SUPERVISED MACHINE LEARNING APPLICATION FOR ESTIMATION OF SHEAR WAVE AND CHARACTERIZATION OF LOWER GORU FORMATION, SAWAN AREA, LOWER INDUS BASIN.

BY

SYED USAMA AFZAAL NOSHAHI

MPhil Geophysics

2021-2023

Department of Earth Sciences

Quaid-I-Azam University Islamabad

CERTIFICATE

This dissertation is submitted by **Syed Usama Afzaal Noshahi** S/O **Syed Afzaal Sabtain Noshahi** is accepted in its present form by the Department of Earth Sciences as satisfying the requirement for the award of M.Phil degree in **Geophysics.**

RECOMMENDED BY

Dr. Anees Ahmed Bangash _____

(Supervisor)

Dr. Mumtaz Muhammad Shah _____

(Chairman Dept. Of Earth Sciences)

External Examiner _____

ACKNOWLEDGEMENT

In the name of Allah, the Most Merciful and Generous. All thanks be to the Almighty, The One, The Everlasting, Who begets no one, is begotten by no one, and has no equivalent. Alhamdulillah. I testify that Holy Prophet Muhammad (PBUH) is the final messenger, whose life serves as a flawless example for all humanity until the Day of Judgment. I am grateful to Allah for the strength and blessing He has given me in finishing this thesis.

Without your assistance, I am nothing. Please keep me prostrated in front of you at all times, and do not allow me to leave in front of anyone else.

First and foremost, I want to convey my heartfelt thanks to my supervisor, **Dr. Anees Ahmed Bangash**, for his unwavering support throughout my M.Phil studies and research, as well as his drive, excitement, vast knowledge, and trust in me. I gained a lot of knowledge from his perspective. His advice was invaluable during the research and writing of this thesis.

Mr. Yawar Amin (PhD) and **Mishal Razaq** have my heartfelt gratitude for their guidance in the preparation of this thesis, as well as their assistance and unconditional support during the two-year journey.

I really appreciate my entire family's prayers and efforts, particularly my parents, for their encouragement, support, and sacrifices throughout the study. I'd also like to express my gratitude to the entire staff of my department for providing me with an academic foundation that has allowed me to pursue this course of study.

Syed Usama Afzaal Noshahi

Dedication

I would like to dedicate this research work to my Parents whom affection, love, encouragement, prayers and believe in me made me able to get such success and honor.

ABSTRACT

Conventional geophysical techniques applied to vintage seismic and wireline log data have been successful in hydrocarbon exploration. However, the use of machine learning techniques has become increasingly important in overcoming the acquisition based data limitations and subsequently improving the efficiency and accuracy of hydrocarbon exploration. This thesis presents the characterization of the Lower Goru Formation in the Sawan Area of the Lower Indus Basin through the integration of supervised machine learning techniques into geophysical analysis. The primary objectives of this research encompass the precise demarcation of the reservoir of interest, the identification of significant subsurface structures, and an in-depth examination of the tectonic regime in the region. These objectives serve as a foundation for reservoir characterization and hydrocarbon exploration.

The methodology employed involves precise tying of horizons to seismic sections, the creation of synthetic seismograms using well data and a process of horizon demarcation on seismic sections. This approach facilitates the generation of spatial time maps for key horizons (D-, C-, and B-sands), providing valuable insights into structural variations.

Moreover, the petrophysical analysis of well Sawan-01 reveals key reservoir parameters, including volume of shale (20%), porosity (17.3%), effective porosity (11.4%), water saturation (34%), and hydrocarbon content (66%). A comparative analysis with well Sawan-07 highlights the reservoir heterogeneity within the study area.

The core of this research lies in the prediction of the DT4S log in well Sawan-01 through a supervised machine learning technique. Gradient Boost Regressor gives 94% accurate prediction for data trained and tested on Wells Sawan-07 and Sawan-08. The model was applied to well Sawan-01, demonstrating its efficacy in estimating critical subsurface properties.

Notably, this research concludes with the successful demarcation of hydrocarbon zones through cross-plot analysis utilizing the predicted shear wave data in conjunction with other elastic parameters. The use of predicted log for such analysis is effective for reservoir evaluation, particularly in scenarios where shear wave velocity data is historically lacking in older wells.

Table of Content

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 Introduction

One of the most crucial aspects of the economy of any nation is the exploration of hydrocarbons (Morea., 2023). As the need for energy increases, the research community focuses on previously unexplored areas in search of fresh energy sources. Geophysical techniques and tools are used to examine the ground, and these include taking measurements and calculating physical characteristics to look for differences both laterally and vertically (Arshad et al., 2013).

For oil and gas field development in the future and prospect assessment, accurate characterization of reservoir parameters is essential (Ashraf et al., 2020) In terms of accuracy, resolution, cost, and deep penetration, seismic methods have emerged as the most significant geophysical method. Seismic interpretations can be used to choose the locations of exploratory wells. Estimating fundamental reservoir characteristics like porosity, permeability, volume, type of subsurface fluids, and the lateral and vertical extent of a reservoir with its boundaries represents reservoir characterization. (Wu et al., 2021b, Huang et al., 2021, Iltaf et al., 2021, Toqeer et al., 2021). Most often, well-logs and seismic data are used to evaluate the reservoir characteristics at various levels (Chopra and Marfurt, 2007).

The 3D seismic approach is now commonly used in the petroleum sectors to explore hydrocarbons because it provides a thorough picture of subsurface information not just in the vertical direction but also for the total volume of the studied area (Manzi et al., 2020). Geoscientists can now interpret and extract useful data relating to the identification of anomalous zones as well as for mapping reservoir quality due to the seismic signal's use to alter by interacting with hydrocarbon bearing zones and other potential traps and reflect back to surface (Bacon et al., 2007).

Petrophysical studies were always utilized to distinguish between reservoir and non-reservoir zones (Amigun et al.,2012). It provides a brief explanation of the interactions between the rocks and fluids present there, or between the hydrocarbons and associated reservoir (Ahammod et al., 2014). The depositional setting of the reservoir can be determined from petrophysical characteristics like as porosity, fluid saturation, and shale volume. The connected pores play an

important part in the transportation and storage of hydrocarbons (Donaldson and Tiab, 2004). A reservoir's permeability is its primary source of efficiency. Effective use of these techniques makes it easy to identify hydrocarbon leads. For identifying lithologies with distinct variations on larger sizes, well log cross plots are important.

Machine learning has helped greatly in geophysical research and analysis like predicting shear wave, reservoir characterization, soil modeling etc. in recent years (Dimililer et al., 2021). Machine learning can be greatly help in this regard as it uses algorithms for well log predictions. Several applications of machine learning in geophysics have transformed how data is analyzed, interpreted, and used in different geophysics subfields (Dramsch, 2020). As in old wells, due to expense and lack of techniques, has missing well logs especially Shear wave (DT4S) (Pham et al., 2020).

Cross-plotting parameters against one another is a useful graphical analysis for finding data clusters in targeted zones. Based on their responses, these data sets can be divided into different lithologies/facies and hydrocarbon zone (Veeken & Rauch-Davies, 2006). Elastic properties (VpVs ratio) enable the differentiation between different types of lithology and hydrocarbon bearing zone in the targeted area. Using P-impedance and the VpVs ratio together, lithology and fluid content can be predicted with high accuracy (Azeem et al., 2017).

1.2 Main Aims and Objectives of Dissertation

General aims of the current research work are as follows

- Demarcation of the reservoir of interest and identification of significant structures along with the tectonic regime in the area.
- Petrophysical analysis using wireline logs of wells Sawan-01 and Sawan-07 in order to identify the reservoir of interest.
- Prediction of DT4S log in well Sawan-01 using machine learning technique trained on wells Sawan-07 and Sawan-08.
- Reservoir characterization based on the cross plots analysis of the elastic parameters including the predicted DT4S log estimated in well Sawan-01.

1.3 Study Area:

1.3.1 Introduction:

The Sawan gas field, which contains about 2 Tcf, is the area of study. It is located in Sindh's Khairpur district. In addition, it is located in the southern portion of the Lower Indus basin and is bordered on the north by Jacobabad High, the east by the Indian Shield Suleiman thrust fold belt, and the south by the Karachi Embayment Zone (Figure 1.1) (Afzal et al., 2009), (Kadri, 1995).

Figure 1. 1 : The research area placed on a map of Pakistan's tectonic framework and basin categorization (Gul et al., 2023).

OMV Pakistan discovered Sawan area in 1998, and production there began in 2003 (Berger et al., 2009). This field, which was the result of a collaboration between OMV Pakistan, ENI Pakistan,

PPL, and Government Holding Private Limited (GHPL), is likewise regarded as a major gas-producing field. The Sawan field is currently in the producing phase, with 14 of the 15 producing wells (Ahmed et al., 2010) providing the majority of the gas to SNGPL and SSGPL. In terms of the region, Sawan Field is situated in an extensional regime that primarily consists of normal faults. abruptly up-dip stratigraphic traps with structural characteristics make up the trapping system. The small four-way dip closure and amplitude anomaly are the main targets of the exploratory well Sawan-01 (Afzal et al., 2005). The primary reservoir in the studied region is C Interval sand, which is limited to northeast-southwest directed fairways and is deposited in shallow marine environments. Major source rock found is Sembar formation deposited in shelf marine environment and in whole region it is proven as organic rich rock most suitable source rock along with Lower Goru member (Ahmad et al., 2004).

1.3.2 Previous Exploration Work:

Early-late Cretaceous Lower Goru sands have been the primary source of production in Pakistan for the past 20 years (Ahmad et al., 2004). The primary areas of production in the entire basin are Miano, Sawan, Badin, and Kadanwari. Most frequently, firms like OGDCL and OMV explored the Lower Goru sediments throughout a wider area. In 2003, Oolithica Geoscience Ltd. developed a proper comprehensive model of the reservoir based on well logs from 7 wells and 12 inlines and crosslines passing through the wells, as well as the various study models developed by the OMV for the Sawan C sand. The main result of this work was the identification of a depositional model in which the entire Sawan field was divided into three prograding major clinoforms (Afzal et al., 2005).

Sawan block is divided into three compartments in general by two strike slip faults in the south, center, and north compartments (Rehman and Ibrahim, 2009). In this study a seismic 3D cube (10 x 10 km) in the southern region has been used. The Sawan field's central area is known as the primary tank of gas since it has the most wells that have been drilled are in the whole Sawan block. Study area in Lower Indus basin having following geographical coordinates ranging from:

1.3.3 Data Source:

To continue research in Sawan block, data was provided by Directorate General of Petroleum Concessions (DGPC). Details of data used for the study are as under:

- Formation tops
- Seismic Header files
- 3D Seismic Cube
- Navigation files
- Well Logs (Las Files)

1.4 Data Set:

1.4.1 Seismic Data:

The seismic dataset used in the study is given in (Table 1.1). 3D seismic cube of (10x10 km^2) is used in this research work.

Table1. 1: Details of inline and crossline of 3D Seismic Data

Lines	Start	End	Total No of Lines
In Line	718	864	146
Cross Line	874	1005	131

1.4.2 Well Data:

To accurately interpret seismic data, precise horizon positions, and other in-situ conditions from well logs, a total of three wells were used in this dissertation which are Sawan-01, Sawan-07 and Sawan-08. Due to its location in the center of a 3D cube and the other two wells, the Sawan-07 well was the most appropriate for achieving research objectives. Sawan-07 well has been used to construct the synthetic seismogram for the accurate identification of horizons by seismic to well tie, Upper Goru and C-Interval are successfully identified. The reservoir rock is analyzed using petrophysical analysis of the three wells Sawan-01, 07, and 08. Each well's coordinates, depth, and formation TLG are all specified (Table 1.2).

Table1. 2: Showing well coordinates, well depths and formation TLG of wells Sawan-01, Sawan-07 and Sawan-08

Well Name	Latitude_N	Longitude_E	Well depth (m)	Formations (m) TLG
Sawan_01	26.991830	68.906994	3583	2692
Sawan_07	26.999291	68.923403	3398	2688
Sawan_08	27.009162	68.933396	3429	2695

Detail of formation tops are given in (Table 1.3)

Table1. 3: Formation tops of wells Sawan-01, Sawan-0 and Sawan-08

Formations	Sawan-01	Sawan-07	Sawan-08
Habib Rahi	274	284	283
Ghazij	301	303	304
Sui Main limestone	1118	1117	1120
Ranikot	1262	1260	1259
Upper Goru	2449	2443	2433
Lower Goru	2697	2690	2715
D Interval	3181	-	3195
C Interval	3247	3242	3240
B Interval	3459	-	-

1.5 Methodology

In this research, a comprehensive methodology aimed at improving reservoir characterization has been used. The approach began with a thorough examination that incorporated multiple potential scenarios and an extensive review of existing literature to mark the specific horizons of interest. These horizons were accurately delineated on in-line seismic sections, with careful consideration of prior information and the consistent patterns of reflection. Moreover, creation of a Synthetic Seismogram has been done by using density and sonic logs extracted from well Sawan-01. Spatial time maps for the D-, C-, and B-Sands horizons has been generated for identification of structures.

Additionally, conducting of in-depth petrophysical analysis of wells Sawan-01 and Sawan-07 has been done for demarcation of hydrocarbon bearing zone by using well log data

The DT4S log in well Sawan-01 has been successfully predicted using a machine learning technique. well log Data taken from the wells Sawan-07 and Sawan-08 were used to train this method.

Reservoir characterization relies on cross-plot analysis of elastic parameters, incorporating the predicted DT4S log from well Sawan-01. This approach provides a comprehensive understanding of the reservoir's properties, facilitating informed decisions in the field of reservoir characterization.

Flowchart of methodology is given in (Figure 1.2).

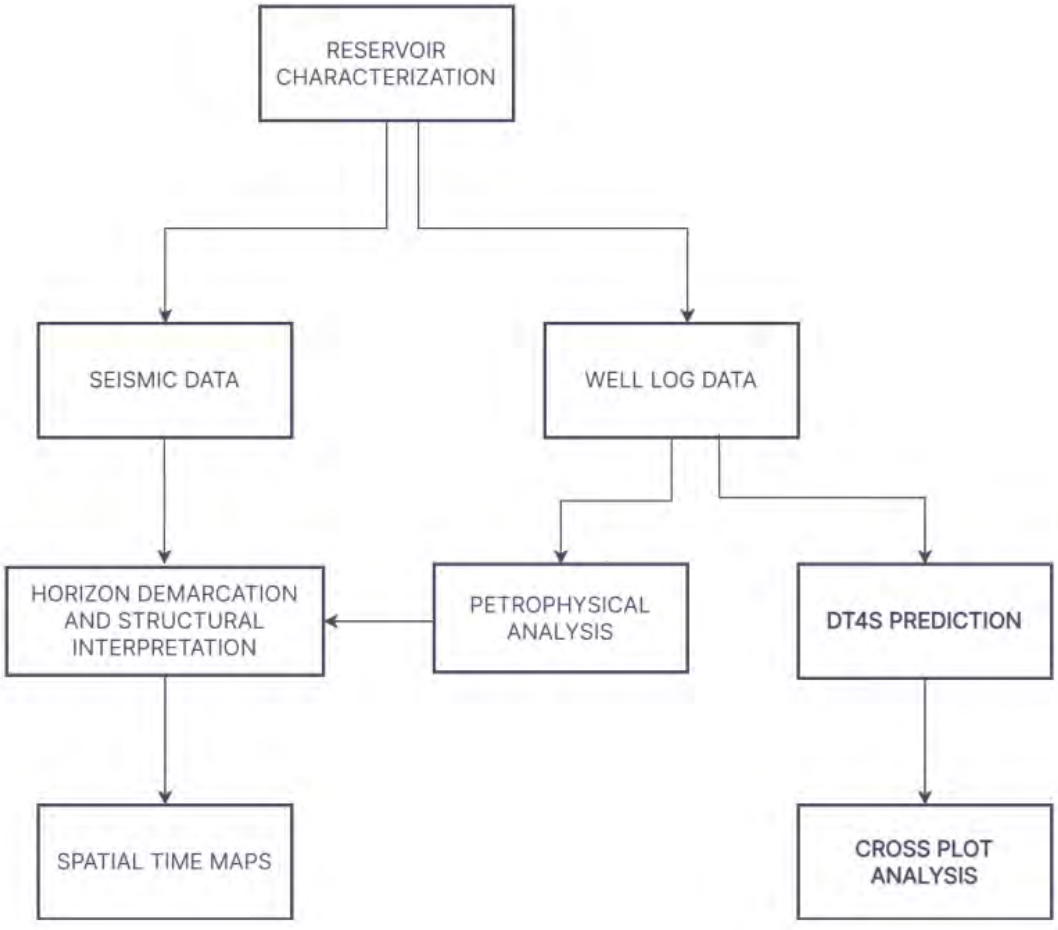

Figure 1. 2: Flowchart of methodology used in this research

CHAPTER 2

REGIONAL GEOLOGY AND TECTONICS

2.1 Introduction

Geological information on any area is essential since it provides the most details about that area's geological features, which may be used to identify depositional sequences, structural types, modes of deformation, and different types of rocks etc. According to the study zone's location, it is in Pakistan's Lower Indus Basin's Khairpur district. In the area of Sukkur and Khairpur, there are some Eocene carbonates and deserts of sand on the eastern edge of the Khairpur zone (Afzal et al., 2005). In general, the Jacobabad High in the north, the India Shield in the east, the Suleiman Thrust and Fold Belt in the west, and the Karachi Embayment in the south of Sawan Area are the boundaries of the study area (Kadri, 1995).

2.2 Petroleum Play

In Lower Goru C interval is producing reservoir which contains sands and in fairway it is trending in northeast-southwest (Afzal et al., 2005). The most recent inversion process occurred in the Eocene, and the trapping mechanism is directly related to generation, migration, and then prevention (Ahmed et al., 2004). Lower Cretaceous shale from a shelf marine deposit serves as the source rock for the Sembar formation. In source rock, which is type-III kerogen, there is nearly 0.5-1.7% total organic content that is present during the gas generation window in the late Cretaceous to early Tertiary (Wandrey et al., 2004).

2.3 Tectonic Settings

Most reservoir zones in the Lower and Central Indus basins are situated on multiple structural highs, including the Mari and Jacobabad highs etc. For a significant accumulation of hydrocarbons during migration, these structures are substantially more important (Kazmi and Jan, 1997). The base unconformity known as the First Tertiary Unconformity was formed during the Cretaceous to Tertiary (K-T) uplift, and nearly all faults end here. The majority of the faults in the Cretaceous strata are NW-SE orientated and range in number from a single Chiltan to many in the Lower and Upper Goru levels. All of these characteristics are the result of the first unconformity generation, which occurred when the Indian and Eurasian plates interacted transtensionally and the Indian plate rotated counterclockwise. Second uplift happened in Late Eocene Oligocene in Central and

Lower Indus basins. Regular episodes of disturbance at the structural highs may have been caused by the subsequent phases of thrust loading in the west and northwest. The potential for secondary migration and reservoir charge, as well as the ultimate advancements trap shapes were most likely occurred at this time (Kazmi & Abbasi, 2008)

Cretaceous times because of extensional regime development of titled fault is evident at large scale in eastern parts of Lower Indus sub-basin. Cretaceous and older strata were destroyed by these tilted fault systems with normal dip according to analysis of seismic reflectors in the area of interest. whereas these faults from Cretaceous eras have strike from N 30° W to N 50° W

The system of tilted fault traps existed at the time of generation of hydrocarbons. The trapping mechanism for hydrocarbons close to the Miano block is also governed by traps created by faults in Lower Goru sandstones (Kemal, 1991).

2.4 Stratigraphic Elements

The generalized regional stratigraphy of the Lower Indus basin and demonstrates that the rocks' ages range between the Jurassic to the Tertiary (Figure 2.1). In Sawan, the Jurassic limestone of Chiltan is overlain by the Triassic rock sequence, and the Chiltan limestone is overlain by thick strata of Sember formation shales (Zaigham and Mallick, 2000). Just above the Sembar formation, which serves as the primary source strata in the area, is the Lower Goru Formation. The Sembar formation is overlain by Basal sands, Basal sands are overlain by Lower shales, Lower shales are overlain by Middle sand, and the fifth unit of Lower Goru is present between Upper and Middle sand as shown in (Figure 2.2). Lower Goru is further divided into five divisions. In the case of the upper sands, the deposition environment was shallow to deltaic marine conditions. As a result, in the lower part of the Indus basin. It is regarded as an excellent reservoir (Alam et al. 2002).

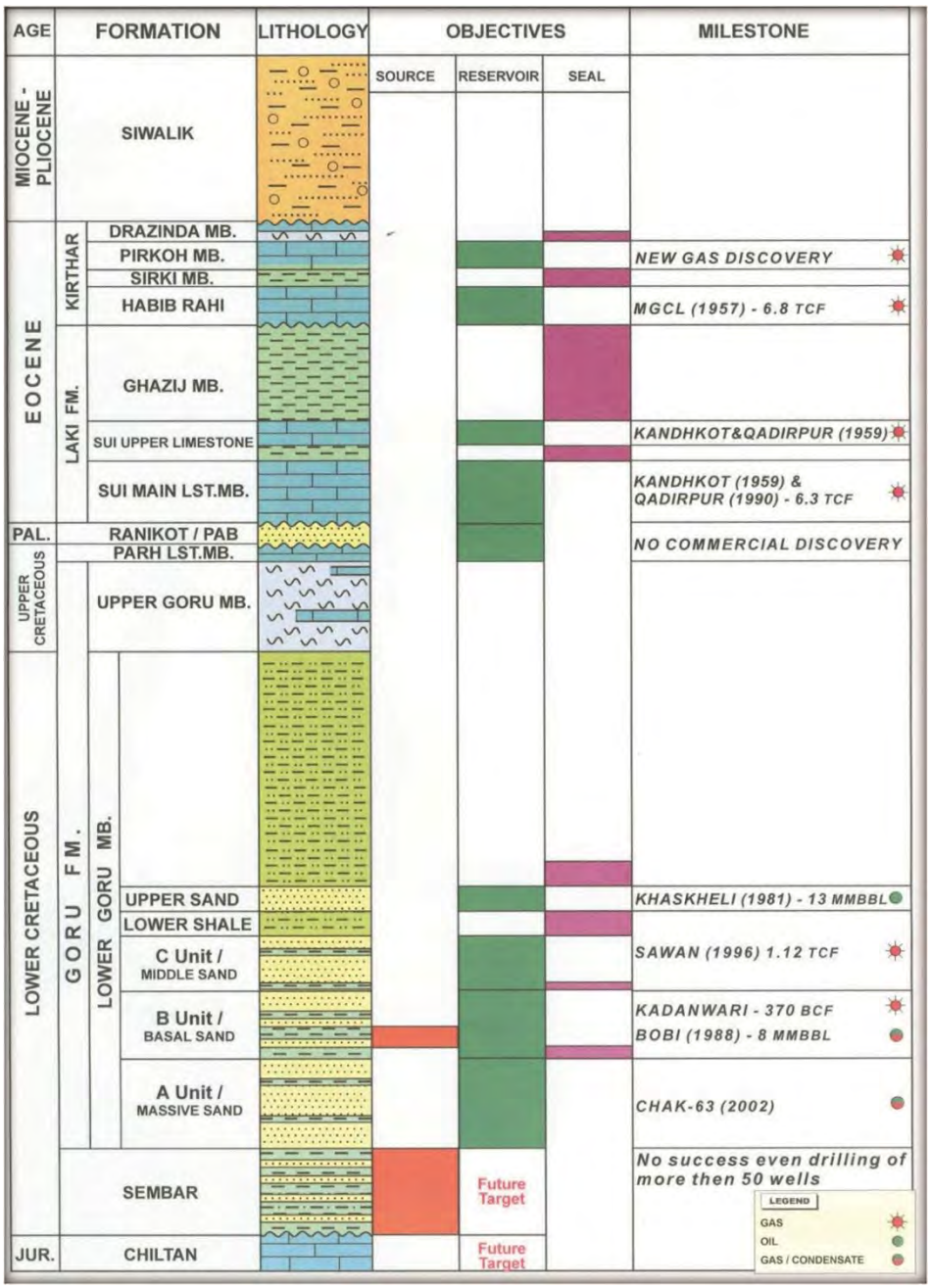

Figure 2. 1: Stratigraphic Column of Sawan area, (Gul et al., 2023)

The main documented source strata of petroleum found in the southern part of the Indus basin are interbedded Lower Goru and Sembar formation shales from the early Cretaceous. In contrast, transgressive marine shales from the Upper Paleocene that are buried make up the secondary source strata in the western region of the Southern Indus Basin (Zaigham and Mallick, 2000).

The Lower Goru Formation is the source of the basal sands in the study area's target formation. Notably, the area has large, varied-thickness sand deposits that act as a significant reservoir and are currently being exploited. Although this theory has not yet been confirmed, it is conceivable that the reservoir rock may have covered the basal sands, a portion of the Lower Goru Formation (Kadri, 1995).

The Cretaceous Sembar Formation, which is made up of transgressive shales, is the main component of the seal rocks in the southern part of the Lower Indus Basin. There are also formations from the Tertiary era, including Kirthar, Laki-Gazij, and Bara-lakhira (Zaigham and Mallick, 2000). The layers of Lower Goru shales function both as horizontal and vertical seals. These shales effectively create a seal even across faults, though it's evident that faults sometimes also contribute to sealing. The Talhar shales, along with the shales from the Upper and Lower Goru, play a crucial role as the primary seal rocks for the reservoir sands of the Lower Goru (Kadri, 1995).

2.5 Hydrocarbon Play Area

Total petroleum play includes three important Formations as follow:

- Sembar Formation
- Goru Formation
- Chiltan Formation

2.5.1 Source Rock

In this region and various other sections of the Indus Basin, the key hydrocarbon generating source rock is the Cretaceous age Sembar Formation. This formation consists of organic rich shales that were deposited in a shelf marine environment (Ashraf et al., 2020).

2.5.2 Reservoir Rock

The Sawan Gas Field contains reservoirs made up of Upper Cretaceous Lower Goru C sands. The structural configuration of the Sawan area, interpreted through seismic analysis, indicates a low stand wedge formation. This wedge is the outcome of a decline in relative sea level or a forced regression phenomenon (Ahmad et al., 2004). This distinctive low stand wedge is readily recognizable in seismic data. Identifying the point where the reservoir sands blend with the more typical shelf facies of the C interval presents more challenges.

In the Sawan region, the reservoir strata consist of Basal Sand, which is a sub-unit of the Lower Goru formation. These layers of sand within the Basal Sand sub-unit exhibit varying thicknesses and are responsible for hydrocarbon production. While the possibility of an overlying reservoir on the Basal sands in the Lower Goru cannot be dismissed, confirming the presence of this reservoir is still a topic requiring further validation (Kadri, 1995).

The main focus for hydrocarbon production reservoirs lies in the Lower Goru sandstones. These sandstones originated from rifting events towards the end of the early Cretaceous period. They formed due to the erosion of the Indian shield and the subsequent depositions of sediments via sequences of barrier bars and deltaic sands across different regions of the middle and lower Indus Basin, extending down-dip to the West (Hussain et al., 1991).

2.5.3 Seal Rock

The Sawan gas accumulation is a trap that combines both structural and stratigraphic elements, identified at the top of the C interval. This trap exhibits a small four-way dip closure oriented northeast-southwest, and its seismic configuration gradually tapers towards the north, northwest, and southeast. Spanning a length of 17 kilometers, this feature is intersected by a minor fault in the southwest and occupies an approximate mapped area of 63 square kilometers. Stacked sand bodies within the Sawan C sand reservoir are composed of distributary channels, mouth bars, and upper shore face belts originating from a forced regressive deltaic setting. These sand bodies are characterized by their arrangement (Dar et al., 2021)

Around the points where freshwater enters the vicinity, early diagenetic Fe-Chlorite coatings envelop coarse and medium sand grains, preserving porosity and permeability as burial deepens. The variation in Paleo-topography and accommodation space resulted in the deposition of a thicker sedimentary pile in the Sawan north area, in contrast to the thinner sand sheets in the south. While a considerable thickness of these stacked sand bodies survived erosion during regression and subsequent transgression in the Sawan north area, only a partial section has been preserved in the south. This discrepancy leads to significant uncertainty in predicting the presence of reservoir sands in the Sawan south region (Ahmad & Ghazi, 2022).

Overlying the Sawan reservoir sands, there exists an extensive sequence of shales and marl that function as a regional top seal. Within the Lower Goru C interval, the shales also serve as bottom and lateral seals (Munir et al., 2011).

The major source of hydrocarbon production in the studied zone is attributed to structural traps. These traps, resulting from both the extensional regime due to rifting events and the geometry of horst and graben features, are formed by tilted faults (Khan et al., 2016). The timeline relationships between hydrocarbon generations, trap formation, exclusion, migration, and the entrapment of hydrocarbons are not uniform across the entire Indus Basin. Consequently, various trapping mechanisms coexist, including negative structures and slanted fault blocks.

In the southern part of the Indus Basin, the seal rocks encompass the Kirthar, Laki-Ghazij, and Bara-Lakhra formations (Zaigham & Mallick, 2000). Effective primary vertical and horizontal sealing in both the Lower and Upper parts of the Indus Basin are provided by the inter-bedded Lower Goru Shales. The uppermost sands of the Lower Goru are capped by the sealing capacity of the Upper Goru formation (Kadri, 1995).

CHAPTER 3

3D SEISMIC DATA INTERPRETATION

3.1 Inroduction

The process of converting seismic reflection data into a structural representation of the subsurface, or extracting valuable information and figures from processed data gathered after conducting and refining seismic surveys, is referred to as Seismic Data Interpretation. This practice holds a vital role in the exploration of hydrocarbons (Avseth et al., 2005), aiming to uncover subsurface details. The insights extracted from seismic data unveil horizon boundaries and the positioning of sub-surface faults (Hilterman, 2001).

Based on the interpreted seismic data, informative models are constructed to enhance the comprehension of the subsurface. This involves identifying and annotating consistent horizons, which aids in the cartography of geological structures and stratigraphy. This in turn contributes to grasping the manner in which reservoirs are formed, the accumulation of hydrocarbons, and the dimensions and volume of reserves (Stewart, 1984). The demarcation of horizons are pivotal tasks carried out during the interpretation of seismic data.

Historically, all reflectors have been marked using shot points in conjunction with vertical seismic sections. The timing for these designated points is determined through the observation of various reflection times. Once these times are established, they are used to create a map that highlights local features with higher elevations or structures, such as local anticlinal features. These features hold the potential for substantial hydrocarbon reserves (Sheriff, 2002).

3.2 Types of Seismic Data Interpretation

Interpretation of processed seismic data extracted from the extensive processing of the acquired seismic data is carried out by two different methods

3.2.1 Structural Interpretation

The primary objective of structural interpretation is to chart structural traps within the seismic section, with the purpose of identifying formations that harbor hydrocarbons. Disrupted reflections distinctly reveal faults, while undulating patterns indicate folded beds. These subsurface

configurations serve as traps capable of containing hydrocarbons, representing a promising reservoir potential (Coffeen, 1986)

3.2.2 Stratigraphic Interpretation

During the process of stratigraphic interpretation of seismic data, seismic sequences are identified and mapped. This method facilitates the comprehension of the depositional setting of various lithological units through an in-depth examination of multiple attributes of seismic facies, which serve as reliable indicators of the depositional environment (Mitchum & Vail, 1997)

Simultaneously, investigating variations in reflection patterns is employed to deduce shifts in both the depositional and stratigraphic context, all while considering the potential for hydrocarbon deposits. Seismic reflection amplitudes and velocities emerge as valuable tools for conducting thorough analyses (Cross & Lessenger, 1988).

In the process of identifying unconformities, drainage patterns play an exceptionally crucial role. Once these unconformities are recognized, it becomes considerably easier to construct a comprehensive picture of the depositional environment within a specific area. Notably, unconformities stand out as significant stratigraphic traps (Telford et al., 1990).

3.3 Seismic Interpretation Workflow

Multiple steps were used in the interpretation of the available seismic data. There are several steps in each phase that were completed using software tools. (Figure 3.1) illustrates in detail how the seismic lines and SEG-Y navigation data were used for interpretation using HRS software. Horizons of interest were demarcated. Horizons were identified using synthetic seismogram derived from well data using careful examination of seismic sections and knowledge of the research area's geologic history.

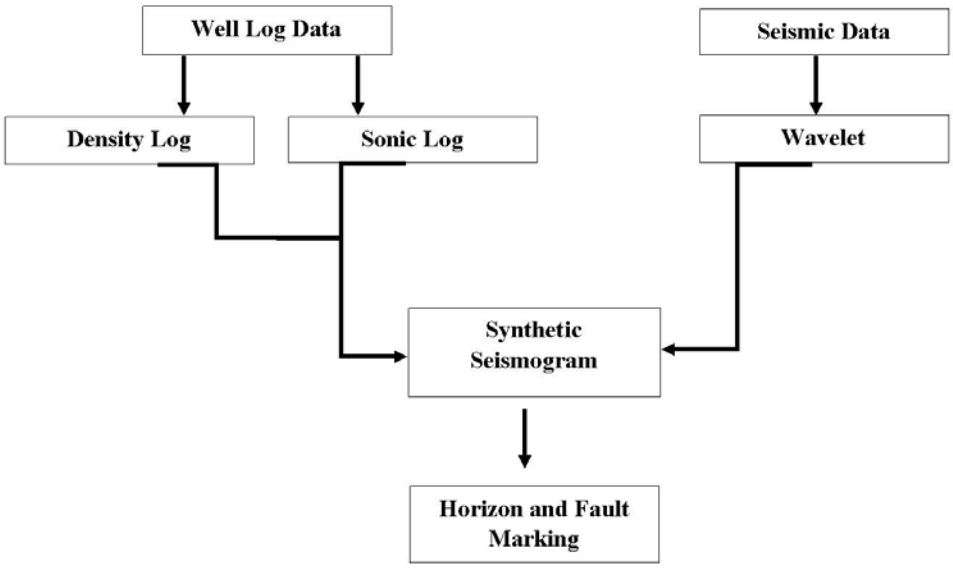

Figure 3. 1: Workflow of Seismic Interpretation

3.4 Base Map:

The base map of the study area is shown (Figure 3.2), where the well locations and the orientation of the inlines and crosslines of the 3D seismic survey are indicated and used for research purposes in the Sawan area.

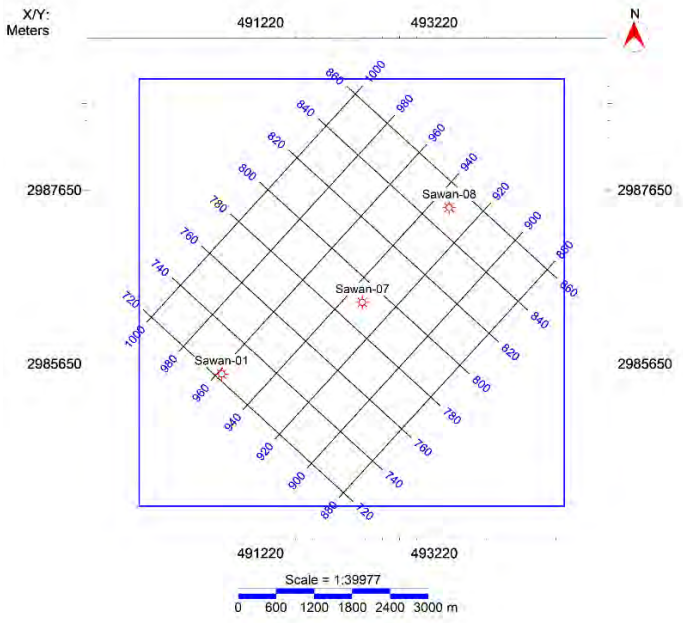

Figure 3. 2: Base map depicting inlines, crosslines, and well locations of the study area

3.5 Synthetic Seismogram

Seismic interpretation involves utilizing modeled traces from synthetic seismograms, which play an important role in establishing correlations between seismic reflections and subsurface stratigraphy encountered. While seismic data is presented in a time scale, well tops are specified in depth, making it challenging to mark horizons in time units. The purpose of generating synthetic seismograms is to establish a correspondence between two-way travel time and depth, enabling the marking of horizons.

Typically, a synthetic seismogram is derived using a sonic log. Ideally, a density log is also recommended, though its availability can be an issue. Reflectivity series is generated using DT (sonic) and RHOB (density) logs. This series is then convolved with the source wavelet to create the synthetic seismogram. In the process of linking seismic sections with borehole geology, synthetic seismograms serve as valuable tools due to their capacity to directly connect patterns of seismic reflections with studied subsurface lithology (Handwerger et al., 2004).

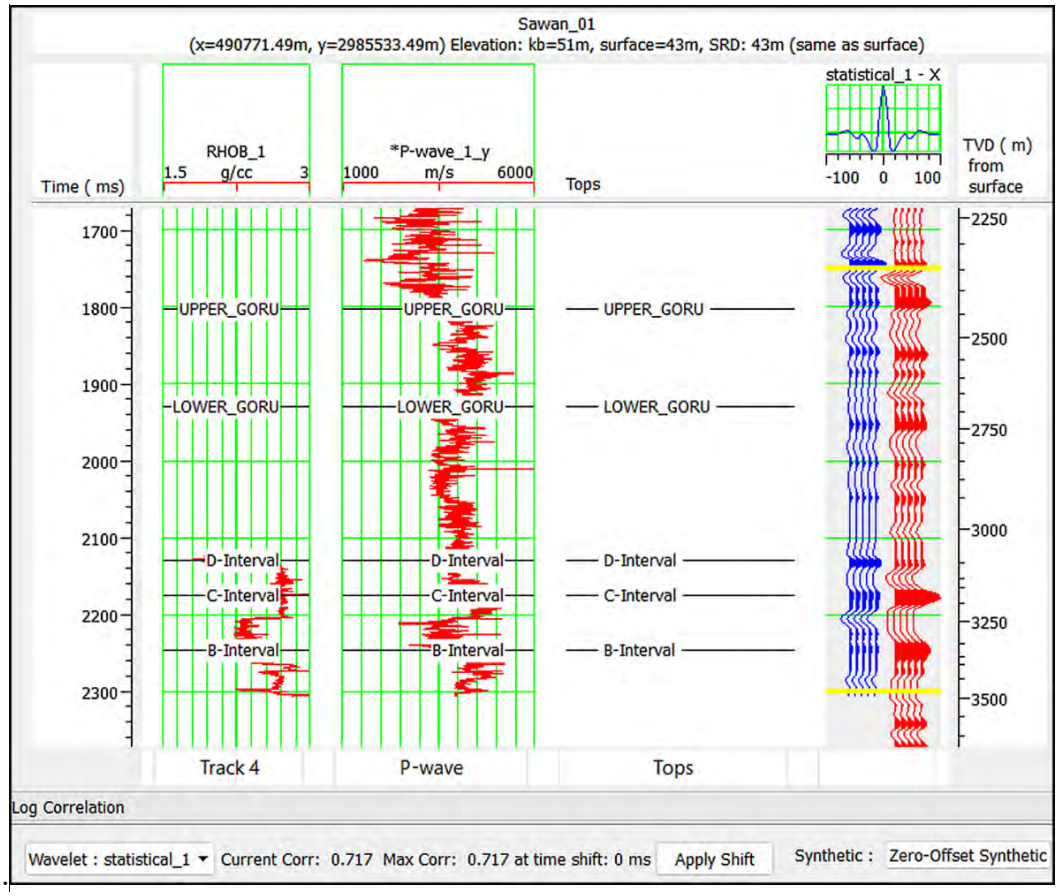

Figure 3. 3: Synthetic seismogram for picking horizons on well Sawan-01

To pinpoint the exact positions of specific horizons in the Sawan area, a synthetic seismogram is generated using data from Sawan-01 located at Inline-721. This synthetic seismogram (Figure 3.3), is prepared along cross-line 960. The primary focus is on horizons identified in a seismic section. The horizon marked at the bottom of the (Figure 3.3) corresponds to a segment of the Lower Goru formation, outlining the positions of the D Interval, C Interval, and B Interval, with the C Interval serving as the main reservoir within these horizons

3.6 Interpretation of Horizons

The fundamental objective of interpretation involves recognizing different horizons that act as boundaries between geological formations (Wrona et al., 2018). A change in lithology results in an acoustic impedance contrast, which is responsible for the initiation of primary reflections. These reflections are designated on seismic sections to understand the connection between geology and the seismic data (Badly, 1985). The seismic data is further utilized to mark the reflectors corresponding to the D Interval, C-Interval, and B Interval after generating synthetic seismograms (Figure 3.4).

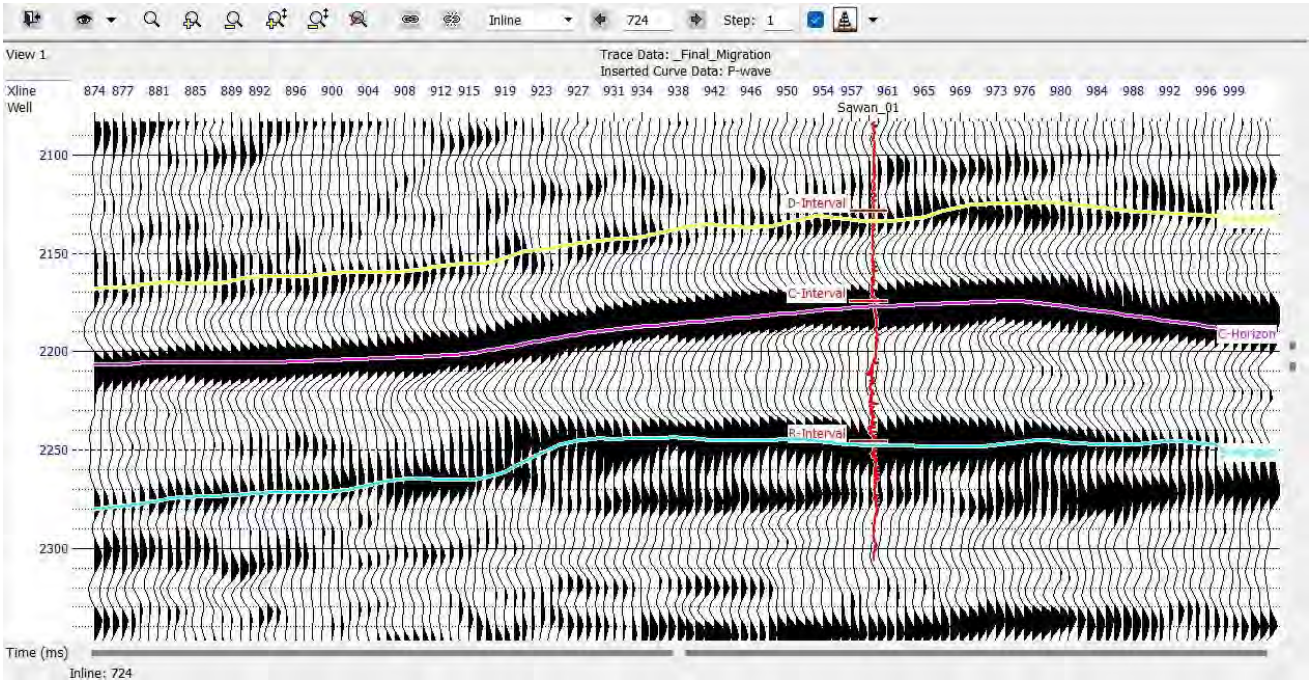

Figure 3. 4: Seismic Section depicting the marked horizons after seismic to well tie with Sawan-01 well. The horizons of interest are D-sand, C-sand and B-sand of the Lower Goru Formation.

3.7 Spatial Time Slices of 3D Seismic Data:

Crucial insights, such as formation dip information, the nature of folding and faulting, structural relief of formations, and reliable data regarding formation slopes, can be extracted from these contour maps. The data used to create these time slices is obtained from the selected times of pertinent reflectors that exhibit potential for hydrocarbons on the seismic section.

The 3D seismic cube (covering 10 square kilometers) provided by DGPC does not prominently exhibit faults or other significant structures. The upper portion of the Lower Goru formation comprises an abundance of shales, which function as a seal across the entire region. The lowermost layer of the Lower Goru formation contains inter-bedded sands and shales. Within the sands of the Lower Goru, various categories like B, C, and D sands are distinguished. The Sembar formation, positioned beneath the Lower Goru, serves as a source rock, while the Upper Goru Formation lies above the Lower Goru.

3.7.1 Time Map of D-Sand

The Lower Goru formation in the Sawan area contains reservoir rock that meets the criteria for being a high-quality reservoir sand in that region, confirming its potential. Additionally, the underlying Sembar Formation serves as the source rock for these sands.

The time slice of the D-Sand is presented. The color gradation from orange to cyan indicates shallower to deeper parts as shown. (Figure 3.5). Notably, the northern to northwest corner and the center of the map display shallower depths, while deeper portions are located in the northeastern and southern corners as shown from color code. This indicates that it is dipping towards south east.

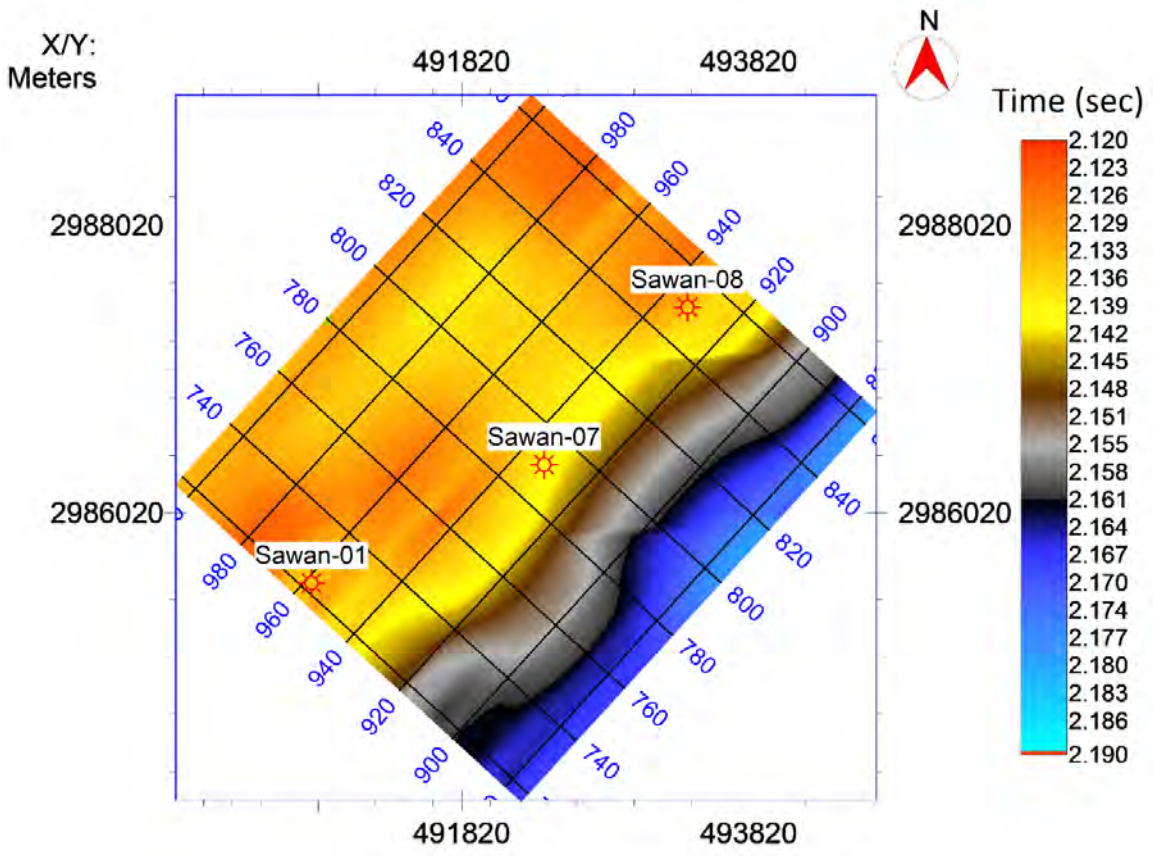

Figure 3. 5: Time Slice of D-Interval Sand depicting dip towards SE

3.7.2 Time Map of C-Sand

The Lower Goru formation in the Sawan area functions as reservoir rock, satisfying the requirements for a reservoir sand with high-quality potential in that particular region, thereby confirming its suitability. Additionally, the underlying Sembar Formation serves as the source rock for these sands.

The time slice of the C-Sand is presented. The color gradation from orange to cyan indicates shallower to deeper parts as shown. (Figure 3.6). Deeper portions are located in the northeastern and southern corners while northern to northwest corner and the center of the map display shallower depths as shown from color code. This indicates that it is dipping towards south east but more towards east.

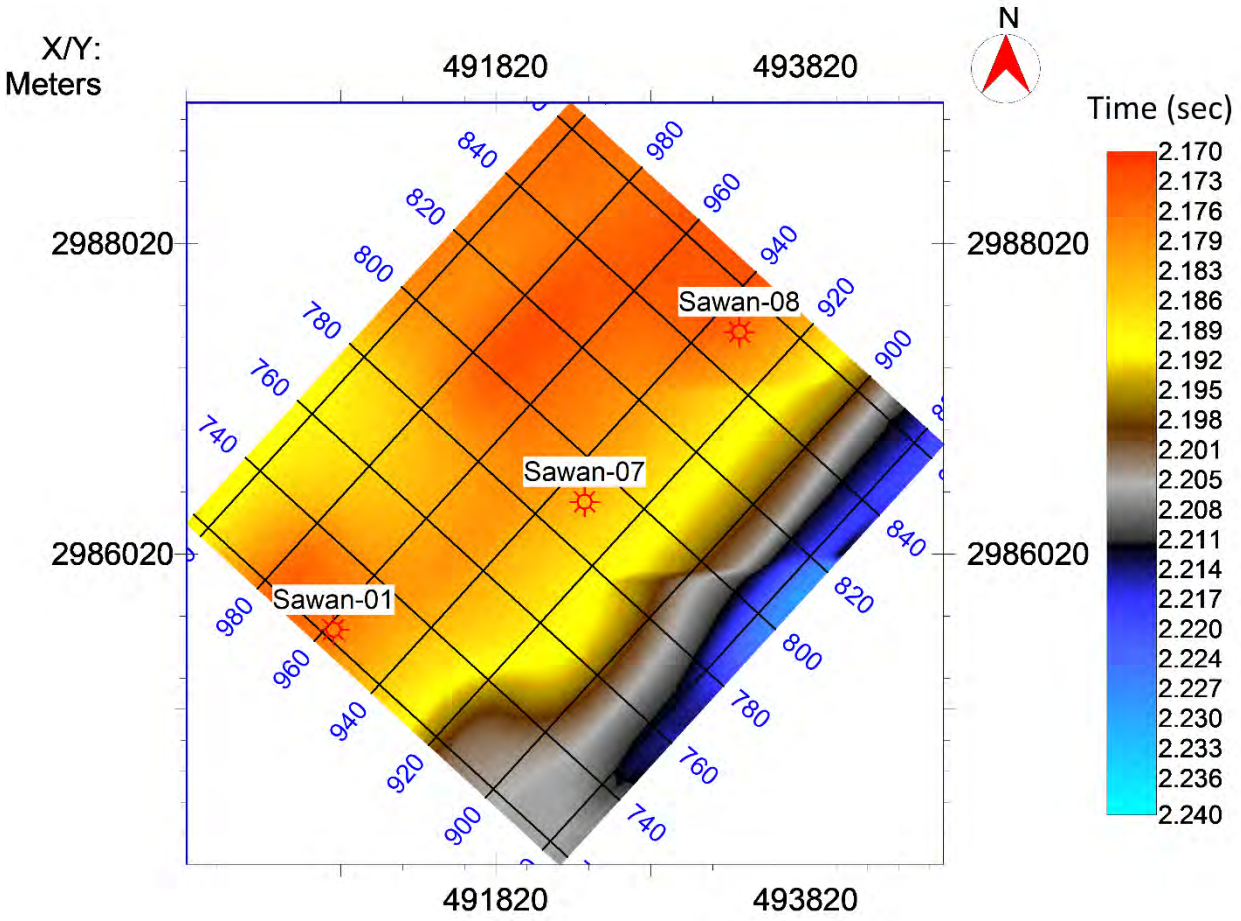

Figure 3. 6: Time Slice of C-Interval Sand depicting dip towards SE

3.7.3 Time Map for B-Sand

In the Lower Goru formation, the B Interval functions as the reservoir stratum in the specified area of interest. The time slice of the B-Sand is presented (Figure 3.7). The color gradation from orange-yellowish to cyan indicates shallower to deeper parts as shown. The northern to northwest corner and the center of the map display shallower depths, while deeper portions are located in the northeastern and southern corners as shown from color code. This indicates that it is dipping towards south east.

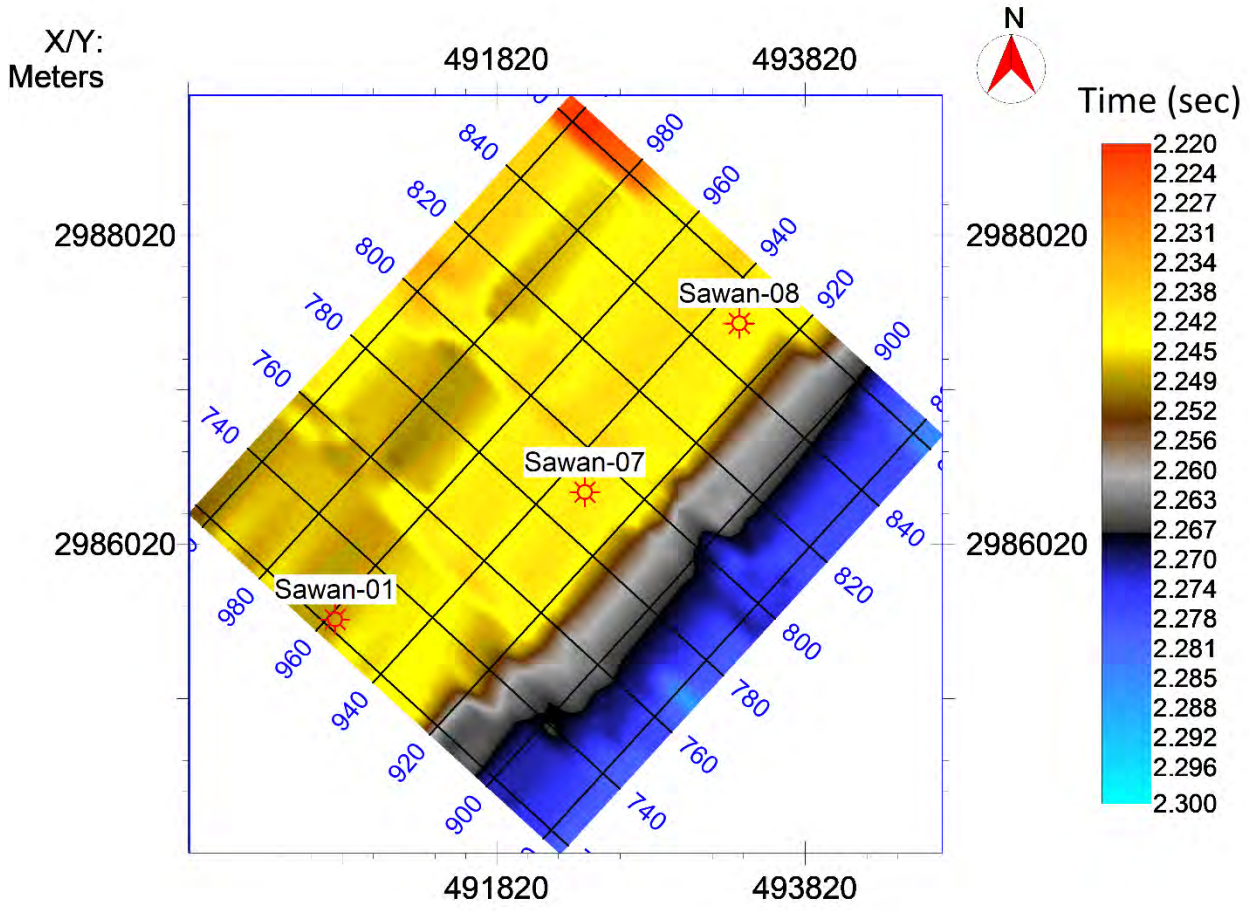

Figure 3. 7: Time Slice of B-Interval Sand depicting dip towards SE

CHAPTER 4

PETROPHYSICAL ANALYSIS

4.1 Introduction

Petrophysical analysis provides a comprehensive understanding of fluids, their identification, and quantification within reservoir rocks (Ali et al., 2015). This analysis revolves around well logs, which are continuous records of various geophysical parameters plotted against depth within the wellbore. Essentially, well logs are employed to verify lithologies and the presence of fluids within rocks. The primary objective of petrophysical analysis is to measure distinct rock properties and their correlations with fluids (Donaldson and Tiab, 2004). By integrating petrophysical outcomes with rock physics, potential prospect zones and non-prospect areas are discerned, enabling geoscientists to derive essential insights. The characterization of reservoirs is founded upon petrophysical findings.

A plethora of newly developed geophysical well logs now exists. Geophysical logs employing highly specialized tools are notably prominent due to their structure. These logs can be conducted prior to well casing, immediately following drilling, a technique termed Measurement While Drilling (MWD), and during the process of formation drilling, referred to as Logging While Drilling (LWD). Typically, MWD logs ascertain the deviation of directional wells, whereas LWD involves measurements such as resistivity and density.

To accurately delineate the likely hydrocarbon zones, one needs to be aware of reservoir characteristics like shale volume, porosity, water saturation, and hydrocarbon. By combining petrophysics and rock physics, geologists and geophysicists can better understand the dangers and potential in the region. Petrophysics is a frequently used method for characterization of reservoirs due to the use of well measurements to enhance reservoir representation (Hussain et al., 1991).

4.2 Objectives for Log Interpretation

The quantitative examination of well logs can be used for estimation of reservoir parameters including:

- Porosity
- Saturation of water
- Fluid type (oil, gas, water)
- Reservoir type (lithology)
- Productivity (permeability)

The major goals of logging are:

- To acquire data for analyzing petroleum reservoirs.
- To assist in well testing, completion, and rehabilitation.

Following parameters has been followed to calculate the oil reserves in a reservoir:

- The oil-bearing formation's thickness.
- Formation porosity.
- Saturation of oil.
- The oil-bearing strata's lateral extension.

4.3 Methodology

Petrophysical studies have been conducted to characterize the reservoir properties in the study area. The well logs data from Sawan-07 well has been employed to analyze various formations. The following parameters have been determined using the log data: shale volume, density porosity, effective porosity, total porosity, water saturation, and hydrocarbon saturation. The wireline log data includes Gamma ray log (GR), Caliper log, Spontaneous log (SP), Laterolog deep (LLD), Laterolog Shallow (LLS), Micro spherically focused log (MFSL), Neutron log (NPHI), Density log (RHOB), and Sonic log (DT). These wireline logs have been categorized into three distinct tracks based on their approximation of reservoir properties and working principles (Figure 4.1). The first track is referred to as the lithology track, the second track is known as the resistivity track, and the third track is designated as the porosity track.

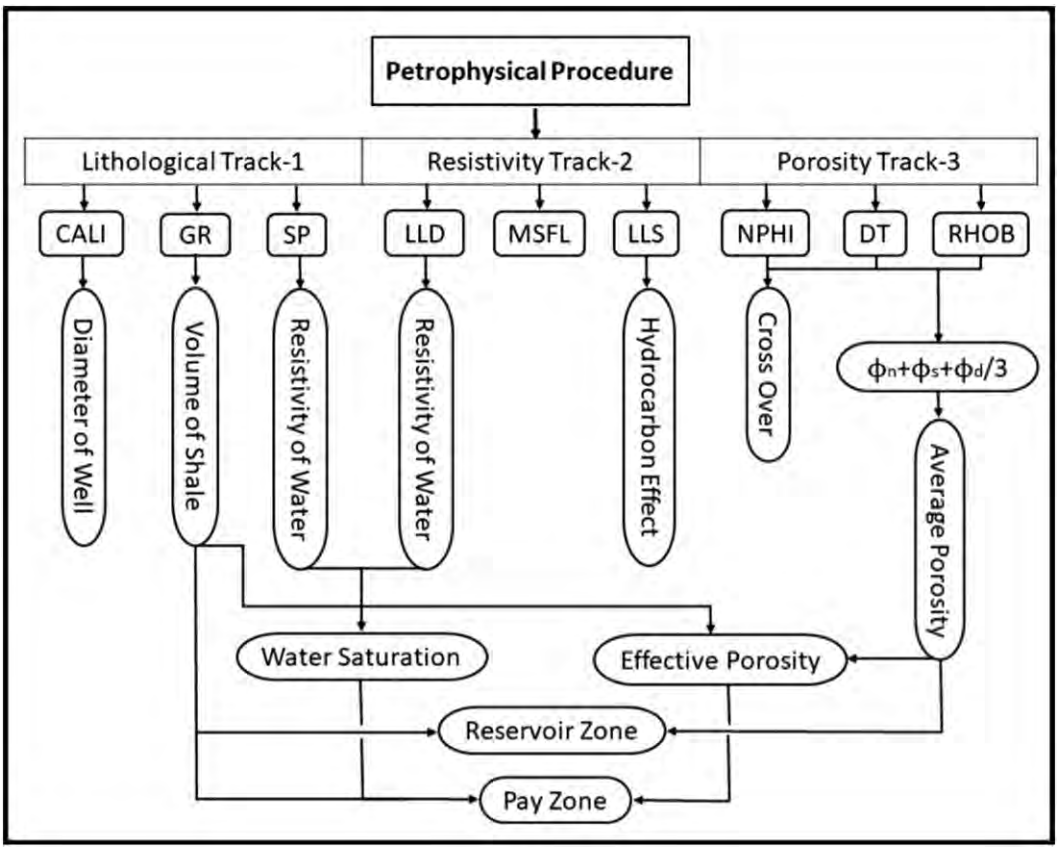

Figure 4. 1: Workflow of petrophysical analysis (Mughal & Akhter, 2021)

The Petrophysical analysis has been carried out in order to measure the reservoir description of the Sawan area using the borehole data of Sawan-01 and Sawan-07 well. The logs defined above will be used in order to calculate the reservoir parameter such as:

- Volume of shale (Vsh)
- Porosities (PHID, PHIT, PHIE)
- Water Saturation (Sw)
- Hydrocarbon Saturation (H.C)
- Net Pay or Net Reservoir

4.4 Volume of Shale

The GR log is used to calculate the volume of shale in petrophysical studies. The GR log is used to calculate the formation's radioactivity (Fanchi, 2005). This makes it very valuable for lithology identification because it provides the concentration of radioactive material that is present in the

formation. In carbonate and sandstone, the gamma ray has a low value, but in shale, it has a higher value. Shale has a higher concentration of radioactive material than sand and carbonates. As a result, it is possible to tell apart reservoir rocks from other types of rocks (Rider, 1990).

The volume of the shale is calculated using the calculations below given by (Miller, 1986)

$$V_{sh} = \frac{GR\log - GR\min}{GR\max - GR\min} \quad (4.1)$$

Where,

V_{sh} = volume of shale

GR_{log} = GR values from log

GR_{min} = value of GR in clean zone

GR_{max} = value of GR in shaly zone.

4.5 Porosity

Porosity is the most important property to estimate for petrophysical analysis. Utilizing density, sonic, and neutron logs, porosity can be calculated. The nuclear measurements are the Neutron and Density logs, whereas the acoustic measurement is the Sonic log (Abdelwahhab & Raef, 2020).

Following are the porosity parameters:

- Density Porosity
- Sonic Porosity
- Total Porosity
- Effective Porosity
- Neutron Porosity (Given)

4.6 Density Porosity

The density log serves as a porosity log, measuring the electron density of the formation (Asquith et al., 2004). The electron density of the formation is interconnected with the bulk density of the formation. This value is derived from the combination of the fluid density, multiplied by its corresponding relative volume, and the matrix density, which is then multiplied by its relative volume.

By leveraging the density log, it becomes possible to determine the accurate porosity of the formation, provided that the matrix densities within the formation or the rock type are already known (Asquith et al., 2004).

$$\text{RHOB } \phi = \frac{\text{RHOB}_{\text{mat}} - \text{RHOB log}}{\text{RHOB}_{\text{mat}} - \text{RHOB}_{\text{fluid}}} \quad (4.2)$$

Where,

RHOB ϕ = the density porosity,

RHOB_{mat} = density log, RHOB mat is value of matrix density,

$\text{RHOB}_{\text{fluid}}$ = density of fluid

The value of RHOB_{mat} 2.65 g/cm^3, which is for sandstone and $\text{RHOB}_{\text{fluid}}$ is 1 g/cm^3 (Glover, 2000).

4.7 Sonic Porosity

Sonic logs are employed to measure the interval transit time (Δ t) of compressional sound waves as they traverse through the formation. This interval transit time is closely linked to the porosity of the formation. The unit of measurement used is typically expressed in microseconds per foot or microseconds per meter (Asquith et al., 2004). The calculation of the porosity of the formation can be carried out using the subsequent formula given by (Asquith et al., 2004):

$$\phi s = \frac{\Delta \text{Tlog} - \Delta \text{Tmatrix}}{\Delta \text{Tfluid} - \Delta \text{Tmatrix}} \quad (4.3)$$

Where

ϕs = the sonic porosity

ΔTmatrix = interval transient time of the matrix

ΔTlog = interval transient time of formation

ΔTfluid = transient time of the fluid

The interval transient time of the formation depends upon the matrix material, its shape and cementation (Wyllie et al., 1956). If fluid (hydrocarbon or water) is present in the formation,

transient interval time is increases and this behavior shows increase in porosity which can be calculated by using sonic log (Asquith et al., 2004).

4.8 Average Porosity

The average porosity is determined by summing up the porosities obtained from various logs and then dividing this sum by the number of logs from which porosity is calculated. In the context of the C-Interval reservoir, which is the focus of interest, all the available logs are interpreted to calculate the average porosity. The relationship for calculating the average porosity is as follows:

$$\phi\text{avg} = \frac{\phi\text{n}+\phi\text{d}+\phi\text{s}}{3} \quad (4.4)$$

Where,

ϕavg = average porosity calculated from the available porosities

ϕn = neutron porosity

ϕd = density porosity

ϕs = sonic porosity.

4.9 Effective Porosity

Effective porosity is defined as the ratio of the volume of interconnected pore spaces within a rock unit to the total volume of the rock, while accounting for the exclusion of shale effects. In areas rich in shale, the effective porosity tends to be zero due to the low permeability of shale. Effective porosity is particularly useful for delineating saturated zones within a reservoir. The calculation of effective porosity follows the formula provided by (Asquith et al. 2004).

$$\phi\text{eff} = \phi\text{avg} * (1 - \text{V}_{\text{sh}}) \quad (4.5)$$

Where ϕeff is effective porosity which is to be calculated, ϕavg represent the average porosity and Vsh represent volume of the shale.

4.10 Neutron Porosity

The neutron log is a measurement that is responsive to the presence of hydrogen atoms within a formation, aiding in the determination of formation porosity. In formations with high porosity, the count rate on the neutron log tends to be low, and conversely, in formations with low porosity, the count rate is higher. The neutron porosity measurement is included in the well log data and is calculated based on the depth of the well (Mills et al., 1988).

4.11 Water Saturation (Sw)

Water saturation in a formation refers to the proportion of the pore volume that is occupied by water within that formation. Archie's equation is particularly accurate for clean sands, providing a reliable result (Figure 4.2). following equation for Sw is given by (Archie, 1952).

$$S_w = \left\{\frac{a*R_w}{\phi^m*R_t}\right\}^{(1/n)} \quad (4.6)$$

Where

S_w = water saturation.

R_t= true resistivity of formation.

ϕ^m = effective porosity.

R_w = water resistivity value.

m = cementation factor.

a = tortuosity factor.

4.11.1 Resistivity of Water (Rw)

Calculation of resistivity of water (Rw) is key for water saturation. Numerous parameters like bottom hole temperature (BHT), surface temperature, water salinity in ppm and SP (Static) are important for valuation of water resistivity (Rw) (Amigun et al., 2012).

Two methods have been applied for resistivity of water:

- SP Method
- Pickett cross plot Method

SP Method

Formula for this method is as follows,

$$Ssp = -K * log\left(\frac{Rmf}{Rw}\right) \quad (4.7)$$

For K

$$K = 65 + 0.24 * T℃ \quad (4.8)$$

Static spontaneous potential (ssp) can be calculated through the equation given by (Miller, 1986) using spontaneous potential log readings. So then saturation of water in the formation can be calculated by the Archie equation (Archie, 1952):

$$S_w = \sqrt[n]{\frac{a * R_w}{\Phi^m * R_t}} \quad (4.9)$$

Where

Rw = resistivity of water calculated from above equation

Rt = true formation resistivity

n = saturation exponent

a = constant, in case of sand represents effective porosity

m = cementation factor

4.12 Hydrocarbon Saturation

Hydrocarbon saturation can be defined as "the pore in formation is filled with hydrocarbon". It can be calculated by using the following mathematical relation given by (Kamel & Mabrouk, 2002):

$$SH = 1 - SW \quad (4.10)$$

Where,

S_H = hydrocarbon saturation

S_W = Water saturation.

4.13 Well Logging Interpretation of Sawan-01

The Lower Goru formation encountered in the Sawan-01 well has been identified as a highly productive sand layer for hydrocarbon exploration. A specific area of interest within this formation has been identified based on distinct characteristics. These characteristics include a clear separation between the Laterolog Deep (LLD) and Laterolog Shallow (LLS) measurements, a significant crossover point between neutron porosity and density readings, and a low volume of shale (Miller, 1986) (Figure 4.2)

By considering these indicators, a zone of interest has been delineated, which is highlighted in the (Figure 4.2). The petrophysical analysis results obtained for this zone of interest have been compiled and presented in (Tables 4.1). These results reveal that the Lower Goru Formation primarily comprises clean sand with a notably high effective porosity ranging from 10% to 16%. Additionally, the water saturation levels are relatively low, falling within the range of 32% to 45%.

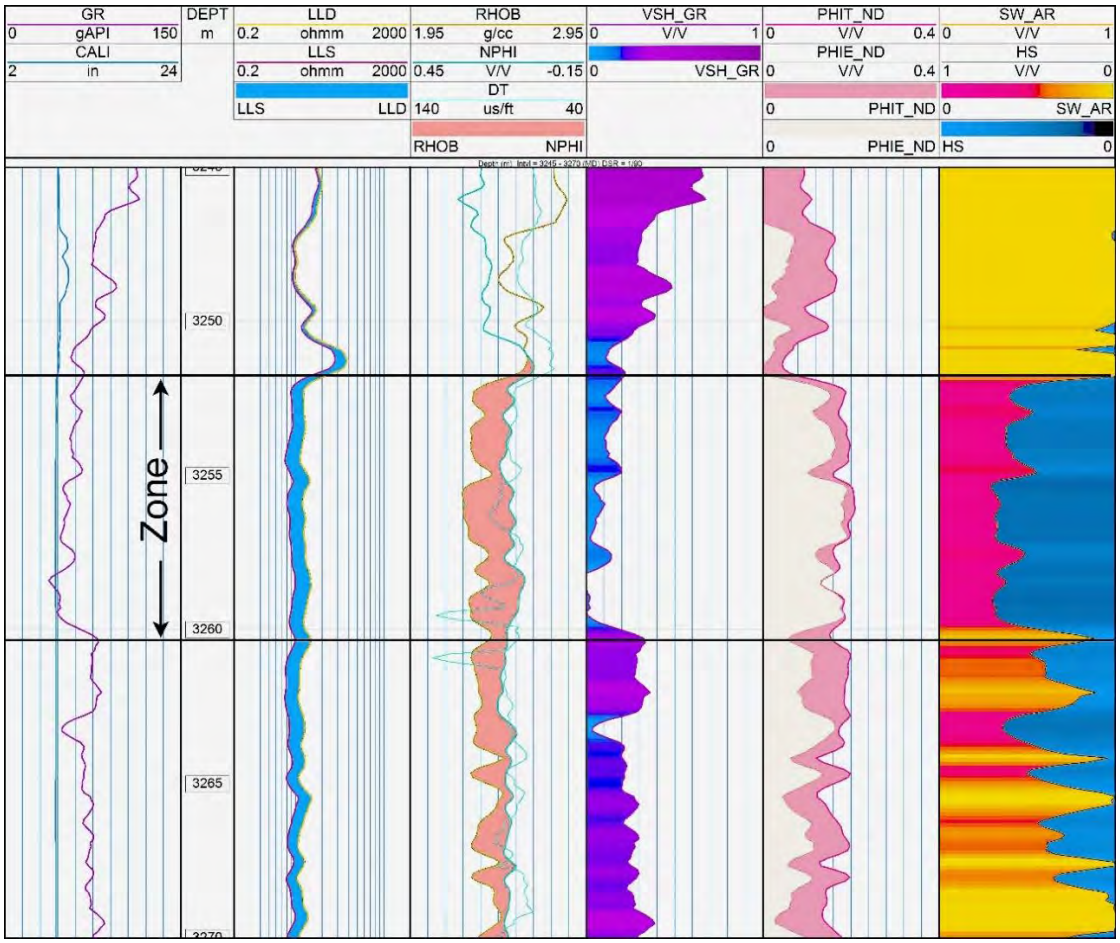

Figure 4. 2: Petrophysical analysis of well Sawan-01 with marked zone of interest

The depth ranges of this particular zone, referred to as "Zone," ranges from 3252m to 3261m are taken (Figure 4.2). This zone holds significant promise for hydrocarbon exploration due to its favorable petrophysical characteristics, which include high porosity and low water saturation levels.

Table 4. 1: Results of petrophysical analysis obtained from Zone of interest (Sawan-01)

Serial Number	Calculation Parameter	Percentage
1	Average Volume of Shale = **Vsh**	20%
2	Average Porosity in(PHIT) Percentage =ϕ_{avg}	17.3 %
3	Average Effective Porosity in Percentage = ϕ_{eavg}	11.4%
4	Average water Saturation in Percentage = S_{Wavg}	34%
5	Average Hydrocarbon in Percentage = $\mathbf{S_{Havg}}$	66%

4.14 Well Log Interpretation of Sawan-07

Based on the log response analysis, the thickest zone within the Lower Goru formation of the Sawan-07 well has been identified. This zone extends from a depth of 3269 meters to 3306 meters, encompassing a substantial thickness of 37 meters. The log responses within this interval provide crucial information regarding the characteristics of the formation.

The log responses reveal specific patterns that contribute to the identification of this thickest zone (Figure 4.3). Notably, the GR log response is observed to be low, indicating sandstone. Additionally, the Caliper log demonstrates a stable trend, without any significant variations that might indicate wash-outs. The separation between the LLS and LLD measurements, as well as the crossover between NPHI and RHOB is an indication of hydrocarbon zone (Miller, 1986). In the (Fig 4.3) it can be observed that there is crossover between NPHI and RHOB and separation between the LLS and LLD measurements, which further contributes to the interpretation that the zone contains a notable presence of hydrocarbons.

The petrophysical analysis of this marked zone of interest reveals favorable attributes. The total and effective porosities within this zone are notably favorable, contributing to the potential for

hydrocarbon accumulation. Moreover, the water saturation levels exhibit a decrease within this marked zone of the Lower Goru formation, further enhancing the prospect of hydrocarbon presence.

The combination of log response patterns, cross-over points, and petrophysical characteristics collectively supports the indication of a substantial hydrocarbon presence within this thickest zone of the Lower Goru formation in the Sawan-07 well (Figure 4.3)

4.14.1 Reservoir Zone using Well Log Interpretation of Sawan-07

The petrophysical analysis results obtained for this zone of interest have been compiled and presented in (Tables 4.2). Reservoir zone is marked which is about 37 m thick having clean sand. Water saturation is 48% and reservoir zone is shown in (Figure 4.3)

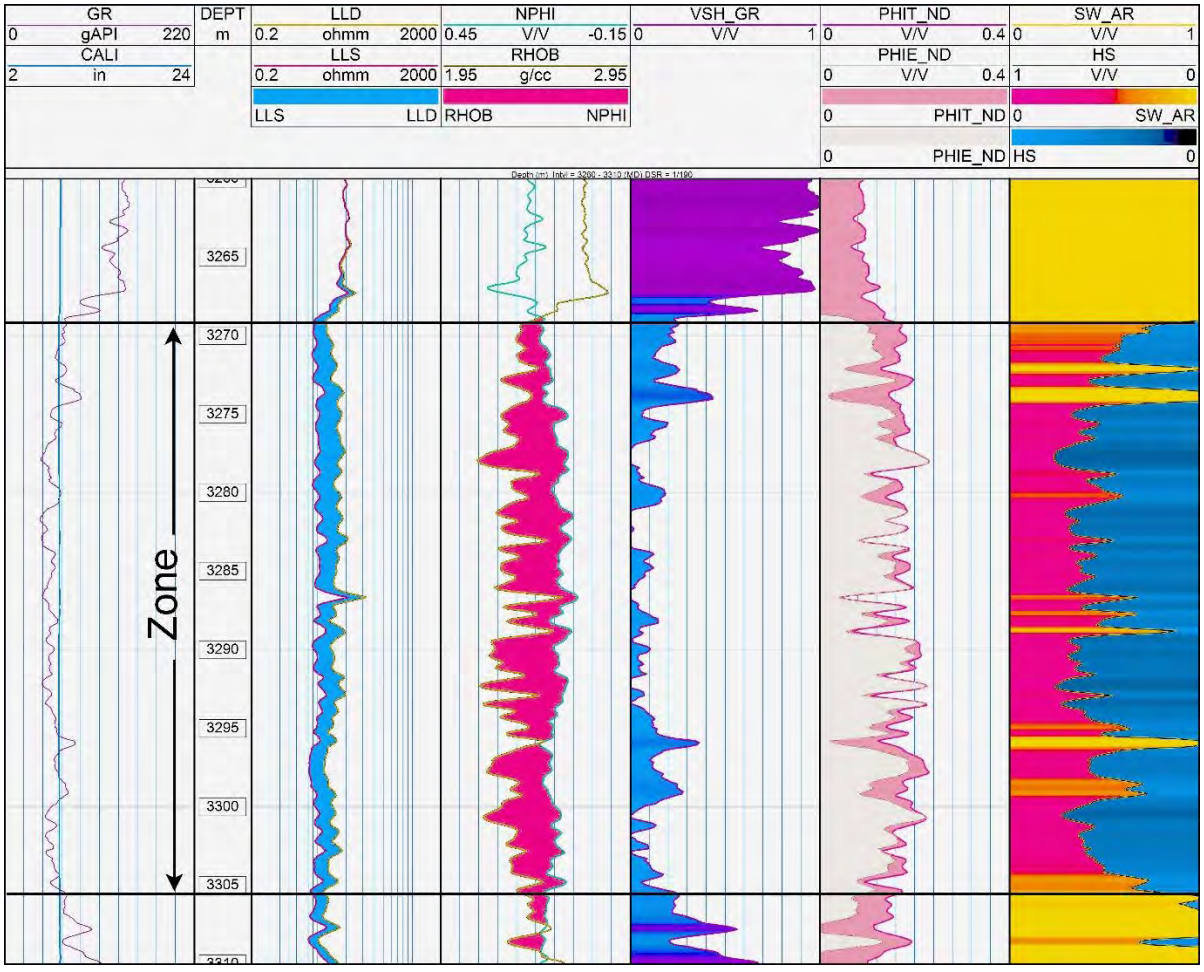

Figure 4. 3: Petrophysical analysis of well Sawan-07 with marked zone of interest

Table 4. 2: Results of petrophysical analysis obtained from Zone of interest (Sawan-07)

Serial Number	Calculation Parameter	Percentage
1	Average Volume of Shale = **Vsh**	14.5%
2	Average Porosity in (PHIT) Percentage =Φ_{avg}	16.6%
3	Average Effective Porosity in Percentage = Φ_{eavg}	12.5%
4	Average water Saturation in Percentage = S_{Wavg}	48%
5	Average Hydrocarbon in Percentage = $\mathbf{S_{Havg}}$	52%

CHAPTER 5

ESTIMATION OF SHEAR WAVE(DT4S) USING MACHINE LEARNING TECHNIQUE

5.1 Introduction

For geoscientists, the fundamental complexity of petroleum reservoir systems remains a challenge. Insufficient comprehension of reservoir behavior and subpar estimations of reservoir potential arise from the absence of high-quality data. (Al-Anazi & Gates, 2010). Numerous geophysical and petrophysical studies use shear wave velocity, particularly when calculating geo-mechanical parameters (Khatibi and Aghajanpour, 2020). Shear waves, in contrast to compressional waves, cannot travel through liquids (Han et al., 2005) knowing the relationship between lithology, pore distribution, and shear wave velocity provides useful information about the reservoir (Anemangely et al., 2019). There are numerous applications of shear wave velocity including reservoir characterization, petrophysical analysis, geomechanical studies, rock engineering applications (Castagna et al., 1985). P-wave velocity and S-wave velocity used in combination can improve reservoir and fluid prediction accuracy while lowering seismic based uncertainty (Zhang et al., 2021). However, shear wave logs are only available in a limited number of wells in an oil field due to the high cost of the log acquisition, thus acquiring shear wave data has become difficult and challenging (Akhundi et al., 2014).

Numerous machine learning algorithms have proven effective at estimating shear wave velocities. Simple models that capture linear relationships are provided by regression techniques like linear, polynomial, and ridge regression. Artificial neural networks (ANNs), which are more sophisticated algorithms, use hidden layers to uncover complex nonlinear patterns in the data. Data is categorized into various velocity categories using support vector machines (SVMs), which are renowned for their robustness. (Xu & Payne, 2009; Russell et al., 2003; Wang et al., 2020; Ghorbani et al., 2012).

Gradient Boosting Regressor (GBR) is a new technique in machine learning techniques. Due to its capability to manage complex relationships within data. GBR has become well-known as an advanced regression algorithm (Maddu et al., 2022). The ability of GBR to quantify the relative importance of input features is a key component of its usage. GBR assigns weights to the features

that have the greatest impact on reducing prediction errors through iterative optimization of loss functions (Khan et al., 2022). For the recent years GBR has been used by experts in various fields for predictions like Energy consumptions, hydrogeological studies, forecasting and in chemical engineering etc (Cai et al., 2020; Otchere et al., 2020; Singh et al., 2021; Nie et al., 2021).

In this dissertation, GBR has been used to predict shear wave for Sawan area. Prior to GBR application, it's very important to analyze the data for it to give best results.

These are following steps in prediction of shear wave:

- Input well log data
- Generation of heat maps
- Generation of Histplot and Boxplot
- Splitting data for training and testing
- Applying algorithm of interest
- Feature importance
- Model validation and testing on a blind well

5.2 Data Set

Only Sawan 7 and Sawan 8 have been used for training of Shear wave prediction among the numerous wells in the study area because both have the full set of necessary logs, including GR, NPHI, RHOB, LLD, LLS, DT4P, and particularly DT4S log as shown (Figure 5.1 and 5.2). 23 logs have been run in both wells in total.

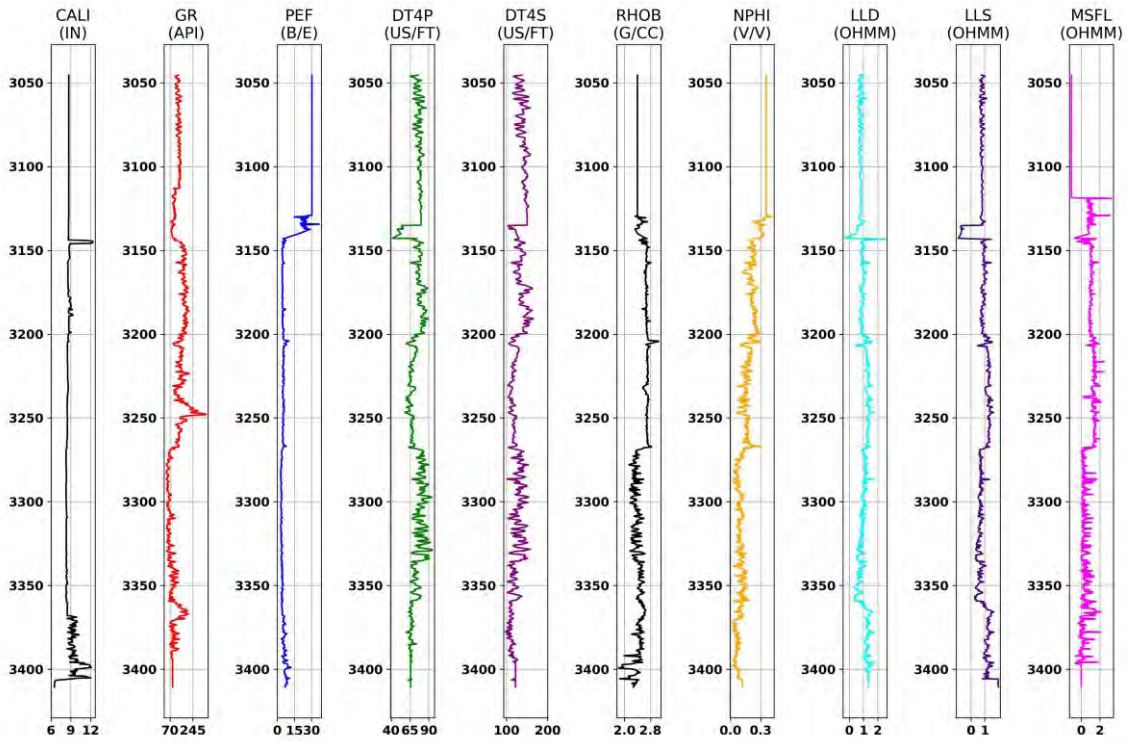

Figure 5. 1: Well logs of Well Sawan-07 initially selected for the prediction of DT4S

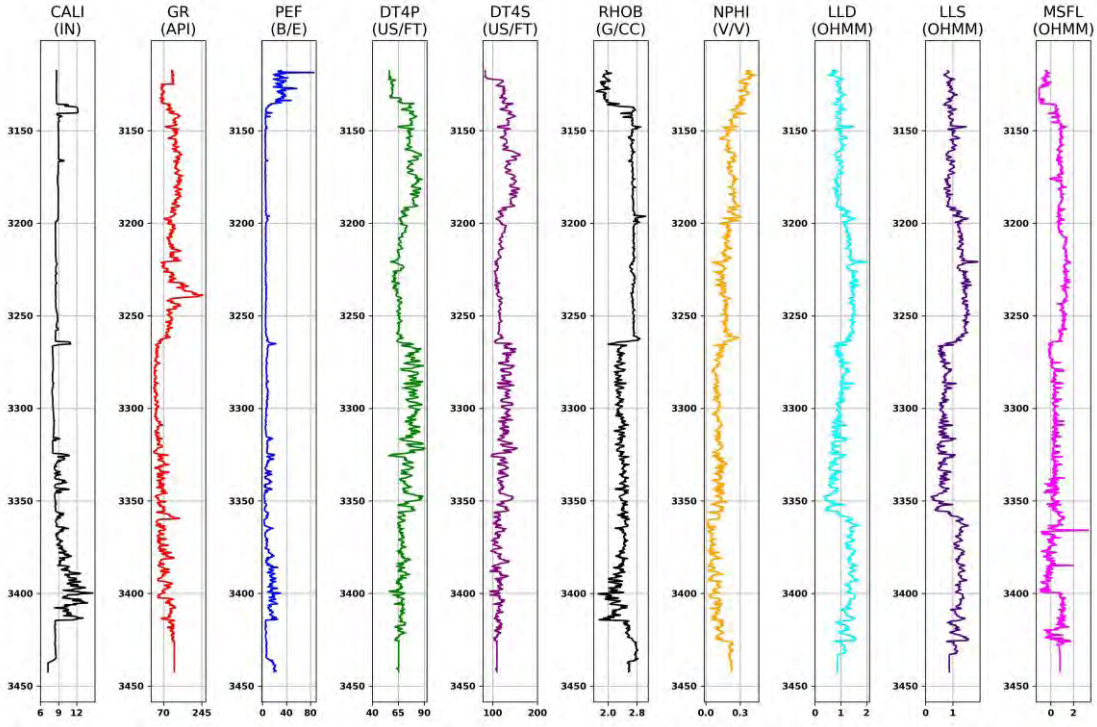

Figure 5. 2: Well logs of Well Sawan-08 initially selected for the prediction of DT4S

5.3 Heat Map Generation

Understanding the connections between various variables is crucial, particularly when predicting shear wave values from well log data. By determining which log have a high correlation with other and vice versa. Correlation analysis enables you to choose features, build models, and interpret data with confidence (Venieri et al., 2021). In this study, heat maps have been generated of Sawan 7 and Sawan 8 to check correlation between wells (Figure 5.3).

These heatmaps show that correlation of MSFL and CALI have bad correlation with all other logs so for data quality, these logs will not be taken for training purposes.

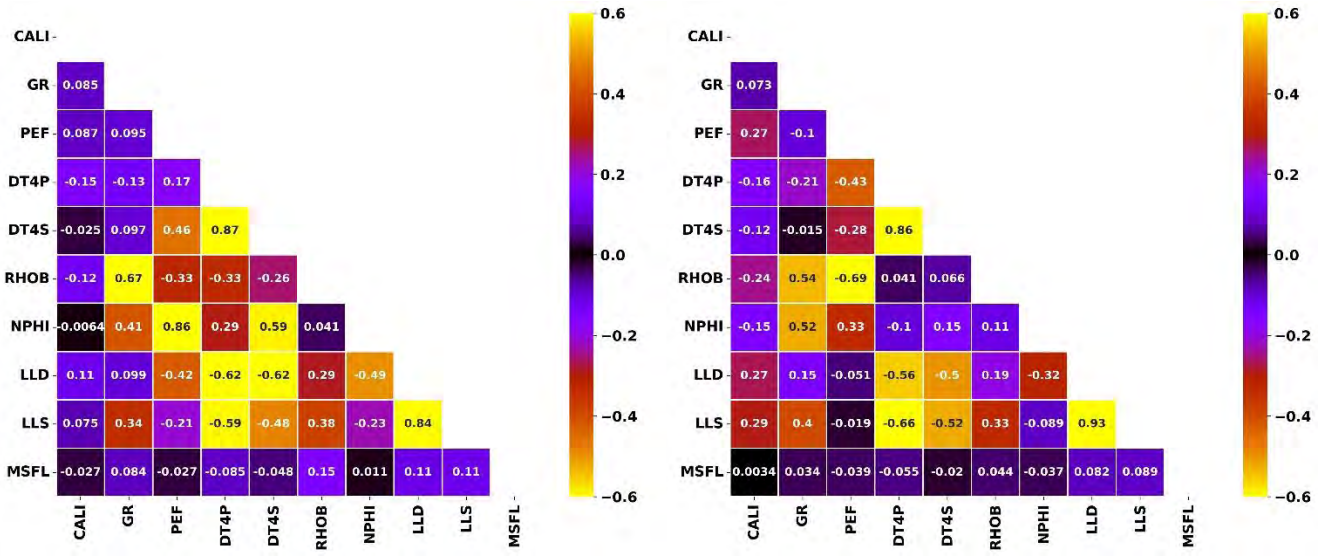

Figure 5. 3: Heat map of Sawan-07 (left) and Sawan-08 (right) indicating correlation values.

After removing bad correlation logs, concatenation of the data has been done from all wells, this means combining respective columns of each well to have one single data set.

5.4 Generation of Histplot and Box plot

Box plots and histograms are crucial tools for data analysis and visualization. They are essential for comprehending the distribution and properties of data, spotting outliers, and selecting models, feature engineering, and data preprocessing in an informed manner (Data et al., 2016).

To detect bad data points i.e. outliers, histplots and boxplots have been generated. (Figure 5.4) shows that there are some outliers in the data. Outliers are bad data points which can be in the data due to noise or some errors in acquisition (Zhang et al., 2010).

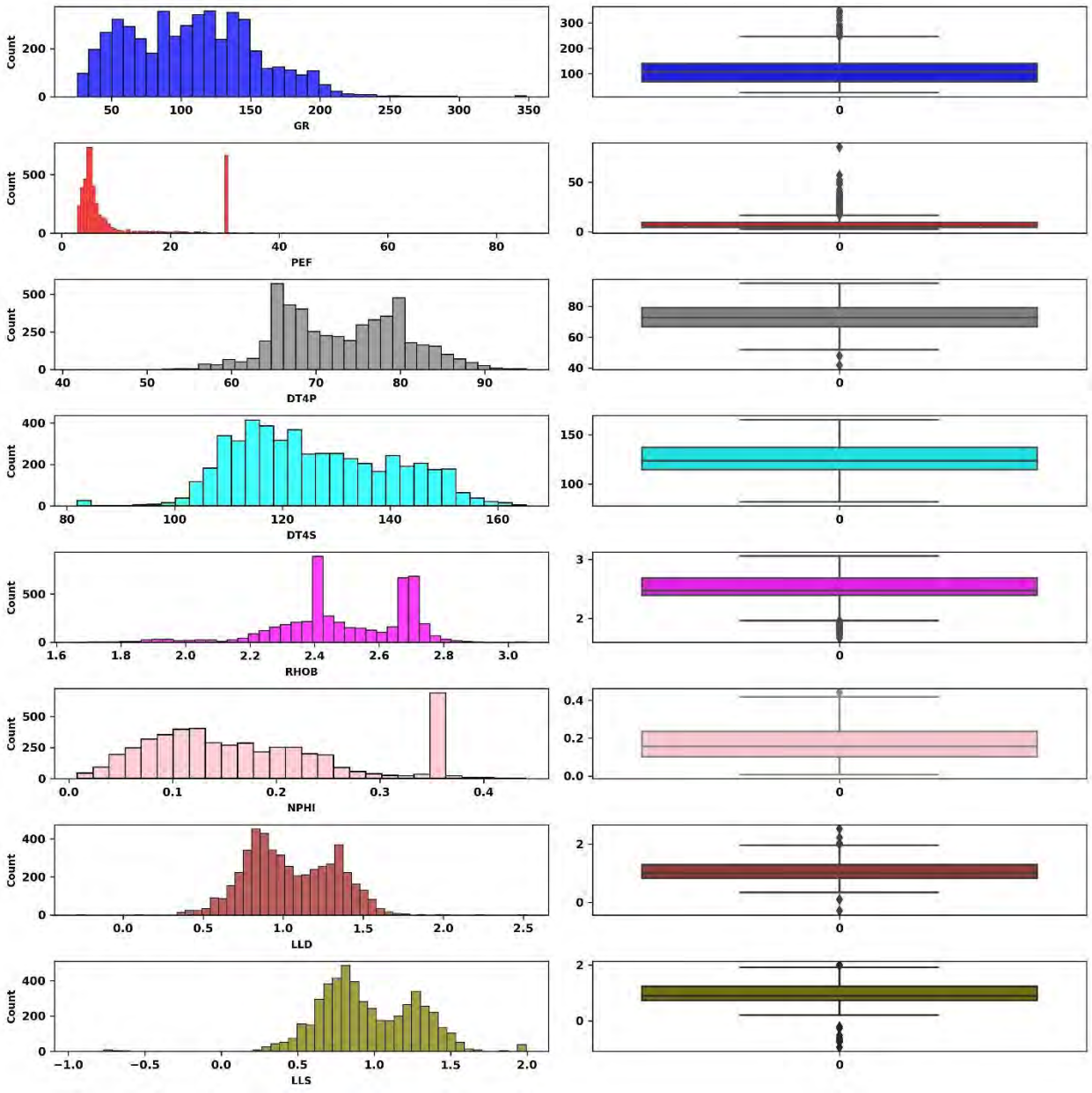

Figure 5. 4: Histogram and box plots showing distribution of data depicting main data and outliers

5.4.2 Removal of Outliers

To increase data quality, maximum possible outliers are to be removed. Larger input values may have a significant impact on a model's outputs, as a result, the coefficients of the parameters may not be correctly determined, which could reduce a model's accuracy and generalizability. Data normalization is therefore required in order to carefully choose the best features. The data is normalized in between 0 and 1 by using Yeo-Johnson transformation (Yeo and Johnson, 2000).

One-Class Support Vector Machine (OC-SVM) has been implemented to remove bad data points. OC-SVM was first given by (Schölkopf et al., 2001). OC-SVM algorithm is used for both classification and regression tasks. It finds an optimal hyperplane to separate data points of different classes in a high-dimensional space.

The mathematical representation of OC-SVM is given by (Schölkopf et al., 2001) as follows:

$$w\hat{}T * x - b = \rho \quad (5.1)$$

where

w = the weight vector of the hyperplane.

x = the input data point.

b = a bias term.

ρ = a threshold.

the optimization problem for OC-SVM given by (Schölkopf et al., 2001) is formulated as:

$$\frac{1}{2}||w||^2 + C\sum_{i=1}^{N}\xi(x_i) - \rho \quad (5.2)$$

Where:

w = weight vector of the hyperplane

C = parameter that controls the trade-off between maximizing the margin and minimizing the outlier errors.

x_i = data points

N = the number of data points

$\xi(x_i)$ = slack variables that allow some points to be on the wrong side of the hyperplane.

ρ = constant representing the threshold for considering a point as an outlier.

Removing of outliers is done by using ‘nu’ = 0.06 and total of 304 outliers are removed. ‘nu’ is to specify the percentage of outliers.

After removing bad data points, De-normalization of data has been done. So every log is back to its original scales. Once again plotting of box and histplots has been done to check if outliers have been removed (Figure 5.5).

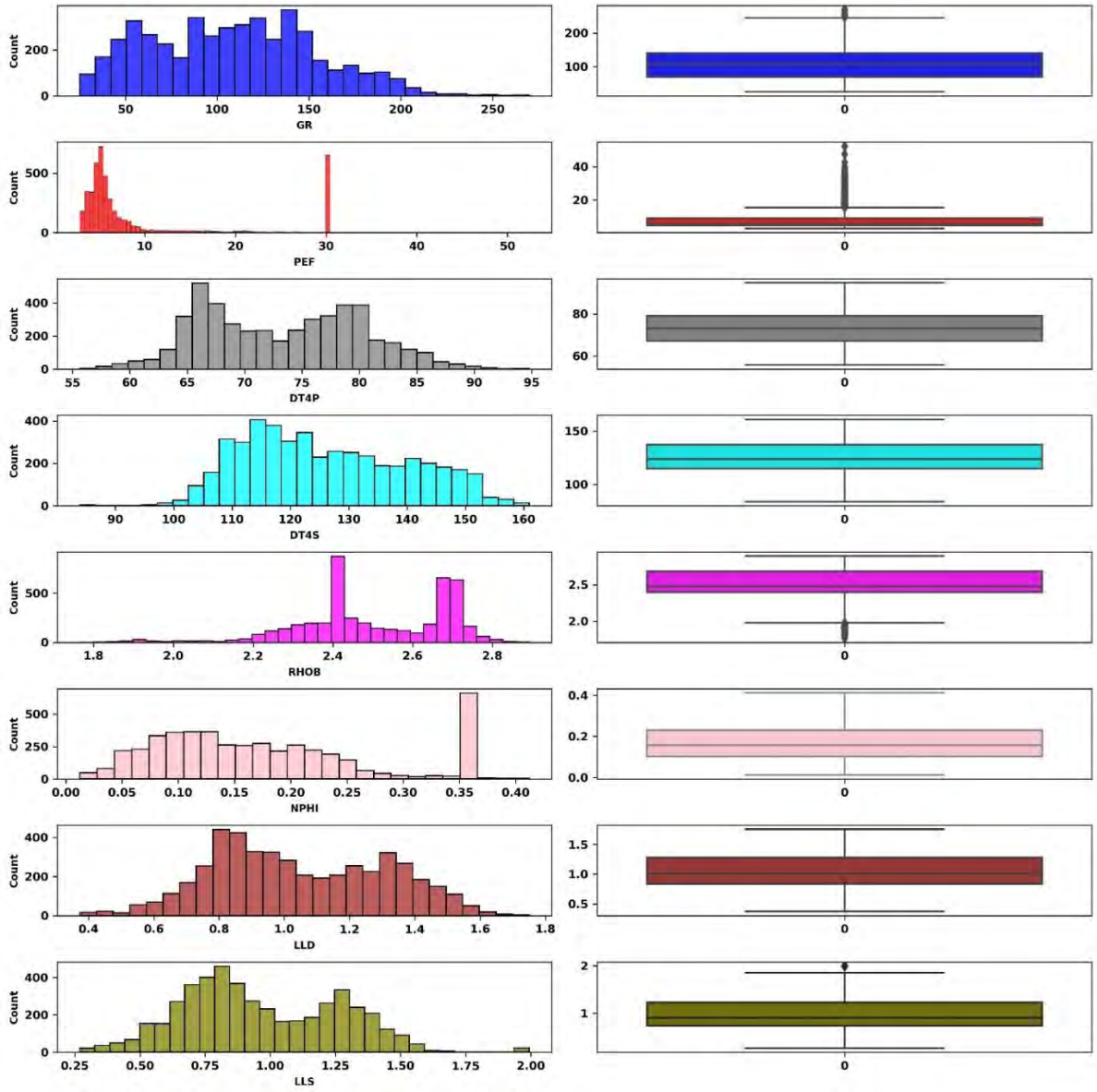

Figure 5. 5: Histogram and box plots showing distribution of data depicting main data points after removal of outliers

It is shown from the figure that maximum possible outliers have been removed and data quality is improved.

5.4.3 Plotting Heat Map of Improved Data

After improving data quality by several methods mentioned above, again heat map has been plotted to check improvement in correlation. As this data is used for training our algorithm for shear wave prediction.

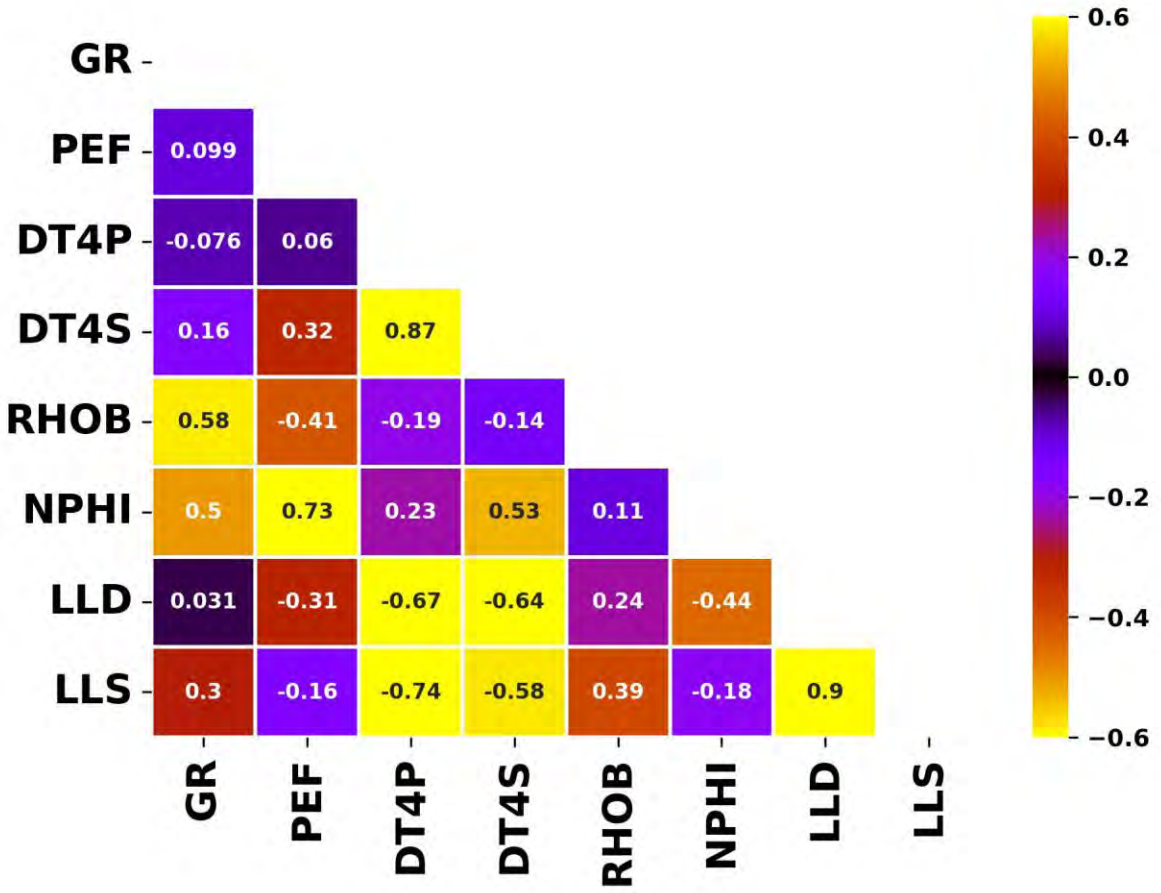

Figure 5. 6: Heat map of the improved data indicating a relatively higher correlation values

5.5 Data Splitting and Algorithm Application

For training and testing purposes, concatenated data has been split by train-test split function by sk-learn library in python. First separate the data into labels (y) and features (X). The data frame is split into the X_train, X_test, Y_train, and Y_test sections. The model is trained and fitted using the X_train and y_train sets. The model is tested to see if it correctly predicts the outputs and labels using the X_test and y_test sets (Buitinck et al., 2013; Pedregosa et al., 2011). Around 80% of the total data was used in training for a best prediction of shear wave, and the remaining 20% was used to test the newly predicted shear wave.

After splitting data for training and testing, GBR is applied on data and shear wave has been predicted. GBR is an iterative collection of sequentially arranged tree models where each successive model learns from the mistakes of the previous model. In order to create a more reliable model, this machine learning model makes predictions by "boosting" the ensemble of weak prediction models, frequently decision trees (Rao et al., 2019). The GBR model used in this study is from the Sk-learn library GBR method (Pedregosa et al., 2011). Following equation is given by (Rao et al., 2019) for GBR.

$$f_M(x_j) = \sum_m^M \Upsilon_m \, h_m(x_j) \tag{5.3}$$

where

h_m = a weak learner that performs poorly individually

Υ_m = is a scaling factor adding the contribution of a tree to the model

5.6 Feature Importance

Understanding which features or attributes have the biggest effects on predicting a target variable is made possible by the fundamental machine learning concept of feature importance (Rajbahadur et al., 2021).

The mathematical representation of feature importance is given by (Ronaghan, 2018) as follow:

$$n_{ij} = w_j c_j - w_{left(j)} c_{left(j)} - w_{right(j)} c_{right(j)} \tag{5.4}$$

n_{ij}= node importance

w_j= weights of nodes

c_j= impurity

Impurity is a measurement of the disorder or impurity of a collection of samples inside a decision tree node. To ascertain how well a specific attribute divides the data into homogeneous subsets, it is used in tree-based algorithms like Decision Trees, Random Forests, and GBR. It can be calculated as entropy. Equation is given by (Rajbahadur et al., 2021).

$$H = -\sum_{K=1}^{K} p_K \log(P_K) \tag{5.5}$$

Where

H = entropy

K = number of classes

P_K = proportion of samples that belong to class in node

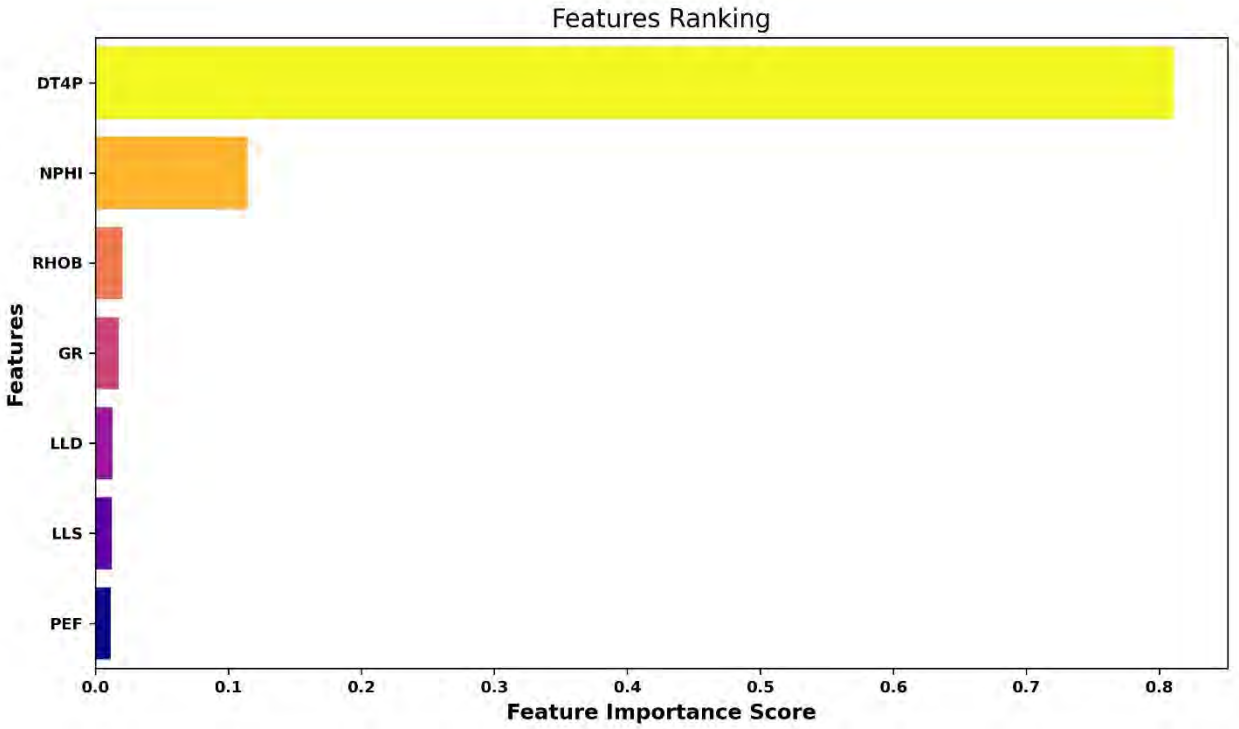

Figure 5. 7: Feature importance showing which log has greatest effect in predicting DT4S

(Figure 5.7) shows that DT4P has greatest effect in predicting DT4S. The reason why it has most feature importance score is because of it has best correlation with all input logs as it is shown in (Figure 5.6). On contrast PEF has least score due to its bad correlation.

5.7 Model Validation and Testing of Model on a Blind Well

Model testing is done by using metrics like R-squared (R^2) and Root Mean Squared Error (RMSE). Especially in regression tasks, the RMSE is an important metric for assessing the precision of a predictive model. It gauges how closely the model's predictions match the data observed by calculating the average magnitude of the errors between predicted and actual values (Chai et al., 2014).

Equation for RMSE is given by (Ćalasan et al., 2020) as follows:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(yi - y\hat{}i)^2} \quad (5.6)$$

Where

n = number of data points

yi = the actual (observed) value of the target variable for the i-th data point.

$y\hat{}i$ = the predicted value of the target variable for the i-th data point.

$\sum$ = summation over all n data points.

The square of the differences between actual and predicted values is summed up, divided by n, and then the square root is taken to compute the RMSE.

R^2 is correlation coefficient, a statistical measure used to evaluate the goodness of fit of a regression model. It displays the percentage of the variance in the target's dependent variable that can be accounted for by the model's independent variables (Weisstein, 2006).

Equation for R2 is given by (Weisstein, 2006) as follows:

$$R^2 = 1 - \frac{SS_{res}}{SS_{total}} \quad (5.7)$$

Where

SS_{res} = sum of squared residuals (also known as the sum of squared errors), which represents the variability that is not explained by the model.

SS_{total} = is the total sum of squares, which represents the total variability in the dependent variable. This gives R^2 value between 0 and 1. The higher the value the better will be the model and vice versa.

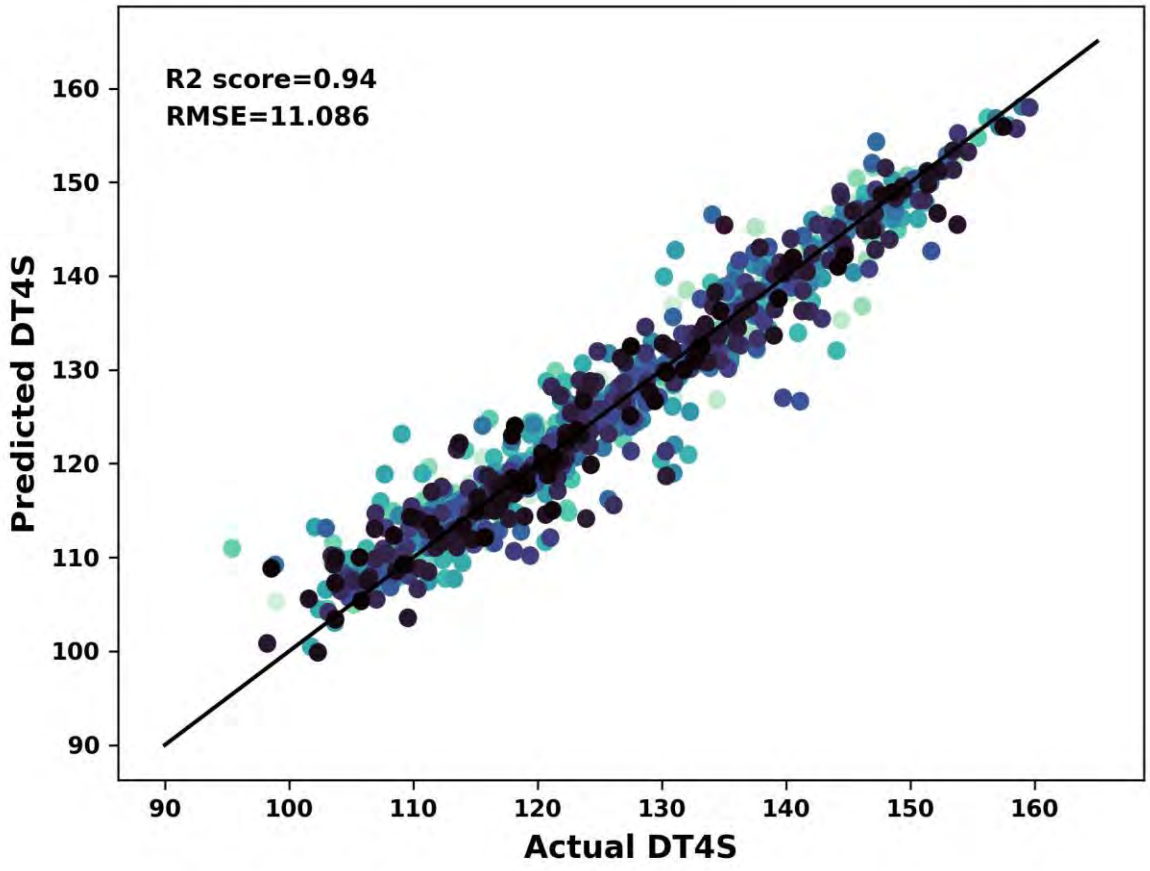

Figure 5. 8: Correlation plot between predicted DT4S and Actual DT4S showing 94% correlation through RMSE and R^2.

20% data which was for testing is being used in this correlation. R2 correlates that data with actual values and the result shows $R^2 = 0.94$, which means the correlation achieved is 94%. This validates the accuracy of the prediction.

For a model testing a blind well data which wasn't used in either training or testing, Sawan-01 having logs GR, PEF, DT4P, RHOB, NPHI, LLD and LLS is being used.

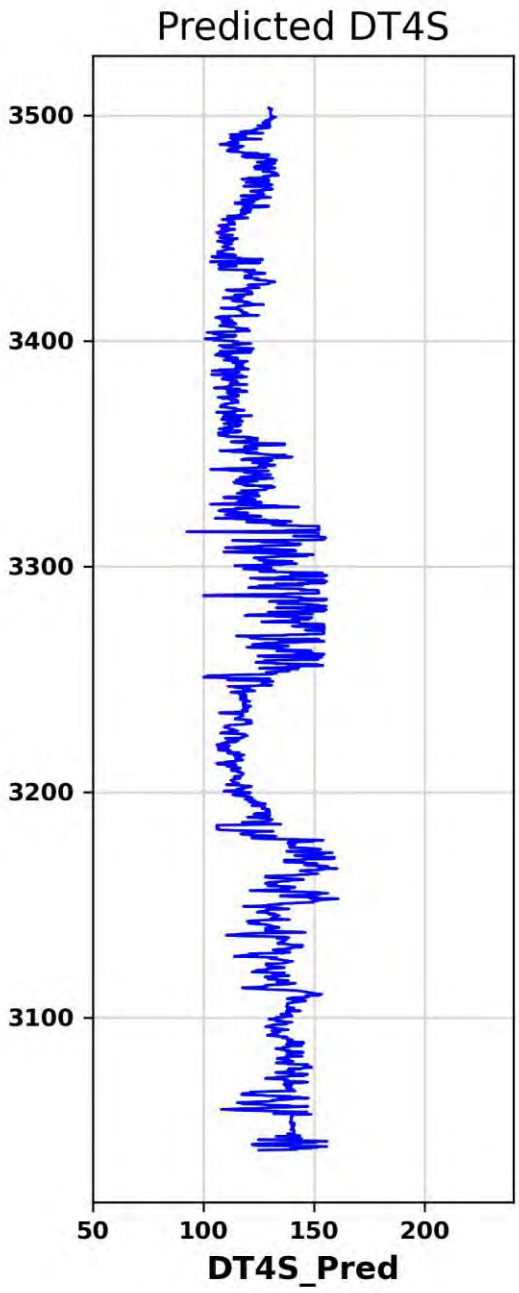

Figure 5. 9: Predicted DT4S log in Sawan-01

(Figure 5.9) shows DT4S log curve which is obtained by using model on a blind well (Sawan-01). To cross check its validation, all well logs are shown below (Figure 5.10). DT4P log and predicted DT4S log is showing almost same which further validates its accuracy and correlation.

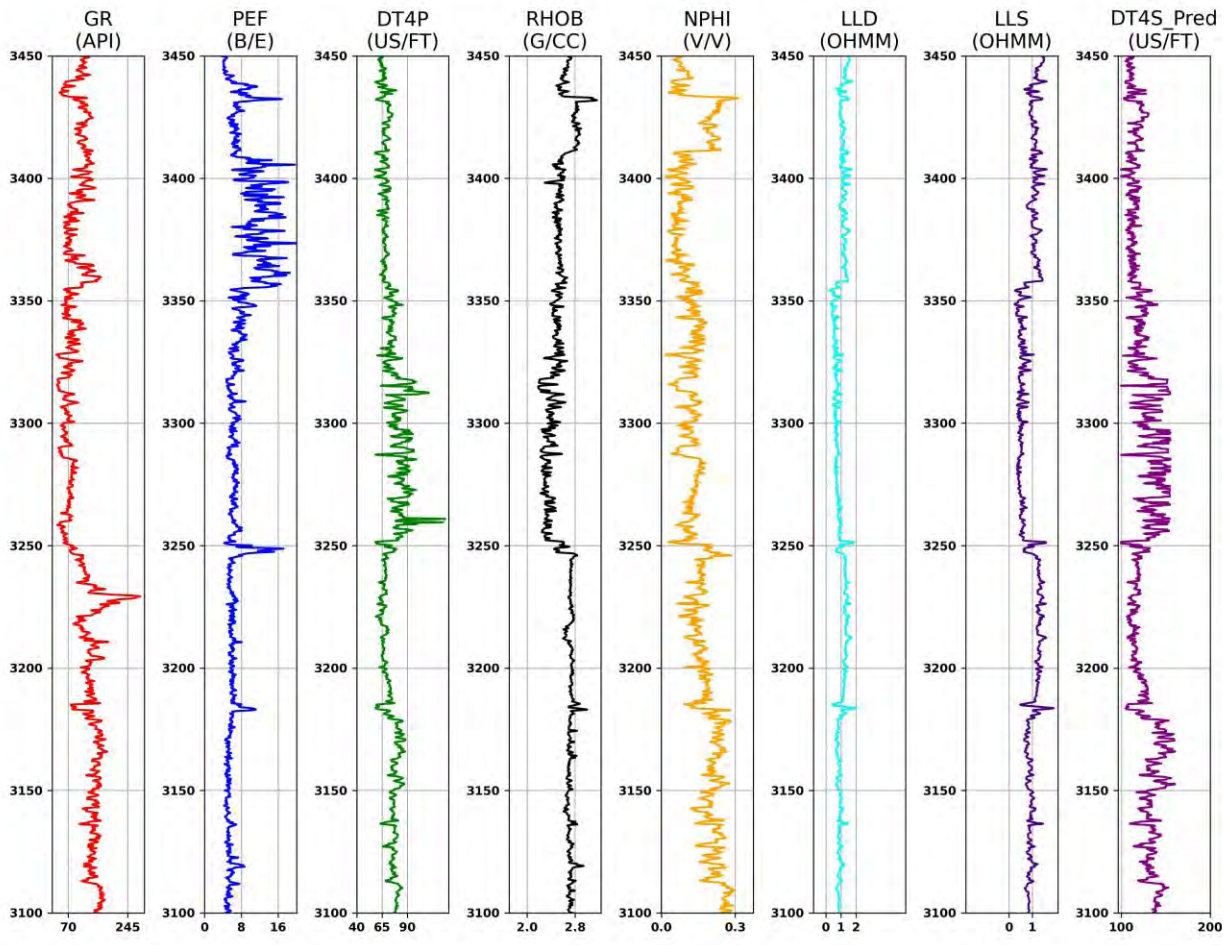

Figure 5. 10: Well logs of Sawan-01 with predicted DT4S indicating its trend along with other logs used in this prediction.

CHAPTER 6

RESERVOIR CHARACTERIZATION AND CROSS PLOT ANYLYSIS USING PREDICTED DT4S

6.1 Introduction

The process of comprehending and describing the heterogeneity of subsurface reservoirs is known as reservoir characterization. Heterogeneity describes the variation in a reservoir's rock properties, including porosity, permeability, lithology, and fluid saturation. The efficient and effective development of hydrocarbon resources, whether in the oil or natural gas industry, depends on accurately describing this heterogeneity (Dominguez, 1992). It is a diverse, integrated task that demands expertise in reservoir geology, well logging, geophysics, geostatistics, petrophysics, and reservoir engineering to produce precise 3D images of the petrophysical parameters necessary to predict reservoir performance (Grana et al., 2013). In particular, identifying hydrocarbon-bearing and non-hydrocarbon-bearing zones within a reservoir requires the use of petrophysical analysis, which is essential for reservoir characterization. To convert these data into key reservoir characteristics, petrophysicists use wireline log data and a variety of methodologies (Lucia et al., 2003). In reservoir characterization, petrophysical analysis is particularly useful for distinguishing hydrocarbon and non-hydrocarbon bearing zones (Yuedong & Hongwei , 2007). In most cases, wireline log data are converted into reservoir parameters like shale volume, porosity, permeability, and water and hydrocarbon saturation using petrophysical analysis (Islam et al., 2021).

Shear (Vs) and compressional (Vp) velocities have great significance in reservoir characterization. With help of machine learning these velocities can be predicted and used. Reservoir characterization relies significantly on shear and compressional velocities. These velocities could be predicted using a machine learning. The reservoir system is an illustration of a physical system that displays both deterministic and random behavior (Fallah-Mehdipour et al 2012).

When combined and analyzed, velocity ratios (Vp/Vs) and density (ρ) are significant seismic attributes that can help estimate many important parameters essential for reservoir analysis and characterization (Wu et al., 2022). Some of these parameters are:

6.1.1 Porosity (φ)

Porosity, which represents the void space within a rock and directly affects fluid storage, flow, and reservoir performance, is a crucial parameter in reservoir analysis. Estimates of porosity are crucial for comprehending reservoir capacity, fluid storage, and permeability (Gunter et al., 1997). Different petrophysical models can be used to calculate porosity by relating the compressional wave velocity (Vp), shear wave velocity (Vs), density (ρ), and porosity (φ) (Hamada, 2004). The Wyllie time-average equation is frequently employed models for determining porosity from seismic data (Helgerud et al., 1999).

Wyllie time-average equation is given by (Wyllie et al., 1956):

$$\frac{1}{v} = \frac{\emptyset}{v_f} + \frac{1-\emptyset}{v_m} \quad (6.1)$$

Where:

V =Velocity

V_f =Velocity of the fluid

Φ = Porosity

V_m = Velocity of the matrix

6.1.2 Lambda-Rho (λρ)

Reservoir characterization both use the seismic attribute lambda-rho (λρ). It is the outcome of the first Lamé parameter (λ) and the density of subsurface rocks (ρ) (Kaczmarczyk-Kuszpit, 2021). In order to calculate λρ, both λ and ρ must be known, which can be calculated using density, compressional (P-wave) velocity (Vp), and shear (S-wave) velocity (Vs) (Obilo et al., 2007).

Equation for lambda-rho (incompressibility factor) is given by (Goodway, 2001) as follows:

$$\lambda\rho = \lambda * \rho = (Vp^2 - 2 * Vs^2) * \rho \quad (6.2)$$

Where:

Vp = compressional velocity

Vs = Shear velocity

ρ = density

λ = Lamé parameter

6.1.3 Mu-Rho

The shear modulus (μ) and density (ρ) product, commonly abbreviated as "μρ" or "MuRho," is a crucial parameter in reservoir characterization that aids in comprehending various aspects of subsurface rock properties. It has implications for reservoir engineering, rock mechanics, and seismic analysis (Guedez, 2019).

Equation for Mu- Rho (μρ) is given by (Goodway, 1997) as follows:

$$\mu\rho = \mu * \rho \tag{6.3}$$

6.1.4 Poisson Ratio

The Poisson's ratio (ν) reveals details about a rock's capacity to change shape under pressure. It is possible to evaluate the mechanical behavior of the rocks inside the reservoir during reservoir characterization, which is essential for comprehending rock deformation, wellbore stability, and hydraulic fracturing procedures. It has several implications in reservoir characterization like rock elasticity, seismic interpretation, fluid saturation, geomechanical modelling, reservoir compaction etc. (Maxwell et al., 2011).

Equation for Poisson ratio is as follows (Sheriff, 2002):

$$\sigma = \frac{Vp^2 - 2Vs^2}{2(Vp^2 - 2Vs^2)} \tag{6.4}$$

Where:

σ = the dimensionless Poisson's ratio.

Vp = compressional wave velocity

Vs = Shear wave velocity

6.1.5 VpVs Ratio

The compressional wave velocity (Vp) to the shear wave velocity (Vs) in subsurface rocks is represented by the seismic attribute known as the VpVs ratio, also referred to as the P-wave to S-wave velocity ratio. Due to the fact that it offers important details regarding the elastic

characteristics of the rocks and their response to seismic waves, this ratio is significant in geophysics and reservoir characterization (Hamada, 2004).

VpVs ratio has implications in lithological determination, fluid identification, fracture detection, reservoir compaction and pressure prediction, hydrocarbon exploration etc. (Hamada, 2004; Wang et al., 2012; Julià & Mejía, 2004; Zimmer et al., 2002).

Equation is given as follows:

$$VpVs\ ratio\ = \frac{Compressional\ Wave\ Velocity}{Shear\ Wave\ Velocity} \tag{6.5}$$

6.1.6 Impedance Log

A vital tool in geophysics, petrophysics, and reservoir characterization is the impedance log, which provides details about acoustic impedance (Z) along a wellbore. Its significance comes from its capacity to offer insightful data on fluid content and subsurface rock properties, both of which are essential for assessing reservoirs for potential hydrocarbons. The density of the rock (ρ) and the seismic wave velocity (V) within it are what determine the acoustic impedance (Z) (Dubrule et al., 1998).

Equation for impedance log or acoustic impedance is given as follows:

$$Z = \rho * V \tag{6.6}$$

6.2 Cross Plot Analysis

An effective graphical analysis for locating data clusters in targeted zones is cross-plotting parameters against one another. These data sets can be categorized into various lithologies/facies based on their response. The ability to distinguish between various types of lithology and payable sand in the targeted area is made possible by elastic properties (VpVs ratio). To accurately predict lithology and fluid content, P-impedance and the VpVs ratio are combined (Azeem et al., 2017).

On the basis of predicted the Vs, Vp and Rhob in the subsurface different parameters which are fundamental to reservoir characterization and geophysical analysis can be estimated. The reservoir and its behavior can be understood by using these predicted values to estimate a variety of parameters and properties. These parameters provide important information that assists in finding and evaluating potential hydrocarbon reservoirs (Veeken & Rauch-Davies, 2006).

In this study, on the basis of cross plot analysis of these parameters which are calculated mathematically, hydrocarbon zone is successfully marked.

6.2.1 Lambda-Rho vs VpVs Ratio

The cross plot between Lambda-rho and VpVs ratio with Rhob on z-axis is shown in (Figure 6.1). From petrophysical analysis done previously, depth of zone was taken which ranges from 3252m to 3261m. Low values of lambda-rho indicate hydrocarbon zone which ranges from 10 – 34 GPa*g/cc and VpVs ratio lies within 1.55 to 1.84. This cross plot shows low values for Rhob in the range of 2.2500 – 2.5000 g/cc (Figure 6.1).

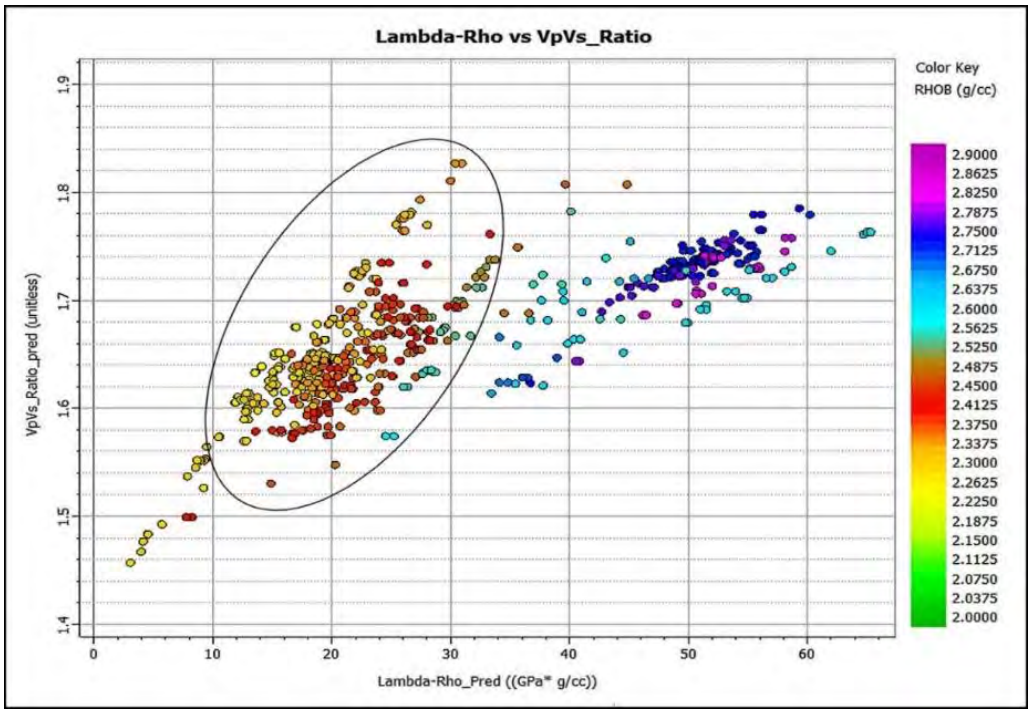

Figure 6. 1: Cross plot between Lambda-rho and VpVs ratio with Rhob on z-axis shows clear indication of hydrocarbon bearing zone.

The cross plot between Lamda-rho and VpVs ratio and color coded with Poisson is shown in (Figure 6.2). Low values of lambda-rho indicate hydrocarbon zone which ranges from 10 – 34 GPa*g/cc and VpVs ratio lies within 1.55 to 1.84. This cross plot shows high values of Poisson ratio that ranges from 0.26 - 0.32 (Figure 6.2).

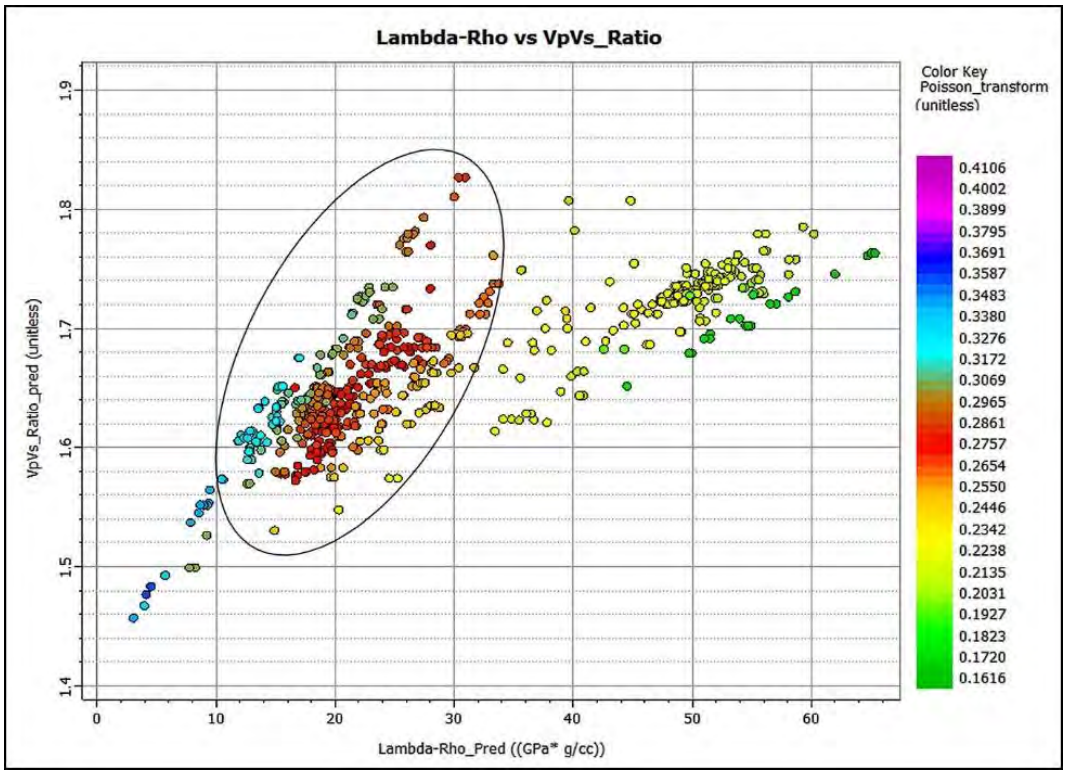

Figure 6. 2: Cross plot between Lambda-rho and VpVs ratio color coded with Poisson ratio shows clear indication of hydrocarbon bearing zone.

The cross plot between Lamda-rho and VpVs ratio with Phie on z-axis shown in (Figure 6.3). Low values of lambda-rho indicate hydrocarbon zone which ranges from 10 – 34 GPa*g/cc and VpVs ratio lies within 1.55 to 1.84. This cross plot shows high values of Phie that ranges from 0.1080 – 0.1660 (Figure 6.3).

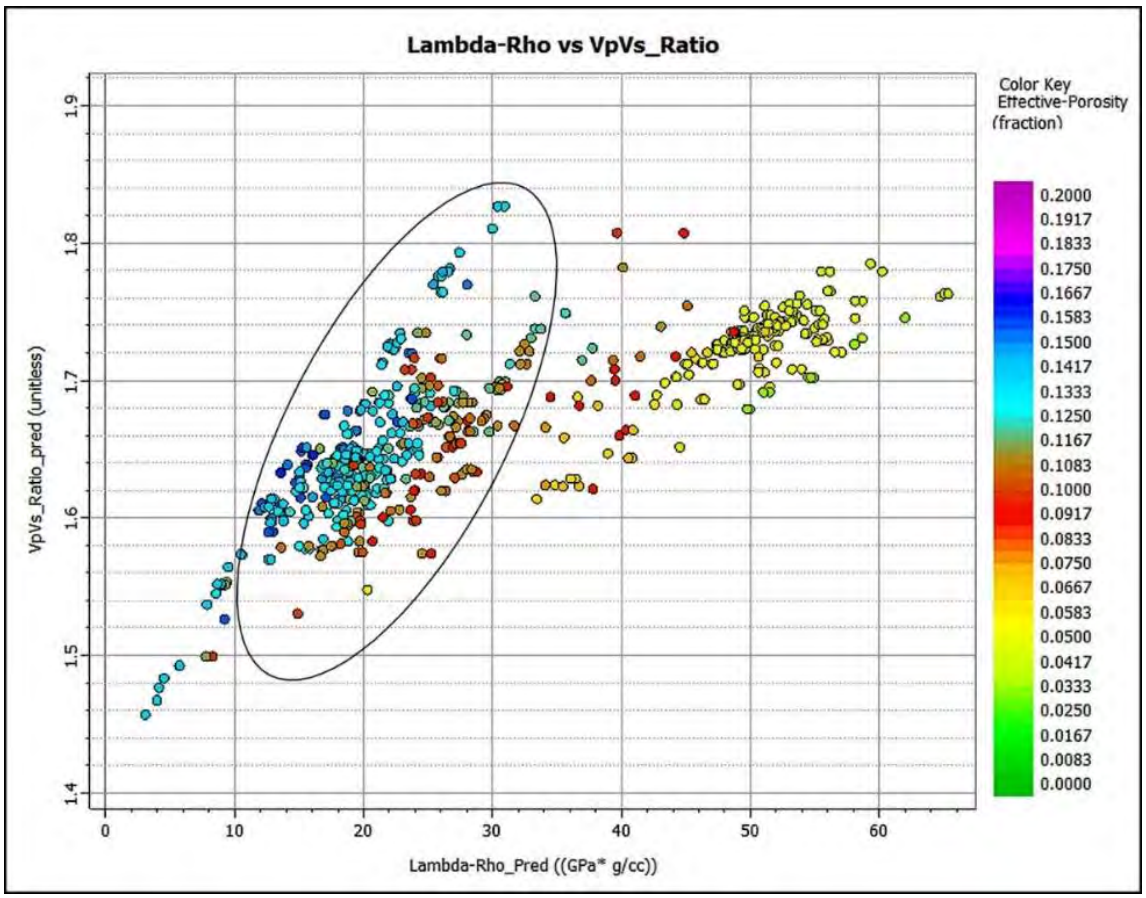

Figure 6. 3: Cross plot between Lambda-rho and VpVs ratio with Phie on z-axis shows clear indication of hydrocarbon bearing zone.

The cross plot between Lamda-rho and VpVs ratio with p-impedance on z-axis shown in (Figure 6.4). Low values of lambda-rho indicate hydrocarbon zone which ranges from 10 – 34 GPa*g/cc and VpVs ratio lies within 1.55 to 1.84. 9000 range value for P-impedance is perfect for hydrocarbons. This cross plot shows high values of P-impedance that ranges from 7700-9500 ((m/s) * (g/cc)) (Figure 6.4).

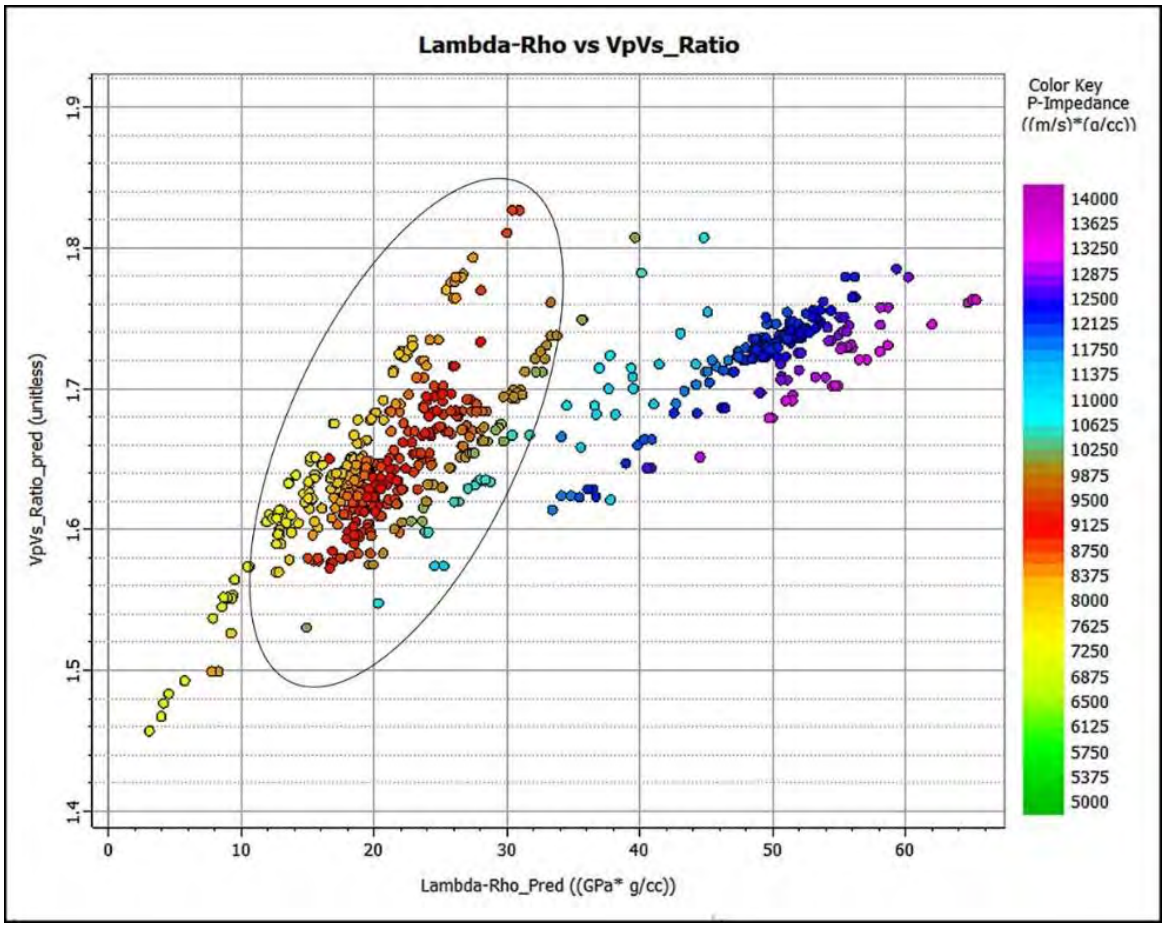

Figure 6. 4: Cross plot between Lambda-rho and VpVs ratio with P-Imepdance on z-axis shows clear indication of hydrocarbon bearing zone

6.2.2 Vp/Vs Ratio vs Impedance

The cross plot VpVs ratio and Impedance color coded with PHIE in (Figure 6.5). Values of Impedance indicate hydrocarbon zone which ranges from 7200-10000 m/s *g/cc and VpVs ratio lies within 1.55 to 1.84. This cross plot shows high values of PHIE that ranges from 0.1080 – 0.1660 (Figure 6.5)

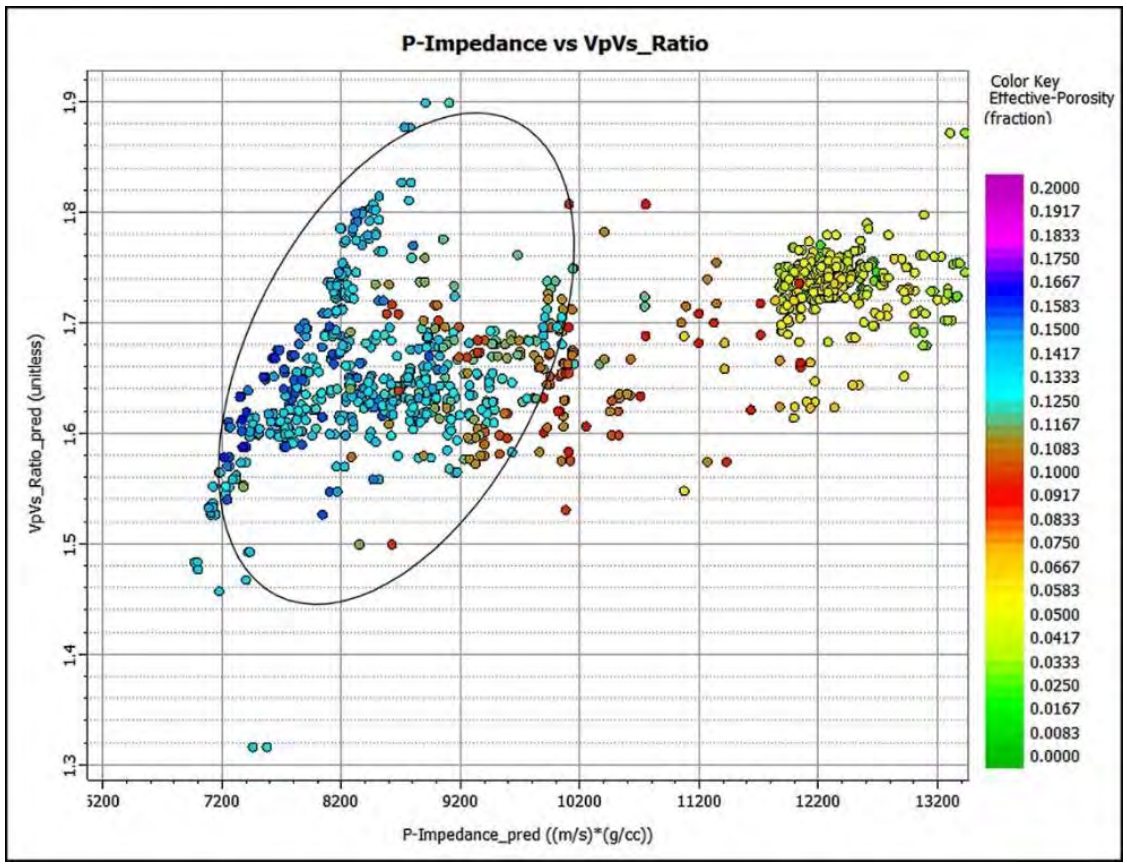

Figure 6. 5: The cross plot VpVs ratio and P-Impedance color coded with Phie shows clear indication of hydrocarbon bearing zone.

The cross plot VpVs ratio and Impedance color coded with Lambda-Rho shown in (Figure 6.6). Values of Impedance indicate hydrocarbon zone which ranges from 7200-10000 m/s *g/cc and VpVs ratio lies within 1.55 to 1.84. This cross plot shows Low values of lambda-rho indicate hydrocarbon zone which ranges from 10 – 34 GPa*g/cc (Figure 6.6).

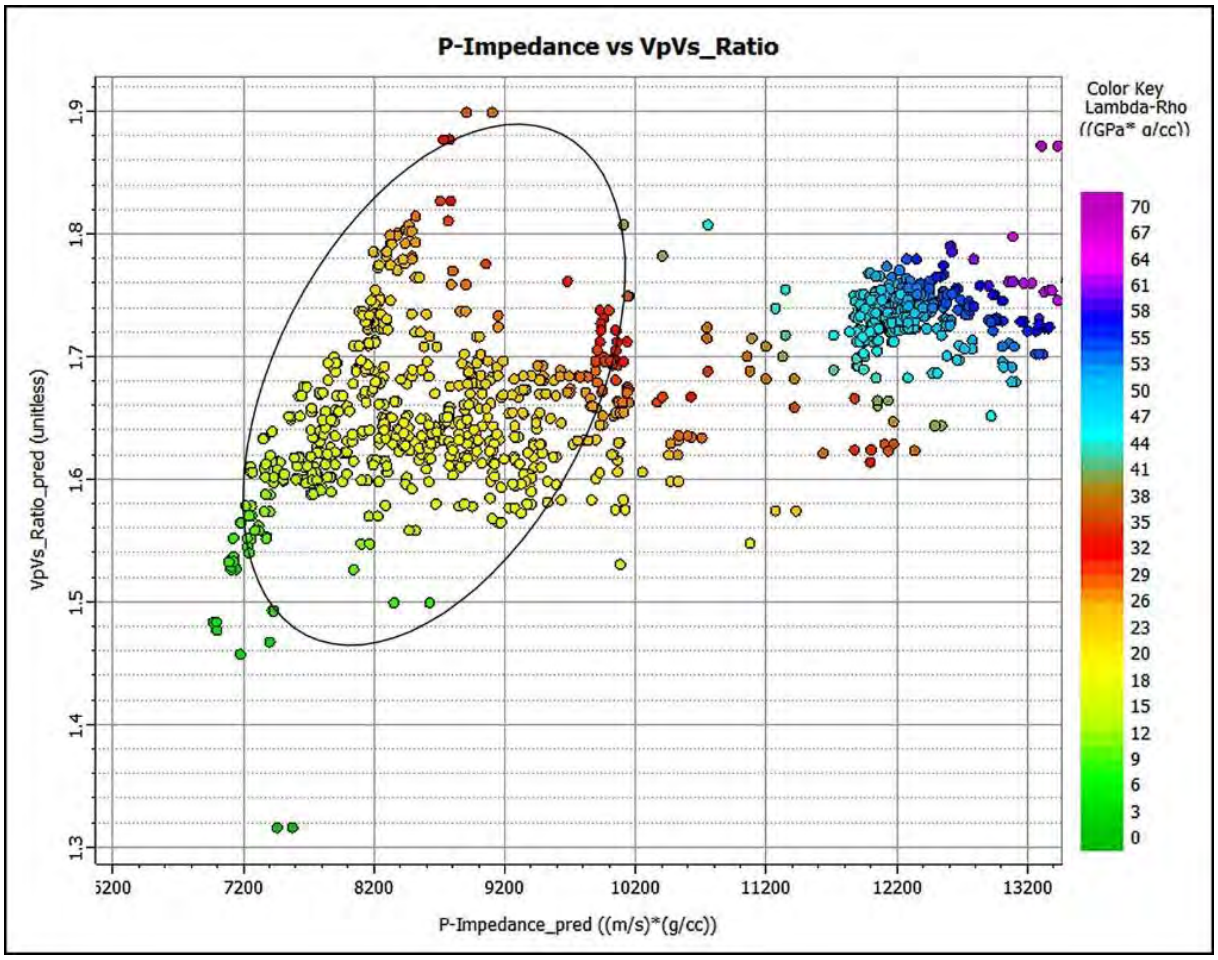

Figure 6. 6: The cross plot VpVs ratio and Impedance color coded with Lambda-Rho shows clear indication of hydrocarbon bearing zone.

The cross plot VpVs ratio and Impedance with Rhob on z-axis shows clear indication of hydrocarbon bearing zone. Values of Impedance indicate hydrocarbon zone which ranges from 7200-10000 m/s *g/cc and VpVs ratio lies within 1.55 to 1.84. Values of Rhob ranges from 2.2500 – 2.5000 g/cc indicates potential hydrocarbon zone, which is marked on cross plot (Figure 6.7).

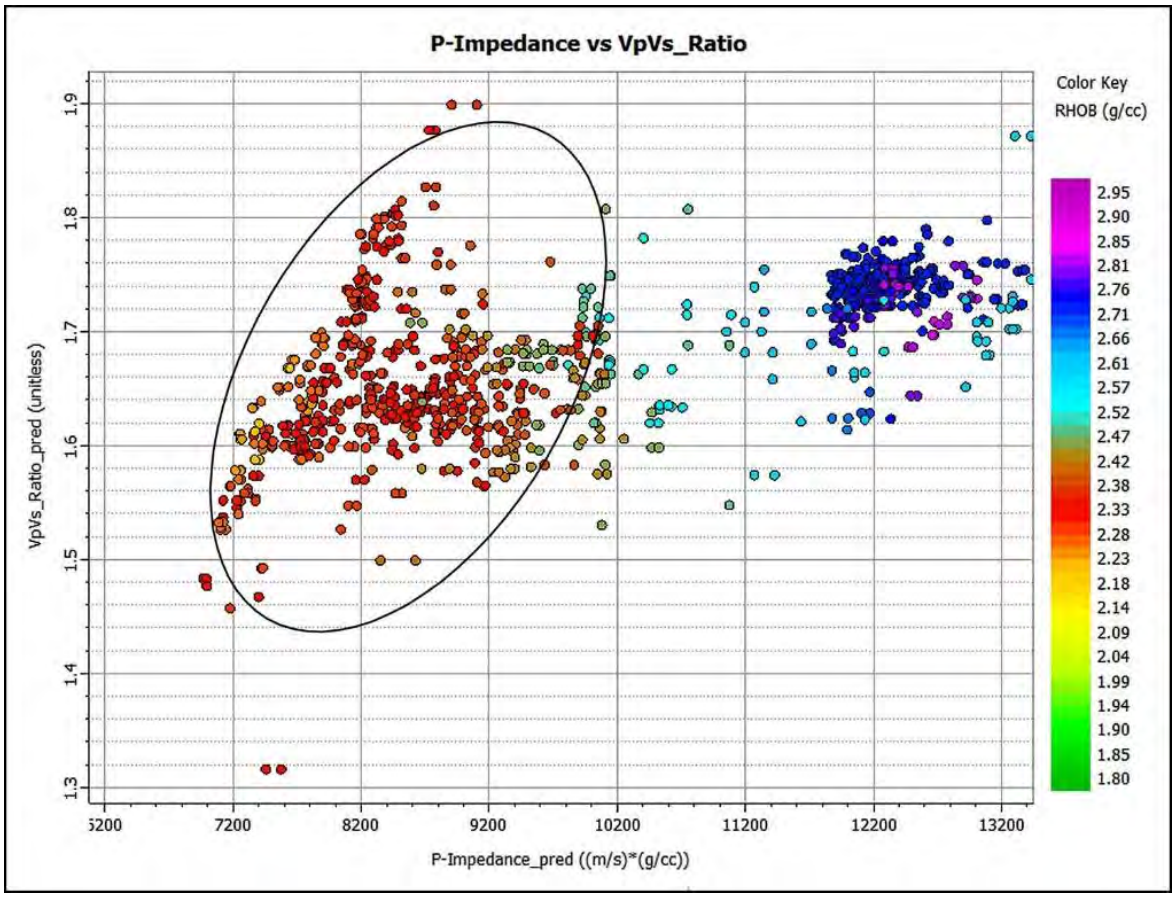

Figure 6. 7: The cross plot VpVs ratio and Impedance with Rhob on z-axis shows clear indication of hydrocarbon bearing zone.

The cross plot VpVs ratio and Impedance with Poisson ratio on z-axis shows clear indication of hydrocarbon bearing zone. Values of Impedance indicate hydrocarbon zone which ranges from 7200-10000 m/s *g/cc and VpVs ratio lies within 1.55 to 1.84. Values of Poisson ratio ranges from 0.26 - 0.32 marks the potential hydrocarbon zone (Figure 6.8).

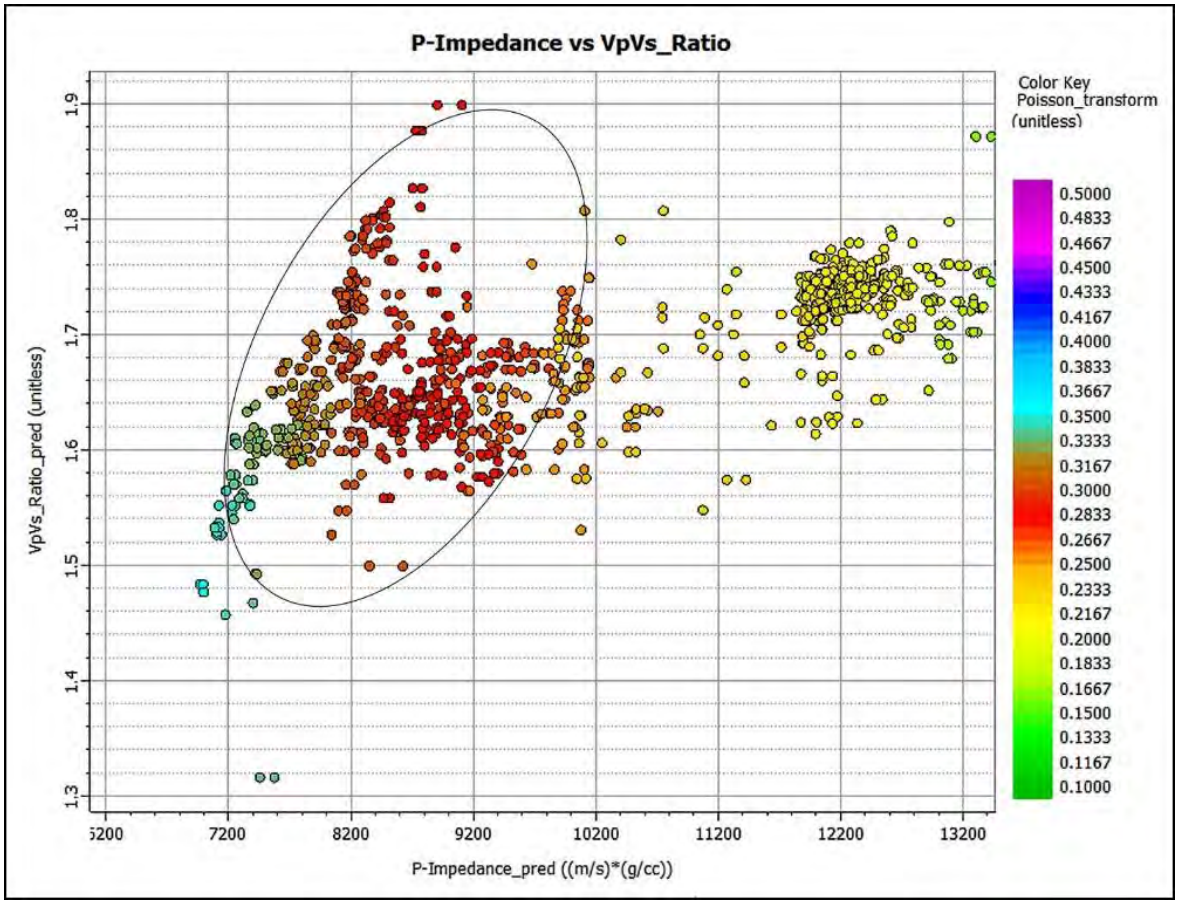

Figure 6. 8: The cross plot VpVs ratio and Impedance with Poisson ratio on z-axis shows clear indication of hydrocarbon bearing zone.

6.2.3: Lambda-Rho vs Mu-Rho

The cross plot Lambda-Rho and Mu-Rho color coded with Poisson ratio on z-axis shows clear indication of hydrocarbon bearing zone. Mu-Rho ranges from 20-40 ((Gpa)*(g/cc)) and Lambda-Rho lies within 10 – 34 ((GPa)*(g/cc)). Values of Poisson ratio ranges from 0.26 - 0.32 marks the potential hydrocarbon zone (Figure 6.9).

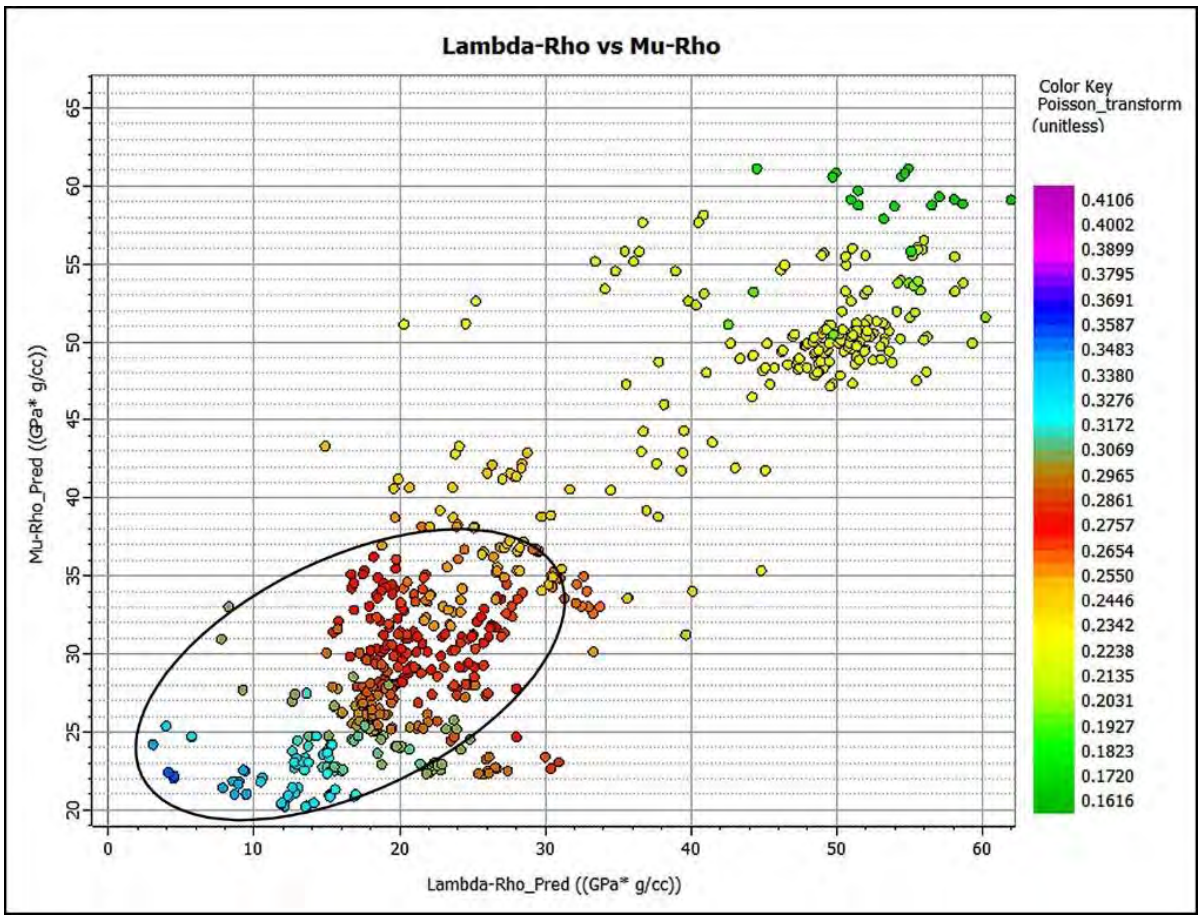

Figure 6. 9: The cross plot Lambda-Rho and Mu-Rho color coded with Poisson ratio shows clear indication of hydrocarbon bearing zone.

The cross plot Lambda-Rho and Mu-Rho color coded with P-Impedance shows clear indication of hydrocarbon bearing zone. Mu-Rho ranges from 20-40 ((Gpa)*(g/cc)) and Lambda-Rho lies within 10 – 34 ((GPa)*(g/cc)). P-Impedance indicate hydrocarbon zone with values ranges from 7200-10000 m/s *g/cc (Figure 6.10)

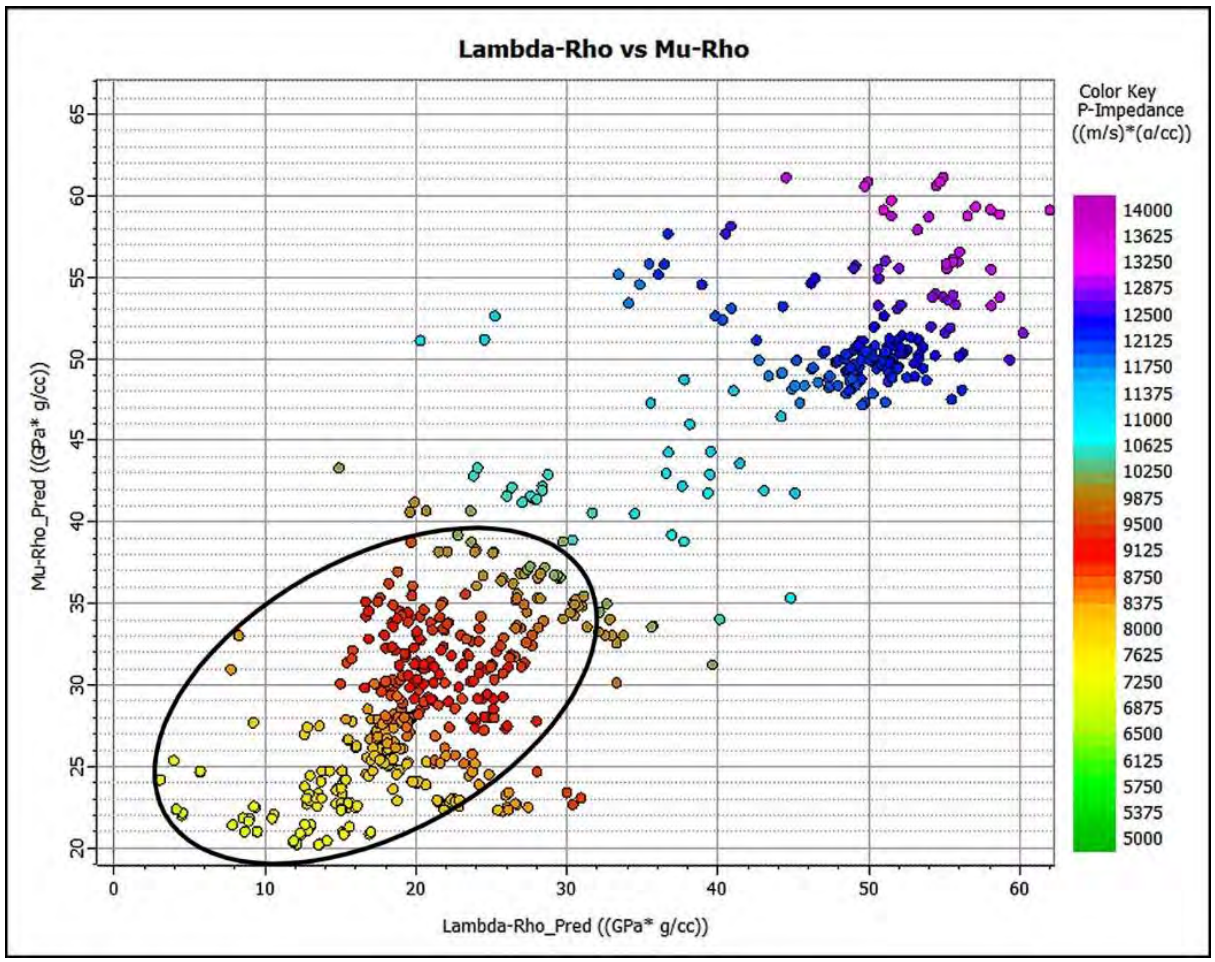

Figure 6. 10: The cross plot Lambda-Rho and Mu-Rho color coded with P-Impedance shows clear indication of hydrocarbon bearing zone.

The cross plot Lambda-Rho and Mu-Rho color coded with Phie shows clear indication of hydrocarbon bearing zone. Mu-Rho ranges from 20-40 ((Gpa)*(g/cc)) and Lambda-Rho lies within 10 – 34 ((GPa)*(g/cc)). Phie marks hydrocarbon zone with values ranges from 0.1080 – 0.1660 (Figure 6.11).

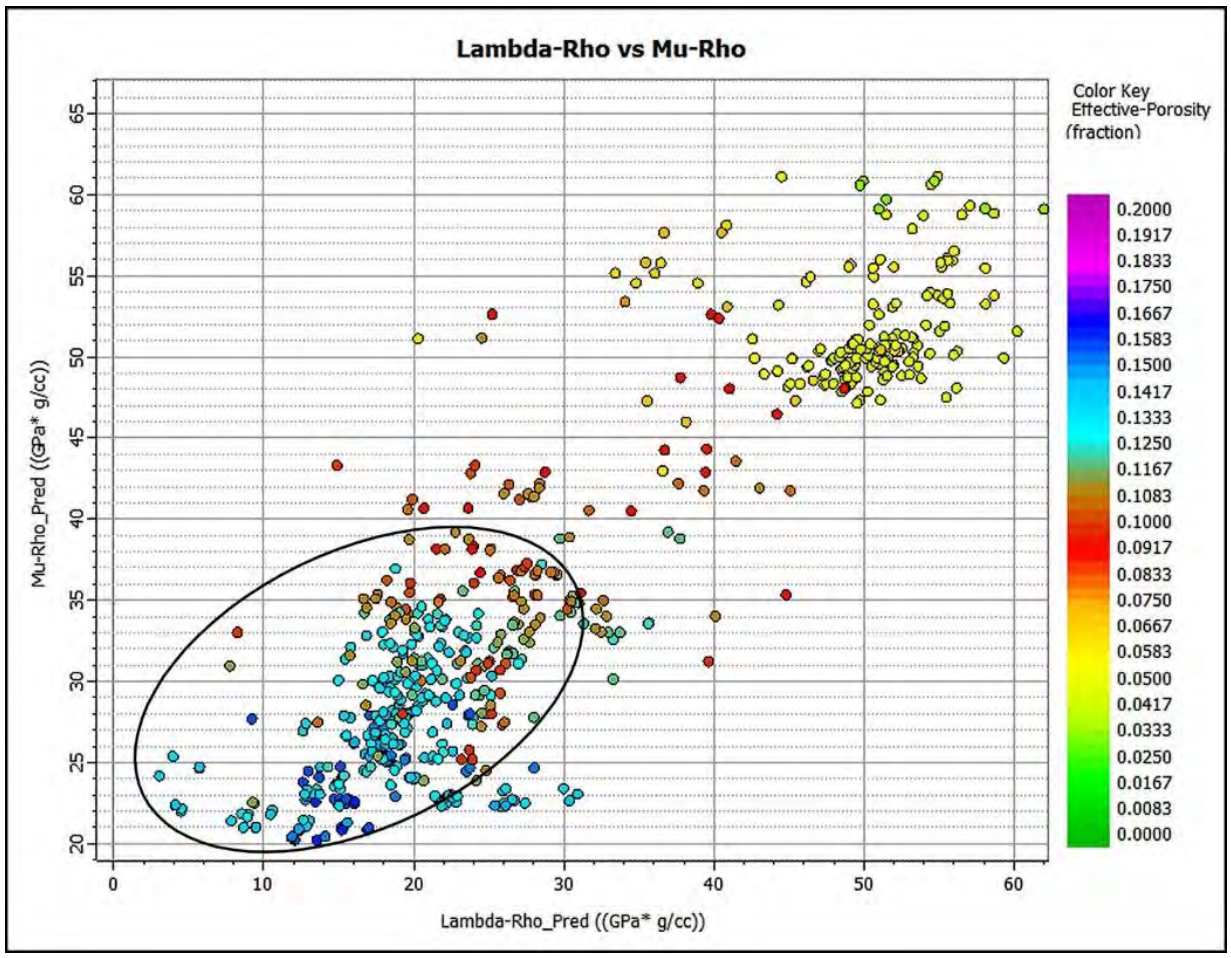

Figure 6. 11: The cross plot Lambda-Rho and Mu-Rho color coded with Phie shows clear indication of hydrocarbon bearing zone.

The cross plot Lambda-Rho and Mu-Rho with Rhob on z-axis shows clear indication of hydrocarbon bearing zone. Mu-Rho ranges from 20-40 ((Gpa)*(g/cc)) and Lambda-Rho lies within 10 – 34 ((GPa)*(g/cc)). Phie marks hydrocarbon zone with values ranges from 2.2500 – 2.5000 g/cc (Figure 6.12).

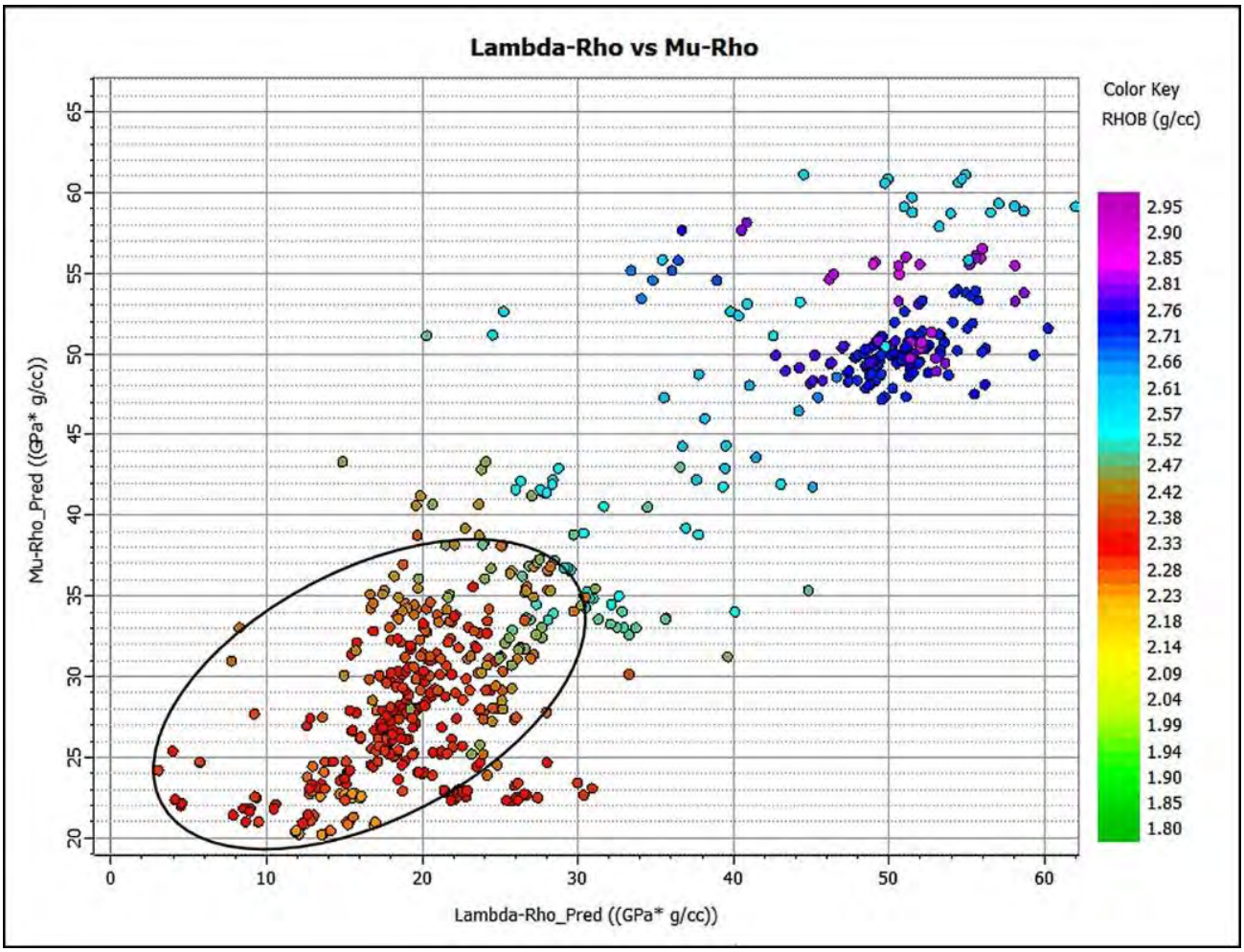

Figure 6. 12: The cross plot Lambda-Rho and Mu-Rho with Rhob on z-axis shows clear indication of hydrocarbon bearing zone.

CHAPTER 7

DISCUSSION AND CONCLUSION

7.1 Discussion

The study marked horizons of D, C and B-sands in the Sawan Area and successfully delineated the reservoir of interest in the Lower Goru Formation (Figure 3.4). This was done by combining 3D seismic data with well log data, allowing for a comprehensive understanding of the subsurface geology. The observed south-east dip of the scattered Sawan field provides important insights into the area's structural aspects (Figure 3.5, 3.6 and 3.7). The primary reservoir rock in the Sawan gas field was identified through a petrophysical analysis of the well logs from Sawan-01, Sawan-07, and Sawan-08. It is from the Upper Cretaceous Lower Goru C-sands (Figure 4.2 and 4.3). In addition to determining the thickness of the reservoir rock, this analysis also set the basis for further reservoir characterization.

The DT4S log in well Sawan-01 was successfully predicted using a machine learning model trained on data from wells Sawan-07 and Sawan-08. The capability of the model to capture challenging relationships between input well log data and the target DT4S log is demonstrated by its ability to achieve more than 94% matching accuracy (Figure 5.8). Despite variations in geological and petrophysical characteristics, this indicates that the model has generalized well to predict shear wave velocities in multiple parts of the reservoir (Figure 5.10). The selection of informative features for the model's input is an important step in machine learning success. The specific well log data features selected for this model and their significance (Figure 5.7), accurately predicting DT4S need to be discussed in depth. This can aid present and future researchers and professionals in comprehending the crucial factors influencing the predictions.

It's critical to emphasize the role of accurate DT4S prediction in overall reservoir characterization. Reservoir engineers can choose where to place wells, how to drill them, and how to optimize production with a precise estimation of this important parameter. This integration of reservoir characterization techniques and machine learning-based predictions emphasizes the study's practical value. Cross-plot analysis of elastic parameters was done after obtaining the predicted DT4S log for well Sawan-01 in order to identify the reservoir and non-reservoir zones. With the

help of this analysis, it is possible to distinguish between productive and non-productive zones within the Lower Goru Formation effectively (Figure 6.1 – 6.12).

It's crucial to talk about the machine learning model's wider relevance because that goes beyond the limits of this study. This method may be used by researchers and professionals to predict shear wave velocities in additional Lower Goru Formation wells or to modify it for use in various geological settings. Additionally, recommendations for potential model enhancements and new directions for future research should be made. This might include investigating different machine learning techniques, adding more data sources, or optimizing hyper-parameters for even more accurate prediction.

7.2 Conclusion

Conclusions of the study are as follows:

- Three dimensional seismic interpretations reveal that Study area is dipping towards south east.
- Petrophysical analysis indicates a zone of interest in Lower Goru Formation with PHIE values 11.4% in well Sawan-01 and 12.5% in well Sawan-07 and hydrocarbon saturation in well Sawan-01 is 66% while in well Sawan-07 is 52%.
- Gradient Boost Regressor gives 94% accurate prediction for data trained and tested on Wells Sawan-07 and Sawan-08. The model was applied to well Sawan-01
- Cross plot analysis based on elastic parameters estimated using Vp Rho and predicted Vs clearly segregates the reservoir and non-reservoir zone.

References

- Abdelwahhab, M. A., & Raef, A. (2020). Integrated reservoir and basin modeling in understanding the petroleum system and evaluating prospects: the Cenomanian reservoir, Bahariya Formation, at Falak Field, Shushan Basin, Western Desert, Egypt. *Journal of Petroleum Science and Engineering*, *189*, 107023.
- Afzal, J., Kuffner, T., Rahman, A., & Ibrahim, M. (2009, November). Seismic and well-log based sequence stratigraphy of the early Cretaceous, Lower Goru "C" sand of the Sawan gas field, middle Indus Platform, Pakistan. In Proceedings, Society of Petroleum Engineers (SPE)/Pakistan Association of Petroleum Geoscientists (PAPG) Annual Technical Conference, Islamabad, Pakistan.
- Ahammod, S., Hai, M. A., Islam, M. R., & Abu, S. M. (2014). Petro-Physical Analysis Of Reservoir Rock Of Fenchuganj Gas Field (Well# 03) Using Wireline Log. American Journal of Engineering Research (AJER), 3(8), 37-48.
- Ahmad, N., Fink, P., Sturrock, S., Mahmood, T., & Ibrahim, M. (2004). Sequence stratigraphy as predictive tool in lower goru fairway, lower and middle Indus platform, Pakistan. PAPG, ATC, 1, 85-104.
- Ahmad, S., & Ghazi, S. (2022). Depositional trends and reservoir geometries of the Early Cretaceous Lower Goru Formation in Lower Indus Basin, Pakistan: evidence from sequence stratigraphy. Journal of Petroleum Exploration and Production Technology, 12(11), 2981-3001.
- Ahmed, A. R., Ahmad, M., & Rehman, A. U. (2010, November). Comparison of Core/Log and Well Test Permeabilities-A Closer Look" Sawan Tight Sands". In SPE/PAPG Pakistan Section Annual Technical Conference (pp. SPE-142836). SPE
- Akhundi, H., Ghafoori, M., Reza, G., & Lashkaripour. (2014). Prediction of Shear Wave Velocity Using Artificial Neural Network Technique, Multiple Regression and Petrophysical Data: A Case Study in Asmari Reservoir (SW Iran). Scientific Research, 4, 11.
- Al-Anazi, A. F., & Gates, I. D. (2010). Support vector regression for porosity prediction in a heterogeneous reservoir: A comparative study. Computers & Geosciences, 36(12), 1494-1503.

- Ali, A., Kashif, M., Hussain, M., Siddique, J., Aslam, I., & Ahmed, Z. (2015). An integrated analysis of petrophysics, cross-plots and Gassmann fluid substitution for characterization of Fimkassar area, Pakistan: A case study. *Arabian Journal for Science and Engineering*, *40*, 181-193.
- Amigun, J. O., Olisa, B., & Fadeyi, O. O. (2012). Petrophysical analysis of well logs for reservoir evaluation: A case study of 'Laja'Oil Field, Niger Delta. Journal of Petroleum and Gas Exploration Research, 2(10), 181-187.
- Anemangely, M., Ramezanzadeh, A., Amiri, H., & Hoseinpour, S. A. (2019). Machine learning technique for the prediction of shear wave velocity using petrophysical logs. Journal of Petroleum Science and Engineering, 174, 306-327.
- Archie, G. E. (1952). Classification of carbonate reservoir rocks and petrophysical considerations. Aapg Bulletin, 36(2), 278-298.
- Ashraf, U., Zhang, H., Anees, A., Ali, M., Zhang, X., Shakeel Abbasi, S., & Nasir Mangi, H. (2020). Controls on reservoir heterogeneity of a shallow-marine reservoir in Sawan Gas Field, SE Pakistan: Implications for reservoir quality prediction using acoustic impedance inversion. Water, 12(11), 2972.
- Asquith, G. B., Krygowski, D., & Gibson, C. R. (2004). *Basic well log analysis* (Vol. 16). Tulsa: American Association of Petroleum Geologists.
- Avseth, P., van Wijngaarden, A. J., Flesche, H., Fristad, T., Rykkje, J., & Mavko, G. (2005, June). Seismic fluid prediction in poorly consolidated and clay laminated sands. In *67th EAGE Conference & Exhibition* (pp. cp-1). European Association of Geoscientists & Engineers.
- Berger, A., Gier, S., & Krois, P. (2009). Porosity-preserving chlorite cements in shallow-marine volcaniclastic sandstones: Evidence from Cretaceous sandstones of the Sawan gas field, Pakistan. AAPG bulletin, 93(5), 595-615.
- Berryman, J. G. (1999). Origin of Gassmann's equations. Geophysics, 64(5), 1627-1629.
- Buitinck, L., Louppe, G., Blondel, M., Pedregosa, F., Mueller, A., Grisel, O., ... & Varoquaux, G. (2013). API design for machine learning software: experiences from the scikit-learn project. arXiv preprint arXiv:1309.0238.

- Cai, J., Xu, K., Zhu, Y., Hu, F., & Li, L. (2020). Prediction and analysis of net ecosystem carbon exchange based on gradient boosting regression and random forest. Applied energy, 262, 114566.
- Ćalasan, M., Aleem, S. H. A., & Zobaa, A. F. (2020). On the root mean square error (RMSE) calculation for parameter estimation of photovoltaic models: A novel exact analytical solution based on Lambert W function. Energy conversion and management, 210, 112716.
- Castagna, J., Batzle, M., & Eastwood, R. (1985). Relationships between compressional-wave and shear wave velocities in clastic silicate rocks. Geophysics, 571-581.
- Chai, T., & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)? –Arguments against avoiding RMSE in the literature. Geoscientific model development, 7(3), 1247-1250.
- Condessa, L. G. (1995, June). Hydrocarbon identification in fresh-water bearing reservoirs using dynamic Poisson's ratio: a case study. In SPWLA Annual Logging Symposium (pp. SPWLA-1995). SPWLA.
- Cross, T. A., & Lessenger, M. A. (1988). Seismic stratigraphy. Annual review of earth and planetary sciences, 16(1), 319-354.
- Dar, Q. U. Z., Renhai, P., Ghazi, S., Ahmed, S., Ali, R. I., & Mehmood, M. (2021). Depositional facies and reservoir characteristics of the Early Cretaceous Lower Goru Formation, Lower Indus Basin Pakistan: Integration of petrographic and gamma-ray log analysis. Petroleum.
- Data, M. C., Komorowski, M., Marshall, D. C., Salciccioli, J. D., & Crutain, Y. (2016). Exploratory data analysis. Secondary analysis of electronic health records, 185-203.
- Dimililer, K., Dindar, H., & Al-Turjman, F. (2021). Deep learning, machine learning and internet of things in geophysical engineering applications: An overview. Microprocessors and Microsystems, 80, 103613.
- Dominguez, G. C. (1992). Carbonate reservoir characterization: A geologic-engineering analysis, Part I. Elsevier.
- Donaldson, E. C., & Tiab, D. (2004). Petrophysics: Theory and Practice of measuring reservoir rock and fluid transport properties. Gulf Professional Publishing

- Dramsch, J. S. (2020). 70 years of machine learning in geoscience in review. Advances in geophysics, 61, 1-55.
- Dubrule, O., Thibaut, M., Lamy, P., & Haas, A. (1998). Geostatistical reservoir characterization constrained by 3D seismic data. Petroleum Geoscience, 4(2), 121-128.
- Fallah-Mehdipour, E., Bozorg Haddad, O., & Mariño, M. A. (2012). Real-time operation of reservoir system by genetic programming. Water resources management, 26, 4091-4103.
- Ghorbani, A., Jafarian, Y., & Maghsoudi, M. (2012). Estimating shear wave velocity of soil deposits using polynomial neural networks: Application to liquefaction. Computers & Geosciences, 86-94.
- Glover, P. W. (2000). Petrophysics. University of Aberdeen, UK, 270.
- Goodway, B. (2001). AVO and Lamé constants for rock parameterization and fluid detection. CSEG recorder, 26(6), 39-60.
- Goodway, B., Chen, T., & Downton, J. (1997). Improved AVO fluid detection and lithology discrimination using Lamé petrophysical parameters; "λρ", "μρ", & "λ/μ fluid stack", from P and S inversions. In SEG technical program expanded abstracts 1997 (pp. 183-186). Society of Exploration Geophysicists.
- Grana, D., Paparozzi, E., Mancini, S., & Tarchiani, C. (2013). Seismic driven probabilistic classification of reservoir facies for static reservoir modelling: a case history in the Barents Sea. Geophysical prospecting, 61(3), 613-629.
- Guedez, L. A. L. (2019). CO 2 Sequestration Assessment Using Multicomponent 3D Seismic Data: Rock Springs Uplift, Wyoming (Doctoral dissertation, University of Houston).
- Gul, M. A., Awan, R. S., Khan, A., Iltaf, K. H., & Butt, S. E. H. (2023). 2D seismic interpretation of Sawan gas field integrated with petrophysical analysis: A case study from Lower Indus Basin, Pakistan. Energy Geoscience, 4(2), 100143.
- Gunter, G. W., Finneran, J. M., Hartmann, D. J., & Miller, J. D. (1997, October). Early determination of reservoir flow units using an integrated petrophysical method. In SPE Annual Technical Conference and Exhibition? (pp. SPE-38679). SPE.

- Hamada, G. M. (2004). Reservoir fluids identification using Vp/Vs ratio. Oil & Gas Science and Technology, 59(6), 649-654.
- Han, D. H., & Batzle, M. L. (2004). Gassmann's equation and fluid-saturation effects on seismic velocities. Geophysics, 69(2), 398-405.
- Han, D. H., Liu, J., & Batzle, M. (2005, November). Measurement of shear wave velocity of heavy oil. In SEG International Exposition and Annual Meeting (pp. SEG-2005). SEG.
- Handwerger, D. A., Cooper, A. K., O'Brien, P. E., Williams, T., Barr, S. R., Dunbar, R. B., ... & Richter, C. (2004). Synthetic seismograms linking ODP sites to seismic profiles, continental rise and shelf of Prydz Bay, Antarctica. In *Proc. Ocean Drill. Prog., Sci. Res* (Vol. 188).
- Helgerud, M. B., Dvorkin, J., Nur, A., Sakai, A., & Collett, T. (1999). Elastic-wave velocity in marine sediments with gas hydrates: Effective medium modeling. Geophysical research letters, 26(13), 2021-2024.
- Hussain, M., Getz, S. L., & Oliver, R. (1991). Hydrocarbon accumulation parameters in the central portion of the lower Indus basin of In Proceedings of an international seminar. Islamabad, Pakistan, Ministry of Petroleum and Natural Resources (pp. 106-119).
- Islam, M. A., Yunsi, M., Qadri, S. T., Shalaby, M. R., & Haque, A. E. (2021). Three-dimensional structural and petrophysical modeling for reservoir characterization of the Mangahewa formation, Pohokura Gas-Condensate Field, Taranaki Basin, New Zealand. Natural Resources Research, 30, 371-394.
- Julià, J., & Mejía, J. (2004). Thickness and Vp/Vs ratio variation in the Iberian crust. Geophysical Journal International, 156(1), 59-72.
- Kaczmarczyk-Kuszpit, W. (2021). Reservoir characterization based on the Lambda-Mu-Rho method–case study. Nafta-Gaz, 77(10), 625-632.
- Kadri, I. B. (1995). Petroleum geology of Pakistan. Pakistan Petroleum Limited.
- Kamel, M., & Mabrouk, W. (2002). An equation for estimating water saturation in clean formations utilizing resistivity and sonic logs: theory and application. Journal of Petroleum Science and Engineering, 159-168.
- Kazmi, A. H., & Abbasi, I. A. (2008). Stratigraphy & historical geology of Pakistan.
- Khan, M. S. I., Islam, N., Uddin, J., Islam, S., & Nasir, M. K. (2022). Water quality prediction and classification based on principal component regression and gradient

boosting classifier approach. Journal of King Saud University-Computer and Information Sciences, 34(8), 4773-4781.

- Khan, M., Nawaz, S., Shah, M., & Hasan, M. (2016). Interpreting seismic profiles in terms of structure and stratigraphy, an example from Lower Indus Basin Pakistan. Universal Journal of Geoscience, 4(3), 62-71.
- Khatibi, S., & Aghajanpour, A. (2020). Machine learning: A useful tool in geomechanical studies, a case study from an offshore gas field. Energies, 13(14), 3528.
- Lucia, F. J., Kerans, C., & Jennings Jr, J. W. (2003). Carbonate reservoir characterization. Journal of Petroleum Technology, 55(06), 70-72.
- Maddu, R., Pradhan, I., Ahmadisharaf, E., Singh, S. K., & Shaik, R. (2022). Short-range reservoir inflow forecasting using hydrological and large-scale atmospheric circulation information. Journal of Hydrology, 612, 128153.
- Manzi, M. S., Cooper, G. R., Malehmir, A., & Durrheim, R. J. (2020). Improved structural interpretation of legacy 3D seismic data from Karee platinum mine (South Africa) through the application of novel seismic attributes. Geophysical Prospecting, 68(1-Cost-Effective and Innovative Mineral Exploration Solutions), 145-163.
- Maxwell, S. C., Cho, D., Pope, T., Jones, M., Cipolla, C., Mack, M., ... & Leonard, J. (2011, January). Enhanced reservoir characterization using hydraulic fracture microseismicity. In SPE hydraulic fracturing technology conference and exhibition (pp. SPE-140449). SPE.
- Miller H, R. (1986). *The geological interpretation of well logs.* U.S.A.
- Mills, W. R., Stromswold, D. C., & Allen, L. S. (1988, June). Pulsed Neutron Porosity Logging. In SPWLA Annual Logging Symposium (pp. SPWLA-1988). SPWLA.
- Mitchum, R., & Vail, P. (1997). Seismic Stratigraphy--Applications to Hydrocarbon Exploration. AAPG.
- Morea, J. P. (2023). The economic, social and environmental impacts of offshore oil exploration in Argentina: A critical appraisal. The Extractive Industries and Society, 15, 101295.

- Mughal, M. R., & Akhter, G. (2021). Predicting the gas resource potential in reservoir C-sand interval of Lower Goru Formation, Middle Indus Basin, Pakistan. Open Geosciences, 13(1), 49-71.
- Nie, P., Roccotelli, M., Fanti, M. P., Ming, Z., & Li, Z. (2021). Prediction of home energy consumption based on gradient boosting regression tree. Energy Reports, 7, 1246-1255.
- Obilo, J. C., Francis, M., Ojo, C., & Licalsi, P. (2007, June). Hydrocarbon Detection Using Lambda–Rho Attribute Computed from Prestack Seismic Data. In 69th EAGE Conference and Exhibition incorporating SPE EUROPEC 2007 (pp. cp-27). European Association of Geoscientists & Engineers.
- Otchere, D. A., Ganat, T. O. A., Ojero, J. O., Tackie-Otoo, B. N., & Taki, M. Y. (2022). Application of gradient boosting regression model for the evaluation of feature selection techniques in improving reservoir characterisation predictions. Journal of Petroleum Science and Engineering, 208, 109244.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. the Journal of machine Learning research, 12, 2825-2830.
- Pham, N., Wu, X., & Zabihi Naeini, E. (2020). Missing well log prediction using convolutional long short-term memory network. Geophysics, 85(4), WA159-WA171.
- Qiao, Y., & An, H. (2007). Study of petrophysical parameter sensitivity from well log data. Applied Geophysics, 4, 282-287.
- Rao, H., Shi, X., Rodrigue, A. K., Feng, J., Xia, Y., Elhoseny, M., ... & Gu, L. (2019). Feature selection based on artificial bee colony and gradient boosting decision tree. Applied Soft Computing, 74, 634-642.
- Rider, M. H. (1990). Gamma-ray log shape used as a facies indicator: critical analysis of an oversimplified methodology. *Geological Society, London, Special Publications*, *48*(1), 27-37.
- Russell, B., Lines, L., & Hampson, D. (2003). Application of the radial basis function neural network to the prediction of log properties from seismic attributes. Exploration Geophysics, 15 - 23.

- Schölkopf, B., Platt, J. C., Shawe-Taylor, J., Smola, A. J., & Williamson, R. C. (2001). Estimating the support of a high-dimensional distribution. Neural computation, 13(7), 1443-1471.
- Sheriff, R. E. (2002). Encyclopedic dictionary of applied geophysics. Society of exploration geophysicists.
- Singh, U., Rizwan, M., Alaraj, M., & Alsaidan, I. (2021). A machine learning-based gradient boosting regression approach for wind power production forecasting: A step towards smart grid environments. Energies, 14(16), 5196.
- Telford, W. M., Geldart, L. P., & Sheriff, R. E. (1990). *Applied geophysics*. Cambridge university press.
- Veeken, P. C., & Rauch-Davies, M. (2006). AVO attribute analysis and seismic reservoir characterization. First break, 24(2).
- Venieri, M., Pedersen, P. K., & Eaton, D. W. (2021). Predicting unconventional reservoir potential from wire-line logs: A correlation between compositional and geomechanical properties of the Duvernay shale play of western Alberta, Canada. AAPG Bulletin, 105(5), 865-881.
- Wang, J., Cao, J., & Yuan, S. (2020). Shear wave velocity prediction based on adaptive particle swarm optimization optimized recurrent neural network. Journal of Petroleum Science and Engineering, 107466.
- Wang, X. Q., Schubnel, A., Fortin, J., David, E. C., Guéguen, Y., & Ge, H. K. (2012). High Vp/Vs ratio: Saturated cracks or anisotropy effects. Geophysical Research Letters, 39(11).
- Weisstein, E. W. (2006). Correlation coefficient. https://mathworld. wolfram. com/.
- Wrona, T., Pan, I., Gawthorpe, R., & Fossen, H. (2018). Seismic facies analysis using machine learning. Geophysics, 83-95.
- Wu, Q., Liu, Q., Liu, S., Wang, S., Yu, J., Ayers, W. B., & Zhu, Q. (2022). Estimating Reservoir Properties from 3D Seismic Attributes Using Simultaneous Prestack Inversion: A Case Study of Lufeng Oil Field, South China Sea. SPE Journal, 27(01), 292-306.
- Wyllie, M. R. J., Gregory, A. R., & Gardner, L. W. (1956). Elastic wave velocities in heterogeneous and porous media. Geophysics, 21(1), 41-70.

- Yeo, I. K., & Johnson, R. A. (2000). A new family of power transformations to improve normality or symmetry. Biometrika, 87(4), 954-959.
- Zaigham, N. A., & Mallick, K. A. (2000). Prospect of hydrocarbon associated with fossil-rift structures of the southern Indus basin, Pakistan. AAPG bulletin, 84(11), 1833-1848.
- Zhang, Y., Meratnia, N., & Havinga, P. (2010). Outlier detection techniques for wireless sensor networks: A survey. IEEE communications surveys & tutorials, 12(2), 159-170.
- Zimmer, M., Prasad, M., & Mavko, G. (2002). Pressure and porosity influences on Vp-Vs ratio in unconsolidated sands. The Leading Edge, 21(2), 178-183.