APPLICATION OF MACHINE LEARNING TECHNIQUES FOR DETAILED FACIES MODELLING OF LOWER GORU FORMATION BY USING THE WELL LOG DATA OF SAWAN GAS FIELD, CENTRAL INDUS BASIN, PAKISTAN

BY

MUHAMMAD UMAIS

MPHIL GEOLOGY 2020-2022

DEPARTMENT OF EARTH SCIENCES
QUAID-I-AZAM UNIVERSITY
ISLAMABAD, PAKISTAN

APPLICATION OF MACHINE LEARNING TECHNIQUES FOR DETAILED FACIES MODELLING OF LOWER GORU FORMATION BY USING THE WELL LOG DATA OF SAWAN GAS FIELD, CENTRAL INDUS BASIN, PAKISTAN

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Philosophy in Geology

BY

MUHAMMAD UMAIS

MPHIL GEOLOGY 2020-2022

DEPARTMENT OF EARTH SCIENCES
QUAID-I-AZAM UNIVERSITY
ISLAMABAD, PAKISTAN

CERTIFICATE

This dissertation is submitted by **Muhammad Umais** S/o Muhammad Yousaf is accepted in its present form by the Department of Earth Sciences, Quaid-i-Azam University Islamabad as it satisfies the requirement for the award of M.Phil. Degree in Geology.

Supervisor: _____

Dr. Matloob Hussain

Chairman: _____

Dr. Aamir Ali

Dated: _____

Dedication

To My Beloved Parents, Siblings, friends and teachers

ACKNOWLEDGMENT

All praise to Almighty Allah who gave me courage and made me able in completing my thesis work successfully. I offer my gratitude to last Prophet Muhammad (PBUH).

I am also extremely grateful to my honorable supervisor **Dr. Matloob Hussain** for their wide knowledge and logical way of thinking have been of great value for me. Their personal guidance has provided a good basis for this study. I must acknowledge the cooperation of all my teachers whose direction and support have been the source of my success.

I would like to pay my colossal respect to Dr. Muhammad Armaghan Faisal. This work could never have been completed without him. Most of the credit of IP software support goes to him.

I would like to acknowledge my seniors and friends for their help and suggestions during my work. I do express my sincere thanks to Taimoor Hassan, Irfan Zia, Muhammad Bilal Malik and Anees-ur-Rahman for their sincere guidance, help, moral support and encouragement

I am thankful to my parents for moral and financial support. I am indebted to my loving parents, who encouraged and motivated me to face the challenges of life during my academic career. I am also highly indebted to my brothers and my sister and specially my seniors for their love and cooperation.

Last but not the least, I am thankful to all my friends and fellows who encouraged and motivated me throughout my academics.

Muhammad Umais
M.Phil. Geology
2020-2022

ABSTRACT

The advancement in computational technology revolutionized the exploration geoscience. Machine learning tools can be used for the analysis of large datasets and to predict the relationships present among the data. Information about subsurface rock formations is gathered during the exploration of hydrocarbons. The subsurface data is quite large and difficult to handle. So, machine learning techniques like K-means clustering and self-organizing maps can be used to find manage subsurface data. The Lower Goru Formation is a prolific reservoir of Middle Indus Basin and is known for its lithological heterogeneity. The lithology of the Lower Goru Formation comprises of alternating layers of sandstone and shale. The sandstone of the Lower Goru Formation is further divided into D, C, B and A interval. Efficient identification of lithology is the key step of reservoir characterization. Present study focuses the petrophysical analysis of the Lower Goru Formation followed by the facies classification by using K-means clustering and self-organizing maps. The study was performed by using the well log data of Sawan-02, Sawan-03 and Sawan-07 wells.

The petrophysical analysis reveals that the C interval of the Lower Goru Formation is a good reservoir having the average volume of shale 15%, effective porosity 12% and water saturation 20%. K-means clustering is used to classify the data into ten clusters which were then consolidated into five groups by using hierarchical clustering. These groups of data were used for facie classification. It is evident from K-means clustering the C interval of the Lower Goru Formation mainly consists of reservoir facies. The self-organizing map were used to display the multidimensional well data into two-dimensional facies map. The result obtained from self-organizing maps confirm the results of K-means clustering. So, these methods can be efficiently used for the lithological identification of reservoir formation which will reduce the risk associated with exploration of hydrocarbons.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Hydrocarbons are the major source of energy in modern world. The economy of a country is also dependent on its hydrocarbon reserves. Effective exploration strategy is thus carried out to explore the sufficient amount of hydrocarbon to fulfil the growing energy demands. The scientific knowledge greatly improves the efficiency of petroleum exploration. The excessive production of hydrocarbon is depleting the existing reservoirs. So, modern methods are adopted to explore new hydrocarbon reservoirs.

Hydrocarbons are accumulations are hosted by the reservoir rocks. The economic viability of hydrocarbon accumulation is greatly controlled by the properties of reservoir rock. Petrophysical analysis is considered as an important step of petroleum exploration because it aids in reservoir characterization. Geophysical properties of rocks are measured by using downhole logging tools such as radioactivity, resistivity and sonic tools. These geophysical properties are then interpreted to give the information about petrophysical properties of rocks. Key petrophysical properties of reservoir includes the volume of shale, average porosity, effective porosity and saturation of water. These properties are used to assess the reservoir character of the rock encountered in wellbore (Fischetti & Andrade, 2002).

Advancement in computational technology has revolutionized the field of hydrocarbon exploration. Different computational tools are thus adopted to enhance the control over subsurface data. Machine learning is a tool that allows to handle the large datasets of subsurface information. It also enables the automatic analysis and model building of data to predict the relationships among data and fetch important information (Alpaydin, 2020).

Facies are the rock units having the specific lithological characteristics which reflect the particular environment of deposition (Nichols, 2013). Identification of lithology is a crucial step of reservoir characterization. Facies are identified on the basis of core and log data. In present research work facies of the Lower Goru Formation are classified by using the two machine learning techniques that are K-means clustering and self-organizing maps. The Lower Goru Formation is comprised of alternating layers of sand and shale. Sand intervals of the Formation are also acting as a proven

reservoir in Lower Indus Basin Pakistan. The lithological heterogeneity of the Lower Goru Formation is characterized by using K-means clustering and self-organizing maps.

K-means clustering classifies the data into clusters on the basis of similarities. The clusters of data are used to classify the heterogeneity of the rock formation which can be interpreted as rock facies. Self-organizing map is unsupervised machine learning technique that is used to display the multidimensional well data in form of two-dimensional facies map. Machine learning techniques can be used in detailed analysis of electrofacies.

1.1.Objectives

The main objectives of this research work are as follows

1. Petrophysical analysis of the Lower Goru Formation using well log data of Sawan-02, Sawan-03 and Sawan-07 to delineate the pay zones and for the estimation of key reservoir properties.
2. K-means cluster analysis for the identification of electrofacies present within the intervals of the Lower Goru Formation.
3. Modeling of electrofacies present within the intervals of the Lower Goru Formation by using self-organizing maps.
4. Comparison of electrofacies obtained by K-means clustering and self-organizing maps to enhance the quality of reservoir characterization.

1.2.Study Area

The study area for present research work is the Sawan Gas Field. Sawan Gas Field is considered as one of major gas fields of Pakistan. It was discovered by OMV Pakistan in 1998 with joint venture of ENI Pakistan and PPL. Sawan Gas Field is located in Khairpur district of Sindh province about 500 km north of Karachi. Geographically, it is located in Thar Desert. The geographical coordinates of Sawan Gas Field are 26°55'34.3452" to 27°17'14.424" N and 68°32'28.9" to 69°18' 40.464" E. Geologically, Sawan Gas Field is located on southeastern part of Jacobabad- Khairpur high. The well data of Sawan-02, Sawan-03 and Sawan-07 is used during the research work. The map showing the location of these wells is shown in Fig.1.1.

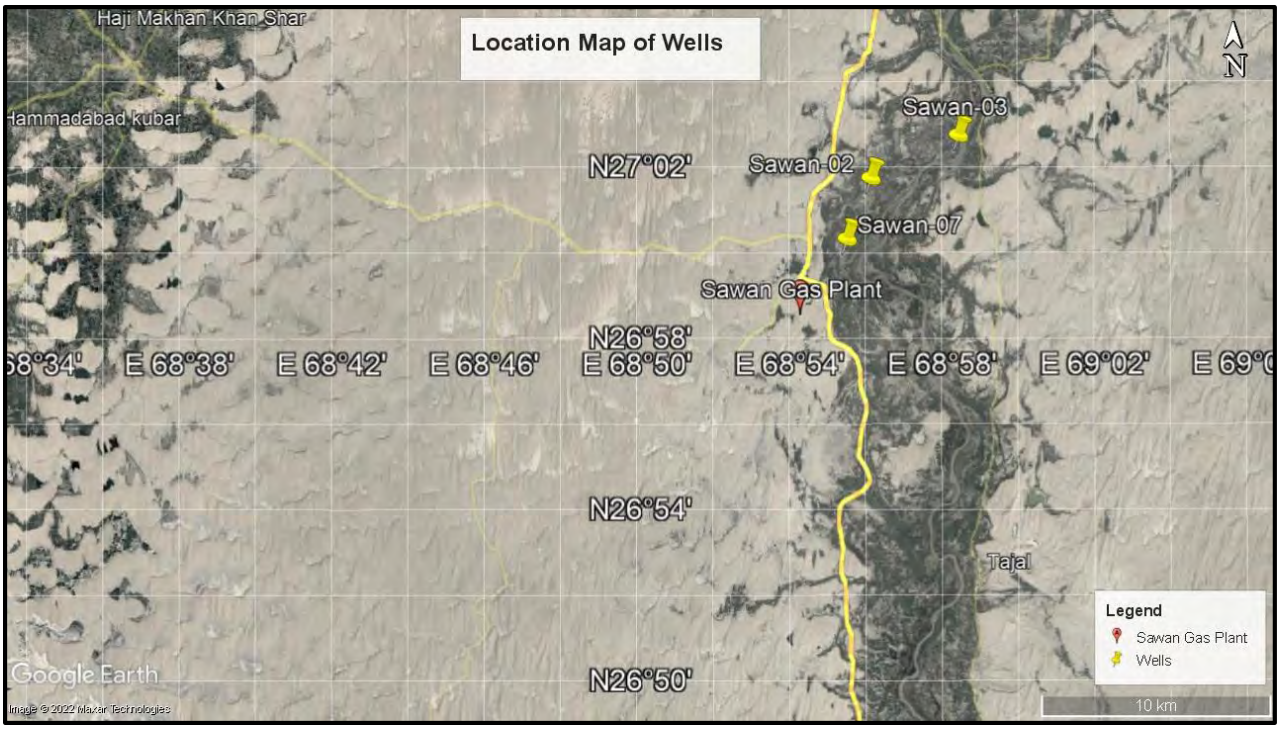


Figure 1.1: Map showing the location of study area (Google Earth).

1.3.Data Sets

Well data of three wells of Sawan Gas Field was used to achieve the objectives of the research works. Well data of Sawan-02, Sawan-03 and Sawan-07 was acquired from LMKR after the approval of DGPC. Acquired data set includes the Las files of well logs, formation tops and well headers. The basic information of each well is shown in table 1.1. The formation tops encountered in each well are shown in table 1.2.

Table 1.1: Showing information about wells including their location, status, class and total depth.

Well Name	Latitude	Longitude	Status	Class	TD(m)
Sawan-02	27 1 22.70	68 56 1.70	Gas	Appraisal	3500
Sawan-03	27 2 22.70	68 58 19.50	Gas	Development	3650
Sawan-07	26 59 57.42	68 55 23.94	Gas	Development	3400

Table 1.2: Showing the information about depths (meters) of formation tops encountered in wells.

	Top Values		
Formation Names	**Sawan-03**	**Sawan-02**	**Sawan-07**
Alluvium-Siwalik	0	12	0
Drazinda Member	68	58	54
Pirkoh Member	----------	178	187
Kirthar Formation	68	----------	----------
Pirkoh Member	198	----------	----------
Sirki Member	304	263	275
Habib Rahi Member	311	273	285
Ghazij Member	334	301	304
Laki Formation	334	----------	304
Sui Main Limestone	1133	1109	1116
Ranikot Formation	1273.3	1251	1261
Upper Goru Formation	2470.8	2412	2444
Lower Goru Formation	2812.4	2683	2691
D Interval	----------	3171	----------
C Interval	3365.6	3259	3239.39
B Interval	3597	3450	3432

1.4.Methodology

The generalized methodology adopted for present research work is shown in Fig. 1.2. The well data of Sawan-02, Sawan-03 and Sawan-07 including the Las files, well header information and formation tops were acquired from LMKR after getting approval from Directorate General of Petroleum Concession (DGPC). The quality of data was then checked thoroughly to filter out

missing and erraneous data. After this the data was loaded into the software for petrophysical interpretation. Petrophysical interpretation is followed by K-means cluster analysis and self-organizing map modelling. The detailed workflow of each analysis is discussed in detail in respective chapters. The results thus obtained are summariezed in discussion section. The conclusion are than drawn from the research work. Recommendations are also suggested at the end to improve the effeciency of exploration strategy.

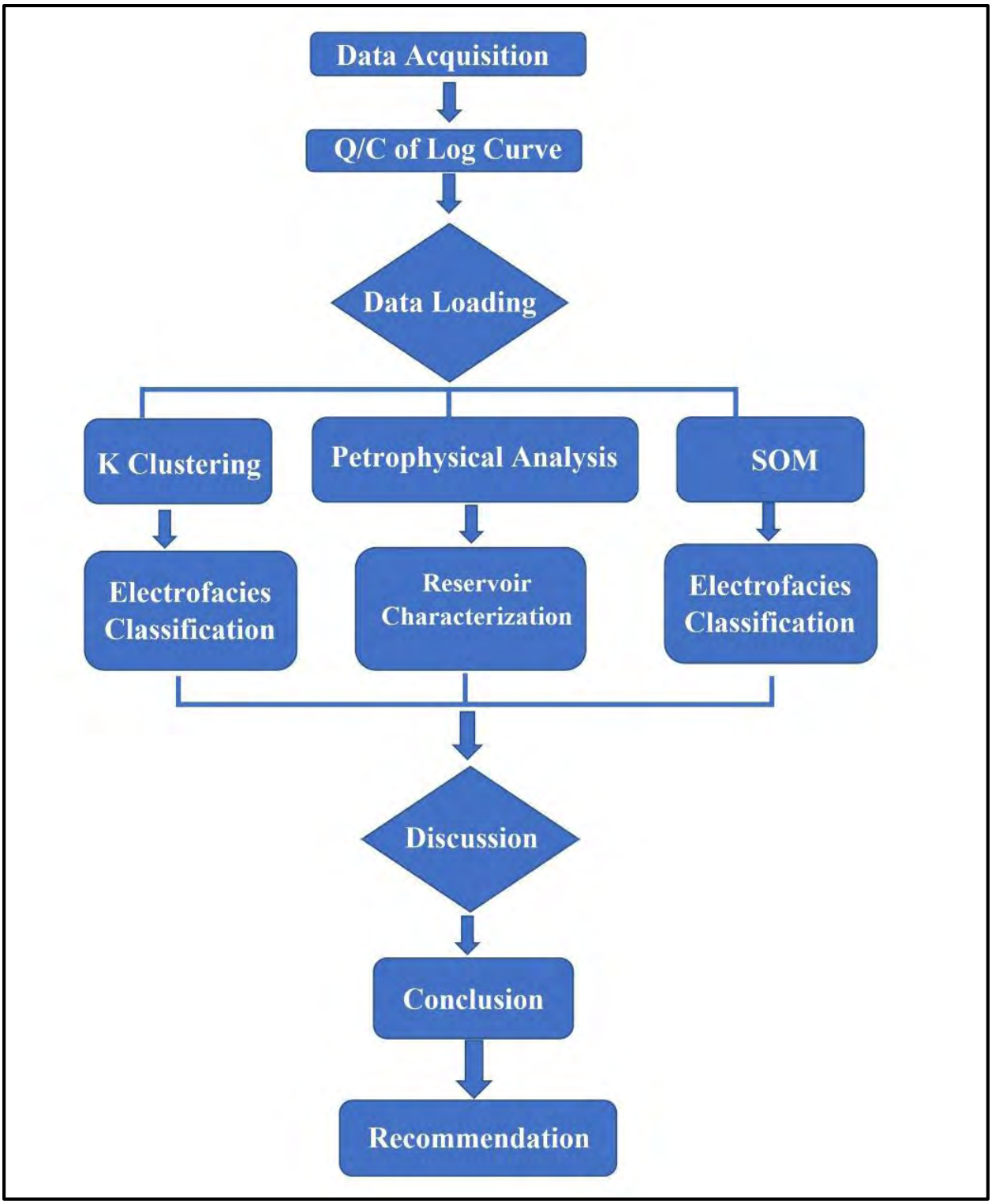


Figure 1.2: Showing the generalized methodology adopted for the electrofacies classification of the Lower Goru Formation.

Chapter 2

Geology and Tectonics

The geology of study area is very important to minimize the risks associated with hydrocarbon exploration and production. A strong command over geology is required for efficient analysis of well log responses. The rock formations present in subsurface are geologically related with each other so, the understanding of these relationships is quite important for efficient petrophysical analysis. The geology of any area is determined by its basin forming processes, basin fill (sediments and rocks) and basin modification processes (Kingston, et al., 1983). The study of geology and stratigraphy aids in better understanding of rock formations and the relationships among these rocks which help in improvement of exploration strategy and predictions of past geological events. This chapter analyzes geologic, tectonics and stratigraphic setting of the study area.

2.1.Geological Setting

The Indus Basin of Pakistan is divided into three main parts that are named as Upper Indus Basin, Central Indus Basin and Southern Indus Basin. Sargodha highs separates the Upper Indus Basin from Central Indus Basin. On the other hand, the boundary of Middle Indus Basin and Southern Indus Basin is marked by Jacobabad-Khairpur and Mari-Kandkot highs. The study area lies within the Lower Indus Basin on the southeastern flank of Jacobabad-Khairpur high as shown in Fig. 2.1. The structural and stratigraphic makeup of Lower Indus Basin yielded several economic petroleum traps. Noticeable discoveries of this area include Sawan, Kadanwari, Kandra, Miano etc. The combined traps along with suitable migration, reservoir charge and accumulation timing are the best traps found in the Central Indus Basin (Kazmi & Jan, 1997).

The geology of study area is mainly fashioned by three episodes of tectonic events that are enlisted below (Ahmad, Fink, Sturrock, Mahmood, & Ibrahim, 2004).

- Late Cretaceous uplift and erosion.
- Late Paleocene tear faulting.
- Late Tertiary uplift of Jacobabad-Khairpur high.

First phase of tectonic activity was observed in Late Cretaceous near K-T boundary. It resulted in uplifting and erosion on stratum. This tectonic event is marked by thinning of Ranikot clasts in the direction of paleohigh. The second tectonic event resulted in wrench faulting induced due to the anticlockwise rotation and collision of Indian Plate with Eurasian Plate (Zaigham & Mallick, 2000). These faults originated as NNW to SSE single rooted faults that transforms into multiple echelon faults near the Upper Goru Formation. These faults originate from basement and after cutting the Cretaceous section they die out against Tertiary unconformity forming the negative flower structure. The final episode of tectonic disturbance resulted in the uplift of Jacobabad-Khairpur high. This event shifted the Lower Goru Formation into the deep marine settings that allowed deposition of the Upper Goru Formation which act as a seal for sands of the Lower Goru Formation.

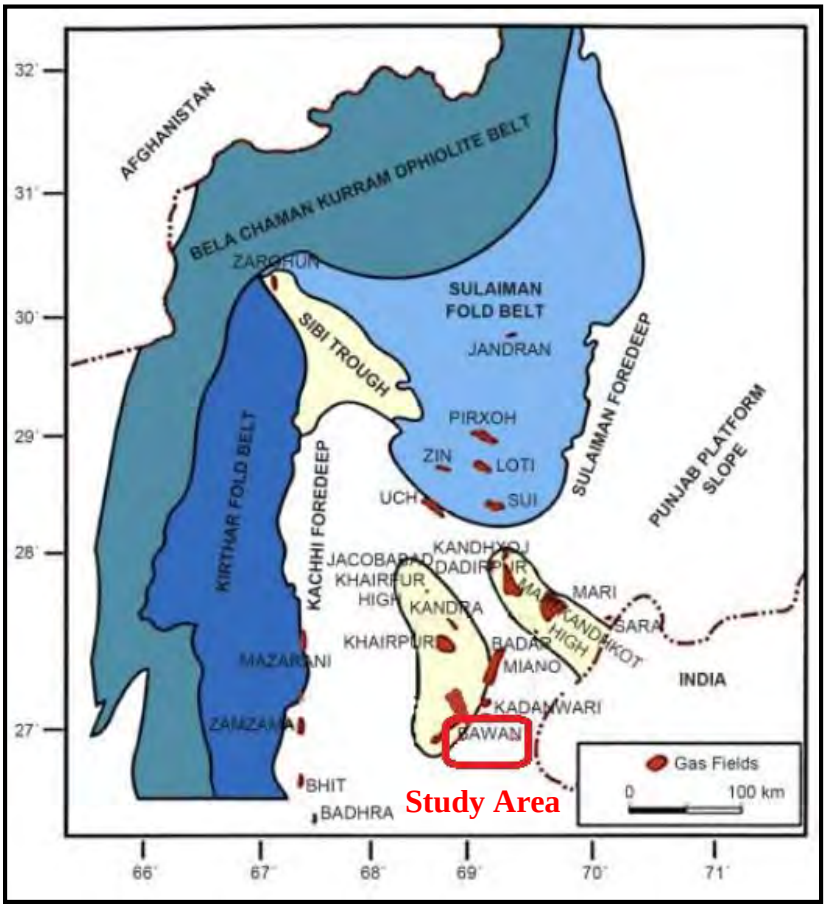


Figure 2.1: Tectonic map of Sawan Gas Field, Geological boundaries and major geological features are highlighted. Also includes the information about the neighboring hydrocarbon fields (Krois et al., 1998).

2.2.Petroleum System

Petroleum system can be defined as a system of source, reservoir, seal rocks along with the migration routes and trap geometry that is capable to yield the economically viable accumulation of hydrocarbons. Petroleum exploration and development data of Sawan area reveals that the suitable petroleum system is hosted by the Cretaceous strata.

2.2.1. Source Rock

The Sembar Formation of Early Cretaceous having TOC value of 1-4% is acting as a proven source rock in the Central and Lower Indus Basin (Aziz, Hussain, Ullah, Bhatti, & Ali, 2018). The predicted environment of deposition of the Sembar Formation is shallow marine environment. Lithology of the Sembar Formation consists mainly of black shale along with nodular argillaceous limestone, siltstone and sandstone. The characterization of organic matter of the Sembar Formation confirms the presence of Type III kerogen which is known for gas production (Wandrey, et al., 2004).

2.2.2. Reservoir Rock

The Lower Goru Formation is the proven reservoir of the Central and Lower Indus Basin. The lithology of the Lower Goru Formation is comprised of intercalated layers of sand and shale. Lower portion of the Lower Goru Formation is sand dominated and hosts the reservoir zones. It is further divided into D, C, B and A-sand Intervals as shown in Fig. 2.2. Studies shows that the Lower Goru Formation was deposited in lowstand shelf edge delta system (Munir, et al., 2011). The porosity value reservoir intervals of Lower Goru sands ranges between 6-25% and permeability value approaches to 2000 mD. Good porosity and permeability value make the Lower Goru a prolific reservoir of Lower Indus Basin.

2.2.3. Seal and Trap

The combination of structural and stratigraphic traps are identified in this area of Lower Indus Basin. Traps were formed during third tectonic episode which marks the phase of tectonic inversion. The termination of shale horizon against prograding wedges of C sand yields the stratigraphic traps. Wrench faulting occurred during the tectonic events also contributed in trap formation.

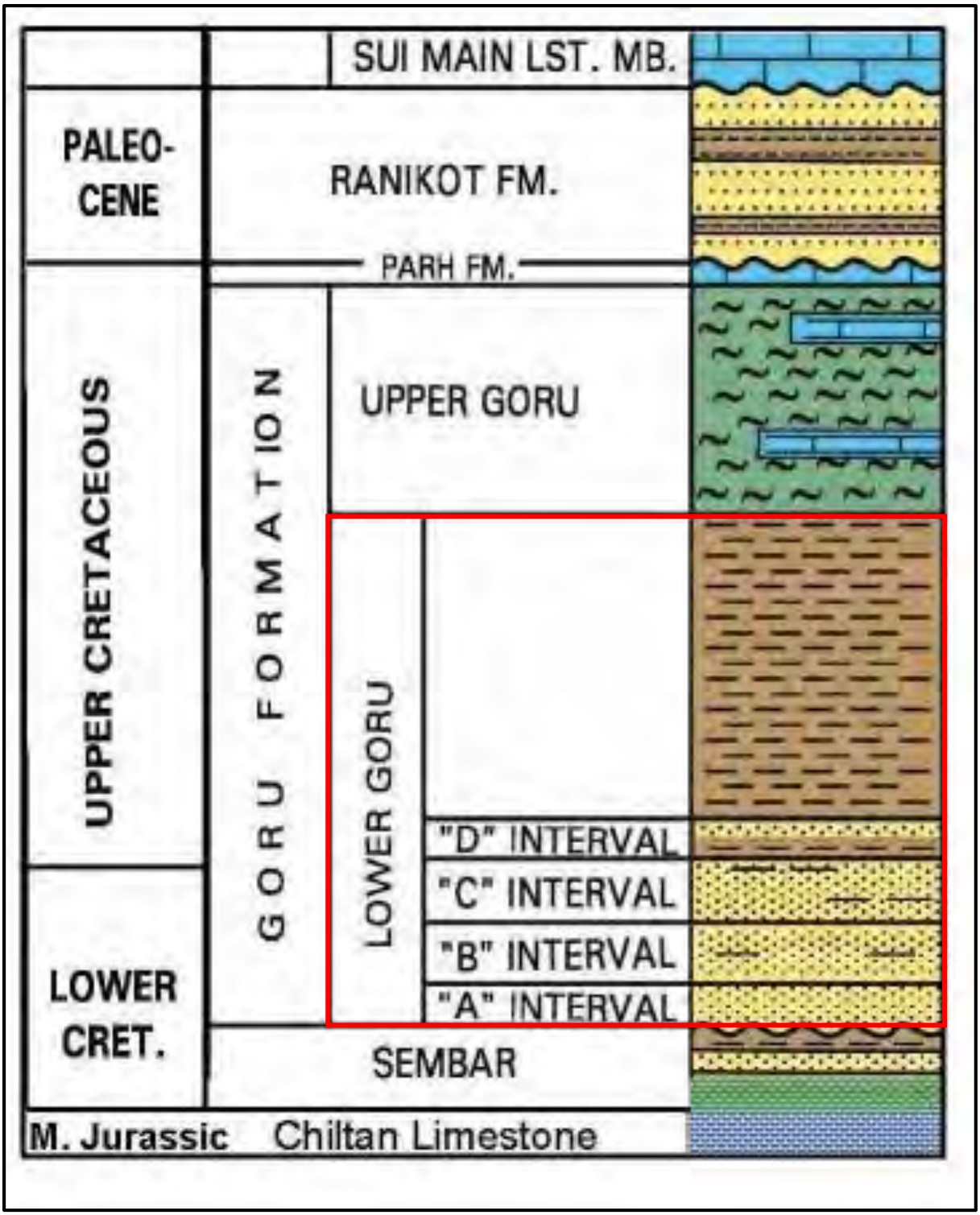


Figure 2.2: Stratigraphic Column of study area (Krois et al., 1998). Showing the stratigraphic succession from Middle Jurassic to Lower Eocene.

CHAPTER 03

PETROPHYSICAL ANALYSIS

3.1 Introduction:

The knowledge of subsurface rocks is very important for adopting the effective strategy of hydrocarbon exploration and production. The physical properties of rocks encountered in wellbore are obtained by lowering the wireline logging tools into the bore hole (Glover, 2000). The petrophysical analysis enables the explorationists to evaluate the reservoir potential of the rocks. So, petrophysical analysis is considered as a crucial step of reservoir characterization. In addition to reservoir characterization the petrophysical analysis has also its application in sequence stratigraphy. It can be used for sequence stratigraphic mapping of subsurface rocks by developing well to well correlation (Rider, 1996).

The petrophysical analysis gives the information about the key reservoir properties like volume of shale, porosity and fluid saturation. Thus, petrophysical analysis can serve as a tool for the quantification of the hydrocarbon reserves present in the area. The comparative analysis of petrophysical properties of subsurface rocks is used to mark the pay zones. The pay zones are the zones that can yield the maximum amount of hydrocarbons within the specific well (Sonnenberg and Selley, 2014).

The petrophysical analysis was carried out by using conventional log suite data of Sawan-02, Sawan-03 and Sawan-07 wells. The conventional log suite comprises of three tracks showing the log curves of different logging tools. The basic logs are the Gamma ray, SP, Caliper, Resistivity, neutron, density, PEF and sonic log. The log curves are displayed in the specific track as shown in table 3.1 (Glover, 2000).

Table 3.1 Log curves assigned to the tracks

Track	Track 1 (Lithology)	Track 2 (Resistivity)	Track 3 (Porosity)
Scale	Linear	Logarithmic	Linear
Curves	Gamma ray (GR) Spontaneous Potential (SP) Caliper (CALI) Bit size (BS)	Shallow resistivity (MSFL) Medium resistivity (LLS) Deep resistivity (LLD)	Density (RHOB) Neutron (NPHI) PEF Sonic (DT)

3.2 Methodology

The petrophysical analysis of Sawan-02, Sawan-03 and Sawan-07 wells was performed in a systematic manner. The major steps involved in the methodology are shown in Fig. 3.1.

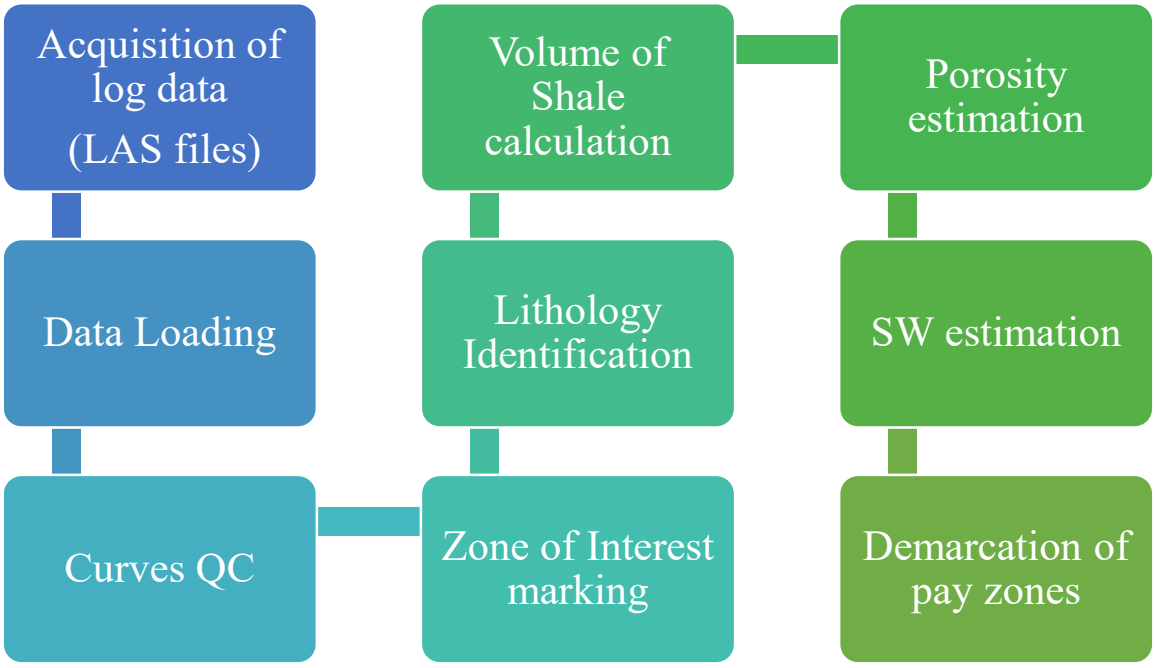


Figure 3.1: Generalized methodology used for petrophysical analysis (adopted from Schlumberger log interpretation manual 1989).

The well log data including Las files, well headers and formation tops was acquired from LMKR after getting approval from Directorate General of Petroleum Concession (DGPC). The well data is then loaded into the software followed by the quick look quality checking of log curves. After the quality checking the petrophysical interpretation was started. The generalized methodology

adopted to complete the petrophysical analysis of the Lower Goru Formation encountered in Sawan gas wells is described in the following sections.

3.2.1 Demarcation of zone of interest

The zone of interest is mainly the possible hydrocarbon bearing zone of the reservoir rock. The zone of interest is marked by quick look method on the basis following log responses of different logging tools.

- The low value of gamma ray tool is observed in the zone of interest. This indicates the low volume of shale and clean lithology. Low shale volume is the clue for the better permeability of the reservoir rock.
- The existence of separation among the shallow (MSFL), medium (LLS) and deep (LLD) resistivity log tool response is an indicator of hydrocarbon bearing zone. The separation among the resistivity curves is resulted due to the presence of resistivity contrast between drilling fluid and the formation fluid. The saline water-based mud is usually used as drilling fluid which has low resistivity. In contrast to this the hydrocarbons possess high resistivity. So, MSFL shows low resistivity and LLD shows high resistivity thus marking the hydrocarbon zone resistivity profile (Darling, 2005).
- Neutron (NPHI), Density (RHOB) crossover is a key indicator of hydrocarbon bearing zone. The crossover is formed due to the low values of neutron porosity (NPHI) and density log (RHOB) (Senosy et al., 2020). Because neutron porosity tool is sensitive to hydrogen index so, presence of hydrocarbons result in low porosity reading of neutron porosity tool.

3.2.2 Lithological identification

Identification of lithology of subsurface geological Formations is an important step of petrophysical interpretation. The response of different logging tools including Gamma Ray Log, Caliper Log, SP Log, Resistivity Log, Neutron and Density log are collectively used to lithological identification of the formations penetrated by the wellbore. The Gamma Ray Log gives the information about shaliness of the formation. Because the radioactive elements (Potassium, Thorium and Uranium) are usually present along with the clay (Rider, 1996). The response of caliper log gives the information about the compactness of the rocks. On gauged caliper log shows the compact and consolidated formation. In contrast to this the out gauged caliper log shows the loose subsurface rocks (Glover, 2000). The resistivity and SP log reflects the information about

the rock lithology along with its permeability. The shale is less resistive as compared to the clean formation. The overlapping of shallow, medium and deep resistivity curves shows the impermeable formation. The impermeability of rock is further confirmed by the uniform response of SP log. Density log directly gives the information about the density of the rock encountered. The density of some minerals is shown in table 3.2.

Table 3.2: Density of common lithologies (Glover, 2000).

S no	Lithology	Density
1	Quartz	2.654
2	Calcite	2.710
3	Dolomite	2.870
4	Halite	2.165

The NPHI/ RHOB overlay also gives the important information about the lithology of the rocks. The overlapping curves of NPHI and RHOB in overlay indicates the presence of limestone. In contrast to this the increase in RHOB and NPHI value is observed when shale is encountered (Hancock, 1992).

3.2.3 Estimation of Volume of Shale

Volume of shale is the quantification of shale content of the reservoir formation. Usually the reservoir rock formation is present in the form of alternating layers of shale and clean lithology (Glover, 2000). The Lower Goru Formation is also comprised of alternating lithological layers of sandstone and shale. GR log is mainly used for the estimation of shale volume of the reservoir formation. The simple equation is used for the computation of shale volume from GR log curve which is shown in Eq. 4.1 (Tiab and Donaldson, 2004).

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

Where GRlog is the value of GR curve at specific depth, GRmin is the minimum value of gamma ray log within the reservoir formation. The calculation of volume of shale enables the explorationists to apply the cut offs for the estimation of total clean rock volume by separating the shale intervals from the clean lithology. In addition to gamma ray log the shale volume can also be calculated by using the spontaneous potential log, density log and neutron log.

3.2.4 Porosity estimation

Reservoir rock is composed of four constituent which are grains, pores, matrix and cement (North, 1985). Pores within the reservoir rocks allows the storage of hydrocarbons to make a viable accumulation. Porosity can be defined as a percentage ratio of total pore volume of a rock to the bulk volume of the rock. Porosity of rock is mainly controlled by the grain size, sorting, packing and cementation of the rock (Selley, 2000). Estimation of porosity is considered as an important step of petrophysical interpretation because, it determines the reservoir character of the rock. Porosity can be determined by using the Density, Neutron and Sonic log data. Because of data limitations and good result of density log only the density and neutron porosity are estimated.

3.2.5 Density porosity

Density log tool is used to measure the bulk density of the rock formation encountered in the wellbore. The density value obtained by density log also includes the density value of matrix and formation fluid (Asquith and Gibson, 1982). By using the density values of formation fluid and rock matrix the density porosity can be calculated by equation shown in Eq. 4.2 (Asquith and Gibson, 1982).

$$\emptyset_{den} = \frac{\rho_{ma} - \rho_b}{\rho_{ma} - \rho_f} \qquad (4.2)$$

Where ρma is the density of matrix, ρb is the density log reading at specific depth and ρf is the density of formation fluid. The density values for common formation fluids is given in table 3.3 (Glover, 2000). The table also shows the density value for the sandstone because, the dominant lithology in the Lower Goru Formation is sandstone.

Table 3.3: Densities of common formation fluids (Glover, 2000)

S no	Rock Component	Density
1	Sandstone	2.65
2	Fresh water	1
3	Brine	1.1
4	Oil	0.8

3.2.6 Neutron porosity

Neutron log is mainly sensitive to the hydrogen present within the rock formations encountered in the wellbore (Asquith and Gibson, 1982). The neutron log tool works by bombarding the formation with fast neutrons. These neutrons are then received at receiver of logging tool after interaction with formation. Neutrons lose their energy during interaction with rocks. The maximum energy is lost when neutron collides with hydrogen atom because the mass of neutron is roughly equal to the mass of hydrogen atom. Inverse relationship exists among the number of received neutrons and porosity. If gas is present in reservoir rock then neutron porosity log underestimates the porosity value because gas has higher hydrogen index which greatly reduces the neutron count rate.

3.2.7 Total porosity

Total porosity can also be termed as average porosity. It is the average of all estimated porosities including density porosity, neutron porosity and sonic porosity. The formula used for estimation of average porosity is given in Eq. 4.3 (Rider, 1996).

$$\emptyset_{Total} = \sqrt{\frac{\emptyset_{density} + \emptyset_{neutron}}{2}} \quad (4.3)$$

Whereas ΦTotal is the total porosity, Φdensity is for density porosity and Φneutron is the neutron porosity. The total porosity quantifies all the porosity of the reservoir rock.

3.2.8 Effective porosity

Effective porosity can be defined as the volume of interconnected pore spaces that has capability to yield economical amounts of hydrocarbon. Effective porosity usually ranges from 40 to 75% of total porosity (North, 1985). Effective porosity can be estimated by excluding the volume of shale from the total porosity. Mathematically, it can be calculated by Eq. 4.4 (Rider, 1996).

$$\varphi_{effective} = \varphi_{Total}(1 - V_{shale}) \quad (4.4)$$

Where, $\Phi_{effective}$ is the effective Porosity, ΦTotal is for total porosity and V_{shale} shows the shale volume.

3.2.9 Estimation of fluid saturation

Identification and quantification of fluid present in the formation penetrated by wellbore is an important step of petrophysical analysis. Resistivity log plays a key role in the identification of formation fluid and also helps in the calculation of fluid saturation (Glover, 2000). Reservoir rock usually contains saline water or hydrocarbon. The resistivity contrast present among saline water and hydrocarbon aids in fluid recognition because, the resistivity value of saline water is very low. On the other hand, the hydrocarbons show very high values of resistivity. The fluid saturation is estimated by using different numerical and analytical methods.

3.2.10 Calculation of formation water resistivity (Rw)

The resistivity of Formation water (Rw) is required for the estimation of water saturation. The Rw can be estimated by using the geothermal gradient, formation temperature, static spontaneous potential and salinity of water (Amigun, et al., 2012). There are two widely used methods for the numerical estimation of resistivity of water. These methods are the SP method and Pickett cross plot method (Greengold, 1986). Pickett cross plot method was adopted for performing the petrophysical analysis of the Lower Goru Formation encountered in Sawan-02, Sawan-03 and Sawan-07 wells.

Pickett plot is quite useful technique for the estimation of Rw. Pickett plot was drawn by plotting deep resistivity on x axis and total porosity of reservoir zone on y axis. The logarithmic scale is used for both axes. In this manner Pickett plot illustrates the relationship among the deep resistivity, porosity, cementation factor and water saturation (El-Khadragy et al., 2014). After plotting the data points the line is drawn on the left most side of plot which shows the 100% water saturation. The data points present above this line depicts the water saturation value less than 100%. Fig. 3.2 is showing the Pickett plot for Sawan-02 well and the estimated Rw value is 0.004. Pickett plot for Sawan-03 well is shown in Fig. 3.3. Reservoir zone lies in the depth range of 3400m – 3600m. The estimated value of Rw is 0.004 which same as that of Sawan-02 well. Fig. 3.4 is illustrating the Pickett plot for Sawan-07 well. Reservoir zone lies within the interval of 3200m – 3400m. The estimated value for Rw in Sawan-07 well is 0.008.

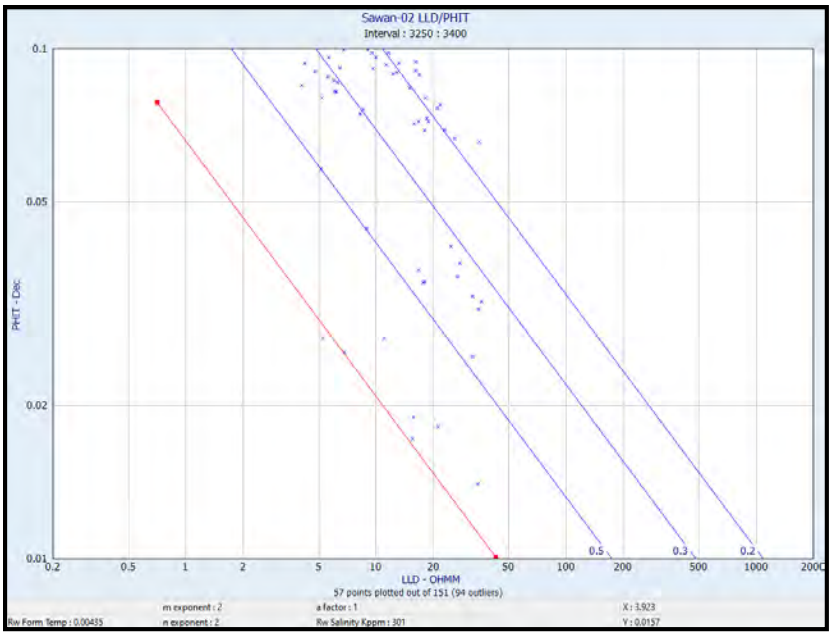


Figure 3.2 Pickett plot of Sawan-02 well. The red colored (left most line) is showing the maximum water saturation while remaining lines account for progressive decrease in water saturation, value of each line is shown at the lower ends of line.

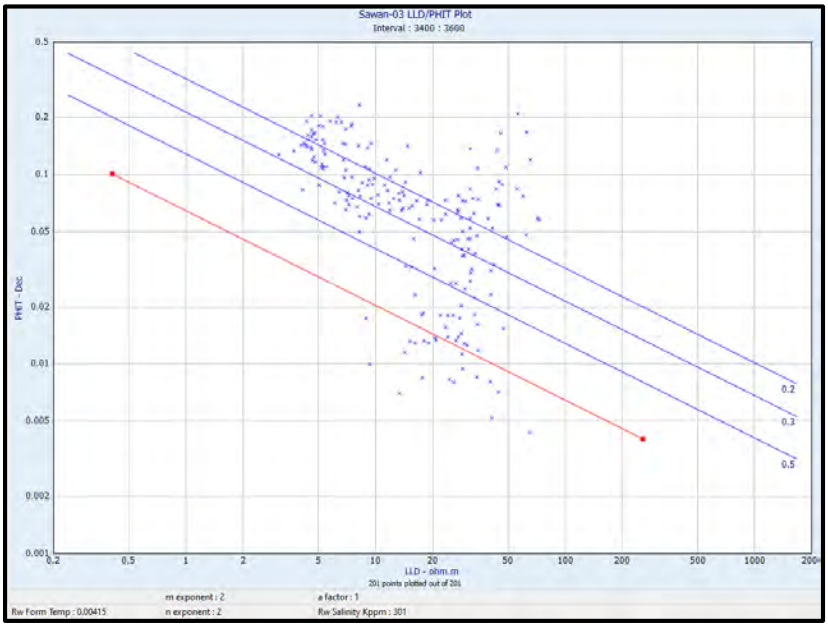


Figure 3.3: Pickett Plot of Sawan-03. The red colored (left most line) is showing the maximum water saturation while remaining lines account for progressive decrease in water saturation, value of each line is shown at the lower ends of line.

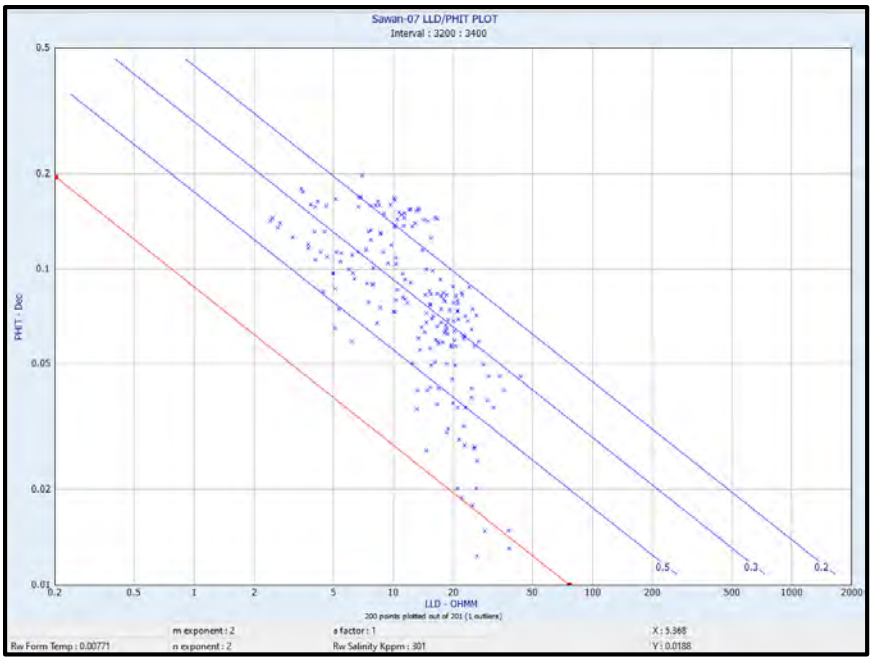


Figure 3.4: Pickett plot of Sawan-07 well. The red colored (left most line) is showing the maximum water saturation while remaining lines account for progressive decrease in water saturation, value of each line is shown at the lower ends of line.

3.2.11 Calculation of water saturation

Estimation of water saturation is an important step of petrophysical analysis (Tiab and Donaldson, 2004). Saturation of water can be defined as the quantification of water present in pore spaces of the reservoir rock. There are different empirical models adopted for the estimation of water saturation. Some well-known water saturation models are the Archie's equation, Dual water, Indonesian equation, Simandoux and modified Simandoux equations. Modified Simandoux equation was used for performing the petrophysical analysis of the Lower Goru Formation using the well data of Sawan-02, Sawan-03 and Sawan-07. The lithology of the Lower Goru Formation is shaly sand so, Archie equation can overestimate the water saturation results (Sam-Marcus et al., 2018). Mathematical form of Modified Simandoux equation is given in Eq. 4.5 (Simandoux, 1963).

$$\frac{1}{Rt} = \frac{\varphi^{m} * Sw^{n}}{\alpha * Rw} + \frac{Vshl * Sw}{Rshl} \tag{4.5}$$

Whereas, Rt shows the value of true formation resistivity (LLD value), Φ is for effective porosity, Sw is saturation of water, m is cementation factor, a is the tortuosity factor, n is the saturation

exponent, Rw is the resistivity of water, Vshl is volume of shale and Rshl shows the resistivity value adjacent to the shale interval of the formation.

3.2.12 Calculation of hydrocarbon saturation

Hydrocarbon saturation can be defined as the pore volume of reservoir rock occupied by the hydrocarbon. Hydrocarbon saturation can be estimated by using formula given in Eq.4.6 (Obeida et al., 2007).

$$S_{HC} = 1 - S_W \tag{4.6}$$

Whereas, S_{HC} shows the saturation of hydrocarbon and Sw is for saturation of water.

3.2.13 Demarcation of pay zones

It is the final and important step of petrophysical interpretation. The pay zones are marked by applying the suitable cut off values of different parameters. Usually the cut offs are applied for height (thickness), volume of shale, effective porosity and water saturation. The cut off height determines the suitable thickness of reservoir which can yield economical production of hydrocarbons. The Vshale cut off separates the clean lithology from shale dominated intervals of the formation. Phi cut off is used to separate reservoir intervals having good porosity from tight and low porosity intervals. Saturation of water cut off is used to differentiate water saturated zones from hydrocarbon saturated zones. The net pay is the net thickness of reservoir rock that is resulted after the application of cut offs and have capability to yield economic production of hydrocarbons.

3.3 Petrophysical interpretation of Sawan-02 well

Petrophysical analysis of Sawan-02 was completed by using the methodology explained in earlier sections of this chapter. The raw log curves and Formation Tops were displayed in the software. The zone interest was identified by using quicklook method. The depth of proposed reservoir zone ranges from 3277m – 3380m. The zone of interest lies in the C interval of the Lower Goru Formation as shown in Fig. 3.5. This zone clearly shows the low values of GR, separation among the shallow, medium and deep resistivity logs. The noticeable neutron density crossover is also present which indicates the presence of gas.

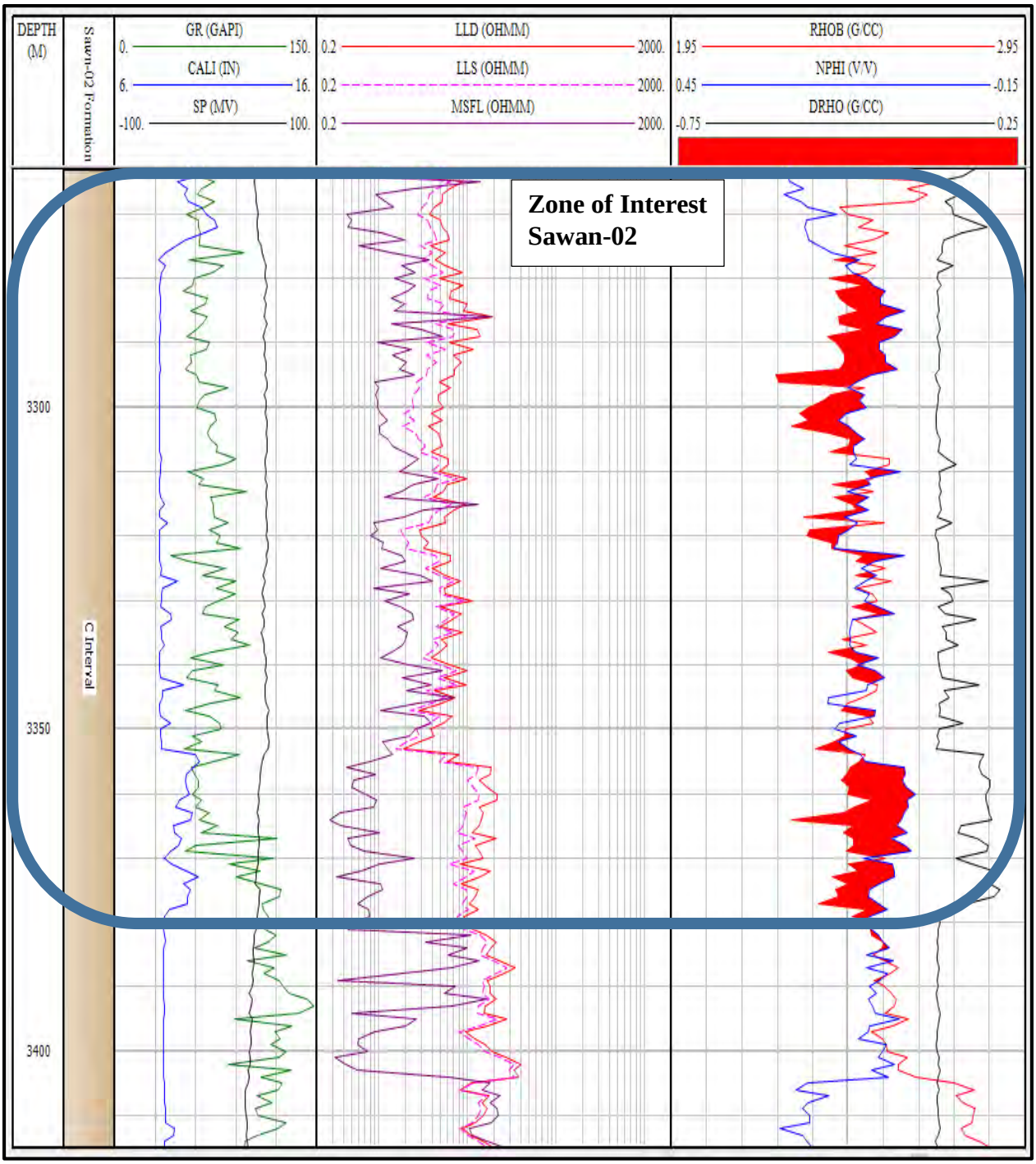


Figure 3.5: Demarcation of Zone of Interest in Sawan-02. The shaded portion is showing the neutron density crossover formed due to low values of both logs. Separation between LLS and LLD is also visible along with low GR values.

The demarcation of zone of interest is followed by the lithological identification of the formation hosting the zone of interest. For lithological identification neutron density crossplot is used. The

NPHI is plotted on x axis and RHOB is on y axis (Clavier & Rust, 1976). The neutron and density data points of C interval are plotted on neutron density crossplot as shown in Fig. 3.6. Schlumberger lithological overlay lines are used for lithological identification. Most of data points lies are located within the range of sandstone. The point present above the sandstone line indicates the underestimated density value because of gas presence. The data points present below the limestone line shows the presence of shale. The presence of shale is also verified by high GR values. Some data points with low GR are also located below limestone line. These points are showing the higher density values mainly due to the variation in mineralogy.

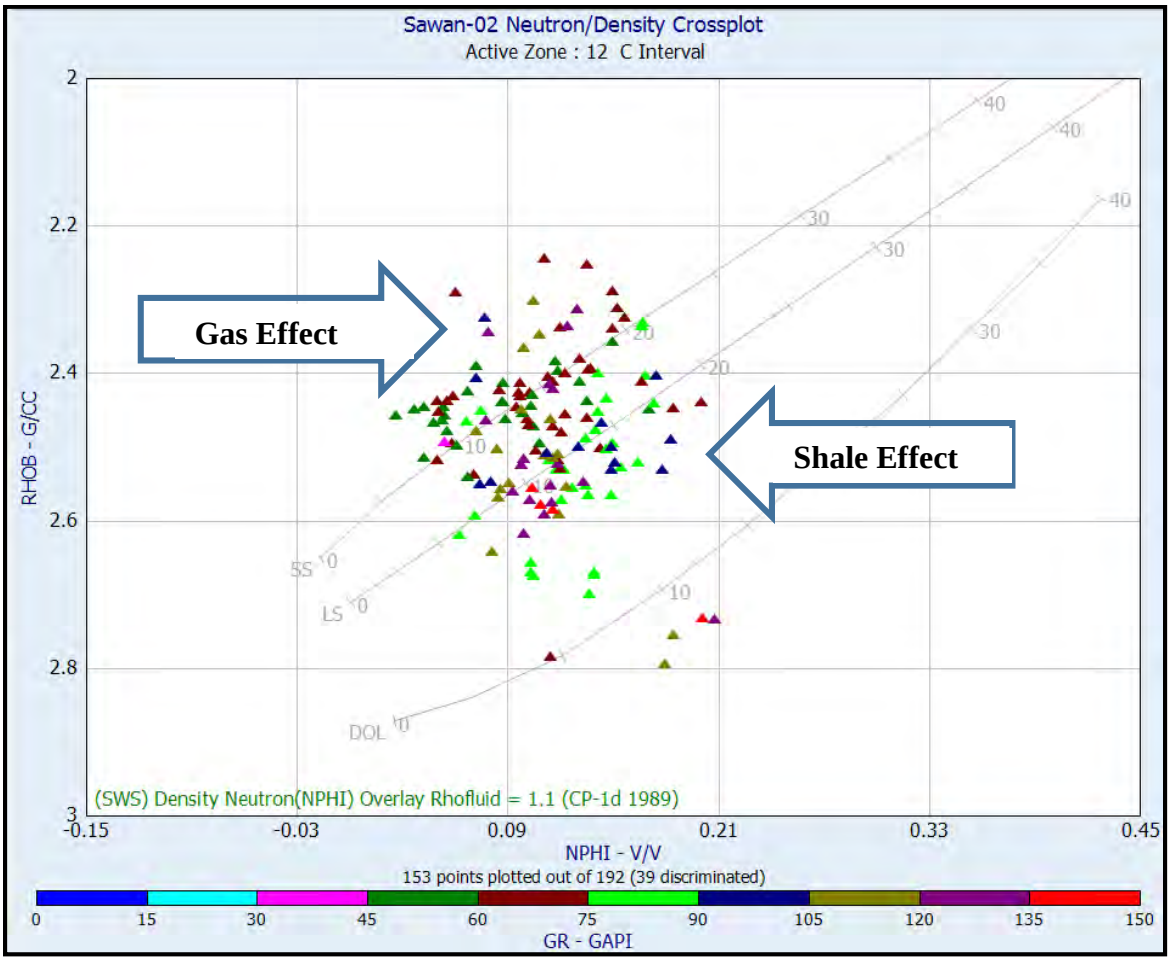


Figure 3.6: Neutron density crossplot for Lithological identification of Sawan-02. Most of data points are present in the vicinity of Sandstone lithological overlay line.

In addition to neutron density crossplot the GR log also gives the information about the lithology of the formation. High GR values indicates the shale while the low GR values shows the presence

of clean lithology. Lithology curve is added adjacent to GR log which indicates that the Lower Goru Formation is composed of alternating bedding of sand and shale as shown in Fig. 3.7. It is evident from Fig. 3.7 that the zone of interest is mainly comprised of clean sandstone layers of the Lower Goru Formation.

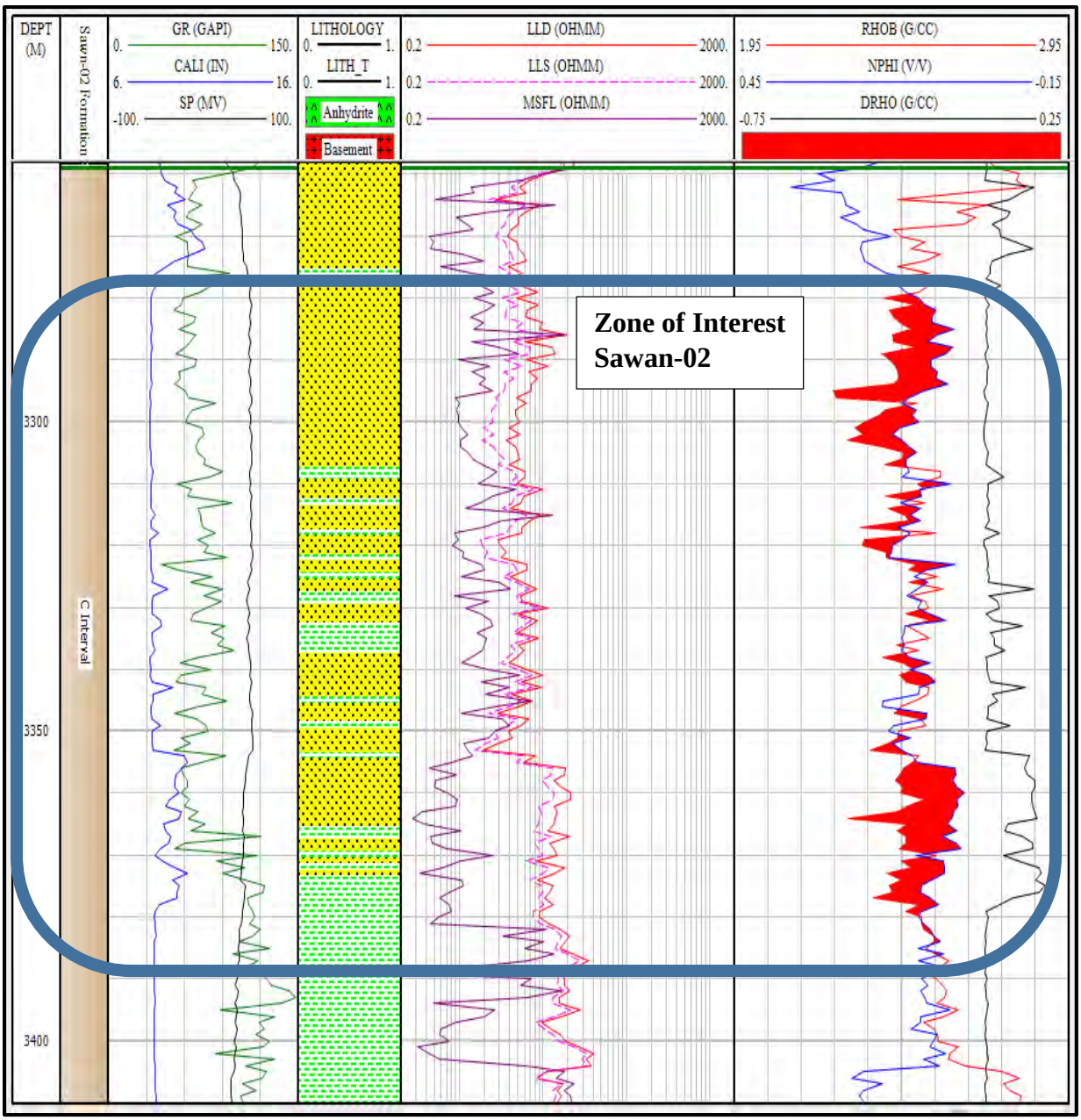


Figure 3.7: Lithology of Reservoir zone in Sawan-02. Lithological curve is differentiating between the sandstone and shale layers of the Lower Goru Formation. Sandstone layers are the major contributor in zone of interest.

The volume of shale is estimated and displayed in form of curve adjacent to the porosity track as shown in Fig. 3.8. The volume of shale is computed by using GR log readings. Higher values of shale volume are observed at depths showing higher value of GR log.

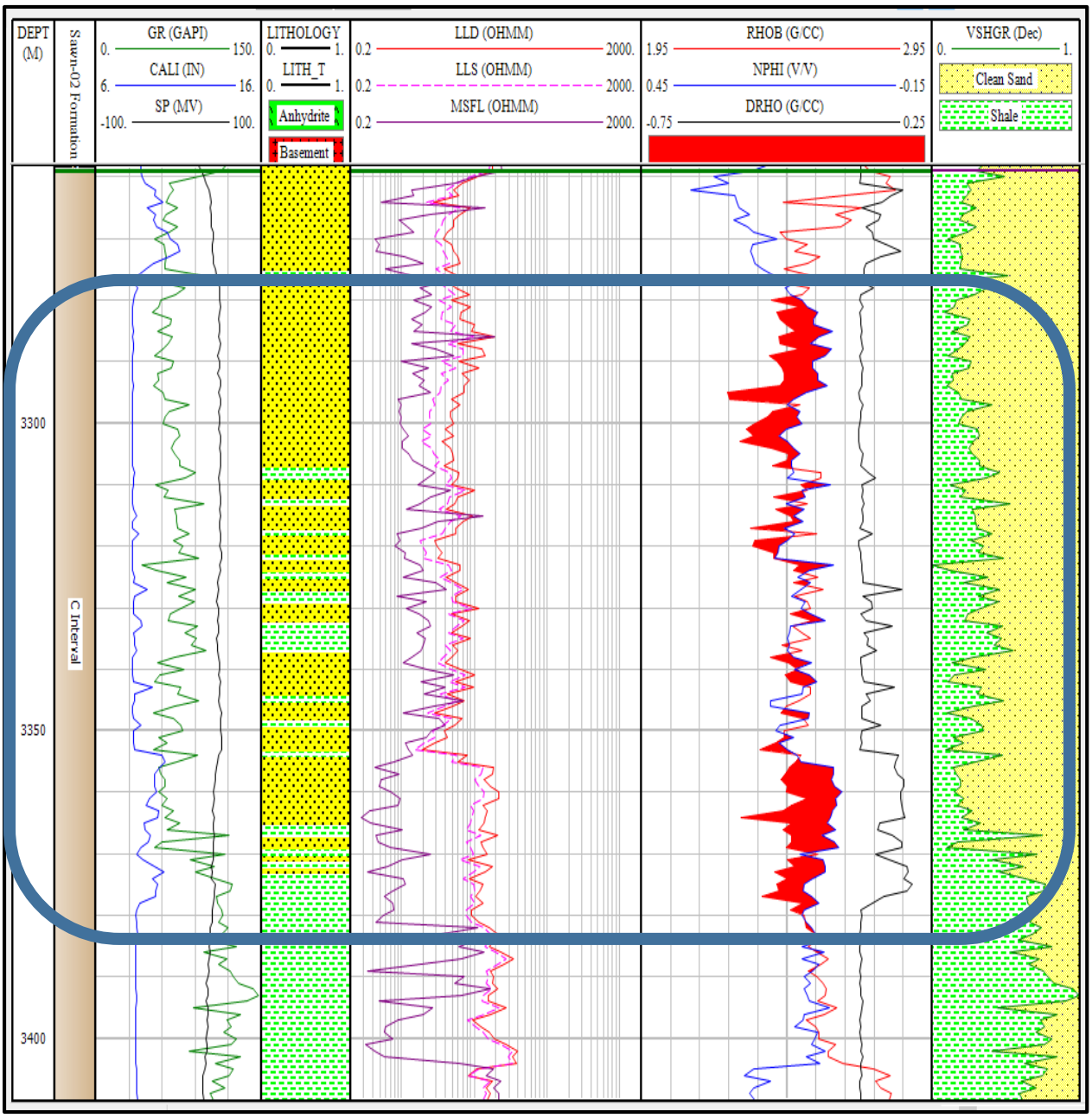


Figure 3.8: Estimated volume of shale in Sawan-02. The curve showing volume shale of C-interval of the Lower Goru Formation is displayed adjacent to lithology track. Low values of Vshale marks the reservoir zone.

After estimation of volume of shale average porosity is estimated by using the neutron and density logs. The effective porosity is than computed by using average porosity and volume of shale. The estimated effective porosity is shown in curve adjacent to Vshale track as shown in Fig. 3.9. Water saturation is a key petrophysical parameter that is estimated after porosity estimation. The water saturation is calculated by using modified Simandoux equation because, the Lower Goru Formation is consisted of sand and shale interval. Due to this the average values of water saturation are around 0.35 as shown in Fig. 3.9.

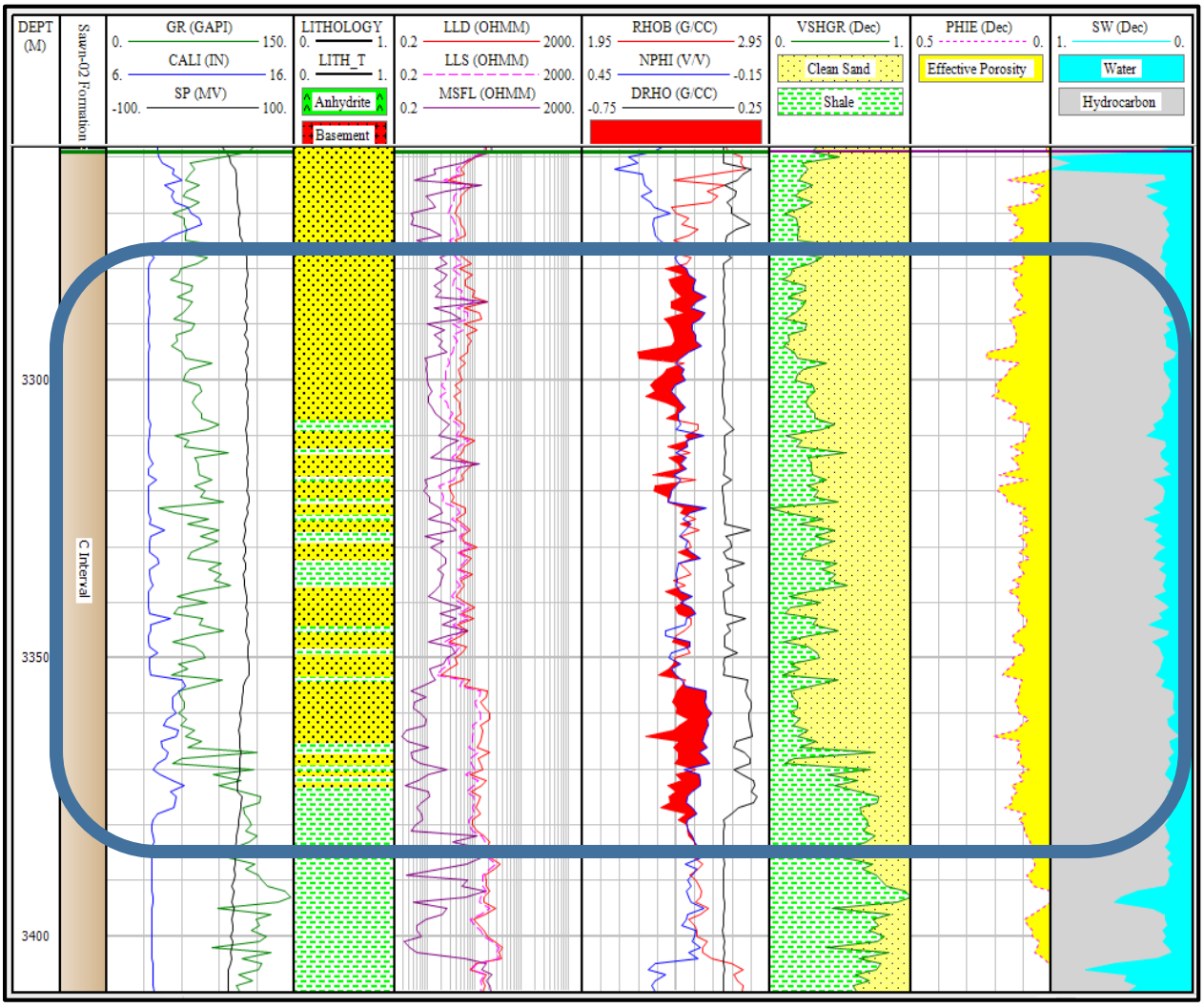


Figure 3.9: Interpreted log plot of Sawan-02. The curves showing the effective porosity, saturation of water and saturation of hydrocarbon are displayed in tracks next to the volume of shale track. Yellow color is indicating effective porosity, blue color is showing water saturation and grey is for hydrocarbon saturation.

The final step of petrophysical analysis is the demarcation pay zones. Suitable cut offs are estimated and applied to efficiently mark the pay zones.

After application of the cut offs the pay zones are marked as shown in Fig. 3.10. The pay zones are indicated by red color flag. Pay zones are the zones of reservoir rock that can yield economic production of hydrocarbons.

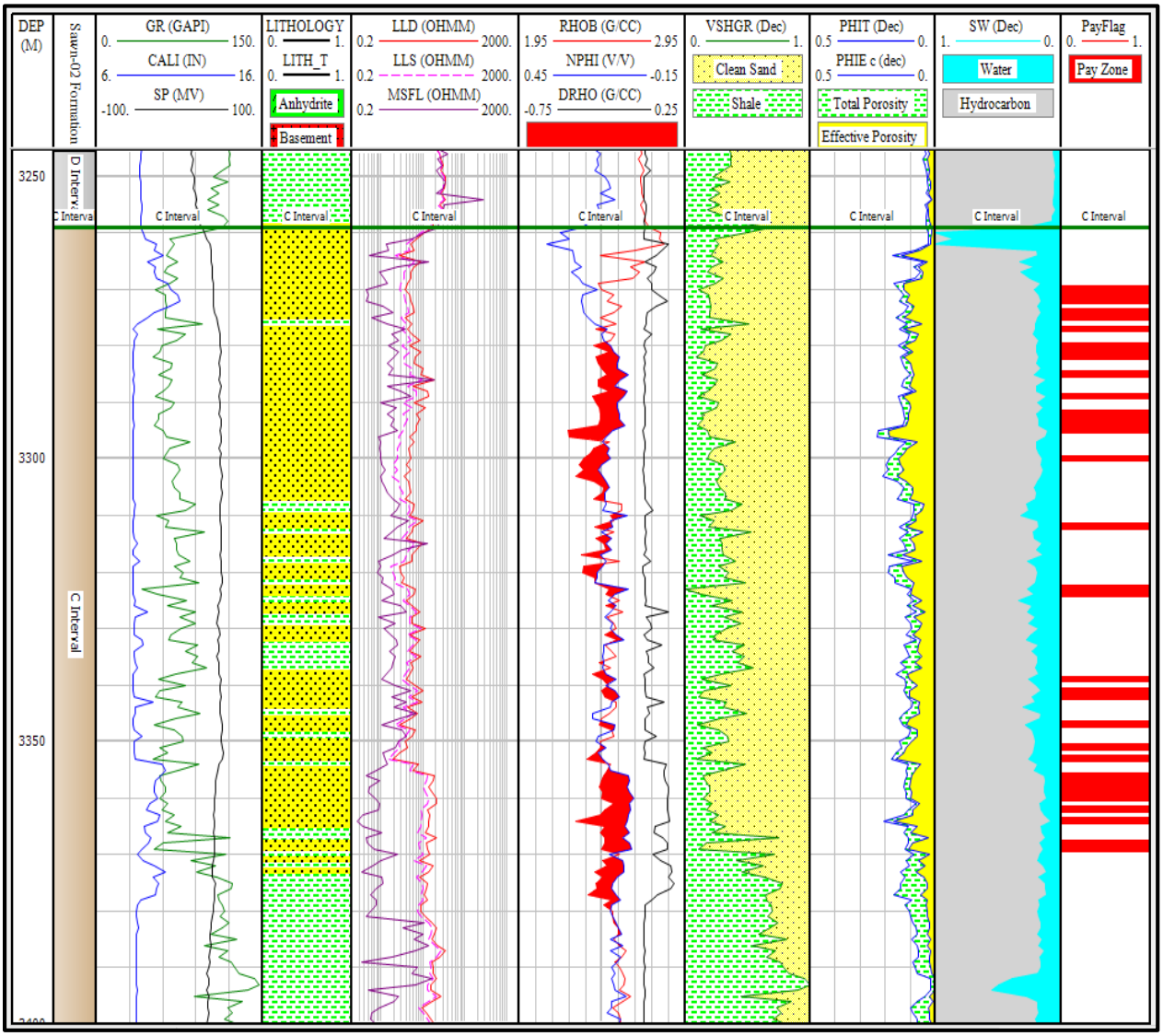


Figure 3.10: Interpreted log plot of Sawan-02 with pay zones. The zones fulfilling the cut off values are shown in red color which are identified as pay zones in C interval.

The summary of pay zone ratios and thicknesses give the information about net to gross ratio as shown in table 3.5.

Table 3.4: Pay zones summary of Sawan-02 including the averages of petrophysical parameters and the reservoir thickness value based on each petrophysical parameter.

Quantity	Value
Zone#	12
Zone Name	C Interval
Top	3259
Bottom	3450
Gross	191
Net	36
N/G	0.188
Av Phi	0.114
Av Sw	0.182
Av Vcl	0.142
Phi*H	4.1
PhiSo*H	3.35
Vcl*H	5.12

3.4 Petrophysical interpretation of Sawan-03 well

Petrophysical analysis of the Lower Goru Formation is also performed in Sawan-03 by using the same methodology. In Sawan-03 well two zones of interest were identified by using the quick look petrophysical indicators. The main zone of interest lies in C interval of the Lower Goru Formation. On the other hand, the second zone of interest is quite small and lies within the B interval of the Lower Goru Formation. The lithological identification was performed by neutron porosity and bulk density log crossplot as shown in Fig. 3.11. The major lithology found in the zone of interest is mainly sandstone along with some layers of shale. The Gas effect and shale effect is clearly visible in the neutron density lithology plot that results in plotting of data points outside the range of sandstone line. The GR log scale is also given at the bottom of lithology chart which shows that points having higher GR values are concentrated below the limestone line. Some data points

having higher GR values are also showing the lower density value this can be due to occurrence of radioactive minerals (mainly Uranium) within the clean lithology.

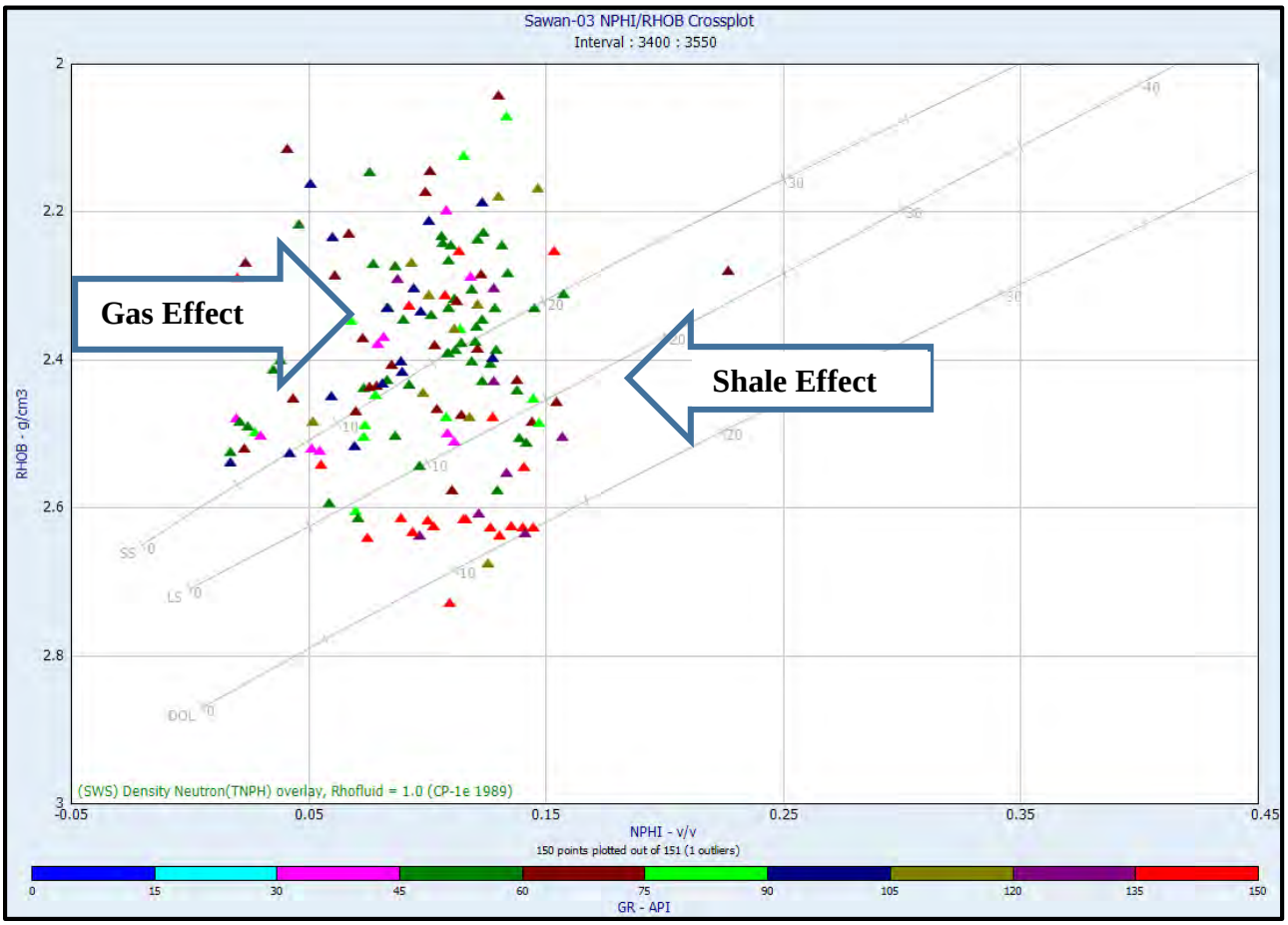


Figure 3.11: Neutron density crossplot for lithological identification in Sawan-03. Most of data points are present in the vicinity of Sandstone lithological overlay line. Gas and shale anomaly is marked by arrows.

The first zone of interest lies in the depth range of 3400-3540m while second zone ranges from 3596m to 3609m. The interpreted log plot of both zones of interest are shown in Fig. 3.12. Interpreted log is showing the curves of lithology, volume of shale, total porosity, effective porosity and saturation of water. These parameters greatly aid in efficient demarcation of pay zones which determines the success or failure of hydrocarbon well in terms of production.

The suitable cut offs are estimated and applied for demarcation of pay zones. The Vshale cut off is kept relatively higher in Sawan-03 well as compared to Sawan-02 because increase in shale content is observed in C interval of the Lower Goru Formation in Sawan-03.

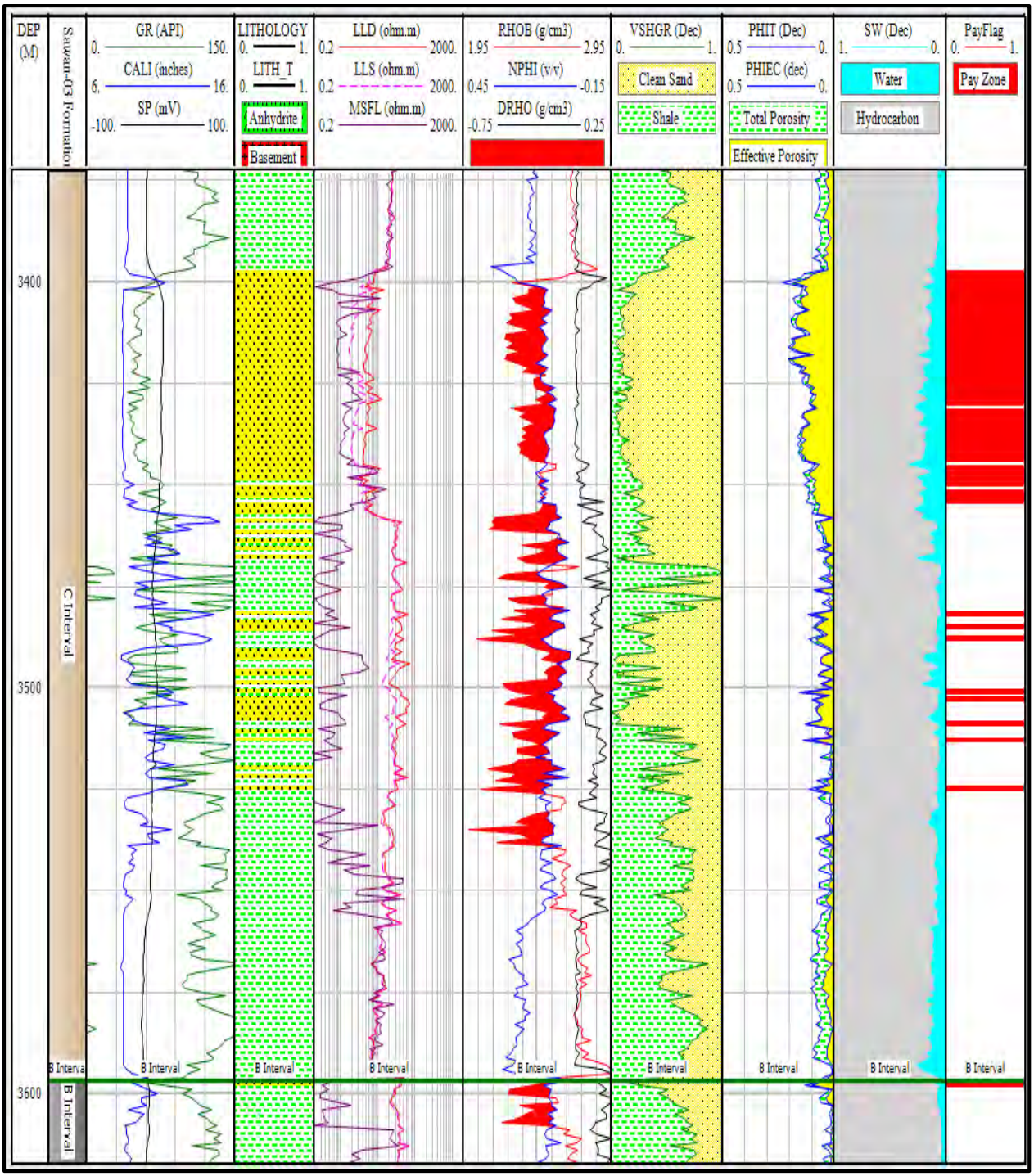


Figure 3.12: Interpreted log plot of Sawan-03. The curves showing the petrophysical properties including volume of shale, average porosity, effective porosity, water saturation and hydrocarbon saturation are given in tracks next to the lithology track. Pay zones are showed in right most track.

The summary of pay zones along with petrophysical averages and net to gross ratio are shown in table 3.7.

Table 3.5: Pay Zones summary of Sawan-03 including the averages of petrophysical parameters and the reservoir thickness value based on each petrophysical parameter are given.

Zone#	**13**	**14**
Zone Name	C Interval	B Interval
Top	3365.6	3597
Bottom	3597	3666
Gross	231.4	69
Net	62	1.00
N/G	0.268	0.014
Av Phi	0.125	0.097
Av Sw	0.158	0.048
Av Vcl	0.13	0.26
Phi*H	7.76	0.1
PhiSo*H	6.53	0.09

3.5 Petrophysical interpretation of Sawan-07 well

Petrophysical analysis of Sawan-07 well is carried out by using the same methodology that was adopted for the petrophysical interpretation of Sawan-02 and Sawan-03 wells. The Lower Goru Formation was also encountered in Sawan-07 well. After demarcation of zone of interest, the lithological identification of reservoir zone was done by using the neutron density crossplot with Schlumberger's lithological overlay lines as shown in Fig. 3.13. The major interpreted lithology of zone is the sandstone along with relatively small quantities of shale. Gas effect and shale effect are clearly shown over the crossplot which are further verified by the GR color scale given at the bottom of the crossplot.

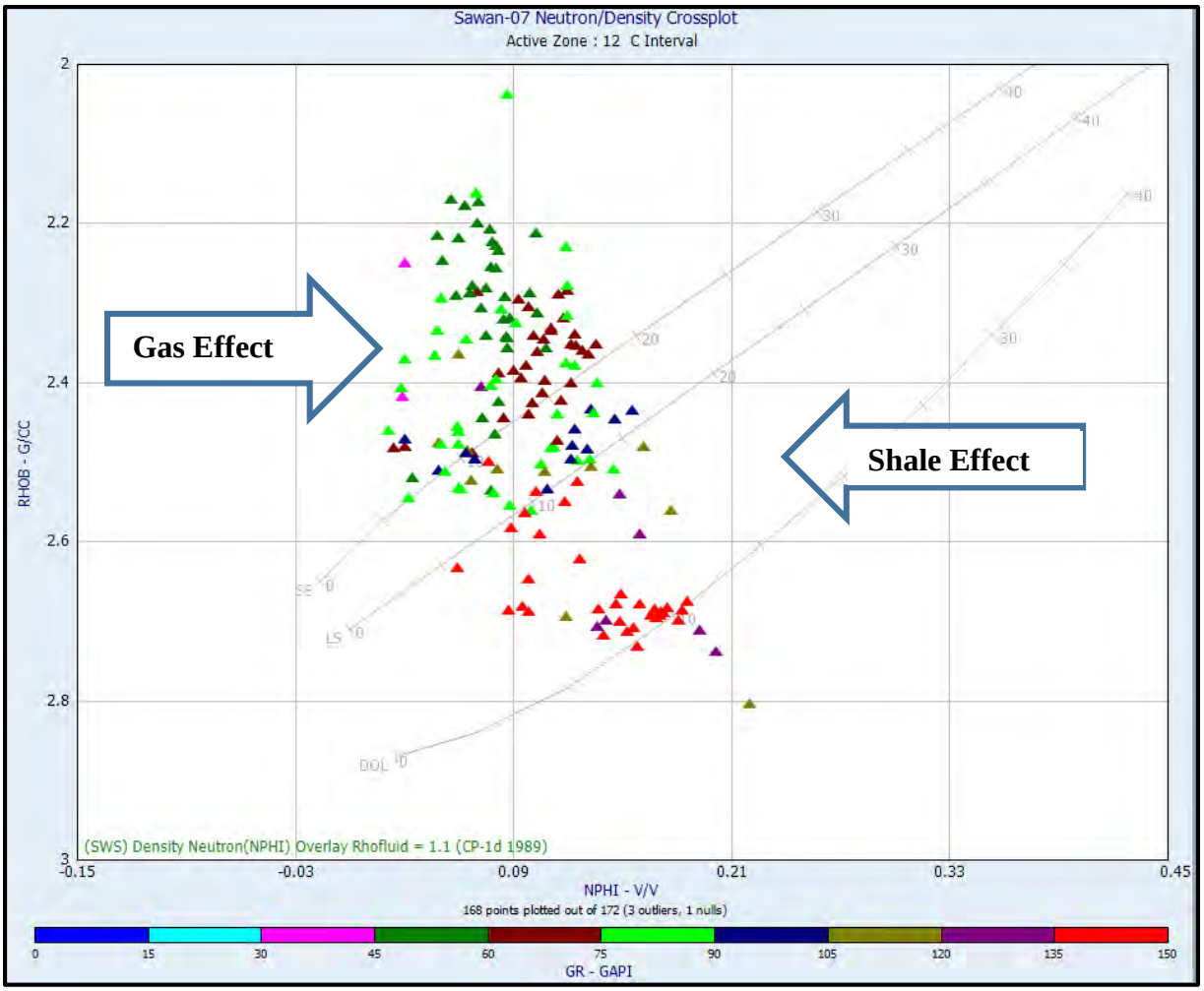


Figure 3.13: Neutron density crossplot for lithological identification of Sawan-07. Most of data points are present in the vicinity of Sandstone lithological overlay line. The gas effect and shale effect are also marked.

Lithological identification of reservoir formation is followed by the petrophysical interpretation. During this step key petrophysical properties like volume of shale, total porosity, effective porosity and saturation of water of the reservoir were estimated. In Sawan-07 well the zone of interest lies in the depth range of 3267m-3360m. The interpreted log plot depicting the reservoir zone of Sawan-07 is shown in Fig. 3.14. The reservoir zone of Sawan-07 is showing the best reservoir characters among all three wells (Sawan-02, Sawan-03 and Sawan-07) because, it has relatively low values of volume of shale, higher effective porosity and lower water saturation.

In light of results obtained from petrophysical analysis the suitable cut off values were estimated. The cut off value adopted for marking the pay zone in Sawan-07 is relatively low because, the

reservoir zone encountered in Sawan-07 showed quite low water saturations. So, the zones having relatively higher water saturation are filtered out as a result of SW cut off. The cut off values are applied to mark the pay zones shown by pay flag in Fig. 3.14.

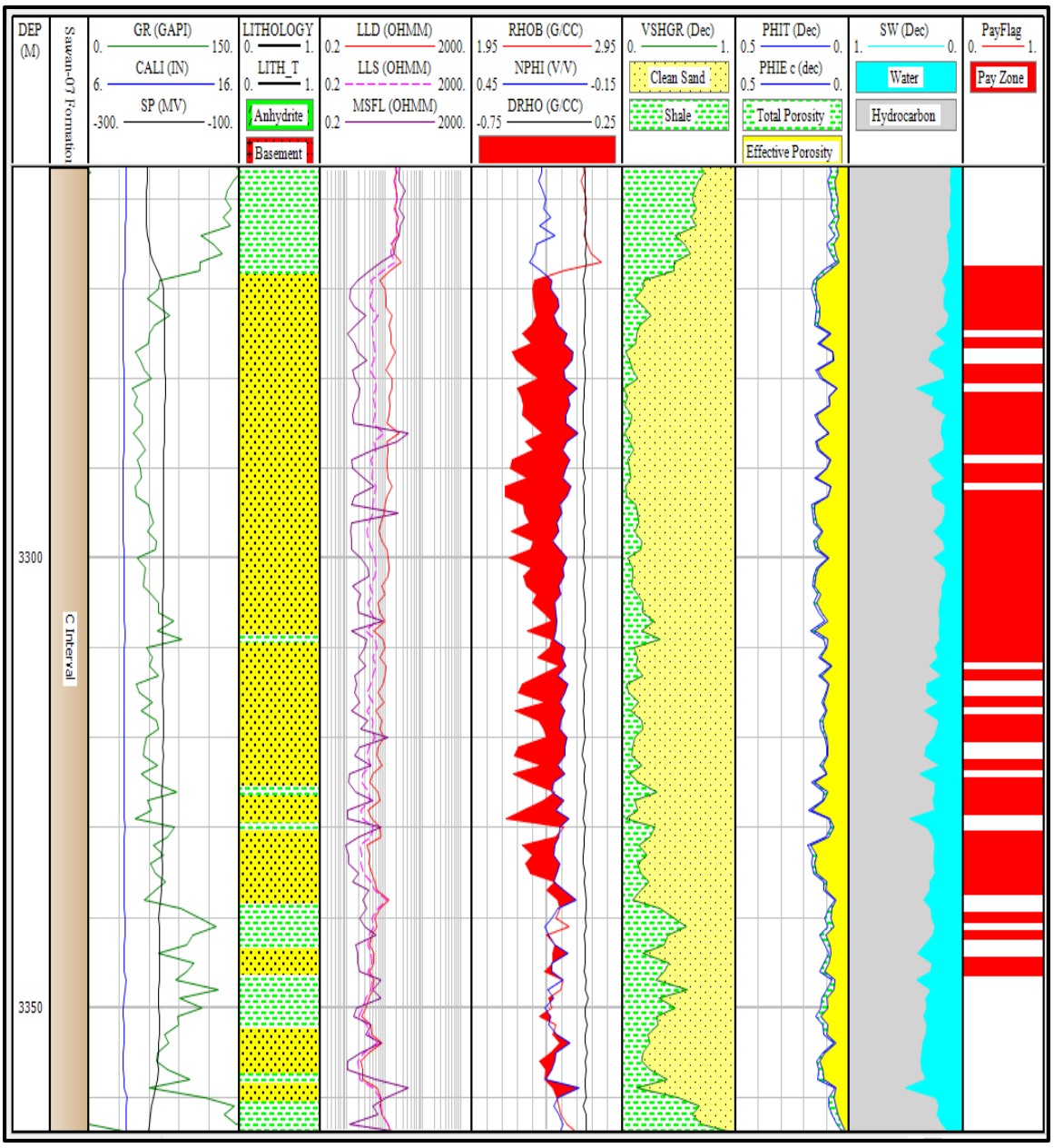


Figure 3.14: Interpreted log plot of Sawan-07. The curves showing the petrophysical properties including volume of shale, average porosity, effective porosity, water saturation and hydrocarbon saturation are given in tracks next to the lithology track. Pay zones are showed in right most track.

The averages of petrophysical properties along with cut off thicknesses and net to gross ratio of Sawan-07 well are shown in table 3.9.

Table 3.6: Pay zone summary of Sawan-07 including the averages of petrophysical parameters and the reservoir thickness value based on each petrophysical parameter are given.

Zone#	**12**
Zone Name	C Interval
Top	3239.39
Bottom	3432
Gross	172.11
Net	67.00
N/G	0.389
Av Phi	0.122
Av Sw	0.196
Av Vcl	0.096
Phi*H	8.15
PhiSo*H	6.55

The petrophysical analysis of the Lower Goru Formation encountered in Sawan-02, Sawan-03 and Sawan-07 reveals that the C interval of the Lower Goru Formation possesses the good reservoir character. It can serve as a prolific petroleum reservoir. The comparison of pay zone thicknesses shows that the pay zones vary from well to well but the maximum pay zone thickness is observed in Sawan-07 which is 67m. In contrast to this the pay zone thickness in Sawan-03 is 63 m and in Sawan-02 is 36 m respectively.

CHAPTER 04

K-MEANS CLUSTER ANALYSIS

4.1 Introduction:

The lithology and fluid content identification of reservoir rock is a very important step of reservoir characterization. Traditionally, the reservoir characterization is performed by petrophysicist using the numerical models of petrophysics. The knowledge and experience of petrophysicist play the key role in quality of reservoir evaluation results (Ali and Sheng-Chang, 2020). Advancement in computational technology also modified the field of petrophysics by introducing machine learning tools and techniques which have ability to refine the process of reservoir characterization. Cluster analysis is a multivariate machine learning technique which is used to classify the large data sets into groups and sub groups on the basis of similarities (Cornish, 2007). So, the data points present in the cluster are more identical to each other as compared to the data points present in other clusters.

Cluster analysis can be performed on well log data for the identification of electrofacies. Electrofacies can be defined as a unique set of log signatures which reflect the specific lithology and fluid type of the rock investigated by logging tool (Euzen and Power, 2012). The lithology of the Lower Goru Formation is quite complex. K-means clustering can serve as a tool to differentiate the various lithological packages of the Lower Goru Formation by using electrofacies analysis. The well log data of Sawan-02, Sawan-03 and Sawan-07 is used for performing the K-means clustering for facies identification of Lower Goru Formation.

4.2 Clustering methods

Cluster defines the group of data points. There are different methods of clustering some of which are listed as follows (Estivill-Castro, 2002).

4.2.1 Hierarchical clustering

Hierarchical clustering is based on connectivity cluster model. It classifies data points on the basis of distance present between the data points. The most commonly used methods for hierarchical clustering are the nearest neighbor method, the furthest neighbor method and the average distance method (Kattan, et al., 2018).

4.2.2 K-meanss clustering

K-meanss clustering is based on centroid model of clustering. This model uses the mean vector to represent the cluster of data points. In K-means clustering data is classified into predefined number of clusters. The classification of data into clusters is based on nearest mean method (Wang et al., 2012).

4.2.3 Biclustering (Two-mode-clustering)

It is based on subspace models of clustering. In this method of clustering the clusters are defined by using the cluster models along with the relevant attributes.

4.3 Workflow

K-meanss clustering is implemented for the facies classification of the Lower Goru Formation encountered in Sawan-02, Sawan-03 and Sawan-07 well. K-means clustering is known for its efficiency and easy implementation (Nazeer and Sebastian, 2009). It is considered as the unsupervised machine learning technique. It classifies the n number of observations (data points) into k number of clusters. The mean of each cluster is represented by its center also known as centroid. Data sets are classified by using nearest mean method because, each data point in cluster is related with each other on basis of nearest mean (Wang et al., 2012). Mathematical expression showing K-meanss clustering is given as (Wang et al., 2012).

$$J = \sum_{j=1}^{k} \sum_{i=1}^{n} x_i^{(j)} - C_j^2 \quad (4.1)$$

Whereas, J is showing the objective function, K is for number of clusters, n is showing the number of data points. $x_i^{(j)}$ is showing the particular data point and C_j is considered as the center of the cluster. The generalized workflow for K-meanss clustering is shown in Fig. 4.1.

K-meanss clustering begins with the random selection of cluster numbers that are desired to be formed from given set of data points. The set of data points is divided into subsets based on already decided cluster numbers. Each subset is assigned with the mean value which is known as centroid or the center point of the cluster. Based on obtained results the cluster numbers are redefined which will redistribute the data again into clusters and assigned with centroids. These steps are iterated until the value of centroid become constant (Ali and Sheng-Chang, 2020).

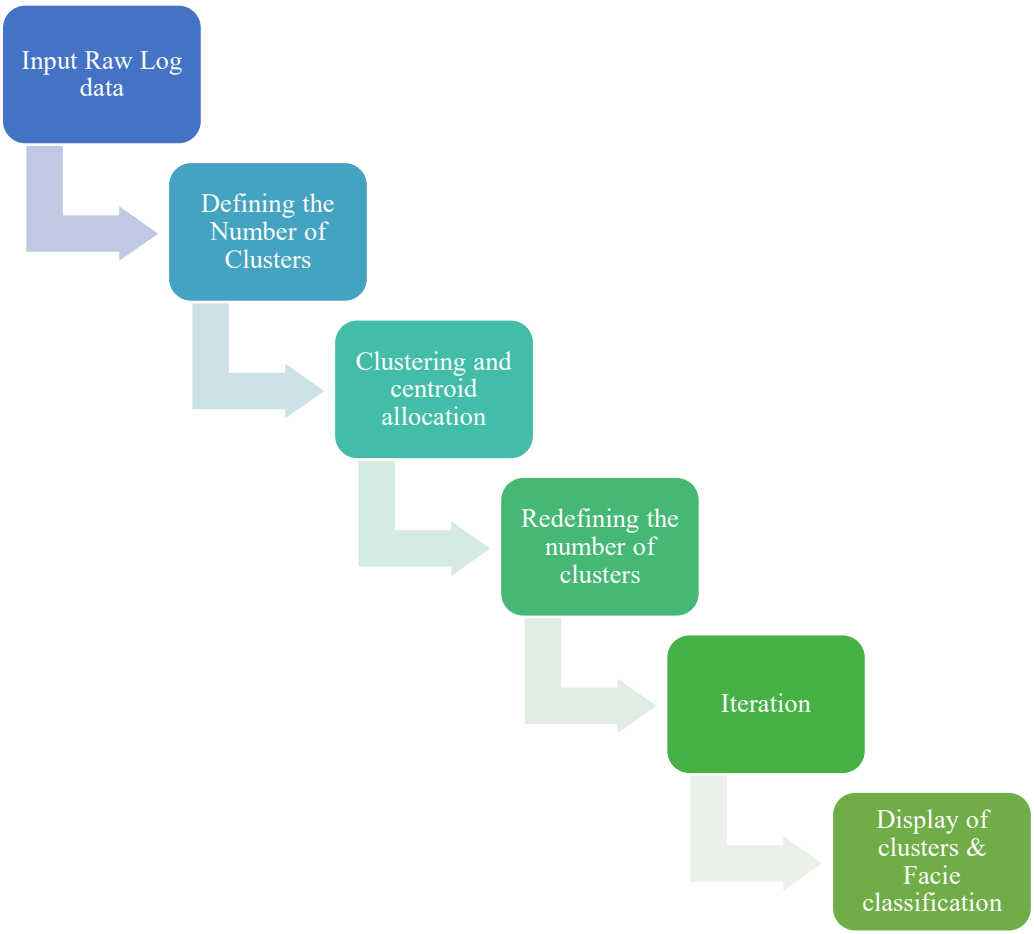


Figure 4.1: Generalized workflow of K-meanss clustering adopted for facies classification of Lower Goru Formation (Doveton, 1994).

4.4 Cluster analysis of the Lower Goru Formation

The Lower Goru Formation is acting as a proven hydrocarbon reservoir in Central and Lower Indus Basin of Pakistan. Lithological heterogeneity present in the Lower Goru Formation leads to its further classification into upper and lower parts. Alternative layers of sandstone and shale marks the upper part of the Lower Goru Formation. On the other hand, the lower parts of this formation also contain thin interbedding of limestone in addition to sandstone and shale. The sandstone layers of the Lower Goru Formation are subdivided into intervals that are D interval, C interval, B interval and A interval (Dar et al., 2021). Detailed facies analysis is required for efficient reservoir characterization. Facies analysis is generally performed by studying outcrop and core samples. Well log responses can also be used for the facie analysis of the rock formations encountered in

the wellbore (Chow, Li and Fuh, 2005). Analysis of conventional log suite of Sawan-02, Sawan-03 and Sawan-07 wells reveals that lithological heterogeneity is also present within the sand intervals of the Lower Goru Formation. So, K-meanss clustering can be used for the further classification of these intervals on basis of lithology. K-means clustering is performed by using the well data of Sawan-02, Sawan-03 and Sawan-07 including the log curves of GR, LLD, RHOB and NPHI tools.

4.4.1 Selection of cluster numbers

The data points obtained from log data curves are plotted on the crossplots. Frequency plot of data points is also created which illustrates the count rate of data points according to specific range as shown in Fig. 4.2.

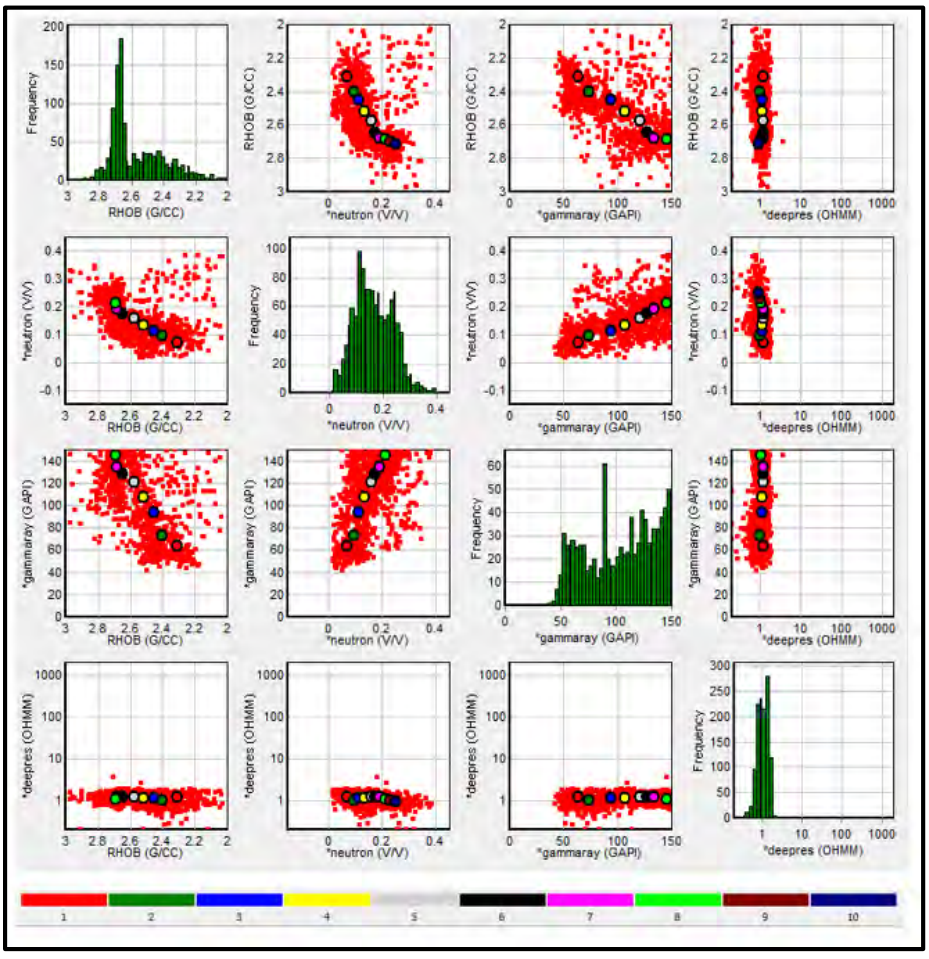


Figure 4.2: Crossplots showing number of clusters. The combination of curves is plotted in the form of crossplots. The red color points are showing data points. The colored points are assumed cluster means and frequency histograms are shown along the diagonal.

Once data is displayed on crossplots then the cluster numbers are selected randomly. The clusters should cover all the ranges containing the densely populated data points. In cluster analysis of the Lower Goru Formation ten clusters are randomly selected. These clusters are shown by colored points in Fig. 4.2. The colored bar scale present at the bottom of figure is the color key for the identification of randomly selected clusters.

4.4.2 Classification of data

The selection of clusters is followed by the initial estimation of mean for each log curve in every single cluster. The mean value for log curves present in clusters are shown in table 4.1. The mean value computed during this step is further used for K-means clustering.

Table 4.1: Computed mean value for clustering. This mean value is based on the random identification of clusters.

Cluster	#	Cluster	RHOB		*neutron		*gammaray		*deepres	
#	Points	Spread	Mean	Std Dev.	Mean	Std Dev.	Mean	Std Dev.	Mean	Std Dev.
1	0		2.3123		0.06908		63.887		1.1826	
2	0		2.4058		0.09661		72.911		0.99207	
3	0		2.4517		0.11177		93.612		1.1167	
4	0		2.5177		0.13528		107.01		1.1209	
5	0		2.5729		0.16011		120.93		1.1932	
6	0		2.6439		0.17515		127.87		1.2126	
7	0		2.6805		0.19214		134.16		1.1634	
8	0		2.6906		0.21315		145.26		1.0259	
9	0		2.7012		0.22921		158.43		0.97431	
10	0		2.7155		0.24945		176.19		0.91299	

The algorithm used for K-means clustering is based on minimum sum of squares of difference among cluster mean value and the specific data point. K-means clustering operates by calculation of sum of squares of difference that exists between each cluster and each data point. The data point is then assigned to a cluster which exhibits least difference of mean value and that specific data point. In this manner all data points are assigned to the suitable cluster. After assigning all the data points to clusters the new mean value of cluster is estimated. Calculation of new mean value is followed by the iteration which re assign the data points to clusters. These iterations continues until the mean value become constant and does not changes with iteration. The results thus

obtained are the clusters of data points based on the mean values of data groups known as K-means clusters. The clustering results of ten pre identified clusters are shown in table 4.2.

The results displayed in table shows the number of points, cluster spread, mean value and standard deviation. These entities are explained as follows.

- Number of points are the count of data points present within the specific cluster.
- Cluster spread is the standard deviation that marks the variation of distance of mean value of cluster from each data point. It also indicates the tightness of cluster. It is inversely related to cluster tightness. Lower the cluster spread, more tightly the data points will be clustered along the mean value.
- Mean value defines the arithmetic mean. It is the average value of log data within specific cluster along which the clustering of data occurred.
- Standard deviation shows the deviation value of the data points from the mean value of clusters. The gamma ray log is showing high value of standard deviation because, the log value of gamma ray curve is relatively high as compared to other logs.

Table 4.2: Computed spread and statistical parameters of clustering. These are the statistical results of cluster analysis.

Cluster	#	Cluster	RHOB		*neutron		*gammaray		*deepres	
#	Points	Spread	Mean	Std Dev.	Mean	Std Dev.	Mean	Std Dev.	Mean	Std Dev.
1	78	1.187	2.2406	0.111	0.08652	0.02965	99.771	34.87	1.3966	0.1915
2	47	1.598	2.2582	0.1321	0.29777	0.04861	110.81	27.77	0.67133	0.4896
3	209	0.8509	2.4109	0.08685	0.11581	0.03309	67.79	14.02	0.76798	0.1658
4	110	0.9206	2.49	0.07007	0.06843	0.02845	82.014	19.91	1.3871	0.2024
5	106	0.7637	2.6069	0.06911	0.14019	0.02469	116.44	16.04	1.0601	0.1584
6	42	1.039	2.6303	0.09958	0.12554	0.02907	201.32	37.21	1.3637	0.1145
7	121	0.9136	2.6314	0.0814	0.12577	0.02928	129.29	22.45	1.5532	0.2421
8	154	0.6146	2.6893	0.03474	0.2295	0.02893	177.16	13.25	0.91625	0.1193
9	183	0.7439	2.6995	0.07103	0.24	0.03223	131.68	19.89	0.86127	0.1479
10	172	0.6963	2.7236	0.06938	0.18967	0.02722	132.09	19.3	1.2868	0.1046

The results of K-means clustering are displayed on the crossplots as shown in Fig. 4.3. The colors of data points are representing the specific cluster of data. The mean value of cluster is plotted at center of each cluster and represented by the bold point of same color. The central point showing

mean of data is known as centroid. The colored scale bar is present at the bottom of the figure. The log data points that were initially plotted on the crossplots are now classified into ten clusters denoted by different colors.

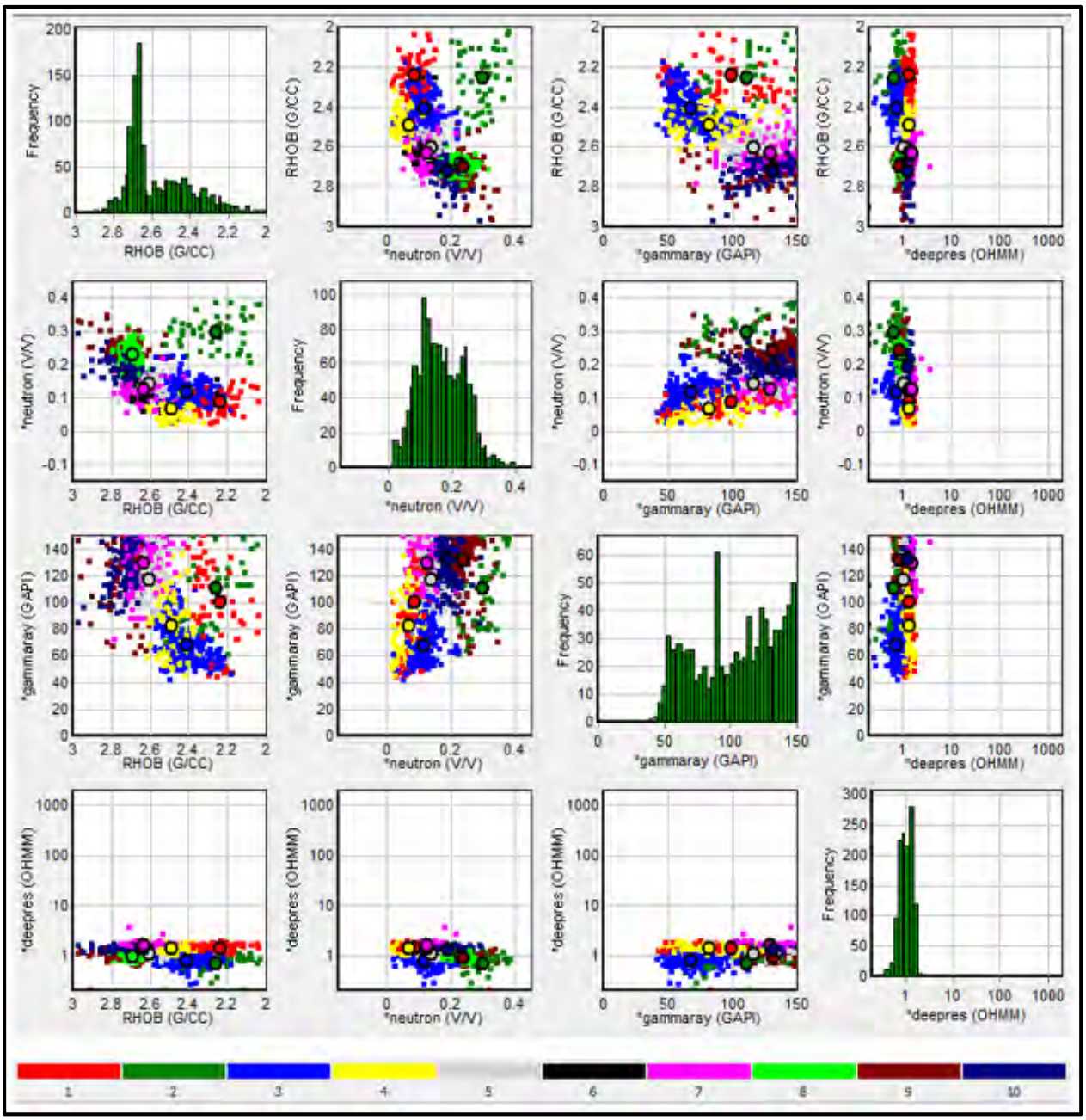


Figure 4.3: Crossplots showing clustered datasets. The combination of curves is plotted in the form of crossplots. The colored points are showing the clusters of well log data. The bold points are representing mean of each cluster. The frequency histograms are shown along the diagonal.

4.4.3 Consolidation of Clusters

Cluster consolidation is the process to merge the identical cluster into groups that define the specific rock facies. Hierarchical clustering method was adopted for consolidation of clusters into groups. This method operates by estimation of distances among all cluster followed by merging together the two closest clusters. After merging the two closest clusters the new distance among the clusters is re estimated and two clusters having minimum distances among them are merged. This process is repeated until only one cluster is left behind (Doveton, 1994). The clusters were consolidated by using average distance between merged clusters method. This method states that the distance from newly merged cluster to another cluster is the roughly equal to the average distance among all objects present in cluster formed by the merging of two clusters.

The results of cluster consolidation can be graphically represented in the form of dendogram. Dendogram is the tree diagram that represents the closeness relationship among the clusters (Doveton, 1994). The dendogram showing the cluster consolidation results of K-means clustering of the Lower Goru Formation is shown in Fig. 4.4. The original number of clusters computed from K-means clustering are shown at the bottom of the dendogram. Merging order is also shown in dendogram. The merging order number is displayed at the top of each branch. The resultant clusters of K-means clustering of the Lower Goru Formation are consolidated into five groups. Each group accounts for specific facie.

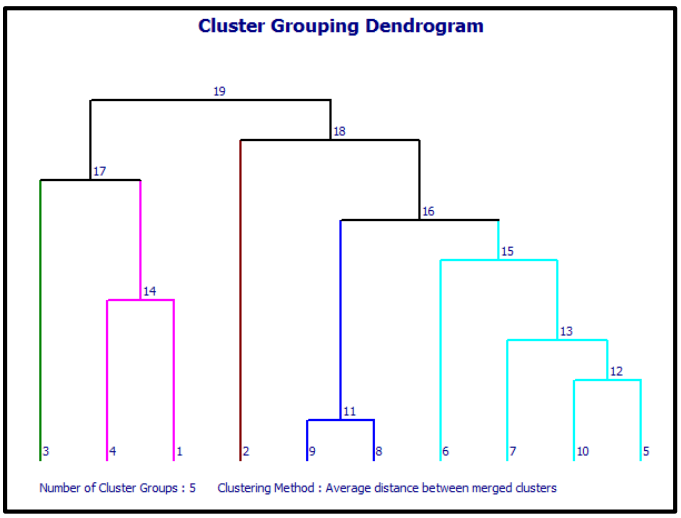


Figure 4.4: Dendogram showing the relationships among clusters. The finalized resultant consolidated clusters are represented by black lines.

The randomness of the cluster levels can be analyzed by using the randomness plot. The randomness of data is plotted along y-axis while the cluster groups are plotted along x-axis of the graph. The randomness of cluster analysis of the Lower Goru Formation is illustrated in Fig. 4.5. The randomness can be estimated by using the average thickness of cluster layer and random thickness. The numerical method for the estimation of randomness is given as (Doveton, 1994). Average thickness in Eq. 4.2. and Random thickness Eq.4.3. is required for the estimation of Randomness index Eq.4.4.

$$Average\ Thickness = \frac{Number\ of\ depth}{Number\ of\ cluster\ layer} \tag{4.2}$$

$$Random\ Thickness = \sum Pi\ /(1 - Pi) \tag{4.3}$$

Whereas pi shows the proportion of depth level given to i^{th} cluster.

$$Randomness\ index = \frac{Average\ Thickness}{Random\ Thickness} \tag{4.4}$$

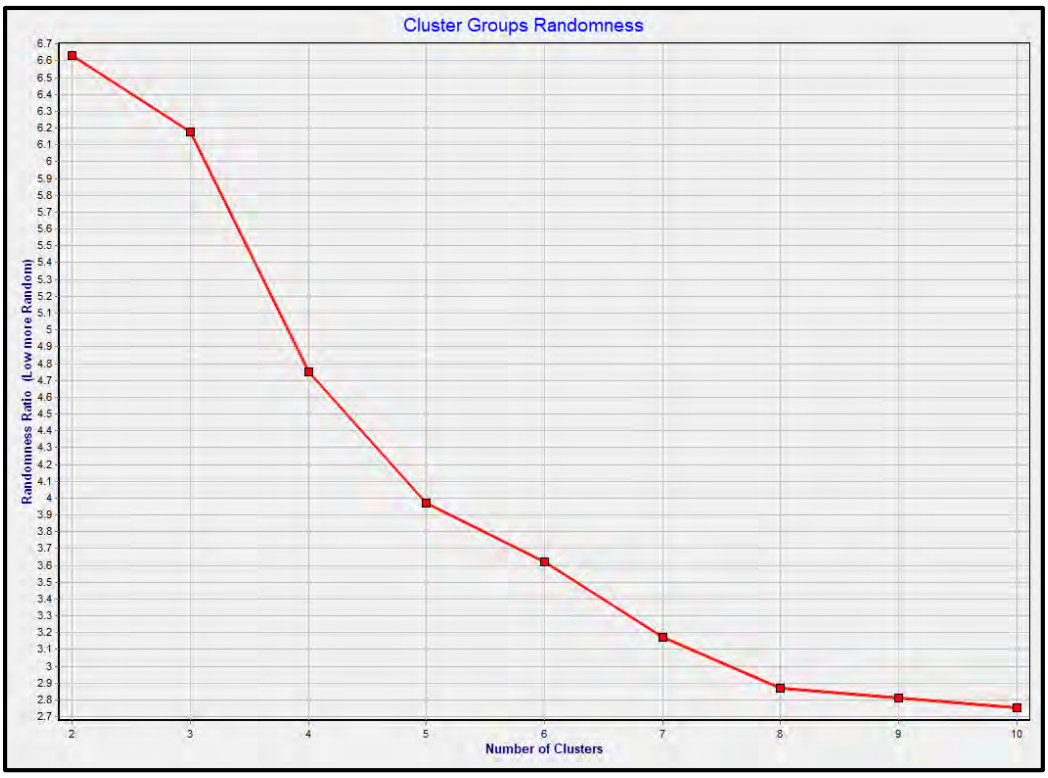


Figure 4.5: Showing cluster randomness. The 5th number cluster is considered as an inflection point. So, initially identified ten clusters are consolidated into five cluster groups.

The randomness plot shows that the most structured data is shown by cluster number 2 and 3. In contrast to this the point of inflection cluster number 5 which can also be termed as the elbow of the plot. The cluster grouping of five (facies) seems optimum in this case because, the graph after this shows the nearly uniform trend of randomness.

4.5 Classification of Facies

The five groups of clusters are classified into five facies on the basis of their log values. Two gas sands, water saturated sand, shaly sand and shale are identified on the bases of cluster analysis. The color pallet showing facies adjacent to their assigned colors is shown in Fig. 4.6.

4.5.1 Gas Sand

It is a medium quality reservoir facie comprised of red and yellow clusters shown Fig. 4.3. This facie shows low density and neutron log values. GR log values vary from low to high in this facie. The high GR log value challenges the quality of this reservoir facie.

4.5.2 Water Saturated Rock

This facie shows the high neutron porosity value, medium to low density value and low values of deep resistivity log. The GR log values are variable ranging from low to high values. These log indicators show that this facie consist of mixed lithology (sand and shale) and possesses high water saturation. This facie comprises of green colored cluster shown in Fig. 4.3.

4.5.3 Gas Sand 2

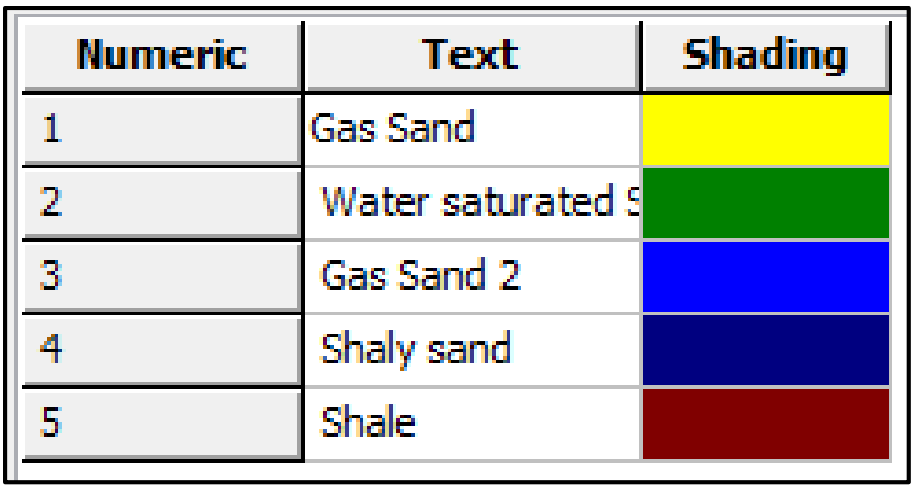
It is a good quality reservoir facie. It is marked by low values of neutron porosity log and density log. It also shows the low values of GR log. So, it is a clean gas bearing formation. It is represented by blue colored cluster in Fig. 4.3.

4.5.4 Shaly Sand

This facie shows the mixed lithology which is the mixture of sandstone and shale. This facie shows the medium to high density log value, medium level neutron porosity and GR log also shows the medium to high values. It is marked by pink, grey and black colored clusters in Fig. 4.3.

4.5.5 Shale

This facie is characterized by high neutron porosity log value, high density log value and very high value of GR log. These log indicators confirm the shale lithology. This facie is marked by dark brown and light green colored clusters in Fig. 4.3.



Numeric	Text	Shading
1	Gas Sand	
2	Water saturated S	
3	Gas Sand 2	
4	Shaly sand	
5	Shale	

Figure 4.6: Color code of facies. These are the five consolidated cluster groups which are classified into facies and each facie is represented by specific color.

4.6 Facies classification of Sawan-02

Facies classified by K-means clustering are plotted against the interpreted log plot of Sawan-02 well as shown in Fig. 4.7. Producing zone of Sawan-02 well is efficiently marked by both Gas Sand Facies. Mostly the pay zone is lying within the Gas Sand 2 facie which is a good quality reservoir facie. The shaly sand is comprising the D interval of the Lower Goru Formation. The shale is the dominant lithology in the upper portion of the Lower Goru Formation. The facies identified by K-means clustering are also verified by using the log response of raw log curves of conventional log suite which includes the GR, Resistivity, Neutron and density log curves.

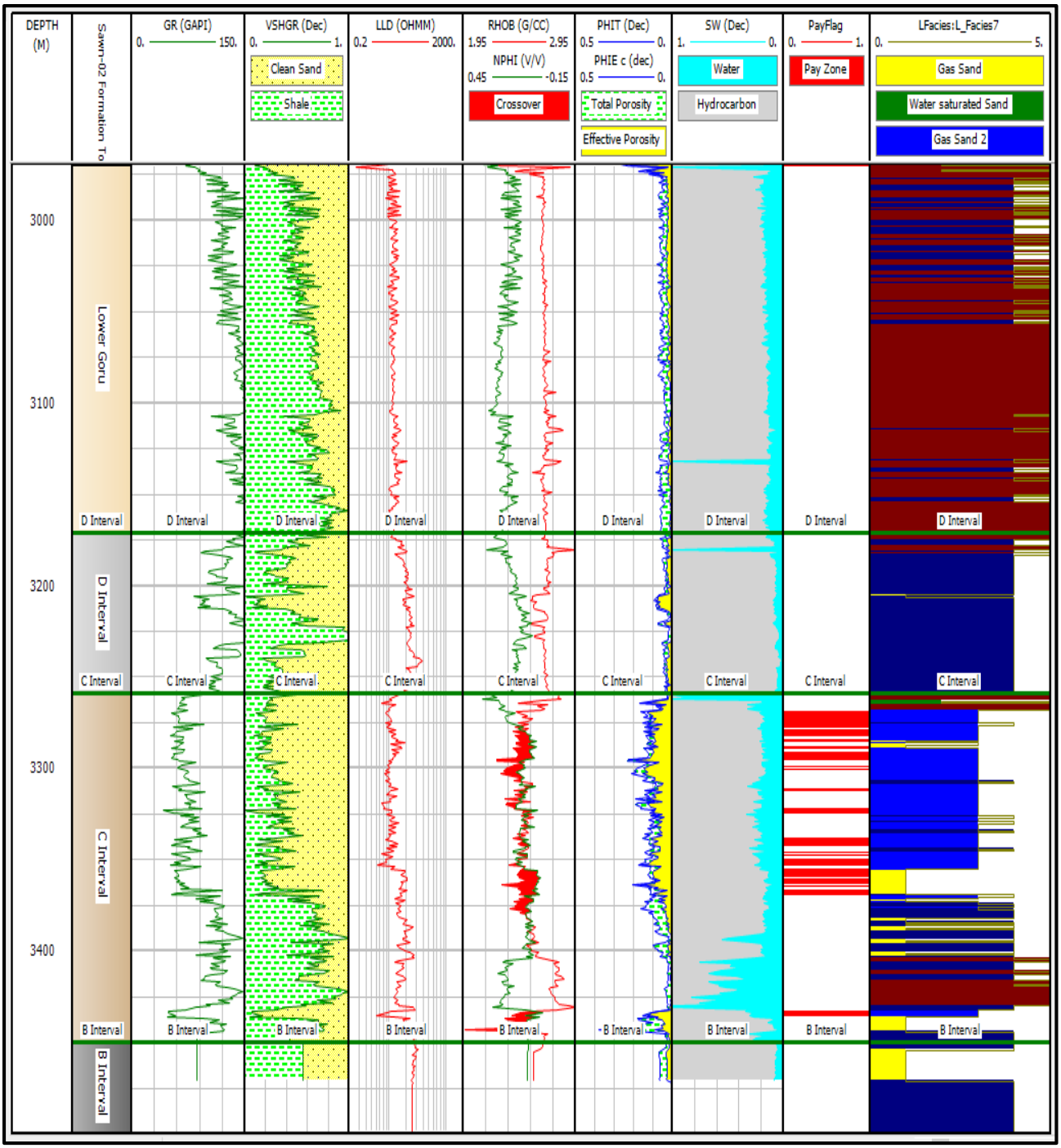


Figure 4.7: Electrofacies classification of Sawan-02. The good quality reservoir zone is marked by blue colored electrofacie (Gas Sand 2). The low-quality reservoir facie is shown by yellow color (Gas Sand). Alternative layers of navy blue and maroon marks sandy shale and shale respectively.

4.7 Facies classification of Sawan-03

Facies of the Lower Goru Formation classified by K-means clustering are plotted along the interpreted log plot of Sawan-03 well as shown in Fig. 4.8. The log responses and the clustered facies are efficiently verifying each other. The two reservoir facies are also shown. Gas sand 2 is hosting relatively thick pay zones as compared to gas sand facie.

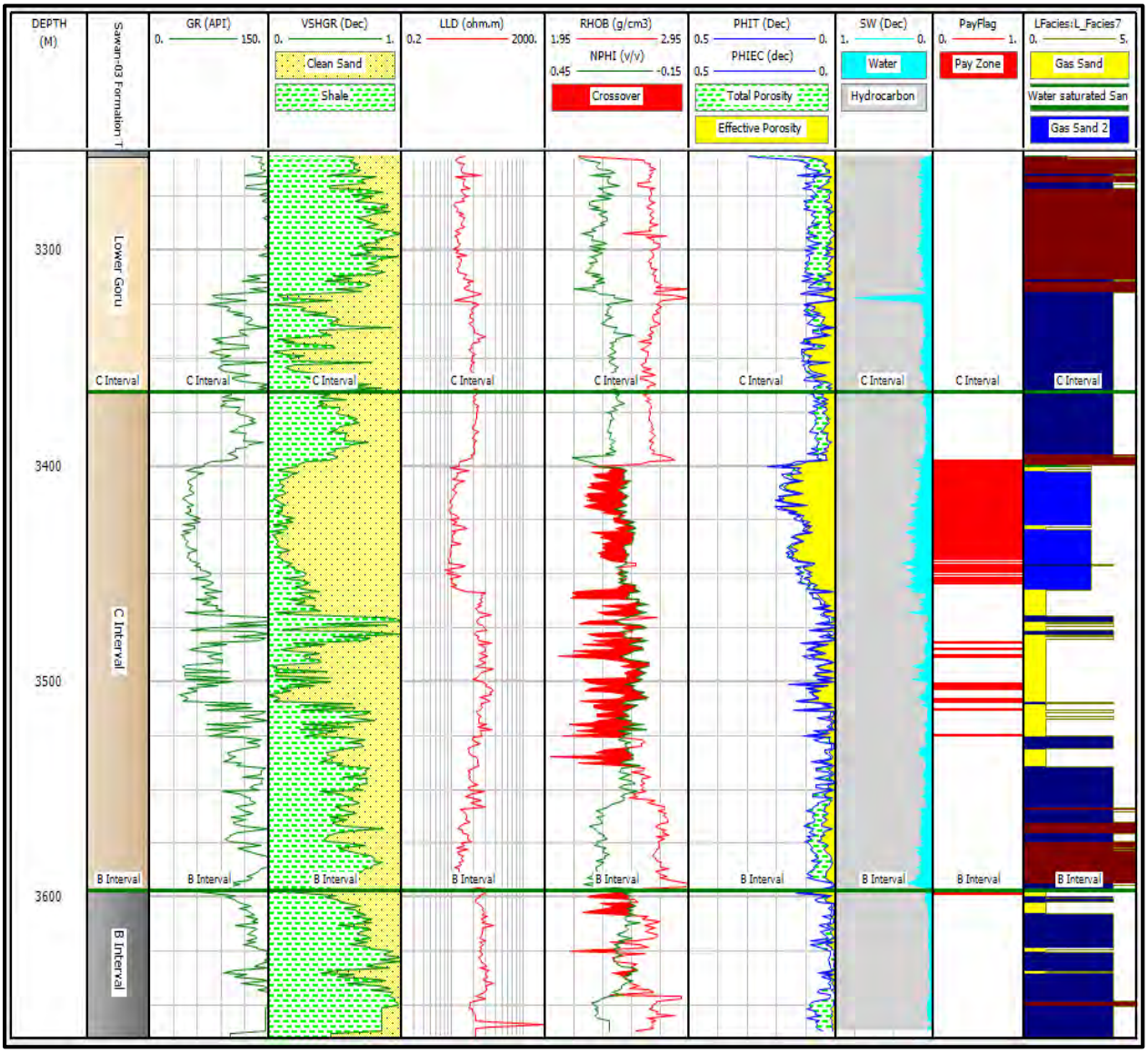


Figure 4.8: Facies classification of Sawan-03. The good quality reservoir zone is marked by blue colored electrofacie (Gas Sand 2). The low-quality reservoir facie is shown by yellow color (Gas Sand). Alternative layers of navy blue and maroon marks sandy shale and shale respectively.

4.8 Facies classification of Sawan-07

The interpreted log plot along with the resultant facies of K-means clustering of Sawan-07 well are shown in Fig. 4.9. The major pay zone thickness lies within Gas Sand 2 facie. The Gas sand facie also hosts the pay zones but their thickness is relatively less. The log responses are also satisfying the resultant facies of K-means clustering.

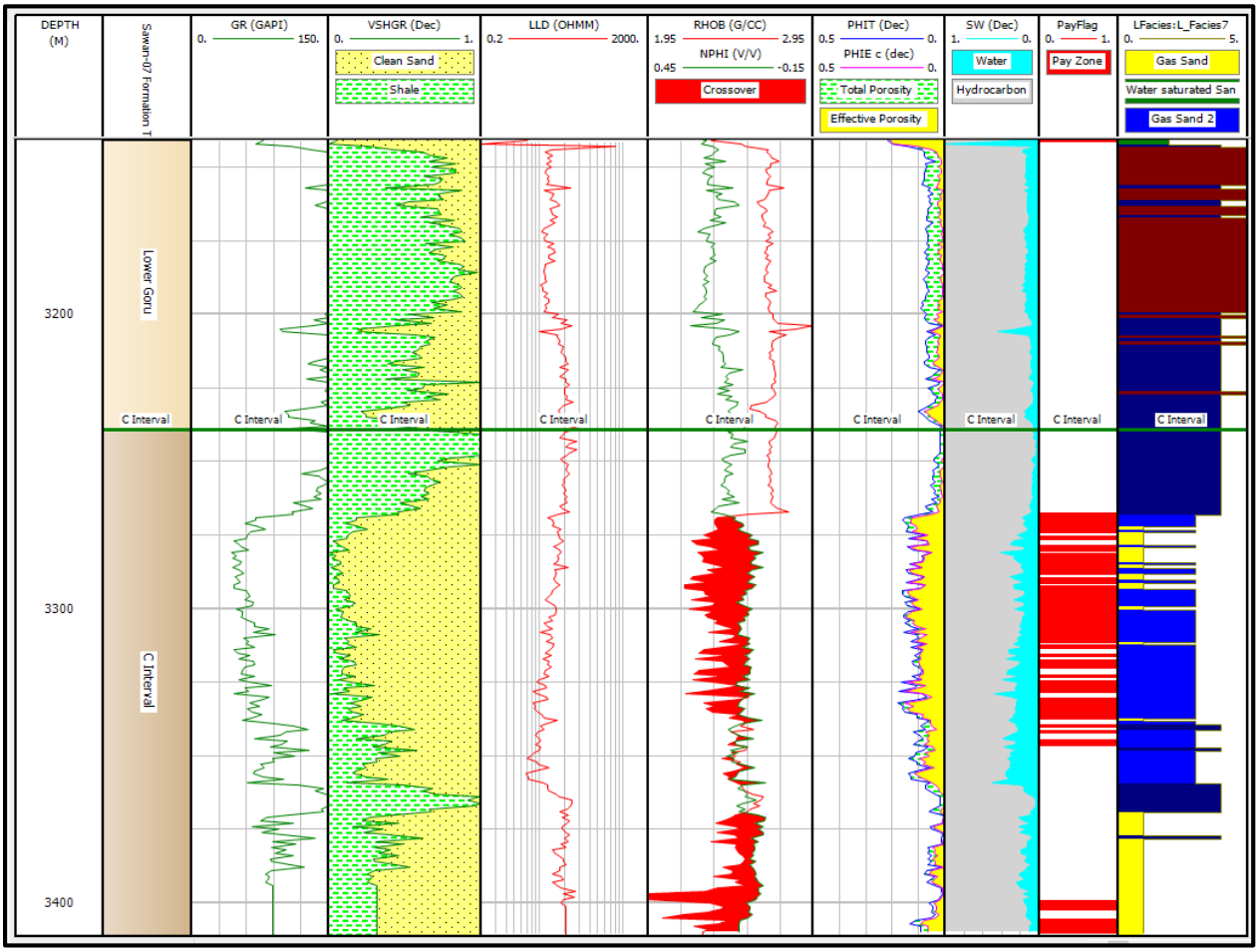


Figure 4.9: Facies classification of Sawan-07. The good quality reservoir zone is marked by blue colored electrofacie (Gas Sand 2). The low-quality reservoir facie is shown by yellow color (Gas Sand). Alternative layers of navy blue and maroon marks sandy shale and shale respectively.

K-means clustering is used for the facie classification of the Lower Goru Formation. The results of facies classification in all three wells (Sawan-02, Sawan-03 and Sawan-07) satisfies the log responses which shows that K-means clustering can aid the explorationists for adopting the efficient exploration strategy.

CHAPTER 05

SELF ORGANIZING MAPS

5.1 Introduction

Well log data is a sophisticated source that brought information about the subsurface rock formations into the access of geoscientists. Data obtained from multiple well logging tools is in the form of high dimensional data. Number of variables owned by a data point is known as the dimension of data (Qian et al., 2019). The dimensionality of data can be reduced by using the self-organizing maps. Self-organizing maps can be defined as the unsupervised machine learning tool having an ability to represent the high dimensional data in form of low dimensional data. During the representation of data, the topological structure is kept constant by using neighborhood function (Kohonen, 2012). SOM is considered as a type of artificial neural networks but it does not involve error correction learning and are trained by competitive learning (Kohonen and Honkela, 2007). The SOM is initially trained over the input data and after training the data is displayed in form of map which can be either in form of square, hexagon or sphere (Villmann and Bauer, 1998). The trained SOM can be calibrated or clustered to give the information about electrofacies.

Self-organizing maps can aid in the classification of electrofacies of geological formations having complex lithological variations like the Lower Goru Formation. Electrofacies can be defined as the subclasses of lithofacies (Khalid, et al., 2020). The Lower Goru Formation encountered in Sawan-02, Sawan-03 and Sawan-07 is classified into electrofacies by using self-organizing maps.

5.2 Workflow

Self-Organizing Map can be used for the prediction of missing log data over the SOM grid by using the trends of input data which can be further used for electrofacies identification. The SOM is the low-cost alternative to the other facie identification method like core studies. Results obtained from SOM can be calibrated with core results for more accuracy (Chang, et al., 2002). The Lower Goru Formation depicts the lithological heterogeneity so, SOM is applied for the electrofacies classification of the Lower Goru Formation encountered in Sawan-02, Sawan-03 and Sawan-07 wells.

The workflow of self-organizing maps begins with the input and normalization of raw log curves. Log curves of GR, RHOB, NPHI and LLD of Sawan-02, Sawan-03 and Sawan-07 are used while performing the electrofacies modeling of the Lower Goru Formation. After normalization of curves the SOM model is trained by using the log data of input curve. Trained SOM model is calibrated by using hierarchical clustering and displayed in the form of two-dimensional facies map. The generalized workflow adopted for the electrofacies modeling of the Lower Goru Formation is shown in Fig. 5.1.

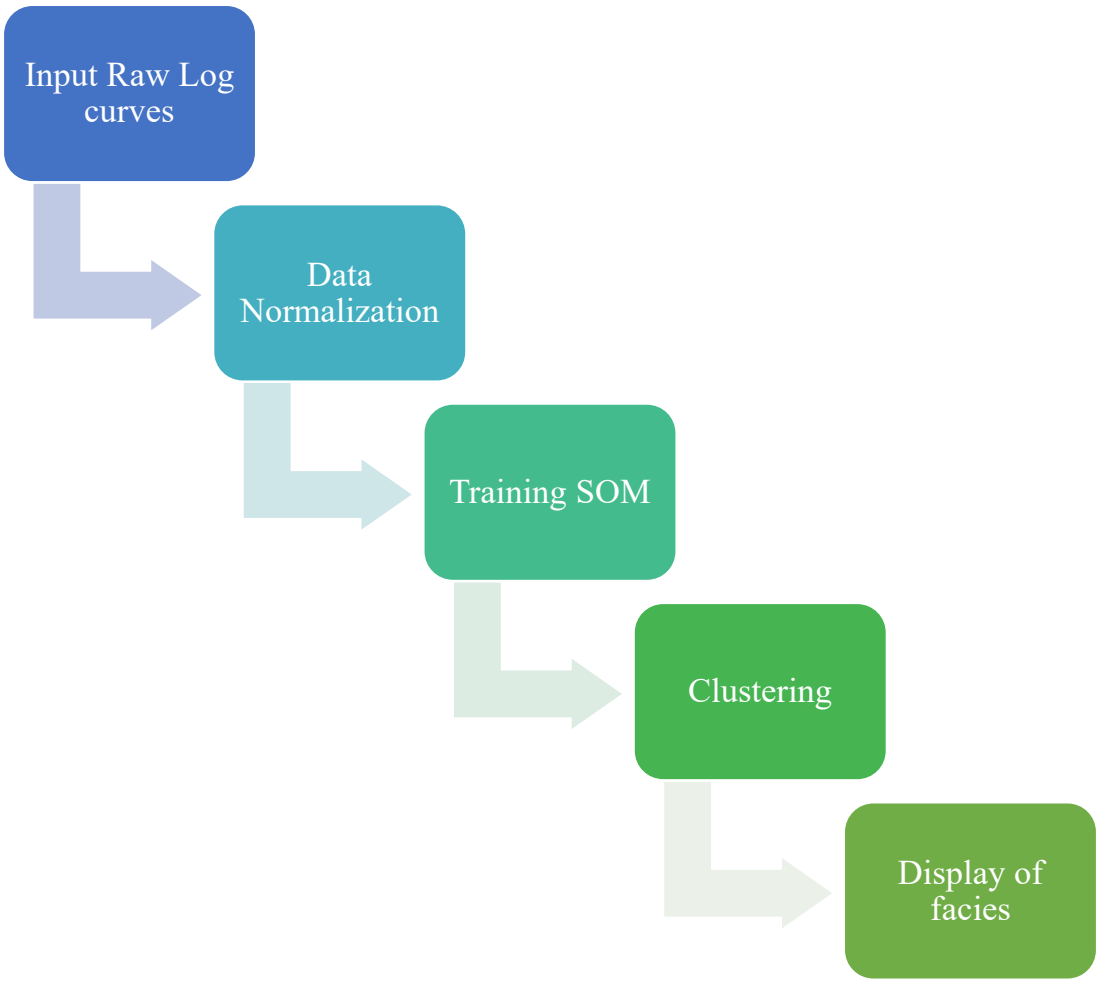


Figure 5.1: Generalized workflow adopted for electrofacies classification of Lower Goru Formation using self-organizing map (Khalid, et al., 2020).

5.3 Electrofacies analysis of Lower Goru Formation using self-organizing maps

The Lower Goru Formation is an important gas reservoir of Lower Indus Basin of Pakistan. Geologically it is distributed in Middle and Lower Indus Basin. The Lower Goru Formation depicts the noticeable lithological heterogeneity. The Lower Goru Formation mainly consists of alternative layering of sandstone, shale and due to this lithological heterogeneity the Formation is further classified into distinctive shale and sand packages by exploration companies (OMV and OGDCL) working in this area (Ahmad et al., 2004). In this research work the OMV classification is used which classifies the sandstone layers of the Formation as D, C, B, and A interval (Dar et al., 2021). Facies identification is a crucial step in reservoir characterization for efficient exploration and field development. Traditionally facies are identified by using the outcrop and core data. The advancement in modern computing techniques also suggest the low-cost model-based facies identification by using self-organizing maps (Chang, Kopaska-Merkel and Chen, 2002). Electrofacies classification by self-organizing maps is based on well log responses of the formation encountered in the wellbore. The raw log curves including GR, RHOB, NPHI and LLD of Sawan-02, Sawan-03 and Sawan-07 were used while performing electrofacies analysis of the Lower Goru Formation by using self-organizing maps.

5.3.1 Training the SOM model

The data obtained from input log curves (GR, RHOB, NPHI and LLD) of Sawan-02, Sawan-03 and Sawan-07 wells is plotted on the two-dimensional square map. The map was initiated with 20 nodes thus the relation among the width of map and total nodes is shown in Eq. 5.1.

$$Total\ nodes = Map\ width^2 \tag{5.1}$$

Initially the random value is assigned to the nodes of map. After this the value of each node in the map is determined by the weight value of each input curve. In our case the nodes are representing the weight of GR, RHOB, NPHI and LLD. The value of each node is represented by the simple bar graph within the specific node as shown in Fig. 5.2. The color key showing each curve is present at the bottom of figure.

Initialization of map is followed by the calculation of best matching unit (BMU) for each input data level. BMU can be defined as the node which closely resembles to the given input data. The color pallet map of input vector is shown in Fig. 5.3. The calculation of BMU is executed by the

calculation of Euclidean distance among the specific input vector and the weight vector of each node. The lowest value of Euclidean distance marks the BMU. The equation used for the computation of Euclidean distance is given in Eq. 5.2. (Kohonen, 2012).

$$Euclidean\ Distance = \sqrt{\sum_{i=0}^{i=n}(V^i - W^i)^2} \tag{5.2}$$

Whereas V shows the input vector and W indicates the vector weight of node.

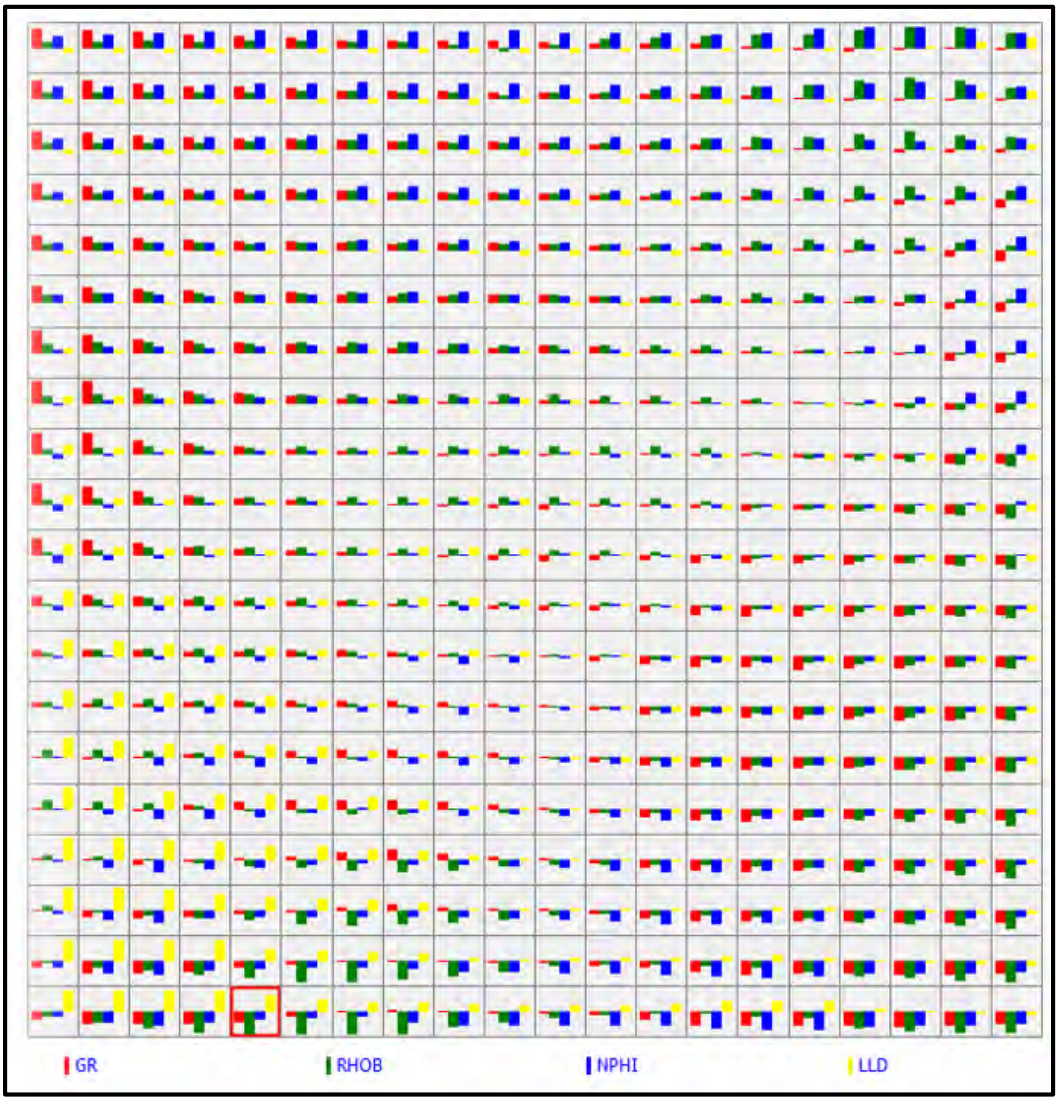


Figure 5.2: Map of well curve data using SOM. Each node is defined by using the values of GR, RHOB, NPHI and LLD log curves. The histograms present in each node are representing the values of specific log curves. The color assigned to each log curve is given at the bottom of map.

After the determination of BMU, the weights of nodes readjusted to develop similarity among nodes and input vector. This process is mathematically given in Eq. 5.3. (Kohonen, 2012).

$$W_{t+1} = W_t + L_t(V_t - W_t) \tag{5.3}$$

Whereas W shows the weight, t is for time step, v is denoting input vector and L is indicating learning rate. The decrease in initial learning rate for each iteration is given in Eq. 5.4. (Kohonen, 2012).

$$L_t = L_o \, exp(-\frac{1}{\lambda}) \tag{5.4}$$

Whereas Lo shows the initial learning rate, t is for training pass iteration and λ is denoting the time constant. The learning rate used for the training of SOM was 0.1 and the number of iterations used were 60,000.

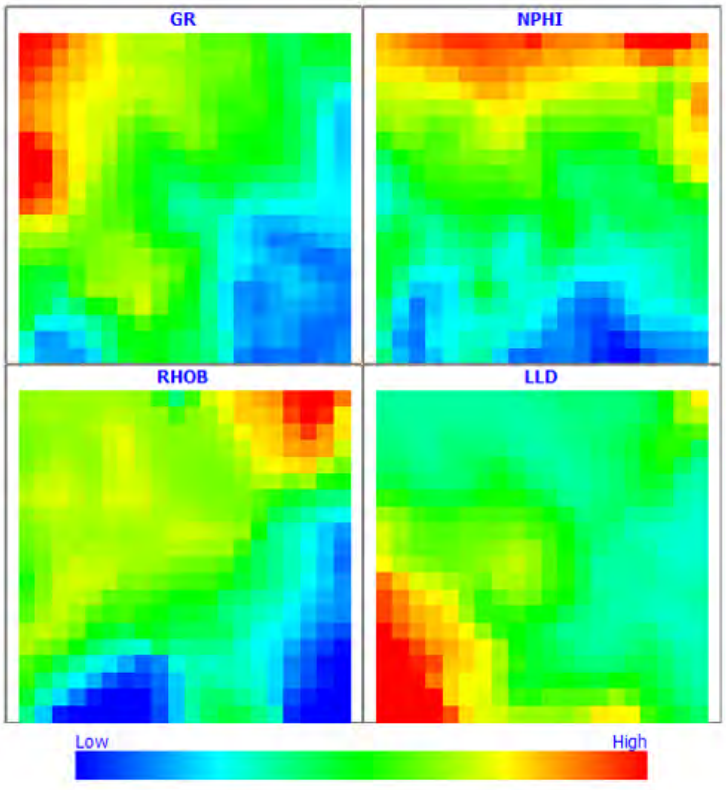


Figure 5.3: Color pallet-based data distribution for Self-organizing map model. The RGB color pallet is used for the display of data values. Red color shows the higher curve values and blue color is for low values.

The adjustment of weight vectors of BMU results in the adjustment of neighboring nodes of BMU. The neighboring radius of BMU can be identified by using Pythagoras theorem. The estimation of neighborhood radius is followed by the initialization of radius as a half mapping grid. The neighborhood radius reduces with each iteration and finally reaches to single node. The equation governing the decrease of neighborhood radius is same as equation 5.4 which is showing decrease in learning rate. The adjustment of weights of neighboring nodes is done by executing the predefined number of iterations. Upon completion of iteration the SOM map is ready for clustering as shown in Fig. 5.2. The distortion of SOM gives the information about the measure of resemblance of trained SOM with the training data. The estimated distortion value of SOM is 1.8 which shows that map is suitable for clustering.

5.3.2 Clustering of data

The training of SOM model is followed the clustering of data points. Hierarchical method was adopted for clustering. This method is already explained in detail in chapter 4.4.3. The classification of data points into five cluster showing five facies were performed by using average distance between merged clusters.

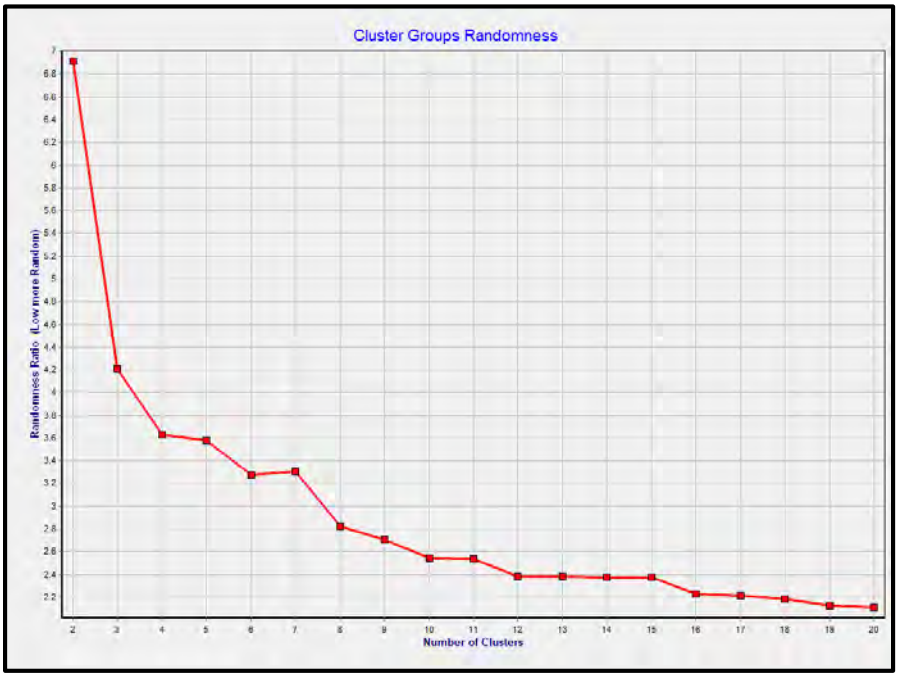


Figure 5.4: Randomness among data clusters. At cluster number five inflection is observed because trend of data is shifting from high values to low. So, data is consolidated into five cluster groups.

The randomness plot of showing randomness ratio of clusters is shown in Fig. 5.4. The figure shows that the five cluster are suitable according to trend of data because, inflection in plot is clearly visible at cluster number five.

SOM resulted in a large volume of data which is clustered into five clusters. The dendogram showing the hierarchical clustering from 60 clusters to five cluster groups is shown in Fig. 5.5. Sixty clusters are shown at the bottom of dendogram because, it is difficult to show all data points in dendogram. So, this dendogram only shows the clustering relationship present among the clusters.

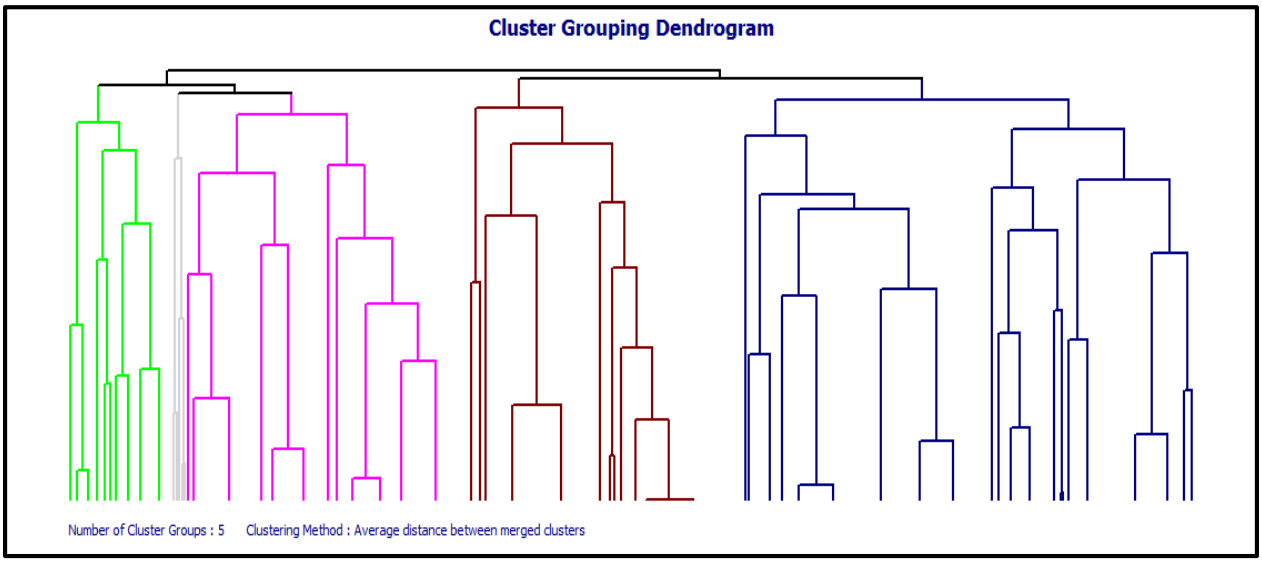


Figure 5.5: Dendogram showing relationship among clusters. The resultant five consolidated cluster groups are marked by black line on the top of dendogram.

The dataset produced as a result of clustering is the classified form of input vector data of SOM. The five clusters of well log data are displayed in the form of multi curve crossplots in Fig. 5.6. Frequency distribution of data is also shown in the form of frequency histograms in Fig. 5.6. Each cluster of data is defined by specific color. The color-coded scale bar is shown at the bottom of figure.

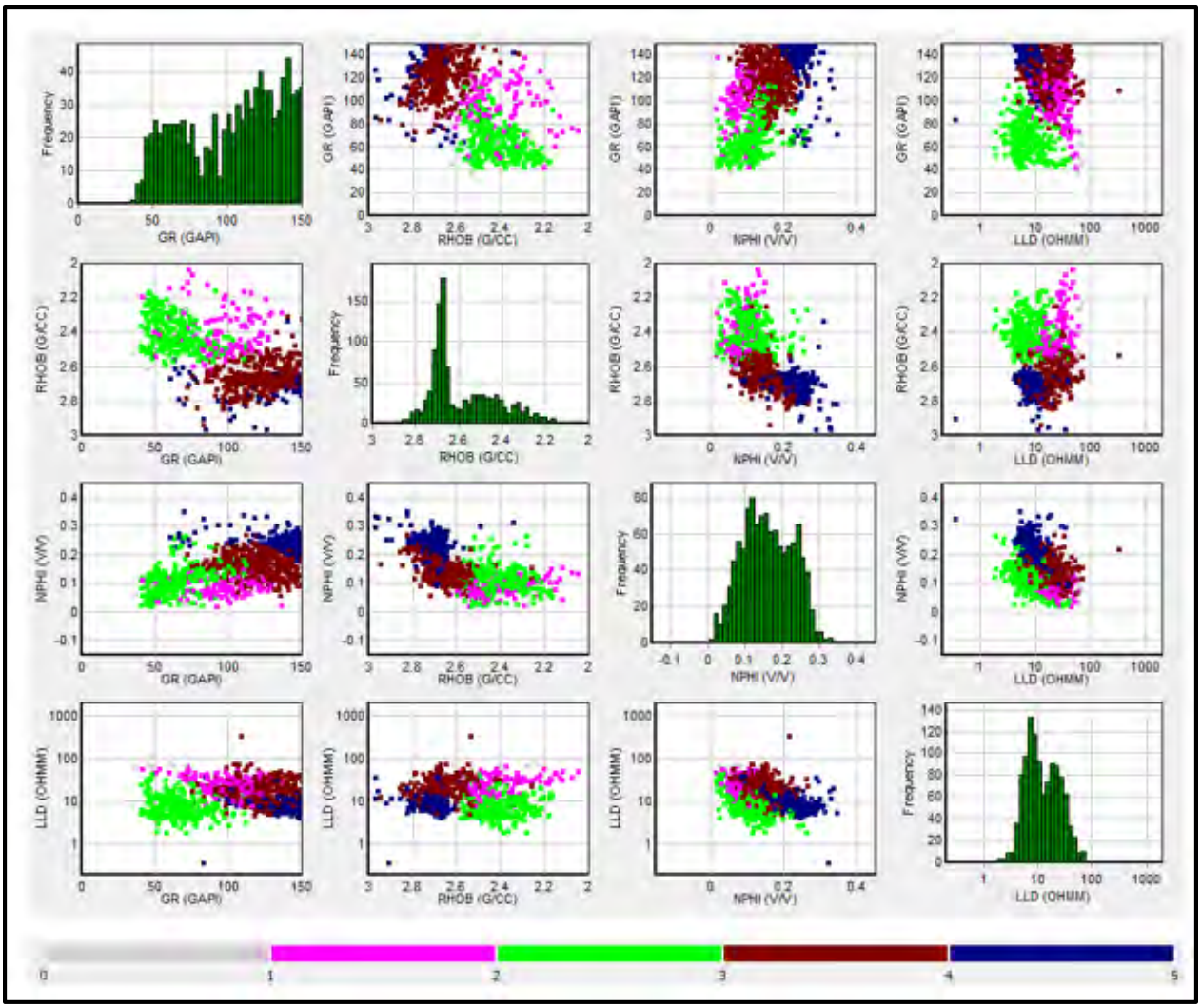


Figure 5.6: Crossplots showing clustered data set. The combination of curves is plotted in the form of crossplots. The colored points are showing the clusters of well log data which is coded according to the scale given at the bottom of figure. The frequency histograms are shown along the diagonal.

5.4 Classification of Facies

The five facies are classified based on SOM results. Two gas sands, water saturated sand, shaly sand and shale are identified by SOM. The color pallet showing facies adjacent to their assigned colors is shown in Fig. 5.7.

5.4.1 Water Saturated rock

This facie shows the high neutron porosity value, medium to low density value and specific response of LLD. The GR log values are variable ranging from low to high. These log indicators show that this facie consist of mixed lithology (sand and shale) and possesses high water saturation. This facie comprises of grey colored cluster shown in Fig. 5.6.

5.4.2 Gas Sand 2

It is a medium quality reservoir facie comprised of pink cluster shown Fig. 5.6. This facie shows low density and neutron log values. GR log values vary from low to high in this facie. The high GR log value challenges the quality of this reservoir facie.

5.4.3 Gas Sand

It is a good quality reservoir facie. It is marked by low values of neutron porosity log and density log. It also shows the low values of GR log. So, it is a clean gas bearing formation. It is represented by light green colored cluster in Fig. 5.6.

5.4.4 Shaly Sand

This facie shows the mixed lithology which is the mixture of sandstone and shale. This facie shows the medium to high density log value, medium level neutron porosity and GR log also shows the medium to high values. It is marked by maroon colored cluster in Fig. 5.6.

5.4.5 Shale

This facie is characterized by high neutron porosity log value, high density log value and very high value of GR log. These log indicators confirm the shale lithology. This facie is marked by blue colored clusters in Fig. 5.6.

Facies	Value	Text Value	Color
1	1	Water Saturated	
2	2	Gas Sand 2	
3	3	Gas Sand	
4	4	Sandy shale	
5	5	Shale	

Figure 5.7: color pallet showing key for electrofacies demarcation. These are the five consolidated cluster groups which are classified into facies and each facie is represented by specific color.

The classified electrofacies are related to each other based on SOM averages. The relationship among these facies is illustrated by using star plot shown in Fig. 5.8.

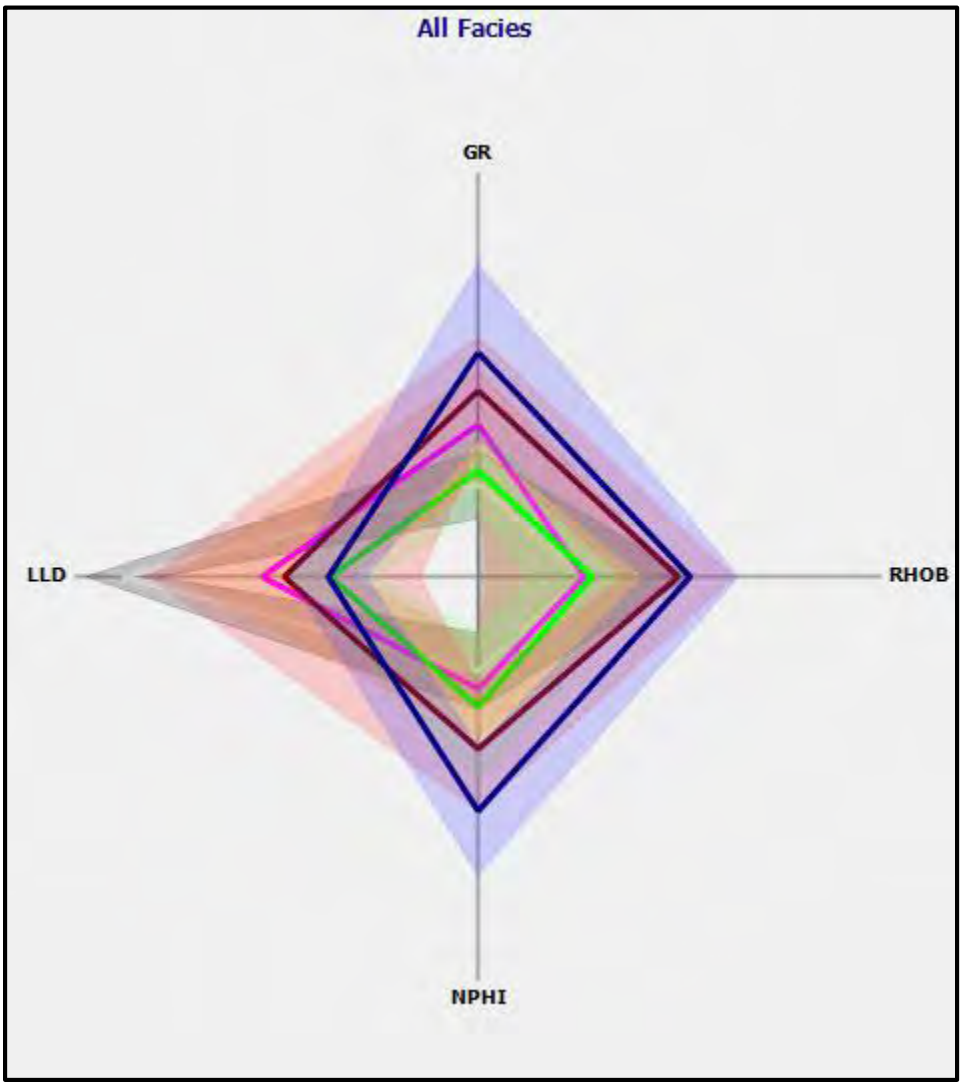


Figure 5.8: Relationship among classified facies. The solid colored line shows the average log value of the specific facie, while the shaded portion shows the standard deviation present within the specific facie cluster.

The facies are clustered on the basis of input log curves of the SOM model. The input vector data of SOM consisted of five raw log curves including GR log, bulk density log, neutron porosity log and deep resistivity log. The average values of each log curves for specific facies are indicated on the arms of star plot shown in Fig. 5.9.

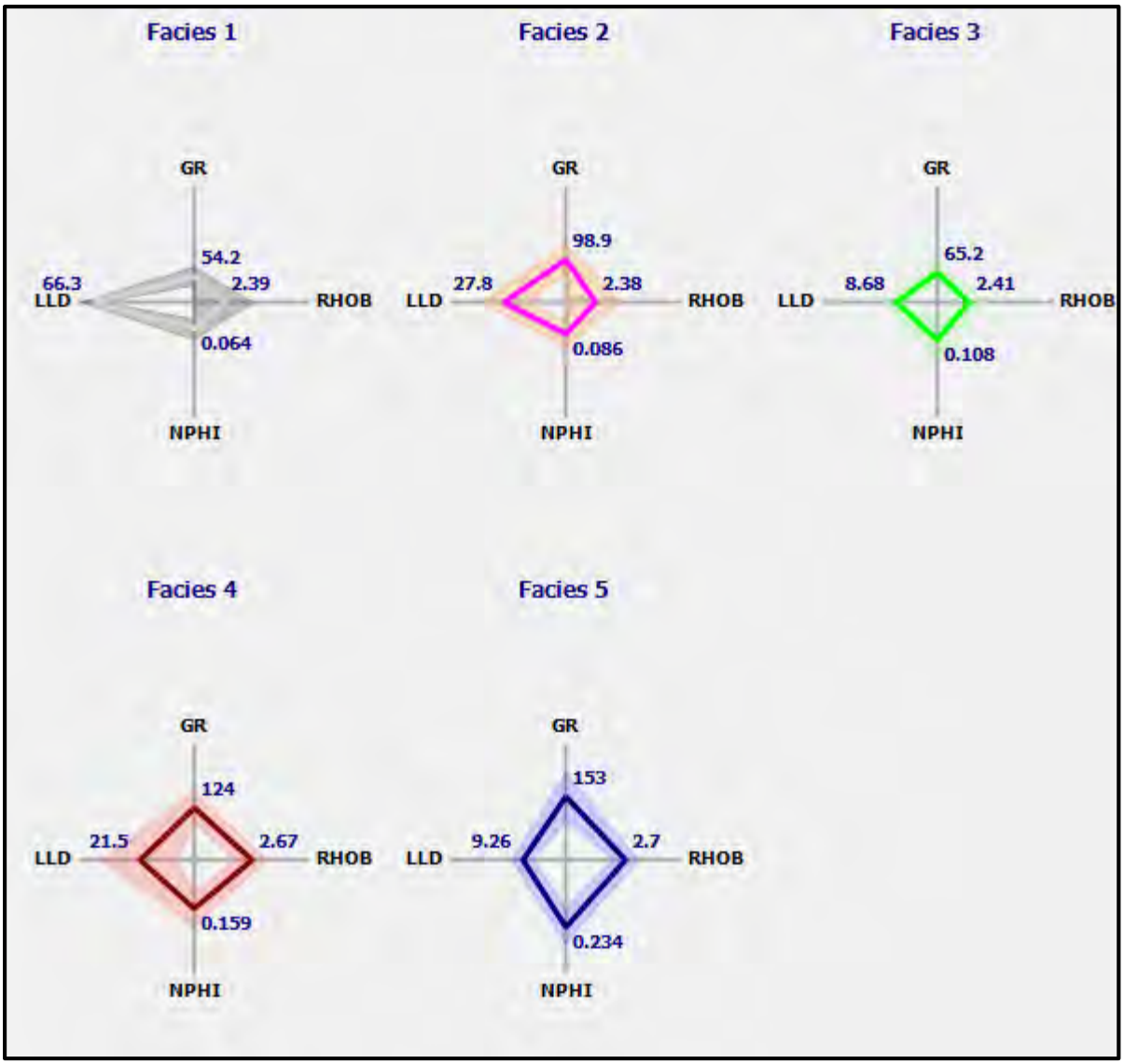


Figure 5.9 Means log curve values for each facie. Facie 1 is the water saturated rock, Facie 2 shows low quality reservoir formation, Facie 3 is the good quality reservoir formation, Facie 4 is the sandy shale and Facie 5 is shale.

The mean values of each log curve for specific facie cluster along with data points and standard deviation is shown in table 5.1.

Table 5.1: Statistical values for clustered facies

	Number of	GR	GR	RHOB	RHOB	NPHI	NPHI	LLD	LLD
	Data Values	Mean Value	Std Deviation	Mean Value	Std Deviation	Mean Value	Std Deviation	Mean Value	Std Deviation
Facies 1	6	54.21	12.69	2.387	0.1561	0.0645	0.0234	66.27	5.012
Facies 2	106	98.94	20.58	2.384	0.1399	0.0862	0.0285	27.82	12.18
Facies 3	277	65.21	15.12	2.412	0.0955	0.1082	0.0458	8.684	6.248
Facies 4	414	124.4	20.63	2.665	0.0799	0.1586	0.0383	21.46	18.66
Facies 5	312	153.4	32.76	2.699	0.0689	0.2343	0.0384	9.257	5.232
All Facies	1115	115	40.76	2.584	0.1594	0.1599	0.0658	15.72	14.98

The classification of facies is followed by the display of facies on two-dimensional self-organizing map. In this step the results of clustering are displayed over the SOM. The calibrated facie map of Sawan-02, Sawan-03 and Sawan-07 is shown in Fig. 5.10.

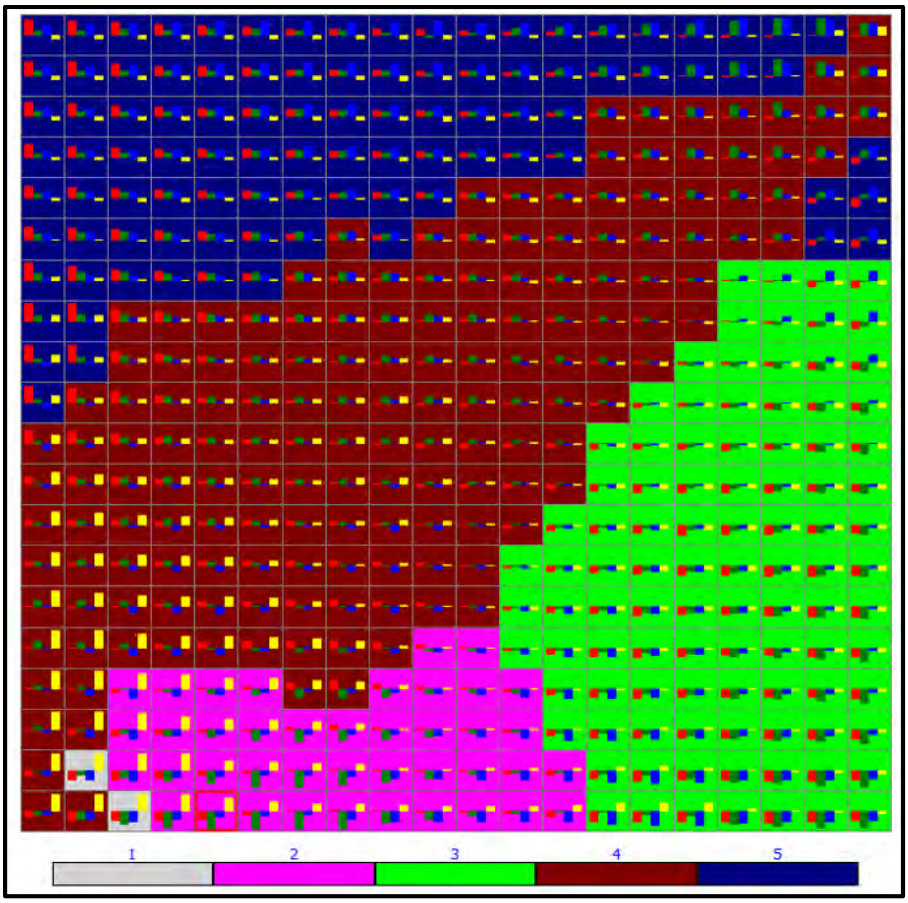


Figure 5.10: Facies calibrated SOM. The map illustrates that the major portion of the formation is consisted of sandy shale (Maroon) and shale (Blue). Good quality reservoir formation (Pale Green) is also present in suitable proportion. The proportion of water saturated formation (Grey) is negligible.

5.5 Facies classification of Sawan-02

Facies classified by K SOM are plotted against the interpreted log plot of Sawan-02 well as shown in Fig. 5.11. Producing zone of Sawan-02 well is efficiently marked by both Gas Sand Facies. Mostly the pay zone is lying within the Gas Sand facie which is a good quality reservoir facie. The shale is the dominant lithology in the upper portion of the Lower Goru Formation. The facies identified by SOM are also verified by using the log response of raw log curves of conventional log suite which includes the GR, Resistivity, Neutron and density log curves.

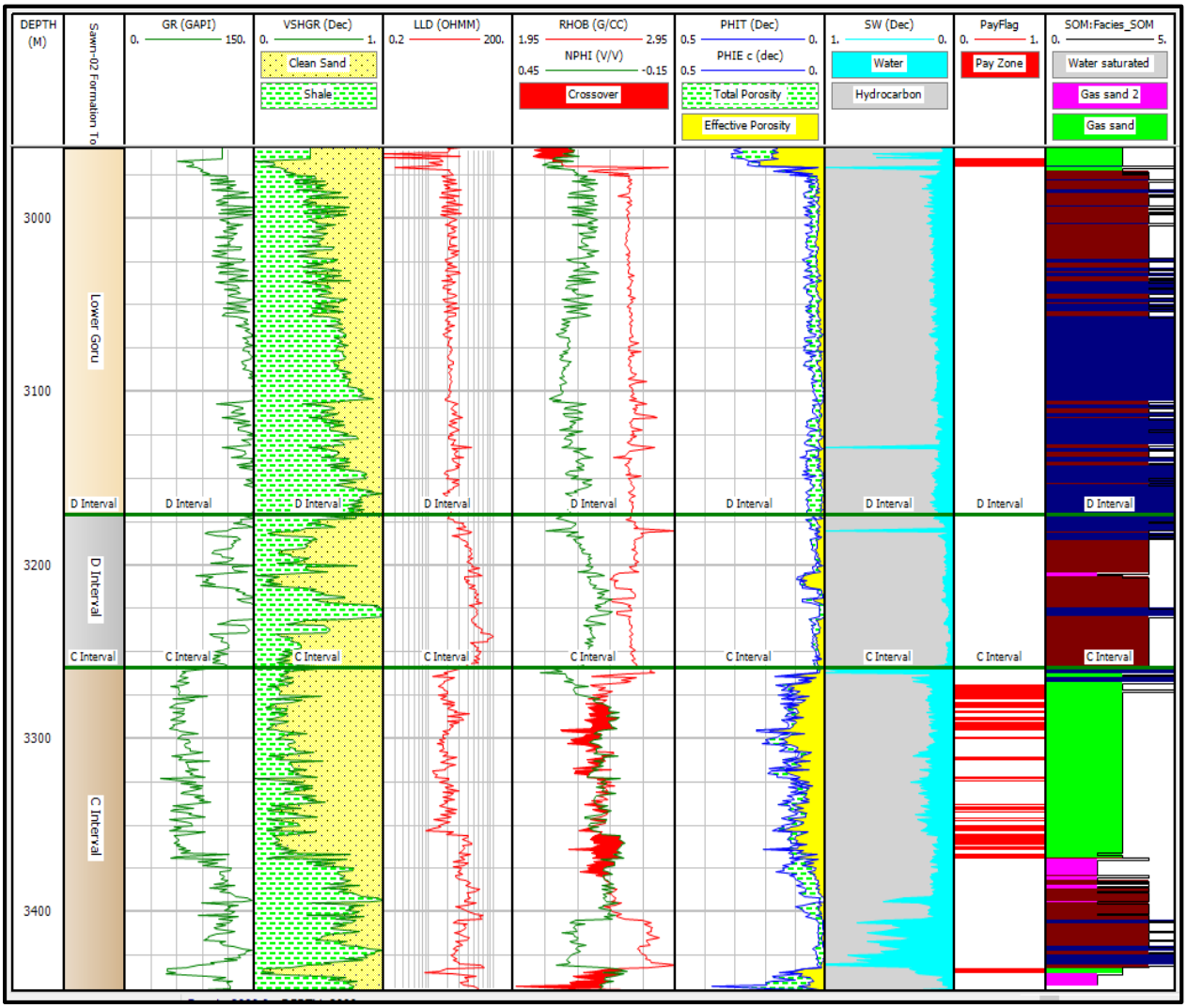


Figure 5.11: Facies classification in Sawan-02. The good quality reservoir zone is marked by pale green colored electrofacie (Gas Sand). The low-quality reservoir facie is shown by pink color (Gas Sand 2). Alternative layers of maroon and navy-blue marks sandy shale and shale respectively.

5.6 Facies classification of Sawan-03

Facies of the Lower Goru Formation classified by SOM are plotted along the interpreted log plot of Sawan-03 well as shown in Fig. 5.12. The log responses and the clustered facies are efficiently verifying each other. The two reservoir facies are also shown. Gas sand is hosting relatively thick pay zones as compared to gas sand 2 facie.

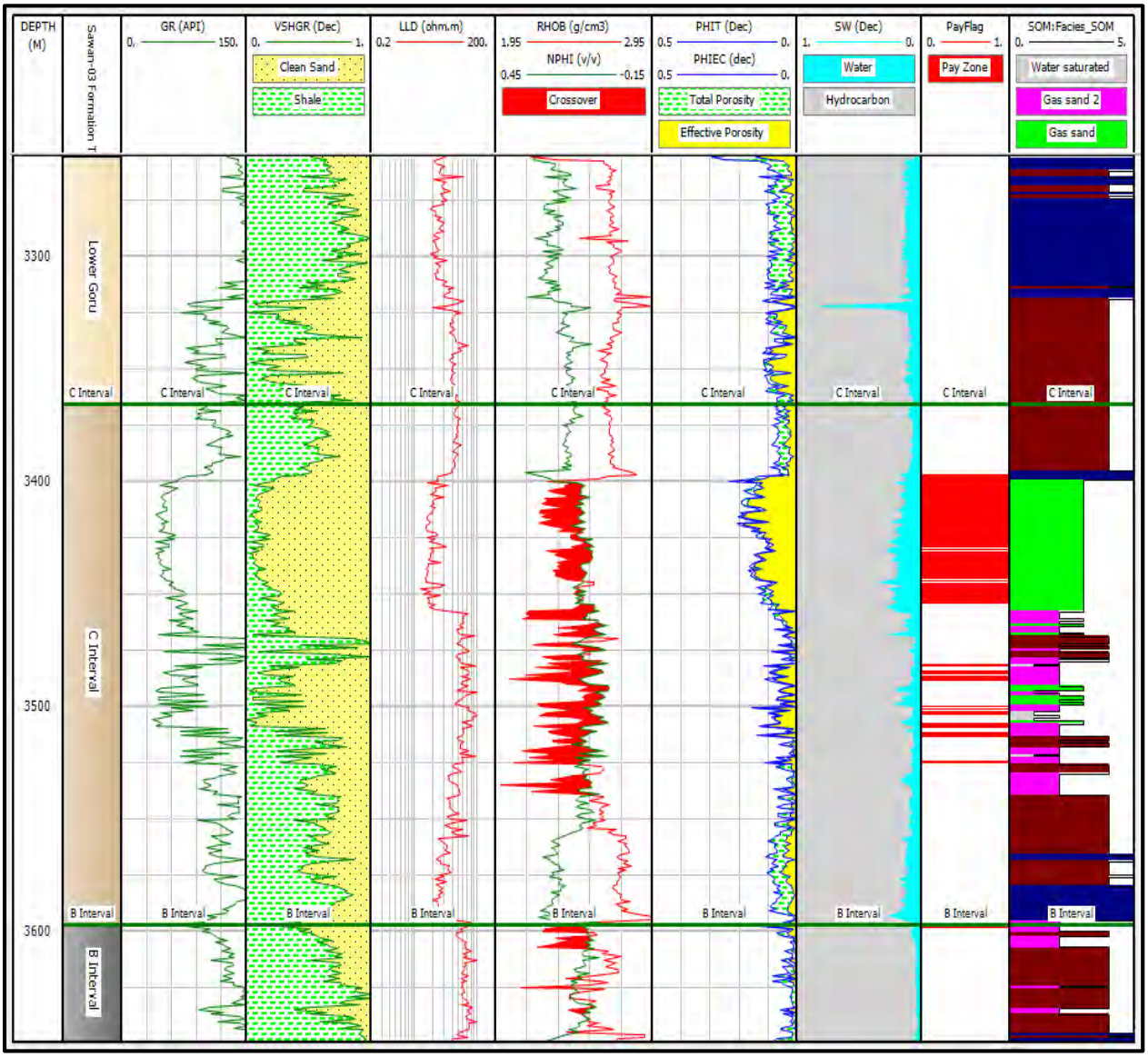


Figure 5.12: Facies classification in Sawan-03. The good quality reservoir zone is marked by pale green colored electrofacie (Gas Sand). The low-quality reservoir facie is shown by pink color (Gas Sand 2). Alternative layers of maroon and navy-blue marks sandy shale and shale respectively.

5.7 Facies classification of Sawan-07

The interpreted log plot along with the resultant facies of SOM of Sawan-07 well are shown in Fig. 5.13. The major pay zone thickness lies within Gas Sand facie. The Gas sand 2 facie also hosts the pay zones but their thickness is relatively less. The log responses are also satisfying the resultant facies of SOM.

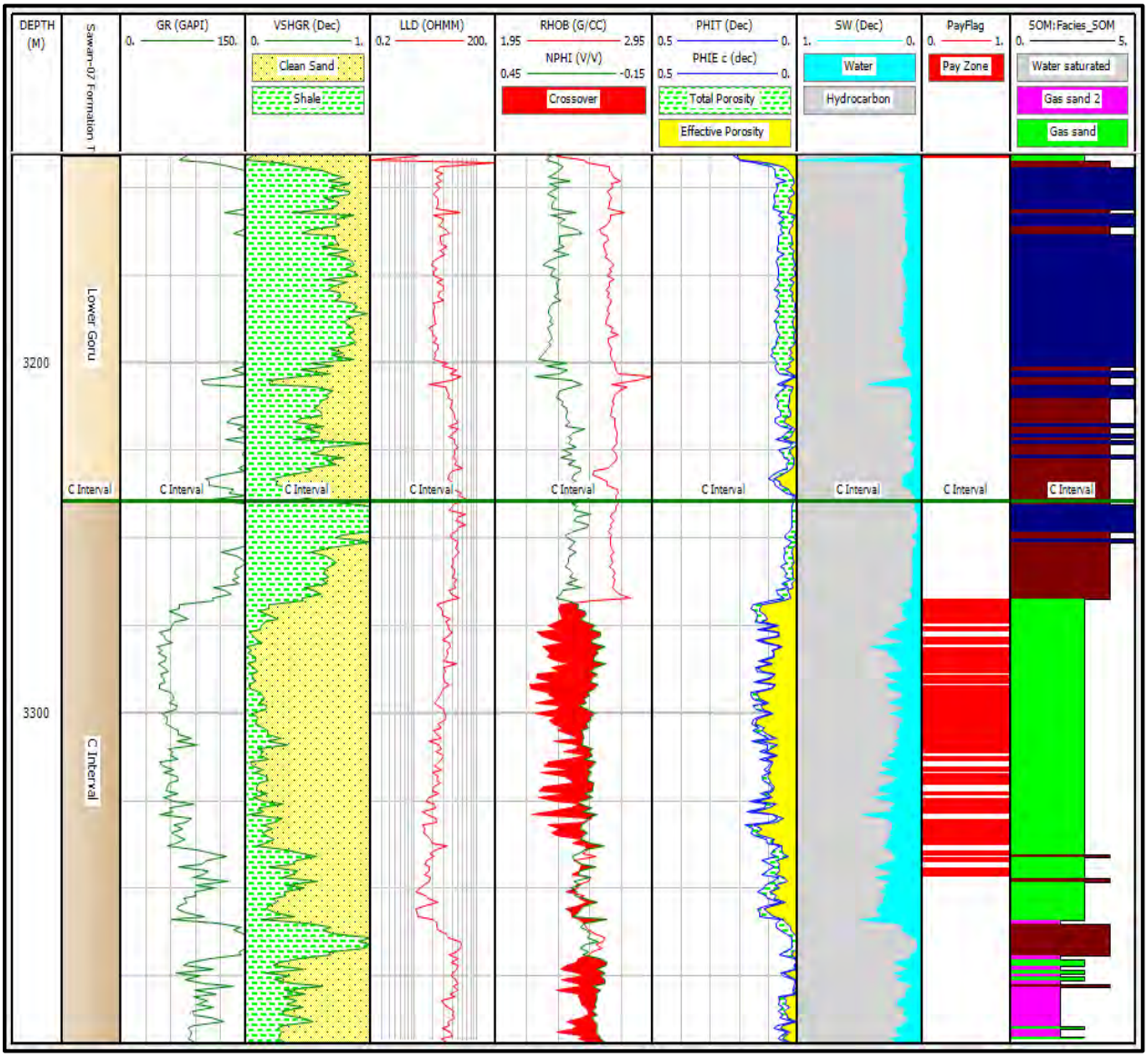


Figure 5.13: Facies classification in Sawan-07. The good quality reservoir zone is marked by pale green colored electrofacie (Gas Sand). The low-quality reservoir facie is shown by pink color (Gas Sand 2). Alternative layers of maroon and navy-blue marks sandy shale and shale respectively.

Chapter 6

Discussion

The main objectives of this research work are the estimation of petrophysical properties of the Lower Goru Formation followed by the electrofacies identification of the formation by using K-means clustering and self-organizing maps. The objectives are accomplished by using the well log data, formation tops and well header information of Sawan-02, Sawan-03 and Sawan-07 gas wells. The systematic workflows of petrophysical interpretation, K-means clustering and self-organizing maps are used for the modelling of lithological heterogeneities present within the Lower Goru Formation.

The research work initiated with the identification of reservoir zone within the Lower Goru Formation by using the well log responses of Sawan-02, Sawan-03 and Sawan-07 wells. The suitable reservoir zone is identified within the C interval of the Lower Goru Formation in all three wells. In addition to this major hydrocarbon bearing zone a small zone is also marked in upper portion of B interval encountered in Sawan-03 well. This step is followed by the petrophysical analysis of reservoir formation.

The petrophysical analysis of C interval reveals that the C interval is a having an excellent reservoir character with average shale volume of 15%, average effective porosity of 12% and average water saturation of 20%. On the other hand, the B interval encountered in Sawan-03 well possesses the relatively low reservoir potential with average shale volume of 26%, average effective porosity of 9% and average water saturation of 16%. The lithology of C interval and B interval comprises of alternating layers of sandstone and shale. Modified Simandoux equation is used for the computation of water saturation thus the value obtained for saturation is quite low because, modified Simandoux subtracts the effect of shale from water saturation. The average values of volume of shale in C interval are shown in Fig. 6.1. The average values of effective porosity in C interval encountered in all three wells are plotted over 3D cube given in Fig. 6.2. Average values of water saturations in C interval penetrated by Sawan-02, Sawan-03 and Sawan-07 are shown in Fig. 6.3.

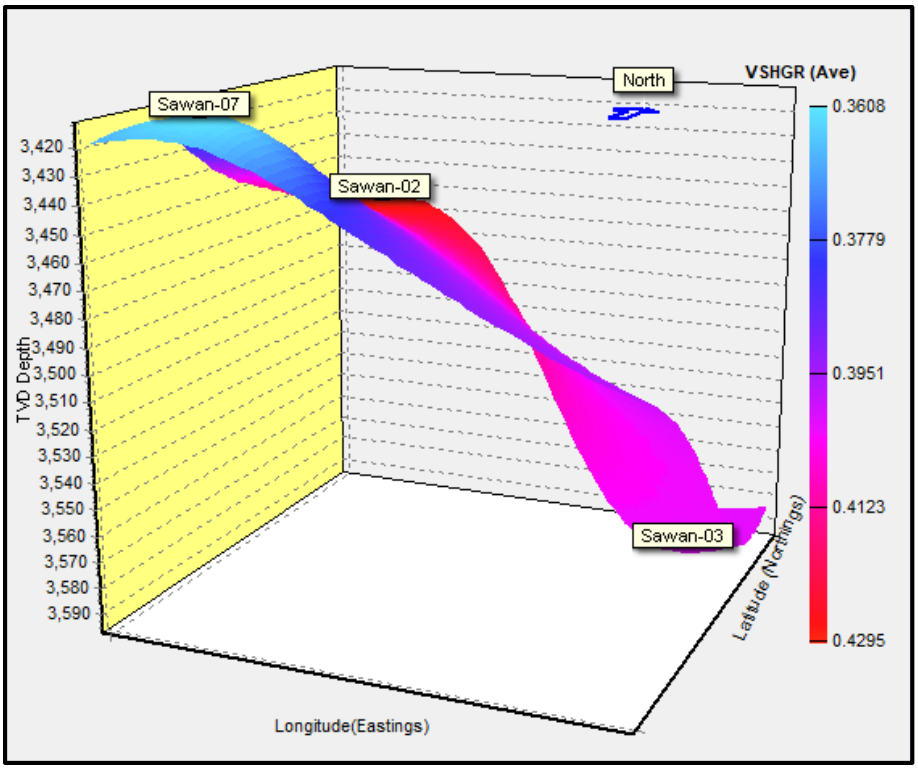


Figure 6.1: 3D cube Showing Average volume of shale in Sawan-03, Sawan-02 and Sawan-07 well. The depth of C interval is also indicated.

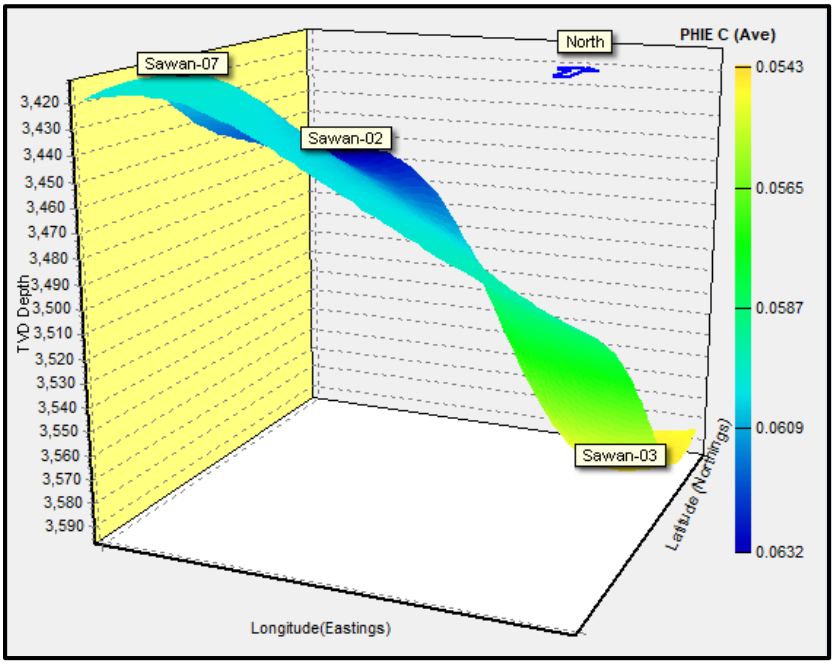


Figure 6.2 3D cube Showing Average effective porosity in Sawan-03, Sawan-02 and Sawan-07 well. The depth of C interval is also indicated.

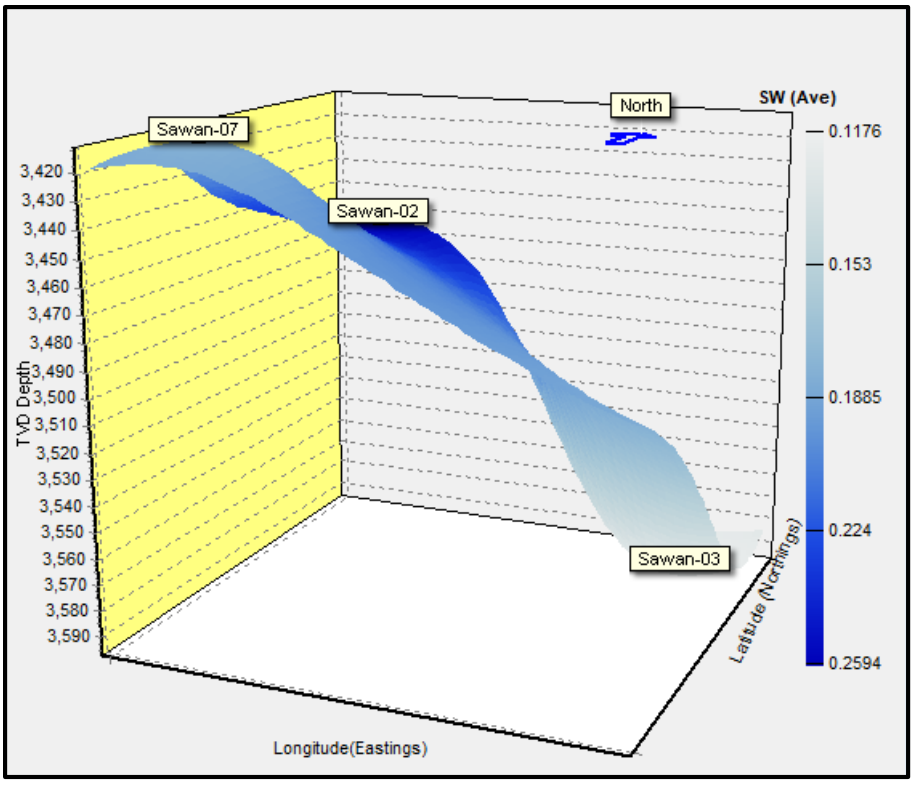


Figure 6.3 3D cube Showing Average water saturation in Sawan-03, Sawan-02 and Sawan-07 well. The depth of C interval is also indicated.

The final and important step of petrophysical analysis is the demarcation of pay zones. The cut off values are specifically decided for each well on the basis of petrophysical properties determined during petrophysical analysis. The cut off values adopted for the demarcation of pay zones in Sawan-02, Sawan-03 and Sawan-07 are shown in table 6.1. In Sawan-02 the cut off value of shale volume is kept very low because the lithology of reservoir formation encountered in the well is dirty. In contrast to this, the reservoir formation encountered in Sawan-03 shows relatively clean lithology. So, the cut off value of Vshale in Sawan-03 is kept higher than Sawan-02. The cut off value of water saturation in Sawan-07 is also kept lower because of overestimated saturation values due to the application of modified Simandoux equation. Efficient demarcation of pay zones is necessary for the success of strategy adopted for the exploration and production of hydrocarbon.

Table 6.1: Showing cut off values used for the demarcation of pay zones in Sawan-02, Sawan-03 and Sawan-07.

Well	Zn #	Zone Name	Top	Bottom	Thickness	VShl	PHIE	SW
Sawan-02	12	C Interval	3259	3450	191	<= 0.2	>= 0.08	<= 0.5
Sawan-03	13	C Interval	3365.6	3597	231.4	<= 0.35	>= 0.08	<= 0.5
	14	B Interval	3597	3666	69	<= 0.35	>= 0.08	<= 0.5
Sawn-07	12	C Interval	3239.39	3432	192.61	<= 0.25	>= 0.08	<= 0.3

Petrophysical analysis of the Lower Goru Formation is followed by K-means cluster analysis for the identification of electrofacies within the Lower Goru Formation. Electrofacies can be defined as the subclasses of lithofacies (Khalid, et al., 2020). Electrofacies distribution analysis can serve as an optimum strategy for dealing with the lithological heterogeneity of the Lower Goru Formation. Efficient lithological identification is a key for the efficient reservoir characterization. The log data comprised of GR, RHOB, NPHI and LLD log curves is classified into ten clusters by using K-means clustering technique. These ten clusters are further consolidated into five clusters by using the hierarchical clustering. The resultant five cluster marks the five electrofacies. Electrofacies marked during research work are named as Gas Sand 2, Gas Sand, Shaly Sand, Shale and Water Saturated Formation. The good quality reservoir zone in all three wells is marked by the Gas Sand 2 facie indicated by blue color. The correlation of Sawan-03, Sawan-02 and Sawan-07 is showing the results of K-means clustering and facie modelling of the Lower Goru Formation which is illustrated in Fig. 6.4.

The electrofacies marked by K-means clustering are further confirmed by using the self-organizing maps. Self-organizing maps are the type of artificial neural networks which are trained by using competitive learning (Kohonen and Honkela, 2007). The data obtained from GR, RHOB, NPHI and LLD log curves of Sawan-02, Sawan-03 and Sawan-07 are used to train the model of self-organizing maps. Once the model is trained the multidimensional data of log curves is plotted on the two-dimensional map. The resultant data obtained from the model of Self-organizing maps is

consolidated into five clusters using hierarchical clustering method. Each cluster of data represents the specific electrofacie. The marked electrofacies are classified as Gas Sand, Gas Sand 2, Shaly Sand, Shale and Water Saturated Formation. The electrofacies marked by using self-organizing maps in Sawan-03, Sawan-02 and Sawan-07 are shown in the well correlation given in Fig. 6.5. The quality reservoir zone of all three wells lies in the C interval of Lower Goru Formation and it is marked by Gas Sand facie which is shown by light green color in Fig. 6.5.

The electrofacies marked by K-means clustering and self-organizing maps are plotted adjacent to each other in a well correlation of Sawan-03, Sawan-02 and Sawan-07 which shows the facies distribution among three available wells as illustrated in Fig. 6.6. This correlation shows that the electrofacies marked by both methods are same. However, the negligible variation is present because of minor distortion in self-organizing maps. The distortion value in self-organizing map is 1.8. This value can be further minimized by using spherical arrangement of nodes. The results obtained from K-means clustering and self-organizing map reveals that these both techniques can serve as an efficient tool for electrofacies modeling which will greatly improve the efficiency of reservoir characterization.

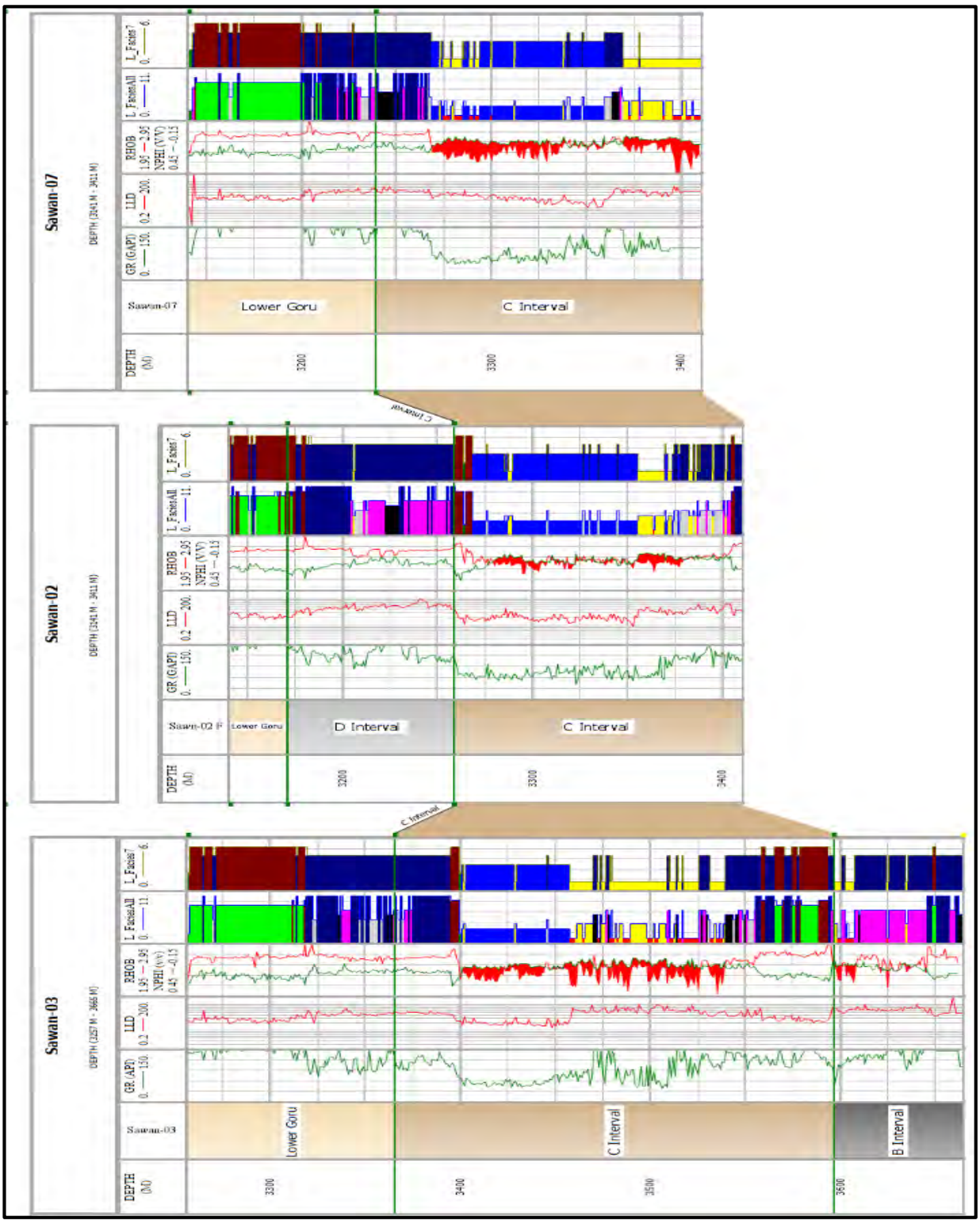


Figure 6.4: Well correlation showing K cluster analysis. Initially identified clusters are displayed in track adjacent to the porosity track (NPHI and RHOB). The grouped clusters which are classified facies are displayed in the last track of each well.

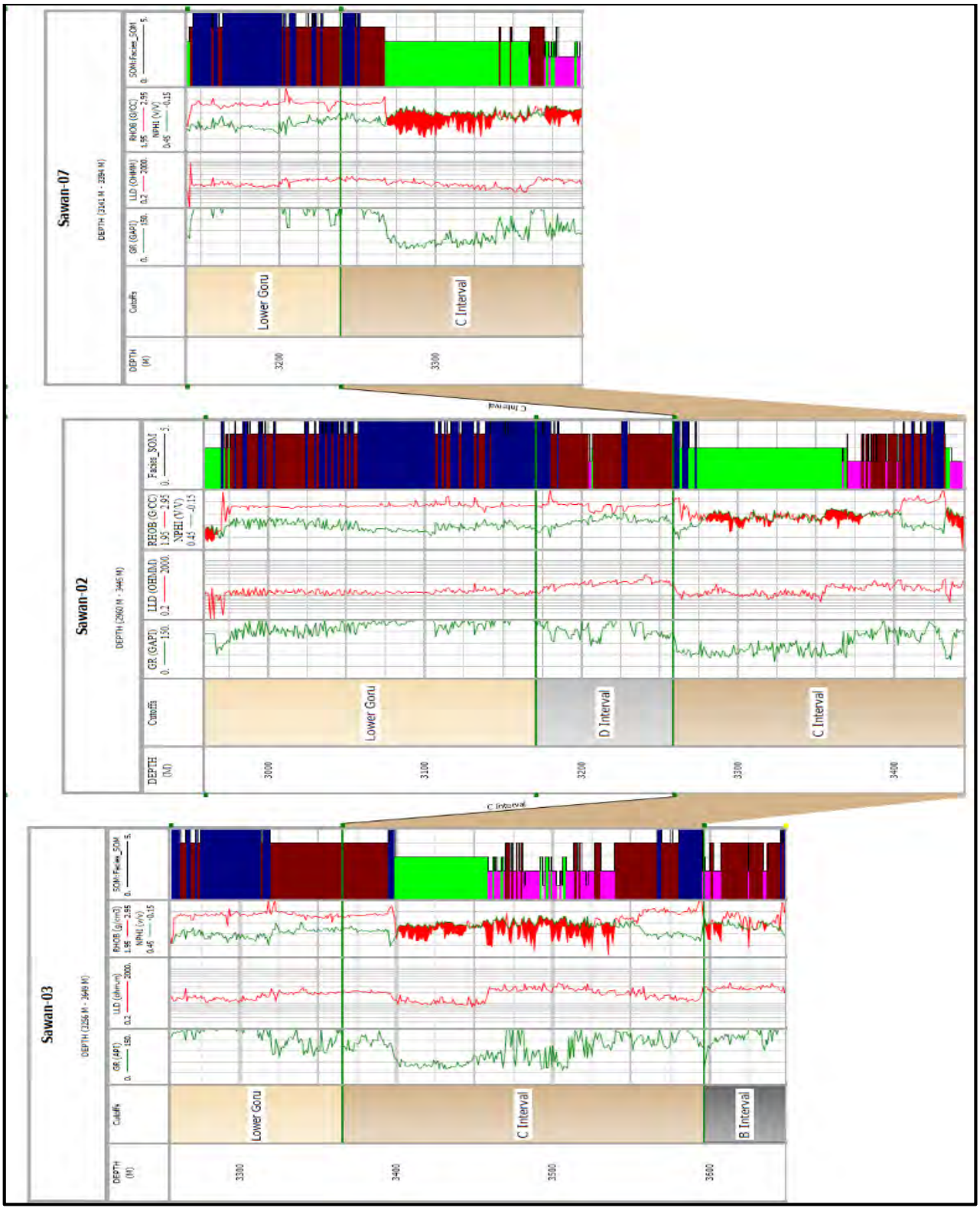


Figure 6.4: Well correlation showing facies modelled by using self-organizing map. The electrofacies are displayed in the last track of each well.

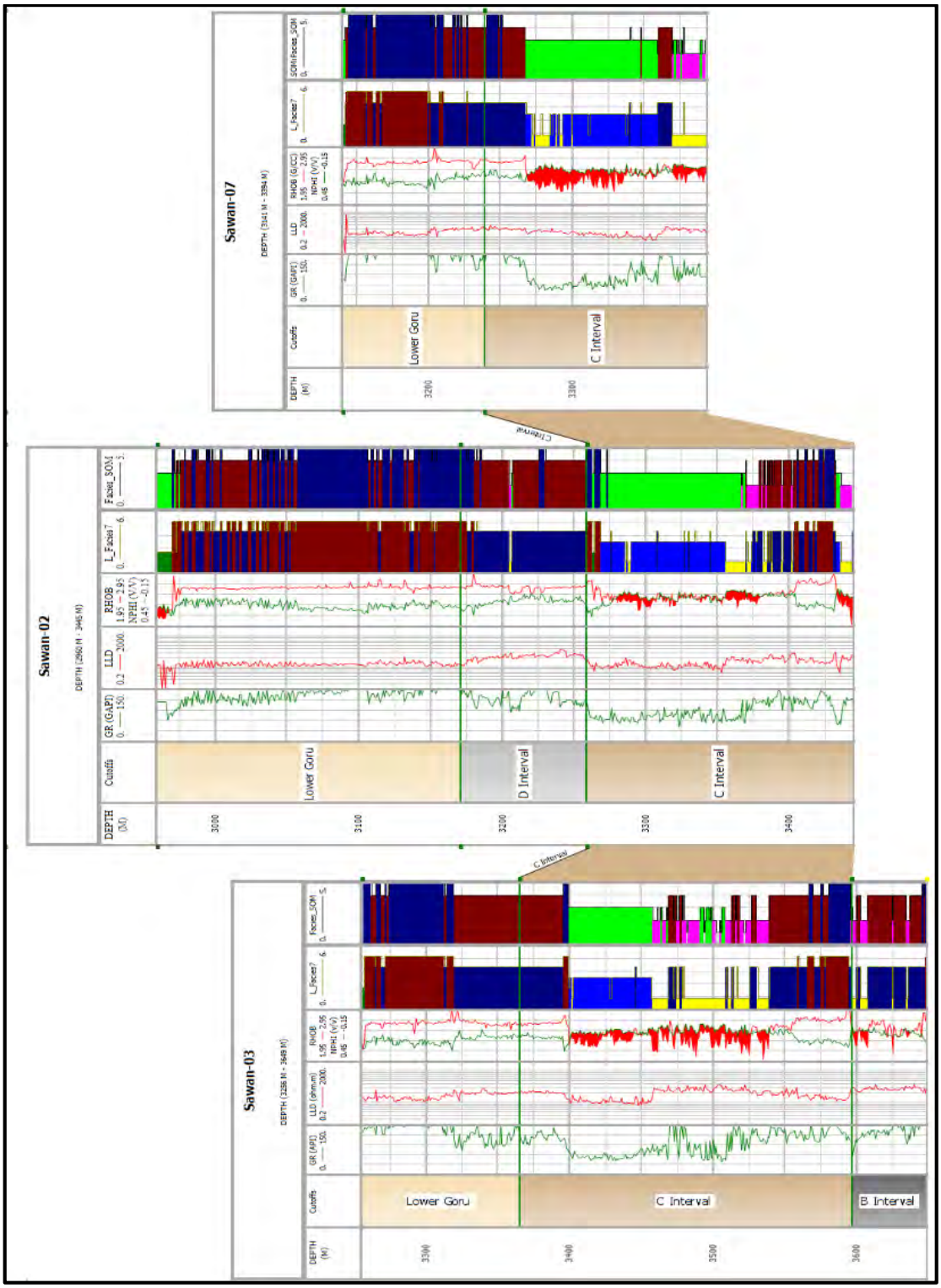


Figure 6.5: Well correlation showing K clustering and SOM comparison. Electrofacies classification obtained from K-means cluster analysis is given adjacent to porosity track (NPHI and RHOB) while SOM modelled facies are plotted in last track of each well.

Conclusion

Following are the conclusion drawn from the present research work.

- The formation evaluation of the Lower Goru Formation shows that it has potential to act as a prolific reservoir of Middle Indus Basin Pakistan. Lithological analysis carried out by using raw log curves confirms the lithological heterogeneity within the formation. The alternating layers of sandstone and shale forms the lithological make-up of the Lower Goru Formation.
- Petrophysical analysis of the Lower Goru Formation reveals that the C-interval of the Lower Goru Formation is the major contributing reservoir with good reservoir properties. The estimated volume of shale is 15%, effective porosity is 12% and water saturation is 20%.
- K-means cluster analysis efficiently marked the five facies within the C interval of the Lower Goru Formation. The major portion of C interval is comprised of two reservoir facies. The alternating layers of sandstone and shale are also marked by using cluster analysis.
- The results obtained by K-means clustering are further verified by using the self-organizing maps which displayed the well data into two-dimensional facies map.
- It is also evident from the present research work that the machine learning tools have ability to identify the facies at very small scale. So, these can be adopted to reduce the risk associated with hydrocarbon exploration.

Recommendations

In the light of conducted research work following are the recommendations for improving the efficiency of lithological identification during the reservoir characterization.

- Lithological identification of reservoir formation should be done efficiently to avoid misinterpretation.
- The identification of lithology performed by one method should be cross checked by using alternative methods to reduce error.
- Machine learning techniques should be employed in exploration workflow to enhance the efficiency.
- K cluster analysis gives the better classification of electrofacies. So, this method should be adopted for lithological identification at very fine scale.
- Multidimensional datasets of subsurface information should be visualized by using self-organizing map to give better understanding of subsurface.
- The data obtained from conventional workflows should be compared with the information obtained from machine learning tools to get better understanding of subsurface.

References

- 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. Pakistan. PAPG, ATC, 1, 85–104.
- Ali, A., & Sheng-Chang, C. (2020). Characterization of well logs using K-mean cluster analysis. Journal of Petroleum Exploration and Production Technology, 10(6), 2245–2256. doi:10.1007/s13202-020-00895-4
- Alpaydin, E. (2020). Introduction to Machine Learning (4th ed.). London, England: MIT Press.
- 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.
- Asquith, G. B., & Gibson, C. R. (1982). Basic well log analysis for geologists (Vol. 3). Tulsa: American Association of Petroleum Geologists.
- Aziz, O., Hussain, T., Ullah, M., Bhatti, A. S., & Ali, A. (2018). Seismic based characterization of total organic content from the marine Sembar shale, Lower Indus Basin, Pakistan. Marine Geophysical Research, 39(4), 491–508. doi:10.1007/s11001-018-9347-6
- Chang, H.-C., Kopaska-Merkel, D. C., & Chen, H.-C. (2002). Identification of lithofacies using Kohonen self-organizing maps. Computers & Geosciences, 28(2), 223–229. doi:10.1016/s0098-3004(01)00067-x
- Chow, J. J., Li, M.-C., & Fuh, S.-C. (2005). Geophysical well log study on the paleoenvironment of the hydrocarbon producing zones in the erchungchi formation, hsinyin, SW Taiwan. Terrestrial Atmospheric and Oceanic Sciences, 16(3), 531. doi:10.3319/tao.2005.16.3.531(t)
- Clavier, C., & Rust, D. H. (1976). Mid plot: A new lithology technique. The Log Analyst, (06).
- Cornish, R. (2007). Statistics: cluster analysis. Statistics: Cluster Analysis. Mathematics Learning Support Centre, 3, 1–5.
- Dar, Q. U. Z. 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. doi:10.1016/j.petlm.2021.09.003

- Darling, T. (2005). Well logging and formation evaluation. London, England: Elsevier Science.
- Doveton, J. H. (1994). Multivariate Pattern Recognition and Classification Methods: Chapter 4.
- El-Khadragy, A. A., Ghorab, M. A., Shazly, T. F., Ramadan, M., & El-Sawy, M. Z. (2014). Using of Pickett's plot in determining the reservoir characteristics in Abu Roash Formation, El-Razzak Oil Field, North Western Desert, Egypt. Egyptian Journal of Petroleum, 23(1), 45–51. doi:10.1016/j.ejpe.2014.02.007
- Estivill-Castro, V. (2002). Why so many clustering algorithms: A position paper. SIGKDD Explorations: Newsletter of the Special Interest Group (SIG) on Knowledge Discovery & Data Mining, 4(1), 65–75. doi:10.1145/568574.568575
- Euzen, T., & Power, M. R. (2012). Well log cluster analysis and electrofacies classification: a probabilistic approach for integrating log with mineralogical data. CSPG CSEG CWLS Convention.
- Fischetti, A. I., & Andrade, A. (2002). Porosity images from well logs. Journal of Petroleum Science & Engineering, 36(3–4), 149–158. doi:10.1016/s0920-4105(02)00292-9
- Glover, P. (2000). Petrophysics MSc course notes. Leeds, England: University of Leeds.
- Greengold, G. E. (1986). The graphical representation of bulk volume water on the Pickett crossplot. The Log Analyst, (03).
- Hancock, N. J. (1992). Quick-Look Lithology from Logs: Part 4. Wireline Methods. In N. J. Hancock (Ed.), ME 10: Development Geology Reference Manual. AAPG.
- Kattan, A., Jawad, W., & Jomaah, S. N. A. (2018). Cluster Analysis Approach to Identify Rock Type in Tertiary Reservoir of Khabaz Oil Field Case Study. Iraqi Journal of Chemical and Petroleum Engineering, 19(2), 9–13.
- Kazmi, A. H., & Jan, M. Q. (1997). Geology and tectonics of Pakistan. Karachi, Pakistan: Graphic Publishers.
- Khalid, P., Akhtar, S., & Khurram, S. (2020). Reservoir characterization and multiscale heterogeneity analysis of cretaceous reservoir in Punjab platform of middle Indus basin,

Pakistan. Arabian Journal for Science and Engineering, 45(6), 4871–4890. doi:10.1007/s13369-020-04443-4

- Kingston, D. R., Dishroon, C. P., & Williams, P. A. (1983). Global basin classification system. AAPG Bulletin.
- Kohonen, T. (2012). Self-Organizing Maps (2nd ed.). Berlin, Germany: Springer.
- Kohonen, T., & Honkela, T. (2007). Kohonen network. Scholarpedia, 2(1), 1568.
- Munir, K., Iqbal, M. A., Farid, A., & Shabih, S. M. (2011). Mapping the productive sands of Lower Goru Formation by using seismic stratigraphy and rock physical studies in Sawan area, southern Pakistan, a case study. Journal of Petroleum Exploration and Production Technology, (1), 33–42.
- Nazeer, K. A., & Sebastian, M. (2009). Improving the accuracy and efficiency of the k-means clustering algorithm.
- Nichols, G. (2013). Sedimentology and Stratigraphy (2nd ed.). Hoboken, NJ: Wiley-Blackwell.
- North, F. K. (1985). Petroleum Geology. London, England: Chapman and Hall.
- Obeida, T. A., Al-Jenaibi, F., Rassas, S., & Serag El Din, S. S. (2007). Accurate calculation of hydrocarbon saturation based on log-data in complex carbonate reservoirs in the middle-east. All Days. SPE.
- Qian, J., Nguyen, N. P., Oya, Y., Kikugawa, G., Okabe, T., Huang, Y., & Ohuchi, F. S. (2019). Introducing self-organized maps (SOM) as a visualization tool for materials research and education. Results in Materials, 4(100020), 100020. doi:10.1016/j.rinma.2019.100020
- Rider, M. H. (1996). The geological interpretation of well logs (2nd ed.). Houston, TX: Gulf Publishing.
- Sam-Marcus, J., Enaworu, E., Rotimi, O. J., & Seteyeobot, I. (2018). A proposed solution to the determination of water saturation: using a modelled equation. Journal of Petroleum Exploration and Production Technology, 8(4), 1009–1015. doi:10.1007/s13202-018-0453-4
- Schlumberger. (1989). Log interpretation principles/applications manual.
- Selley, R. C. (2000). Applied Sedimentology (2nd ed.). San Diego, CA: Academic Press.

- Senosy, A. H., Ewida, H. F., Soliman, H. A., & Ebraheem, M. O. (2020). Petrophysical analysis of well logs data for identification and characterization of the main reservoir of Al Baraka Oil Field, Komombo Basin, Upper Egypt. SN Applied Sciences, 2(7). doi:10.1007/s42452-020-3100-x
- Simandoux, P. (1963). Mesures dielectriques en milieu poreux, application a mesure des saturations en eau, Etude du Comportment des massifs Argileux. 193–215.
- Sonnenberg, S., & Selley, R. (2014). Elements of Petroleum Geology (3rd ed.). San Diego, CA: Academic Press.
- Tiab, D., & Donaldson, C. (2004). Petrophysics Second Edition-Theory and Practice of Measuring Reservoir Rock and Fluid Transport Properties.
- Villmann, T., & Bauer, H.-U. (1998). Applications of the growing self-organizing map. Neurocomputing, 21(1–3), 91–100. doi:10.1016/s0925-2312(98)00037-x
- Wandrey, C. J., Law, B. E., & Shah, H. A. (2004). Sembar Goru/Ghazij Composite Total Petroleum System, Indus and Sulaiman-Khirthar Geologic Provinces. Geological Survey Bulletin.
- Wang, Q., Wang, C., Feng, Z. Y., & Ye, J. F. (2012). Review of K-means clustering algorithm. Electronic Design Engineering, 20(7), 21–24.
- Zaigham, N. A., & Mallick, K. A. (2000). Prospect of hydrocarbon associated with fossil-rift structures of the southern Indus basin, Pakistan. Am Assoc Pet Geol Bull, 84(11), 1833–1848.