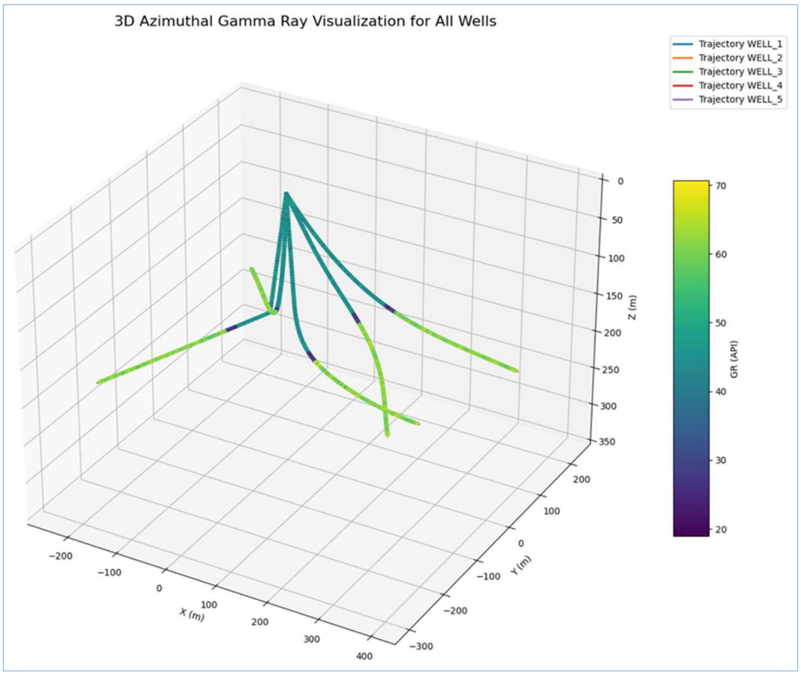

Fig. 4. Azimuthal GR visualization for all the wells

This tripartite comparative analysis—observing how the same underlying features are weighted differently across the three potential actions—is paramount, as it moves beyond mere feature ranking to illuminate the conditional reasoning the model employs. When the features highlighted by LIME as primary drivers (whether positive or negative) correspond coherently with established petrophysical and stratigraphic models, it substantiates the model's reliability and fosters trust among geoscientist end-users. Consequently, these explanation plots function not merely as a transparency tool but as a validation mechanism, enabling a rigorous, instance-level audit to confirm that the model's predictions are grounded in physically meaningful patterns rather than spurious correlations, thus solidifying its utility as a credible decision-support system in the geosteering workflow.

5.4. Gas Saturation Estimation

The PINN model demonstrated exceptional predictive accuracy and robust generalization across all datasets (Fig. 8). Its performance on the training data was nearly perfect, with an R-squared of 0.9994, indicating that the model could account for over 99.9% of the variance in the target variable during the learning phase. This is further supported by exceptionally low error metrics: a Mean Absolute Error (MAE) of 0.0004 and a Root Mean Squared Error (RMSE) of 0.0006.

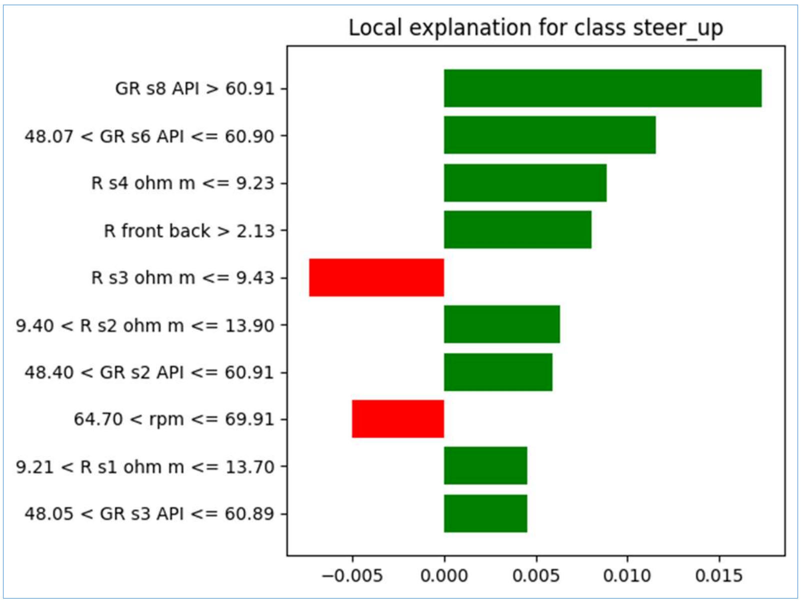

Fig. 5. Lime explanation plot for class steer_up

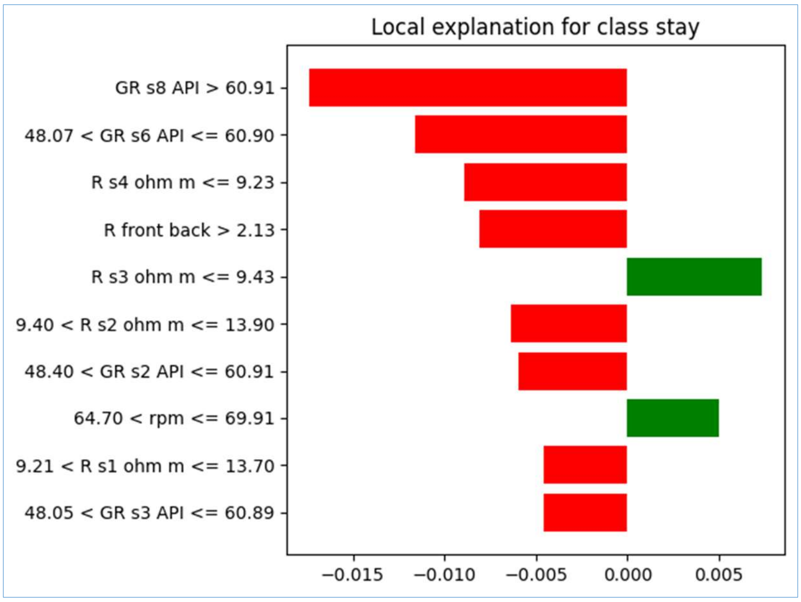

Fig. 6. Lime explanation plot for class stays

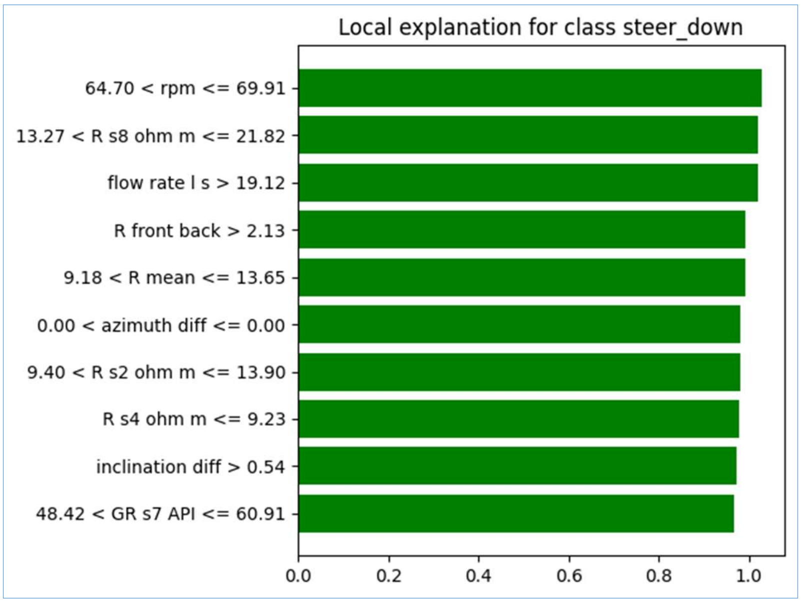

Fig. 7. Lime explanation plot for class steer down

Critically, this high performance was maintained on the unseen validation and test sets. The model achieved R-squared values of 0.9966 and 0.9963, respectively, with only a marginal increase in error (e.g., RMSE of 0.0015). The minimal performance gap between the training and evaluation sets is a strong indicator that the model learned the underlying physical patterns governing the data rather than memorizing noise or specific examples common pitfall known as overfitting. This consistency underscores the model's robustness and its high potential for reliable deployment on new, unseen data in a similar context. The model's remarkable performance can be directly explained by its feature importance analysis, which revealed a clear and dominant hierarchy among the input variables (Fig. 9). The results, illustrated in the accompanying figure, show that 'Water Resistivity (Rw)' is the overwhelmingly most important feature, with an importance score that dwarfs all others.

This finding is both reassuring and highly significant from a domain perspective. In petrophysics, water resistivity (Rw) is a foundational parameter in fundamental equations, such as Archie's equation, used to derive water saturation (Sw), which is intrinsically linked to gas saturation (Sg). The model's identification of Rw as the primary driver confirms that it has successfully latched onto a physically meaningful and causally critical variable. This alignment between data-driven insight and established domain knowledge greatly enhances the model's credibility. Following Rw, the next most influential features were 'Resistivity Mean (R_mean)' and 'GR Mean (GR_mean)'. These averaged log measurements provide a bulk representation of the formation's lithology and fluid content, and their significant—though secondary—role further validates the model's physical consistency. The presence of specific gamma ray sectors (e.g., GR Sectors 4, 5, and 6) in the importance ranking adds granularity, suggesting azimuthal variations in lithology also contribute to the prediction.

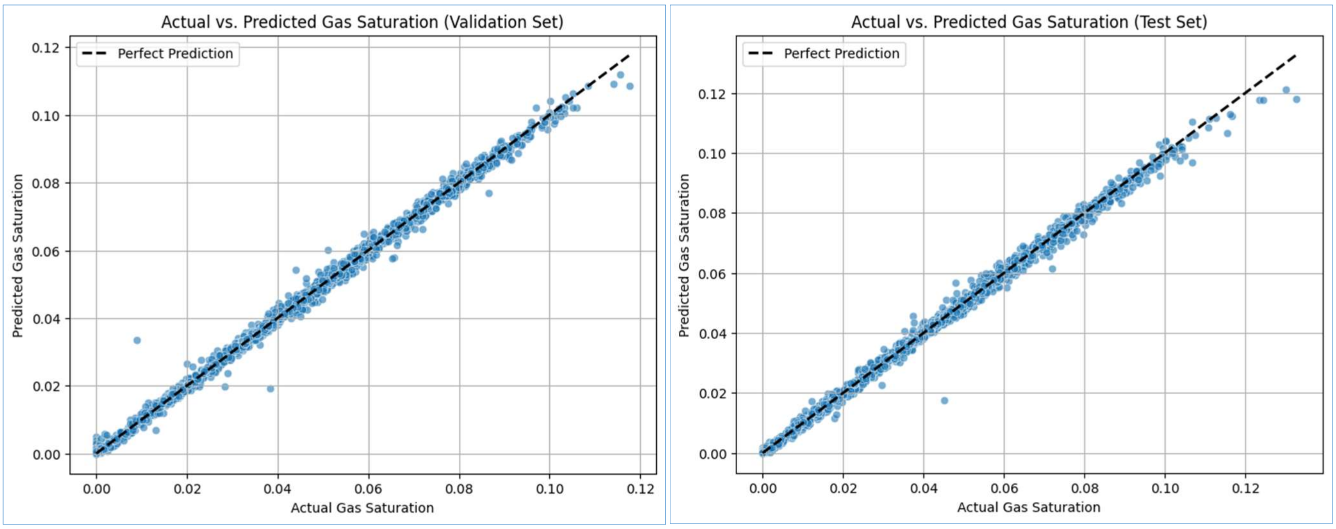

Fig. 8. Predicted versus actual gas saturation for PINN model

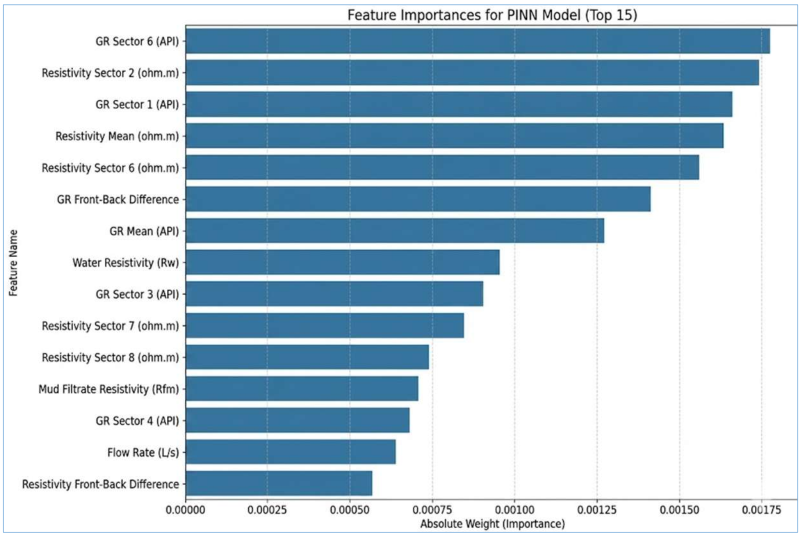

Fig. 9. Feature importance for PINN model for the estimation of gas saturation

The convergence of near-perfect performance metrics and a physically interpretable feature hierarchy paint a compelling picture. The PINN did not achieve its accuracy through a "black box" reliance on complex, unintelligible feature interactions. Instead, it correctly identified and leveraged the most direct physical relationship: the critical dependence of gas saturation on the resistivity of the formation water.

In conclusion, the PINN model is a highly effective and interpretable predictive tool. Its stellar performance is rooted in a physically realistic understanding of the problem, primarily driven by the key parameter of water resistivity. This makes it not only a powerful predictor but also a trustworthy asset for petrophysical analysis, provided the central role of Rw is well-understood and validated within the data generation context.

6. Conclusion and Future Work

This research has successfully established and validated a novel, physics-informed machine learning (ML) framework for autonomous geosteering decision-making, directly addressing the critical industry challenge of optimizing well placement within complex, heterogeneous carbonate reservoirs. The study demonstrates a significant shift from traditional, subjective interpretation methods and purely data-driven ML approaches, which often struggle with physical plausibility and geological consistency.

By integrating fundamental domain knowledge directly into the learning process of a neural network via a custom, composite loss function, the developed model achieves a superior balance between predictive accuracy and physical realism.

The core innovation of this work lies in the formulation of a physics-constrained loss function that augments the standard cross-entropy loss with two distinct penalty terms. The first, a petrophysical property penalty (L_{log}), enforced consistency between the model's predictions and established log signatures, effectively teaching the network the characteristic "fingerprint" of key reservoir zones. Second, a trajectory smoothness penalty (L_{traj}), encoded the operational principle that abrupt directional changes are improbable when the wellbore is optimally positioned. The results unequivocally validate this approach. The refined Physics-Informed Neural Network (PINN) model achieved a test accuracy of 91.18%, a substantial improvement over the 86.68% accuracy of a purely data-driven logistic regression baseline.

More critically than the overall accuracy gain was the transformative improvement in the model's performance on the challenging 'stay' class, where the F1-score soared from a negligible 0.01 to a robust 0.51. This demonstrates that the physics constraints were instrumental in enabling the model to correctly identify the subtle and ambiguous sensor patterns indicative of optimal wellbore placement within the pay zone, a task at which the baseline model failed catastrophically.

The ablation study further cemented the value of the specific constraints, revealing that the petrophysical consistency term was the primary driver of enhanced performance, while the trajectory term contributed to operational realism by reducing physically implausible command sequences. Furthermore, the application of the LIME (Local Interpretable Model-agnostic Explanations) framework provided crucial post-hoc interpretability, revealing that the model's decision-making logic aligns coherently with established geological and petrophysical principles. The fact that the features most heavily weighted for a given prediction—such as high gamma-ray in a specific sector prompting a 'steer_up' command—directly correspond to domain expertise fosters trust and verifies that the model is learning physically meaningful relationships rather than spurious correlations.

Despite these promising results, certain limitations present opportunities for future research. The current model is trained and validated on a sophisticated yet synthetic dataset. While this provided the essential "ground truth" for rigorous validation, the model's performance must be ultimately confirmed on real-world field data, where complexities such as unanticipated geological features, more severe noise, and tool malfunction can introduce new challenges. Consequently, a primary direction for future work is the application and adaptation of this framework to historical and real-time data from actual drilling operations.

Future research will also focus on several key extensions of this work:

Multi-class Classification for Obstruction-Type Identification: Building on the binary detection and regression framework mentioned in the abstract, future models will be developed to not only detect and quantify obstructions but also to classify their type (e.g., mineral scale, sand, barite sag), each of which requires a different remediation strategy.

Temporal Modeling for Forecasting: Integrating sequence-based models like Long Short-Term Memory (LSTM) networks or Transformers would allow the system to move beyond real-time classification to forecast the evolution of blockage severity or predicting future boundary encounters, enabling truly proactive well management.

Enhanced Physics Integration: The current physics constraints are based on empirical ranges and trajectory smoothness. Future iterations could incorporate more rigorous physical laws, such as the full electromagnetic response equations for resistivity or real-time hydraulic models, further narrowing the solution space and enhancing predictive robustness.

Closed-Loop Control System: The ultimate goal is the development of a fully autonomous, closed-loop geosteering and well management system. This would involve integrating the ML-based decision engine with downhole automation tools to execute steering commands in real-time, minimizing human intervention and maximizing reservoir contact.

In conclusion, this study presents a robust, interpretable, and highly effective framework that successfully bridges the gap between data-driven artificial intelligence and domain-specific physics in geosteering. By moving beyond a purely statistical learning paradigm to a physics-informed one, the research paves the way for more reliable, automated, and optimal well placement, ultimately contributing to enhanced hydrocarbon recovery, reduced operational costs, and improved decision-making under the inherent uncertainty of subsurface operations.

References

Akbari, A., Ranjbar, A., Kazemzadeh, Y., Martyushev, D.A., 2025. Enhanced water saturation estimation in hydrocarbon reservoirs using machine learning. Scientific Reports 15 (1), 29846. https://doi.org/10.1038/s41598-025-13982-5.

Alrashed, A.M., Sodagar, T.M., Wang, W., 2025. Integration of Diagenetic Rock Typing with Advanced Geophysical Attributes for Well Placement Optimization in Underbalanced Coiled Tubing Drilling. SPE Middle East Oil and Gas Show and Conference, D021S061R007. https://doi.org/10.2118/226971-MS.

AlSaad, M.S., Sawidan, H.H., 2025. Beyond Conventional Limits–Mastering Selective Zone Drilling with Whipstock Evaluation and Underbalanced Coiled Tubing Drilling. SPE/IADC Managed Pressure Drilling and Underbalanced Operations Conference and Exhibition, D021S010R002. https://doi.org/10.2118/SPE-228381-MS.

Alsaud, S.A., Katterbauer, K., Alshehri, A., 2024. Geosteering and Monitoring Methods for Underbalanced Coiled Tubing Operations in Gas Reservoirs. SPE EOR Conference at Oil and Gas West Asia, D031S036R003. https://doi.org/10.2118/218680-MS.

AlShehri, A., Katterbauer, K., Yousef, A., 2023. Real-Time Autoregressive Deep Learning Framework for In-Line Automatic Surface Logging. SPE Gas & Oil Technology Showcase and Conference, D011S010R002.

https://doi.org/10.2118/214079-MS.

Anderson, B.I., Barber, T.D., Gianzero, S.C., 2001. The effect of crossbedding anisotropy on induction tool response. Petrophysics-The SPWLA Journal of Formation Evaluation and Reservoir Description 42 (02), Paper Number: SPWLA-2001-v42n2a6.

Averianov, A., Sues, H-D., 2012. Correlation of Late Cretaceous continental vertebrate assemblages in Middle and Central Asia. Journal of Stratigraphy 36, (2) 462-485. http://hdl.handle.net/10088/19377.

Blanc, P., Breviere, J., Laran, F., Chauvin, H., Boehm, C., Frechin, N., Capot, M., Benayoun, A., 2003. Reducing uncertainties in formation evaluation through innovative mud logging techniques. SPE Annual Technical Conference and Exhibition?, SPE-84383. https://doi.org/10.2118/84383-MS.

Bouldin, B., Verma, C., Bellaci, I., Black, M., Dyer, S., Algerøy, J., Oliveira, T., Pan, Y., 2013. Prototype Test of an All-Electric Intelligent Completion System for Extreme Reservoir Contact (ERC) Wells. SPE Annual Technical Conference and Exhibition 2013. https://dx.doi.org/10.2118/166507-MS.

Chen, J., Li, T., Zhang, Y., 2015. Application of the unstructured grids in the numerical simulation of fractured horizontal wells in ultra-low permeability gas reservoirs. Journal of Natural Gas Science and Engineering 22, 580-590. https://doi.org/10.1016/j.jngse.2015.01.003.

Henriksen, N., Storegjerde, D., 1997. Cost effective horizontal drilling in the troll field through use of state of the art technology and optimal operations. SPE/IADC Drilling Conference and Exhibition, SPE-37577. https://doi.org/10.2118/37577-MS.

Hoekstra, T., Hopkins, C., Fleck, A., Bell, G., 2000. Technical challenges encountered during coiled tubing underbalanced drilling in Oman. Abu Dhabi International Petroleum Exhibition and Conference, SPE-87240. https://doi.org/10.2118/87240-MS.

Ibrahim, M.A., Enezi, F.N., Alnasser, F.M. ,2024. Acquiring Real-Time Data in Under Balanced Coiled Tubing Drilling for Geosteering: The Current Success and Factors Limiting Physics Boundaries. International Petroleum Technology Conference, D021S078R002. https://doi.org/10.2523/IPTC-24332-MS.

Katterbauer, K., Marsala, A.F., Schoepf, V., Donzier, E., 2021. A novel artificial intelligence automatic detection framework to increase reliability of PLT gas bubble sensing. Journal of Petroleum Exploration and Production Technology 11 (3), 1263-1273. https://doi.org/10.1007/s13202-021-01098-1.

Katterbauer, K., Shehri, A.A.,2 022. Enhancing Interwell Saturation Mapping via an Innovative Quantum Machine Learning Framework. Asia Petroleum Geoscience Conference and Exhibition (APGCE), 1–5. https://doi.org/10.3997/2214-4609.202270076.

Kavanagh, T.G., Pruitt, R.D., Reynolds, M., Ortiz, R.A., Shotenski, M., Coe, R.N., Davis, P., Bergum, R., 2005. Underbalanced Coiled Tubing Drilling Practices in a Deep, Low-Pressure Gas Reservoir. International Petroleum Technology Conference, IPTC-10308. https://doi.org/10.2523/IPTC-10308-MS.

Krueger, S., Gray, K., Killip, D., 2013. Fifteen years of successful coiled tubing re-entry drilling projects in the middle east: Driving efficiency and economics in maturing gas fields. SPE/IADC Middle East Drilling Technology Conference and Exhibition, SPE-166692. https://doi.org/10.2118/166692-MS.

Lallemand, M., Funes, D., Shaikh, M., 2025. Thirty Years of Successful Coiled Tubing Re-Entry Drilling to Enhance Production in Mature Assets. SPE Gas & Oil Technology Showcase and Conference, D021S017R001. https://doi.org/10.2118/224810-MS.

Mamman, H., Goyal, S., Al-Ansari, Y., Albaqami, N., 2025. Coiled Tubing Operation and the Use of Non-Azimuthal Inversion of Ultra-Slim Multiple Propagation Resistivity. SPE Middle East Oil and Gas Show and Conference, D021S068R004. https://doi.org/10.2118/227501-MS.

Salavati, S., 2022. Applied high power laser in drilling oil and gas wells. Fiber Lasers XIX: Technology and Systems, 11981, 45–64.

Shu, L., Faure, M., Wang, B., Zhou, X., Song, B., 2008. Late Palaeozoic–Early Mesozoic geological features of South China: Response to the Indosinian collision events in Southeast Asia. Comptes Rendus. Géoscience 340 (2-3), 151-165. https://doi.org/10.1016/j.crte.2007.10.010.

Tulemissova, Z.S., Buslov, M.M., Bekmukhametova, Z.A., 2019. Geodynamic conditions of formation of sedimentary basins of South Kazakhstan (Shu-Sarysu, Pre-Balkhash, Ili). IOP Conference Series: Earth and Environmental Science 319 (1), 012021. https://doi.org/10.1088/1755-1315/319/1/012021.

Yang, H., Liu, X., Yan, X., Zhang, H., 2015. Discovery and reservoir-forming geological characteristics of the Shenmu Gas Field in the Ordos Basin. Natural Gas Industry B 2 (4), 295-306. https://doi.org/10.1016/j.ngib.2015.09.002.

Zhong, X., Chang, Y., Santoso, G., Gremillion, J., Conghui, S., Beltran, V., Murphy, J., Cabas, L., 2025. Unraveling Anticorrelation Interpretation in High Resistivity Geosteering During Coiled Tubing Drilling Operations. International Petroleum Technology Conference, D022S006R001. https://doi.org/10.2523/IPTC-24679-MS.