

Predicting Permeability from Well Logs in Carbonate Formation Using Machine Learning

Ikeh, Lesor^{1*}; Asimiea, N. W.²

^{1,2}University of Port Harcourt, Faculty of Engineering, Department of Petroleum and Gas Engineering, East-West Road, Choba, Port Harcourt, Nigeria, P M B 5323, Choba.

Corresponding Author: Ikeh, Lesor^{1*}

Publication Date: 2026/04/10

Abstract: The rate of fluid flow through reservoir rocks is determined by permeability, one of the key characteristics of a reservoir. Production forecasts, history matching, and robust reservoir simulation all depend heavily on accurate permeability estimates. It can be difficult to build trustworthy permeability models because of the inherent variety of permeability at various sizes and the scarcity of core data. To get above these obstacles, this work uses a variety of machine learning techniques, including Support Vector Regression (SVR), Random Forest (RF), XGBoost, and LightGBM, to predict lab-measured core permeability from frequently obtained well logs. A datasets that represented a carbonate platform (Y-Field) was considered. The resilience of this technique under various geological settings could be assessed using the Y- field dataset, which consisted of 17 wells spread across a single reservoir. This approach relies heavily on feature engineering, especially when it comes to integrating vertical variability. Taking into account the smoothing effect of well logs over small-scale heterogeneities and the significance of spatial context, measurements from nearby well log readings were added into the models. By taking into consideration nearby depositional environments and shared geological history, this increased prediction accuracy. Results shows that the R2 values in Y- Field's blind tests were as high as 0.64, and the leave-one-well-out cross-validation technique produced validation R2 values as high as 0.8. The Yorla, Kpean, and Teera-ue formations had blind test R2 ratings of up to 0.82, 0.74, and 0.80 for the Y- Field, respectively. Even if these results are satisfactory, they demonstrate how machine-learning techniques can be used to accurately estimate permeability and emphasize the need for feature engineering. This work argues that although automated feature engineering using machine learning shows potential, human intervention more especially, the incorporation of geographical context can still greatly improve predictions. It may be the goal of future developments to incorporate this spatial awareness into machine learning algorithms.

Keywords: Permeability, Formation, Well Logs, Reservoir, Fluid, Porous Rock, Carbonate.

How to Cite: Ikeh, Lesor; Asimiea, N. W. (2026) Predicting Permeability from Well Logs in Carbonate Formation Using Machine Learning. *International Journal of Innovative Science and Research Technology*, 11(3), 3770-3786. <https://doi.org/10.38124/ijisrt/26mar2050>

I. INTRODUCTION

Permeability is a measure of how well a porous rock or soil allows fluids, such as oil and natural gas, to pass through it. It is an important property to consider in the oil and gas industry, as it determines the ease with which fluids can be extracted from the ground. The higher the permeability, the easier it is to extract the oil or gas. Permeability is typically measured in units of darcies or millidarcies. A rock or soil with high permeability, such as sandstone, may have a permeability of hundreds or thousands of darcies, while a rock with low permeability, such as shale, may have a permeability of less than one darcy. Permeability is determined by several factors, including the size, shape, and distribution of the pores in the rock or soil, as well as the nature of the fluids being extracted. For example, oil and gas can have different permeabilities due to their different physical and chemical properties, this is known as relative

permeabilities. In the oil and gas industry, permeability is an important factor to consider when selecting a site for drilling. High permeability rocks or soils are more likely to yield significant amounts of oil or gas, while low permeability rocks or soils may require more advanced extraction techniques, such as hydraulic fracturing, to extract the fluids. Overall, permeability is a crucial property to consider in the oil and gas industry, as it determines the ease with which fluids can be extracted from the ground and the potential productivity of a particular drilling site.

The traditional method of determining permeability is by conducting laboratory tests on samples of the rock or soil from the drilling site. These tests may involve measuring the flow of a fluid through the sample under controlled conditions, such as constant pressure, to determine the permeability. Several different methods can be used to measure permeability in the laboratory, including Darcy's

law, the flexible wall permeameter, and the constant head permeameter. Each method has its advantages and disadvantages, and the choice of method may depend on the specific characteristics of the rock or soil being tested.

Permeability can also be determined by conducting field tests, such as injection tests or production tests. Injection tests involve injecting a fluid into the ground and measuring the flow rate, pressure, and other factors to determine the permeability of the rock or soil. Production tests involve extracting fluids from the ground and measuring the flow rate and other factors to determine permeability.

Another method of determining permeability is the use of well logs. Well logs are a type of record that is used in the oil and gas industry to provide information about the geologic characteristics of a drilling site. Well logs typically include measurements of various physical and chemical properties of the rock or soil, such as porosity, density, and resistivity, as well as information about the location and depth of the drilling site. From well logs, information about the size, shape, and distribution of the pores in the rock or soil can be determined. For example, well logs can show the presence of large, interconnected pores, which may indicate high permeability, or the presence of small, isolated pores, which may indicate low permeability. In addition to providing information about permeability, well logs can also be used to identify potential reservoirs of oil and natural gas, as well as to evaluate the feasibility of drilling at a particular site. Well logs can be used in conjunction with other geologic data, such as seismic data, to create a more complete picture of the subsurface conditions at a drilling site. Overall, well logs are an important tool in the oil and gas industry, as they provide valuable information about the geologic characteristics of a drilling site, including permeability, that can be used to determine the feasibility and potential productivity of a well.

Factors that determine which method to use in determining permeability can be the level of accuracy required, phase of the field (exploration or production), and cost. Generally speaking, the core analysis provides the best result because it is a direct analysis of the physical rock being studied. Although, it is an expensive process because of the number of people and steps involved (site, logistics, laboratory). It also has a direct impact on drilling time due to coring trips and the result takes time to be available. The major advantage of well testing is that it can be done during the development or production phase of the well. This can give us the current well conditions. This process is also expensive because it involves obstructing production and running pressure gauges to the bottom hole for data gathering, and the test takes time depending on the design. The analysis process involves a lot of parties (data gathering team, data analysis team, and other teams providing other useful data like geological data) and it is error-prone. Well logging has proven over the years to be the most used method of determining the permeability of wells in the oil and gas industry. The data is obtained during drilling operations, and traditionally, permeability has been calculated using complex mathematical models, correlations, and manual analysis of

well logs. However, this process can be time-consuming and subject to human error.

In recent years, advances in machine learning have made it possible to automate and improve the analysis of well logs. Machine learning is a type of artificial intelligence that involves training a computer model to make predictions or decisions based on data. It is based on the idea that systems can learn from data, identify patterns, and make decisions or predictions without being explicitly programmed to do so. By using machine learning algorithms, it is possible to quickly and accurately analyse large amounts of well log data, which can be difficult and time-consuming to do manually. In addition, machine learning algorithms can be trained to recognize patterns and trends in the data that may not be immediately obvious to a human analyst, which can help improve the accuracy of results. Machine learning is used in the oil and gas industry in a variety of ways, including predicting the production of wells, identifying potential drilling sites, and optimising the extraction of oil and natural gas. For example, machine learning algorithms can be trained on large datasets of geologic and production data to identify patterns and trends that can be used to improve the efficiency of drilling operations. Additionally, machine learning can be used to improve the accuracy of subsurface modelling, which is the process of creating a virtual representation of the subsurface conditions at a drilling site. This can help oil and gas companies better understand the geology of a site and make more informed decisions to improve the efficiency and accuracy of extraction operations. Machine learning algorithms can be trained to recognize patterns and trends in data, and can be used to quickly and accurately analyse large amounts of well log data. This has the potential to improve the accuracy and efficiency of permeability calculations and to help geologists and other oil and gas professionals make more informed decisions about the productivity of a well.

II. METHODS OF PERMEABILITY DETERMINATION

➤ Core Analysis

The standard method for determining permeability is the direct method called core analysis which is achieved by taking samples of 3 feet of the zone of interest and drilling plugs of 5 inches from every thirty centimeters along the longitudinal axis of the core sample. The plugs are analyzed under controlled laboratory conditions replicating that of the reservoir. Through core analysis, small-scale heterogeneity that might not affect flow at the reservoir scale is measured and upscaled (Tagavi, 2005). Permeability measurements in core samples are based on the observation that under steady-state flowing conditions, there exists a constant pressure gradient and it is directly proportional to the fluid's flow rate. The relationship between pressure gradient and flow rate is given by Darcy's law (de Marsily, 1986). Relative permeability is also measured through core analysis. However, the permeability obtained from core analysis can be misleading because core permeability is only a representation of a particular aspect of the reservoir sample and not the entire reservoir or well as the case may be. This is also because the laboratory condition does not effectively

duplicate the conditions inherent in the reservoir. Core analysis is also an expensive method of determining formation permeability.

➤ *Well Logs*

An essential instrument in the geo-energy sector, well logging yields a multitude of information that may be utilised to evaluate a reservoir's qualities. Numerous physical characteristics of the formations and the fluids they contain are measured by several sorts of logging, and many of these characteristics are related to the reservoir rock's permeability.

For well-logs derived permeability, this is gotten by measuring physical properties of the formation rock matrix

and pore fluids and also provide formation data not readily accessible by any means other than coring. Utilizing these measured physical properties makes it practical to determine porosity, water and hydrocarbon saturations, shale content, etc. from which permeability can be estimated, over-pressured zones detected and residual oil saturation calculated. Various correlations have been proposed over the years and under certain conditions and have been able to provide acceptable approximations of hydraulic conductivity (single phase intrinsic permeability). It is inexpensive to determine permeability from well logs compared to core analysis and this method also provides permeability values on foot-by-foot basis, hence a permeability that is representative of the formation.

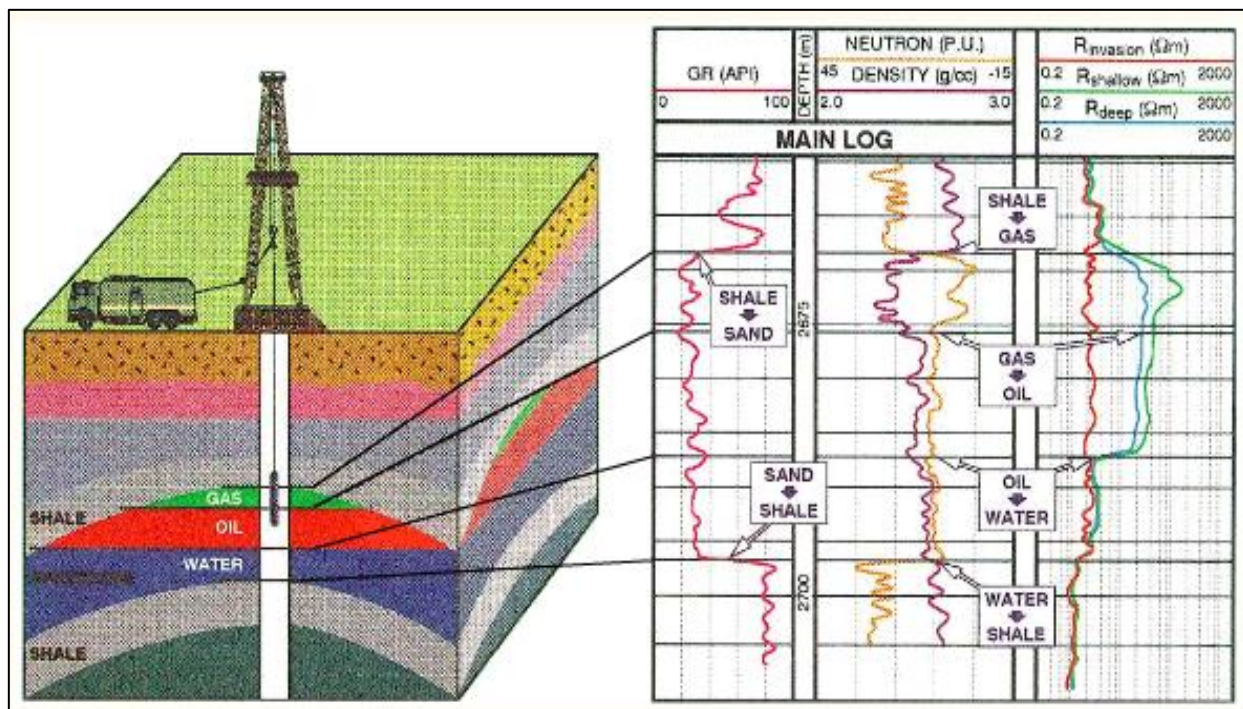


Fig 1 Field Setup of Logging Equipment and Sample of a Well Log

➤ *Empirical Approach*

Permeability is estimated from well logs using a variety of empirical models. One of these is the semi-empirical Kozeny-Carman equation, which connects a porous medium's permeability to its porosity and specific surface area (Gholanlo, 2021). Aigbedion (2007) states that additional empirical models including the Tixier, Timur, and Coates & Dumanoir techniques are also frequently used. The latter models rely on relationships between porosity, irreducible water saturation, and permeability. Assuming specific values for the cementation factor and/or saturation exponent, they mainly use resistivity measurements (Mohaghegh et al., 1997). When there is residual water saturation, these techniques are very useful for clean sand formations. It is crucial to remember that because heterogeneous formations are complicated, none of these empirical models may work effectively in them (Mohaghegh et al., 1997).

➤ *Machine Learning Approaches for Permeability Prediction*

Rezaee and Ekundayo (2022) examine how well logs and core data from five boreholes in the Surat Basin, Australia, can be used to forecast permeability using machine learning approaches. The authors discovered that the Artificial Neural Network that utilised all seven input logs had the best performance, with an approximate coefficient of determination (R²) of 0.93 for the training data set and 0.87 for the data set used for blind testing. Nevertheless, the lack of feature engineering and data standardisation in the study may have hindered the performance of other machine learning models and decreased their capacity to correctly estimate permeability.

A machine learning method for forecasting reservoir permeability and water saturation based on wireline records from the North Sea's Volve field is presented by Ottere et al. in 2021. The authors forecasted these parameters using statistical techniques and Extreme Gradient Boosting regression. They then suggested a novel ensemble model that

combines the predictions of Random Forest and Lasso Regularisation and prioritises input features according to their significance for permeability prediction. It is crucial to remember, nevertheless, that the permeability employed in this investigation is KLOGH, a permeability computed from the Statoil model, rather than the real permeability values discovered from core sampling. The interpretation of the study's findings and their direct comparability with those of other studies may be impacted by this factor. Nevertheless, it was discovered that the ensemble model that was suggested performed better in terms of accuracy, consistency, and precision than the conventional XGBoost and hybrid PCA-XGBoost models. The ensemble model, which uses the predictions of several models to improve overall forecast accuracy, is probably the study's strongest point.

The goal of (Z. Zhang et al., 2021) is to provide trustworthy techniques for estimating porosity and permeability in the Lower Cambrian dolomite reservoir in China's Tarim Basin. To assess their effectiveness, the authors gathered 253 core samples with logging data, used six machine learning techniques, and five common permeability-porosity correlations. The authors found that each discrete rock type (DRT) displayed a distinct permeability-porosity connection and that the Flow Zone Indicator (FZI) had a greater capacity to classify samples when used in conjunction with DRT analysis. These results led to the development of an integrated indirect permeability prediction approach that combines the Particle Swarm Optimization-Support Vector Machines (PSOSVM) algorithm with the physical rock typing methods (FZI). When compared to direct machine learning prediction methods, the suggested integrated methodology showed better permeability prediction accuracy, peaking at an R2 of 0.869. This study's use of a randomised holdout validation based on individual samples rather than a full fully blind test to assess the efficacy of the authors' suggested method is one possible shortcoming. It is also important to keep in mind that the 253-core sample size might not be enough to fully capture the range of complexity and variability found in the reservoir.

The goal of (Elkatatny et al., 2017) is to create an artificial neural network (ANN) model that uses well logs to estimate permeability in heterogeneous carbonate reservoirs. To train the model and evaluate its performance on unobserved data, the authors used real core data points with three logs (resistivity, bulk density, and neutron porosity). The mobility index (MI), which they coined, has the following definition:

$$MI = \frac{1}{\sqrt{R_{fz}}} - \frac{1}{\sqrt{R_{dp}}} \quad (1)$$

Where R_{fz} is the normal value of the flushed zone resistivity and R_{dp} is the normal value of the deep resistivity. The mobility index was found to improve the permeability prediction compared to the actual measured data. The ANN model showed an R2 of 0.95 and a root-mean-square error of 0.28 mD and was comparable to support vector machine and adaptive neuro-fuzzy inference system models.

Regression of logging sequences is frequently used to forecast parameters like porosity, permeability, and water saturation, and the majority of petrophysical models used for traditional logging interpretation rely on linear or nonlinear relations with Well Logs, according to (S. Zhang et al, 2022).

However, ensemble learning techniques like Light gradient boosting machine (LightGBM) have emerged as a potentially useful solution to this problem. The effectiveness of the proposed CRBM-Bayes-LightGBM hybrid predictor, which uses LightGBM as a state-of-the-art ensemble learning model to anticipate the target reservoir features, was highlighted by (S. Zhang et al., 2022). To verify the effectiveness of this model, (S. Zhang et al., 2022) carried out experiments in the Jiyuan Oilfield, Ordos Basin, Northern China, using data from the Chang 8 member. The findings show that the LightGBM-cored predictor outperforms three rival models in terms of accuracy and that small volumetric datasets can benefit from transfer learning for more precise prediction-making.

A unique method for showing the predictive patterns discovered by machine learning models for permeability prediction is put forth by (G. Zhang et al., 2021). They use deep residual neural networks (ResNet; He et al., 2016), support vector regression random forest and wireline logs and porosity as input features in addition to ResNet features of thin sections to construct machine learning models. With an R2 of 0.88 for testing and 0.9 for training, the ResNet-infused model outperforms the others in terms of permeability prediction, according to the data. Given that ResNet is a well-known pre-trained model for image classification, its adoption is notable.

III. METHODS AND PROCEDURES

In geological setting, the datasets that was utilized to train and assess the machine learning models for permeability prediction utilizing well logs and core data are the Y-field datasets. The datasets comprises stratigraphic well tops and core permeability data together with well logs with resistivity, gamma ray (GR), density (RHOB), neutron porosity (NPHI), sonic (DT), and resistivity measurements. The machine learning algorithms created in this work were trained and tested on these datasets

A. Machine Learning

➤ Support Vector Regression

An ideal hyperplane is created by the machine learning algorithm Support Vector Machine (SVM) regression to predict output variables. The data points, referred to as support vectors, that are closest to the decision boundary define this hyperplane (Kramer, 2016). The SVM regression algorithm seeks to minimize the error between predicted and actual values and maximize the margin between these support vectors. The term "margin" refers to a range around the hyperplane where errors are allowed. In SVM regression, the epsilon-insensitive loss function, represents the trade-off between maximizing the number of data points falling inside

the range and minimizing the error of data points falling outside (G. Zhang et al., 2021).

To generate a linear decision boundary in the new feature space, the SVM regression algorithm uses a kernel function to convert the input data into a higher dimensional feature space. According to Vapnik (1998), the most often utilized kernel functions are the radial basis function (RBF), polynomial, and linear.

After the data is processed, the algorithm finds the coefficients of the hyperplane that best divides the data by choosing the support vectors, a subset of the training data. By maximizing the margin between the support vectors and lowering the sum of squared errors between the expected and actual values, this is accomplished. Gradient descent is one method used to optimize the coefficients of the hyperplane (G. Zhang et al., 2021).

During training, the SVM regression algorithm uses a regularization parameter called C. The goal of this parameter is to strike a balance between increasing the margin and lowering the error. According to Brereton and Lloyd (2010), a greater value of C produces a narrower margin but may cause the data to be overfit. A smaller value of C produces a broader margin but may increase mistakes in the training data.

The SVM regression model can be used to predict the output variable on fresh, unknown data after training. The input data is translated into the same feature space using the kernel function, and the position of the input data concerning the hyperplane is used to generate the expected output (G. Zhang et al., 2021).

➤ *Random Forest Regression*

By randomly choosing a subset of characteristics from the training data, the random forest regression technique constructs a decision tree. Until a stopping condition, such as a maximum tree depth or a minimum number of data points at a leaf node, is satisfied, the process of recursively splitting the data into smaller subsets based on the values of the chosen features continues (Feng et al., 2021). Several decision trees are created by repeating this procedure, and the average of all the trees' forecasts yields the final prediction.

The algorithm's biggest strength is its ability to handle noisy and missing data, as well as complicated and nonlinear interactions between input and output variables. Furthermore, the approach can pinpoint the most important input parameters for the forecast, offering a critical understanding of the underlying geologic processes (Feng et al., 2021).

A random forest regression model's hyperparameters need to be properly adjusted to maximize performance. According to Feng et al. (2021), hyperparameters are parameters that are set before the algorithm begins to learn and have a significant influence on its behaviour and performance.

➤ *Extreme Gradient Boosting (XGBoost) Regression*

Regression tasks are the main application for the potent machine-learning algorithm Extreme Gradient Boosting (XGBoost). XGBoost works by building an ensemble of decision trees and iteratively improving the model by training each tree to rectify the mistakes caused by the previous trees (Chen and Guestrin, 2016).

Regression analysis is frequently performed using LightGBM, an open-source gradient boosting framework, in a variety of fields, including geoscience and petrophysics (S. Zhang et al., 2022). It is specifically made to handle datasets with millions or more instances with efficiency, albeit the definition of "large" will change based on the task and environment. LightGBM, however, is accurate and computationally efficient over a large range of dataset sizes. Iteratively, the LightGBM regressor builds a sequence of decision trees, each of which aims to fit the residual errors of the preceding tree (Ke et al., 2017). The residual error fitting technique is used by both LightGBM and XGBoost in their gradient-boosting algorithms. But one thing that sets LightGBM apart from XGBoost is that it uses the leaf-wise growth strategy instead of the level-wise one (Saha, 2023). The efficiency and complexity of the model can be greatly impacted by these variations in tree development tactics.

➤ *Hyperparameter Tuning*

Random grid search was used to find the best hyperparameters for machine learning models, which is essential for enhancing model performance. Using a predetermined range of hyperparameters, this technique selects them at random and analyzes them over a predetermined number of iterations (Kavzoglu and Teke, 2022). With some distributions on a log scale (such as `reg_lambda`), the ranges correspond to the lowest and maximum values of uniform distributions utilized for hyperparameter adjustment. Because random grid search saves computing time and ensures a wide range of parameter exploration, it is chosen over exhaustive grid search.

➤ *Cross-Validation Techniques*

An essential method for assessing a model's performance and avoiding overfitting is cross-validation. To measure the generalization error, it entails dividing the dataset into subsets, training the model on X-1 subsets, and making predictions on the remaining subset. Understanding how effectively the model will function on unknown data is aided by this procedure. K-fold cross-validation and leave-one-well-out cross-validation are two popular forms of cross-validation.

➤ *K-fold Cross-Validation*

Depending on how the data are distributed, the dataset is divided into k equal-sized or nearly equal-sized folds using K-fold cross-validation (Pedregosa et al., 2011). Every fold is used as the test set just once during the k repetitions of this operation. The performance of the model is then calculated by averaging the outcomes of each fold. Using several folds for validation, this approach lowers the risk of overfitting and guarantees that every sample is used for both testing and training applications.

➤ *Leave-One-Well-Out Cross-Validation*

In geoscience applications where spatial autocorrelation may be present in the data, Leave-one-well-out Cross-validation (Pedregosa et al., 2011) is a common type of k-fold cross-validation. A grouping variable, like the well name or location, is used to separate the data, and the model is trained on all but one of the groups before being assessed on the remaining group.

To get an estimate of the model's performance, this process is repeated for every group in the dataset. The average of the findings is then used. This guarantees that the model can more precisely capture spatial variance in the data and does not show bias towards any particular well or location.

The decision between k-fold and leave-one-well-out cross-validation is influenced by the model's practical use in addition to its pros and cons. K-fold cross-validation may be

a better fit if an algorithm is created to interpolate between several permeability samples along a well since it manages big datasets that are randomly scattered with computational efficiency. On the other hand, the Leave-one-well-out method offers a more accurate performance estimate in an exploration scenario where the objective would be to predict values for a new well, particularly if spatial autocorrelation is present in the data. As a result, choosing among various validation techniques is contingent upon the particular research question at hand as well as the attributes of the dataset.

IV. RESULTS

➤ *Train/Test Splitting Strategy*

The data were split into training and testing wells using two distinct techniques. As shown in Table 1, the first method comprised selecting 3 wells at random for testing and the remaining 14 for training.

Table 1 Wells Used for Training and Testing

| Training Wells | Testing Wells |
|----------------|---------------|
| HW-10 | HW-26 |
| HW-24 | HW-29 |
| HW-25 | HW-7 |
| HW-27 | |
| HW-28 | |
| HW-3 | |
| HW-30 | |
| HW-31 | |
| HW-32 | |
| HW-4 | |
| HW-5 | |
| HW-6 | |
| HW-8 | |
| HW-9 | |

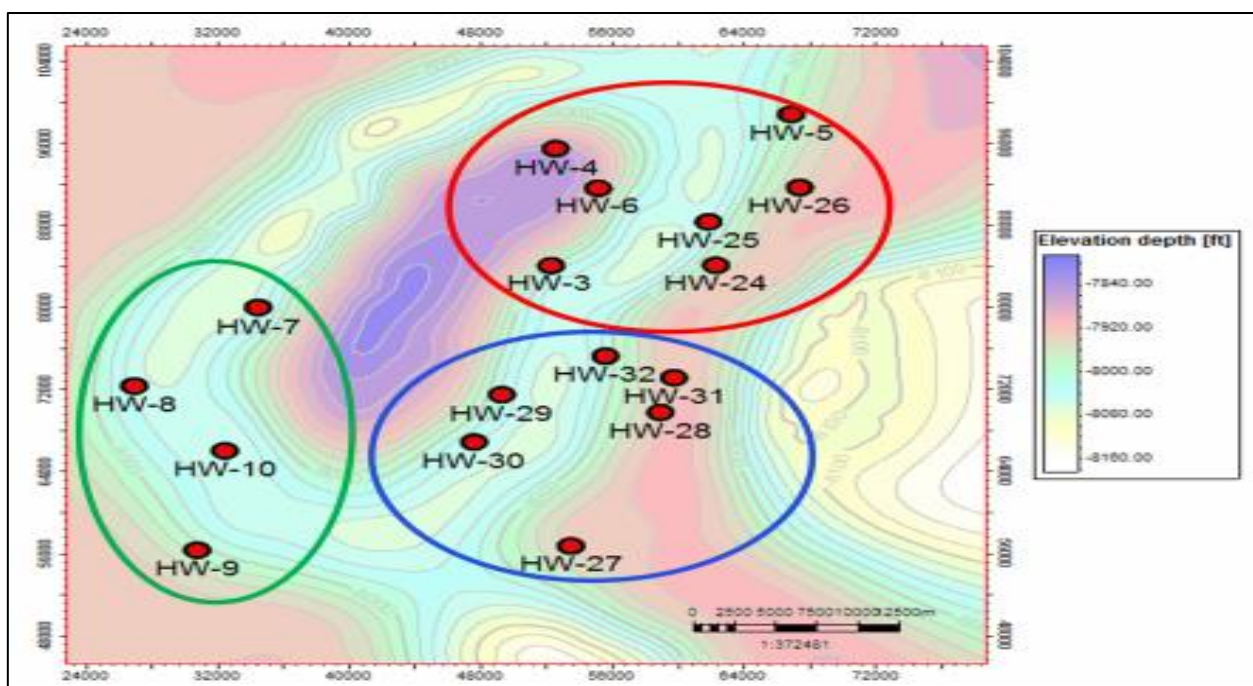


Fig 2 Contour Map Showcasing the Top of the Upper Kharib Member with Marked Well Locations, Based on Data from (Costa Gomes et al., 2022).

As shown in (Figure 2), the second method involved a more intensive process in which the data was divided into three clusters based on geographic position. These location-based clusters then defined the training and testing criteria. This strategy used wells from two chosen clusters for testing and the remaining cluster for training. Furthermore, the "leave one well out" cross-validation technique was used to assess the model's performance within each cluster. With fewer data point, this approach made it easier to evaluate the model's robustness within a geographically limited area. Stated differently, every well in a cluster was treated as a

blind well in turn, allowing us to compare the model's performance in situations with limited available data.

➤ *Data Exploration*

An important stage in the process is called exploratory data analysis (EDA), during which the datasets are examined and deconstructed to highlight their key statistical features, often with the aid of data visualization techniques. With the aid of EDA, the datasets are better managed to obtain the required answers, facilitating the process of identifying patterns, and irregularities, testing assumptions, and verifying hypotheses.

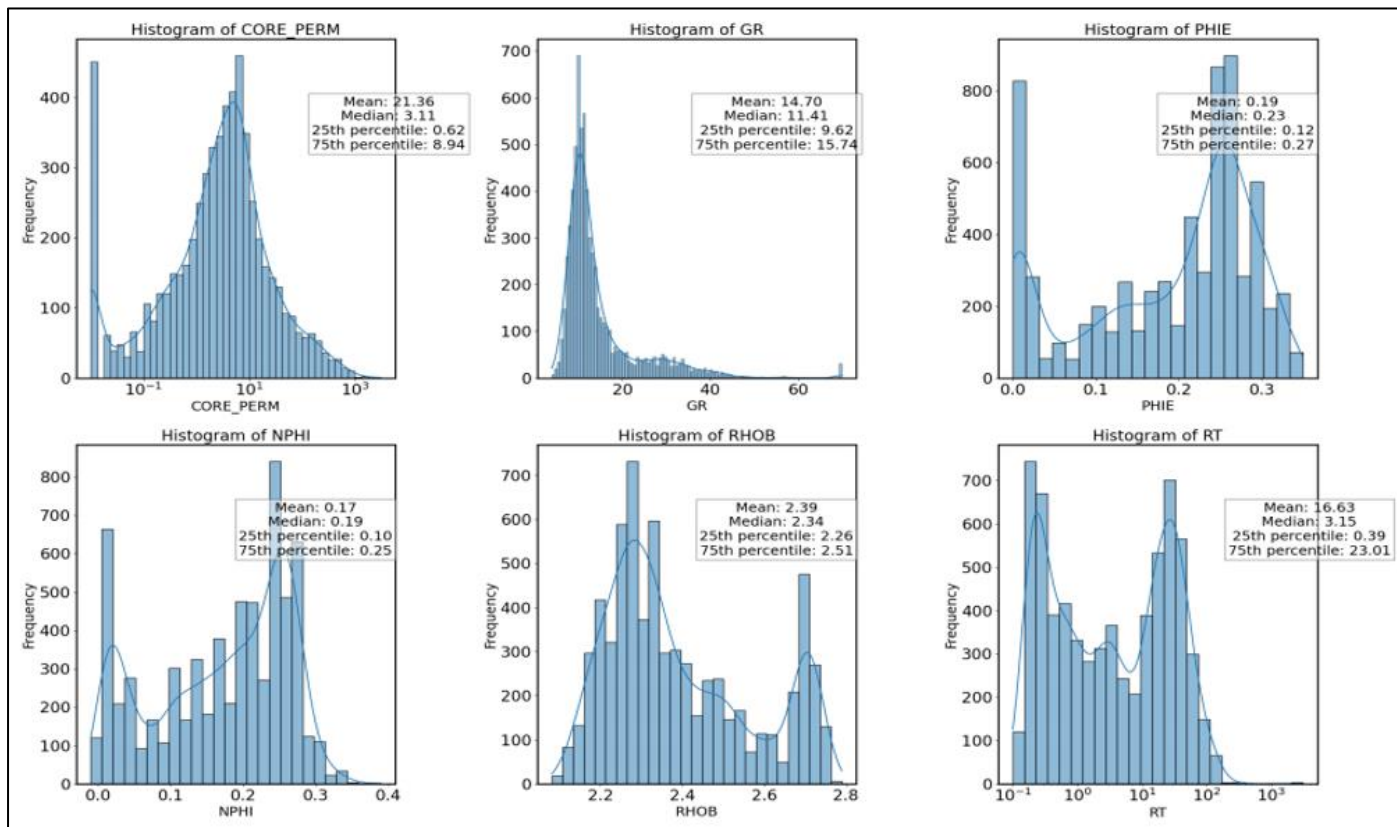


Fig 3 Summary Statistics of the Y-Field Dataset

➤ *Features Evaluation*

Generally speaking, the input features employed for a model's training and validation greatly influence how well it performs (Janiesch et al., 2021). To forecast permeability, identifying the most pertinent features is essential. The significance of feature quality lies in the fact that subpar features might lead to redundancy and complicate the interpretation of the model. The Pearson correlation coefficient (p), which evaluates linear correlations particularly, was used in the investigation to assess the correlation between various characteristics and log-transformed permeability (Cohen et al., 2009). Although this strategy might not fully capture non-linear correlations, every feature in the model was used to guarantee a thorough analysis.

The relevance criterion was set at p values of around ± 0.6 , which indicates a reasonably strong link between the feature and log-transformed permeability. When examining certain features, the largest associations were seen with PHIE, NPHI, and RHOB, with p -values of roughly 0.72, 0.68, and -0.68, respectively. A moderate correlation was seen by GR, with a p -value of about -0.43. NDS and $\log_{10}RT$ showed a modest connection with log-transformed permeability, with p values of about -0.23 and -0.06, respectively, despite their significance in locating hydrocarbon pay zones. Nevertheless, as can be seen in (Figure 4), every feature was included in the finished model.

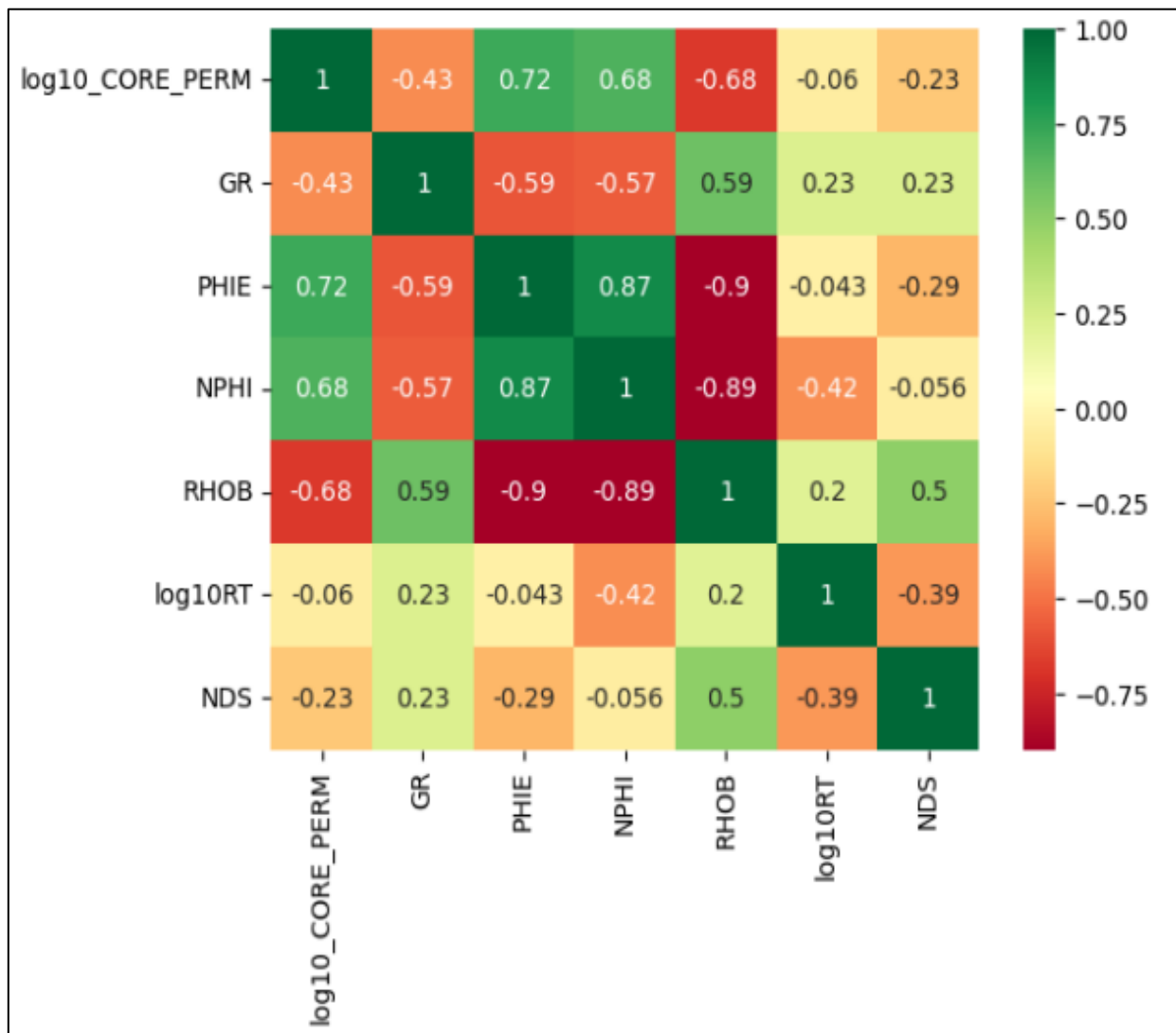


Fig 4 Pearson Correlation Matrix for the Y-Field Dataset.

➤ *Leave One Well Out Validation Results*

This section discusses the model findings obtained from the Y-field dataset, which show a discernible improvement throughout the course of the investigation (Figure 5). With a mean R2 score of 0.368, it is evident from the base model performance—where feature engineering is not used—that the SVM model yields the most accurate predictions. This suggests that SVM was the most effective base model for identifying the underlying relationships in the Y-field dataset.

Feature engineering was implemented, and all of the models' performances showed significant gains. The Random Forest and SVM models (Figure 5a) gained the most from this.

Specifically, Random Forest showed the greatest improvement in mean R2, rising from 0.281 to 0.432. This suggests that the designed characteristics were successfully encapsulating the fundamental data structures, which in turn improved the Random Forest's capacity to forecast core permeability based on well logs. (Bilding 4.4).

When the feature-engineered models' hyperparameters were adjusted, more gains in the models' predictive power were seen (Figure 5c). The mean R2 scores of the XGBoost and LightGBM models showed notable improvements. Among all the models, the tweaked LightGBM model had the greatest mean R2 of 0.510, indicating that feature engineering and the right hyperparameters allowed the LightGBM model to produce the most accurate permeability predictions for the Y-field dataset. This emphasizes how crucial of model optimization is in geology jobs via machine learning.

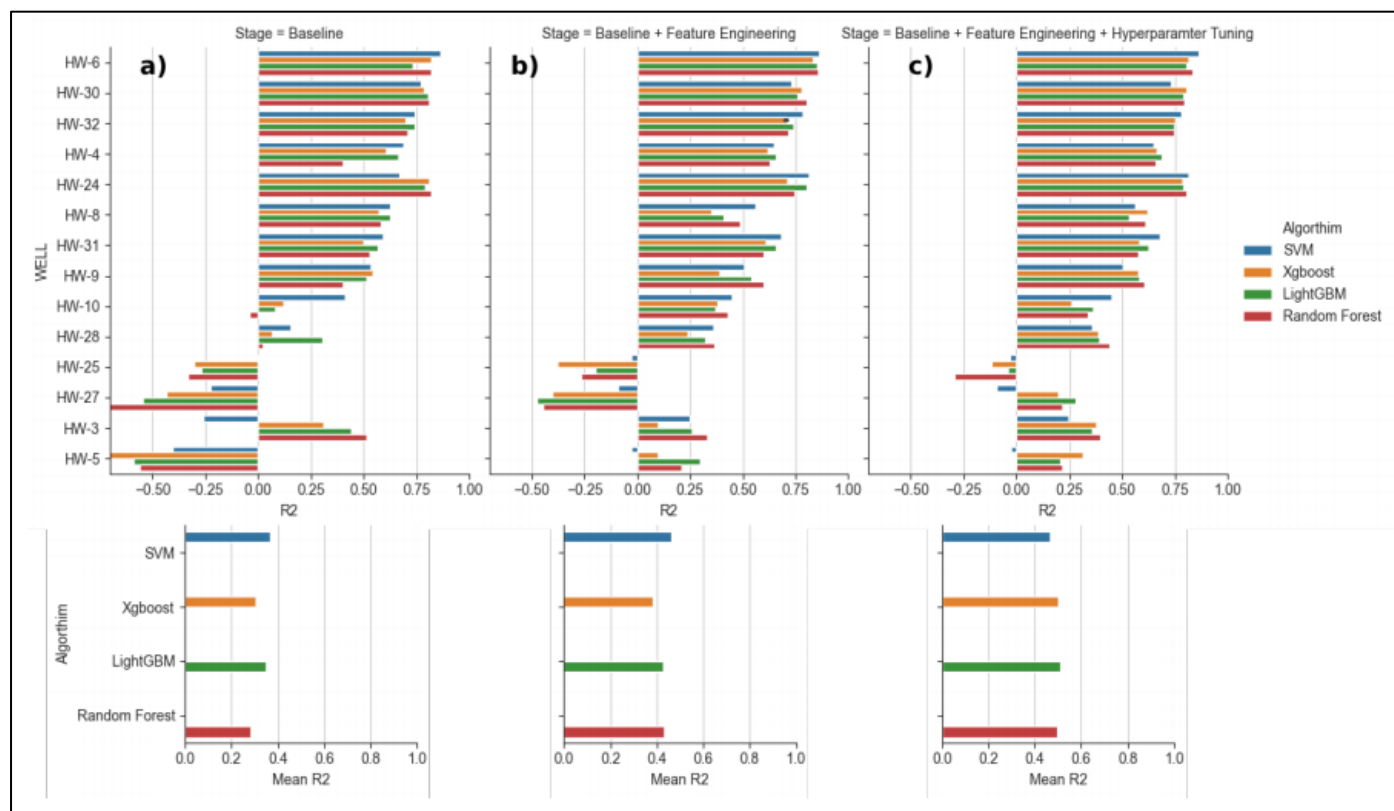


Fig 5 Comparative Validation Results at Various Processing Levels

It is important to note that one well in this figure performs poorly.

During the course of the investigation, some wells showed a notable change in performance, especially when assessed using particular algorithms. For example, HW-10 showed a significant improvement when evaluated using the Random Forest model. It started with a baseline stage R2 score of -0.0422, but once feature engineering and hyperparameter adjustments were integrated, the score shot up to 0.3377. In a similar vein, HW-27 showed an impressive improvement in its R2 score after optimization, going from an initial -0.5447 to 0.2820 when assessed using the LightGBM model. In conclusion, the Xgboost model for HW-5 demonstrated a noteworthy change in R2 score from -0.8357 to 0.3160. These notable advancements highlight the significant role that feature engineering and model optimization play in boosting machine learning's capacity for prediction.

This study shows that machine learning models can help predict core permeability from well logs in carbonate rocks, like those found in the Y-field dataset, when feature-engineered and hyperparameter-tuned. This further emphasizes the potential of machine learning as a useful tool in petrophysics, and future studies could explore other optimization techniques to further enhance prediction accuracy.

➤ *Sensitivity Analysis of Vertical Variability*

The performance of all four machine learning models is significantly impacted by the vertical variability feature

engineering technique, as demonstrated by the sensitivity analysis of the technique (Figure 6). The predictive capability of the model appears to be improved by including values from both above and below the data points. This suggests that when estimating core permeability from well logs, stratigraphic consistency and spatial continuity are important considerations.

The findings show that the mean R2 value climbs with increasing vertical variability (from 0 to 3 meters), suggesting improved model performance. Notably, this feature engineering strategy was especially beneficial to the Support Vector Regression (SVR) model, as demonstrated by a noteworthy increase in mean R2 from 0.3891 (with no vertical variability) to 0.4941 (with 3 meters of vertical variability). This improvement implies that the SVR is more capable of capturing the stratigraphic dependencies in the permeability data by adding vertical variability.

Furthermore, confirming its efficacy, the Random Forest Regressor and the Light Gradient Boosting Machine both performed better when vertical variability was added. However, the Extreme Gradient Boosting model appears to gain less from this, perhaps because the additional features are not properly utilized by the model's inherent setup. This demonstrates that whereas vertical variation generally enhances model performance, the degree of benefit varies depending on the model.

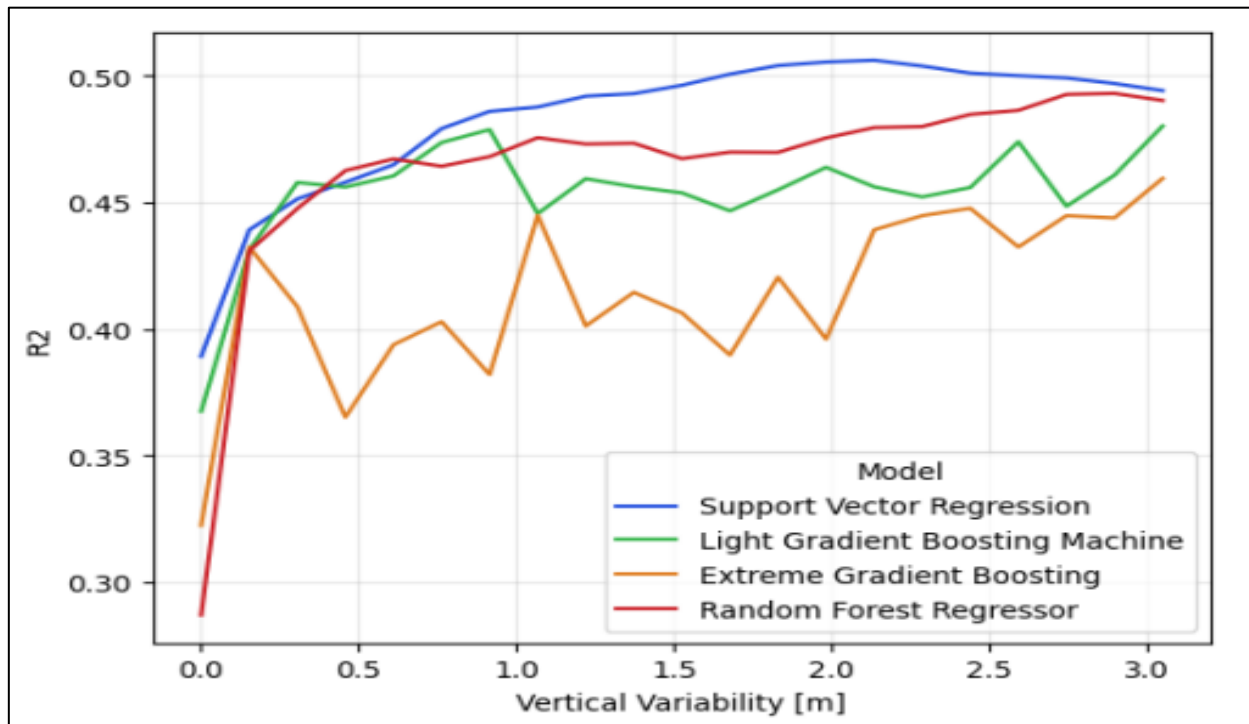


Fig 6 Vertical Variability Sensitivity Analysis.

➤ *Blind Testing Results on The Test Wells*

When assessing how generalisable machine learning models are, blind testing is essential. Wells that were not included in the training dataset are referred to as blind tests in this study. The effectiveness of the feature-engineered and hyperparameter-tuned models in forecasting unseen data is being tested on these wells. Logs from these tests are displayed in Figures 7, 8, and 9. The average R2 value, which

is over 0.6 as shown in Figure 10, is a positive finding. This implies that not only are the improved models fitting the training set, but they are also making good generalizations to brand-new, untested data. To emphasize that all results displayed are from feature-engineered and hyperparameter-tuned models, each result is colour-coded according to the particular machine-learning model that was employed.

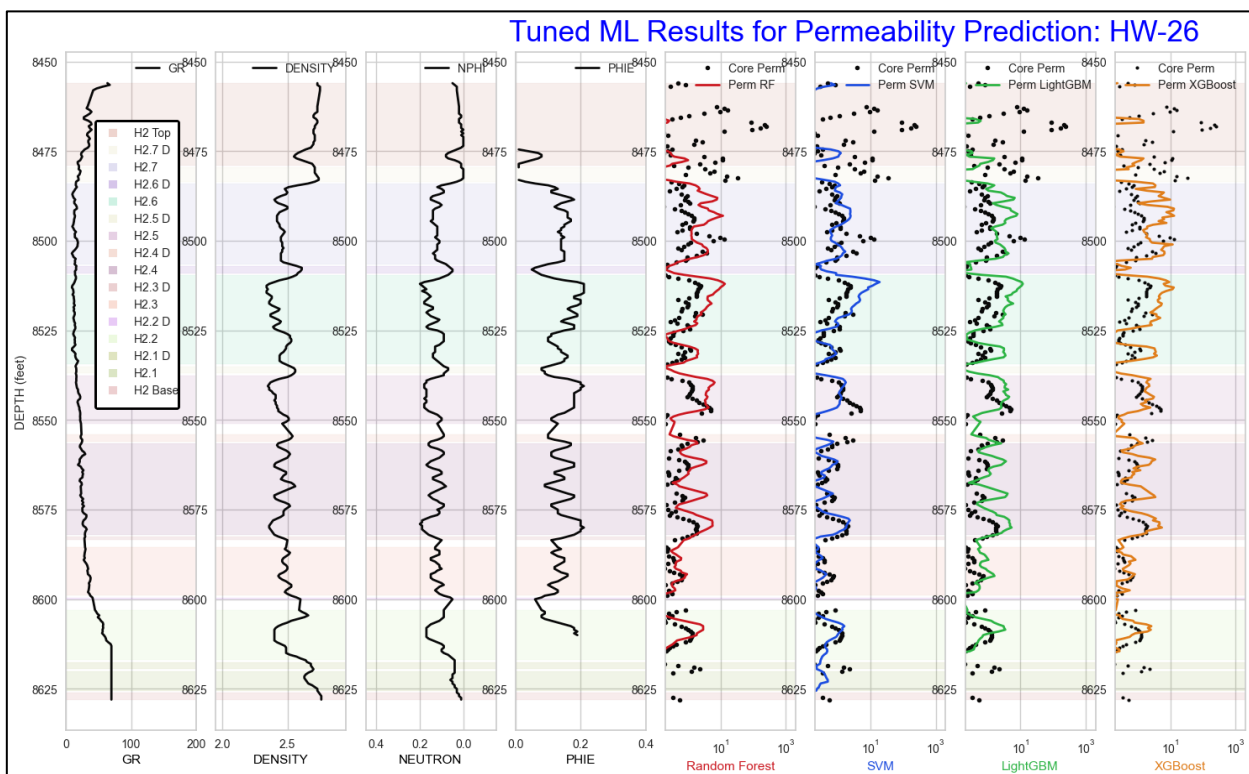


Fig 7 Prediction Results for Blind WELL: HW-26.

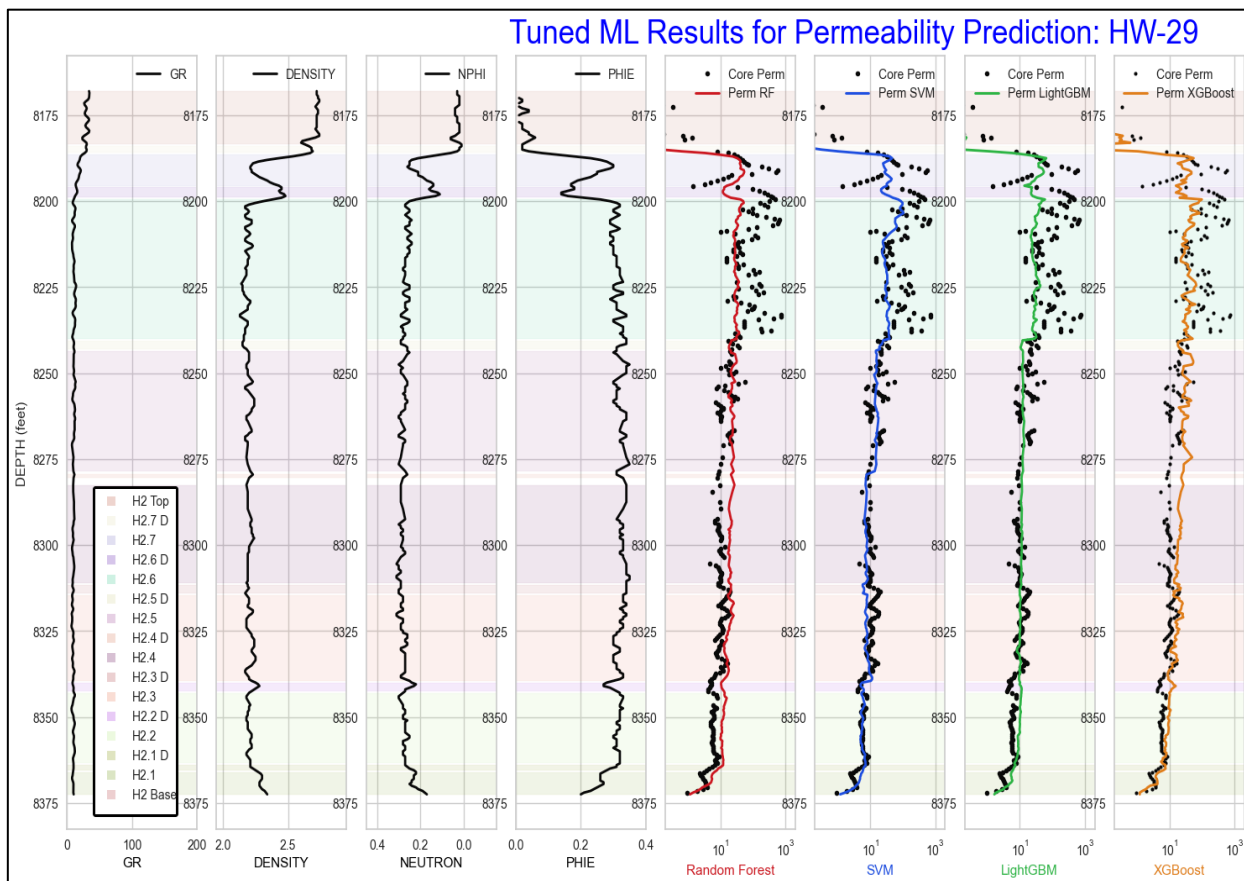


Fig 8 Prediction Results for Blind WELL: HW-29.

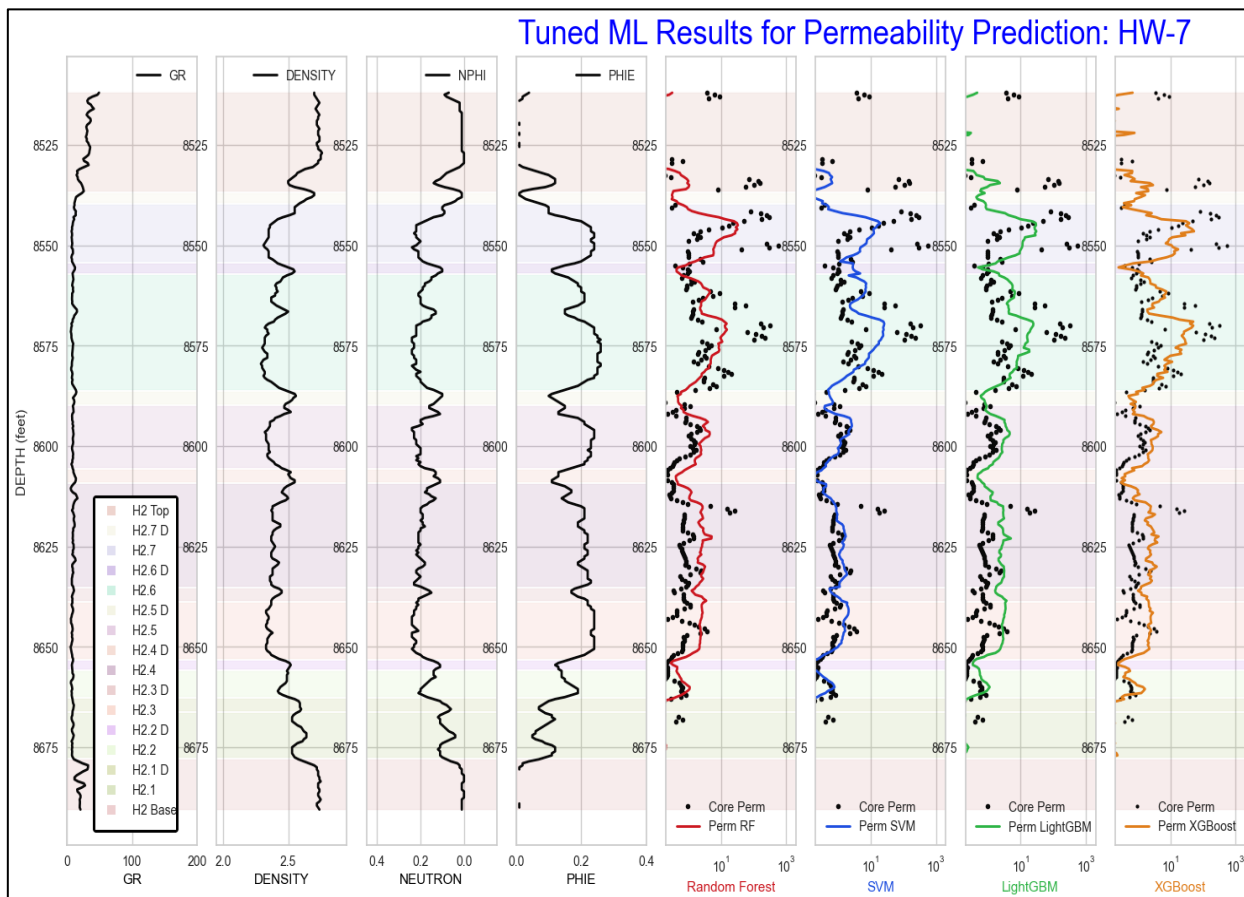


Fig 9 Prediction Results for Blind WELL: HW-7.

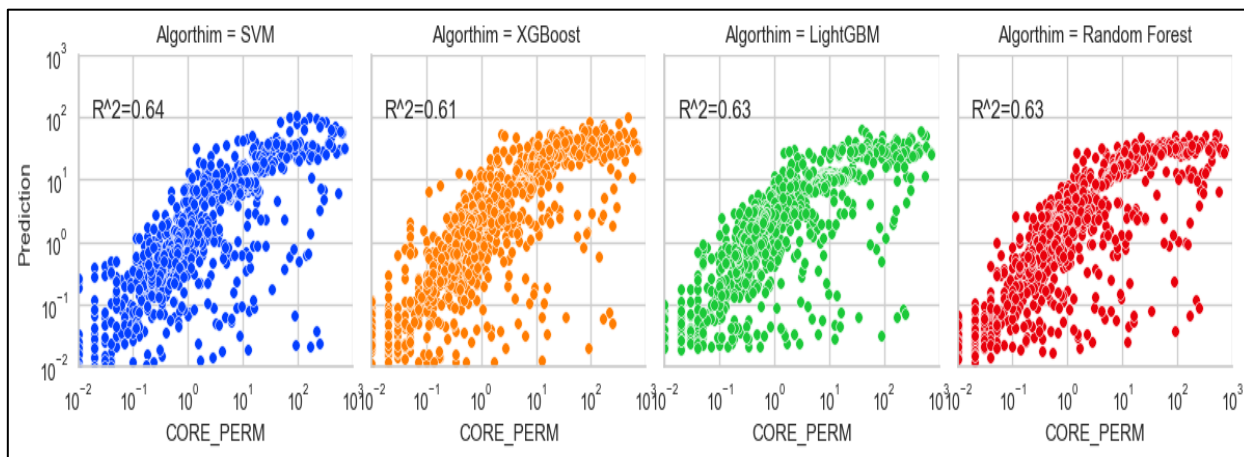


Fig 10 Scatter Plots of Predicted Permeability vs. Measured Permeability for the Blind Test Wells.

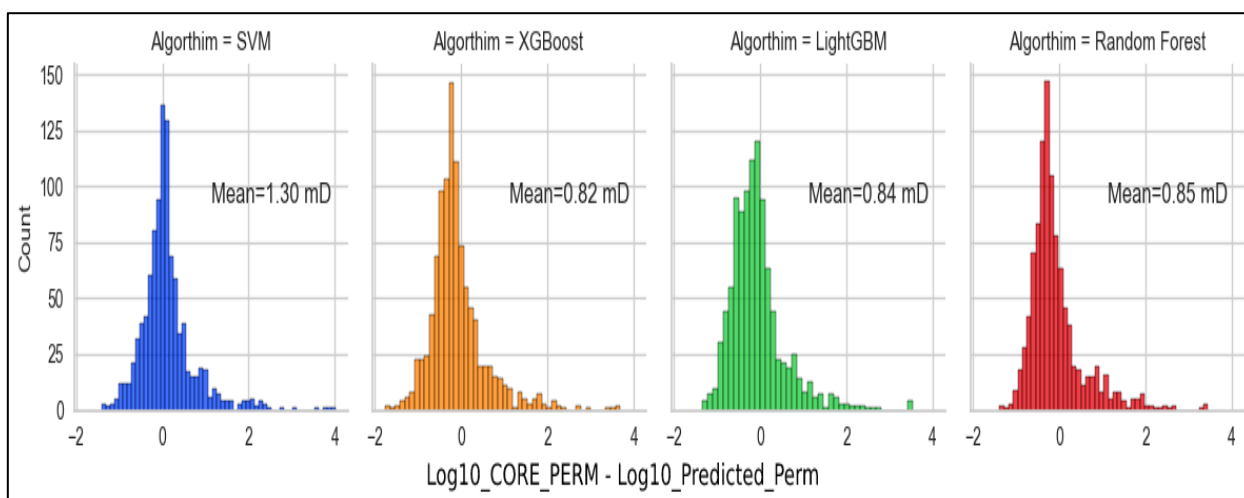


Fig 11 Histograms Displaying the Absolute Error Between Predicted Permeability and Measured Permeability for the Blind Test Wells.

➤ *Geographical Location Clustering Results*

The carbonate platform was split into three separate geographic clusters for this study's validation technique: northeast (Cluster1), southeast (Cluster2), and southwest (Cluster 3). As shown in (Figure 12), these clusters offer an

important testing ground for assessing the generalization capacities of machine learning models under geological fluctuations. The results have a significant impact on these models' predicted dependability.

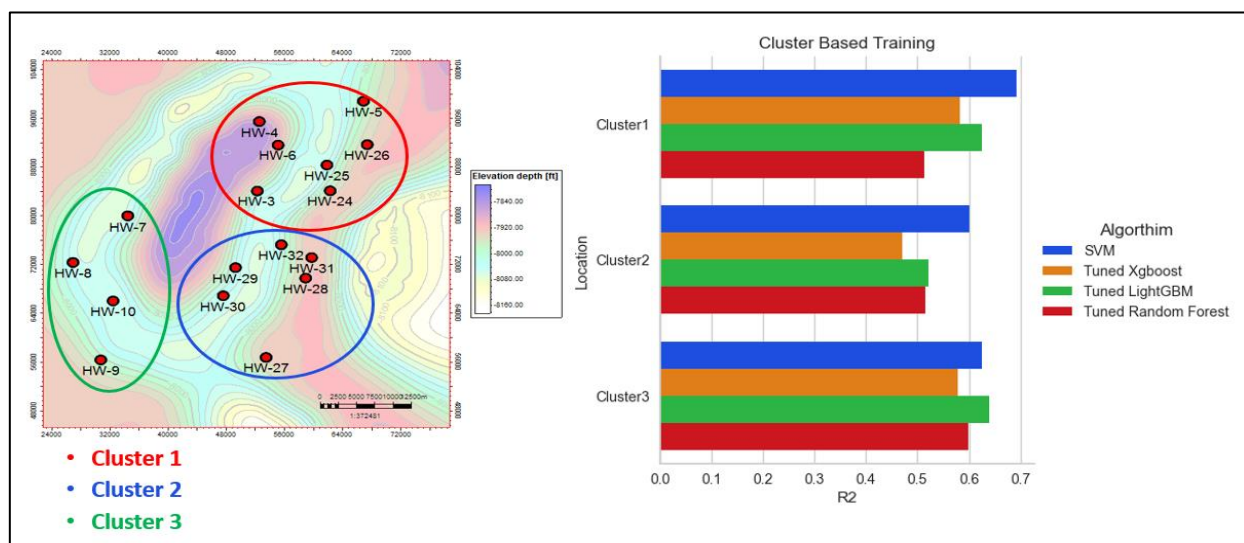


Fig 12 Tuned Prediction Results Using Clusters of Wells to Predict a Cluster.

Throughout the three geographically dispersed clusters in this experiment, the Support Vector Machine (SVM) model consistently showed the most resilient performance. In every situation, it registered the greatest R2 values (Figure 12). The consistent performance of SVM indicates that it is resistant to changes in geology, making it an effective tool for predicting permeability in a variety of geologic settings within the carbonate platform. SVM performs better than other methods, as shown in Figure 12. This is probably because it is better at managing smaller datasets, which are more common in geological studies.

In contrast, as shown in (Figure 12), the results of the Tuned LightGBM, Tuned Xgboost and Tuned Random Forest models showed a high level of variability among the three clusters. Optimized LightGBM performed best in Clusters 3 and 1, which most likely have similar geographical features. When applied to Cluster 2, however, its effectiveness decreased, indicating possible constraints in generalizing across different geological formations. In addition, the Tuned Xgboost model performed worse in Clusters 2 and 3 than it did in Cluster 1, suggesting that there

may be limitations to its ability to adapt to different geological environments.

With the lowest R2 values, the Tuned Random Forest model demonstrated with each unique geological cluster acting as a test set. It also emphasizes how various machine learning algorithms may produce the best results depending on a range of factors, including dataset size and the geographical heterogeneities that are recorded in the data. Consequently, it is imperative to investigate several models in various scenarios and choose the model that yields the highest mean metric in these various scenarios. This technique is supported by the analysis and findings shown in (Figure 12). Each cluster's "leave-one-well-out" cross-validation provides insightful information about the performance of each model. With this approach, every well in a single cluster was trained, except one, which was subsequently evaluated for predictive accuracy using the R2 statistic the least consistency across the three clusters. This discrepancy, as shown in (Figure 12), implies that although Random Forest is an effective tool for managing intricate non-linear interactions and avoiding overfitting, it may have trouble generalizing when trained on smaller datasets.

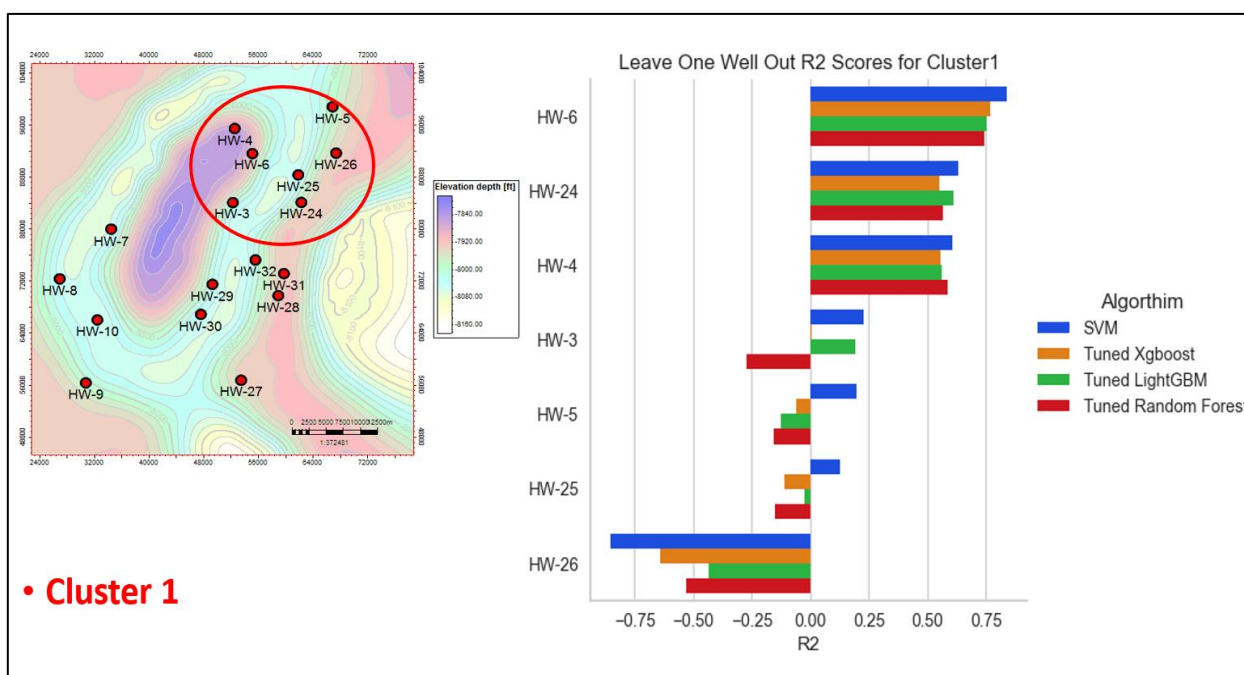


Fig 13 Tuned Prediction within Cluster 1.

According to the statistics, the performance of Tuned Random Forest, Tuned Xgboost, Tuned LightGBM, and SVM varies depending on the specific wells within the clusters. This fluctuation highlights how geological heterogeneities and geographic correlation affect permeability prediction.

For instance, low R2 values indicate that wells like HW-25, HW-26, HW-5, and HW-3 continuously underperformed across all models within Cluster 1 (Figure 13). This persistent underperformance raises the possibility that there are local heterogeneities present, which the models may find difficult to detect because of the limited training set. These heterogeneities could include sudden changes in the properties of the rock, variations in the saturation of hydrocarbons, or other geological characteristics that affect the estimates of permeability.

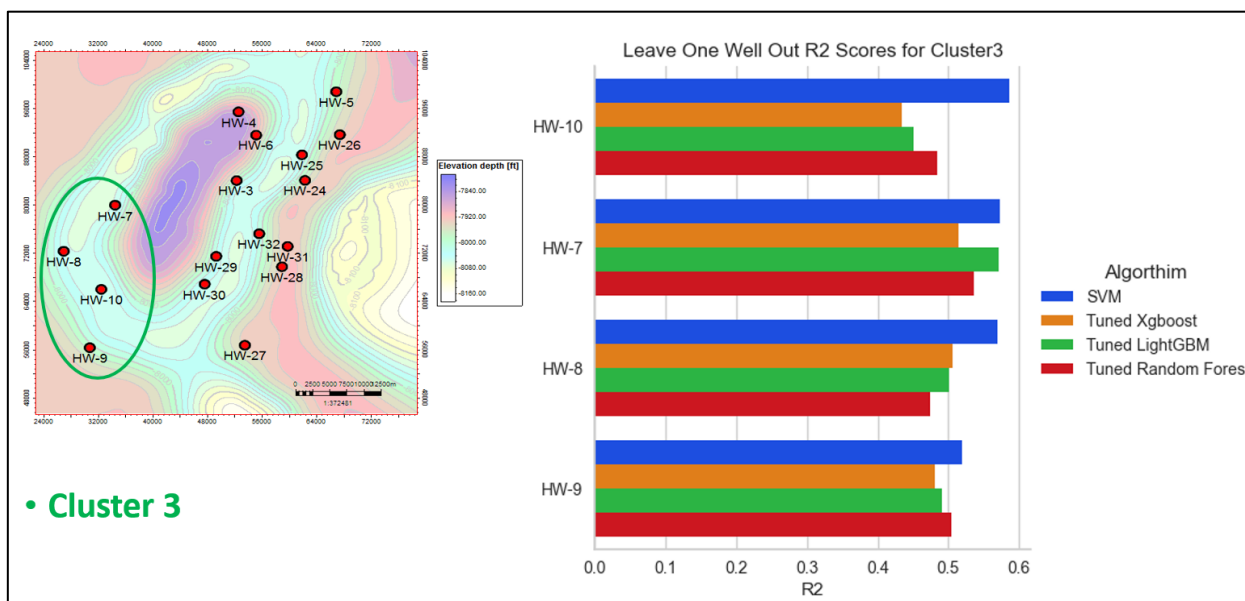


Fig 14 Tuned Prediction within Cluster 3.

On the other hand, Cluster 3 did well overall (Figure 14), especially for the wells that were not previously mentioned, suggesting a significant level of spatial correlation between these wells. The cluster's well-rounded performance highlights the efficacy of machine learning models in high-spatial correlation scenarios.

Comparably, well HW-27 in Cluster 2 (Figure 15) likewise performed poorly in all modes, suggesting the existence of local heterogeneities. This observation highlights the enormous impact that singular outliers or wells possessing distinct geological characteristics can have on the prediction capacity of machine learning models.

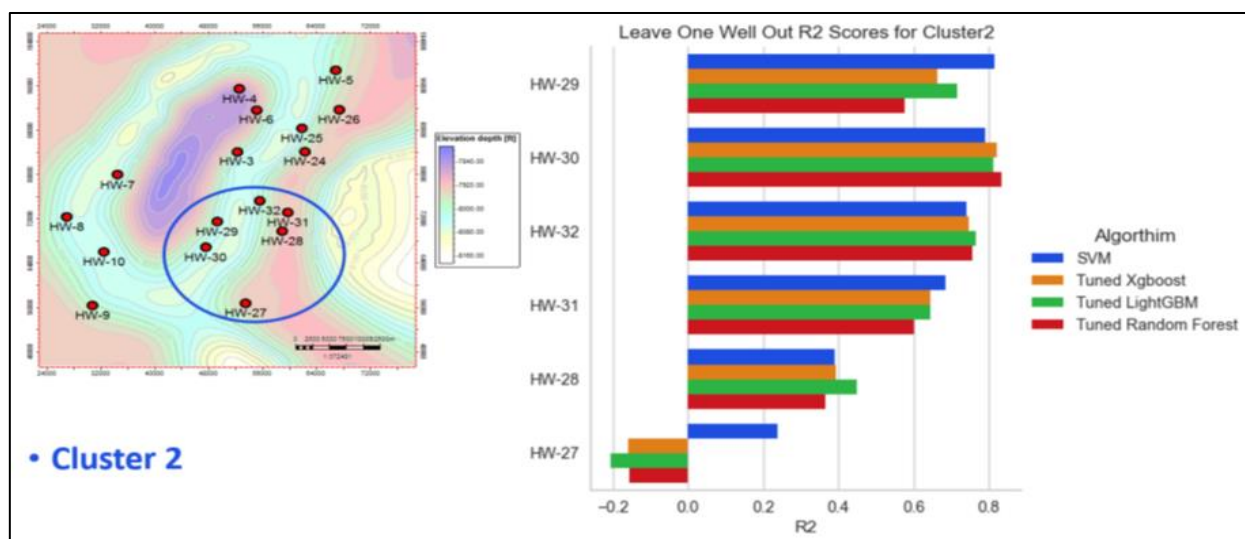


Fig 15 Tuned Prediction within Cluster2

➤ *SHAP (SHapley Additive exPlanations)*

The XGBoost machine learning model was utilized in this investigation to forecast permeability, offering noteworthy perspectives on the significance of several attributes, as measured by Shapley Additive Explanations (SHAP) values. The average marginal contribution of each feature to the model's prediction, across all potential feature combinations, is measured by SHAP values.

Given its high mean SHAP value of 1.065, the XGBoost model's permeability predictions were significantly influenced by the Gamma Ray (GR) feature. The Upper Kharaib Member's concentration of stylolites or the influence

of shaly interbeds could both be explained by the GR log, which offers information on lithology and shale composition. Effective Porosity (PHIE) had a high mean SHAP value of 1.395, indicating significant influence. The XGBoost model effectively captures the variability brought about by varying degrees of pore system interconnectivity as well as the impact of porosity on permeability. The XGBoost model had a significant influence on the DEPTH feature on permeability predictions, as seen by its mean SHAP value of 1.199. In line with the knowledge of how depth affects compaction and diagenesis, the model found a relationship between permeability and depth.

As demonstrated in (Figures 16 and 17), a SHAP contribution plot visualizes the impact of each feature on the model's prediction for each instance, with the x-axis representing the SHAP value (contribution to total permeability) and each point representing a specific instance. This helps us understand how the presence or value of a

feature influences the prediction. X_LOG and Y_LOG, representing spatial coordinates of well logs, exhibited a moderate influence on the permeability predictions, suggesting spatial variability in depositional and diagenetic events within the reservoir.

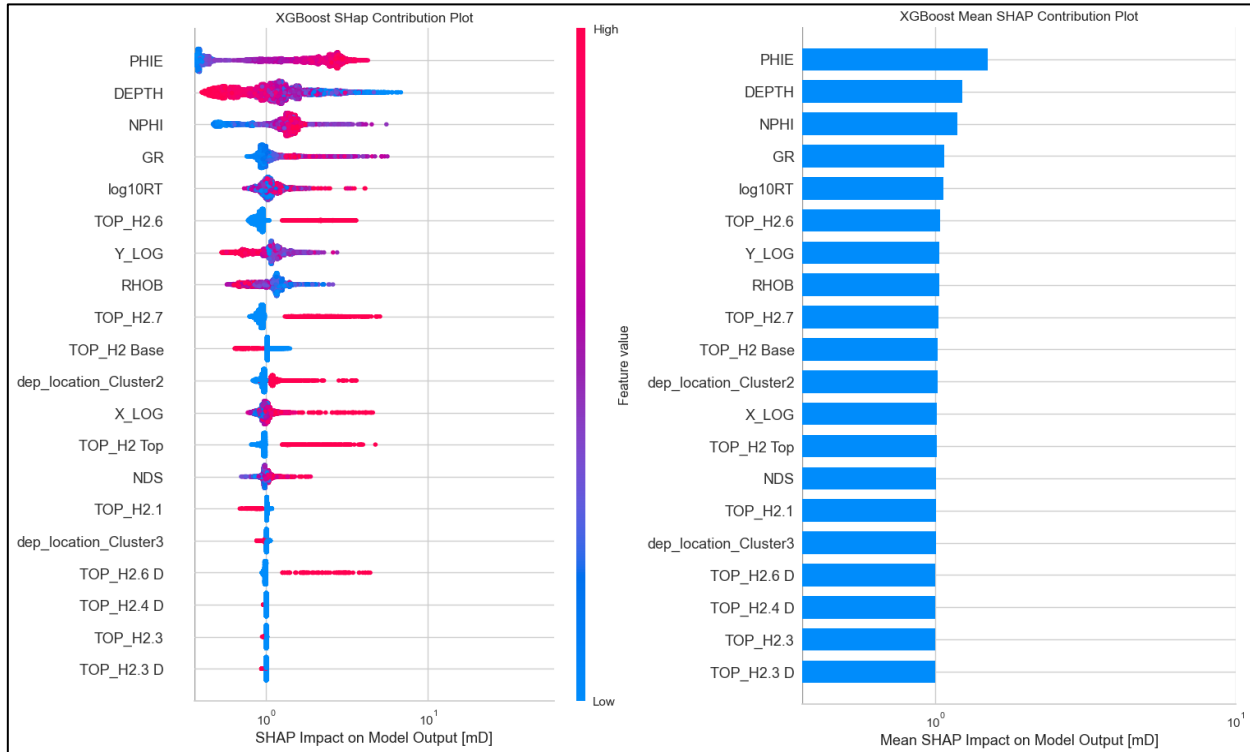


Fig 16 SHAP Summary Plots for the XGBoost Model Using the Y-Field Dataset.

In the plot, each dot represents a data point. The x-axis indicates the SHAP value, which shows the impact of the feature on the model's prediction. Features are listed on the y-axis, with the most influential feature at the top. The colour

of the dots represents the feature's value, with red indicating higher values and blue indicating lower values. A dot's position on the x-axis shows whether the feature pushes the prediction higher (to the right) or lower (to the left).

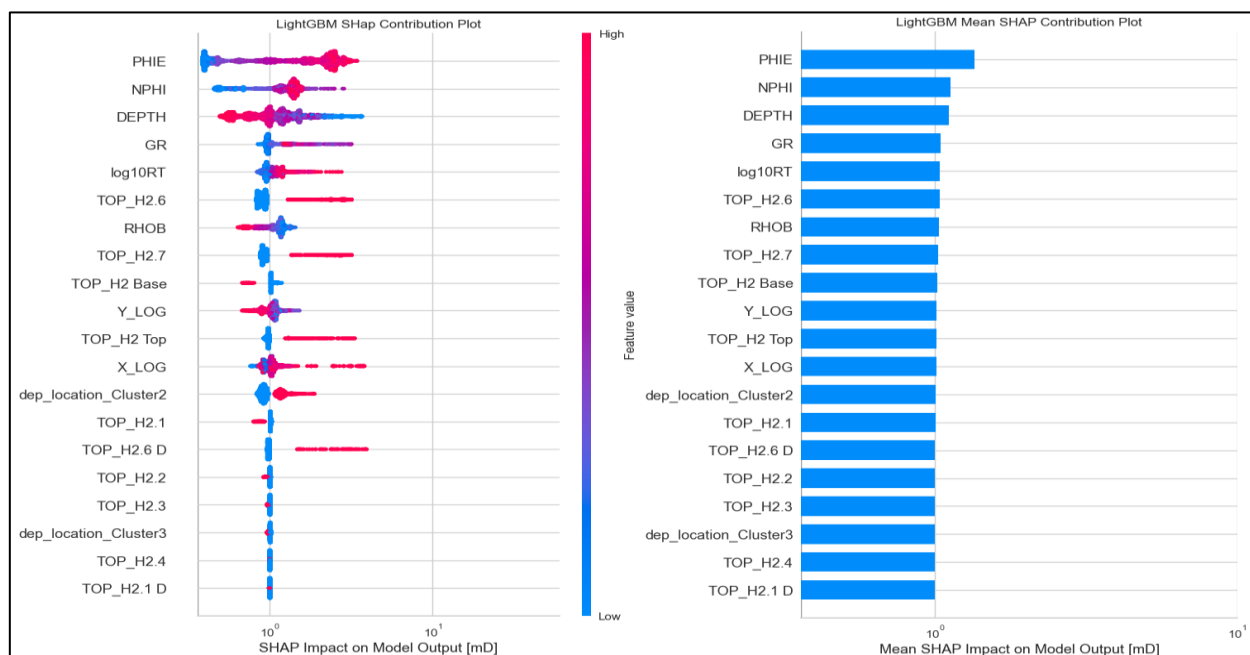


Fig 17 SHAP Summary Plots for the Lightgbm Model Using the Y-Field Dataset.

V. DISCUSSION

The process of measuring core permeability in a laboratory is rife with uncertainties that arise from a variety of sources (McPhee et al., 2015). These uncertainties include the inherent heterogeneity of core samples, potential alterations from sample handling and preparation, the precision of instruments, the methodologies used, and the particular conditions of the laboratory, such as stress and temperature. Furthermore, variables such as pore connectivity, tortuosity, and potential alterations in rock fabric due to chemical reactions or physical changes during drilling can affect the results. According to MCPhee et al. (2015), misinterpretations and mistakes might result from inaccuracy in matching core measurements to the proper depth in the wellbore, which presents a substantial problem. It is crucial to understand that these laboratory uncertainties—including depth alignment errors—are intrinsically carried over when using machine learning techniques to propagate permeability as a well log. This may have an impact on how well predictions correspond with in-situ data, which can affect prediction reliability.

Porosity-based variables are frequently given priority in permeability prediction by both machine learning techniques and conventional petrophysical methods such as the Tixier, Timur, and Coates models (Figure 15). Conventional techniques depend on parameters generated from the core, such as the cementation factor and saturation exponent, as well as well log data, such as resistivity and porosity (Mohaghegh et al., 1997). However, in diverse reservoirs, their applicability may be restricted, frequently necessitating zone-specific modifications. Machine learning, on the other hand, not only evaluates these standard characteristics but also integrates a varied variety of information, delivering refined forecasts for each particular geological context.

As this studies and the literature have shown, feature engineering is essential to the creation of reliable and effective models in petrophysics. The performance of machine learning models can be greatly improved by the process of generating new features from preexisting ones. A lot of attention has been focused on feature engineering in this work, and the particular requirements of the problem at hand, as well as the characteristics of the data, have influenced the decisions. Adding Vertical Variability was one of the main strategies employed to take into consideration the unique qualities of well logs and core samples. The volume captured by well logs is seen in Figure 4.35, thereby smoothing out local variation. On core samples, the smoothing effect is substantially weaker, and as a result, measurements tend to reveal greater variability at smaller scales. The performance metrics of the models shows that the prediction outcomes were greatly enhanced by integrating Vertical Variability.

VI. CONCLUSIONS

➤ *Effectiveness of Machine Learning Models*

The study effectively compared the performance of various machine-learning models in predicting permeability from well logs. Tree-based models, particularly Random

Forest and XGBoost, demonstrated superior accuracy and robustness. The blind test R2 scores in the Y-field dataset reached up to 0.74, while validation R2 scores for individual wells increased to 0.8. These results underscore the models' high accuracy and their potential for practical applications in reservoir characterization.

➤ *Impact of Feature Engineering*

Feature engineering was a crucial component of the study, significantly enhancing the performance of the machine learning models. By incorporating features like Vertical Variability, the models could better account for the inherent differences between well logs and core samples. This approach led to more accurate permeability predictions and highlighted the importance of tailored feature engineering in improving model performance.

➤ *Hyperparameter Tuning*

The study highlighted the importance of hyperparameter tuning in achieving optimal model performance. Proper tuning of hyperparameters resulted in substantial improvements in R2 scores, demonstrating the critical role this process plays in refining machine-learning models. The significant performance boost observed post-tuning underscores the necessity of meticulous hyperparameter optimization in developing reliable predictive models.

➤ *Generalisability and Robustness*

The models' performance across the Y-field datasets demonstrated their generalisability and robustness. With R2 scores ranging from 0.74 to 0.84 in the Y-field dataset, the models proved capable of adapting to different geological contexts. This versatility is crucial for practical applications, ensuring reliable predictions across various reservoirs and formations.

➤ *Interpretability of Models*

The study emphasized the importance of model interpretability, particularly in the context of petrophysical applications. Tree-based models, such as Random Forest and XGBoost, not only provided high prediction accuracy but also allowed for easier interpretation of the decision-making processes. The use of interpretability methods, like SHAP, provided valuable insights into the factors influencing the models' predictions, thereby enhancing their transparency and trustworthiness.

REFERENCES

- [1]. Aigbedion, I. (2007). A case study of permeability modeling and reservoir performance in the absence of core data in the niger delta, nigeria. *Journal of Applied Sciences*, 7(5), 772–776.
- [2]. Asquith, G. B., Krygowski, D., Gibson, C. R., & Asquith, G. B. (2004). *Basic well log analysis*. American Association of Petroleum Geologists.
- [3]. Brereton, R. G., & Lloyd, G. R. (2010). Support vector machines for classification and regression. *The Analyst*, 135(2), 230–267.
- [4]. Carcione, J. M., Gei, D., Picotti, S., Misnan, M. S., Rashidi, M. R., Bakar, Z. A., Harith, Z. Z., Bahri, N.

- H., & Hashim, N. (2020). Porosity and permeability of the overburden from wireline logs: A case study from offshore malaysia. *Geomechanics and Geophysics for Geo-Energy and Geo-Resources*, 6(3).
- [5]. Chaya. (2022). Random forest regression. <https://levelup.gitconnected>.
- [6]. Chen, T., & Guestrin, C. (2016). XGBoost. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*.
- [7]. Cohen, I., Huang, Y., Chen, J., Benesty, J., Benesty, J., Chen, J., Huang, Y., & Cohen, I. (2009). Pearson correlation coefficient. *Noise reduction in speech processing*, 1–4.
- [8]. Costa Gomes, J., Geiger, S., & Arnold, D. (2022). The design of an open-source carbonate reservoir model. *Petroleum Geoscience*, 28(3).
- [9]. Crampin, T. (2008). Well log facies classification for improved regional exploration. *Exploration Geophysics*, 39(2), 115–123.
- [10]. Deegan, C. E., & Scull, B. J. (1977). A standard lithostratigraphic nomenclature for the central and northern North Sea. UK Institute of Geological Sciences, Report, 1, 36.
- [11]. Diniz Ferreira, E. L., & Torres-Verdin, C. (2012). Improved estimation of pore connectivity and permeability in deepwater carbonates with the construction of multi-layer static and dynamic petrophysical models.
- [12]. Elkatatny, S., Mahmoud, M., Tariq, Z., & Abdulraheem, A. (2017). New insights into the prediction of heterogeneous carbonate reservoir permeability from well logs using artificial intelligence network. *Neural Computing and Applications*, 30(9), 2673–2683, Equinor. (2022). Volve field data set.
- [13]. Fadairo, A., & Awuyo, O. (2019). Achieving best practices in log pre-processing for facies and permeability modeling. *International Journal of Mechanical Engineering and Technology*, 10.
- [14]. Feng, R., Grana, D., & Balling, N. (2021). Imputation of missing well log data by random forest and its uncertainty analysis. *Computers & Geosciences*, 152, 104763.
- [15]. Gholanlo, H. H. (2021). Analysis of permeability based on petrophysical logs: Comparison between heuristic numerical and analytical methods. *Journal of Petroleum Exploration and Production Technology*, 11(5), 2097–2111.
- [16]. Gray, E., Hartley, A., & Howell, J. (2020). The influence of stratigraphy and facies distribution on reservoir quality and production performance in the triassic skagerrak formation of the uk and norwegian central North Sea. *Geological Society, London, Special Publications*, 494(1), 379–409.
- [17]. Grinsztajn, L., Oyallon, E., & Varoquaux, G. (2022). Why do tree-based models still outperform deep learning on tabular data?
- [18]. Han, D., & Kwon, S. (2021). Application of machine learning method of a data-driven deep learning model to predict well production rate in the shale gas reservoirs. *Energies*, 14(12), 3629.
- [19]. Hansen, S. (1996). A compaction trend for cretaceous and tertiary shales on the Norwegian shelf based on sonic transit times. *Petroleum Geoscience*, 2(2), 159–166.
- [20]. He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition.
- [21]. Ibrahim Ahmed Osman, A., Najah Ahmed, A., Chow, M. F., Feng Huang, Y., & El-Shafie, A. (2021). Extreme gradient boosting (xgboost) model to predict the groundwater levels in Selangor malaysia. *Ain Shams Engineering Journal*, 12(2), 1545–1556.
- [22]. Katahara, K. (2008). What is shale to a petrophysicist? *The Leading Edge*, 27(6), 738–741.
- [23]. Kavzoglu, T., & Teke, A. (2022). Advanced hyperparameter optimization for improved spatial prediction of shallow landslides using extreme gradient boosting (xgboost). *Bulletin of Engineering Geology and the Environment*, 81(5).
- [24]. Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., & Liu, T.-Y. (2017). Lightgbm: A highly efficient gradient boosting decision tree. *Advances in Neural Information Processing Systems 30 (NIP 2017)*.
- [25]. Kieft, R. L., Jackson, C. A.-L., Hampson, G. J., & Larsen, E. (2010). Sedimentology and sequence stratigraphy of the hugin formation, quadrant 15, norwegian sector, south viking graben. *Geological Society, London, Petroleum Geology Conference Series*, 7(1), 157–176.
- [26]. Kramer, O. (2016). Scikit-learn. In *Machine learning for evolution strategies* (pp. 45–53). Springer International Publishing.
- [27]. Kumar, A. (2022). Xgboost vs lightgbm.
- [28]. Strohmenger, C. J., Weber, L. J., Ghani, A., Rebelle, M., Al-Mehsin, K., Al-Jeelani, O., Al-Mansoori, A., & Suwaina, O. (2004). High-resolution sequence stratigraphy of the kharab formation (lower cretaceous, u.a.e.) All Days.
- [29]. Talebi, H., Peeters, L. J., Otto, A., & Tolosana-Delgado, R. (2021). A truly spatial random forests algorithm for geoscience data analysis and modelling. *Mathematical Geosciences*, 54(1), 1–22.
- [30]. Vapnik, V. (1998). The support vector method of function estimation. In J. A. K. Suykens & J. Vandewalle (Eds.), *Nonlinear modeling: Advanced black-box techniques* (pp. 55–85). Springer US.