MINISTRY OF EDUCATION AND TRAINING
HANOI UNIVERSITY OF MINING AND GEOLOGY

NGUYEN TRUNG DUNG

RESERVOIR CHARACTERIZATION OF MIDDLE MIOCENE CARBONATE RESERVOIR OF CX FIELD

SUMMARY OF THE PhD THESIS

HANOI – 2024

The dissertation has been completed at **Department of Geophysics, Faculty of Petroleum and Energy, University of Mining and Geology**

Scientific supervisor:

1. Assoc. Prof. Dr. Phan Thien Huong
The University of Mining and Geology

2. Dr. Cu Minh Hoang
Petrovietnam Exploration and Production Corporation

Reviewer 1: **Assoc. Prof. Dr. Nguyen Van Phon**
Vietnam Association of Geophysicists

Reviewer 2: **Assoc. Prof. Dr. Hoang Van Long**
Vietnam Union of Geological Sciences

Reviewer 3: **Dr. Ha Quang Man**
Petrovietnam Exploration and Production Corporation

The dissertation will be defended at the University-level Dissertation Evaluation Council Meeting at the University of Mining and Geology at ……hr, dd mm yyy

The dissertation can be found at:

- **National Library**
- **The University of Mining and Geology Library**

INTRODUCTION

Carbonate reservoirs play an important role in the hydrocarbon exploration and production accounting for approximately 60% of the world's hydrocarbon reserves. In Vietnam, despite of limited hydrocarbon discoveries in carbonate reservoirs, significant important discoveries have been made in fields such as Ham rong, Blue Whale, Dai Hung, Phong Lan Dai, Lan Tay, Lan do, Sao Vang Dai Nguyet, Thien Ung and Ca Map Trang. Research on reservoir characteristics aiming to optimize the HIIP calculation, 3D geological model and 3D dynamic model poses numerous challenges.

Due to the complexity of petrology, porosity, and other factors, describing the relationships in the petrophysical model of carbonate rocks is a specific concern for oil companies. Studies indicate that the use of machine learning techniques helps describe reservoir properties with higher reliability than traditional methods.

The CX field, located in the southern part of the Song Hong Basin, on the Tri Ton uplift with Middle Miocene carbonate reservoir, faces challenges due to the high heterogeneity of reservoir rocks. Therefore, research to elucidate the petrophysical parameters of CX carbonate rocks, utilizing comprehensive core sample analysis and wireline logging data, along with modern soft computing methods, including machine learning, is indispensable and meaningful.

1. Overview of the previous reserch status and the new issues posed by this reserch topic in comparison to earlier studies.

1.1 Research situation abroad

Due to the importance of carbonate reservoirs and their complexity, research on describing the characteristics of carbonate reservoirs is highly regarded. Studies by Archie (1952), Dunham (1962), Lucia (1987), G.V. Chilingarian (1992, 1996), Pittman (1971, 1992), Amaefule (1993), and others emphasize the heterogeneity of carbonate reservoirs and the necessity of their classification.

Research published in reputable oil journals such as SPE, AAPG, etc., focuses on assessing risks in water saturation calculations and proposes alternative methods to replace the use of the Archie formula, suggesting different approaches to increase accuracy in predicting permeability in carbonate rocks.

Over time, advancements in science and technology have prompted the application of soft computing tools in describing the characteristics of carbonate reservoirs, utilizing intelligent technology to forecast reservoir properties.

1.2 Domestic Research Situation

Although the number of hydrocarbon discoveries in carbonate reservoirs in sedimentary basins in Vietnam is limited, numerous studies have focused on methods for describing the reservoir characteristics of this rock type. These studies have highlighted the necessity of classifying carbonate reservoirs, especially classifying them based on hydraulic flow units (HFU), along with predicting reservoir rock groups and permeability. However, most of these studies have primarily relied on methods such as linear regression and artificial neural networks.

With the advancement of science and technology, machine learning methods have gained increasing attention and development. Machine learning algorithms not only aid in data grouping but also provide more reliable predictions of reservoir rock groups. In the CX field, although the characteristics of Middle Miocene carbonate reservoirs have been studied, partially reflected in the field's reserve reports, in-depth research has mainly focused on studying the diagenesis processes of the early Miocene carbonate system, related studies on biofacies and depositional environments and etc. In-depth studies related to reservoir rock grouping, permeability determination, and water saturation of the reservoir rock have not yet received sufficient attention.

Therefore, within the scope of this research, the author will explore methods to classify carbonate reservoir rocktyping based on hydraulic flow units, to predict reservoir permeability using machine learning methods, and to apply the results to determine water saturation for the Middle Miocene carbonate reservoirs in the CX field.

2. Research Objectives

Applying machine learning methods for classifying and predicting hydraulic flow unit groups, permeability, and utilizing the results to forecast water saturation in carbonate reservoirs. The outcomes will enhance the efficiency of geological modeling, as well as reservoir dynamic models, improving the accuracy of production forecasts and reserve assessments.

3. Defense Arguments

Argument 1: Applying machine learning is of significant importance in reservoir characterization. Unsupervised machine learning methods have enabled the improvement of efficiency and optimal classification of the Middle Miocene carbonate reservoir into 5 Hydraulic Flow Units within the CX field.

Argument 2: Applying supervised machine learning methods in conjunction with the determined HFU results allows for more accurate prediction of permeability, water saturation and their variations within the Middle Miocene carbonate reservoir in the CX field. The permeability varies widely, ranging from below 1 mD to over 2,000 mD; water saturation (Sw) gradually decreases with the gas column height, reaching up to 6% depending on the HFUs

4. Recent findings in the thesis

Carbonate reservoirs exhibit high heterogeneity, in-depth studies related to describing the characteristics for this reservoir type in Vietnam, in general, and in the Tri Ton high region, in particular, are currently limited. For the CX field, this is the first in-depth study on reservoir characterization.

The application of Machine Learning in reservoir characterization study in Vietnam is relatively novel. The thesis has established a workflow to enhance the results of describing carbonate reservoir characteristics through Machine Learning tools:

- Carbonate reservoir rocktyping based on the hydraulic flow unit (HFU) concept using unsupervised machine learning, including retionale for machine learnining method and optimal number of HFU selection;
- Proposing an optimal workflow for HFU/permeability prediction based on wireline logging and core data using supervised machine learning methods.
- Applying the results of HFU classification and HFU/permeability prediction to contruct water saturation prediction model using saturation height function method for each HFU. This helps minimize uncertainty caused by the influence of Archie parameters on the water saturation calculation results."

5. Database

In the process of conducting the thesis, various types of data were collected:

- Studies related to the characteristics of Middle Miocene carbonate sedimentary features in the southern part of the Song Hong Basin;
- Well-related documents including well logging data, core sample analysis data (comprising both regular core sample analysis and special core sample analysis), cuttings, side-wall core samples, formation test documents, and completion reports for the four wells drilled in the CX field;
- Evaluation reports of the research area.

6. Dissertation structure

The dissertation is presented across 127 pages, 71 figures, 12 tables and 1 appendix. Beyond the introduction and conclusion, the thesis is organized into 05 chapters corresponding to the relevant publications, structured as follows:

• Chapter 1: Provides an overview of the research area, highlighting the characteristics of the Middle Miocene carbonate reservoir in the CX field.

• Chapter 2: Presents the approach and methods of machine learning; application of machine learning methods in describing reservoir characteristics, outlining the research framework; the database used for the research.

• Chapter 3: Presents the results of unsupervised machine learning-based hydraulic flow unit grouping.

• Chapter 4: Presents the implementation of supervised machine learning methods and the results of predicting HFU/permeability based on well logging data and core sample analysis.

• Chapter 5: Presents results of building a water saturation prediction model based on the saturation height function models for each HFU.

CHAPTER 1
OVERVIEW OF THE STUDY AREA

1.1 Location and Geological characteristics of study area

The CX field is located in the southern part of the Song Hong Basin, extending along the Tri Ton uplift in a northwest - southeast direction in the offshore of Da Nang. It is approximately 80 km from the coastline between Quang Nam and Quang Ngai provinces (Figure 1).

1.1.1 Tectono-structural characteristics

The southern part of the Song Hong basin comprises of second-order structural units such as the Danang Shelf, Quang Ngai graben, Tri Ton uplift and East Tri Ton rift, influenced by tectonic factors from major fault zones like the Red River, Ma River and Rao Nay. Among them, the Tri ton uplift is a basement high, spanning over 500 km, covered by Oligocene clastics sediments and Miocene carbonate buildups.

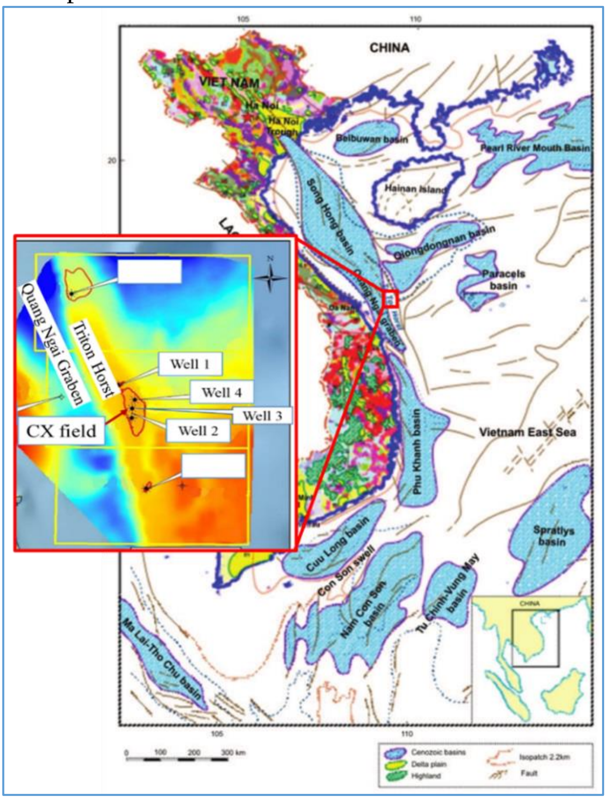


Figure 1.1: CX field location and adjacent wells (on the background map of Kainozoic sedimentary basins in Vietnam - according to Phan Trung Dien)

1.1.2 Litho-Stratigraphy

The stratigraphic succession of Southern part of Song Hong basin comprises of the followings: Pre-Kainozoic basement; Oligocene clastics – Bach Tri formation; Song Huong formation – Lower Miocene; Tri ton formation – Middle Miocene; Quang Ngai formation – Upper Miocene; Pliocene clastics – Quaternary/Bien Dong formation. Among that, the study objective is located in the Tri ton formation.

The Tri Ton Formation unconformably overlies the Song Huong Formation. It includes fine-grained clastic sediments on both sides of the Tri Ton uplift and carbonate rocks, reaching several hundred meters in thickness on the Tri Ton uplift. The depositional environment primarily consists of shallow marine, marine platforms, and coastal plains.

1.2 The characteristics of the Middle Miocene carbonate in CX field

1.2.1 Formation mechanism

The Tri Ton Uplift was formed during the Eocene - Early Oligocene period through the extensional processes opening the Song Hong Basin, giving rise to rifts and uplifts. The tectonic activity of the Early Miocene is characterized by seafloor spreading and the expansion of the East Sea, which led to a reduction in temperature and a rise in sea levels, creating favorable conditions for carbonate formation in the Song Huong Formation. During the Middle Miocene, the continental shelf continued to subside, and sea levels rose, resulting in the development of Mioxen buildups in the Tri Ton Formation.

Diagenetic processes, including leaching, dissolution, compaction, fracturing, recrystallization, and dolomitization, led to the formation of various types of porosity, such as interparticle porosity, intraparticle porosity, crystalline porosity, and fracture porosity.

1.2.2 Middle Miocene carbonate characteristics on well data.

Up to now, a total of 4 wells have been drilled in the CX field, with wells GK2, GK3, and GK4 having comprehensive core sample analysis and high-quality well log measurement data.

The Middle Miocene carbonate reservoir rocks in the CX field exhibit good porosity and permeability, with a wide range of variations (porosity ranging from a few percent to over 30%, featuring various types of porosity: interparticle, vuggy, fractures; permeability ranging from below 1 mD to over 2,000 mD), displaying high heterogeneity (Figure 1.8).

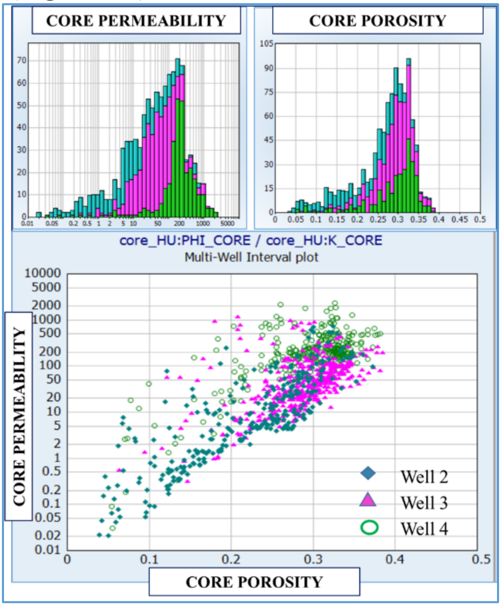


Figure 1.8: Porosity-Permeability Relationship from Core Sample Analysis Results in the CX Field.

The results of special core sample analysis reveal that the cementation exponent m and the saturation exponent n vary widely: m ranges from below 2 to over 2.7, with the majority concentrated in the range of 2.2-2.5; n varies from 1.9 to over 3.0, with the primary concentration in the range of 2.2-2.5. This indicates the complexity of the pore morphology.

CHAPTER 2
STUDY METHODOLODY AND DATABASE

2.1 Carbonate reservoir classification

i) Carbonate classification

Carbonate rocks are differentiated based on sedimentary characteristics, grain types, depositional environments, energy levels of the depositional environment, rock composition, mud-to-grain ratio, and other factors. Among the various methods of carbonate rock classification, the approaches proposed by Folk (1959) and Dunham (1962) are widely recognized.

Regarding pore space, three (03) classification systems for carbonate rock porosity are commonly used in the petroleum industry: Archie's classification (1952), Choquette and Pray (1970), and Lucia (1983, 1995).

ii) Carbonate rocktyping

The carbonate rocktyping is generally complex due to the intricate effects of the lithification process on the pore network. Each contractor, each researcher addresses the challenge of rocktyping in their own way. Figure 2.3 illustrates trends in carbonate rocktyping along with the advantages and disadvantages of these methods.

Several grouping methods are commonly used to describe the properties of carbonate rocks, including the Lucia method (1983, 1995), the Winland R35 method, the Pittman method, ... and the most common is based on Hydraulic Flow Units (HFU)

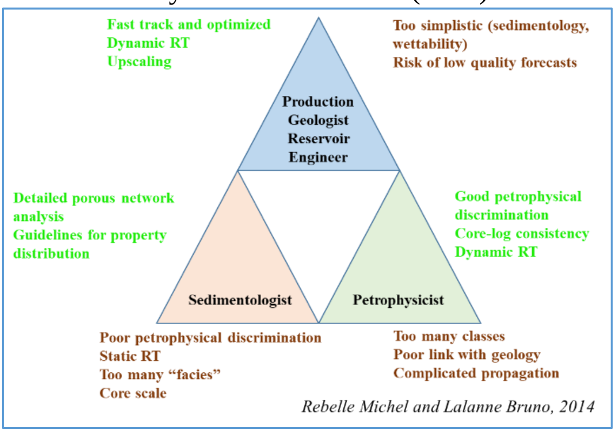


Figure 2.3: Advantages and limitations of each direction in carbonate rock grouping (Rebelle Michel and Lalanne Bruno, 2014).

Carbonate Rocktyping by Hydraulic Flow Unit - HFU

A Hydraulic Flow Unit (HFU) is defined as a volume of reservoir rock in which the geological characteristics control fluid flow consistently and predictably and are distinct from the characteristics of other units within the rock (Ebanks and nnk, 1984).

From the original Kozeny-Carman equation, Amaefule and colleagues (1993) introduced two supporting factors: PHIZ - normalized porosity and RQI - rock quality index, and proposed a new formula to determine the Flow Zone Indicator (FZI) based on the relationship between porosity and permeability. Simultaneously, it accurately estimates the reservoir rock quality for each specific HFU.

Permeability is calculated for each HFU using the mean FZI value (FZI_mean) of that HFU in the following formula:

$$k = 1040 * FZI\_mean^2 * \frac{\emptyset_e^3}{(1-\emptyset_e)^2} \qquad (2.10)$$

The author chose to approach the classification of carbonate reservoirs based on Hyraulic Flow Units for the research object due to the following reasons:

1. HFU classification provide an intimate connection between the spatial distribution of pore space and its impact on fluid flow.
2. Each HFU is controlled by specific geological characteristics and is distinct from others.
3. Once HFUs are classified and their distribution is predicted, it becomes convenient to describe reservoir characteristics. This includes predicting permeability, constructing water saturation models for each HFU, and subsequently building geological and dynamic models for accurate reservoir volume calculations and production forecasts.
4. In the CX area, there is a lack of research on classifying reservoirs based on HFUs."

2.2 The application of machine learning for HFU classification and permeability prediction in carbonate reservoir

For more than two decades, the FZI method has been widely used for HFU classification. Although traditional methods using a probabilistic plot or histogram of FZI is still being used for HFU classification, it shows the limitation that one can not see clearly on the chart the classified HFU as the groups of points are not clearly separated, the histogram does not clearly reveal the distribution of the groups.

Previously, for the prediction of rock types and petrophysical properties from well logs, experimental models or simple linear regression models (logK = aPHI + b) were often used. However, these models may introduce significant errors when applied to complex reservoirs such as carbonates. Nowadays, advancements in Machine Learning have been applied to describe reservoir characteristics, especially in the classification and prediction of rock types, yielding positive results.

Machine learning (ML) is a type of artificial intelligence (AI) that enables computer systems to classify, cluster, identify and analyze vast and complex sets of data while eliminating the need for explicit instructions and programming.

There are two common approaches to categorize Machine Learning algorithms: one based on function, and the other based on the learning style of each algorithm (Figure 2.6).

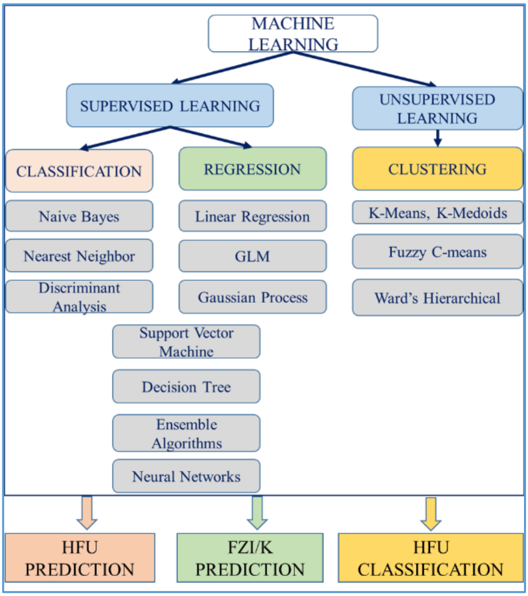


Figure 2.6: Machine learning categories

According to the learning styles, Machine Learning algorithms are generally divided into two main groups: Supervised Learning and Unsupervised Learning.

Supervised Learning is an algorithm that predicts the outcome of new input data based on known pairs of (*input, outcome*) from previous data. This pair of data is also referred to as *(data, label).* Supervised learning algorithms are further categorized into *Classification*, applied to discrete data such as HFU, and *Regression*, applied to continuous data like Flow Zone Indicator FZI or Permeability.

Unsupervised Learning is an algorithm that does not know the outcome or label in advance, only having input data. The algorithm relies on the data structure to perform tasks such as clustering or reducing the dimensionality of the data for easier storage and computation. These algorithms are applied to HFU clustering based on core sample analysis data FZI_core.

2.3 Water Saturation prediction

The accurate modeling of water saturation is one of the factors that affects the hydrocarbon in place estimation, the prediction of recoverable hydrocarbon, the recovery process and the future plans of developing such reservoirs. However, the water saturation is not only controlled by the porosity, lithofacies and other parameters in Archie formula, but also by pore structure, shale content and wettability that are not consistent in carbonate reservoir and all of these factors will lead to the high uncertainty in the traditional method for water saturation estimation, the alternative method using the Saturation Height Function (SHF) based on the relationship between water saturation (Sw) and the capillary pressure (Pc) was introduced.

With Saturation Height Function (SHF), the geologist or reservoir engineer is able to predict the saturation anywhere in the reservoir for the given height above the free water level (FWL) and for a given reservoir permeability or porosity independently of Archie parameters.

To ensure accuracy in predicting Sw for complex carbonate rocks, the Pc = f(Sw) relationship must be constructed for each Flow Unit (Figure 2.14).

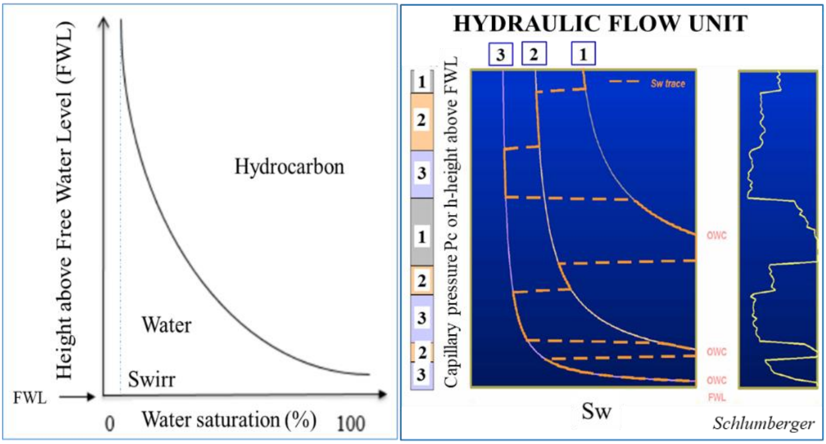


Figure 2.14: The relationship of Sw and Pc or Height above Free Water Level

2.4 Workflow

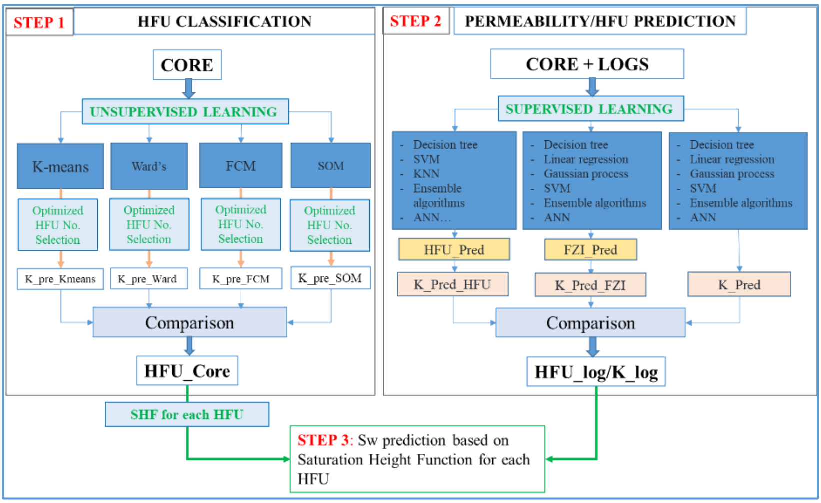


Figure 2.15: Workflow applied for the research thesis

Based on the analysis of the research methods mentioned above, the author proposes a workflow to be applied in the thesis with 3 main steps as follows (Figure 2.15) :

- Step 1: HFU classification using unsupervised ML (K-Means, Ward's Hierarchical clustering, FCM, SOM)
- Step 2: Permeability/HFU prediction using supervised ML (Decision Tree, SVM, Ensemble Algorithm, Gaussian Process Regression, Neural Networks, ...)
- Step 3: Sw prediction using SHF model constructed for each HFU

2.5 The research database

The primary data utilized for the research includes data collected from wells drilled within the CX field, comprising: i) traditional wireline logging measurements; ii) core sample analysis (including over 1,000 RCAL, 152 Pc measurements from SCAL. Additionally, the author consulted various other research documents, including the thin-section biostratigraphy reports.

CHAPTER 3
APPLICATION OF MACHINE LEARNING METHODS IN HFU CLASSIFICATION

3.1 Application of unsupervised Machine Learning in Carbonate Rocktyping by HFU

Unsupervised machine learning with widely used methods such as K-means, Ward's hierarchical, Self-Organizing Maps and Fuzzy-C Mean has been used. These methods aim to ensure the objectivity and accuracy in HFU clustering.

The optimal number of HFUs is selected based on the elbow method on the graph of the relationship between the Squared Correlation Coefficient - R2 or the Root Mean Square Error - RMSE and the number of clusters (n). These graphs (n-R2, n-RMSE) are constructed for each algorithm to determine the optimal number of clusters that provides the best fit for permeability predictions after HFU clustering compared to the permeability values of core samples.

3.2 The HFU classification result

After removal of outliers, a total of 997 sample points from 3 wells, GK2, GK3, GK4, were used for HFU classification.

Table 3.1 shows that although there is not a significant difference in R2 and RMSE among the methods, all methods provide high correlation (R2) and low RMSE values between the predicted permeability (K_pred) and core permeability (K_core). Among them, the K-means clustering method with 5 HFUs has the highest R2 value and the lowest RMSE value. Therefore, the results of grouping based on the K-means method with 5 HFUs will be used for the subsequent steps.

Table 3.1: Comparison of Flow Unit (FU) grouping results using Machine Learning algorithms

	K-means (5 HFUs)	**FCM (4 HFUs)**	**Ward (5 HFUs)**	**SOM (4 HFUs)**
R2	0.9730	0.9562	0.9708	0.9558
RMSE	0.1459	0.1840	0.1505	0.1851

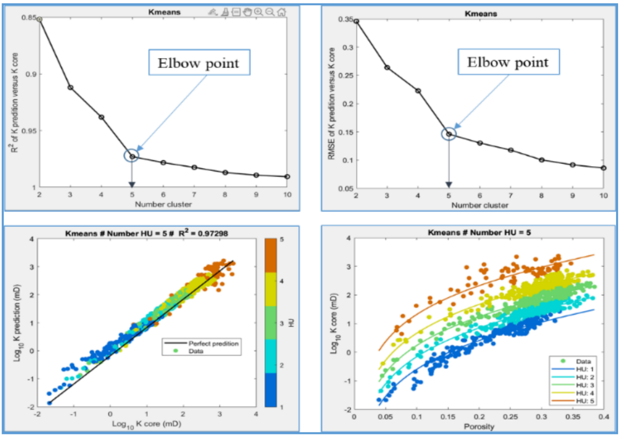


Figure 3.4: Result of HFU's number selection and the HFU classification method based on K-means method

The FZI-mean value for each HFU are shown in the table 3.2 below:

Table 3.2: HFU classification using K-means method

	HFU 1	**HFU 2**	**HFU 3**	**HFU 4**	**HFU 5**
FZI range	0.2243-0.5795	0.5795-0.9374	0.9374-1.50	1.50-2.811	2.811-10.2688
FZI-mean	0.4433	0.7566	1.1915	1.9679	4.4991

Poro-perm relationship	K=204.37 5*PHI^3/(1-PHI)^2	K=595.34 1*PHI^3/(1-PHI)^2	K=1476.4 6*PHI^3/(1-PHI)^2	K=4001.3 8*PHI^3/(1-PHI)^2	K=21051. 6*PHI^3/(1-PHI)^2

The differences between HFUs can be identified in thin section analysis. Figure 3.9 shows some images of thin section analysis representing different HFUs.

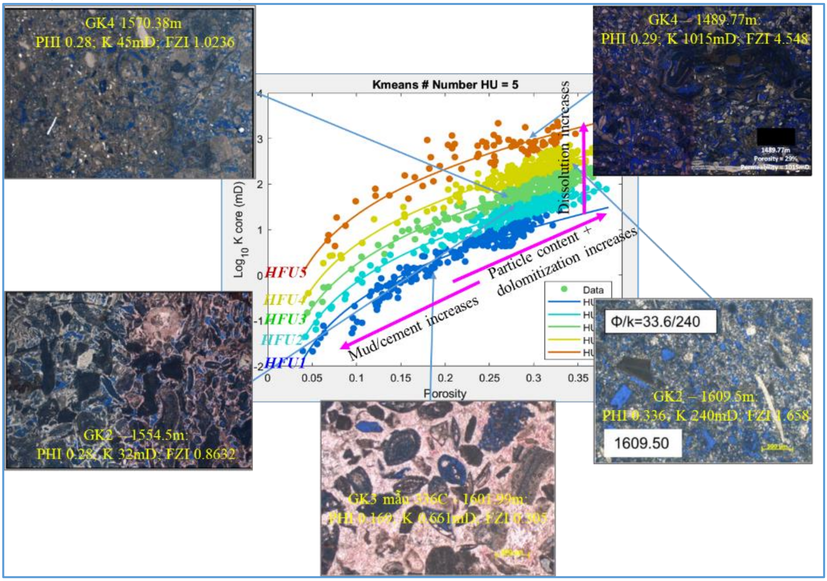


Figure 3.9: Representative thin section analyses for the classified HFUs at the CX wells.

The relationship between porosity and permeability is controlled by the type of pore and the connectivity between pores. In Figure 3.9, it can be observed that as the mud/cement increases, the porosity decreases; as the particle content increases/the degree of dolomitization increases, leading to an increase in porosity. However, this does not necessarily correspond to a proportional increase in permeability, as this relationship is non-linear and depends on each flow unit. Within the same range of porosity, the permeability values tend to increase gradually from HFU 1 to HFU 5. This occurs due to the increasing dissolution process from HFU 1 to HFU 5, resulting in enhanced connectivity between pore spaces. The rock quality increases gradually from HFU 1 to HFU 5 according to the increasing values of the Flow Zone Indicator (FZI).

CHAPTER 4
PERMEABILITY PREDICTION USING MACHINE LEARNING METHODS

4.1 Application of supervised machine learning methods in HFU/Permeability prediction

With supervised machine learning algorithms, permeability (K_pred) can be directly predicted from the core-log relationships, or indirectly through predicting the values of flow zone indicator (FZI_pred) or hydraulic flow units (HFU_pred). Subsequently, the relationship between porosity and permeability for each HFU can be used to calculate the permeability. To ensure the most objective results, the author performs permeability prediction using all these approaches, and then compares and selects the best-performing method.

Almost all common supervised machine learning methods were utilized in this research including decision tree models, boosted tree methods, artificial neural networks (ANN), support vector machines (SVM), Gaussian process regression, and others.

4.2 HFU/Permeability prediction results

The author sequentially performs the prediction of flow zone indicator values (FZI_pred), HFU predictions (HFU_pred), and permeability predictions (K_pred) based on the basic well logs data through the following steps:

- ***i)*** ***Training dataset preparation:*** Core data and well log data are carefully checked to remove outliers and depth matched. The correlation between core data and log curves was also checked to ensure the machine learning methods will work with this dataset.
- ***ii)*** ***Training:***

The inputs consist of various well log curves, including Gamma Ray (GR), Deep Resistivity (RD), Shallow Resistivity (RS), Micro Spherically Focused Log (MSFL), Bulk Density (RHOB), Neutron Porosity (NPHI), Compressional Sonic Log (DTC), and Shear Sonic Log (DTS). Supervised data includes FZI_core/HFU_core/K_core, and the output data includes FZI_pred/HFU_pred/K_pred. Tables 4.1 and 4.2 present the verification and testing results of different algorithms used for predicting Flow Zone Indicator (FZI_Pred), permeability (K_Pred), and Hydraulic Flow Unit (HFU_Pred).

Table 4.1: FZI_Pred, K_Pred Prediction results using ML methods

No	Model		FZI_PRED PREDICTION				K_PRED PREDICTION			
			RMSE		R2		RMSE		R2	
			Validation	Test	Validation	Test	Validation	Test	Validation	Test
1	Regression	Linear	0.895	0.732	0.502	0.484	0.516	0.463	0.679	0.717
2		Interactions	0.933	0.678	0.459	0.557	0.504	0.418	0.693	0.77
3		Robust	1.120	0.843	0.220	0.315	0.518	0.459	0.677	0.723
4		Stepwise	0.905	0.687	0.491	0.545	0.512	0.463	0.684	0.718
5	Decision Tree	Fine	0.616	0.549	0.764	0.709	0.481	0.332	0.722	0.855
6		Medium	0.736	0.615	0.663	0.636	0.481	0.355	0.721	0.834
7		Coarse	0.816	0.672	0.587	0.564	0.543	0.437	0.644	0.749
8	SVM	Linear	1.002	0.755	0.376	0.450	0.518	0.461	0.676	0.72
9		Quadratic	0.872	0.609	0.527	0.642	0.489	0.394	0.712	0.795
10		Cubic	0.904	0.490	0.492	0.769	0.614	0.32	0.546	0.865
11		Fine	0.802	0.453	0.600	0.803	0.432	0.282	0.775	0.896
12		Medium	0.871	0.567	0.528	0.690	0.457	0.35	0.748	0.839
13		Coarse	1.026	0.761	0.345	0.442	0.523	0.454	0.67	0.729
14	Ensemble algorithms	Boosted	0.583	0.507	0.789	0.752	0.428	0.343	0.779	0.845
15		Bagged	0.605	0.494	0.773	0.764	0.416	0.287	0.792	0.892
16	Gaussian Process Regression	Square Exponential	0.539	0.487	0.819	0.772	0.416	0.286	0.791	0.892
17		Matern 5/2	0.494	0.479	0.848	0.779	0.361	0.267	0.843	0.906
18		**Exponential**	**0.459**	**0.395**	**0.869**	**0.850**	**0.320**	**0.235**	**0.870**	**0.927**
19		Rational Quadratic	0.463	0.423	0.867	0.828	0.323	0.235	0.860	0.927
20	Neural Network	Narrow	0.694	0.593	0.701	0.661	0.454	0.352	0.752	0.837
21		Medium	0.669	0.555	0.722	0.703	0.452	0.337	0.754	0.851
22		Wide	0.632	0.458	0.752	0.798	0.5	0.341	0.699	0.846
23		Bilayered	0.882	0.593	0.517	0.661	0.449	0.38	0.757	0.81
24		Trilayered	0.654	0.834	0.734	0.329	0.481	0.315	0.721	0.869

Table 4.2: HFU_Pred Prediction results using ML methods

No	Model		Accuracy (%)	
			Validation	Test
1	Decision Tree	Fine	70.7	79.1
2		Medium	63.1	64.7
3		Coarse	54.9	51.4
4	Discriminant	Linear	55.9	55.8
5		Quadratic	58.7	60.6
6	Logistic	Efficient Logistic Regression	40.8	45.0

No	Model		Accuracy (%)	
			Validation	Test
7	Naive Bayes	Gaussian	52.1	57.0
8		Kernel	58.2	63.9
9	SVM	Linear	57.4	58.6
10		Quadratic	69.8	77.5
11		Cubic	78.9	80.3
12		Fine Gaussian	75.8	82.3
13		Medium Gaussian	64.4	66.3
14		Coarse Gaussian	51.1	52.2
15	KNN	Fine	80.5	85.9
16		Medium	62.7	67.5
17		Coarse	49.2	51.8
18		Cosine	65.4	67.9
19		Cubic	63.0	69.9
20		Weighted	80.7	85.9
21	Ensemble	Boosted tree	69.1	73.9
22		Bagged tree	81.7	84.3
23		Subspace discriminant	55.2	55.0
24		Subspace KNN	78.7	85.1
25		RUSBoosted tree	67.6	70.3
26		Optimizable	82.5	85.9
27	Neural Networks	Narrow	69.0	70.7
28		Medium	77.0	81.5
29		Wide	79.4	85.9
30		Bilayered	72.2	77.5
31		Trilayered	78.1	79.5

Based on the machine learning results, the Gaussian Exponential Regression method was selected for predicting the Flow Zone Indicator (FZI_pred) and permeability (K_pred) due to its highest correlation (R2) and lowest Root Mean Squared Error (RMSE) values for both verification and testing data. The Combined Optimizable method was chosen for predicting the Hydraulic Flow Unit (HFU_Pred) as it demonstrated the highest accuracy in both verification and testing datasets.

iii) Applying the trained models to the entire carbonate section

The selected models were applied to the entire carbonate section of the research subject. The correlation coefficients between the predicted permeability from different methods in the step 2 of the workflow are shown in the tables 4.4 below:

Table 4.4: Comparison of correlation coefficients between predicted permeability from different methods and the permeability values from core samples (K_core).

	Premeability prediction directly using ML (K_Pred)	Premeability prediction using predicted FZI using ML (K_Pred_FZI)	Premeability prediction using predicted HFU using ML (K_Pred_HFU)
R2	**0.876**	**0.780**	**0.803**

The results of directly predicted permeability K_pred from machine learning methods will be used in the subsequent steps of the study as they exhibit the highest correlation.

Figure 4.13, 4.17 illustrate the permeability prediction results (K_pred) directly from the well log data using the machine learning method based on the Gaussian Exponential regression algorithm. There is a better correlation between the predicted results and the core analysis data in comparison with the result using the traditional linear regression method.

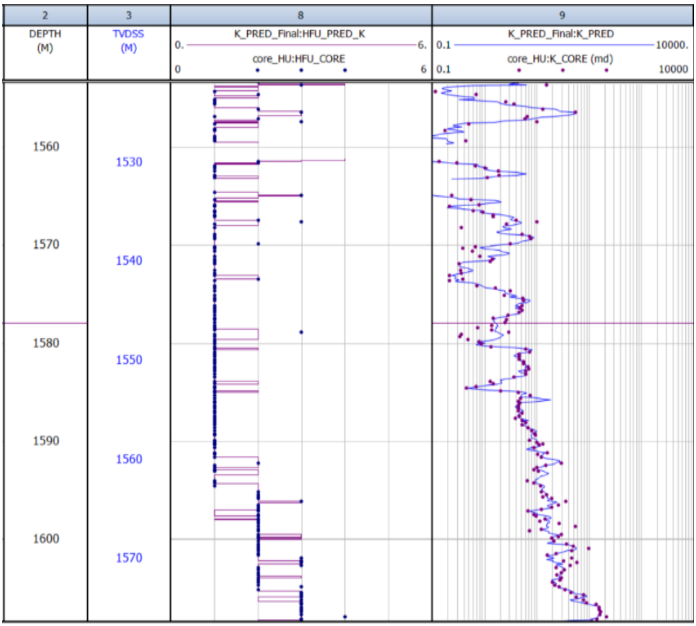


Figure 4.13: The results of permeability and HFU prediction in Well 2

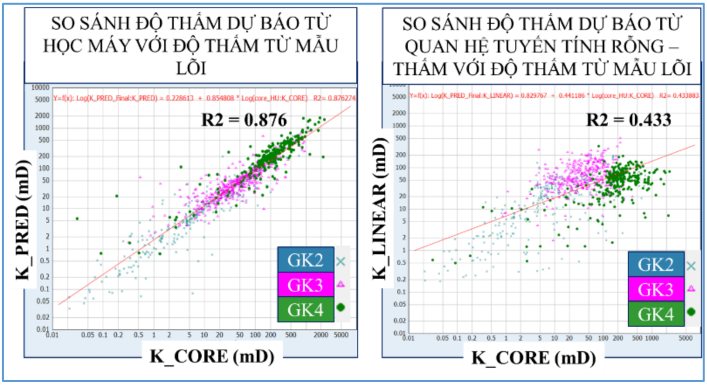


Figure 4.17: Comparison permeability prediction results with ML method (left) and with traditional linear regression method (right)

CHAPTER 5
WATER SATURATION PREDICTION BASED ON HFU AND PERMEABILITY PREDICTION RESULTS

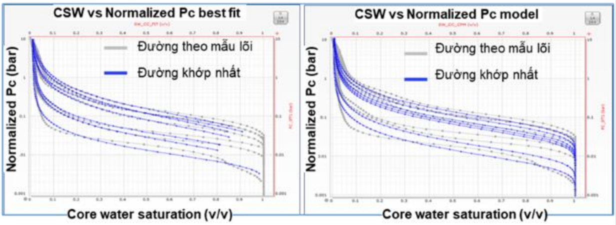
5.1 SHF model construction

The relationship between water saturation (Sw) and capillary pressure (Pc) is detailed for HFU. A total of 5 models were tested, including Leverett-J, Brook Corey, Lambda method, Thomeer method, and Skelt-Harrison method. According to the model fitting results, the Skelt-Harrison model gave the smallest error; therefore, this model is used to build the Sw prediction model for each hydraulic flow unit (Table 5.1).

Table 5.1: The fitting error of Sw prediction models

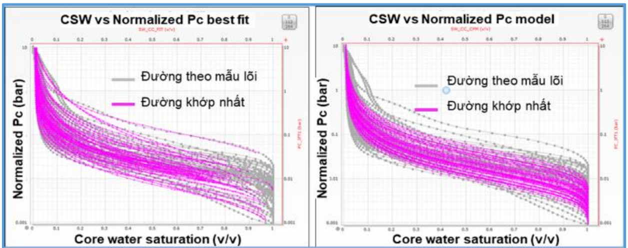
	Brook Corey model	**J-function model**	**Lamda model**	**Thomeer model**	**Skelt Harrison model**
Error	0.0428	0.039	0.0359	0.0878	**0.0349**

The results of the equation construction for Sw calculation based on the Skelt Harrison model for each HFU are illustrated in the figures below:



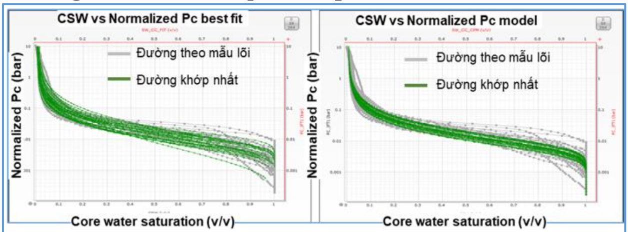
- PC_IFT1 = (HAFWL * (0.98152 - 0.18014) * 0.0980665 * (1.0 / (50 * 1)))
- A = 0.992609
- B = (0.0230608 * Pow(PERM, -0.298812))
- C = 1.02147
- D = 0
- SW = Min(1, Max(0, If((B < 0 Or PC_IFT1 <= -D), 1, 1 - A * Exp(-Pow(B / (D + PC_IFT1), C)))))

Figure 5.7: Sw equation prediction for HFU 1



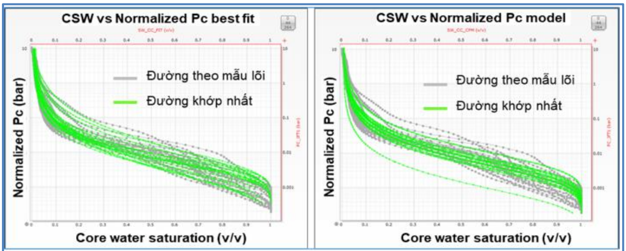
- PC_IFT1 = (HAFWL * (0.98152 - 0.18014) * 0.0980665 * (1.0 / (50 * 1)))
- A = 0.992609
- B = (0.0230608 * Pow(PERM, -0.298812))
- C = 1.02147
- D = 0
- SW = Min(1, Max(0, If((B < 0 Or PC_IFT1 <= -D), 1, 1 - A * Exp(-Pow(B / (D + PC_IFT1), C)))))

Figure 5.8: Sw equation prediction for HFU 2



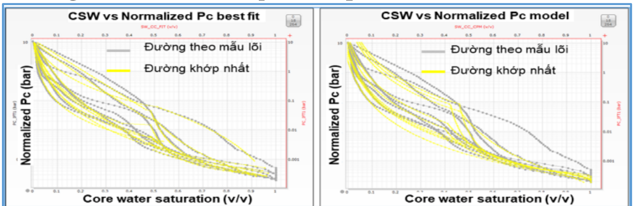
- PC_IFT1 = (HAFWL * (0.98152 - 0.18014) * 0.0980665 * (1.0 / (50 * 1)))
- A = 0.992855
- B = (0.0252263 - 0.014156 * Log(Sqrt(PERM / POR)))
- C = 1.00401
- D = 0
- SW = Min(1, Max(0, If((B < 0 Or PC_IFT1 <= -D), 1, 1 - A * Exp(-Pow(B / (D + PC_IFT1), C)))))

Figure 5.9: Sw equation prediction for HFU 3



- PC_IFT1 = (HAFWL * (0.98152 - 0.18014) * 0.0980665 * (1.0 / (50 * 1)))
- A = 0.995728
- B = (0.0176328 - 0.00560397 * Log(PERM))
- C = 0.762647
- D = 0
- SW = Min(1, Max(0, If((B < 0 Or PC_IFT1 <= -D), 1, 1 - A * Exp(-Pow(B / (D + PC_IFT1), C)))))

Figure 5.10: Sw equation prediction for HFU 4



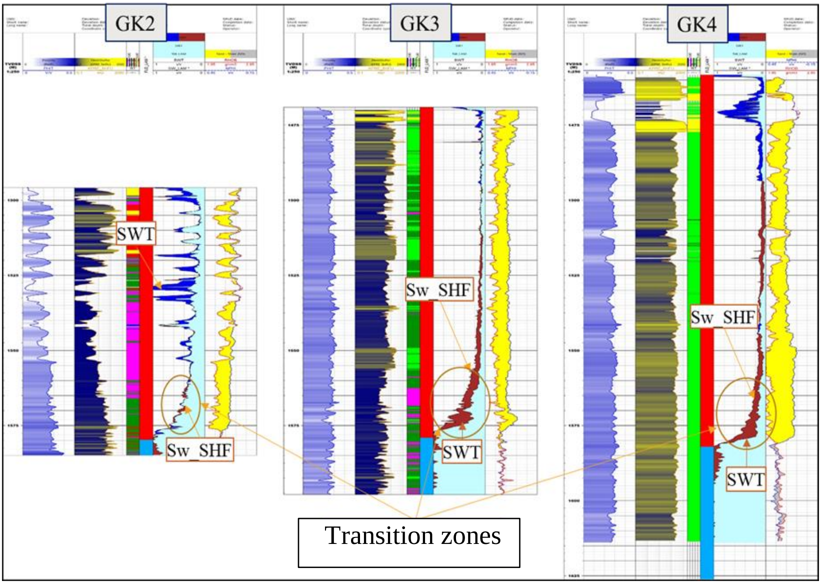
- PC_IFT1 = (HAFWL * (0.98152 - 0.18014) * 0.0980665 * (1.0 / (50 * 1)))
- A_PT1 = Pow(10, -0.490762 - 2.37374 * POR)
  Lambda_PT1 = 0.332301
- B_PT1 = -0.04702
  SWi_PT2 = (0.482889 - 0.098745 * Log(PERM))
- PCe_PT2 = Pow(10, -3.55136 - 0.000188502 * PERM)
- G_PT2 = (0.685344 + 0.000782752 * PERM)
- SW = Min(Min(1, Max(0, A_PT1 * Pow(PC_IFT1, -Lambda_PT1) + B_PT1)),Min(1, Max(0, If(PC_IFT1 < PCe_PT2, 1, SWi_PT2 + (1 - SWi_PT2) * (1 - Pow(Exp(1), (G_PT2 / Ln (PCe_PT2 /PC_IFT1)))))))

Figure 5.11: Sw equation prediction for HFU 3

5.2 Sw prediction result using constructed SHF

The estimated Sw_SHF was plot against the log-derived water saturation SwT (log-derived Sw) in Figure 5.12. From the plot, we can see that above the transition zone, the Sw_SHF and SwT values have many similarities, except for points with low porosity values. The permeability values at these points according to the core sample analysis results can still allow gas to flow (considering gas flow where permeability >0.1 mD), but the SwT results are quite high (>0.7), which is quite inappropriate. Thus, the result of estimated Sw_SHF in this zone is more optimal than log-derived SwT.

In the transition zone there is a relatively large difference between Sw_SHF and SwT espescially in well 3 (marked brown in track water saturation in Figure 5.12). The water saturation distribution within the transition zone is controlled by the distribution of rock types (by HFU in this study). The height of transition zone depends on the reservoir rock types. Better the reservoir quality, thinner the transition zone and vice versa. With dominant HFU 2 in this zone of well 3 the water saturation will change gradually, not change dramatically as SwT. So the Sw_SHF shows a more reasonable behavior in this zone.



Hình 5.12: Sw estimation result using SHF model built for 5 HFUs

CONCLUSION AND RECOMMENDATION

1. CONCLUSION

Given below are some conclusions drawn from the study results:

- Unsupervised machine learning methods are highly useful in objectively and accurately clustering hydraulic flow units. The K-means method was chosen to apply in dividing the Miocene carbonate reservoir in the CX field into 5HFUs.

- Based on supervised machine learning methods, permeability was directly predicted from the wireline logging data with high correlation with the results of core sample analysis.
- Applying the 5 hydraulic flow units combined with permeability directly predicted from the wireline logging data into the saturation height function building process has been found to be entirely suitable for carbonate rocks in CX field.

Based on the research results, the author observed that:

- The CX reservoir carbonate rocks exhibit high heterogeneity, with permeability ranging widely from very poor to very good (from below 1 mD to over 2,000 mD), porosity-permeability relationship is non-linear and dependent on each hydraulic flow unit.
- Water saturation decreases with height relative to the free water level; the water saturation value and the depth of transition zone are influenced by the hydraulic flow units.

Applying machine learning techniques to describe reservoir characteristics has numerous advantages, but several considerations need to be taken into account:

- Machine learning methods often require a large amount of input data to generate accurate predictive models. Input data must be carefully selected and cleaned to ensure quality. With core sample data, there needs to be a sufficient representation of each HFU to ensure accurate and objective HFU classification and prediction.
- Supervised and unsupervised machine learning methods yield vastly different results depending on the dataset. Therefore, it's necessary to experiment with various methods to choose the most suitable one for the dataset obtained in the research area.

2. RECOMMENDATION

Based on the research findings, the author recommends:

- Applying the hydraulic flow unit classification using unsupervised machine learning combined with permeability prediction results from supervised machine learning to determine water saturation for carbonate rocks in other regions.
- Conducting experimental studies to apply this method to other types of heterogeneous rocks (fractured basement rocks, highly heterogeneous sandstones, etc.).

LIST OF PUBLISHED ARTICLES AND SCIENTIFIC WORKS BY THE AUTHOR

Vietnamese

1. Nguyễn Văn Hoàng, Hoàng Việt Bách, Nguyễn Trung Dũng, Lê Trung Tâm và nnk (2021), "Tối ưu mô hình tính độ bão hòa nước cho tầng chứa Miocene khu vực Đông bắc bể Cửu Long, Việt Nam", *Tạp chí dầu khí*, số 7 - 2021, trang 16-22, ISSN 2615 - 9902.

English

2. Nguyen Trung Dung, Ha Quang Man, Nguyen Hong Viet, Phan Thien Huong, Cu Minh Hoang, Truong Khac Hoa (2023), "Developing a Saturation Height Function (SHF) for the classified Hydraulic Flow Unit, a case study from Middle Miocene Carbonate Reservoir in the southern part of Song Hong basin, Vietnam", *Journal of Mining and Earth Sciences* Vol. 64, Issue 6 (2023) 1-10.
3. Nguyen Trung Dung, Ha Quang Man, Truong Khac Hoa, Phan Thien Huong, Cu Minh Hoang, Nguyen Hong Viet (2024), "Applying machine learning for hydraulic flow unit classification and permeability prediction: case study from carbonate reservoir in the southern part Song Hong basin", *Vietnam Journal of Marine Science and Technology*, 4/2024.

Russian

4. G.G.Djalylov, Vu Tat Thang, Nguyen Trung Dung, Vu Nguyen Khang (2012), "An new insight on the tectono-structural characteristics of the Central and Northern regions of the South Usturst Area and the Hydrocarbon Potential of the Region", *Uzbek journal of oil and gas*, (3/2012), tr. 39-42.
5. Nguyen Tri Dung, I.T. Boykobilov, Vu Tat Thang, Nguyen Trung Dung and others, (2014), "New hydrocarbon exploration direction in pre-Jurassic formation of Bukhara Khiva Region, Uzbekistan" - *Current Issues in Petroleum Geological Science, Deep Drilling Techniques, and Well Research*, p. 4-9.
6. G.S. Abdullaev, Nguyen Tri Dung, Nguyen Trung Dung, I.T. Boykobilov, M.R.Nurmatov (2014), "Geological Structure and Petroleum Potential of the Permian Formation in the Kossor Investment Block", *Uzbek Journal of Oil and Gas*, Special Edition, (OGU 5/2014).