

Article

Machine Learning-Based Prediction of Well Logs Guided by Rock Physics and Its Interpretation

Ji Zhang , Guiping Liu, Zhen Wei, Shengge Li, Yeheya Zayier and Yuanfeng Cheng *

School of Geology and Mining Engineering, Xinjiang University, Urumqi 830046, China; 18799632186@163.com (J.Z.); ping0991@163.com (G.L.); weizhen@xju.edu.cn (Z.W.); shenggeli0428@163.com (S.L.); yeheya0926@163.com (Y.Z.)

* Correspondence: chengyuanfeng@aliyun.com

Abstract: The refinement of acquired well logs has traditionally relied on predefined rock physics models, albeit with their inherent limitations and assumptions. As an alternative, effective yet less explicit machine learning (ML) techniques have emerged. The integration of these two methodologies presents a promising new avenue. In our study, we used four ML algorithms: Random Forests (RF), Gradient Boosting Decision Trees (GBDT), Multi-layer Perceptrons (MLP), and Linear Regression (LR), to predict porosity and clay volume fraction from well logs. Throughout the entire workflow, from feature engineering to outcome interpretation, our predictions are guided by rock physics principles, particularly the Gardner relations and the Larionov relations. Remarkably, while the predictions themselves are satisfactory, SHapley Additive exPlanations (SHAP) analysis uncovers consistent patterns across the four algorithms, irrespective of their distinct underlying structures. By juxtaposing the SHAP explanations with rock physics concepts, we discover that all four algorithms align closely with rock physics principles, adhering to its cause–effect relationships. Nonetheless, even after intentionally excluding crucial controlling input features that would inherently compromise prediction accuracy, all four ML algorithms and the SHAP analysis continue to operate, albeit in a manner that seems irrational and starkly contradicts the fundamental principles of rock physics. This integration strategy facilitates a transition from solely mathematical explanations to a more philosophical interpretation of ML-based predictions, effectively dismantling the traditional black box nature of these ML models.



Academic Editor: Richard J. Povinelli

Received: 27 November 2024

Revised: 15 January 2025

Accepted: 27 January 2025

Published: 30 January 2025

Citation: Zhang, J.; Liu, G.; Wei, Z.; Li, S.; Zayier, Y.; Cheng, Y. Machine Learning-Based Prediction of Well Logs Guided by Rock Physics and Its Interpretation. *Sensors* **2025**, *25*, 836. <https://doi.org/10.3390/s25030836>

Copyright: © 2025 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 (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: machine learning; well logs; rock physics; SHapley Additive exPlanations (SHAP); explainable artificial intelligence (XAI)

1. Introduction

Well logs are capable of capturing a diverse array of geophysical parameters within the penetrated formations, encompassing caliper measurements, acoustic velocity, gamma-ray radioactivity, neutron density, spontaneous potential, resistivity, nuclear magnetic resonance, and more. These vast well logging datasets are extensively employed in comprehensively analyzing the physical attributes of the target strata, offering indispensable insights into the geological composition and potential of hydrocarbon reservoirs [1,2].

Rock physics models are typically employed in the refinement of initially acquired well logs, aimed at correcting inaccuracies or predicting missing data segments [3,4]. These models bridge the gap between macroscopic measured seismic properties and microscopic rock features, such as porosity, resistivity, mineralogy, fluid saturation, and pore pressure [5–7]. These models can be broadly categorized into two types: empirical and

physically based [8]. Physically grounded models, such as Gassmann's equation, often hinge on certain assumptions, thus restricting their scope of application [9–12]. In contrast, empirical models, exemplified here by the Gardner relations and the Larionov relations, are founded on experimental data and stem from straightforward regression analyses [13–17]. Despite their simplicity, these empirical models have demonstrated remarkable practicality and widespread applicability in the petroleum industry [18,19].

Alternatively, machine learning (ML) algorithms offer a solution for correcting errors or filling in gaps in well logs [20,21]. In recent years, the application of ML technology has experienced a dramatic surge in popularity and adoption across the logging industry, marking a significant leap in technological progress [22–25]. While rock physics relies on explicit knowledge gained through extensive experience and experimentation, ML algorithms depend on data training [26,27]. This often renders ML a black box with limited interpretability, significantly reducing its perceived enhanced value and credibility within the industry [28–31]. Explainable artificial intelligence (XAI) serves as an umbrella term, encompassing the incorporation of human expertise into ML workflows [32–35]. This integration spans a wide spectrum, from precisely defined equations (PINN) where human knowledge is explicit [36–38], to more implicit and less readily articulated experience that individuals possess [39,40].

In this study, we propose to integrate the SHapley Additive exPlanations (SHAP) framework and rock physics into the workflow for predicting well logs, leveraging four distinct ML algorithms. This approach strives to dismantle the traditional opacity of ML models, thereby fostering unprecedented transparency. Specifically, SHAP allows us to quantify, on a localized level, the precise impact of each input feature on individual predictions, while also presenting a holistic, global overview across the entire dataset [41–43]. Furthermore, rock physics serves as a pivotal guide, quality guardian, and reference point throughout the ML-based workflow, spanning from feature engineering to results interpretation. By leveraging this integration, we aim to not only bolster the accuracy but also elevate the interpretability of ML-based well log predictions, thus propelling the field forward towards more reliable and intelligible outcomes.

2. Methodology

2.1. Rock Physics

To describe the dependency between bulk density and compressional velocity, a set of relations were formulated as [16]

$$\rho_b = aV_p^b \quad (1)$$

where ρ_b is bulk density in g cm^{-3} , V_p is the P-wave velocity in ft s^{-1} , and a and b are lithology-specific constant coefficients that vary across different rock types.

In this study, the Gardner relations were initially used to calibrate the correlation between acoustic logs and density logs. Subsequently, these relations were used to generate density logs in cases of suspicion or inaccuracies. It is noteworthy that, besides acoustic data, density can also correlate with shear wave velocity, neutron porosity, or gamma ray [44]. Therefore, in addition to the Gardner relations, a machine learning-driven approach can be implemented to generate density logs from a diverse range of inputs.

To describe the clay volume fraction in a rock, the gamma index, I_c , was introduced [14], which is simply a normalized representation of the recorded nature gamma logs. Specifically,

$$I_c = \frac{GR - GR_{\min}}{GR_{\max} - GR_{\min}} \quad (2)$$

where GR is the measured value of gamma logs expressed in American Petroleum Institute (API) units, and GR_{\max} and GR_{\min} are the respective maximum and minimum

values encountered in shales and sands, with these thresholds herein specified as 140 and 20 API units.

Since the gamma index tends to exceed the actual clay volume fraction, V_c , alternative modified equations are often employed [14]:

$$V_c = 0.083 \left(2^{3.7I_c} - 1 \right) \quad (3)$$

The formula holds applicability for both the Tertiary period and more recent geological epochs. The gamma logs capture the natural gamma radiation emanating from radioactive elements such as potassium (K), uranium (U), and thorium (Th), which are primarily present in clays. However, it is worth noting that non-shale radioactive minerals like Sylvite, feldspars, and micas can also contribute to gamma logs, potentially introducing inaccuracies [45–48]. Therefore, for specific geological problems, the above equations may be further modified or machine learning algorithms can be employed, using either gamma logs alone or a combination of multivariate inputs including gamma logs, density, P-wave velocity, and acoustic impedance, among others.

2.2. Data

Nestled in the southern North Sea basin, the F3 block (54.8669° N, 4.8131° E) in the Dutch sector lies atop the Central Graben, bordered by the Ringkøbing-Fyn High to the east and the Mid-North Sea High to the west. Geologically, the Mid-Miocene Unconformity (MMU) serves as a demarcation line, separating the Cenozoic into the underlying Paleogene units and the overlying Neogene units [49]. The Paleogene units beneath have been structurally influenced by the syn-depositional halokinetic movements of the underlying Zechstein evaporates, manifesting in the formation of faults and drapes. In contrast, the Neogene units overlaying them consist primarily of coarse-grained prograding sediments with an overall high porosity of 22–30%. These sediments have formed immense polycyclic fluvio-deltaic systems, a direct response to high-frequency relative sea-level cycles [50]. Evidence of these cycles can be seen in large-scale sigmoidal beddings of truncations, with onlaps onto them to the east, and downlaps to the west [51,52]. Within this geological context, our study focuses specifically on a select interval of log data (FS2–FS6) obtained from four wells within the F3 block, providing insights into a subset of these remarkable fluvio-deltaic systems.

The 3D seismic data and log data for this study originate from the F3 block, with the seismic data acquired in 1987 by NAM for petroleum exploration purposes [53]. dGB Earth Sciences has kindly provided the upper segment of the seismic data, encompassing a time range of 0 to 1.848 s. The inline range ranges from 100 to 750, while the crossline range extends from 300 to 1250, both with a 25 m line spacing, covering a total area of 386.93 km². An inline is shown in Figure 1b. Four vertical wells inside the survey are provided: F02-1, F06-1, F03-2, and F03-4, all measured by NLOG. These wells are equipped with sonic and gamma ray logs. However, it is noteworthy that according to dGB's records, only F02-1 and F03-2 have density logs available. For F03-4 and F06-1, density logs were predicted using a neural network model trained on data from the first two wells, using sonic and gamma logs as input. Additionally, the original porosity logs for all wells were calculated from density data based on a linear formula [3,13]:

$$\phi = \frac{\rho_{ma} - \rho_b}{\rho_{ma} - \rho_{fl}} \quad (4)$$

where ϕ is porosity, and ρ_{ma} and ρ_{fl} are the densities of the matrix and mud filtrate respectively, with ρ_{ma} set to 2.65 g cm⁻³ and ρ_{fl} set to 1.05 g cm⁻³. Furthermore,

for this study, the clay volume fraction was calculated from gamma logs using the Larionov relations.

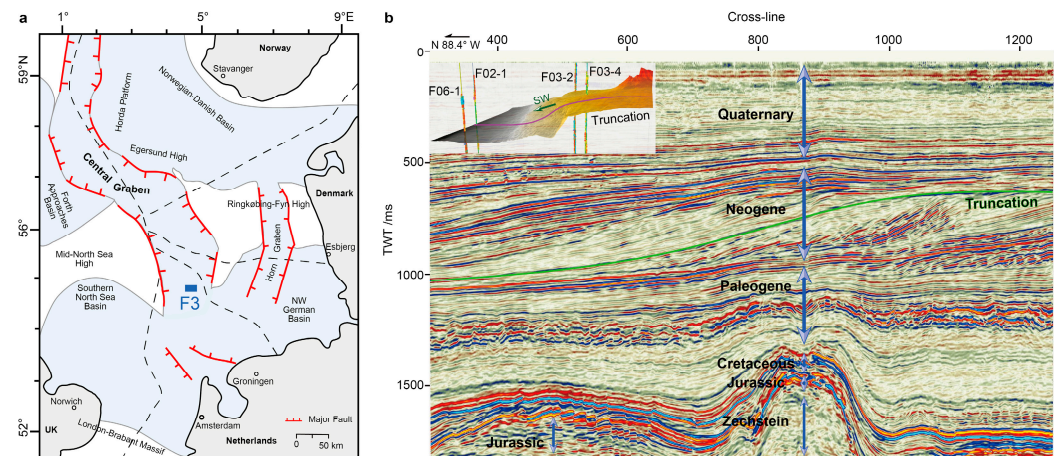


Figure 1. Geology setting of the F3 block. (a) Structural setting of the North Sea basin. The location of the F3 block is highlighted by a blue rectangle. (b) Seismic section of the F3 block (Inline 362). The geological periods of strata are annotated. The green line is the Truncation horizon. The inset is the Truncation horizon in 3D, with the four wells and the prograding direction annotated. The red line serves as the intersection between the horizon and the seismic section.

Figure 2 gives a comprehensive overview of the seven logs corresponding to the four distinct wells. For each well, a specific interval ranging from FS2 to FS6 is selected and presented. In comparison to F03-2 and F03-4, the depths recorded for F02-1 and F06-1 are notably deeper and shorter. This difference is due to their geographic position, which lies in the southwestern periphery of the survey area, precisely at the foreland of the prograding delta, as depicted in Figure 1.

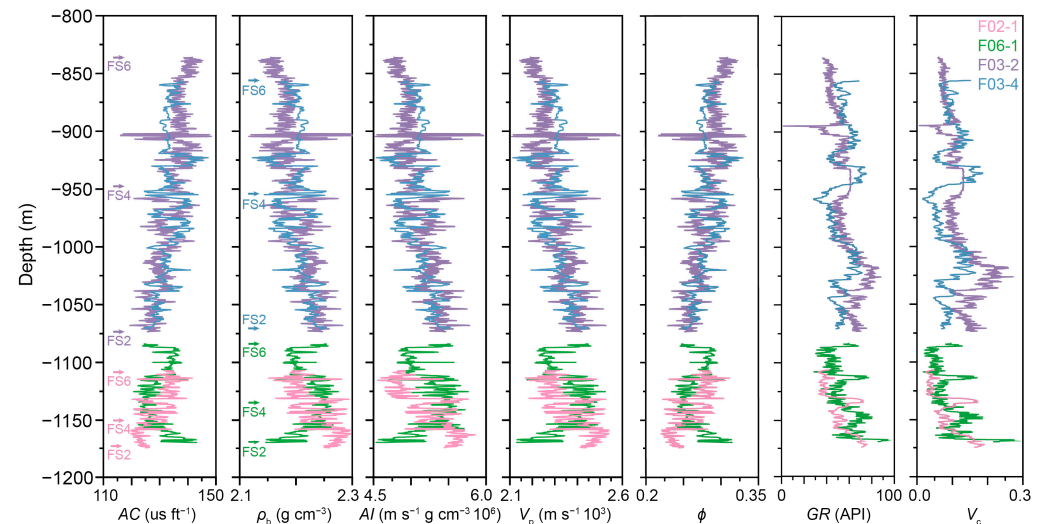


Figure 2. Log data obtained from the four wells. From left to right: acoustic, bulk density, acoustic impedance, P-wave velocity, porosity, gamma, clay volume. The top of each interval in each well is denoted using the same color as the line of the logs.

2.3. Machine Learning

Four different popular ML algorithms were used, comprising Linear Regression (LR), Multi-Layer Perceptron (MLP), Random Forests (RF), and Gradient Boosting Decision Tree (GBDT). LR, without any regularization, as the first type of regression analysis to undergo rigorous study and extensive practical application, is widely used owing to its simplicity, interpretability, and ease of implementation [54]. MLP, a commonly applied artificial

neural network model, boasts robust expressive and generalization capabilities, enabling it to tackle non-linear problems and high-dimensional data [55–57]. Both GBDT and RF are ensemble learning techniques rooted in decision trees. GBDT sequentially constructs multiple decision trees, with each subsequent tree aiming to correct the prediction errors of its predecessor, optimizing predictive performance by gradually minimizing residuals [58]. Conversely, RF concurrently builds numerous decision trees, utilizing averaging for regression and voting for classification, thereby mitigating the risk of overfitting and often yielding more stable and accurate predictions [59].

Due to the limited availability of only four well logs, we opted to utilize three of them for training the ML algorithms, reserving the remaining well for validating the predictions of the trained models. However, in scenarios where there were suspicions of inaccuracies in these logs, we relied on rock physics principles to calibrate and reconstruct the logs as replacements. As detailed in the Data section, we were able to generate density logs employing the Gardner relations, while V_c logs were derived utilizing the Larionov relations.

Standard scaling was applied to the input features for MLP and LR, whereas it was omitted for the two tree-based algorithms. The coefficient of determination (R^2) and root mean squared error (RMSE) were used as metrics to evaluate the accuracy of predictions made by ML algorithms. Furthermore, to interpret the mathematical principles of how these ML algorithms operate, both locally and globally, we used the SHapley Additive exPlanations (SHAP) framework [41]. Based on the solid theoretical foundation of cooperative game theory and the utilization of Shapley values, SHAP offers quantitative insights into the specific contributions of each feature to individual predictions, as well as their collective influence across the entire dataset [60,61]. This approach not only facilitates the assessment of ML algorithm performance but also serves as a valuable guide for feature selection and model design optimization.

The Python scripts used in this study for ML and SHAP analysis were based on the Scikit-learn and SHAP libraries [62,63].

Our primary objective is to elucidate the mathematical frameworks that underpin ML algorithms and interpret the causality behind their predictions within the context of rock physics. As such, we prefer to use standard and straightforward ML methods to facilitate their explanations and to derive conclusions that are broadly applicable. In interpreting SHAP, it is crucial to concentrate solely on the influence of input features by maintaining the consistency of the ML models. Our benchmarks, detailed in Appendix A, reveal that optimizing the models' hyperparameters did not significantly improve their performance compared to using fixed settings (Tables A1 and A2), but it did lead to notable changes in SHAP explanations (Figure A1). Therefore, we consistently avoid making any modifications based on optimization outcomes.

To address variability from random initialization, we tested 40 random seeds and found consistent results within each stochastic model, except for the MLP which showed significant fluctuations. This is shown in Figure A2.

3. Results and Discussions

3.1. Porosity

First, we show the results of ML-based prediction of porosity logs.

Figure 3 shows the density logs of the four wells, before and after correction using the Gardner relations. The original density of F02-1 is significantly lower than the other three wells. Since density logs were calculated from measured acoustic logs, we used the Gardner relations to fit the density of the three wells, where in the exponential function we

optimized the coefficient to be 0.1018 and the index to be 0.3445. Relying on the Gardner relations, density logs of F02-1 were corrected using acoustic logs as inputs.

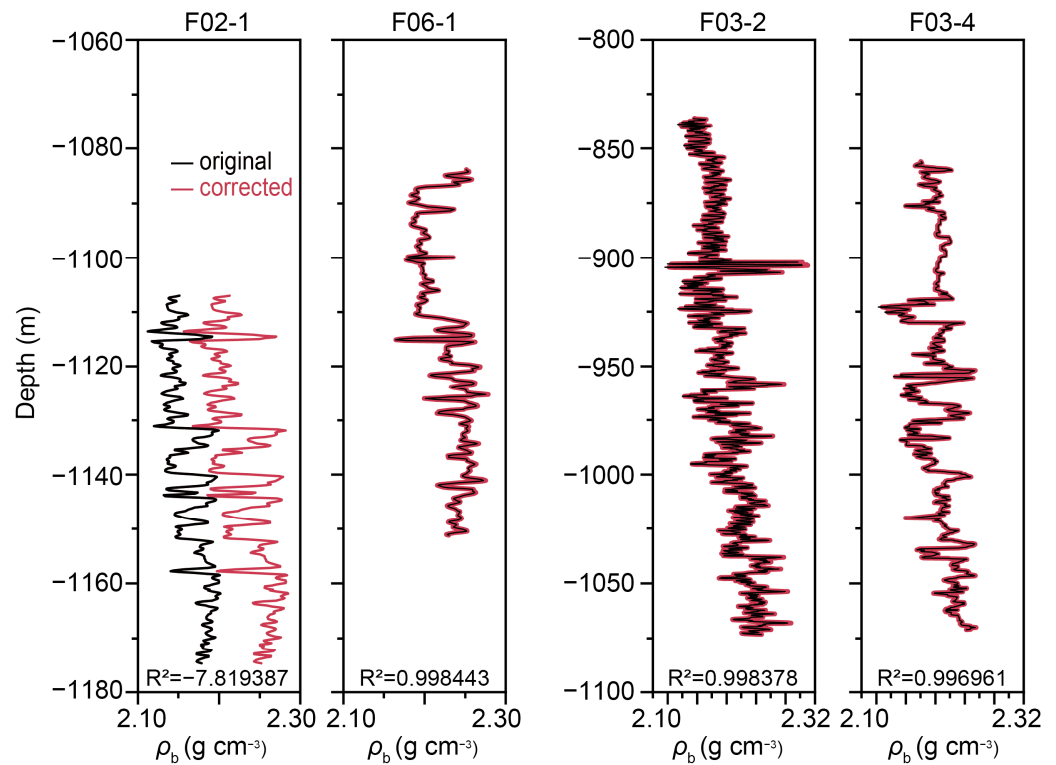


Figure 3. Density logs of the four wells before and after correction using the Gardner relations.

To gain a deeper understanding, we compiled a crossplot that illustrates the interrelationships among the seven log data sets, as shown in Figure 4. Notably, we have incorporated an additional well, F02-1Cor, which is a corrected version of F02-1. In this updated version, the density was fine-tuned using the Gardner relations, leading to a subsequent recalculation of the porosity. As observed in the crossplot comparing ρ_b and AC , the data points of F02-1 stand out conspicuously, indicating the need for further scrutiny. The corrected F02-1Cor, however, exhibits a seamless alignment with the other wells, demonstrating the effectiveness of the corrections. Moreover, we recalculated V_p as the reciprocal of AC , and consequently, AI was recalculated as the product of ρ_b and V_p . This corrected well, and F02-1Cor, exhibits an exceptional correlation across all seven logs, testament to the rigor and reliability of our findings. However, as Figure 4 indicates, the correlations are poor between gamma logs (and consequently V_c) and elastic logs, including AC , V_p , ρ_b , AI , and ϕ . This incongruence in characteristics could potentially influence ML-based predictions, and its implications will be discussed in the next section.

Figure 5 presents a comparative analysis of the ML-based prediction of porosity logs, highlighting the difference between the pre- and post-correction states of density logs. In this process, four parallel ML algorithms were used, specifically RF, GBDT, MLP, and LR. The training of these ML algorithms encompassed data from three wells, whereas the remaining well served as the testing ground. Notably, the problematic density logs from F02-1 posed a challenge.

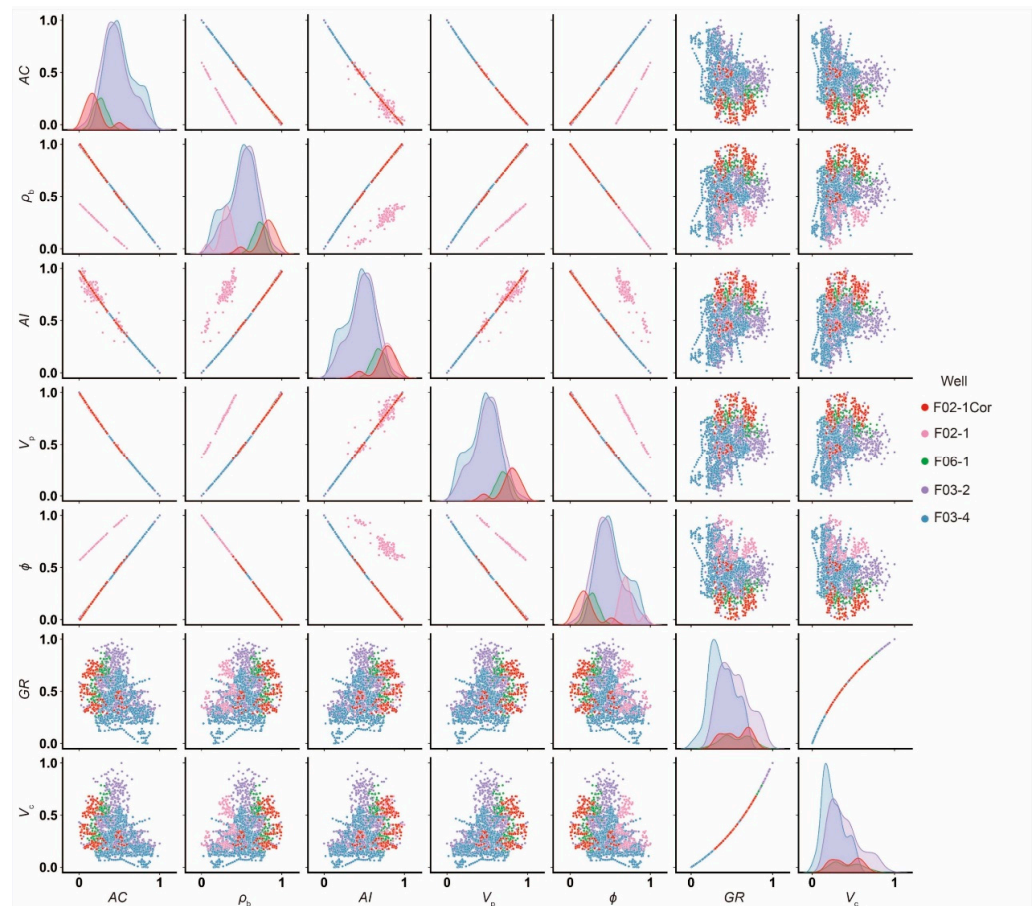


Figure 4. Crossplot of the seven logs of the four wells and F02-1Cor. From left to right, AC , ρ_b , AI , V_p , ϕ , GR and V_c .

In the pre-correction scenarios, two distinct approaches were adopted for ML training and testing: either including F02-1 in the training set or reserving it for testing. Given that porosity measurements were not readily available for all wells and were instead estimated from density logs using a linear formula, density itself was excluded as an input variable for ML training. Similarly, V_p and AI were also excluded, leaving only the measured logs, comprising AC and GR , as the input features.

Remarkably, for all four ML algorithms, the prediction accuracy of pre-correction states was deemed unsatisfactory. However, an interesting trend emerged when the problematic F02-1 was included in the training dataset, resulting in improved prediction performance compared to when it served as the testing well. Furthermore, the prediction accuracy of all four algorithms underwent a significant enhancement when both acoustic and gamma logs were used as inputs, surpassing the accuracy achieved by relying solely on acoustic logs. Despite the fact that gamma logs display a notably feeble correlation with porosity, as depicted in Figure 5, their integration with acoustic logs contributes to a more precise prediction.

After the correction of density logs, however, the prediction accuracy for all four ML algorithms experienced a substantial surge, achieving exceptionally high levels of accuracy, with R^2 values approaching 1 for each algorithm. This trend was observed when F03-4 served as the testing well, and the corrected F02-1 was incorporated into the training dataset, as shown in Figure 5c. It is crucial to emphasize that the achieved optimal performance hinges critically on the establishment of the ideal dataset setting, where the porosity data is systematically derived from the measured acoustic data through the application of a distinct and carefully calibrated formula. This underscores the paramount importance of a

rigorous and systematic approach to data processing and model training, which is essential for attaining predictions of both high fidelity and precision.

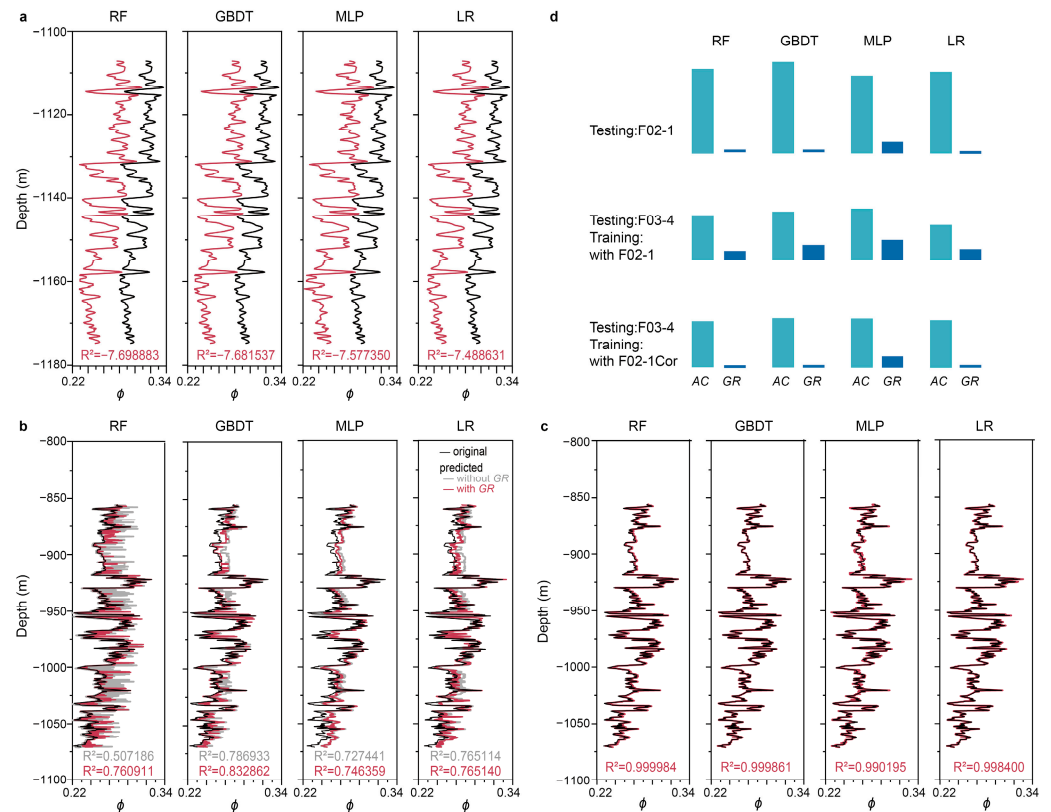


Figure 5. Prediction of porosity logs with original and corrected density logs using four ML algorithms. (a) F02-1 as the testing well. (b) F03-4 as the testing well. F02-1 was included in the training dataset. (c) F03-4 as the testing well. F02-1Cor was included in the training dataset. (d) SHAP explanation for these predictions. Top, for cases in both (a) and the cases where F02-1Cor served as the testing well. Middle, for cases in (b). Bottom, for cases in (c).

To gain a deeper insight into the mechanisms underlying ML-based predictions, global SHAP values were calculated to interpret the results. These analyses consistently pointed to acoustic logs as having an overarching influence on predicting porosity, regardless of the ML algorithm, training and testing scenarios, or prediction accuracy, as summarized in Figure 5d. Conversely, the other input logs, namely gamma logs, were largely disregarded, with the exception of one instance when the problematic F02-1 log was incorporated as an input. In this particular scenario, the inclusion of gamma logs resulted in a marginal improvement in prediction accuracy. This interpretation aligns with the weak correlation exhibited between gamma logs and porosity logs, as illustrated in Figure 4. Generally, MLP exhibits a greater reliance on gamma logs compared to the other three algorithms.

3.2. Clay Volume Fraction

Here we present the outcomes of ML-based prediction of clay volume fraction (V_c) logs and their interpretations.

Typically, V_c is derived from gamma logs, which records the natural gamma radiation emitted from radioactive elements present in clays. In this study, the Larionov relations were used to calculate V_c from the measured gamma logs. As depicted in Figure 4, V_c theoretically exhibits a strong correlation with gamma logs, but it correlates weakly with acoustic logs and other logs derived from acoustic logs, such as density, acoustic impedance, P-wave velocity, and porosity. Figure 6c shows the prediction of V_c for F03-4 when the in-

puts for training the ML algorithms included logs that correlate poorly with V_c . Specifically, the ML algorithm used was RF. The training data originated from the other three wells, spanning the interval between FS2 and FS4. As anticipated, the prediction performance was unsatisfactory. However, when gamma logs were incorporated into the training process, the prediction accuracy was improved significantly, as shown in Figure 6c. Similar results were observed when F03-2 served as the testing well and the other three wells as the training well, as shown in Figure 6e. Nonetheless, there were sections of the V_c logs for both F03-4 and F03-2 that remained inaccurately predicted.

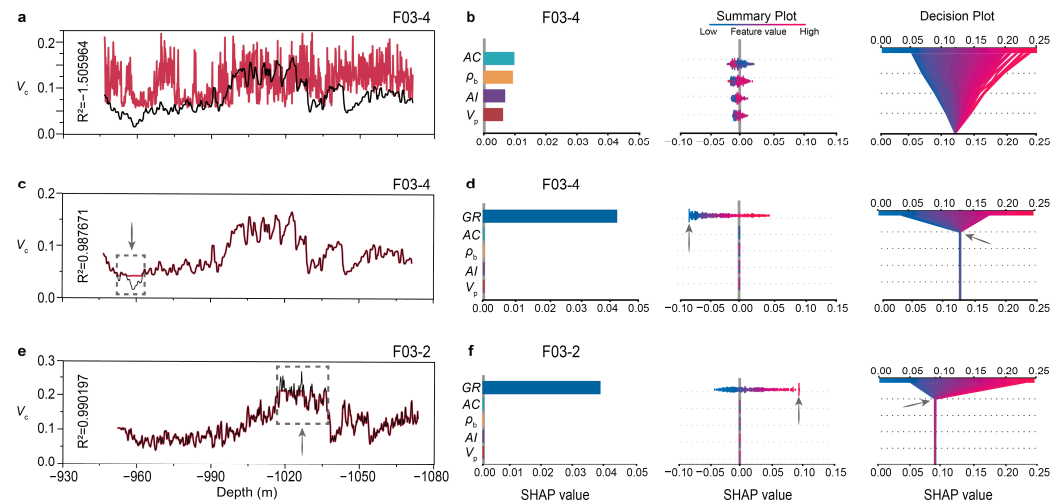


Figure 6. Prediction of clay volume fraction using RF and the SHAP explanations. (a) Prediction for F03-4 where gamma logs were excluded as inputs. (b) Feature importance for (a) as explained by SHAP summary plot and decision plot. (c) Prediction for F03-4 where gamma logs were incorporated as inputs. (d) SHAP explanations for (c). (e) Prediction for F03-2 where gamma logs were incorporated as inputs. (f) SHAP explanations for (e). Arrows indicate unpredictable samples.

To gain a more comprehensive understanding of the prediction's performance using RF and the other three ML algorithms, we calculated both global and local SHAP values for each individual prediction. In the case of predicting V_c for F03-4 using RF, gamma logs emerged as the predominant input feature, as explained by both the SHAP summary plot and decision plot (comparing Figure 6b with Figure 6d). This finding aligns with the anticipation based on human knowledge, as gamma logs are the physical trigger for the effect, namely V_c . Meanwhile, the impact of the other four input features is minimal. However, in scenarios where gamma logs are absent from the training dataset, the accuracy of V_c predictions declines significantly, rendering the importance of the four remaining input features indispensable. In such a situation, irrespective of the prediction inaccuracy, and regardless of the weak correlation between the input features and the prediction target, the ML algorithm relies heavily on the provided inputs, as it lacks alternative data sources. When assessing the impact on the inaccurate prediction results, the four input features are approximately of equal significance. This trend is observed across the other three algorithms, as shown in Figure 7b. However, if a ranking is necessary, acoustic logs emerge as the most crucial input except for LR, likely attributed to their status as the only measured feature among the inputs.

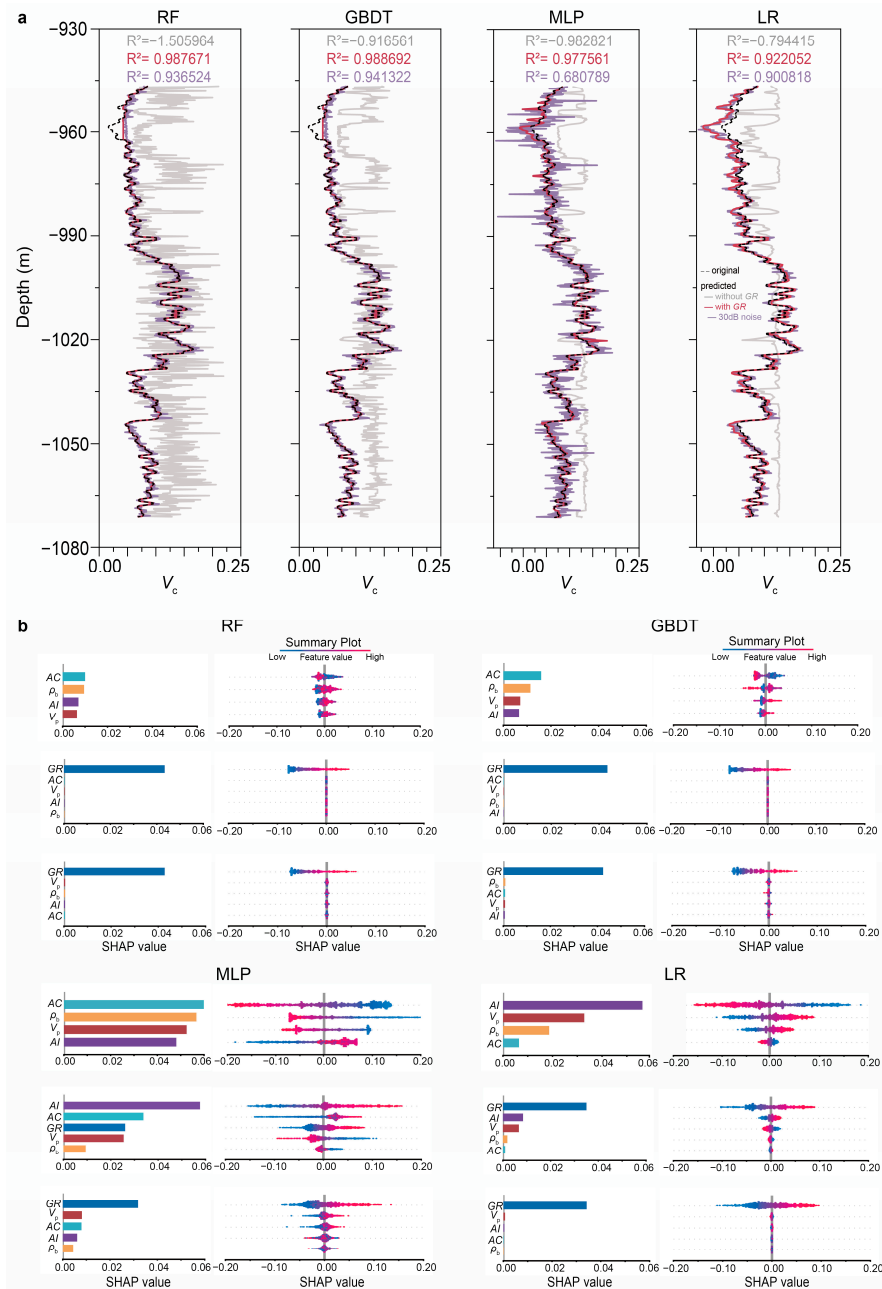


Figure 7. Accuracy and feature importance in predicting clay volume fraction for F03-4 as explained by SHAP summary plot for four ML algorithms. (a) Predicted V_c logs under three different scenarios. In contrast to a standard scenario, two variations were tested: one excluded GR in training, and the other employed a noise-robust training and testing approach with an SNR of 30. (b) Both local and global SHAP explanations for these three different prediction scenarios. For each algorithm: Top, training without GR. Middle, standard training and testing with GR. Bottom, noise-robust training and testing.

The interpretation of the significance of various features across all four ML algorithms is summarized in Figure 7. Notably, although MLP initially failed to prioritize gamma logs as the primary input feature, it regained its rationality after undergoing a noise-robust training procedure. This can be observed in Figure 7b for MLP, which demonstrates the algorithm's adaptability. Within this noise-robust training paradigm, the predictions attained satisfactory levels, albeit with a notable decline in MLP's performance due to a relatively low signal-to-noise ratio (SNR) of 30, as evident in Figure 7a for MLP. The SHAP analysis revealed consistent patterns among all four algorithms, regardless of their

varied underlying architectures—whether tree-based, neuron network-driven, or linear regression-oriented. By comparing the SHAP explanations with rock physics principles, we observed a close alignment of the four algorithms with the fundamentals of rock physics, indicating adherence to its cause relationships.

To interpret the sections of the V_c logs that exhibited unpredictable patterns at a depth of approximately -960 m, a review of the training and the testing dataset was undertaken. Specifically, considering the scenario where F03-4 serves as the testing well and the other three wells as the training well, we note that the V_c range for the other three wells ranges from 0.050 to 0.300, whereas the V_c range for F03-4 to be predicted is 0.025 to 0.200. Therefore, the predicted lowest values of V_c for F03-4 exceed the bounds of the training dataset, which is indicative of an Independent and Identically Distributed (IID) issue. Analogously, reversing the roles of the testing and training well also leads to an inaccurate prediction of the higher V_c values in F03-2. The lowest and highest unpredictable V_c values are positioned at the periphery of both the SHAP summary plot and the decision plot, as indicated by the grey arrows in Figure 6c,d, and the yellow arrows in Figure 6e,f.

To further elaborate on the intricacies of the IID problem and explain the inability to make accurate predictions, we employed a straightforward binary logarithm function of $y = \log_2 x$, with the range of x set from 1 to 11. We subjected the prediction of y to a rigorous test using four different ML algorithms, contrasting the results against those obtained from cubic spline interpolation and the ground truth. As shown in Figure 8a at $x = 2.4$, where the true value of y is denoted by black crosses along the black line, RF predicted $y = 1.291$, marked by a black dot, exceeding the actual value. From $x = 2.0$ to 2.8, the predictions either exceeded or fell below the true values, as evident in the lower inset. Notably, as we incrementally removed one or two neighboring labels from the training dataset, the predicted values rose significantly, depicted by blue dots. Conversely, the predicted values decreased progressively when one or two labels to the right were eliminated. A comparison with the ground truth is illustrated in the upper inset. Similar trends were observed for the x range from 9.0 to 9.8, as shown in Figure 8b. The performance of the remaining three algorithms is presented in Figure 8c–e. Despite the remarkable success ML has achieved owing to its numerous advantages, the predictions generated by the four ML algorithms can still trail behind the accuracy of cubic spline interpolation, even with a substantially vast training dataset.

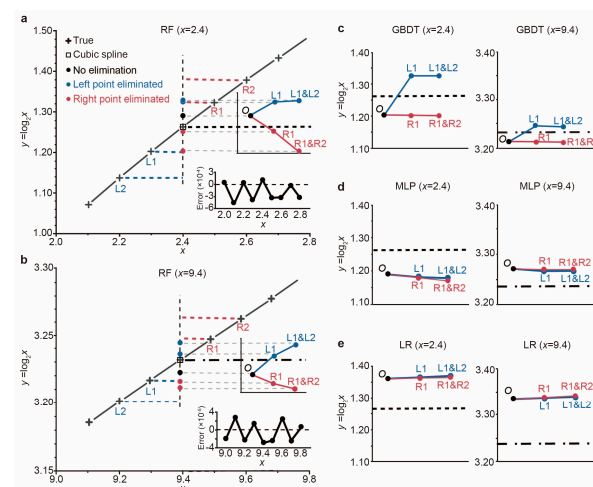


Figure 8. ML-based prediction of the binary logarithm function. (a) For x range from 2.0 to 2.8. Black line, theoretical line. Black crosses, theoretical values. Dots, ML-based predictions. Upper inset, predictions versus truth as neighboring samples eliminated from training. Lower inset, prediction errors at each testing point. (b) For x range from 9.0 to 9.8. (c) For GBDT. **Left**, $x = 2.4$. **Right**, $x = 9.4$. (d) For MLP. (e) For LR.

These findings underscore the sensitivity of ML-based predictions to the training dataset. Furthermore, as neighboring samples are gradually removed, the IID assumption becomes invalid, leading to an out-of-distribution (OOD) problem that ultimately compromises the predictive accuracy.

4. Conclusions

We have integrated SHAP and rock physics into the ML-based approach for predicting well logs, specifically targeting porosity and clay volume fraction. While ML excels in handling complex, nonlinear challenges that involve diverse input features and extensive datasets, its operational mechanisms are often less mathematically transparent compared to traditional algorithms such as least squares and Bayesian methods. Conversely, rock physics, with its clear and intuitive foundations, faces constraints in its applicability due to the numerous inherent assumptions, hampering its capacity to address ambiguous or uncertain scenarios.

The union of these two methodologies has yielded significant advantages in our study. Rock physics has played a pivotal role in feature engineering, selection, algorithm comparison, and the comprehensive evaluation of four distinct algorithms: RF, GBDT, MLP, and LR. Despite their varied architectures—tree-based, neutron network-driven, or linear regression-oriented—the SHAP analysis has uncovered consistent operational patterns across these algorithms in noise-robust training and testing scenarios. Moreover, the SHAP explanations closely align with the causal relationships outlined in rock physics principles. However, when crucial controlling input features were intentionally omitted, affecting prediction accuracy, all four ML algorithms and the SHAP analysis continued to operate in a seemingly irrational manner, starkly contrasting with the fundamental principles of rock physics. By comparing SHAP explanations with rock physics concepts, we have established parallels and contrasts with human knowledge, inherently rooted in causal relationships. The invaluable guidance provided by rock physics has empowered us to intervene in ML-based predictions, not only enhancing our comprehension of the strengths and limitations of ML models but also potentially paving the way for optimizing the performance of these workflows. This approach has propelled a transition from purely mathematical explanations to a more philosophical interpretation of ML-based predictions, effectively dismantling the traditional black-box nature of ML models.

Author Contributions: Methodology, J.Z. and Y.C.; Data analysis, J.Z. and Z.W.; Results interpretation, J.Z., G.L., S.L., Y.Z. and Y.C.; Writing—original draft, J.Z., G.L. and Y.C.; Funding acquisition, G.L. and Y.C.; Supervision, Y.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Natural Science Foundation of Xinjiang Uygur Autonomous Region (2022D01C422); the Tianchi Scholarship of Xinjiang Uygur Autonomous Region (Chengyuanfeng2022).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The custom Python scripts and data associated with this research are available and can be accessed via the following URL: <https://gitee.com/zhangjixju/ml-code/tree/master> (accessed on 13 January 2025).

Acknowledgments: The authors are very grateful to Xingmou Wang of SINOPEC for valuable scientific discussions on field applications of MLs. We thank dGB Earth Sciences for making the data available as an OpendTect project via their TerraNubis portal terranubis.com.

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

Appendix A

In this appendix, we provide a comprehensive account of the execution of the ML and SHAP methodologies.

To assess the effectiveness of hyperparameter optimization, two comprehensive benchmark tests were conducted for the three stochastic models, RF, GBDT and MLP. One test, termed case- ϕ (corresponding to Figure 5c), evaluated the porosity prediction with exceptional accuracy across all four algorithms. The other test, termed case- V_c (corresponding to Figure 6a), evaluated the V_c prediction and found significantly lower accuracy for all four algorithms.

Table A1 summarizes all hyperparameters and their respective final values for each model. Table A2 provides a detailed description of the hyperparameter tuning process, which utilized random search and k-fold cross-validation.

The benchmark tests indicate that optimizing the models' hyperparameters does not yield a substantial improvement in their performance metrics, specifically R^2 and RMSE, compared to using fixed settings. Furthermore, as illustrated in Figure A1, while optimization has a negligible effect, it significantly modifies the SHAP explanations for case- V_c . Both the magnitude and the relative importance of input features exhibit notable variations.

To assess the effect of random initialization on the performance of the three stochastic models with fixed hyperparameter settings, we conducted experiments using 40 different random seeds. Our findings indicated that the results were generally consistent for each model for case- ϕ . Random initialization exhibits minimal influence on the performance of all three models. Nevertheless, for case- V_c , there are substantial fluctuations in prediction accuracy, especially for MLP, as illustrated in Figure A2.

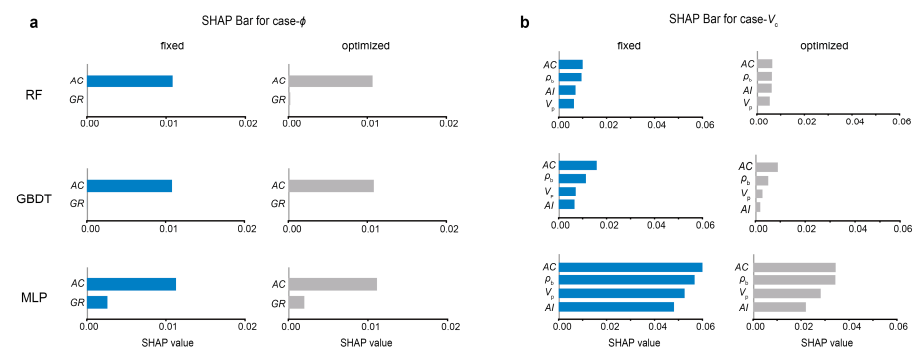


Figure A1. Global SHAP explanations for the two benchmark tests. (a) case- ϕ and (b) case- V_c . Model hyperparameters are fixed (blue) and optimized (grey).

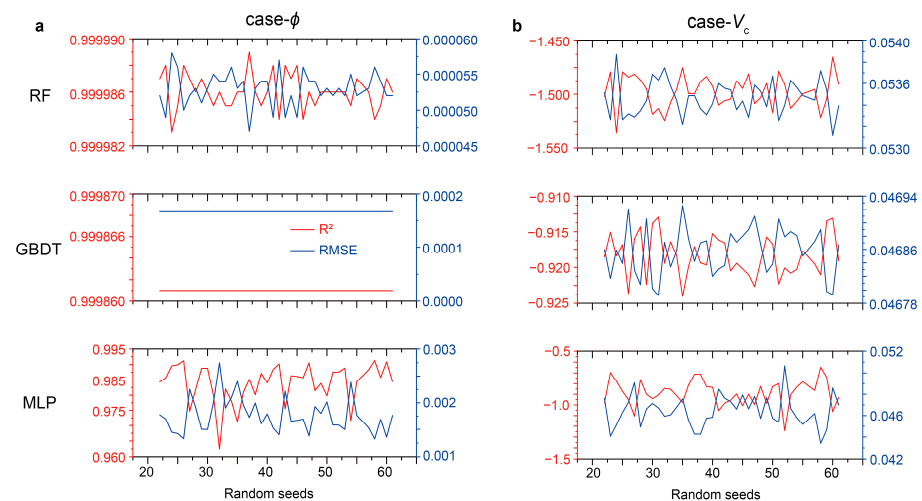


Figure A2. Performance of the three stochastic models' initialization randomly with 40 different seeds for the fixed hyperparameter settings. (a) case- ϕ . (b) case- V_c .

Table A1. Hyperparameters for the three stochastic ML models.

Model		Hyperparameters															Performance			
		n_estimators	max_depth		min_samples_splits			bootstrap		min_samples_leaf		random_state		R ²	RMSE					
RF	fixed	case- ϕ case- V_C	200	none		2			true		1		42		0.999984	0.000057				
		optimization range	100–500	10–50, none		2–10			true, false		1–4		42		–1.505964	0.053555				
	opt.	case- ϕ	200	20		2			false		1		42		0.999977	0.000068				
		case- V_C	300	10		9			true		1		42		–1.110164	0.049144				
GBDT			n_estimators	max_depth		learning_rate			subsample		min_samples_leaf		random_state		R ²	RMSE				
	fixed	case- ϕ case- V_C	200	3		0.1			1.0		1		42		0.999861	0.000167				
		optimization range	100–500	3–7		0.01–0.3			0.6–1		1–6		42		–0.916561	0.046836				
	opt.	case- ϕ	300	6		0.03			0.6		1		42		0.999992	0.000039				
case- V_C		200	3		0.01			0.7		3		42		–0.823654	0.045686					
MLP			hidden layers	number of neurons		activation	batch size	learn. rate	solver	beta1	beta2	epsilon	tol	max iter	n_iter no change	early stop.	alpha	random state	R ²	RMSE
	fixed	case- ϕ case- V_C	2	L1 100	L2 50	relu	auto	0.001	adam	0.9	0.999	10 ^{–8}	0.0001	200	50	true	0.0001	42	0.990195	0.001403
		optimization range	1–3			logistic, relu, tanh	16–128	0.001–0.1	adam, sgd	0.9	0.999	10 ^{–8}	0.0001–0.01	200–2000	50	true	0.0001–1	42	–0.982821	0.047638
	opt.	case- ϕ	2	L1 100	L2 100	logistic	40	0.003	adam	0.9	0.999	10 ^{–8}	0.01	200	50	true	0.0001	42	0.998986	0.000451
case- V_C		1	50		relu	126	0.05	sgd	0.9	0.999	10 ^{–8}	0.01	1200	50	true	0.9474	42	–0.883199	0.046426	

Table A2. Optimization of model hyperparameters via random search (200 iterations) with k-fold cross-validation, using RMSE as the performance metric.

Models	K-Fold (n_Folds = 5)		Testing: F03-4		
	Training: F02-1, F06-1, F03-2		R ²	RMSE	
		Mean_R ²			Mean_RMSE
RF	case- ϕ	0.999910	0.000186	0.999977	0.000068
	case- V_c	0.020244	0.046256	-1.110164	0.049144
GBDT	case- ϕ	0.999965	0.000114	0.999992	0.000039
	case- V_c	0.110894	0.044113	-0.823654	0.045686
MLP	case- ϕ	0.999084	0.000615	0.998986	0.000451
	case- V_c	0.122800	0.043823	-0.883199	0.046426

References

- Ellis, D.V.; Singer, J.M. *Well Logging for Earth Scientists*, 2nd ed.; Springer: Berlin/Heidelberg, Germany, 2008.
- Selley, R.C.; Sonnenberg, S.A. *Elements of Petroleum Geology*, 3rd ed.; Academic Press: San Diego, CA, USA, 2015.
- Tiab, D.; Donaldson, E.C. *Petrophysics: Theory and Practice of Measuring Reservoir Rock and Fluid Transport Properties*, 4th ed.; Gulf Professional Publishing: Boston, MA, USA, 2015.
- Wang, H.; Ma, J.; Li, L. Pore pressure prediction based on rock physics theory and its application in seismic inversion. *J. Appl. Geophys.* **2024**, *229*, 105494. [[CrossRef](#)]
- Aki, K.; Richards, P.G. *Quantitative Seismology*, 1st ed.; W.H. Freeman and Co.: San Francisco, CA, USA, 1981.
- Avseth, P.; Mukerji, T.; Mavko, G. *Quantitative Seismic Interpretation*, 1st ed.; Cambridge University Press: Cambridge, UK, 2005.
- Ettayebi, M.; Wang, S.; Landrø, M. Time-Lapse 3D CSEM for Reservoir Monitoring Based on Rock Physics Simulation of the Wisting Oil Field Offshore Norway. *Sensors* **2023**, *23*, 7197. [[CrossRef](#)] [[PubMed](#)]
- Mavko, G.; Mukerji, T.; Dvorkin, J. *The Handbook of Rock Physics*, 1st ed.; Cambridge University Press: Cambridge, UK, 1998.
- Hertz, H. Ueber die Berührung fester elastischer Körper. *J. Reine Angew. Math.* **1882**, *92*, 156–171. [[CrossRef](#)]
- Zoeppritz, K. Erdbebenwellen VIII B. Über reflexion und durchgang seismischer wellen durch unstetigkeitsflächen. *Göttinger Nachr.* **1919**, *1*, 66–84.
- Gassmann, F. Über die elastizität poroser medien. *Vierteljahrsschr. Naturforschenden Ges. Zur.* **1951**, *96*, 1–23.
- Kuster, G.T.; Toksöz, M.N. Velocity and attenuation of seismic waves in two-phase media: Part I. Theoretical Formulations. *Geophysics* **1974**, *39*, 587–606. [[CrossRef](#)]
- Wyllie, M.R.J.; Gregory, A.R.; Gardner, L.W. Elastic wave velocities in heterogeneous and porous media. *Geophysics* **1956**, *21*, 41–70. [[CrossRef](#)]
- Larionov, V.V. *Borehole Radiometry*; Nedra: Moscow, Russia, 1969.
- Poupon, A.; Gaymard, R. The evaluation of clay content from logs. In Proceedings of the SPWLA Annual Logging Symposium, Los Angeles, CA, USA, 3–6 May 1970; SPWLA-1970-G.
- Gardner, G.H.F.; Gardner, L.W.; Gregory, A.R. Formation velocity and density—the diagnostic basics for stratigraphic traps. *Geophysics* **1974**, *39*, 770–780. [[CrossRef](#)]
- Han, D. Effects of Porosity and Clay Content on Acoustic Properties of Sandstones and Unconsolidated Sediments. Ph.D. Thesis, Stanford University, Stanford, CA, USA, 1986.
- Fjaer, E.; Holt, R.M.; Horsrud, P.; Raaen, A.M.; Risnes, R. *Petroleum Related Rock Mechanics*; Elsevier: Amsterdam, The Netherlands, 2008.
- Simm, R.; Bacon, M. *Seismic Amplitude: An Interpreter's Handbook*, 1st ed.; Cambridge University Press: Cambridge, UK, 2014.
- Pham, N.; Wu, X.; Naeini, E.Z. Missing well log prediction using convolutional long short-term memory network. *Geophysics* **2020**, *85*, WA159–WA171. [[CrossRef](#)]
- Feng, R.; Grana, D.; Balling, N. Variational inference in Bayesian neuralnetwork for well-log prediction. *Geophysics* **2021**, *86*, M91–M99. [[CrossRef](#)]
- Chen, B.; Zeng, X.; Fan, L.; Li, K.; Zhang, W.; Cai, G.; Wang, Y.; Du, R.; Chen, T.; Zhang, B.; et al. Adaptive spatiotemporal neural networks based on machine learning for missing well-log prediction. *Geophysics* **2023**, *88*, V431–V443. [[CrossRef](#)]
- Lin, L.; Wei, H.; Wu, T.; Zhang, P.; Zhong, Z.; Li, C. Missing well-log reconstruction using a sequence self-attention deep-learning framework. *Geophysics* **2023**, *88*, D391–D410. [[CrossRef](#)]
- Pan, W.; Torres-Verdín, C.; Duncan, I.J.; Pyrcz, M.J. Improving multiwell petrophysical interpretation from well logs via machine learning and statistical models. *Geophysics* **2023**, *88*, D159–D175. [[CrossRef](#)]

25. Anjom, F.K.; Vaccarino, F.; Socco, L.V. Machine learning for seismic exploration: Where are we and how far are we from the holy grail? *Geophysics* **2024**, *89*, WA157–WA178. [CrossRef]
26. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [CrossRef] [PubMed]
27. O’Callaghan, J. How OpenAI’s text-to-video tool Sora could change science-and society. *Nature* **2024**, *627*, 475–476. [CrossRef] [PubMed]
28. Quiñero-Candela, J.; Sugiyama, M.; Lawrence, N.D.; Schwaighofer, A. *Dataset Shift in Machine Learning*, 1st ed.; The MIT Press: Cambridge, MA, USA, 2009.
29. Pearl, J.; Mackenzie, D. *The Book of Why: The New Science of Cause and Effect*, 1st ed.; Basic Books: New York, NY, USA, 2018.
30. Zhang, Y.; Tino, P.; Leonardis, A.; Tang, K. A survey on neural network interpretability. *IEEE Trans. Emerg. Top. Comput. Intell.* **2021**, *5*, 726–742. [CrossRef]
31. Wang, J.; Lan, C.; Liu, C.; Ouyang, Y.; Qin, T.; Lu, W.; Chen, Y.; Zeng, W.; Yu, P.S. Generalizing to unseen domains: A survey on domain generalization. *IEEE Trans. Knowl. Data Eng.* **2023**, *35*, 8052–8072.
32. Schölkopf, B. Causality for machine learning. *arXiv* **2019**, arXiv:1911.10500v2.
33. Schölkopf, B.; Locatello, F.; Bauer, S.; Ke, N.R.; Kalchbrenner, N.; Goyal, A.; Bengio, Y. Toward causal representation learning. *Proc. IEEE* **2021**, *109*, 612–634. [CrossRef]
34. Khan, I.U.; Aslam, N.; AlShedayed, R. A proactive attack detection for heating, ventilation, and air conditioning (HVAC) system using explainable extreme gradient boosting model (XGBoost). *Sensors* **2022**, *22*, 9235. [CrossRef] [PubMed]
35. Mustafa, A.; Koster, K.; AlRegib, G. Explainable machine learning for hydrocarbon prospect risking. *Geophysics* **2024**, *89*, WA13–WA24. [CrossRef]
36. Karniadakis, G.E.; Kevrekidis, I.G.; Lu, L.; Perdikaris, P.; Wang, S.; Yang, L. Physics-informed machine learning. *Nat. Rev. Phys.* **2021**, *3*, 422–440. [CrossRef]
37. Li, P.; Liu, M.; Alfarraj, M.; Tahmasebi, P.; Grana, D. Probabilistic physics-informed neural network for seismic petrophysical inversion. *Geophysics* **2024**, *89*, M17–M32. [CrossRef]
38. Zhao, T.; Liu, C.; Song, C.; Waheed, U.B.; Zhang, X. Smoothness: The key factor in well-log information-assisted PINNtomo. *J. Appl. Geophys.* **2024**, *226*, 105417. [CrossRef]
39. Liu, T.; Ungar, L.; Kording, K. Quantifying causality in data science with quasi-experiments. *Nat. Comput. Sci.* **2021**, *1*, 24–32. [CrossRef]
40. von Rueden, L.; Mayer, S.; Beckh, K.; Georgiev, B.; Giesselbach, S.; Heese, R. Informed machine learning—a taxonomy and survey of integrating prior knowledge into learning systems. *IEEE Trans. Knowl. Data Eng.* **2023**, *35*, 614–633. [CrossRef]
41. Lundberg, S.M.; Lee, S.I. A Unified Approach to Interpreting Model Predictions. *Adv. Neural Inf. Process. Syst.* **2017**, *30*, 4765–4774.
42. Lee, J.; Lumley, D.E. Interpreting the effects of shale rock properties on seismic anisotropy by statistical and machine learning methods. *Geoenergy Sci. Eng.* **2023**, *224*, 211631. [CrossRef]
43. Nazat, S.; Arreche, O.; Abdallah, M. On Evaluating Black-Box Explainable AI Methods for Enhancing Anomaly Detection in Autonomous Driving Systems. *Sensors* **2024**, *24*, 3515. [CrossRef]
44. Oloruntobi, O.; Butt, S. The new formation bulk density predictions for siliciclastic rocks. *J. Pet. Sci. Eng.* **2019**, *180*, 529–537. [CrossRef]
45. Tudge, J.; Tobin, H.J. Velocity-porosity relationships in smectite-rich sediments: Shikoku Basin, Japan. *Geochem. Geophys. Geosystems* **2013**, *14*, 5194–5207. [CrossRef]
46. Wang, H.; Liu, T.; Tang, T.; Shi, Y. A unified model to evaluate shaliness in compacted and soft formations using downhole GR log. *J. Pet. Sci. Eng.* **2017**, *156*, 877–883. [CrossRef]
47. Aftab, F.; Zafar, M.; Hajana, M.I.; Ahmad, W. A novel gas sands characterization and improved depositional modeling of the Cretaceous Sembar Formation, Lower Indus Basin, Pakistan. *Front. Earth Sci.* **2022**, *10*, 1039605. [CrossRef]
48. Zhu, L.Q.; Sun, J.; Zhou, X.Q.; Li, Q.P.; Fan, Q.; Wu, S.L.; Wu, S.G. Well logging evaluation of fine-grained hydrate-bearing sediment reservoirs: Considering the effect of clay content. *Pet. Sci.* **2023**, *20*, 879–892. [CrossRef]
49. Steeghs, P.; Overeem, I.; Tigrek, S. Seismic volume attribute analysis of the Cenozoic succession in the L08 block (Southern North Sea). *Glob. Planet. Change* **2000**, *27*, 245–262. [CrossRef]
50. Sørensen, J.C.; Gregersen, U.; Breiner, M.; Michelsen, O. High-frequency sequence stratigraphy of Upper Cenozoic deposits in the central and southeastern North Sea areas. *Mar. Pet. Geol.* **1997**, *14*, 99–123. [CrossRef]
51. Mojeddifar, S.; Kamali, G.; Ranjbar, H. Porosity prediction from seismic inversion of a similarity attribute based on a pseudo-forward equation (PFE): A case study from the North Sea Basin, Netherlands. *Pet. Sci.* **2015**, *12*, 428–442. [CrossRef]
52. Ishak, M.A.; Islam, M.A.; Shalaby, M.R.; Hasan, N. The Application of Seismic Attributes and Wheeler Transformations for the Geomorphological Interpretation of Stratigraphic surfaces: A Case Study of the F3 Block, Dutch Offshore Sector, North Sea. *Geosciences* **2018**, *8*, 79. [CrossRef]
53. dGB. Available online: <https://terranubis.com/datainfo/F3-Demo-2020> (accessed on 26 June 2024).
54. Gross, J. *Linear Regression*; Springer: Berlin/Heidelberg, Germany, 2003.

55. Rosenblatt, F. The perceptron: A probabilistic model for information storage and organization in the brain. *Psychol. Rev.* **1958**, *65*, 386–408. [[CrossRef](#)]
56. Minsky, M.; Papert, S. *Perceptrons: AN Introduction to Computational Geometry*, 1st ed.; The MIT Press: Cambridge, MA, USA, 1969.
57. Rumelhart, D.E.; Hinton, G.E.; Williams, R.J. Learning representations by back-propagating errors. *Nature* **1986**, *323*, 533–536. [[CrossRef](#)]
58. Friedman, J.H. Greedy function approximation: A gradient boosting machine. *Ann. Stat.* **2001**, *29*, 1189–1232. [[CrossRef](#)]
59. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
60. Shapley, L.S. A value for n-person games. In *Contributions to the Theory of Games*; Princeton University Press: Princeton, NJ, USA, 1953; pp. 307–317.
61. Lundberg, S.M.; Erion, G.; Chen, H.; DeGrave, A.; Prutkin, J.M.; Nair, B.; Katz, R.; Himmelfarb, J.; Bansal, N.; Lee, S.-I. From local explanations to global understanding with explainable AI for trees. *Nat. Mach. Intell.* **2020**, *2*, 56–67. [[CrossRef](#)] [[PubMed](#)]
62. Scikit-Learn. Available online: <https://scikit-learn.org> (accessed on 12 January 2025).
63. Lundberg, S.M. Available online: <https://pypi.org/project/shap> (accessed on 26 June 2024).

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.