# SEISMIC DRIVEN THIN RESERVOIR FACIE CLASSIFICATION USING ADVANCED MACHINE LEARNING ALGORITHMS: A RESEARCH ON LOWER RANIKOT SANDSTONE RESERVOIR, KIRTHAR FOLDBELT, LOWER INDUS BASIN, PAKISTAN

**Umar Manzoor**

**01-286202-002**

**Bahria University, Islamabad**

# SEISMIC DRIVEN THIN RESERVOIR FACIE CLASSIFICATION USING ADVANCED MACHINE LEARNING ALGORITHMS: A RESEARCH ON LOWER RANIKOT SANDSTONE RESERVOIR, KIRTHAR FOLDBELT, LOWER INDUS BASIN, PAKISTAN


Umar Manzoor

01-286202-002

A thesis submitted in fulfilment of the requirements for the award of the degree of Doctor of Philosophy (Geophysics)

Department of Earth and Environmental Sciences

Bahria University, Islamabad

OCTOBER 2023

**APPROVAL FOR EXAMINATION**

Scholar's Name: Umar Manzoor

Registration No. 26259

Programme of Study: PhD Geophysics

Thesis Title: Seismic Driven Thin Reservoir Facie Classification Using Advanced Machine Learning Algorithms: A Research on Lower Ranikot Sandstone Reservoir, Kirthar Foldbelt, Lower Indus Basin, Pakistan.

It is to certify that the above scholar's thesis has been completed to my satisfaction and, to my belief, its standard is appropriate for submission for examination. I have also conducted plagiarism test of this thesis using HEC prescribed software and found similarity index _____% that is within the permissible limit set by the HEC for the PhD degree thesis. I have also found the thesis in a format recognized by the BU for the PhD thesis.

Principal Supervisor's Signature_____

Date: ___12-10-2023_____

Name: Dr. Muhsan Ehsan

## AUTHOR'S DECLARATION

I, Umar Manzoor, hereby state that my PhD thesis titled "Seismic Driven Thin Reservoir Facie Classification Using Advanced Machine Learning Algorithms: A Research on Lower Ranikot Sandstone Reservoir, Kirthar Foldbelt, Lower Indus Basin, Pakistan" is my own work and has not been submitted previously by me for taking any degree from this university Bahria University, Islamabad or anywhere else in the country/world. At any time if my statement is found to be incorrect even after my graduation, the University has the right to withdraw/cancel my PhD degree.

Name of scholar: Umar Manzoor

Date: 12-10-2023

**PLAGIARISM UNDERTAKING**

I, solemnly declare that research work presented in the thesis titled "Seismic Driven Thin Reservoir Facie Classification Using Advanced Machine Learning Algorithms: A Research on Lower Ranikot Sandstone Reservoir, Kirthar Foldbelt, Lower Indus Basin, Pakistan." is solely my research work with no significant contribution from any other person. Small contribution/help wherever taken has been duly acknowledged and that complete thesis has been written by me.

I understand the zero tolerance policy of the HEC and Bahria University towards plagiarism. Therefore, I as an Author of the above titled thesis declare that no portion of my thesis has been plagiarized and any material used as reference is properly referred/cited.

I undertake that if I am found guilty of any formal plagiarism in the above titled thesis even after award of PhD degree, the university reserves the right to withdraw/revoke my PhD degree and that HEC and the University has the right to publish my name on the HEC / University website on which names of scholars are placed who submitted plagiarized thesis.

Scholar / Author's Sign: _____

Name of the Scholar: Umar Manzoor

## DEDICATION

To my loving parents, wife, and kids

## ACKNOWLEDGEMENTS

Thank you to Almighty Allah for all of His blessings and for providing me with the strength to accomplish this study project. I will be eternally grateful to my parents for their prayers and encouragement throughout my life. A special thanks to my wife for her consistent and unwavering support, which enabled me to complete this manuscript. Thank you especially to my children for their patience and understanding.

First and foremost, I would like to express my sincere gratitude to my supervisor, Dr. Muhsan Ehsan, for his invaluable guidance, support, and encouragement throughout my Ph.D. journey. His expertise, wisdom, and patience have been instrumental in my success. A special thanks to Mr. Atif Rais Khan, Mr. Iftikhar Atif Khan, Mr. Sohail Rashid, Mr. Muhmmad Khizer Iftikhar, Dr. Urooj Shakir, Mr. Bilal Ahmed Malik, Ms. Hina Wajid Satti and Mr. Farooq Arshad for helping me out during the research work.

I would like to express my sincere gratitude to my colleagues Dr. Muyyassar Hussain and Dr. Zahid Ullah Khan for their unwavering support throughout this Ph.D. journey. Their encouragement and motivation kept me going during the most challenging times and helped me maintain my focus and determination toward completing this noble task.

I am grateful to the Directorate General of Petroleum Concessions, Ministry of Energy (Petroleum Division), for providing the dataset to carry out this study. I would also like to thank LMK Resources (GVERSE), and GeoSoftware (HampsonRussell), for providing the necessary software and technical assistance needed to complete this research work. AI tools were used only to polish the English language of the thesis and not for scientific writing.

I am also thankful to Higher Education Commission (HEC) as this study is funded by the HEC, Pakistan under Grant No. 20-14925/NRPU/R&D/HEC/2021/2020.

## ABSTRACT

This study addresses the crucial challenge of characterizing thin gas sand reservoirs in the Lower Ranikot/Khadro Formation of Pakistan's Lower Indus Basin, a reservoir with varying thicknesses 4 to 7 m below seismic resolution. Previous studies have struggled to produce precise results due to reservoir heterogeneity, data limitations, and associated uncertainties. An optimized, integrated approach, combining seismic attributes, petrophysical properties, advanced machine learning (*ML*) algorithms, and continuous wavelet transform (*CWT*) addresses thin gas sand facies and pore pressure challenges comprehensively. Among several employed ML algorithms gradient boosting regressor (*GBR*) accurately predicted thin sands (>90%), reducing uncertainty in hydrocarbon-bearing sand distribution. A delicate *ML* approach has been broadly applied to analyze the potential and robustly interpret well-logs while addressing the associated challenges. Support vector machine (One-class-*SVM*) helps to reduce outliers with great certainty while the missing log's sonic and density are precisely predicted via *GBR* and extra tree regressor (*ETR*) with the highest $R^2$ respectively. Likewise, random forest regressor (*RFR*) performed exceptionally well for water saturation modeling expressing the highest 0.93 correlation among *ML* and conventional results. Finally, the decision tree classifier (*DTC*) modeled reservoir facies with the best 91% accuracy and 0.935 F1 measures at the blind well. Additionally, an optimized workflow generates high-frequency acoustic impedance synthetics by utilizing a deep neural network (*DNN*) integrated with *CWT* components at the reservoir level vis-a-vis validating the results with existing geological facies to resolve thin beds without introducing noise. The shale layers of the formation are quite problematic and complex geological variations exhibit pore pressure discrepancy making drilling operations crucial. Among all conventional methods for pore pressure prediction, *GBR* integrated with *CWT* has provided very good results after validation. The study characterizes reservoirs below seismic resolution, enabling more efficient resource exploration and development. It outperforms previously done conventional approaches by delivering higher accuracy, reducing uncertainty, and unlocking valuable insights using advanced *ML* and *CWT* techniques. It offers broad applicability to other complex, thin-bed reservoirs worldwide, optimizing field development and maximizing hydrocarbon recovery.

**Key Words.** Machine learning, Continuous wavelet transform, Conventional neural network, Reservoir Characterization, Thin-heterogeneous sand, Pore pressure

# TABLE OF CONTENTS

# LIST OF TABLES

## LIST OF FIGURES

## LIST OF ABBREVIATIONS

| | | |
|---|---|---|
| *1D* | - | One Dimensional |
| *2D-3D* | - | Two Dimensional -Three Dimensional |
| *ACE* | - | Alternating Conditional Expectations |
| *AI* | - | Acoustic impedance |
| *APHI* | - | Average porosity |
| *AUC* | - | Area Under Curve |
| *AVO* | - | Angle Versus Offsets |
| *BMU* | - | Best Matching Units |
| *CALI* | - | Caliper Log |
| *CMP* | - | Common Midpoint Gathers |
| *CWT* | - | Continuous Wavelet Transformation |
| *DBN* | - | Deep Belief Networks |
| *DGPC* | - | Directorate General of Petroleum Concession |
| *DHI* | - | Direct Hydrocarbon Indicators |
| *DL* | - | Deep Learning |
| *DNNs* | - | Deep Neural Networks |
| *DPHI* | - | Density porosity |
| *DT* | - | Sonic Compressional Wave/ Decision Tree |
| *DTC* | - | Decision Tree Classifier |
| *DTR* | - | Decision Tree Regressor |
| *ETR* | - | Extra Tree Regressor |
| *FL* | - | Fuzzy Logic |
| *GAM* | - | Generalized Additive Models |
| *GBR* | - | Gradient Boosting Regressor |
| *GBR* | - | Gradient Boosting Regressor |
| *GI* | - | Geophysical Inversion |
| *GL* | - | Genetic Algorithms |
| *GR* | - | Gamma ray |
| *GWC* | - | Ground Water Contact |
| *HFSSV* | - | High-Frequency Synthetic Seismic Volume |

| | | |
|---|---|---|
| *HI* | - | Hydrogen Index |
| *IQR* | - | Interquartile Range |
| *K/T* | - | Cretaceous/Tertiary |
| *KFTB* | - | Kirthar Fold and Thrust Belt |
| *LGBM* | - | Light Gradient Boosting Machine |
| *LLD* | - | Laterolog Deep |
| *LLS* | - | Laterolog Shallow |
| *LSSVM* | - | Least Squares Support Vector Machines |
| *MAE* | - | Mean Absolute Error |
| *MAPE* | - | Mean Absolute Percentage Error |
| *MCC* | - | Matthews Correlation Coefficient |
| *ML* | - | Machine Learning |
| *MMBBLS* | - | Million Barrels of Oil |
| *MMT* | - | Main Mantle Thrust |
| *MSFL* | - | Micro Spherical Focused Log |
| *NCTL* | - | Normal Compaction Trend Line |
| *NNET* | - | Neural Networks |
| *NPHI* | - | Neutron porosity |
| *NTG* | - | Net-to-Gross Ratio |
| *OBG* | - | Overburden Gradient |
| *PCA* | - | Principal Component Analysis |
| *PHIE* | - | Effective porosity |
| *PHIE* | - | Effective porosity |
| *PMBSI* | - | Post-Stack Model Based Inversion |
| *PNN* | - | Probabilistic Neural Network |
| *PP* | - | Pore pressure |
| *PSI* | - | Pounds Per Square Inch |
| *PSO* | - | Particle Swarm Optimization |
| *PSSI* | - | Pre-Stack Simultaneous Seismic Inversion |
| *Q* | - | Mud Flow Rate |
| *Q1* | - | First Quartile |
| *Q3* | - | Third Quartile |
| *QC* | - | Quality Control |

| | | |
|---|---|---|
| *RC* | - | Reflection Coefficient |
| *RES* | - | Resistivity Log |
| *ResNET* | - | Residual Network |
| *RF* | - | Random Forest |
| *RTR* | - | Random Tree Regressor |
| *RHOB* | - | Bulk Density Log |
| *RMS* | - | Root Means Square |
| *RMSE* | - | Root Mean Squared Error |
| *ROP* | - | Rate of Penetration |
| *RPM* | - | Rock Physics Modeling |
| *RPT* | - | Rock Physics Template |
| *RS* | - | Rotation speed |
| $R_w$ | - | Water Resistivity |
| *SOM* | - | Self-Organizing Maps |
| *SPHI* | - | Sonic porosity |
| *SPP* | - | Standpipe Pressure |
| *SVM* | - | Support Vector Machines |
| *SVR* | - | Support Vector Regression |
| $S_w$ | - | Water Saturation |
| *TCF* | - | Trillion Cubic Feet |
| *TOC* | - | Total Organic Carbon |
| *TVD* | - | True Vertical Depth |
| *VCL* | - | Volume of Clay |
| $V_p$ | - | P-wave Velocity |
| $V_s$ | - | S-wave Velocity |
| $V_{SH}$ | - | Volume of Shale |
| $Z_p$ | - | P-impedance |
| $Z_s$ | - | S-impedance |

## LIST OF APPENDICES

# CHAPTER 1

# INTRODUCTION

The analysis and characterization of reservoirs are crucial in developing an accurate reservoir model for flow simulation, as well as assessing the field's performance and improving hydrocarbon productivity while reducing costs. The reliability of the reservoir model is influenced by the availability and accuracy of data, as well as the collection, interpretation, and assignment methods used in reservoir characterization. In reservoir modeling, uncertainties can be reduced by improving the reservoir description through the integration of various available data. This includes static data such as geological information obtained from laboratory analyses and geologic studies, as well as dynamic fluid flow data like field production. In reservoir simulation, particularly for enhanced oil recovery, it is critical to ensure that the heterogeneity of petrophysical properties reflects field data for accurate flow simulation. This is necessary to generate a reliable reservoir model that can accurately predict reservoir behaviour and enhance oil recovery (Teh et al., 2012). Several data correlation and integration techniques, along with assessments of uncertainty, have been proposed in recent times to improve the estimation of reservoir properties, including porosity, permeability, and saturation. By using these techniques, the reliability of the reservoir model can be improved, resulting in more accurate predictions and reduced uncertainty. Integrated data models are more predictive in history matching compared to traditional approaches, as they maintain geoscience continuity (Huang et al., 2011).

Conventional seismic interpretation techniques were developed over four decades ago for exploratory plays to describe the reservoir by detecting hydrocarbon seismic responses (Maurya et al., 2020). These seismic amplitude responses were based on procedures that were based on Post-Stack seismic amplitude analysis and were suggestive of direct hydrocarbon indicators (*DHI*). *DHI* plays are defined as shallow, high-porosity reservoirs with much lower acoustic impedance than the strata around them. These hydrocarbon-filled high porosity reservoirs seismic signatures might include aberrant high-amplitude reflections known as "bright spots". Nowadays, the oil industry is focused on exploring and developing deeper and more complex reservoirs. The traditional seismic reservoir characterization techniques show limitations in such scenarios; hence, advanced research and development methods are considered for better reservoir understanding (Nanda, 2021).

Petrophysical analysis involves the examination of the physical and chemical properties of rocks and sediments, particularly their capacity to store and transport fluids like oil, and gas. This analysis plays a crucial role in the exploration and production of hydrocarbons and the evaluation of water resources. It involves utilizing various types of data, such as well logs (including resistivity, porosity, and sonic logs), core data, seismic data, and other relevant information, to comprehend the physical properties of rocks and the distribution of fluids within them (Ali et al., 2021). The petrophysical analysis aims to assess the hydrocarbon potential and fluid flow properties of subsurface rocks (Amigun et al., 2012). Petrophysical analysis can be broadly classified into two types: conventional and unconventional methods. Conventional petrophysical analysis is a traditional approach that evaluates the physical and chemical properties of subsurface rocks to assess their ability to store and transmit fluids, particularly hydrocarbons. On the other hand, unconventional petrophysical analysis is used to evaluate the physical and chemical properties of subsurface rocks that have smaller, more dispersed reservoirs of hydrocarbons and are referred to as unconventional resources (Aghli et al., 2016). The unconventional petrophysical analysis is typically applied to resources like shale gas and tight oil. Although both types of petrophysical analysis aim to determine the physical and chemical properties of subsurface rocks, their focus and methodology differ depending on the type of resource being evaluated (Muther et al., 2022).

Since the 1970s, acoustic impedance has become a fundamental quantity that is still utilized in seismic reflectivity inversion and interpretation. The multiplication of density and acoustic velocity is commonly referred to as acoustic impedance and it is closely related to various rock parameters (Barclay et al., 2008). Unlike seismic reflectivity, which occurs at the intersections of different strata, acoustic impedance takes on consistent values within rock layers, providing a simple and clear link with geology and stratigraphy, but has a temporal limit, whereas the corresponding wells in the vicinity allow for the injection of higher resolution data with limited spatial coverage. Therefore, integrating seismic and well data is crucial for establishing a Petro-elastic relationship that can enhance reservoir characterization. The density, *P-wave* velocity ($V_p$), and *S-wave* velocity ($V_s$) of the earth's subsurface, which determine its elastic characteristics, are often modeled using an inversion approach (Mahgoub et al., 2018). However, this simple technique requires extensive computation power and time in order to resolve reservoirs deposited within complex geological settings. The use of advanced *ML* algorithms can help address the complexities involved in characterizing reservoir. Hence,

in recent times, oil and gas industries are rapidly adopting *ML* based solutions to help characterize such complex reservoirs (Liu, 2017).

Rock physics is a field of geophysics that focuses on the study of the physical properties of rocks and their correlation with geological structures. It involves the study of how rock properties such as density, elasticity, and porosity respond to changes in temperature, pressure, and fluid saturation, and how they can be used to infer information about subsurface geology. Rock physics is a vital field in the petroleum industry, with applications including hydrocarbon exploration and production, as well as reservoir characterization (Ghosh et al., 2018; Jensen et al., 2016). Rock physics is based on the idea that the physical properties of rocks, such as density, elasticity, and porosity, are dependent on the rock's composition, structure, and geological environment. Geophysicists can learn about the subsurface geology and predict how rocks will respond to various stimuli, such as changes in temperature, pressure, or fluid saturation by analyzing these properties and their changes under different conditions (Sohail and Hawkes, 2020; Avseth and Skjei, 2011). Rock physics helps to quantify important properties of a reservoir, including not only seismic data such as velocity and reflection amplitude, but also critical well log data like porosity, lithology, pore fluid type, saturation, and *PP* (Karimpouli et al., 2018; Sayers and den Boer, 2011). This quantitative analysis is crucial for the exploration and production of hydrocarbons. Well logs have a higher temporal (vertical) resolution, making them more accurate and consistent for inferring true reservoir properties in situ (Dou et al., 2009).

Reservoir characterization is the procedure of gathering and evaluating data to identify the physical properties of a reservoir, including its geology, fluid content, and potential for production (Muther et al., 2022). In the context of reservoir characterization, *ML* techniques can be used to analyze large datasets and extract insights that might not be readily apparent through traditional methods. This can help to improve predictions about reservoir performance and increase our understanding of subsurface reservoirs. Some common applications of *ML* in reservoir characterization include seismic interpretation, well log analysis, and production forecasting (Anifowose et al., 2017).

In very recent times, *ML* has taken over the world of technology by the swarm. Every industry has incorporated complex algorithms to optimize and enhance previously incorporated workflows. Similarly, in the world of geoscience, *ML* employs advanced computer algorithms to assist geoscientists in comprehending the links between enormous volumes of geophysical or geological data or information with the complex

subsurface geological settings at the reservoir level (Mustafa et al., 2022; Radwan et al., 2022; Ali et al., 2021; Ashraf et al., 2021; Rajabi et al., 2021; Naseer, 2020a; Gorain and Thakur, 2015). The computer algorithm is trained using input data and then adapts autonomously to deliver repeatable and trustworthy quantitative seismic interpretation findings. The two approaches *ML* and geophysical inversion (*GI*) work on principle based on subsurface knowledge and feature selection to understand reservoir potential by integrating the sciences of statistics and physical domains (Reading et al., 2015) *ML* techniques refer to a set of algorithms that can turn data into meaningful insights, without relying on physics-based models. While they hold tremendous potential, *ML* also has its limitations. For instance, it often struggles to accurately predict outcomes outside the range of data it has been trained on (Lantz, 2019). Moreover, dealing with large amounts of data, which may contain irrelevant or redundant information, poses a significant challenge for *ML* applications (Géron, 2022; Lantz, 2019). Datasets used in *ML* often contain obscure or redundant information that can have a negative impact on the performance of the algorithms, even though this issue is not unique and *ML* can handle it in an optimized way by removing noise (Wuest et al., 2016). *ML* methods can provide practical alternatives, as these are more precise than physics-based solutions. Furthermore, most *ML* applications require a large number of features to reproduce the output response more effectively but might raise three primary concerns. First, the higher the dimensionality of the feature space that becomes prone to extrapolation. Second, the need for a large number of features results in an exponential increase in computational resources required. Third, these techniques are more susceptible to overfitting due to reliance on larger parameter spaces. Another significant concern with *ML* applications is the realism of the predictions (Nwachukwu et al., 2018).

Types of *ML* systems are classified based on the amount of supervision with which they are trained, including supervised, unsupervised, and semi-supervised learning. In supervised learning, the algorithm is trained on a set of labelled data to generate a function that can map inputs to target variables. Examples of algorithms in this category include linear regression, K-nearest neighbours, and random forest (*RF*) (Chen et al., 2020; Zhang et al., 2020). In unsupervised learning, the training data is completely unlabelled, and the goal is to identify natural groupings in the dataset, such as K-means, hierarchical clustering, and the apriori algorithm (Liu et al., 2019). Semi-supervised learning deals with partially labelled data where the amount of unlabelled data is significantly greater than the amount of labelled data (Chapelle et al., 2006). Clustering

strategies are typically employed with limited guidance, and some techniques such as deep belief networks (*DBN*) combine supervised and unsupervised algorithms (Huang et al., 2017). However, this simple technique requires extensive computation power and time to resolve reservoirs deposited within complex geological settings. Advanced *ML* algorithms have emerged as an efficient means of addressing these complexities. As a result, the oil and gas industry has increasingly turned to *ML* based solutions to characterize complex reservoirs (Liu, 2017).

Deep Learning *(DL)* is a branch of Artificial Intelligence that employs neural networks (*NNET*), featuring multiple layers to discover intricate patterns and correlations within datasets. In geophysics, *DL* has been found to be useful in a variety of applications, particularly in seismic and well data analysis (Korjani et al., 2016). Seismic data, which is collected through the study of seismic waves, can be used to create subsurface images of geological structures and is a crucial tool for the exploration and production of hydrocarbons. Applications of *DL* algorithms include improving the resolution of seismic images and enhancing their interpretation. In well data analysis, *DL* is used to analyze logs, core data, and other well measurements to better understand the subsurface geology and make more informed decisions about drilling and production (Zhang and Alkhalifah, 2020).

The Continuous Wavelet Transformation *CWT* is a method of analysing signals based on their frequency and scale, using mathematical principles. This characteristic makes it a useful technique in a variety of applications such as signal processing, time-frequency analysis, and image processing. It is an extension of the wavelet transform that allows for the analysis of signals with infinite length, as opposed to signals of finite length analyzed using the discrete wavelet transform (Naseer, 2020b). The *CWT* is a mathematical tool that can be used to analyze signals in both the frequency and time domains simultaneously. It can be applied in various fields, such as signal processing, image processing, and time-frequency analysis, as it provides information on the frequency content and temporal localization of the signal. *CWT* is widely used in geoscience for analyzing signals and images in various geological and geophysical applications, especially in seismic and well data, reservoir characterization and monitoring and evaluation application. In seismic data analysis, the *CWT* can be used to analyze seismic data to detect and analyze subsurface geological structures, including faults, salt domes, and hydrocarbon reservoirs (Jayaswal et al., 2021). The *CWT* can also be used to improve seismic imaging by suppressing noise and enhancing signals. In the

context of well data analysis, *CWT* can be utilized to examine various logs, including resistivity, density, and sonic logs, to determine geological features beneath the earth's surface, such as lithology, fluid saturation, and porosity. In reservoir characterization, *CWT* can be used to analyze both seismic and well data to better understand subsurface geological features and properties, such as lithology, porosity, and fluid saturation. As concerned for monitoring and evaluation, it is used to monitor and evaluate production from oil and gas wells by analyzing time-delay seismic data and well data to identify changes in subsurface geological features and fluid levels (Naseer, 2020b)

In this research an integration of *ML* and *CWT* is focused by using the seismic and well data to resolve the thin sands of Ranikot/Khadro Formation of the study area successfully. This approach proved robust, state of the art and economical as compared to the conventional approaches for reservoir characterization with optimized results.

## 1.1 Research Area

The research area is located in the eastern portion of Kirthar Fold and Thrust Belt (*KFTB*). *KFTB* is spread roughly an area of 60,000 $km^2$ in the west side of the frontal *KFTB,* eliminating the Kirthar mountain front in the east where Bhit, Badhra, Zamzama and Mehar gas fields are located. The *KFTB* spans from Khojak Segment and Sibi Trough in the north to the Karachi Arc and Porali Trough in the south. The *KFTB* area is surrounded by Zarghun and Ziarat fields in the north and Mazarani, Mehar, Zamzama, Bhit and Badhra fields in the east (Khan et al., 2011).

The Zamzama Gas Field is situated in the Lower/Southern Indus Basin, in the Kirther Foredeep, which spans around 120 $km^2$, and is considered the fourth largest gas field found in Pakistan. It is located about 200 km North of Karachi, and approximately 8 km west of the existing Sui-Karachi pipeline as shown in Figure 1.1. The neighbouring Bhit gas field is located to the southwest, while the Kadanwari, Sawan, and Miano fields are situated to the northeast of Zamzama. On its northern side, the Mehar wells can be found (Qureshi et al., 2021).

The secondary Palaeocene sandstone reservoir of Ranikot/Khadro Formation is focussed in this research due to challenges regarding its proper thickness, heterogeneous nature, and deprived quality (Jackson et al., 2004). Khadro Formation is conformably lying over the major Pab sandstone reservoir in the Zamzama gas field that has produced enormous gas and now facing depletion. Therefore, a detail assessment of Khadro

Formation is necessary for the optimized field production as it has produced from couple of wells (Zamzama-03 and Zamzama-05) in the field.

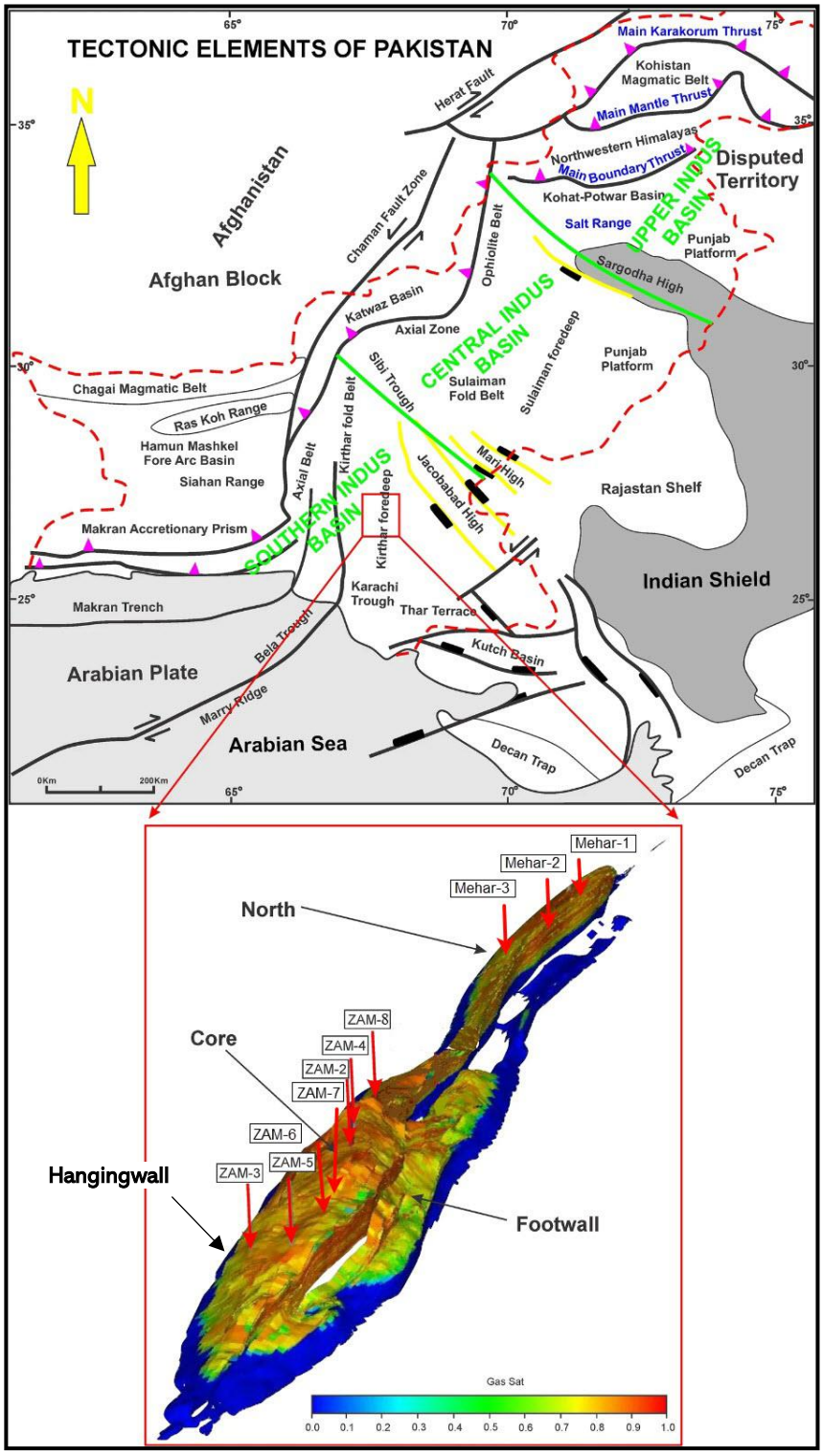



**Figure 1.1** Regional tectonic map of Pakistan along with division of basin. The area of interest is shown by red box. The subset part of picture shows the complete structure style of Zamzama Field along with the foot and hanging wall structure. The Mehar wells lies on the norther side of structure geometry (Khan et. al, 2021).

## 1.2 Significance of Study

This research has a direct impact on the energy and the petroleum upstream sector of Pakistan. Focusing on the petroleum sector, the primary reservoir in this region was previously Pab Formation, which is now mostly depleted, and only small pockets of gas or condensate are being discovered. This mostly is economically non-viable for most companies.

It was observed during production and development of the fields that other reservoirs such as Ranikot/Khadro Formation have proven to produce handsome volumes of hydrocarbons and were previously not among the primary focus for explorations; due to its limited thickness of deposition which is 7 m to 15 m average. However, with advancements in technology, it is now easier to delineate such kinds of thin beds, which are below seismic resolution. Despite the thin vertical thickness of the reservoir beds, the lateral extension of this thin bed reservoir extends basin-wide, hence has a huge potential for exploration.

Identifying the regions for plausible areas of the hydrocarbon bearing facies, shall not only help our petroleum industry flourish but also attract international companies to come to this region for further exploration and development.

Secondly, the most important issue in Pakistan during these trying times is the energy crisis. The government has declared an emergency in this sector to help overcome this issue as soon as possible, leading to a smooth run for the different sector industries of Pakistan. Exploring gas from such thin reservoirs using the advanced technologies available today, shall not only prove to be a key pillar for the economy of Pakistan but also has the potential alone to affix the dreadful energy crisis prevailing today.

Moreover, the successful application of the technology on the thin reservoirs of Lower Ranikot/Khadro Formation can later be extended across all sedimentary basins of Pakistan. Having the most complex geology in the world, Pakistan faces a similar challenge in all the sedimentary basins to help delineate such thin beds. Allowing to work on this reservoir can help exploit the very similar issues being faced by the E&P companies, and this can drastically impact the economics of the petroleum industry.

## 1.3 Scope of Study

The research has included three phases, initiating from the first phase that included data purchase related to reports of core and well-cutting analysis. During this

first phase conventional geophysical interpretation approach has used to analyze the subsurface geology. Followed by the second phase, where detailed petrophysical interpretation and lithofacies identification of Lower Ranikot/Khadro Formation has been carried out. Whereas the final stage of the study has look into the application of advanced seismic inversion techniques to highlight potential sweet spots within Lower Ranikot/Khadro Formation across the entire region. These steps have been performed on the seismic cube, with well coverage, consisting of complete log suite data. The results obtained from this study has helped to devise and plan for further development or exploration wells in the recommended areas. This study has proved that the application of advance integrated *ML* approach is the recommended interpretative solution for such complex geological areas, which contain thin reservoirs, especially in Pakistan. The key steps included in this study are as follows:

I. Seismic and well data preparation
II. Seismic interpretation and data conditioning
III. Well logs conditioning
IV. Well logs calibration and wavelet extraction
V. Rock physics modelling for wells
VI. Wedge modelling
VII. Relative impedance inversion
VIII. Low frequency modelling
IX. Simultaneous inversion
X. Stochastic inversion

*ML* driven seismic geostatistical inversion has improved results that helped to identify between the reservoir and non-reservoir zones along with any plausible tuning effects generated by thin-bed reservoirs. This technique has also assisted in delineation the facie distribution across the study area while demarking the potential sweet spots within the reservoir under study.

### 1.4 Research Questions

There are only 2% of the wells in Lower Indus Basin that are drilled till the reservoir under study (i.e., Ranikot/Khadro Formation) (Ehsan et al., 2018). Hence, it is vital for geoscientists to explore and develop this formation potential. The aim of this study is to identify potential reservoirs that have promising prospects within the Lower

Indus Basin, which supports the integration of petrophysical and elastic properties, focusing to perform reservoir characterization of the Lower Ranikot/Khadro Formation thin sands reservoir deposited within the *KFTB* region. Following questions and challenges have been comprehensively answered in the research:

I. Thin reservoir, below tuning thickness.
II. Non-availability of complete dataset within study area.
III. Non-availability of core analysis reports for key wells used in the study.
IV. Wellbore stability with the help of *PP* prediction.

## 1.5 Problem Statement

The key problems present in study area that are addressed in the research:

I. Structural complexities of the anticline.
II. Resolution limitation of seismic data for thin and heterogenous Khadro/Ranikot Formation.
III. Separation of plausible sands from the non-reservoir for optimized field production.

## 1.6 Research Objective

Objectives of the study have focused on:

I. To build a relationship between elastic properties (impedances, $V_p/V_s$ ratio, etc.) and petrophysical properties (porosity, water saturation ($S_w$), etc.) to determine sweet spots.
II. To generate facies and fluid probabilities volumes for quantitative interpretation.
III. To analyze the lateral porosity distribution across the field, using the artificial neural networking technique, and correlate with the core analysis reports purchased.

Apart from the primary objectives, this study has achieved the following sub goals:

a. Improve the usefulness of seismic data by strengthening the physical basis of the use of inverted properties.
b. Well data conditioning and calculate essential petrophysical properties (e.g., porosity and saturation).
c. Build petro-elastic models to predict elastic properties ($V_p$, $V_s$, Density, etc.).

d. Determine the expected seismic and petro-elastic parameter response to hydrocarbon, lithological, and porosity variations.
e. Estimation of Net-to-Gross (N/G) ratio using a quantitative method calibrated with log-derived seismic rock properties.
f. Lithofacies estimation from inverted seismic attributes and respective well log data to quantify the uncertainty in the seismic lithology prediction (hydrocarbon saturation).

## 1.7 Adopted Research Techniques Significance and Importance

In recent years, *ML* has become increasingly essential in reservoir characterization due to reduced cost and time. Large datasets of seismic and wells can be used to train *ML* algorithms to uncover patterns and relationships that would be difficult or impossible to detect using conventional methods. *ML* improves accuracy and resolution while reducing uncertainty even when the data is incomplete or noisy. The conventional reservoir characterization approach is more laborious and time consuming, it is usually used pre-stack gathers a lot of work to complete the procedure, whereas Post-Stack seismic data has done the same work thoroughly. Similarly, this simple and comprehensive strategy can be used in reservoirs with the same challenges all over the world.

## 1.8 Integration of multi-scale datasets

The data sets employed in the research work mainly comprised of high frequency well logs and band-limited seismic data obtained from the Directorate General of Petroleum Concession (*DGPC*). The 3D seismic data of Zamzama Gas Field (15 square kilometers) is integrated with seven wells lying within this seismic cube for structural mapping and stratigraphic distribution of properties while three wells from Mehar gas field is employed for 1D assessment of properties across region (figure-1.2). Therefore, 10 wells are utilized to conduct the research work such that seven from Zamzama block and the remaining three wells from Mehar block. Table 1.1 & 1.2 shows the seismic data and names of well used in studies. The *ML* approaches comprehensively optimized the frequency content of seismic data through advance non-linear relationships with wells and assess the tuned thin sands more profoundly. The heterogeneities of potential sands are evaluated through enhanced elastic and petrophysical properties throughout the field.

**Table 1.1:** Details of seismic data that are utilized for seismic driven thin reservoir classification.

| **Seismic Data** |
|---|
| *3D* Post-Stack seismic data of Zamzama Block covering the area of 15 square kilometers along with partial angle stacks (Near, Mid and Far Angle stacks). |

**Table 1.2:** Details of wells data that are utilized for research purpose.

| **Well Data** | |
|---|---|
| Zamzama-02 | Zamzama-07 |
| Zamzama-03 | Zamazam-08-ST-02 |
| Zamzama-04 | Mehar-01 |
| Zamzama-05 | Mehar-02 |
| Zamzama-06 | Mehar-03 |

## 1.9 Thesis Outline

The thesis comprises the following components for assessment of potential sands of Ranikot/Pab Formation.

I. Detail structural and stratigraphic analysis of the study area.
II. Evaluate the presence of Petroleum system surrounding KFTB.
III. Well properties assessment through petrophysics of Ranikot/Khadro Formation.
IV. Well-to-seismic relationship for structural and stratigraphic framework delineation.
V. Seismic attributes analysis for the Zamzama gas field having 3D seismic availability.
VI. *ML* based rock physics modeling
VII. *ML* based elastic and petrophysical properties approximation having optimized frequency contents.
VIII. Facies and pore-pressure prediction and distribution using enhanced elastic and petrophysics properties compensating the thin reservoir heterogeneities of Ranikot/Khadro Formation.

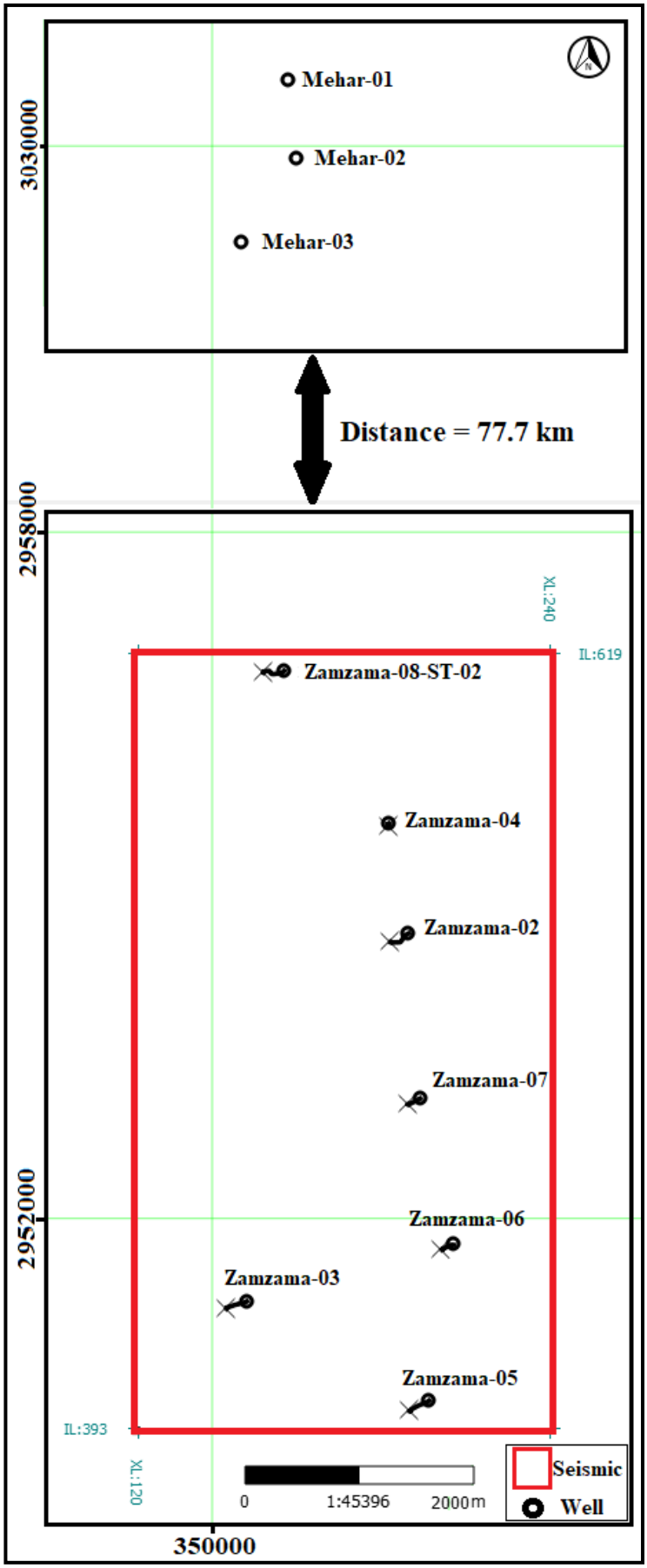



**Figure 1.2** The base map highlights the seismic data set (red rectangle) along with Zamzama (Z) and Mehar (M) wells utilized for the research work.

# CHAPTER 2

# LITERATURE REVIEW

The Kirthar Fold Belt is part of the mountain belts in Pakistan linking the Himalaya orogeny with the Makran accretionary wedge. This region is deforming very oblique/nearly parallel to the regional plate motion vector. The Kirthar Fold Belt is subdivided into several crustal blocks/units based on structural orientation and deformation style (e.g., Kallat, Khuzdar, Frontal Kirthar).

Existing structural concepts for the area were developed for the northern Kirthar Fold Belt and the Sulaiman Fold Thrust Belt and based on the early concepts that proposed a passive roof duplex model (Banks and Warburton, 1986; Jadoon et al., 1994) For the southern Kirthar Fold Belt a model of basement inversion with folds in the sedimentary cover had been put forward (Fowler et al., 2004). However, detailed kinematics on how thick-skinned and thin-skinned structures link are not obvious.

The Ranikot/Khadro Formation is of early Danian age and comprises a more heterogeneous sequence. It is 52.6 m– 55.4 m thick and developed during a period of rising base level and basin re-organization. The lower half (K3 & K4 Layers) is essentially non-reservoir, and dominated by well drained coastal plain muds, with minor ephemeral meandering stream and crevasse deposits. Estuary – shoreline – bay paleo-environments were established in the upper half of the Ranikot/Khadro Formation (K1 & K2 Layers), through progressive marine inundation of the coastal flood plain. A variety of reservoir sand bodies are developed here. These depositional systems probably retained multiple riverine point sources along the coastline, as in the Pab Formation, leading to complex and heterogeneous reservoir architecture. These are capped by transgressive shallow marine limestone of the Ranikot/Khadro Formation and succeeded by deeper marine planktic mudstones of the Girdo Shale (Ooilithica Geoscience, 2004).

A major change in sand provenance characterizes the Ranikot/Khadro Formation, with a major influx of volcanic rock fragments. These sediments were chemically unstable and prone to alteration, principally to chamosite / chlorotic fabrics and amorphous pellets. The resulting lithology comprise very fine to very coarse grained chamositic sublitharenites, which exhibit further decrease in mineral maturity within the finer grained sandstones (Ooilithica Geoscience, 2004).

Many of the *ML* algorithms are employed throughout the world for its robustness and accuracy. Al-Mudhafar (2016) integrated *ML* and data analytics for clastic reservoir

facies and discovered that LogitBoost is the most accurate algorithm, with 100% accuracy in total correct facies prediction. However, the total correct percentages for Multinom and XGBoost were 80.24% and 70.83%, respectively. Rafik and Kamel (2017) employed a combination of *PCA*, model-based cluster analysis, and discriminant analysis to estimate permeability and porosity from well-log data. Three non-parametric techniques are investigated: *ACE*, *GAM*, and *NNET* to predict permeability. The strategy of *ACE* strategy demonstrates superior performance as compared to the other two methods. Ahmadi et al (2014b) conducted research to forecast the permeability and porosity of petroleum reservoirs in northern Persian Gulf oil fields using petrophysical logs. To overcome the aforementioned issue, various *AI* approaches, such as *FL* and *LSSVM*, were used. It is suggested that *LSSVM* and *FL* be used in conjunction with a *GA*. It has been observed that the correlation coefficient between model estimates and relevant real data is greater than 0.96 for the *GA-FL* technique and 0.97 for the *GA-LSSVM* approach and their result is more trustworthy in the case of porosity and permeability predictions. Ahmadi and Chen (2019) evaluated various *ML* methods and suggested that the hybridized technique could predict the reservoir's petrophysical parameters with high accuracy. In the Central Indus Basin, Pakistan, few studies evaluated *ML* methods.

The analytical comparison of seismic inversion with several other multivariate prediction methodologies based on various *ML* algorithms, such as linear regressions, *RF*, and *NNET*, was covered by Priezzhev et al., 2019. To estimate petrophysical and petro-elastic parameters, seismic and well data were merged. In some circumstances, the authors found it feasible to estimate the petrophysical reservoir parameters more correctly and with less interpretational bias, thanks to the nonlinear predictive operator that *ML* techniques extract. According to Otchere et al. (2022), it is essential for *ML* algorithms to have the ability to preprocess the data with the aid of feature selection, removing extraneous factors and effective results produced after detail training and predictions of *ML* algorithm. Ali et al (2021) used the *RFR* to forecast facies with an accuracy of 83.85%, and Ahmed et al (2022) used a stacking method to combine the outputs of numerous models, including the *ETR* reservoir, with an accuracy of 87.23%. Otchere et al. (2022) preferred to use *GBR* and *RF* on eight feature selection technique to model and characterize the shallow marine reservoir with better accuracy.

Yasin et al. (2021a) consolidated rock physics modelling with Post-Stack seismic data, as well as the use of a joint inversion *ML* strategy that combines *SVM* and particle swarm optimization (*PSO*), which assisted the author in not only capturing the thin layers

of clastic packages/channel sand bodies, but also providing a good match of sand and shale layers using a combination of low and high acoustic impedance. In contrast, the use of the Gaussian simulation technique allowed for the analysis of spatial changes in *3D* porosity models displaying consistency at wells. Finally, the author summarized the findings by predicting sweet spots identified with high porosity, permeability, and gas sands in direct comparison to low elastic property values, i.e., *P-wave*, *S-wave*, and $V_p/V_s$ ratio; inside Lower Goru Formation Sand-03.

Hussain et al. (2022) used the basic concept of self-organizing maps (*SOM*), as well as the U-matrix and best matching units (*BMU*), to predict lithofacies and rock types in Zamzama Gas Field log curves. Curve normalization techniques such as normalization and standardization were applied to eliminate the scale value gap between various curves. Later, the use of a clustering technique within *SOM* assisted in providing realistic and optimized results to predict four facies' classes validated at the well location. Likewise, Ahmed et al. (2022) performed *ML* base reservoir characterization in the nearby Zamzama Gas Field, which lies in Lower Indus Basin.

Furthermore, Shakir et al. (2021) applied the Probabilistic Neural Network approach successfully to predict the porosities within Paleocene/Cretaceous age reservoirs of Mehar field, Middle Indus Basin. More than 30 internal seismic attributes were utilized along with external attributes of P-impedance ($Z_p$), S-impedance ($Z_s$), & $V_p/V_s$ ratio in order to train the algorithm on the dataset. The results obtained were verified and tested on the blind wells. Hence, the results were spatially calculated for the whole survey area with maximum accuracy.

In addition, Khan et al. (2022) applied probabilistic neural networking (*PNN*), a non-linear interpolation mathematical approach, within the Zamzama Gas Field. The study demonstrated algorithm training using input petrophysical parameters as seismic internal attributes (envelope, amplitude weighted cosine phase, amplitude weighted frequency, and instantaneous phase) and elastic properties/amplitude variation with offset (*AVO*) attributes as external attributes. Training and testing were conducted at specific wells, while petrophysical parameters such as shale volume, porosities, and saturations were afterward projected across the field.

*ML* enables the manipulation of nonlinear data without a prior specification of the exact nonlinearity, offering a flexible and adaptable approach to data analysis (Poulton and Raiswell, 2002). *ML* algorithms can learn from experience by receiving multiple inputs without being expressly programmed to do so. The constructed model can be used

to make predictions such as classification, continuous value prediction, and forecasting future events or performance. *ML* and statistics have been widely used as regression tools (e.g., *RF*, one-class-support vector machine, support vector regression (*SVR*), genetic algorithm, etc. for predicting petrophysical properties such as shale volume, porosity, $S_w$, and permeability, among others (Zhang et al., 2021; Al Khalifah et al., 2020; Male et al., 2020; Miah et al., 2020; Zhong et al., 2019; Fattahi and Karimpouli, 2016; Hampson et al., 2001; Helle et al., 2001a;). Among these *ML* tools, the *RFR* and *ETR* are highly effective that have seen widespread use in the estimation of reservoir properties through using well log variables (Zhang et al., 2021; Erofeev et al., 2019; Ahmadi et al., 2014a; Helmy et al., 2010; Helle et al., 2001b). There is a lack of literature comparing various *ML* approaches against conventional methods in the prediction of reservoir properties in carbonate reservoirs, as demonstrated by studies such as (Male et l., 2020; Helle et al., 2001a).

Many of the supervised/unsupervised *ML* algorithms approaches are employed for the assessments of the field's secondary reservoir, which is the primary exploration target at the time in the area and nearby blocks. The advance algorithms include *SVM*, *DTR*, *RF*, and *GBR* to highlight Ranikot/Khadro's Formation heterogeneous thin gas sands facies. Based on statistical metrics, the *GBR* algorithm generated the best results by identifying and describing the thin heterogeneous sand facies. The main principle underlying *GBR* success is to generate new base learners that are maximally correlated with the ensemble's negative gradient of the loss function, which is achieved by a sequential error fitting strategy.

According to the previous results, the developed approach can be used as an analogue for the regional development of thin-bedded sandstone systems in the studied basin. This workflow may serve as a vital example for the exploration of the remaining gas-bearing stratigraphic systems within the studied basin and similar basin settings. The Ranikot/Khadro's Formation heterogeneous thin gas sands facies are an important rock unit in the Indus Basin petroleum system because it contains an economically significant volume of potential reservoir rocks that are composed of sandstones. As a result, investigating the reservoir facies and seismic signature of the Paleocene Ranikot/Khadro Formation thin-bedded reservoir is useful for effective reservoir quality prediction, which contributes to the overall understanding of the Paleocene hydrocarbon plays. A better understanding of these reservoirs can be gained by accurately predicting the sand

distribution probability trends of these deposits using advanced imaging techniques, which will aid in further exploration and reservoir modeling.

ML assisted facies classification has yielded highly effective and promising results, particularly for evaluating unexplored regions within the Indus Basin. The generated probability distribution maps can pinpoint and guide the potential exploitation of these reservoirs across regional and subsurface scales. Crucially, identifying the most suitable ML technique has the potential to significantly improve reservoir prediction and analysis, ultimately facilitating better reservoir management and recovery in these heterogeneous, thin-bedded sandstone formations. Furthermore, our analysis of physical property prediction indicates the potential to classify the sandstone reservoirs into three distinct units traceable across the entire study area. These three generated facies maps can serve as a valuable foundation for field development strategies, enabling more efficient exploitation of these thin-bedded secondary reservoirs.

## 2.1 Generalized Geology

Pakistan is situated in the western part of the Indian Plate is geologically complex, with a varied geological history. The country is characterized by several tectonic plates, faults, and folded mountain ranges. Pakistan is home to several sedimentary basins, which have been formed by tectonic activity and subsequent subsidence. Pakistan's geology is primarily composed of two sedimentary basins: the Balochistan Basin and the Indus Basin. These basins formed over a series of geological periods and came together as a consequence of the collision between the Indian Plate and the Eurasian Plate during the Paleogene era (Ghazi et al., 2015; Zhuang et al., 2015). The Indus Basin belongs to the Indian Plate, whereas the Baluchistan Basin is part of the Eurasian Plate. In addition to these two major basins, there are several smaller basins, such as the Peshawar Basin, the Hazara Basin, and the newly discovered Kakar Khorasan Basin (Pishin Basin), which is a pull-apart basin (Kazmi and Abbasi, 2008; Kazmi and Jan, 1997).

The Indus Basin, located in the northwestern region of Pakistan, is the largest sedimentary basin in the country, covering an area of over 300,000 square kilometers. The basin is composed of sedimentary rocks that range in age from the Early Jurassic to the Late Tertiary period. The basin is an important source of petroleum, natural gas, and coal. The Balochistan Basin, which occupies an area of approximately 100,000 square kilometers, is situated in the western part of Pakistan (Busby and Pérez, 2011). The basin

is predominantly composed of sedimentary rocks that range in age from the Early Permian to the Late Tertiary period. The basin contains significant reserves of natural gas and petroleum. The Potwar Basin is located in northern Pakistan and covers an area of about 30,000 square km. It is filled with sedimentary rocks ranging in age from Late Cretaceous to Late Tertiary (Bender and Raza, 1995). The basin is an important source of petroleum and natural gas. SIB is located in southern Pakistan and covers an area of about 10,000 square kilometers and comprises sedimentary rocks that span from the Early Cretaceous to the Late Tertiary period. The basin contains significant reserves of natural gas. These sedimentary basins are an important part of the geology of Pakistan, providing valuable resources and contributing to the country's economy (Busby and Pérez, 2011).

The main sedimentary basin in Pakistan is the Indus Basin has been shown in Figure 2.1a. The primary element in the proto-Indus Basin that governed the sedimentation up to the Jurassic age is the Precambrian Indian Shield rocks of Kirana Hills (Sargodha High) (Asim et al., 2014). The main geographic characteristic that separates the Upper Indus Basin from north of Sargodha High and Lower Indus Basin from southern side of Sargodha High (Kemal, 1991). The Kohat sub-basin, located western part of the Indus River, and the Potwar sub-basin, situated in eastern side of the Indus River, are additional divisions of the Upper Indus Basin (Kadri, 1995).

The Khairpur-Jacobabad High is an important underlying topographic feature that divides the Lower Indus Basin into the Central Indus Basin and the Southern Indus Basin. Southern Indus Basin is separated from Central Indus Basin by Upper Indus Basin via Sargodha High in the northern and Khairpur-Jacobabad High in the southern side. Southern Indus Basin is the basin that spreads into Offshore Indus Basin from Khairpur-Jacobabad High in the southward side (Kadri, 1995). The Upper Indus basin is characterized by compressional tectonic forces, whereas the Central and Southern Indus basins are mainly affected by extensional forces in the eastern half and a fold and thrust belt along the western boundary (Asim et al., 2014).

## 2.2 Southern Indus Basin

The Zamzama Block is situated within the Kirthar Sub basin, which is part of the Southern Indus Basin (Figure 2.1a, b). It is located to the south of the Sukkur Rift, which acts as a dividing line between the Southern and Central Indus Basins. The Southern Indus Basin is bounded to the east by the Indian shield and to the west by the marginal

zone of the Indian plate. Its southern limit is determined by the offshore Murray Ridge (Kadri, 1995). Pakistan's Southern Indus Basin is a sedimentary basin that spans an area of around 95,000 square kilometres, stretching from the southern reaches of the Thar Desert to the Arabian Sea. The basin takes the form of a trough-like depression in the Earth's crust that has accumulated sediment over millions of years (Qureshi et al., 2021).

The sedimentary formations within the Southern Indus Basin consist mainly of sandstones, shales, and limestones, and were laid down in a marine setting. The Southern Indus Basin is known to contain significant reserves of oil and gas, as well as coal and other minerals (Ehsan et al., 2018). The exploration and production of these resources is an ongoing activity and is a major contributor to the economy of Pakistan. Southern Indus is classified as:

a. Thar platform.
b. Karachi Trough.
c. Kirthar Fold Belt.
d. Kirthar Foredeep.
e. Offshore Indus basin.

### 2.2.1 Thar Platform

The Thar Platform is a geological feature located in the Thar Desert of Pakistan. It is a large, flat, and relatively stable area of land that has been elevated due to tectonic uplift. The Thar Platform is composed of Precambrian rocks, including granites, gneisses, and metamorphic rocks. The platform is surrounded by the younger, less stable sediments of the Indus Basin. The Thar Platform is of great geological and economic significance (Siddiqui, 2012). It contains rich deposits of coal, as well as other minerals such as copper, gypsum, and salt. The Thar coal fields are considered one of the largest coal reserves in India and have been extensively mined. The Thar Platform is also an important source of groundwater, with several large-scale irrigation and drinking water projects underway (Qayyum et al., 2016).

The Thar Platform is a gently sloping monocline, similar in nature to the Punjab Platform, and influenced by underlying basement topography. The sedimentary layers thin out towards the Indian Shield, which is marked by the Nagar Parkar High. Unlike the Punjab Platform, the Thar Platform exhibits buried structures caused by extension tectonics resulting from the latest counter-clockwise movement of the Indian plate

(Siddiqui, 2012). It is bounded by the Indian Shield to the east, merges with the Kirthar and Karachi Troughs to the west and is bordered to the north by the Mari-Bugti Inner Folded Zone. A strati-structural cross-section that traverses the Thar Platform, Karachi Trough, and Offshore Indus reveals the stratigraphic and structural differences between the two sub-basin (Quadri and Shuaib, 1986).

### 2.2.2 Karachi Trough

The Karachi Trough is a geological feature located off the coast of Karachi, Pakistan. It is a depression in the Earth's crust that extends from the coast of Pakistan to the Arabian Sea. Geological evidence suggests that the Karachi Trough was produced due to tectonic processes occurring at the interface between the Indian Plate and the Eurasian Plate. The trough is filled with sediments, primarily sandstones and shales, which were deposited in a marine environment. The Karachi Trough is of geological and economic significance. It has been identified as a promising area for oil and gas exploration and has attracted numerous drilling and exploration initiatives (Hussain et al., 2017).

The Karachi Trough is an embayment that extends towards the Arabian Sea, featuring thick early sediments that indicate the final stages of marine sedimentation. The area encompasses numerous narrow chains of anticlines, some of which contain gas deposits. The rocks of the early, middle, and late Cretaceous periods are well-preserved in this region, which has remained a trough throughout its geological history. The Upper Cretaceous is characterized by the westward progradation of a marine delta (Nabi et al., 2019).

One of the most intriguing aspects of the Karachi Trough is the occurrence of continued sediment deposition across the Cretaceous/Tertiary (*K/T*) boundary. The Korara Shales were deposited during this period, with the basal section representing Danian sediments. This localized phenomenon is considered a rare example of uninterrupted sedimentation at the end of the Cretaceous era (Shuaib, 1973).

### 2.2.3 Kirthar Foredeep

The Kirthar Foreland Basin, also known as the "Foredeep," is situated in Pakistan and is part of the Kirthar Fold Belt, a region of folded and thrust-faulted mountains and sedimentary basins spanning Pakistan and western India. The basin is a substantial

depression in the Earth's crust that has accumulated sediment from the Kirthar Mountains to the west. The sedimentary rocks in the basin consist mainly of sandstones, shales, and limestones that were deposited in a marine environment (Asim et al., 2014). The Kirthar Foreland Basin is considered an important source of oil and gas, with several exploration and production activities underway. The basin is also an important area for agriculture, with large-scale irrigation projects providing water for crops. The Kirthar Foreland Basin is a geologically complex area and has been the subject of extensive geological and geophysical studies (Smewing et al., 2002).

The Kirthar Foreland Basin is situated in Pakistan and is part of the Kirthar Fold Belt, a region characterized by folded mountains and sedimentary basins. The basin is aligned in a north-south direction and has an eastern boundary that is faulted with the Thar Platform. Sediment accumulation in the basin has resulted in a thickness of more than 15,000 m, with the predominant sedimentary rock types being sandstones, shales, and limestones. Although continuous sedimentation is thought to have occurred, the Upper Cretaceous appears to be absent based on well correlations between Mari Khairpur and Mazarani. The Palaeocene is well developed in the depression, although it is missing in the Khairpur-Jacobabad High area. The basin is believed to have significant potential for the maturation of source rocks, similar to the Suleiman Depression (Smewing et al., 2002).

### 2.2.4 Kirthar Fold Belt

The geological formation known as the Kirthar Fold Belt was created by the collision of the Indian Plate and the Eurasian Plate. This complex structure is made up of parallel mountain ranges that were formed by the folding and thrust faulting of sedimentary rocks. The rocks in the Kirthar Fold Belt are estimated to be between Jurassic and Tertiary in age, and consist of sandstones, shales, and limestones.

The Kirthar Fold Belt is a north-south trending geological feature that shares structural similarities and stratigraphic equivalence with the Suleiman Fold Belt. Sedimentary rocks ranging in age from the Triassic to the recent were deposited in this region. The configuration of the Kirthar Fold Belt also signifies the closing of Oligocene-Miocene seas. The western part of the Kirthar Fold Belt adjacent to the Balochistan Basin, which demarcates the western edge of the Indus Basin, has undergone significant tectonic disturbance. Hydrothermal activities associated with this western margin have resulted in

the formation of economically significant mineral deposits of baryte, fluorite, lead, zinc, and manganese.

### 2.2.5 Offshore Indus

The geology of the offshore Indus region is characterized by a thick sequence of sedimentary rocks that were deposited in the shallow waters of the delta and the surrounding coastal areas. These rocks contain important hydrocarbon reserves, as well as other important minerals. The Indus River also brings large quantities of sediment into the Arabian Sea, creating a dynamic and constantly changing landscape that is shaped by the interplay of river discharge, waves, and ocean currents (Nabi et al., 2019).

The offshore region of the Indus basin is situated on a passive continental margin and has undergone two major geological periods: the Cretaceous-Eocene and the Oligocene-recent. Sedimentation began in the region during the Cretaceous era, with deltaic and submarine fan sedimentation occurring since the middle Oligocene following the inception of the Proto-Indus system (Siddiqui, 2012).

The offshore Indus region is divided up into two areas, the platform and depression, along the Hinge Line which is parallel to 67°E Longitude. The platform consists of the Karachi Trough and the Thar Platform deltaic area, which are separated by a line that marks the boundary between the Karachi Trough and the Thar Slope onshore (Busby and Pérez, 2011).

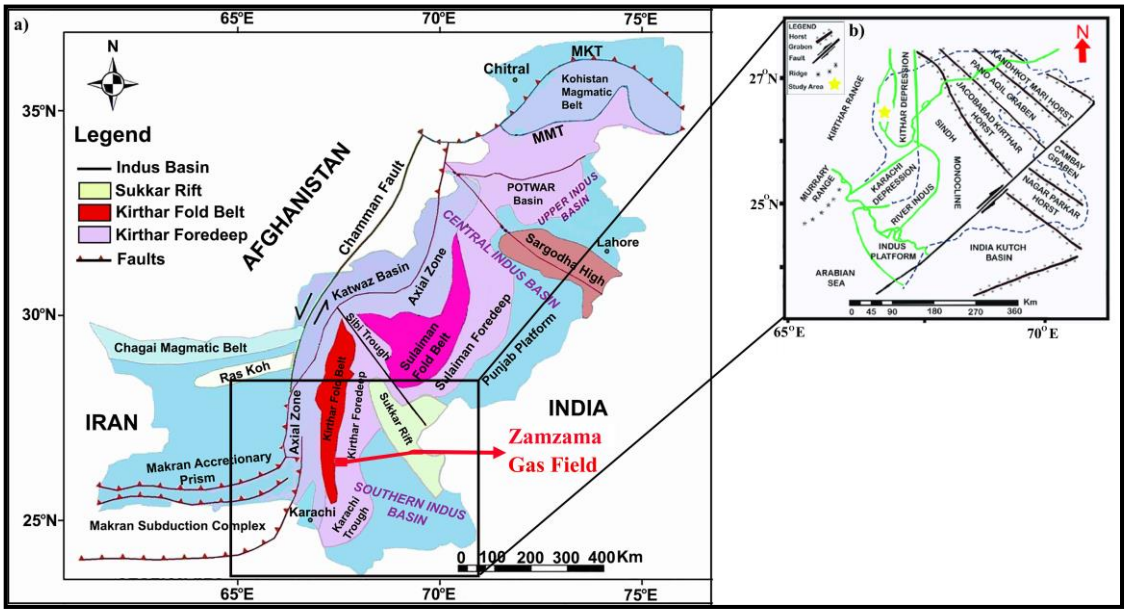



**Figure 2.1** Geology and tectonic features of a) Pakistan and b) Southern Indus basin (Ahmed et. al., 2022).

## 2.3 Tectonic Framework

The area of investigation is tectonically complex (Kazmi and Jan, 1997). The Gondwana fragments began to drift in the Jurassic period, with the Indian plate moving rapidly northward during the Cretaceous and then undergoing a counterclockwise rotation as part of the collision process. This led to the formation of the Chaman rift, a transform zone that resulted from renewed spreading in the Indian Ocean during the post-collisional northward movement of the Indian plate. The compressional forces that were produced likely gave rise to various syntaxes, arcs, and oroclines throughout Pakistan, possibly in the Early Oligocene, with the pre-collisional tectonic elements of the mio-geosyncline and the shield areas of the Indian plate likely having played a role in the formation of these structures. The formation of the structure was possibly influenced by tectonic elements that existed prior to the collision between the Indian plate and the Eurasian plate, particularly within the mio-geosyncline and shield areas of the Indian plate. These tectonic elements included the development of narrow, welt-like structures within geo-anticlinal areas in the western part of the Indian plate, which were initiated in the Jurassic period and fully emerged during the Cretaceous period. These structures, along with their associated intrusive and volcanic emplacements, have been compared to island arcs by Hunting Survey Corporation (1961) and their origin has been linked to worldwide oceanic ridges (Kazmi and Jan, 1997).

## 2.4 Stratigraphic and Structural Framework of Zamzama Block

The Kirthar Fold belt area experienced various phases of syn rift and post-rift sedimentation since the Jurassic period, and its rift-drift-collision history is well-documented (Besse and Courtillot, 1988; Patriat and Achache 1984; Searle et al., 1997). The oldest stratigraphic unit in the region is the Shirinab Group, which includes shallow marine clastic and limestone formations that can be considered as syn-rift deposits (Figure 2.2). During the Middle-Late Jurassic, the Chiltan Limestone was deposited on a broad carbonate shelf of the Indian subcontinent (Dolan, 1990). In the Early Cretaceous, sedimentation occurred in a passive margin environment with the rising of the emergent Indian continent to the southeast. The deposition of thick sediments of the Goru and Sembar Formations took place along the edge of the passive margin due to the large prograding delta systems (Richard et al., 2001; Bender and Raza, 1995). During the deposition of the Sembar Formation, which is the most abundant source rock in the Lower

and Middle Indus basins and the Kirthar Fold belt, restricted marine, anoxic environments prevailed in the Early Cretaceous.

In the Middle Cretaceous, the eastern part of the Indus Basin experienced continued transgressive-regressive cycles, resulting in the deposition of shoreface sand in the Lower Goru Formation. The Kirthar foredeep and fold belt area received only thin turbidites from sediment delivery channels to the basin floor, as the study area was too far from the input feeders. Consequently, the Goru Formation only deposited marine shales and carbonates. A major marine transgression occurred in the Turonian, which further reduced the clastic sediment supply to the basin, leading to sediment starvation and condensation in deeper areas. In the Kirthar Fold belt area, deeper water carbonates (Parh Limestone) were deposited throughout the Coniacian-Companian period (Richards, Boyce and Pringle, 2001)

In the Late Cretaceous period, the Kirthar Foredeep and frontal section of the fold belt regions received a considerable amount of sand-rich turbidities known as the Mughalkot. These formations are clearly visible in the Southern Kirthar Fold Belt due to the uplifting of the Indian Shield, which occurred as the Indian Plate moved across the Reunion Hotspot (Gnos, 1998). During the Early Paleocene, the Indian Shield was uplifted, leading to the creation of shallow marine environments in multiple regions of the Kirthar Fold belt. This resulted in the selective accumulation of fluvio-deltaic and shallow marine sands from the Pab Formation and Lower Ranikot/Khadro Formation. After the deposition of shallow marine and fluvio-deltaic sands of the Pab Formation and the Lower Ranikot/Khadro Formations in many parts of the Kirthar Fold belt, there was a decrease in the sea level rate during the Late Palaeocene-Early Eocene, which allowed for the deposition of Dunghan/SML carbonates in the study area. Throughout the basin, there were regional transgressive and backstepping events during the Late Early Eocene, with thin shales containing alternating limestone units of the Ghazij Formation deposited in the eastern part of the Kirthar Fold belt, and thick shales deposited in deep to shallow marine environments in the western part of the Kirthar Fold belt area. The Kirthar Limestone was deposited throughout the basin during the Late Eocene under calm marine conditions. Due to plate collision, the Western Kirthar part of the basin was restricted for sedimentation during the Oligocene, which led to the development of small, isolated basins where sedimentation continued until the Quaternary period.

The Zamzama structure is a large anticline that verges eastward and is oriented north-south. As older sediments of the Lower Indus Basin are exposed in the Kirthar

foredeep to the west of Zamzama, the degree of thrusting and folding increases. The uplift and erosion at the unconformity led to the formation of extensional faults, some of which might have been reactivated during the development of the Zamzama structure. The exact onset timing of folding at Zamzama is difficult to determine but is thought to have occurred in the Plio-Pliestocene.

## 2.5 Stratigraphic and Structural Framework of Mehar Block

On eastward portion of the Kirthar fold, there exists a movement along the Ornach-Nal and Chaman transform faults which occurred through collision of Oligocene Miocene plate has produced several different anticline structures. The presence of ophiolites on west slope of the Kirthar fold belt indicates occurrence of an active plate boundary, which can be seen through seismic events. Mafic to ultramafic rocks make up the rock type along plate border (Jadoon et al., 1994).

Pre-existing normal faults with a *NW-SE* orientation run virtually parallel to the Jacobabad highs and interrupt the foredeep area in research region on south side of Middle Indus basin. Mazarani Thrust is the fundamental flaw that has resulted in the building of thrust in this area. This thrust belt has a north-south trend and is west dipping. Extensional faults that existed previously were reactivated and can be found between the cores of the Mehar Mazarani Fold structures. The trend along these former normal faults indicates that the thrusting in the Mehar Mazarani Fold is not happening close to the pre-existing normal faults, since if it were, the orientation of the main mantle thrust (*MMT*) would be alongside the fault plane of the prior extensional faults (*NW-SE*) (Ahmed and Ali, 1991).

From the Infra-Cambrian to the late Cretaceous, the Mehar Mazarani Fold Belt has been subjected to many tectonic processes that result in uplifting, rifting, and erosion. Tectonic activity also reactivated a fault that existed in the basement of this location. The major cretaceous source and reservoir were accumulated in a passive margin environment, whereas drifting in cretaceous may revive the preexisting normal strength and induce the rising of the area east toward north of the Mehar Mazarani Fold belt obsessed by the Jacobabad Highs belt. The Jacobabad highs have been eroded by late cretaceous sediments. Beginning from the Mehar Mazarani fold belt region towards the high, a wedge-shaped of oligoacene and Paleocene strata show that the paleo high phenomena occur during this period. During the Miocene period, the collision between

the Indian Plate and the Asian Plate resulted in the deposition of thick sedimentary layers that caused strata inversion and the formation of the Mitto anticlinal fold in the Mehar Mazarani fold belt (Ali et al., 2018).

The reason behind this is separation of thick-settled Infra-Cambrian, the Mehar Mazarani fold belt appears to have a folded structure. Basement wrench faults in the fold belt's western portions, supposed towards portion of the Ornach-Chaman transform fault. Because of the fault and west side of the inclined basement, and the shortening caused by Jacobabad paleo highs, such Mehar Mazarani thrust have got an upward ramping structure. The ramping shape structure in this location identical to that seen in the Potwar basin Dhurnal section. The change of incompetent shale lithology of the Goru Formation, which generates narrow resultant detachment zone, causes another occurrence of back thrusting. To understand the complexity of this area, significant amount of research using *3D* data is required for future development and exploration (Ali et al., 2018).

## 2.6 Generalized Stratigraphy of Zamzama Block

The Kirthar Foredeep Basin in Pakistan is a large sedimentary basin. The stratigraphy of the basin is characterized by several geological formations ranging from Late Cretaceous to Holocene in age. The following is a more detailed generalized stratigraphy of the Kirthar Foredeep Basin:

a. Upper Cretaceous: The uppermost unit in the Kirthar Foredeep Basin is composed of sandstone and siltstone formations of Late Cretaceous age. The sandstone formations in this unit are typically cross-bedded and well-sorted, indicating depositional environments such as alluvial fans and deltaic systems. The siltstone formations are typically finer grained and less well-sorted, indicating depositional environments such as lacustrine or shallow marine environments.
b. Eocene - Oligocene: The next unit is characterized by the presence of sandstone, siltstone, and shale formations of Eocene to Oligocene age. The sandstone formations in this unit are typically cross-bedded and well-sorted, indicating depositional environments such as alluvial fans and deltaic systems. The siltstone formations are typically finer grained and less well-sorted. The shale formations in this unit are typically dark and organic-rich, indicating depositional environments such as deep marine environments.

c. Miocene: The sedimentary deposits of the Miocene epoch are prominent in the Kirthar Foredeep Basin and comprise sandstone, siltstone, shale, and limestone formations. This period is considered a significant episode of deposition in the region. The sandstone formations in this unit are typically cross-bedded and well-sorted, indicating depositional environments such as alluvial fans and deltaic systems. The siltstone formations are typically finer-grained and less well-sorted. The shale formations in this unit are typically dark and organic-rich, indicating depositional environments such as deep marine environments. The limestone formations in this unit are typically biogenic in origin and indicate depositional environments such as shallow marine or reef environments.
d. Pliocene - Pleistocene: The Pliocene to Pleistocene unit is characterized by the presence of alluvial and deltaic deposits, as well as sandstone, siltstone, and shale formations. The alluvial and deltaic deposits in this unit are typically coarse-grained and well-sorted, indicating depositional environments such as river and delta systems. The sandstone, siltstone, and shale formations in this unit are typically finer-grained and less well-sorted.
**e.** Holocene: The Holocene unit is characterized by the presence of recent alluvial and deltaic deposits, as well as wind-blown sand and silt deposits. The alluvial and deltaic deposits in this unit are typically coarse-grained and well-sorted, indicating depositional environments such as river and delta systems. The wind-blown sand and silt deposits in this unit are typically fine-grained and well-sorted, indicating depositional environments such as desert and coastal environments.

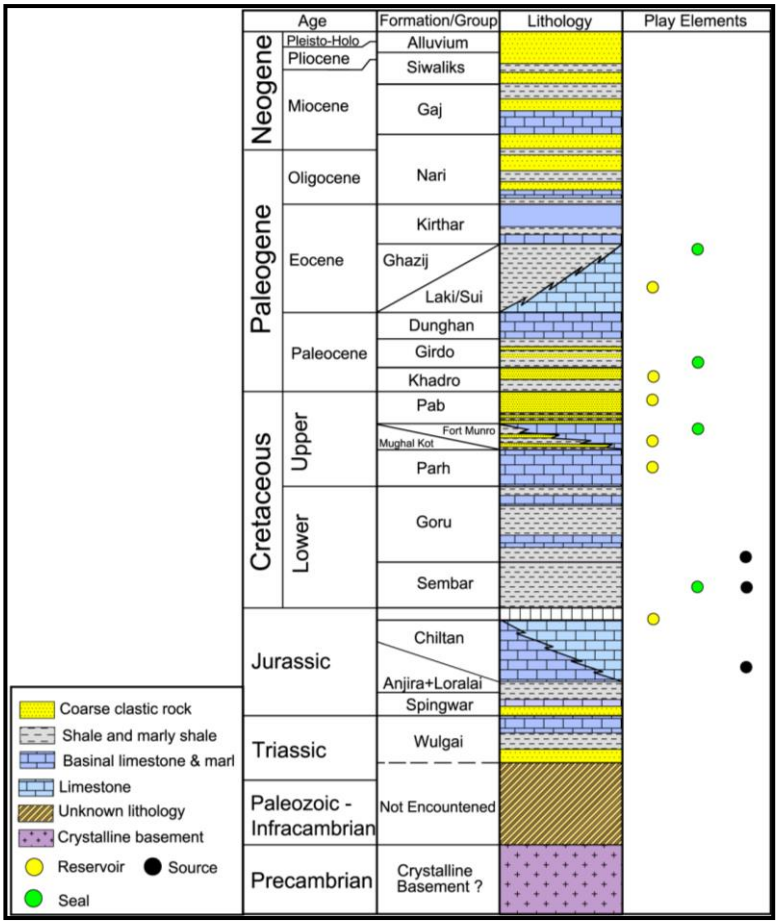



**Figure 2.2** Generalized stratigraphy of the Zamzama Block (Zafar et al., 2018).

### 2.6.1 Sembar Formation

In the Kirthar Sulaiman Region, the Sembar Formation is the lowermost unit of the Cretaceous Sequence. This formation is primarily composed of shale blocks, interbedded with silt stone and nodular argillaceous limestone. The shale and silt typically contain glauconite. The formation is 133 m thick in the type of area and 262 m thick in the Mughalkot section. A gradational contact with the overlying Goru formation has been reported, though in some places an unconformity has also been noted by (Williams, 1959).

### 2.6.2 Goru Formation

The Goru Formation in the Kirthar Fold Belt is primarily made up of interbedded limestone, shale, and siltstone. The lower part is predominantly shaly and consists of thin-bedded, light-colored limestone interbedded with calcareous, hard, splintery gray to olive-green shale. The upper part is mostly thin-bedded, light-colored porcelaneous

limestone with subordinate shale. It grades into the overlying Parh Limestone. At the type locality in Goru Village, the formation has a thickness of 538 m according to (Shah, 2009).

### 2.6.3 Parh Limestone

Parh Limestone is a distinctive and consistent rock formation that can be found extensively throughout the Kirthar Sulaiman region. The limestone is thin-bedded and comes in light grey, white, or cream colours, often with a characteristic pink, purple, or maroon interbed of variegated shales and marls. The limestone itself is hard and can be lithographic to porcelaneous, often appearing as flat slabs with a concoidal fracture. In the type area of the Parh range, the formation is 268 m thick, while elsewhere its thickness ranges from 300 to 600 m. It is in conformable contact with the underlying Goru Formation and is overlain conformably by the Mughalkot Formation, according to (Kazmi, 1988).

### 2.6.4 Mughalkot Formation

The Mughalkot Formation is a geological formation found in Pakistan. The Mughalkot Formation is predominantly made up of sandstone and siltstone and is believed to have formed during the Eocene epoch. The sandstone typically has a medium to coarse grain size and consists of well-sorted, subangular to rounded grains. The siltstone is typically fine-grained and composed of clay and silt-sized grains. The formation is considered to be of fluvial origin, with the sandstone representing channel deposits and the siltstone representing floodplain and over bank deposits. The Mughalkot Formation is an important source of water and minerals in the region and is also considered to be a potential source of hydrocarbons.

### 2.6.5 Fort Munro Formation

This Formation is, located in the Sulaiman Range & Toba Kakar Range. It is a Late Eocene to Early Oligocene age formation and is composed of sandstone, shale, and limestone. Due to its unique lithology and extensive range, the "Fort Munro limestone member" is given the distinction of a formation. The rocks in the area around Bara Nai in the Kirthar Province consist mostly of limestone, which display a range of colors including blue, cream, grey, light brown, and beige. The limestone is thickly bedded,

massive, and has reef-like structures in the lower section. The elevated portion of the limestone is sandy and contains intercalations of shale and brown-weathering sandstone (Williams, 1959).

### 2.6.6 Pab Formation

The Pab Formation is a geological formation in Pakistan that is present in the Salt Range and the Potwar Plateau regions. It is a part of Middle Indus Basin and is composed of sandstone, shale, and siltstone. The formation is believed to have been deposited in deltaic and lacustrine environments during the Late Cretaceous to Early Eocene time period. The sandstone units in the Pab Formation are typically well-sorted and cross-bedded, while the siltstone and shale units are typically finer-grained and less well-sorted. The Pab Formation holds significant importance as it serves as a major reservoir for petroleum and natural gas resources in Pakistan, thus being regarded as a crucial geological formation in the country.

### 2.6.7 Ranikot/Khadro Formation

The primary objective of this study is to characterize this Paleocene reservoir. Different nomenclatures have been utilized by various operators while working on this key reservoir. The most common name is Ranikot Formation while some called it as Khadro Formation. In this study the researcher used Ranikot/Khadro Formation for this key reservoir under study. It is abundantly spread in Kirthar and the surrounding area. On the Laki Cretaceous Pab Formation, it is found unevenly. The Bara Nai area in the Laki Range is the designated type locality for the limestone formation. The formation's lowermost part at the type locality is characterized by the presence of oyster shells and reptile bones (Shah, 2009).

### 2.6.8 Bara Formation

The Bara Formation is a significant geological unit in Pakistan that is located in the Potwar Plateau and Salt Range region. It dates back from the Middle to Late Eocene period and consists of sandstone, shale, and limestone. The formation was deposited in a shallow marine environment and is notable for its well-preserved fossils of foraminifera, bivalves, and gastropods.

Regarding its stratigraphy, the Bara Formation is commonly subdivided into distinct subunits based on variations in rock type and fossil content. The lower part of the formation is usually composed of sandstone and shale, while the upper part is mostly made up of limestone and shale.

In the Kirthar Range and surrounding regions, this formation is distributed widely and conformably encloses the Ranikot/Khadro Formation. The type locality is located in the Laki Range's Bara Nai. Shale and sandstone are scattered within the formation. The sandstone is cross-bedded, ripple-marked, calcareous, ferruginous, and frequently glauconitic, with fine to coarse grain. The thickness of beds can range from a few milli meters to over three meters. The shale is typically carbonaceous, mushy, and gypsiferous (Shah, 2009).

### 2.6.9 Laki Formation

The formation is often found overlying the Ranikot/Khadro Formation Group and is prominently visible in the southern Kirthar and Sulaiman ranges. The impart information is in the northern Laki Range, close to Mari Nai. The formation is made up of lateritic clay, marl, calcareous shale, sandstone, and cream to grey limestone (Shah, 2009).

### 2.6.10 Kirthar Formation

The Kirthar Formation is situated in the Kirthar fold and thrust belt of Pakistan. This formation dates back to the Late Cretaceous to Early Tertiary period and comprises intercalated sedimentary rocks such as sandstone, siltstone, shale, and limestone. It was formed in a shallow marine and deltaic setting and is notable for its exceptional fossil preservation, including ammonites, foraminifera, and bivalves. Moreover, the Kirthar Formation is an important source of groundwater in the region, providing water for various purposes, especially for irrigation.

The Kirthar Formation is widely distributed in the Kirthar region and can be traced to the Waziristan and Suleiman areas of Pakistan. It conformably overlies the Laki Formation in the Kirthar range and the Ghazij Group in other locations. The formation primarily consists of fossiliferous limestone, interspersed with shale and marl. The limestone ranges from grey to white, is thickly bedded to massive, and exhibits nodules and localized algal formations. The thickness of the formation varies considerably,

ranging from 15 to 30 m in the western Kirthar range to as much as 1270 m in the Gaj River type section (Shah, 2009).

### 2.6.11 Nari Formation

The Nari Formation is widely exposed in the Kirthar and Sulaiman regions, with scattered outcrops present in tectonised thrust blocks of the Balochistan Ophiolites and thrust Belt. In the Kirthar province, it conformably overlies the Kirthar Formation, except in the Hyderabad anticlinorium, where it unconformably overlies the Kirthar and Laki Formations. The type section of the Nari Formation is situated in the Gaj River Gorge in the Kirthar Range. The thickness of the formation varies from 1045 m to 1820 m in the Kirthar area (Shah, 2009).

### 2.6.12 Gaj Formation

The Gaj Formation is a geological formation in Pakistan. It is a late Miocene to early Pliocene age formation composed of sandstone, siltstone, and claystone. The Gaj Formation was deposited in a fluvial environment and is known for its well-preserved fossil record, including the fossils of mammals, reptiles, and plants. The sandstone units in the Gaj Formation are typically well-sorted and cross-bedded, with high porosity and permeability. These sandstone units have the potential to act as reservoir rocks for hydrocarbons, although the Gaj Formation is not considered to be a major source of petroleum in Pakistan.

However, in the southern portion of the Kirthar Range near Karachi, the Nari Formation mainly comprises yellow to brown sandstone and cream-colored or pinkish-white argillaceous limestone. The thickness of the formation varies from around 90 m in Quetta to 600 m in the Kirthar area (Shah, 2009).

### 2.6.13 Siwaliks Formation

The Siwalik Formation is a sedimentary rock formation located in Pakistan and is part of the Siwalik Group. The group consists of a series of sedimentary rocks that were deposited during the Neogene period, spanning approximately 23 to 2.6 million years ago. The Siwalik Formation is primarily located in the northern region of Pakistan and comprises alternating layers of sandstone, conglomerate, shale, and siltstone.

Siwalik Formation was formed by the deposition of sediments in large alluvial fan systems that originated from the rising Himalayas. The sediments in the Siwalik Formation are typically well-sorted and contain abundant fossils, including those of mammals, reptiles, and plants. The fossils in the Siwalik Formation have provided important information about the evolution of life in the region and the environmental conditions that existed during the Neogene period.

## 2.7 Petroleum System of Zamzama

The Southern Indus Basin is a recognized hydrocarbon producing province. The proven reservoirs in the basin are the clastics of Cretaceous (Lower Goru and Pab Formation) and Palaeocene (Ranikot/Khadro) Formations, which have been charged by Lower Cretaceous source rocks.

### 2.7.1 Source Rock

Early Cretaceous fine-grained basinal shales representing distal undifferentiated time equivalents of the Sembar and Goru deltas are interpreted to be the source rock section that provides gas to the Pab Formation reservoir.

Source rock is more than 6 km deep in Kirthar foredeep. Goru/Sembar basinal shales are mixed type II/III gas prone kerogen. They have good total organic carbon (*TOC*) values of 1-2 % occasionally 4 %. *HI* is generally <150 mg/g.

### 2.7.2 Reservoir Rock

The primary hydrocarbon reservoir in the Kirthar Fold belt is formed by the late Cretaceous age Pab Formation, and it also serves as the primary reservoir for the Zamzama and nearby Bhit gas fields. The gas is also present in the sandstones of the overlying Ranikot/Khadro Formation, which is of Palaeocene age. The Pab Formation, around 220 m thick in the Zamzama area, is part of a sand-rich fluviodeltaic, coastal plain, and shore face depositional system at a regional scale. The Pab Formation exhibits moderate porosity and permeability (around 8% and 150 mD, respectively) with a Net-To-Gross (*NTG*) of approximately 60%.

The Ranikot/Khadro Formation is secondary reservoir in Zamzama Gas Field. It is generally around 54 m gross thick and heterogeneous with poor sandstone development

in the upper part of the formation (*NTG* c. ~10-15%). Permeability in the Ranikot/Khadro Formation is severely affected by digenetic alteration of volcanoclastic components.

### 2.7.3 Top Seal and Cap Rock

The marine shales of the Girdo (Ranikot/Khadro) Formation act as the top seal for the Ranikot/Khadro Formation and Pab Formation reservoirs. The main thrust fault extends close to surface and clearly seals, otherwise the Pab Formation reservoir would not contain gas. There does appear to be fault leakage when permeable facies are juxtaposed against permeable facies. Major source, reservoir and seal rock of Zamzama is presented in Table 2.1.

**Table 2.1:** Petroleum play of the Zamzama Block.

| Play Elements | Formations | Age |
|---|---|---|
| Seal | Ranikot/Khadro Formation | Paleocene |
| Reservoir | Pab Formation | Upper Cretaceous |
| Source | Sembar Formation | Lower Cretaceous |

## 2.8 Generalized Stratigraphy of Mehar Block

The Mehar block is located in the western part of the Kirthar Foredeep, which is a sub-basin of the Lower Indus sedimentary basin. Sediments between Triassic to present age are exposed at various places. Sedimentation in this part of the fold belt commenced during the Paleozoic. This is recognized in Kalat, Wulgai and Ghazaband areas where Permo-Triassic boundary has been observed and documented by the hunting survey corporation 1960. In the study area the oldest sediments recorded belong to the Permian, but these sediments are never drilled by any well. Magnetic data of the study area indicate several basement highs and lows. Although the exact nature of the sediments overlying the basement is not known, it may be assumed that it is overlain by Paleozoic sediments (Abbasi et al., 2016).

The study regions stayed active tectonically and consist of various phases of syn-rift and post-rift sedimentation since Triassic time. Oldest recognized stratigraphic units are open marine shales and limestones of Triassic Wulgai Formation, which might be

considered the first syn rift deposition in area of interest. At the end of Triassic, the Wulgai Formation was uplifted and eroded, followed by another phase of syn rift sedimentation in the Early Jurassic, resulting in the deposition of the Shirinab Formation. There are three member of Shirinab Formation Jurassic age, named as Spingwar, Loralai and Anjira. Syn-rift deposition of Shirinab Formation started with the deposition of fluvial sandstones and shoreface and inner carbonates shelf of the Spingwar Member, covered by limestones and shales of the Loralai and Anjira members (Ahmad et al., 2013).

In the period of Middle Jurassic, the accommodation space reduced probably due to a fall in relative sea level or basin subsidence which create a space for deposition of Chiltan Formation. Deposition of Chiltan Formation remained followed by an uplift and erosion in the period of early Cretaceous, which similarly terminated syn-rift sedimentation phase. The Spingar sands are considered primary reservoir target where Chiltan Formation is exposed on the surface, whereas Chiltan Formation is primary target in some parts of Khuzdar and south Khuzdar areas (Abbasi et al., 2016).

Fluvial systems and Marginal marine predominate in far eastern side of basin throughout the Early Cretaceous, depositing Sembar Formation shoreface sands and fluvial channels. Along through entrance of turbidites in front portion of the Kirthar Fold Belt, marine conditions occurred foredeep. In initial Cretaceous age, a restricted marine, anoxic environment occurred during the accumulation of the Sembar Formation that might be a more productive source rock in the Kirthar Fold Belt as well as in Lower and Central Indus basin (Abbasi et al., 2016).

Throughout the Middle Cretaceous, the transgressive-regressive cycles continued, depositing Lower Goru Formation shoreface sand in east section of Indus Basin. Only thin turbidites were transported to the basin floor through strong channels in the Kirthar Foredeep and Fold belt. Only carbonate and shale of marine deposit from the Goru Formation, mud, and carbonated calcite from the Parh Formation have been accumulated in Kirthar Fold belt and Foredeep since the research area was too far from the input sources (Bannert and Raza, 1992).

In the Late Cretaceous, the Kirthar Foredeep and frontal sections of the fold belt experienced an influx of sand-rich turbidites from the southeast due to the uplift of the Indian Shield. Until the early Paleocene, continued raising of the shield generated shallow marine conditions in the frontal area of the Kirthar Fold Belt, which favored accumulation of near surface marine and sands of fluvial deltaic environment in Ranikot/Khadro

Formation and Pab Formation. Middle and Late Cretaceous layers were extensively eroded in east side of the Indus Basin, resulting in base Tertiary unconformity that covers the whole basin. A flooding episode occurred during the early Paleocene, depositing thick Ranikot/Khadro Formation shale in the Kirthar Foredeep and Fold Belt. This interval was observed with a fall in ocean stage and then rise through late Paleocene, permitting precipitation of carbonates of Dunghan in Kirthar Folded belt (Kazmi and Jan, 1997).

In the west Kirthar, Dunghan carbonate are primary target in the Kalat and Margand area (North Kalat) where these are most likely composed of slope facies as the shelf edge is probably marked in the immediate northeast of the study area. Regional flooding was occurred throughout the basin during the Eocene, resulting in the deposition of thick shale from the Ghazij Formation, as well as limestone in shallow platform settings at highs (e.g., Khairpur Jacobabad High). The Indo-Pakistan Plate's strike with the Afghan Plate through the Oligocene resulted in the closure of the West Kirthar basin to sediment deposition. This led to the formation of isolated basins where deposition continued until the Quaternary period, as reported by (Kazmi and Jan, 1997).

## 2.9 Petroleum system of Mehar block

The petroleum system of Mehar Block is considered as on the key petroleum systems surrounding *KFTB* as briefly explained as below with major elements.

### 2.9.1 Source Rock

The northwest rim of Kirthar was the first proof of hydrocarbons in the Lower Indus Basin, and it predicted deep oil seeps. The Tajedi, Tando Alam, Khaskeli, Laghari, Dhabi and the Golrachi gas fields were discovered in Lower Goru sandstone. The Sari Song, Hundi, and Kirthar gas fields were discovered in Paleocene sandstone, while gas condensate was found in the late Eocene Laki Formation near Mazarani (Sultan and Gipson Jr, 1995).

As a result, production, movement, and accumulation of hydrocarbons must followed in Lower Indus Basin. Broad research show the Sembar Formation of late Cretaceous age is the vital hydrocarbon generating source rock which is organic rich shales deposited in shelf marine environment in this region and other areas of the Indus Basin (Kazmi and Jan, 1997).

### 2.9.2 Reservoir Rock

The Lower Goru Formation, which is of early Cretaceous age, is a significant producing reservoir. The principal reservoir of Mehar blocks found in the Pab Formation and Lower Goru formations is deltaic and shallow-marine sandstones with beds of shale. Based on such layers, the Lower Goru is divided into upper, middle, and basal and massive sands. Reservoir quality disintegrates as you travel west, while reservoir thickness increases (Khan et al., 2016).

### 2.9.3 Seal Rock

The system's known seals are made up of shales interbedded in (Lower Goru Shales) and overlaying (Upper Goru Shales) the reservoirs. Thin beds of shale are consisted of variable thickness are also acting as good seals in many producing fields (Alam et al., 2003). There is very thick sequence of shales and marl overlying the Sawan reservoir sands which act like local seal. Shales present in Goru Formation similarly behave as base and horizontal seal (Jadoon et al., 2020).

### 2.9.4 Trap

Trap is an essential part; without it the hydrocarbon petroleum play would be incomplete. The movement of hydrocarbons is stifled and sealed by impermeable rocks. The Zamzama structure, which is north-south trending, is a massive structure (thrusted anticline) that leans east, making it a big trap (Jadoon et al., 2020). The major elements of the Mahar Block are enlisted in Table 2.2.

**Table 2.2:** Petroleum play of Mehar Block.

| Play Elements | Formations | Age |
|---|---|---|
| Seal | Upper Ranikot/Khadro | Paleocene |
| Reservoir | Pab Formation /Lower Ranikot/Khadro | Late Cretaceous/Paleocene |
| Source | Sembar Shales | Early Cretaceous |

## 2.10 Regional Correlation of the Study Area

The regional petroleum system covering the areas surrounding the *KFTB* is presented in Table 2.3. The key source, reservoir, seal and traps along with their ages are shown. The petroleum system is producing in many major fields including Bhit, Bhadra, Haleel, Mehar, Mubarak, etc. The study area comprising the two fields including Mehar and Zamzama Gas Fields focusing on the Paleocene reservoir Ranikot/Khadro Formation.

**Table 2.3:** Petroleum Play of the study area (Zafar et al., 2018).

| Age | Source Rocks/Formation | Reservoir Rocks/Formation | Seal Rocks/Formation |
|---|---|---|---|
| Eocene | Kirthar (Limestone/ Shale) | Kirthar (Limestone) | Kirthar/Laki (Shale) |
| | Laki/Ghazij (Limestone/ Shale) | Laki/Ghazij (Limestone) | Ghazij (Shale) |
| Paleocene | Lakhra (Shale) | Lakhra (Sandstone) | Lakra/Bara (Shale) |
| | | Ranikot/Khadro Sandstone | Ranikot/Khadro shale |
| | Bara (Shale) | Bara (Sandstone) | Ranikot/Khadro (Shale) |
| Cretaceous | Mughalkot (Limestone) | Pab (Sandstone) | Mughalkot (Shale) |
| | Goru (Shale) | Goru (Sandstone) | Goru (Shale) |
| | Sembar (Shale) | Sembar (Sandstone) | Sembar (Shale) |

The regional correlation is constructed between Mehar-01, 02, 03, and Zamzama-08-ST2, 04-ST3, 07, 05, 03, based on gamma ray (*GR*) and Sonic Compressional Wave (*DT*) logs (Figure 2.3). The main aim is to assess the depth variations of Paleocene reservoir Ranikot/Khadro Formations along with the sand's depictions based on low *GR* and high *DT* logs. The complete suite of Mehar-01 and Mehar-02 is available that is utilized further in petrophysical properties estimation while Mehar-03 has limited logs. Similarly, all of the Zamzama wells (Zamzama-02, 03, 04, 05, 06, 07, and 08) have basic suit of logs therefore all the wells are employed in the processes including, petrophysics, elastic, and their mutual petro-elastic relationship.

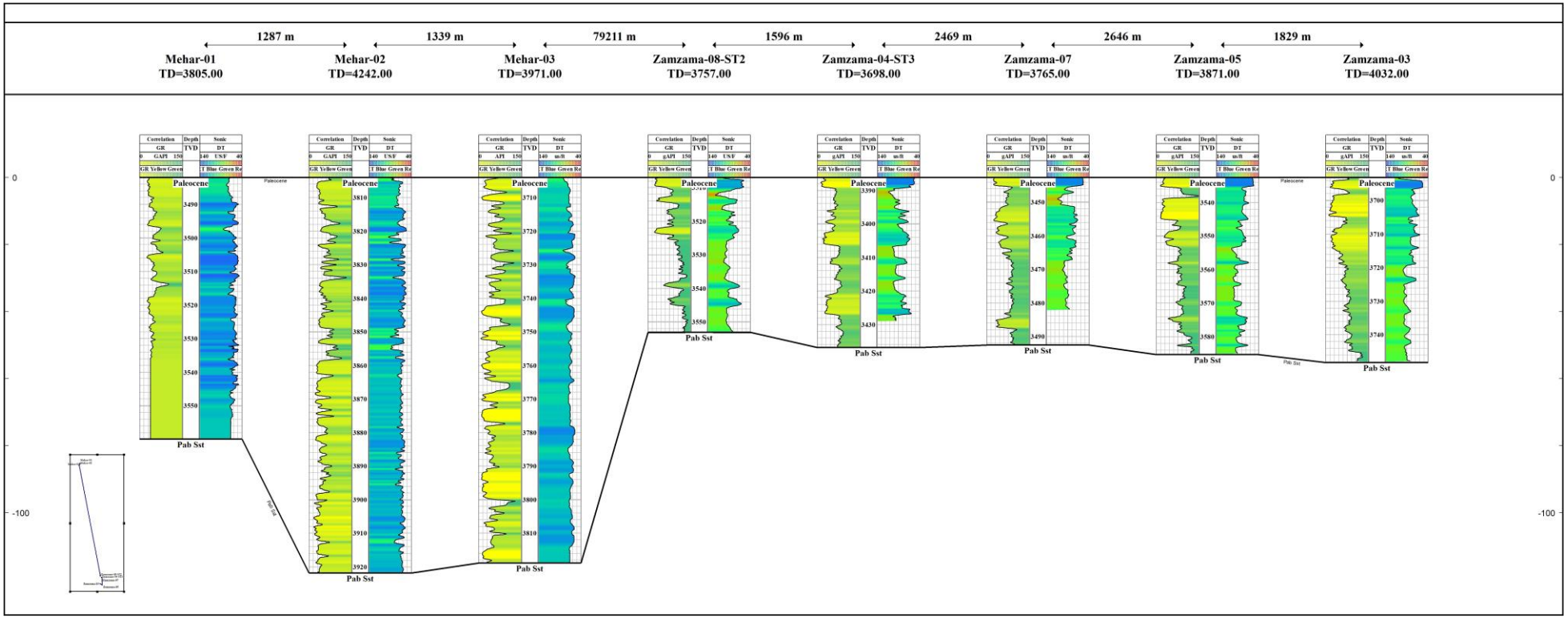



**Figure 2.3** Well correlation connecting wells of Mehar and Zamzama Blocks based on *GR* and *DT* sonic logs.

## 2.11 Summary

The chapter comprised of the geologic and tectonic of the Mehar and Zamzama gas fields. The fields are present in the KFTB of the Lower Indus Basin which is significant for many of the major gas producing fields of Pakistan. Tectonically, anticline is cut by major thrust fault in north-south direction along with the sub-thrust in the similar orientation. The lateral ramps connecting the hanging and footwall caused differential water encroachment and pressure depletions in the gas producing wells. A valid petroleum system is present in the fields i.e., Mehar gas field (seal: Upper Ranikot/Khadro, reservoir: Pab Formation/Lower Ranikot/Khadro, source: Sembar Sahle) and Zamzama gas field (seal: Ranikot/Khadro, reservoir: Pab Formation, source: Sembar shale). The latest and robust *ML* techniques are employed in the recent times for improving the frequency contents that enable the illumination of below resolution beds and captured the heterogeneities. The study is focused to implement the latest techniques to assess the Ranikot/Khadro Formation for its thickness issues along with its heterogeneities. The pressure depletion needs to be evaluated by much reliable, non-linear relationships of *ML*, therefore a novel approach for PP prediction is to be employed.

# CHAPTER 3

# RESEARCH METHODOLOGY

With the advent of technology and advancement in the computational world, today researchers are incorporating new algorithms to help solve key real-life challenges which were not thinkable couple of decades back. Statistical techniques have allowed to analyze subsurface in a manner which allows estimations of potential reservoirs facies which were not possible using conventional techniques. Furthermore, integration of big datasets over statistical approach via *ML* made it possible to minimize uncertainties associated with high success probabilities of discovering reservoirs with their true potentials.

## 3.1 Research Approach

Implementation of advanced geophysical interpretation techniques over a geologically complex reservoir has become a norm in the industry. Advancements in technology and algorithms have allowed geophysicists to analyze regions that were impossible to image and interpret a decade back. However, such advanced algorithms were based on sophisticated statistical techniques, which require intense computation power to execute to obtain reliable results. In this study, an optimized approach is devised to help address the problem statement, thereby, for the same, and elaborated workflow shown in Figure 3.1 which highlights the integral steps opted to successfully accomplish this study and a flowchart outlines the core workflow of supervised ML utilized here. The process starts with a dataset containing both the independent variables (input features) that the model will analyze and the dependent variables (target variables) that we aim to predict. The raw data undergoes a crucial step called feature engineering, where relevant features are selected, transformed, or even newly created to optimize the model's ability to learn patterns. Next, a portion of the data, known as the training set, is used to train the model. During this training phase, the model essentially learns the relationships between the input features and the target variables. The model's performance is then evaluated on a separate, unseen portion of the data called the testing set. This evaluation ensures the model can generalize its knowledge and doesn't simply memorize the training data. Finally, if the evaluation is satisfactory, the trained model can be used for prediction. In the prediction stage, the model takes entirely new, unseen data as input and makes predictions for the target variables based on the relationships it learned during training.

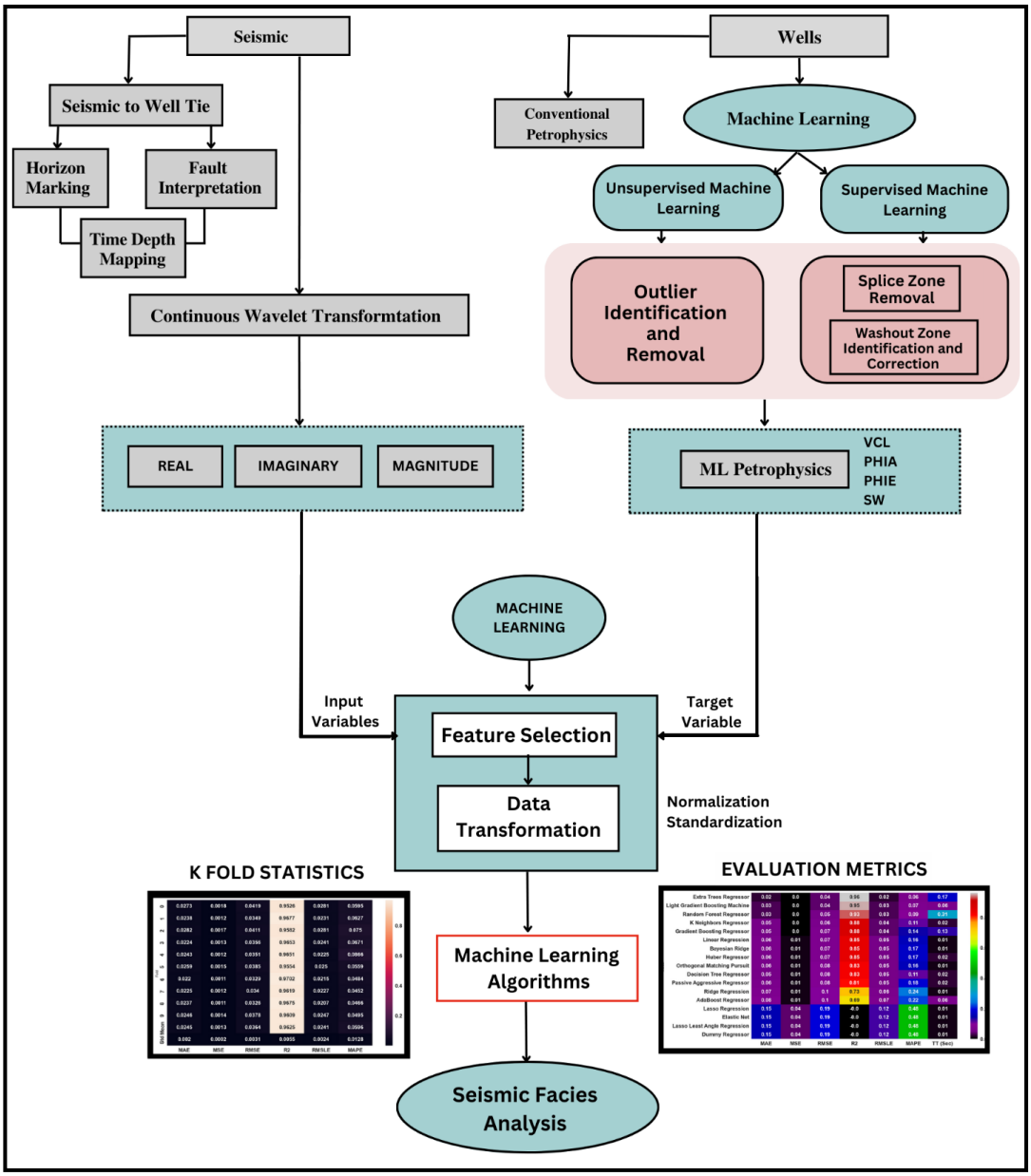



**Figure 3.1** Systematic flow of the adapted step wise approach to perform the research along with the input data set.

### 3.1.1 Petrophysics

The objective of petrophysics is to assess the reservoir characteristics more reliably with the benefit of making the right decisions regarding drilling operations based on the data gathered, completion, and testing. It provides the necessary parameters that enable more realistic models of reservoirs, i.e., static model. Overall, the petrophysicists has a significant role for the success of wells and bridges various fields operation such as production geology, production technology, reservoir engineering, and exploration seismology. The rock's physical and chemical characteristics are analyzed during

petrophysics as well as the fluids within the rock pores. The petrophysical properties are estimated by employing the well logs, while they can also be estimated in the laboratory by the core data as well. The petrophysics analysis retrieves both qualitatively and quantitatively properties quite remarkably. The volumetrics, such as water and hydrocarbons, the presence of shale, the identification of source rock, and the aquifer quantities, can be attained by the relationships among the properties obtained from petrophysics. The properties attained are of keen interest for petroleum engineers, geologists, geophysicists, and core analysists.

Many times the seismic data becomes unable to explain various questions in geophysics regarding zone of interest (MacGregor and Andreis, 2012). The petrophysics is the study that employed variety of dataset including well logs, cores, and production data to obtain more realistic physical properties by integrating all the available data. The approach is majorly concerned with hard constraint data of well bore at a certain location therefore having not much reliance on seismic. The key reservoir characteristics included saturations, permeability, porosity, net pay thicknesses, shale volumetrics, and the contacts between fluids. Nowadays, the comprehensive assessment of reservoirs is made by carefully selecting the wireline suits for interpretation that optimize the accuracy in saturations, porosity, lithologies, permeability, etc.

Primarily on evaluating the advanced *ML*-based models to determine the reservoir's petrophysical characteristics and thus build a locally validated petrophysical model that can be employed in future wells with good data quality control. The novel approach uses two wells to predict the petro-physical properties precisely and accurately. Several *ML* algorithms are used to identify the best model for accurate interpretation of well logs at different well locations. One-class-*SVM* and tree-based approaches are employed to analyze and quantify the uncertainties associated with borehole log data. *ML*-based algorithms such as *RF* and Tree-based Regression have been employed to estimate the reservoir properties. The heat map of *ML* algorithms performance has helped to identify the best model for accurate predictions of blind wells. Based on the statistical performance metrics, it was determined that the Extra Trees Regressor performed exceptionally well in the estimation of volume of shale and porosities as compared to conventional methods. This is attributed to its ability to effectively model reservoir properties by properly identifying patterns in the training data. Our results indicate that amongst tree-based regressions, the Extra trees regressor and *RF* are superior in terms of accuracy and reliability in reservoir properties prediction. The applicable *ML* algorithms

have successfully overcome the limitations of conventional methods, such as generalization and data range, for petrophysical prediction without the need for extensive use of geological or lithological characteristics of the reservoir formation with time efficient approach and optimized results.

The primary objective of this research is to determine the most optimal *ML* model for efficient quality control of well log datasets and accurate prediction of petrophysical properties and a good comparative analysis with the conventional techniques of petrophysical interpretation. *ML* algorithms can generally be classed into supervised learning, unsupervised learning, and reinforcement learning.

In this case study, an unsupervised *ML* approach was applied in order to identify the outliers. Outliers are data points usually called noise that do not fit within the data set's normal statistical distribution. Measurement errors, poor data acquisition methods, washed-out boreholes, or problems with tools and sensors can all lead to outliers, causing our results to be skewed. Here, several algorithms such as least squares support vector machines (*LSSVM*), isolation forest, minimum covariance, etc. were employed on the raw logs to calibrate the appropriate model for outlier identification and removal. The One-Class *SVM* was chosen as the best outlier identifier because it removed the appropriate (not less, not much) number of outliers. After detecting the outliers, supervised algorithms such as Gradient Boosting, Extra Tree Regression, *RF*, Ada Boost etc. were used to train the inliers and predict the outliers (Chen et al., 2020). The washout zones affected raw well logs have been corrected via supervised *ML* algorithms. In the case of the washout zone affected density log, *ETR* predicts the density log, while the *GBR* predicts the missing *DT* log in the case of the splice zone affected *DT* log. Furthermore, the petro-physical properties like the volume of shale, effective porosities, and water saturations are predicted via raw logs after quality control through integrated approach by utilizing both unsupervised and supervised *ML* algorithms. The performance of others *ML* algorithms has been outdated by the ensemble (*ETR* and *GBR*) *ML* algorithm’s performance at this step (Gültekin et al., 2019).

#### 3.1.1.1 Marking Zone of Interest

There are three important criteria that are demonstrated to recognize the zone of interest in a reservoir.

First criteria is to analyze the value of *GR*. It is one of the important factor that is used for prediction of hydrocarbons zones. A lower value of the *GR* shows the presence of clean lithology, which is the characteristic of the excellent reservoir. If the *GR* value is high in a certain zone then it cannot be marked as a clean/sand/hydrocarbon bearing zone (Khan et al., 2021).

Second criteria is the identification of good porous zone which can be highlighted by the crossover between the neutrons and density log for identifying the area of hydrocarbon. The reservoir's porosity is determined with neutron and density log data. If both the neutron and density values continue to decrease, the crossover with both logs decreasing is obtained, indicating the presence of the hydrocarbon zone (Ali et al., 2019).

Third criteria is the evaluation of resistivity log (*RES*). The RESs are significant because it reflects the presence of a reservoir filled with hydrocarbons that depend on electrical resistivity. There are three valuable *RES* logs known as Laterolog Shallow (*LLS*), Laterolog Deep (*LLD*) and Micro Spherical Focused Log (*MSFL*). If the *LLD* is greater than *LLS* and *LLS* is greater than *MSFL* in a certain zone than it shows the presence of hydrocarbon. In other word there should be a segregation among the *LLD* and *LLS* for existence of hydrocarbon. If there is no separation between them, that area may hold water (Ahmed et al., 2022).

#### 3.1.1.2 Calculation of Volume of Shale ($V_{SH}$)

The shale volume is computed from *GR* log data. The amount of natural radioactivity in the formation influences the *GR* log value. Because shale contains many radioactive materials, it has a high value of *GR* log.

Shale volume can be estimated by using the method (equation 3.1) (Rider, 1986)

$$\boldsymbol{V_{SH}} = \frac{\boldsymbol{Gr_{log} - Gr_{cln}}}{\boldsymbol{Gr_{shl} - Gr_{cln}}} \qquad (3.1)$$

Where,

$V_{SH}$ = Volume of shale.

$Gr_{log}$ = Value of *GR* at interested depth.

$Gr_{cln}$ = Minimum *GR* value (Generally an average lowest value *GR* log).

$Gr_{shl}$ = Maximum *GR* Value (Generally an average highest value of *GR* log).

### 3.1.1.3 Estimation of Volume of Clean ($V_{CL}$)

The formula in equation 3.2 is used for calculation of clean volume (Rider, 1986)

$$V_{CL} = 1 - V_{sh} \quad (3.2)$$

### 3.1.1.4 Calculation of Porosity ($\emptyset$)

Porosity with the help of density can be calculated with the help of density log. Following formulation (equation 3.3) is used to evaluate porosity (Schlumberger, 1974).

$$\Phi = \frac{\rho m - \rho b}{\rho m - \rho f} \quad (3.3)$$

Where,

$\rho m$ = Density of matrix.

$\rho b$ = Formation bulk density.

$\rho f$ = Fluid Density.

$\Phi$ = porosity of rock.

### 3.1.1.5 Estimation of Neutron Porosity (*NPHI*)

The *NPHI* value is brought straight from the neutron log, which goes corresponding to the density log.

### 3.1.1.6 Estimation Average Porosity (*APHI*)

The total number of pores present in a rock is referred to as its porosity. These pores are a critical component of the reservoir for hydrocarbon buildup. The *APHI* is calculated using the density and neutron log values. The *APHI* of a reservoir is computed applying the formula given below (equation 3.4) (Rider, 1986).

$$APHI = \frac{NPHI + DPHI}{2} \quad (3.4)$$

### 3.1.1.7 Estimation of Sonic Porosity (*SPHI*)

Sonic log is used for calculation of porosity. Need of *SPHI* can be used when the condition of bore hole is too much bad. So, the sonic log is only log that helps to know about correct porosity. The formula in equation 3.5 of calculation of *SPHI* (Rider, 1986).

$$SPHI = \frac{\Delta t - \Delta t\ ma}{\Delta t fld - \Delta\ tma} \quad (3.5)$$

Where,

$\Delta t$ = sonic log values (μs/feet).

$\Delta tma$ = matrix travel time (μs/feet).

$\Delta tfld$ = fluid travel time (μs/feet).

#### 3.1.1.8 Estimation of Effective Porosity (*PHIE*)

Fraction of connected pore spaces in relation to the bulk volume is known as *PHIE*. The *PHIE* is a parameter utilized in computations as it characterizes the interconnected pore spaces that comprise of recoverable hydrocarbon fluids. *PHIE* is computed using the equation 3.6 (Kamel and Mohamed, 2006).

If condition of borehole is good

$$\boldsymbol{PHIE = APHI * Vcln} \quad (3.6)$$

Use equation 3.7 if condition of borehole is bad (Caving),

$$\boldsymbol{PHIE = SPHI * Vcln} \quad (3.7)$$

As our zones of interest have undergone caving, so *PHIE* is calculated using sonic log porosity.

#### 3.1.1.9 Resistivity of water ($R_w$) Calculation

Resistivity of water ($R_w$) is calculated with help of Pickett plot method. Pickett plot method represent deep resistivity on x-axis and porosity log on y-axis. The plot is based on the logarithmic of Archie equation. Water Saturation points ($S_W$) will be plotted on a straight line with a negative slope of value m. The lower most line on the plot is defined by water zones. The $R_w$ may be estimated from a location on the line because $S_W$ = 1. Other parallel lines with different $S_W$ can be drawn once the water line is formed, assuming a constant *n* (usually 2). Then, in terms of $S_W$, other data can be plotted and evaluated. The Pickett plot is used to estimate the resistivity of water Ranikot/Khadro Formation. The Ranikot/Khadro Formation resistivity is estimated to be 0.076 ohm-m.

#### 3.1.1.10 Saturation of Water

$S_w$ refers to the amount of water in a formation that is not equal to the amount of hydrocarbon present. $S_W$ denotes the saturation of water. It's computed by adding 1 by a $S_W$. In our case Archie equation 3.8 is used for calculation of saturation of water.

$$Sw = \sqrt[n]{\frac{a \times Rw}{R_t \times PHIA^m}} \quad (3.8)$$

$S_W$ = Saturation of Water.

$a$ = Tortuosity factor.

$m$ = Cementation factor.

$n$ = Saturation exponent.

$R_w$ = Resistivity of formation and water.

$PHIA$ = Average porosity

#### 3.1.1.11 Well Data Quality Control

Quality control is necessary for accurate interpretations of well logs. Since well logs contain uncertainties that need to be removed through conventional methods or *ML* methods. Through unsupervised *ML* algorithms and supervised *ML* algorithms, the quality control of the well logs is performed for accurate interpretation of well logs.

#### 3.1.1.12 Outliers Identification and Removal through Unsupervised *ML* algorithm

Unsupervised learning is a type of *ML* algorithm that operates on data without requiring explicit labels or supervision. Unsupervised learning algorithms are employed to find underlying or hidden patterns in the data that are difficult to notice. It includes clustering algorithms, or dimensionality reduction algorithms, which project high-dimensional data onto a lower-dimensional space while preserving as much of the original structure as possible. The detail unsupervised *ML* algorithm is depicted in Figure 3.2. There are several unsupervised algorithms such as One-Class *SVM*, K-means, Principal component analysis (*PCA*) etc. (Ghahramani, 2003; Dreiseitl et al., 2010; Alloghani et al., 2020 ; Batta et al., 2022).

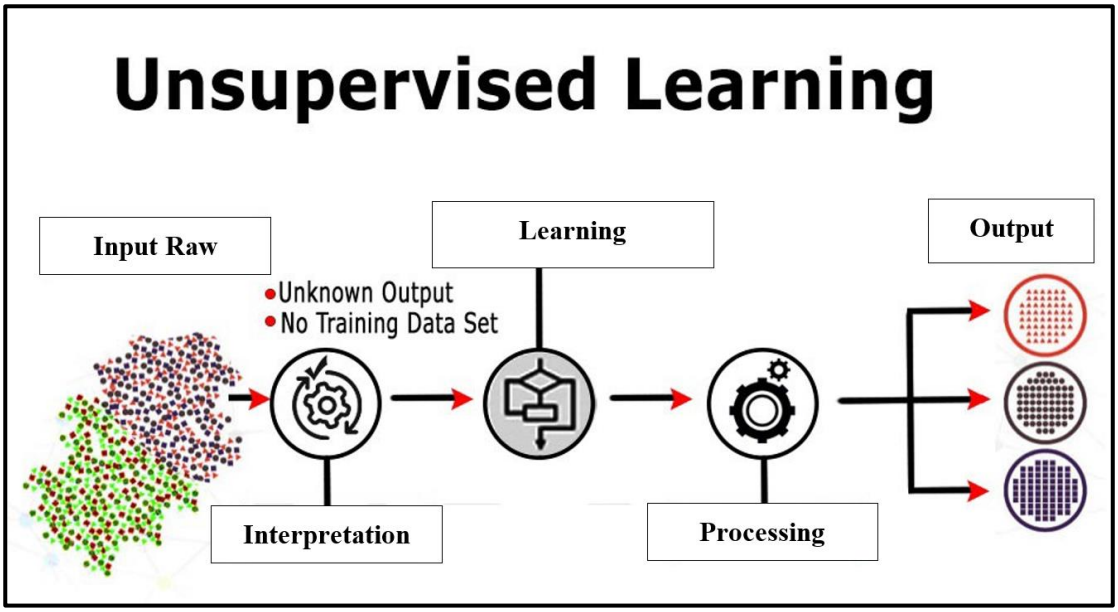



**Figure 3.2** Unsupervised learning approach inputs the data, recognized the data patterns after its training and evaluation and finally predict the objective features (Batta et al., 2022).

#### 3.1.1.12.1 One Class *SVM*

One-Class *SVM* is an unsupervised learning technique used for novelty detection or identifying instances that are significantly different from the majority of the data. The basic aspect behind One-Class-*SVM* is to minimize the hyperplane of the specific class in the training data and to consider all other samples outside of the hyperplane to be outliers or outside the distribution of the training data. Positive numbers are treated as class +1, while negative integers as class -1. Samples in the training dataset are always taken as positive (Küçükdemirci and Sarris, 2022).

In this case study, different well logs such as *GR*, Bulk Density Log (*RHOB*), *NPHI*, sonic (*DT*) and *RES* specifically *LLD* were taken to visualize the presence of outliers using the Pair-Plot method. The red circles show the presence of outliers, which can affect the petrophysical interpretation of the well logs.

### 3.1.1.13 Splice Zone Identification and Removal through Supervised *ML* algorithms

Supervised learning is the most prevalent machine-learning task. It learns a function that relates an input to an output using sample input-output pairs. It derives a function from labeled training data, which consists of a set of training instances as given in Figure 3.3. Classification and regression are the two subtypes of supervised learning.

Based on the data used to train the model, classification tasks convert discrete input values to output categories or labels (e.g., facies classification, reservoir rock typing). Regression models translate the input data to a continuous numerical output (e.g., continuous reservoir properties prediction) (McDonald et al., 2021). There are several supervised learning algorithms such as *DTR*, *SVM*, *RF*, *BT* etc. (Batta et al., 2022; Kotsiantis, Zaharakis and Pintelas, 2007; Caruana and Niculescu-Mizil, 2006).

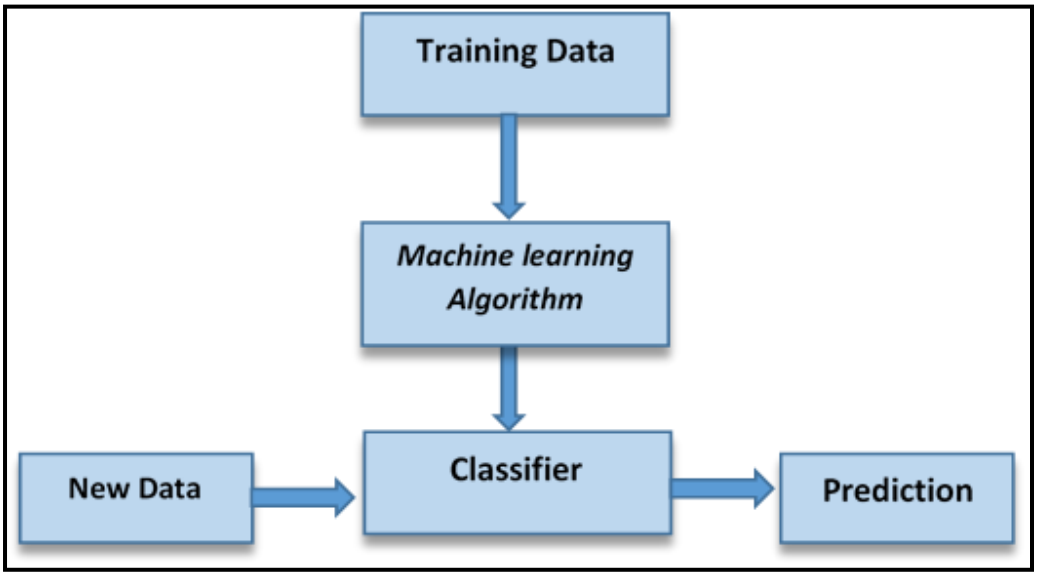



**Figure 3.3** The generalized diagram representing the workflow of supervised *ML* approach. Input the data, train the algorithm based on provided features then classify it and predict the new/unknown data (Batta et al., 2022).

#### 3.1.1.14 Washout Zones correction through Supervised *ML* algorithms

The logs normal data range is distributed in many loosely or poorly cemented sedimentary formations along with other conditions prevailed during wells boring. These scenarios create the well bore abnormality in size and hence creating the washout zones. The washout zones of all the wells are corrected via supervised *ML* algorithms. The raw logs are brought to their normal ranged according to the lithological characteristics.

#### 3.1.1.15 Advanced *ML* Methods

Petrophysics is a crucial discipline for determining the characteristics of reservoirs and developing new fields. Improved drilling efficiency, data repair, reservoir property prediction, reservoir rock type, and other petrophysics-related tasks have all

benefited from the use of *ML* (Banas et al., 2021; Akkurt et al., 2018). The integrated workflow based on the advance *ML* technique for the prediction of petrophysical properties is depicted in Figure 3.4.

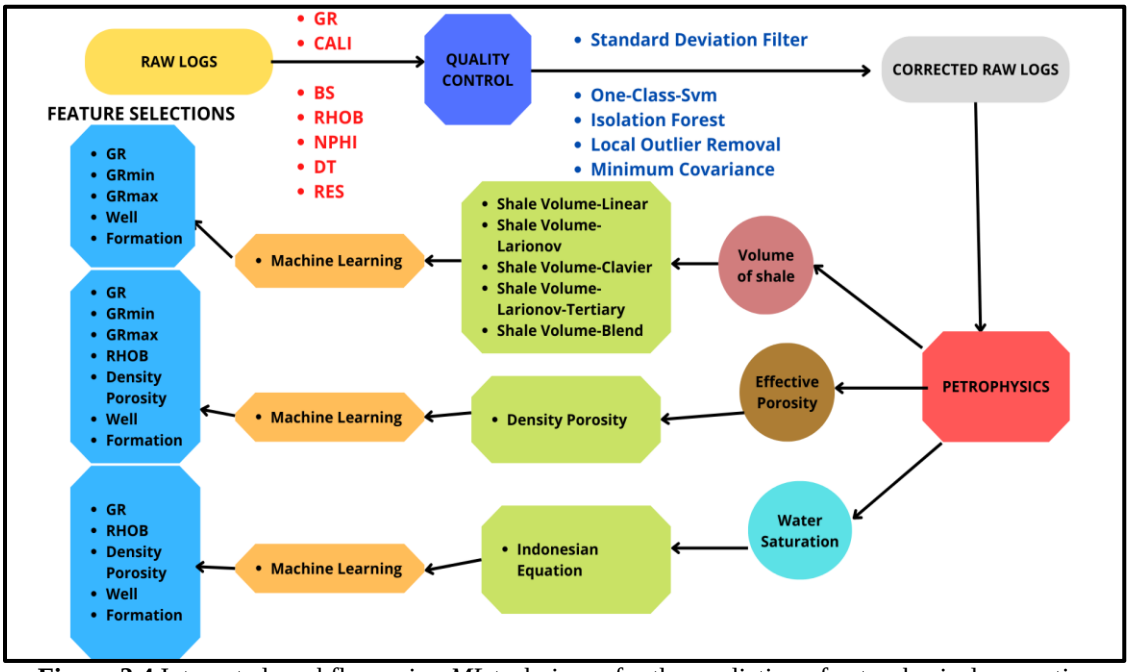



**Figure 3.4** Integrated workflow using *ML* techniques for the prediction of petrophysical properties.

### 3.1.2 Seismic Data Interpretation

Seismic data interpretation is important in the field of geophysics and geology because it allows us to gain insight into the subsurface of the Earth. Seismic data is collected by sending sound waves (or seismic waves) into the ground and measuring the response of these waves as they bounce off different layers of rock and sediment. It involves analyzing these measurements to create images of the subsurface structure of the Earth. This can help geoscientists to understand the location and characteristics of different rock formations, faults, and other geological features that may be important for a variety of applications, including exploration for oil and gas and mineral exploration.

The seismic surveying approach is undoubtedly the most employed geophysical hydrocarbon exploration technique. This is primarily because it can detect subsurface structure that range in size from small to large (Björk et al., 2010). The basic aim of seismic interpretation is the consistent characterization of geology, and all aspects of interpretation work towards this objective. Refraction seismic data interpretation must be correct for oil and gas exploration to be effective (Herron, 2011).

### 3.1.2.1 Workflow for Seismic Interpretation

Following workflow is adopted for completion of conventional seismic interpretation (Figure 3.5).

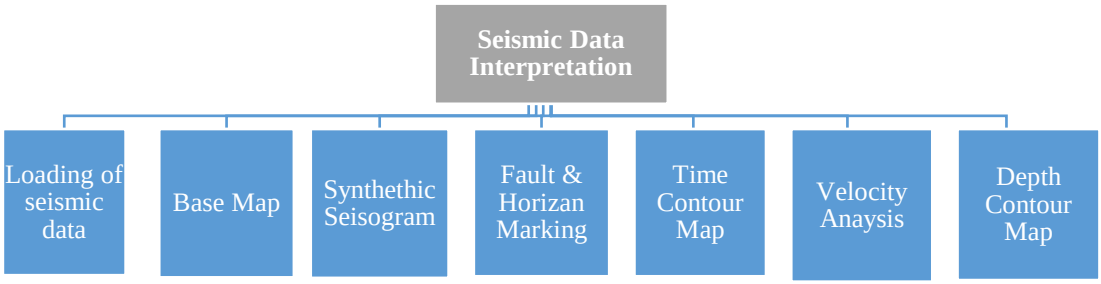



**Figure 3.5** Workflow adopted for Conventional seismic interpretation.

### 3.1.2.2 Loading of Seismic Data

A *3D* seismic data cube in SEG Y format for the Zamzama Gas Field is utilized in the present study. The data is loaded into Discovery Geographix 2019.4 using SEG Y headers for inline and cross line loading. For this study, a fifteen-square-kilometer *3D* seismic cube of the Zamzama Block is provided among its seven wells, i.e., Zamzama - 02, Zamzama-03, Zamzama-04, Zamzama-05, Zamzama-06, Zamzama-07, and Zamzama-08-ST-2. Also, the well data sets of Mehar-01 and Mehar-02 are employed in the interpretation using their log curves. The integrated workflow comprises all the available data sets utilized for the identification of litho-stratigraphic units as well as seismic tie correlations.

### 3.1.2.3 Base Map

After seismic and well data loading is complete, the base map of the given seismic cube of Zamzama *3D* along with its wells is generated whereas Mehar wells are outside of this *3D* seismic cube. The base map shows the orientation of the overall grid. In the *3D* cube, there are two types of lines: one is known as an inline, and the other is a cross line. The Cross lines start from 40 to 300, and the In lines range between 350 and 650 shown in Figure 3.6. The orientation of inlines is from east to west, and for a cross line, the orientation is from north to south.

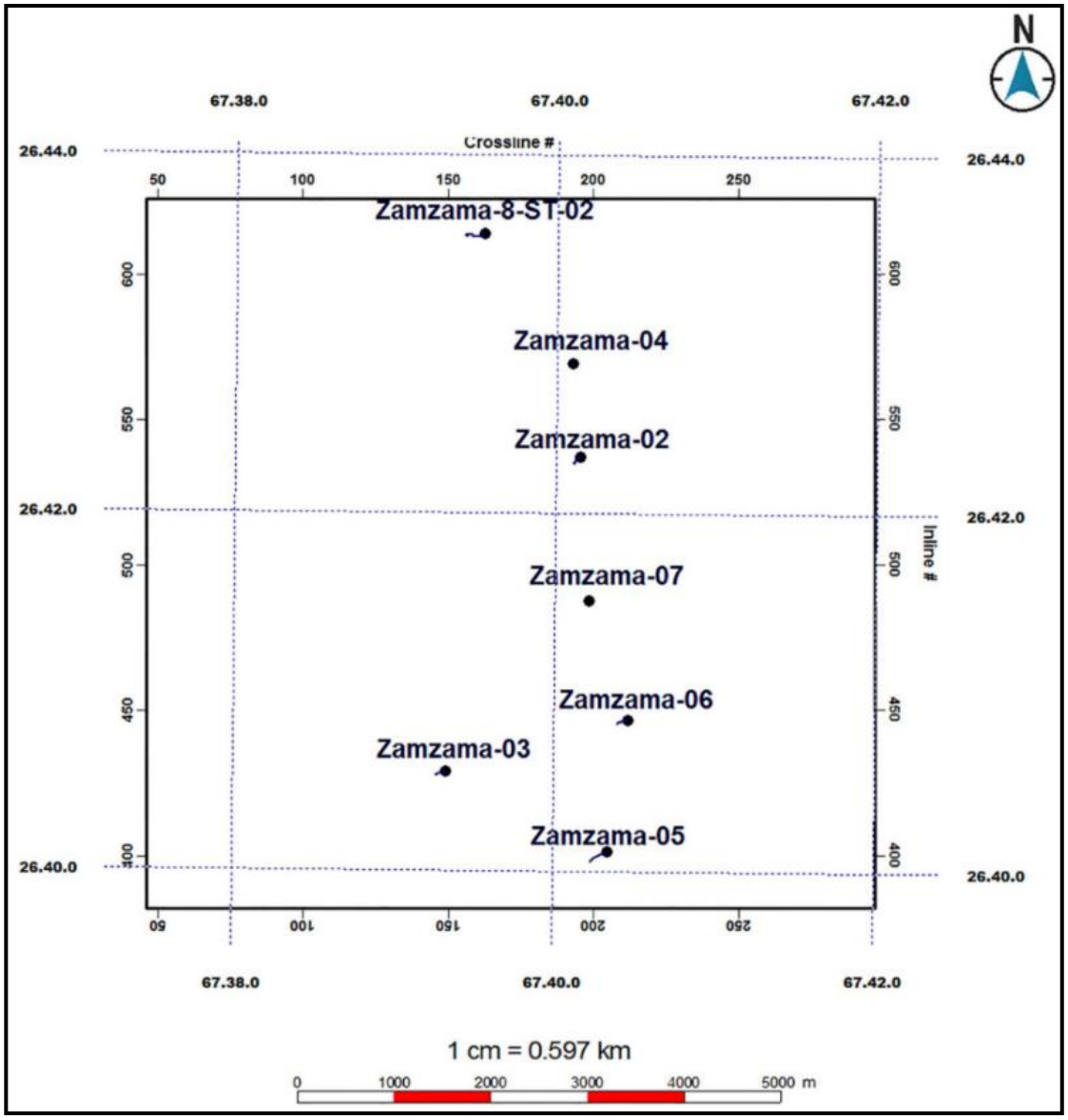



**Figure 3.6** Extents of *3D* seismic cube of Zamzama Gas Field with available wells of Zamzama (inside *3D* seismic cube).

### 3.1.2.4 Generating Synthetic Seismogram

A synthetic seismogram is a 1-D forward model of acoustic energy that travels through the earth's horizons and has an acoustic impedance contrast. The synthetic, using Zamzama wells, was created by the merging of earth reflectivity series that were obtained from the logs such as sonic and density. The wavelet required for synthetic generation was extracted using *3D* seismic cube of Zamzama. The degree of correspondence between a seismic section and a synthetic seismogram is dependent on the quality of well log data. *DT* and density (*RHOB*/*ROHZ*) logs are required to produce a synthetic seismogram. Before marking the horizons, a synthetic seismogram of the given wells is generated (Hampson et al., 2001).

For a well to seismic tie, a seismogram is required. Therefore, sonic and density logs from the Zamzama-03 and 05, along with their *GR* logs as references, are employed in well to seismic correlation. The seismic data traces surrounding the wellbore were used to form a wavelet, which was then convolved with the impedance log generated from the well's sonic and density logs to produce a final synthetic seismogram. After that, the synthetic was compared to the seismic data, and changes were performed using shift and stretch. The well-to-seismic relation was completed once the nature of synthetic and seismic data were matched, and work moved on to the next phase, which was the selection of horizons using the time depth chart correlation produced during the tie.

The synthetic seismograms based on Zamzama-03 and Zamzama-05 are displayed in Figure 3.7 and Figure 3.8, respectively. The Figure 3.7 demonstrations that the time of Ranikot/Khadro Formation is 2195 msec in Zamzama-03 and Figure 3.8 demonstrate the time of Ranikot/Khadro Formation that is 2125 msec for Zamzama-05 well.

#### 3.1.2.5 Seismic to Well Tie

The synthetic seismogram creation and execution of seismic to well tie is due to the reason of marking the exact reflector of horizons on the seismic section. Figures 3.7 & 3.8 show the exact location of formations after the accomplishment of well to seismic ties, along with extracted wavelets, their amplitude, and the correlation coefficient. Therefore, the targeted formations, i.e., Ranikot/Khadro Formation, Pab Formation, and Fort Munro, are marked at their corresponding times throughout the seismic data of the research area.

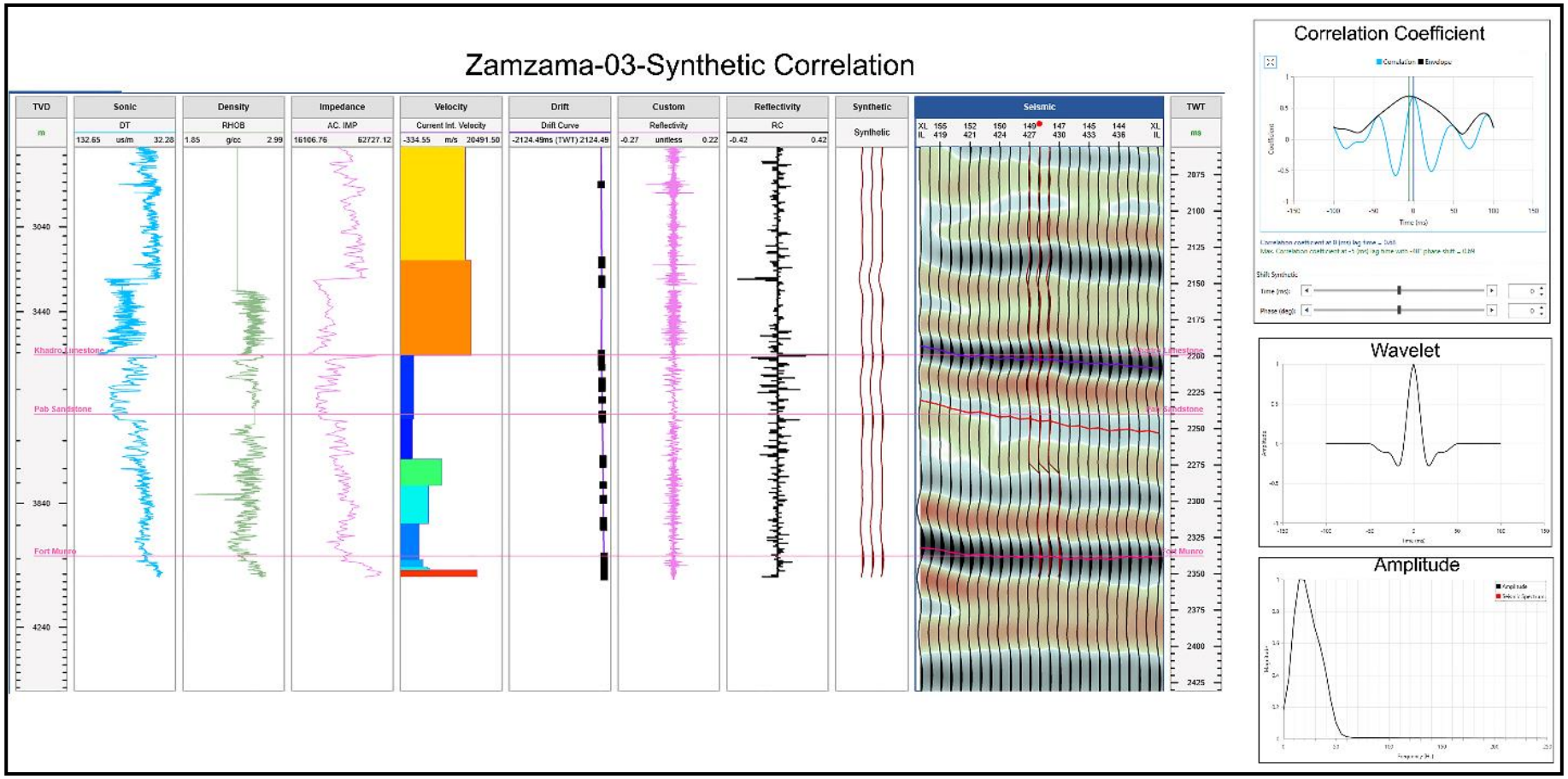



**Figure 3.7** Synthetic seismogram based on Zamzama-03 along with correlation co-efficient, wavelet and amplitude spectrum.

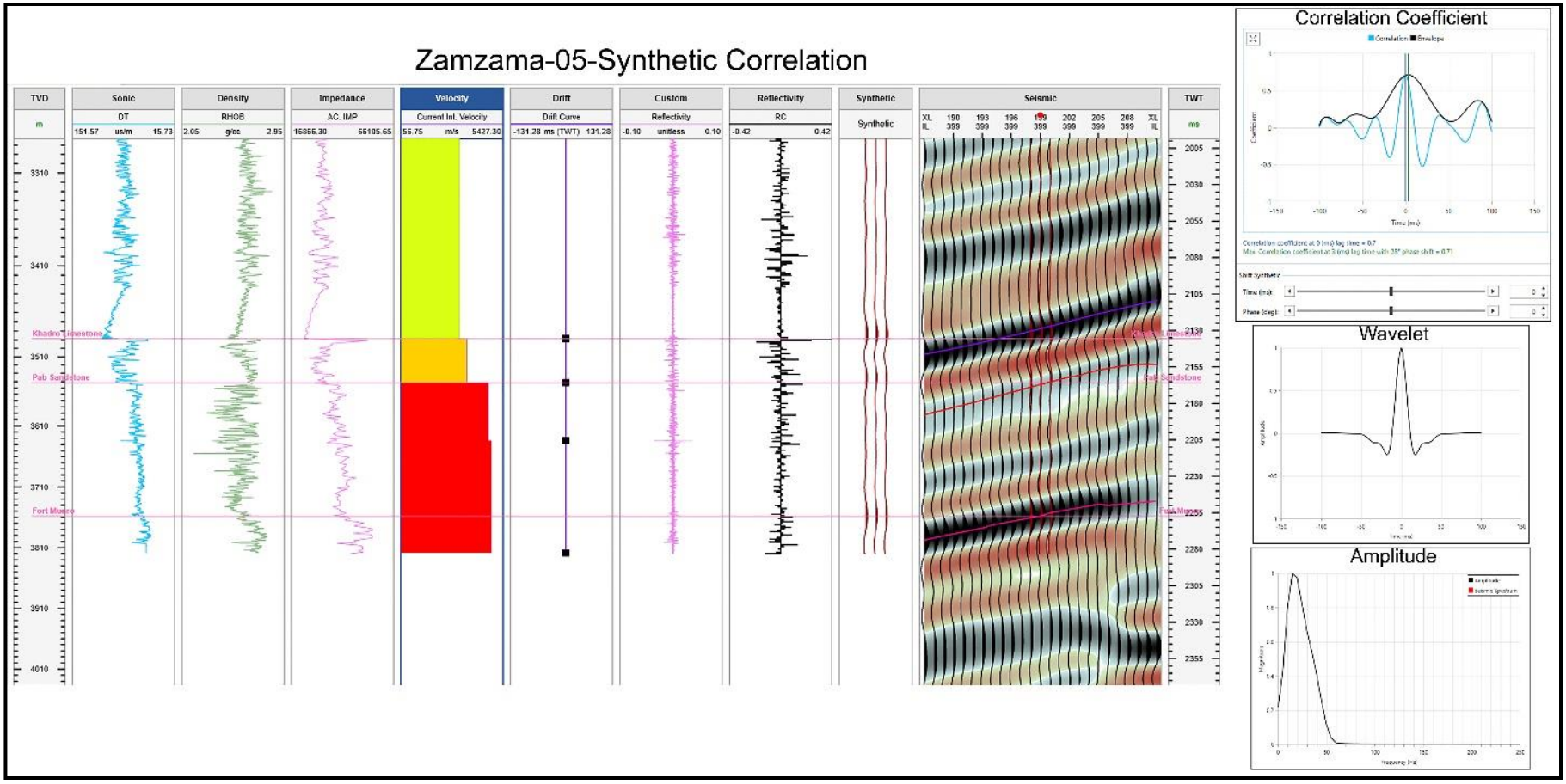



**Figure 3.8** Synthetic seismogram based on Zamzama-05 along with correlation co-efficient, wavelet and amplitude spectrum.

#### 3.1.2.6 Marking of Fault and Horizons

Following the ties, the marking of the concern reflector is initiated, that is in correspondence with the tops present in the wells. The Time-Depth graph created through the seismic to well tie will then be used by software to show these formations top (depth domain data) on the seismic (time) accurately. The primary target of structure interpretation is to mark the Ranikot/Khadro Formation, Pab Formation, and Fort Munro formations. The research is focused on the thin bed of Ranikot/Khadro Formation but to fully justify the structure and to achieve the best results, Pab Formation and Fort Munro Formation are also interpreted. The reflectors on the seismic data are highly visible and can easily be identified and recognized, thus aiding in data matching, particularly during seismic to well tie (Khan et al., 2021).

After the approval of reflectors at the concerned formation tops, horizons and faults have been marked at specific intervals in the *3D* seismic cube. For consistency of interpretation, often every 5th line is interpreted. As every 5th line was interpreted, a hunting tool was used to interpolate the data across the entire cube, allowing the software to automatically select the appropriate horizons based on marked interpretation and similar reflector characteristics. Similarly, faults were interpreted along with the interpretation of horizons throughout the entire cube, and then fault surfaces for the entire data cube were created.

### 3.1.3 Seismic Attributes

The physical properties (amplitude, phase, frequency, etc.) extracted from the seismic data set are designated as seismic attributes and analyzed for enhanced information that might be subtle in conventional data. Such extraction of information leads to improved geophysical and geological interpretation (Hampson et al., 2001; Chopra and Marfurt, 2005). The attribute extraction is simply the boosted characteristic of seismic data that not only improves the reservoir's understanding qualitatively but also improves the quantification of reservoir properties in the area of interest. The effectiveness of seismic attributes is that their sensitivity toward the geological feature or their enhanced property allows the interpreter to delineate the trends in deposition, structure, etc., thereby elaborating practicably on the reservoir properties (Raef et al., 2015; Marfurt et al., 1998; Taner et al., 1994).

The quantifiable characteristics within seismic data, including the dip, polarity, amplitudes, frequency, and their combinations, are assessed through the implementation of various attributes, i.e., trace-based, frequency-based, etc. The attributes are implemented over a single instant of time or by selecting a window of time. It is extracted by using a single trace or traces' set or on the interpreted horizon or surface employing a seismic dataset. The major aim of the interpreter is the assessment of enhanced reservoir properties that ultimately demarcate hydrocarbon potential, even by employing *AVO* attributes (Hussein et al., 2019).

The fundamentals of seismic attribute extraction is based on Hilbert Transformation along with analysis of complex trace that was initially introduced by (Taner and Images, 2001). According to the complex trace theory, the real component is considered to be the actual trace while the Hilbert transformed (shift in the phase upto 90°) version is taken as imaginary or conjugate trace component h (t). The complete description is elaborated in the following equation 3.9.

$$\boldsymbol{F(t) = f(t) + ih(t)} \qquad (3.9)$$

Where:

*F(t)*= Complex Trace

*f(t)* = Seismic Trace

*h(t)*= Hilbert's transform 90 phase shift of *f(t)*.

At reservoir resolution, the extracted properties act as an influential tool for evaluating the fluids within pores and recognizing reservoirs' in-situ circumstances (Zhang et al., 2019). The amplitude characteristic of the reservoir is a reliable predictor of the presence of gas in pore spaces. In addition, it provides an excellent notion of gas-water contact.

The average energy seismic attribute represents the sum of the envelopes among paraphase events (Ismail et al., 2020). Each time sample's average energy attribute is calculated by equation 3.10.

$$\boldsymbol{E(avg) = \frac{1}{N} \sum_{n=1}^{N} (Xn^2)} \qquad (3.10)$$

Where, *E(avg)* is defined by Average energy, *N* designates number of the samples, whereas *Xn* represents the values of amplitude.

The root means square (*RMS*) of the amplitude (*A*) is directly related to reservoir parameters and stratigraphic events. The attribute is computed i.e., the square root of the

average energy sum discovered within the analysis time window. It can be stated mathematically as in the equation 3.11.

$$\boldsymbol{A(RMS)} = \sqrt{\frac{\boldsymbol{1}}{\boldsymbol{n}} \sum_{\boldsymbol{i=1}}^{\boldsymbol{n}} \boldsymbol{A(i)^2}} \qquad (3.11)$$

Where n is total samples' number and A is the sample's amplitude. The above relation demonstrates property, which is tremendously profound to huge values of amplitude, as squaring of amplitude is taken during its extraction before its averaging.

The normalized seismic amplitude attribute is the cosine of the instantaneous phase and shows the lateral continuity of a bed. A vertical window is used for smoothing the computed variance and the observed amplitude is normalized.

The sweetness attribute is analyzed as the proven technique for potential sandstone detection (Hart, 2008). It has the ability to assess sweet spots in sandstone with strong reflection and distinguish sands in the surrounding of shale (Zelenika et al., 2018). Sweetness S(t) is calculated by the implementation of equation 3.12 (Radovich and Oliveros, 1998).

$$\boldsymbol{Sweetness} = \frac{\boldsymbol{Ins(A)}}{\sqrt{\boldsymbol{Ins(f)}}} \qquad (3.12)$$

### 3.1.4 Rock Physics Modeling and Seismic Inversion

Rock physics modeling (*RPM*) plays a key role in the creation of various cross-models based on the elastic properties to assess the rock behavior and check the quality of the petrophysical results. The major objective regarding *RPM* is to increase confidence in evaluating the physical properties of a hydrocarbon bearing zone, including porosity, formation lithologies and their permeability, or dynamic characteristics including the contents of pore-fluids or other pressure conditions. Therefore, these properties can also be used to construct an important rock-physical bridge that connects the different geophysical domains with one another. Petrophysics and rock physics both perform critical role in the assessment and identification of reservoir behavior and field potential. Together these techniques clearly estimate rock properties necessary to generate a model of the subsurface. Resistivity, *GR* and porosity tool measurements are basically transformed into reservoir properties by using Petrophysics whereas the use of Rock physics typically converts these petrophysical outcomes into compressional and shear velocities, acoustic impedance, and density (Yasin et al., 2021b). Rock's elastic

properties are being created and then compared with predicted logs to get accuracy and calibrate elastic property log trends when desired and to validate the correctness of the petrophysical results (Azeem et al., 2017).

The precise and consistent *RPMs* effectively segregate the producing and non-producing zones within the reservoir while handling the complications of log data due to poor hole conditions, i.e., washouts or missing data during logging, mud-filtration problems, etc. (Avseth et al., 2010). Once an accurate *RPM* is made, its successful usage is to develop the elastic properties, i.e., $V_p$, $V_s$, density, etc., of recognizing the problems present in the well logs, improve seismic to well tie, enhance reservoir characterization through seismic inversion techniques, and hence reduce the risk of doubt (Sams and Carter, 2017; Bisht et al., 2013). These *RPMs* bear the capability of reliably predicting the variabilities regarding lithologies and nature of pore-fluids (Mavko et al., 2020; Grana et al., 2012; Avseth and Odegaard, 2004), which is performed using Gassmann's equation (Gassmann, 1951).

The methodology applied in this study was to construct predictive and steady *RPM* with the employment of high-quality and consistent petrophysical properties. The suitable inputs used for *RPM* are seismic and petrophysics to integrate volumes of different fluids and mineral assemblages (lithologies). Elastic attributes ($V_p/V_s$ ratio) can discriminate different types of lithologies and plausible sand bodies by using Gassmann's relationship in the process (Reine, 2017).

The conventional *RPM* technique is used to create a model of $V_p/V_s$ ratio based on the input parameters including total porosity, volume of clay, rock-lithologies, pore-fluids properties, and pore shapes intended for different minerals (quartz and clay) with substitutions employing through Gassmann's theory (Gassmann, 1951). The appropriate tuning of *RPM* at a well can be applied throughout the field to synthesize elastic logs to replace poor quality data and any data that were not measured as long as the required petrophysical values are present. The significant reservoir properties taken into the *RPM* model comprises of volumetric of clay, various porosity types i.e., total, effective, density, etc. and total water percentages in the formation depicted in Table 3.1 (Babasafari et al., 2020).

**Table 3.1:** Input parameters used to compute conventional *RPM*.

| Physical Property | Value |
|---|---|
| Temperature (°C) | 130 |
| Pressure (*PSI*) | 3400 |
| Brine salinity (g/l) | 0.15 |
| Gas gravity | 0.689 |
| Avg. porosity (%) | 10 |
| $S_w$ (%) | 45 |
| Avg. Shale | 25% |

The change of reservoir parameters over an area of interest, rock physics template (*RPT*) is normally constructed that delineates the following effect (Babasafari et al., 2020):

a. An increase in gas saturation will reduce the acoustic impedance and the $V_p/V_s$ ratio.

b. An increase in the cement volume will increase the acoustic impedance and the $V_p/V_s$ ratio.

c. An increase in porosity percentage will decrease the acoustic impedance and the $V_p/V_s$ ratio.

d. An increase in shale will increase the acoustic impedance and the $V_p/V_s$ ratio.

e. Increase in the formation pressure will reduce the effects of impedances along with attribute values of $V_p/V_s$ ratio.

The advantage of *RPT* is to plot the basic characteristics of seismic data on the crossplot at the available well points and color code the petrophysical logs of *GR*, porosity, $S_W$, etc. to evaluate the elastic behavior in relation to petrophysical properties. Such relationships depict the lithologic properties, especially the potential sands (Figure 3.9). The standard lines of various lithologies, along with their values, i.e., x-axis, y-axis, and z-axis, distinguished the lithologies and greatly affected by various parameters. Similar to the plotting at well location, the complete seismic cube is plotted and

distribution of identified lithologies throughout the area is demarcated. The cut-off values of elastic and petrophysical attributes at the crossplot aim to illuminate the reservoir in the zones of interest that is separable from the non-reservoir in the petro-elastic domain. Anomalies concerned with the hydrocarbon's manifestation can be readily assessed and recognized over the entire seismic data. The first to introduce the *RPT* in relation to *RPM* was (Avseth and Odegaard, 2004).

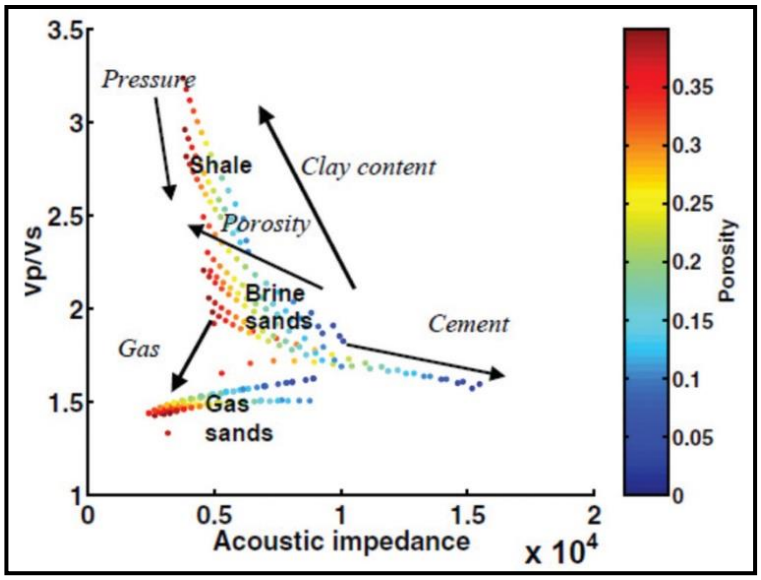



**Figure 3.9** *RPT* to plot the elastic properties of identified lithologies at a given well location (Avseth and Odegaard, 2004).

The basic purpose of this research is to acknowledge the rock *RPM* application for the missing shear sonic approximation and improvement in the poor zones of the density log, ensuring the accuracy of the petro-elastic relationship. It also bears the benefit of evaluating the borehole enlargement impacts on the density log, such as caving, washouts, etc. The accurate *RPM* of missing or poor elastic properties segregates the hydrocarbon/oil saturated sands from non-hydrocarbon bearing lithology.

Hampson Russell software and *ML* techniques are used in combination for the modeling of elastic properties i.e., $V_p$, $V_s$, density, and their combination. *RPM* in conventional way chiefly needs components of rocks including minerals composition, pore-fluids nature, the rock frame, and their combination along with reservoirs physical conditions such as pressure, temperature, gas/oil densities, litho-static pressure, and pore-pressure etc. The *RPM* generated in the research area comprised an unconsolidated model of sandstone; the pores were saturated with saline water along with gas; the Ruess-Voigt

equation was implemented to compute the various elastic moduli of minerals; and density was calculated as their volumetric average (Shakir et al., 2021).

#### 3.1.4.1 *ML* Based *RPM*

Traditional *RPM* introduces significant errors due to the number of input requirements and many assumptions in rock physics theories, making difficulty in the reliable application of interpretation methods quantitatively. *ML* entails the research and development (*R&D*) procedures that initially learn by employing the training data set and then predict data. Sometimes the possibility of forecasting the best method becomes a challenge, while the properties of data models is heavily influenced by the characteristics of the data collection. To achieve the best results, each model must be trained numerous times with varied hyperparameter settings in order to achieve the best accuracy and predictability (Jiang et al., 2020). A *ML* approach has been adopted by (Yasin et al., 2021a) within Central Indus Basin (*CIB*) to characterize the deposited reservoirs of Cretaceous age. (Al-Anazi and Gates, 2010) incorporated the *SVM* technique into the heterogeneous sandstone reservoir for the prediction of permeability and electrofacies.

Following basic data conditioning, a detailed petrophysical analysis was performed, which assisted in characterizing the Ranikot/Khadro gas-bearing sands as having low clay volumes ($V_{CL}$) ($\leq$30%) and $S_w$ ($\leq$45%). True elastic logs are essential for establishing accurate petro-elastic relationships. So, the missing and poor elastic logs have been generated through integration of conventional *RPM* along with the advanced *ML* techniques. The *ML* outputs have been authenticated through conventional *RPM* and published work, and a comprehensive petro-elastic relationship has been successfully developed. The conclusive elastic attribute of the available wells data is plotted that describes the gas-sand facies by the exhibition of low $Z_p$ along with the low values of velocities ratio ($V_p/V_s$ ratio). Finally, based on the best *ML* algorithm, the elastic volumes ($Z_p$, $Z_s$) and the key petrophysical property (*PHIE*) cube has been generated that successfully delineates the further potential present in the thin and complex reservoir.

##### 3.1.4.1.1 *ML* Algorithms Accuracy Measures

The process of *ML* is based on the generation of algorithms allowing the computer to learn. It incorporates the processes of assembling algorithms which utilize the computer's intelligence. Learning is a process of detecting statistical regularities or other

patterns in data rather than a conscious activity. Conclusively, numerous *ML* techniques will only tangentially similar to the human attitude toward learning a task. Learning algorithms, on the other hand, might give you an idea of how tough it is to learn in different settings. The expected output of *ML* algorithms is used to classify them into taxonomies (Ayodele, 2010). Despite having several algorithms to choose from which align and provide the desired results as anticipated, only one algorithm was finalized to ensure the highest correlation and accuracy is maintained while minimizing the uncertainties in later processes. Therefore, the following algorithms were selected and evaluated:

a. Decision Tree
b. Regression with Gradient Boosting
c. *SVM*
d. *RF*

#### 3.1.4.2 *AVO* Modeling and Sand Class Identification

*AVO* modeling is implemented as an important tool for lithology and fluid characterization with in reservoir zone (Rizwan et al., 2018). It is most commonly referred as *DHI* and employed in the assessment of reservoir to reduce the risk associated with drilling the prospective zones (Hussain et al., 2021; Xu et al., 2019). The ideology of *AVO* modeling is based on the evaluation of variations in the reflection amplitude as the waves travel from source to receiver covering the offset distances (Young and LoPiccolo, 2003) .The amplitude variation in relation to offset occurs because of the changing reflection coefficients (*RC*) as the seismic wave's incident angle varies at the reflector. The Zoeppritz equation (Zeoppritz, 1919) is based on the approximation of *RC* along with its linear approximation i.e., the Aki-Richards relationship that is used widely (Aki, 1980) .The equations employed incident angles therefore the offsets of the field are converted prior to the *RC* estimation. The *AVO* classification of sand is done on the intercept-gradient crossplot, which delineates the rock behavior among its fluid content on the basis of *RC* that portrays variations with changing offset or angles (Durrani et al., 2022).

### 3.1.4.3 Seismic Inverted Properties

Seismic inversion is a method that utilizes well log data in conjunction with seismic reflection data to approximate subsurface acoustic and elastic characteristics. These properties are crucial in the interpretation of seismic sections (Krebs et al., 2009). The hydrocarbon industry utilizes seismic inversion methods to identify subsurface depths that may contain hydrocarbon-rich layers. This technique serves as a valuable tool in locating hydrocarbon-bearing strata (Morozov and Ma, 2009). When performing inversions, a geoscientist is interested in physical parameters such as impedances ($Z_p$, $Z_s$), *P-wave* ($V_p$), *S-wave* ($V_s$), as well as density ($\rho$), therefore *RPM* is normally carried for consistent set of elastic properties in wells to compensate any missing or poor property along with the comprehensive understanding of reservoir. In addition, Lame parameters are also considered as they are more responsive to fluid content in the rock's pores (Clochard et al., 2009). Resulting parameters such as lithology can be obtained from inverted impedance, while other petrophysical properties such as pores, ratio between sand/shale, and gas saturation can be estimated by utilizing inverted volumes (Goodway, 2001). Obtaining a trustworthy assessment regarding reservoir's characteristics is a crucial component on which the decisions are taken, particularly throughout the developmental stages (Pendrel, 2006).

Seismic inversion practices are widely categorized between two types: Post and Pre-Stack. Post-Stack Inversion is the more commonly used approach, which involves removing the wavelet's impact from the seismic data as well as generating enhanced images of subsurface lithological properties (Chen and Sidney, 1997) .Additional method of Pre-Stack Inversion depends mainly on constructing models by employing well logs, attributes of seismic data, and incorporating important geological information (Downton, 2005).

The Post-Stack Inversion has many advantages. Firstly, strata delineation is simpler on inverted impedance in comparison to seismic dataset. Secondly, reducing wavelet's effect, side-lobes, and tuning improves the subsurface layer's resolution. Thirdly, the acoustic impedance is approximated directly from seismic data and is compared to the corresponding property of well i.e., impedance, which aids its connection to reservoir's features. Fourthly, geostatistical methods can be utilized to relate porosity to acoustic impedance, and these impedance volumes can be transformed into porosity volumes. Fifthly, acoustic impedance can be utilized for the identification of discrete

reservoir's region. Sixthly, it requires minimum time-frame than the process of Pre-Stack Inversion. The major limitation of the Post-Stack seismic inversion is the lack of shear wave computation; therefore, it shows the limitations to distinguish fluid effects more appropriately (Maurya et al., 2020; Morozov and Ma, 2009; Russell and Hampson, 1999).

Similarly, the elastic attributes that employ the $V_s$ are profound to the pore's fluid and are assessed from Pre-Stack Inverted properties (Moncayo et al., 2012). Pre-Stack Inverted elastic characteristics involve transforming seismic reflection data to impedances (p, and s), density (ρ), and $V_p/V_s$ ratio by the incorporation of well logs along with the interpreted horizon on the refection amplitudes of seismic. The two key properties including $Z_p$ and $V_p/V_s$ ratio are regarded as trustworthy for predicting reservoir properties away from well locations, depending on the target depth and acquisition configuration (Carrazzone et al., 1996). The inverted attributes obtained from Pre-Stack technique offers numerous assistances, i.e., (i) providing impedances including both "*p*" and "*s*" along with density ($\rho$) of the layers simultaneously, overcoming the interface property of seismic data, (ii) enhancing the resolution of subsurface layers by reducing wavelet effects, tuning, and side lobes, (iii) allowing direct comparison of approximated acoustic impedance to similar well-log quantities, therefore creating a bridge between well and seismic and hence the reservoir properties, and (iv) evaluating additional information based on petro-elastic relationship to discriminate various lithologies and fluids effects compared to Post-Stack Inversion techniques (Gholami, 2016).

#### 3.1.4.4 Pre-Stack Inversion

In simultaneous inversion, the vital elastic characteristics of rocks including compressional and shear impedances ($Z_p$, $Z_s$), $V_p/V_s$ ratio, as well as density (ρ) are approximated. Considering the saturated clastic rocks, the shear and compressional impedances studied together provide efficient identification of reservoir lithological and fluid contends differentiation. The Castagna equation (equation 3.13) is based on the development of a linear relationship among the velocities i.e., *P* and *S*-wave regarding the background trend (Castagna et al., 1985). Similar case is developed in the Gardner's equation (Gardner et al., 1974) that depicted a linear association between velocity and density (background trend) (equation 3.14). Thereby, during the process of simultaneous inversion these linear relationships are employed to build a connection among the variables (equations 3.15 and 3.16).

$$\boldsymbol{V_p(m/s) = 1.16V_s + 1360} \quad (3.13)$$

$$\boldsymbol{\rho(g/cc) - 1.74Vp0.25} \quad (3.14)$$

$$\boldsymbol{ln\ Zs = kln\ Z_p + kc + \Delta L_s} \quad (3.15)$$

$$\boldsymbol{ln\ \rho = mln\ Z_p + mc + \Delta L_D} \quad (3.16)$$

Where coefficients (*k, kc, m and mc*) are approximated by the help of well log properties available in the drilled wells. In the above equations, $\Delta L_S$ and $\Delta L_D$ delineate the deviance from background trend due to presence of hydrocarbons.

From the combination of all the above equations and employing Fatti's version (Fatti et al., 1994) of the Aki-Richards equations (Aki, 1980) the equation for simultaneous inversion is derived. The equation estimates the refection amplitudes that are based on the incident angles. The relationships between the $Z_p$, $Z_s$ and density according to Fatti's equation is presented as follows for the simultaneous inversion:

$$\boldsymbol{T(\theta) = \hat{c}_1 W(\theta)\, DL_P + \hat{c}_2 W(\theta)\, DL_S + W(\theta) c_3\, DL_D} \quad (3.17)$$

Where,

$$\hat{c}_1 = \left(\frac{1}{2}\right) c_1 + (1/2)\, k_{c2} + m_{c3} \quad (3.18)$$

$$\hat{c}_2 = (1/2)\, c_1 \quad (3.19)$$

$$c_1 = 1 + tan^2\theta \quad (3.20)$$

$$c_2 = -8\gamma\, tan^2\theta \quad (3.21)$$

$$c_3 = -0.5 tan^2\theta\ 2\gamma sin2\theta \quad (3.22)$$

$$\gamma = V_p/V_s \quad (3.23)$$

W($\theta$) is the wavelet at angle $\theta$, D is the differentiation derivative operator, and $L_p = l_n (Z_p)$, $L_s = l_n (Z_p)$ and $L_D = l_n (D_n)$. Equation (3.17) represents the linear equation system that solves through the matrix inversion technique to give the vector $[L_P\ \ \Delta L_S\ \ \Delta L_D]^T = [ln\ (Z_{P_n})\ \ 0\ \ 0]^T$, in which $Z_p0$ (initial-impedance) is iterated to get a solution by the method of conjugate gradient (Hampson et al., 2005). The components of equation 3.17 are given in equations 3.18-3.23 respectively. Note that if the angle is zero then this equation reduces to zero-offset (Post-Stack) Inversion. In equation, we invert for $L_p$, $L_s$ and $L_D$. In practice, simultaneous inversion involves the following steps:

1. From common midpoint gathers common midpoint gathers (*CMP)* gathers, we have the following information:
   i. A set of *N* angle traces.
   ii. A set of *N* wavelets for each angle set

iii. Initial model values for $Z_p$

2. Calculate the coefficients values for k and m using well-log data.
3. Start with initial model guess i.e., $\left[L_p\ \Delta L_S\ \Delta L_D\right]^{\mathrm{T}} = \left[ln\left(Z_{P_n}\right)\ 0\ 0\right]^{\mathrm{T}}$.
4. Apply the inversion.
5. Calculate the final values of $Z_p$, $Z_s$ and density ($\rho$), equations 3.24-3.26:

$$\boldsymbol{Z_p = exp\ (L_p)} \qquad (3.24)$$

$$\boldsymbol{Z_S = exp\ (kL_p + kC + \Delta L_S)} \ (3.25)$$

$$\boldsymbol{\rho = exp\ (mL_p + mC + \Delta L_D)} \ (3.26)$$

Note that the initial guess model representing the initial model of $Z_p$, while $\Delta L_S$ and $\Delta L_D$ are initialized with zero values in this iteration.

### 3.1.4.5 Low Frequency Model

The computation of low frequency model (*LFM*) is the next step of the inversion. It was created by incorporating low frequencies from well data interpolated within stratigraphic framework following marked horizons (Ali et al., 2018; Russell and Hampson, 1991; Russell, 1991). Thus, the low frequency component (0-15 Hz) is incorporated to obtain the absolute acoustic impedance in inversion algorithm (Russell and Hampson, 1991) . The model based seismic inversion technique uses an initial model (LFM) on the basis of impedances information (using combination of sonic *DT*, sonic DTS, and density logs), structural information (horizon data) and the interpolation of properties between wells following the structural and stratigraphic patterns. The initial model contributes to the structural delineation and overall compaction trend without incorporating the seismic data (Ray and Chopra, 2016). The initial model is established by using all wells of Zamzama Gas Field along with the developed stratigraphic model using low frequencies upto 15 Hz.

### 3.1.4.6 Pre-Stack Simultaneous Seismic Inversion (*PSSI*)

*PSSI* delineated an effective benefit over the process that involves two-steps, such as Post-Stack model based inversion (*PMBSI*), and produces a wide range of information that can be reliably considered to differentiate between producing and non-producing reservoir zones. The process of *PSSI* depends on the angle and generates inverted elastic properties including $Z_p$, $Z_s$, density ($\rho$), and ratio of the p-wave to the *S-wave* ($V_p/V_s$ ratio).

It is a proven technique for assessing reservoir's complex scenarios that are hindered by the application of Post-Stack (Sams and Carter, 2017; Hampson et al., 2005). Generally, the *PSSI* technique handles the effects among the variables that are noise sensitive and creates solutions that are non-unique. The implication of Fatti's equation in the research area for *PSSI* provided beneficial results as it has the ability for efficiently handling the attenuation effects at larger angles by extracting and utilizing the wavelets at that angles (Hampson et al., 2005 ; Fatti et al., 1994 ; Russell and Hampson, 1991).

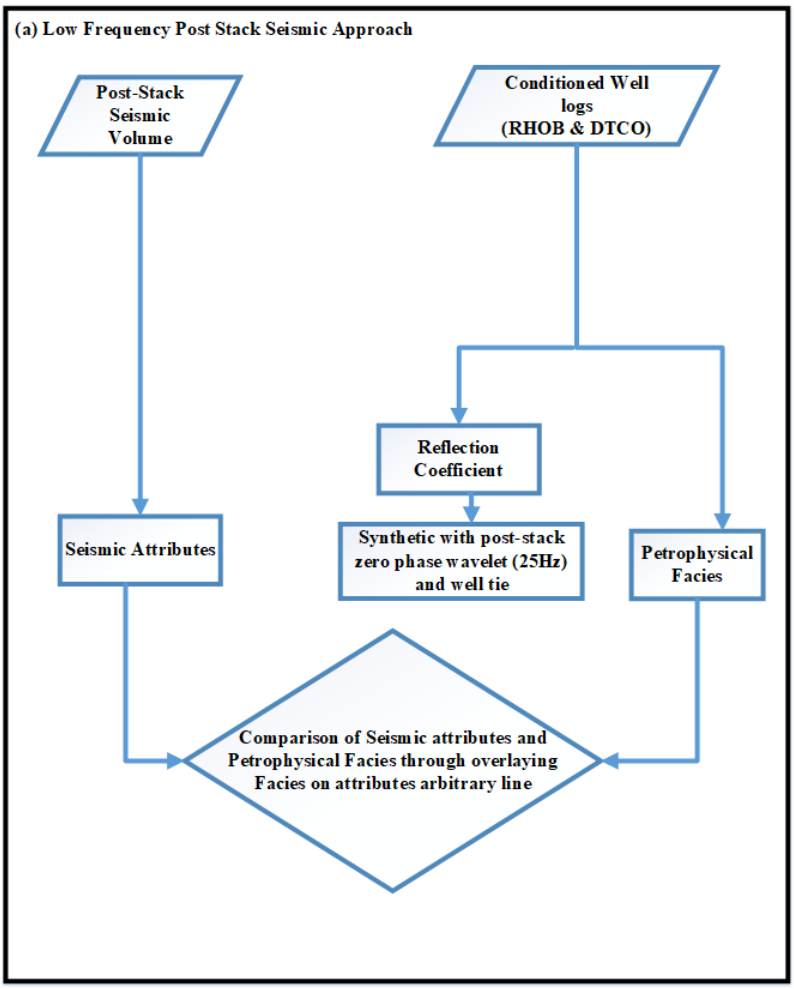
For the background ratio $V_s/V_p$ ratio, expressed as c, is selected as 0.5 depicting the wet sands of the reservoir along with the shales, and it provided effected outcomes in the inversion process. Employing the method of pre-whitening was preferred as it provided individual values considering each parameter. For each elastic property, the adjusted values are $Z_p$=1, $Z_s$=0.01, and density (ρ)=0.01 and the process was run with considering 25 iterations. Various quality control (*QC*) steps have been taken for the accuracy and validity of the inverted elastic volumes, including error analysis, estimated by the difference of actual and measured properties in the presence of background trend at wells while displaying the respective well properties on the inverted model for qualitative *QC* on the inverted sections.

#### 3.1.4.7 Injecting High Frequencies using *DL* and *CWT*

Another artificial intelligence method based on *DL* integrated with *CWT* is employed for attaining enhanced seismic resolution with improved frequency ranges. In this integration of *DL* with *CWT* incorporated seven wells that is available in a *3D* seismic cube, containing suitable information including well logs, formation tops, reservoir parameters, etc. Two wells Zamzama 05 & 07 has kept blind in this process while rest of the wells such as Zamzama 02, 03, 04, 06 & 08-ST-02 have been used in training. As the wells are penetrated through the Ranikot/Khadro Formation with two wells i.e., Zamzama-03 & 05 found commercial gas pay 6.5 m and 6 m respectively, establishing reservoir potential. The thin sands are evaluated for qualitative and quantitative interpretation such as seismic attributes and inverted properties. The seismic attributes extracted from conventional seismic data along with inverted attributes revealed limitations in the thin gas sands facies delineation. The *ML* approach applied in the above sections comprehensively modeled the elastic logs, the inverted attributes and petrophysical properties (*PHIE* and gas sands). The seismic frequency is improved by the

employment of *DL* techniques in combination with the seismic real, imaginary, and magnitude attributes assessed through *CWT*. The synthetic wedge model, delineating tunning thickness of about 60 m, utilizing the wavelet extracted from seismic data proves the seismic resolution limitation for thin Ranikot/Khadro Formation sands (~7 m) reservoir characterization. An advanced technique of *DNNs* is combined with *CWT*'s real, imaginary, and magnitude components that built a relationship between the well's high-frequency *RC* and creates high-frequency synthetic seismic.

The high-frequency synthetics are matched with seismic traces at well locations and populated throughout the cube by having a relationship through *DNNs* with seismic traces. The high-resolution synthetics become capable of resolving the tuned thin Ranikot/Khadro Formation sands which are confirmed by creating a new synthetic wedge modeling. Seismic attributes extracted after *CWT* & Deep Neural Network (*DNN*) approach has reflected an improved image of thin sands after injecting high frequencies into seismic dataset. The detailed workflow for attaining high-frequency volumes using *CWT-DNNs* relation is depicted in Figure 3.10a & b.



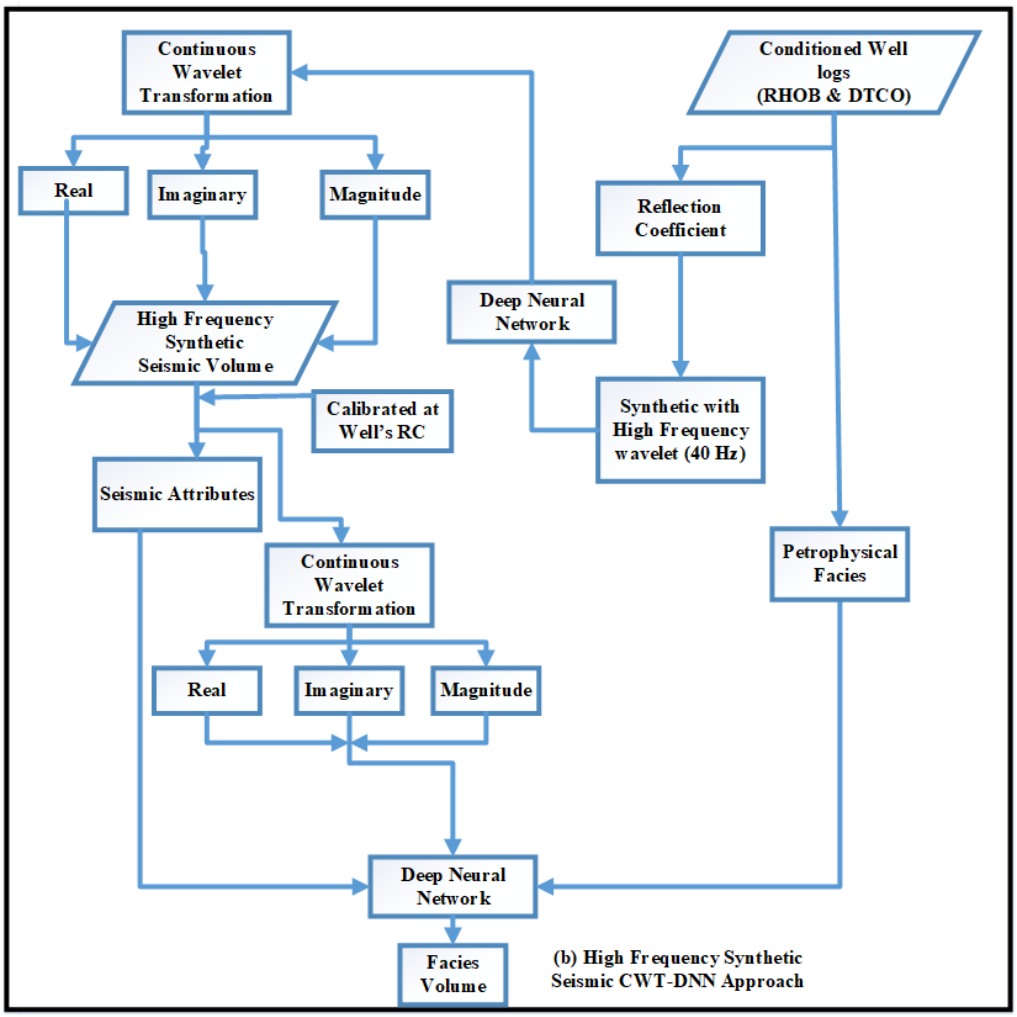



**Figure 3.10** Workflow employed in current research a) to produce results from field seismic. b) high-frequency synthetic seismic volume (*HFSSV*) from *CWT* and *DNNs* to produce high frequency attributes & facies through validation at well locations.

Figure 3.11 demonstrates a computational neural network that is employed in the *DL*. The intermediate network layer also referred to as a "hidden layer," contains neurons that receive values from the input layer and transmit them to those intermediate neurons. The output layer receives the sums of one or maybe more hidden layers and presents the final output to users. In order to harmonize terminology with *NNET*, neurons' outputs are referred to as activations, and the synapses are often indicated by weights.

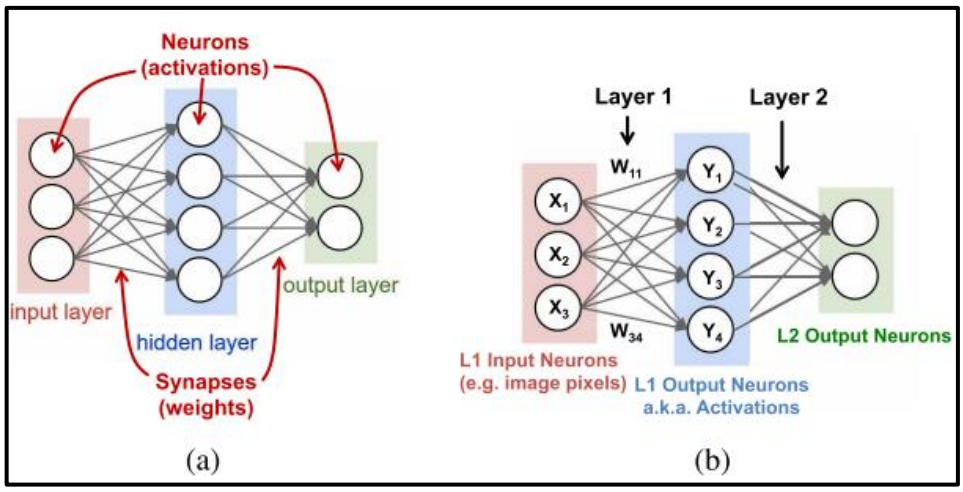



**Figure 3.11** Computational Neural Network (a) Neurons and Synapses (b) Compute weighted sum for each layer.

### 3.1.5 A Novel approach to *PP* Prediction based on Conventional Well logs using Advanced *ML* Approach

*PP* is crucial for evaluating geo-mechanical properties of any reservoir. Similarly, ensuring safe and efficient drilling operations and optimizing reservoir characterization and production are equally imperative.

This study highlights the effects of *PP*, vertical stress and horizontal maximum and minimum stress that are identified through empirical equations and calibrated with measured pressure data at well locations, these geo-mechanical properties are then identified on other wells through advanced *ML* algorithms such as *ETR*, *RFR*, *GBR* and *DTR*. In second stage, for the spatial variations of *PP* in the thin Ranikot/Khadro Formation Sand Reservoir across the entire subsurface area, a *3D PP* prediction is conducted using *CWT*. The relationship between the *CWT* and geo-mechanical properties is then established through supervised *ML* models on well locations to predict the uncertainties in *PP*.

To assess the effectiveness of *ML* algorithms each of seven wells' datasets is trained individually, with one well used as a blind for validation while the others serve for training. The *ML* model divided the augmented logs into two groups: 70% training data and a 30% testing dataset. Among all intelligent regression techniques developed using petrophysical properties and elastic properties for *PP* prediction, *GBR* has provided exceptional results which have been validated by evaluation metrics based on $R^2$ score i.e., 0.41 between the calibrated and predicted *PP*. Via *DNN*, the relationship is

established between *CWT* resultant traces and predicted *PP* to analyze the spatial variation. As a result, *PP* or other geo-mechanical properties are stimulated through the *3D* seismic data accurately. The algorithms have also contributed in predicting the uncertainty of *PP* between the wells spatially which can also be affirmed through comparison between the predicted and measured *PP*. The comprehensive approach presented here provides strong and reliable for the evaluation of modern formations, effectively minimizing uncertainties. It also enables efficient multi-well interpretation in complex reservoirs, paving the way within the realm of *ML*.

A key variable in several drilling and exploration procedures, such as well design, well stability analysis, and mud program design, is *PP* (Yu et al., 2020; Bahmaei and Hosseini 2020; Zhang et al., 2020; Hu et al., 2013). When drilling, imprecise estimation of *PP* can lead to critical consequences, including unwarranted drilling mud seepage into the formation being drilled, unexpected influxes of formation fluid into the well, and the eventual occurrence of well blowouts, posing significant risks to both the drilling rig and the well crew (Zhang et al., 2020 ; Mahetaji et al.,2020 ; Maddahi et al., 2020) . *PP* can be considered a safe pressure when the hydrostatic pressure of the drilling fluid within the wellbore falls below the formation pressure and fracture pressure (*FP*) of the formation (Richards et al., 2020; Darvishpour et al., 2019).

#### 3.1.5.1 *PP* Importance

Numerous researchers have used a variety of techniques recently to precisely forecast and estimate the *PP*. (Terzaghi et al., 1996) marked a significant milestone by initiating the earliest research into *PP* prediction. As a pivotal part of this research, an empirical equation was formulated to compute *PP*. An experiment was devised to explore the influence of rock compaction on overburden pressure (*OBP*), with the aim of comprehending its effects. The proposed theory posited that all stress effects, encompassing alterations in elastic wave velocity, were classified as effective stresses. It further emphasized that the *OBP* was offset by the *PP* generated by vertical stresses.

In 1975, Eaton introduced a crucial advancement by presenting two empirical models geared towards *PP* prediction. These models hinged on data from compressional pressure wave, shear wave, and *RES* (Farsi et al., 2021; Shen et al., 2017; Yoshida et al., 1996).

As per the review of previous literature, it came to light that these equations only accurately forecast the key parameter for the field whose data were utilized to build the empirical equations. (Abad et al., 2021 ; Naveshki et al., 2021). In order to address the problem, plenty of research has been recently conducted to create predictive models for predicting various factors in the energy sector using *AI* approaches (Hazbeh et al., 2021). Abdelaal et al. (2021) formulated three distinct models, encompassing *SVM*, functional networks, and *RF*, for the purpose of projecting *PP* during the drilling process. This was achieved by leveraging a dataset containing 3100 drilling records. The input variables employed in the proposed model encompassed the rate of penetration (*ROP*), mud flow rate (*Q*), standpipe pressure (*SPP*), and rotation speed (*RS*). Upon comparing the predictive outcomes of the diverse models suggested, it was discerned that the *RF* algorithm exhibited the highest degree of prediction accuracy across all four models (with an $R^2$ value of 0.98 and AAPE of 2%) (Abdelaal et al., 2021).

The focus of this study is on pre-drill planning and improved geo-mechanical parameters estimation for monitoring the reservoir (Ranikot/Khadro Formation) and field development in the study region. Estimating changes in the stresses of subsurface geological strata, OBP, *PP*, and FP for subsequent drillings, as well as analyzing borehole stability. The information that is available to calculate the reservoir geo-mechanical parameters consists of *CWT* traces of *3D* seismic data and wireline logs containing the Caliper (*CALI*), *GR*, *RHOB*, $V_p$, $V_s$ of seven wells drilled in the Ranikot/Khadro Formation Reservoir Zone of Zamzama. Initially the *CWT* is implemented on *3D* seismic data. At different in-lines and crosslines, the information of *CWT* traces consists of real, imaginary and their corresponding magnitudes are extracted in time domain. On the second hand, wireline logs play a role in computing *PP* using empirical equations such as the Eaton method, which relies on sonic transit time measurements.

In this research, five intelligent regression techniques for *PP* prediction using petrophysical properties and elastic properties are developed. The five advanced *ML* algorithms are *SVM*, multi-linear regressions, *GBR*, *RFR* and *DTR*. The models are created using a set of nine input parameters that have the greatest impact on the *PP*, followed by a feature importance analysis. To assess the efficacy of *ML* models, each dataset from the seven wells is individually used for training. During this process, one well is kept as a blind dataset for validation, while the remaining wells are utilized for training the models, which are further validated through evaluation metrics consisting of root mean squares, $R^2$ score and absolute mean squares. In order to select the predictive

model that offers the highest precision in terms of predictions, a comprehensive assessment is undertaken. This evaluation involves a thorough comparison of the statistical metrics employed to gauge the models' performance throughout both the training and testing phases. For spatial variations of *PP* through the seismic volume, the relationship is established between *CWT* resultant traces and predicted *PP* through *DNN* and the *PP* or other geo-mechanical properties are stimulated through the *3D* seismic data accurately.

#### 3.1.5.2 Dataset and Methodology

Following are the steps starting from data set to the integration of *DL* with *CWT* for assessing *PP* and geomechanical properties of Ranikot/Khadro Formation.

##### 3.1.5.2.1 Data Acquisition

This study aims to predict *PP* in the Zamzama Field, utilizing valuable well log data from seven wells i.e., Zamzama-02, Zamzama-03, Zamzama-04, Zamzama-05, Zamzama-06, Zamzama-07, and Zamzama-08. This analysis focuses on the Ranikot/Khadro Formation thin bed of the reservoir zone, where *GR*, *RHOB*, *CALI*, and $V_p$ (*DT*) well logs are evaluated. For the spatial distribution of *PP*, a *3D* Post-Stack seismic cube containing 244 in-lines ranging from 382 to 626 and 147 in-lines ranging from 113 to 260 is utilized. The well log and seismic *3D* data were provided by *DGPC*. For, *CWT*, sampling rate is taken as 2.0ms, frequency band ranges from 4Hz to 35Hz and no of scales for is 31.

##### 3.1.5.2.2 Data Processing

In the training dataset, all seven wells are utilized, with one of them kept as a blind well to evaluate the model's performance. This research applies a data split of 70-30% is utilized: 70% for training the model and the remaining 30% for prediction evaluation. The main aim of this study is to accurately predict geo-mechanical properties, using *ML* based method, compared with that of conventional techniques. These predictions are then extended to the seismic volume to quantify uncertainties in *PP*.

The quality and reliability of the input data used to train the algorithms is critical to the accuracy of the predictions, therefore we have used conditioned logs i.e., rock

physics based logs to develop a model predicting *PP*. Several supervised *ML* algorithms i.e., regression techniques such as *RF*, *GBR*, *ETR*, and *DTR* are employed in order to predict the geo-mechanical properties and *PP*. Furthermore, spatial variations of *PP* within the seismic volume are analyzed by establishing a relationship between *CWT* resultant traces and predicted *PP* via *DNN*. The *3D* seismic data effectively stimulates *PP* and other geo-mechanical properties.

#### 3.1.5.3 Supervised *ML*

In the realm of *ML*, supervised learning is a prominent task, in which function is acquired to establish the link between inputs and outputs using the pair. This function is built using labeled training data and relevant features, enabling predictions of unknown attributes (Manzoor et al., 2023).

#### 3.1.5.4 Input Data

The *ML* model development demands the data samples along with multiple features and target variable. After initial quality check, we can optimize our data which helps in improving the accuracy and effectiveness of the models. Furthermore, it aids in simplifying the model's complexity, enhancing its interpretability and suitability for real-world applications (Kuhn and Johnson, 2013).

#### 3.1.5.5 Feature Extraction

Within the realm of *ML*, the selection of training features profoundly influences a model's performance. When selecting them, it is important to consider the target feature and the correlation between the training and the target feature (Zheng and Casari, 2018; Müller and Guido, 2016).

#### 3.1.5.6 Predictions and validation

In *ML*, predictions are made by applying a trained model to a dataset and using the model's output to make a prediction for each data point in the dataset. The quality of these predictions can be evaluated using a variety of performance metrics, depends upon the data nature (Hastie et al., 2009).

The workflow for supervised *ML* algorithm is shown in Figure 3.12.

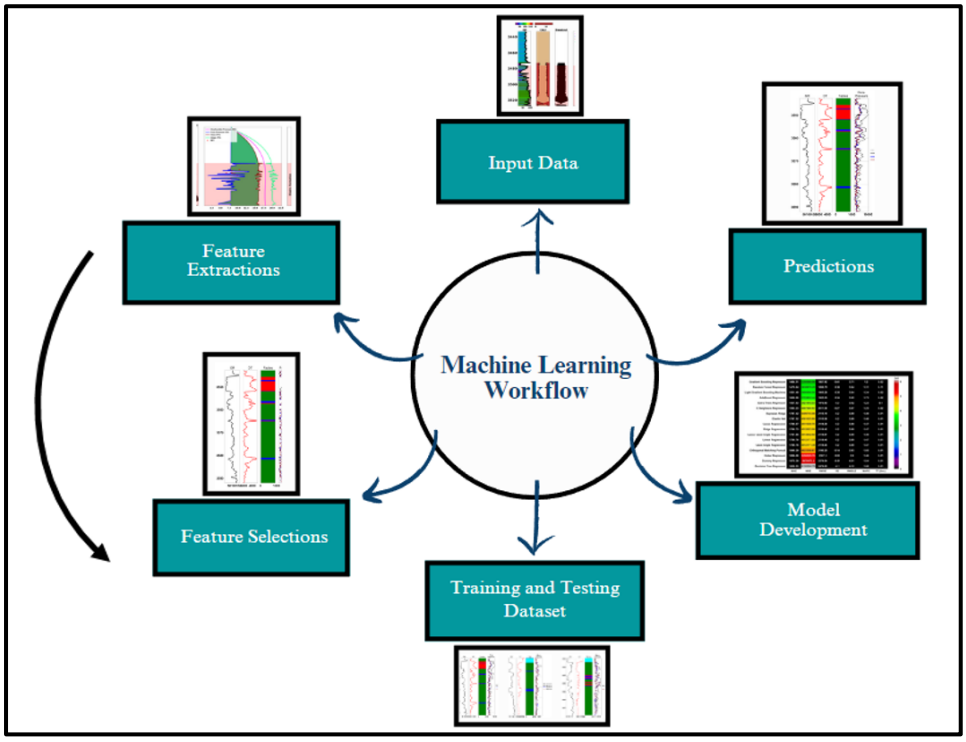



**Figure 3.12** Flowchart Illustrating the Process of Supervised *ML*. After taking an input dataset, requires features are extracted, learning algorithm. Model is developed using training data and evaluated based on testing data used as blind dataset, thus predictions are made.

### 3.1.5.7 Continuous Wavelet Transform (*CWT*)

In order to structure Post-Stack seismic data into vectors or matrices suitable for *NNET*. *CWT* is performed on the raw Post-Stack seismic data, generating seismic spectra through the application of. Prior to executing the wavelet transformation, several pertinent parameters concerning the Post-Stack seismic data are supplied to the *CWT* process.

Sampling rate of Post-Stack seismic data: If the Post-Stack seismic data sampling rate is 2.0ms, then frequency sampling rate which is the required parameter for *CWT* will be 200s.

Frequency band of Post-Stack seismic data: The amplitude spectrum provides information about the frequency bands (lower and higher limit). In this case, lower and higher frequency is 4 and 35Hz respectively.

Number of scales: Number of scales is dependent upon the quantity of data. If the rows of data for *DL* model are fewer, the number of scales should be lower in order to avoid overfitting in data. Constraining the number of scales will give less certain results.

In this case, the optimal number of scales is considered as 31scales. The *CWT* performs a convolution with the Post-Stack seismic data using a wavelet function, like a Ricker wavelet. The *CWT* spectra contain the energy of the signal both in time and in frequency and provide the hidden information inside the Post-Stack seismic data in the form of an imaginary part of seismic data and the real part of seismic data and their corresponding magnitudes. In this way, we can make full use of the frequency content of the original Post-Stack seismic data. The spectrum ranges from 8–50 Hz at a frequency within a specific time window as shown in Figure 3.13.

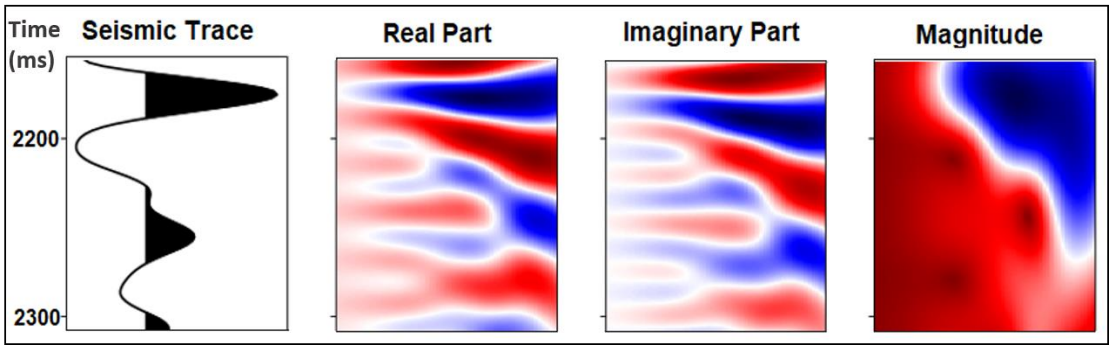



**Figure 3.13** *CWT* of the seismic traces i.e., real, imaginary, and magnitude part.

### 3.1.5.8 Conventional Approach for *PP* Prediction

*PP* is a critical parameter essential for drilling planning, as well as geo-mechanical and geological analyses. It characterizes the pressures exerted by fluids within the pore spaces of permeable formations. It spans a spectrum from hydrostatic pressure to significantly elevated levels, even reaching 48% to 95% of the overburden stress. Instances of abnormal *PP* occur when it departs from the hydrostatic pressure, taking the form of either lower or higher values. The term "overpressure" is used to denote situations where *PP* surpasses the standard levels (Zhang, 2011).

The anticipation of *PP* builds upon the foundational principles encapsulated within Biot's effective stress laws (Biot, 1941; Terzaghi et al., 1996). These principles propose that the *PP* within a formation is influenced by both the overburden stress and the effective stress. The relationship among *pp* and overburden stress ($\sigma V$) is given as follows (equation 3.27):

$$\boldsymbol{pp = \sigma V - \sigma_e/\alpha} \qquad (3.27)$$

Here, $\sigma e$ represents the vertical effective stress, and α corresponds to the Biot effective stress coefficient, traditionally considered as 1.

For the efficient computation of hydrostatic pressure ($p_n$), equation 3.28 is used (Zhang, 2011).

$$\boldsymbol{p_n = p_f g h} \qquad (3.28)$$

In this equation, *pn* represents hydrostatic pressure, *g* is the acceleration due to gravity, *ρf* stands for fluid density, and h indicates the vertical height of the fluid column.

While dealing with one-dimensional scenarios, the formulations for vertical effective stress and vertical *PP*, initially introduced by Biot in 941, can be expressed as follows (equation 3.29): (Terzaghi et al., 1996; Biot, 1941):

$$\boldsymbol{\sigma_e = \sigma V - \alpha p} \quad (3.29)$$

Here, p denotes *PP*, *σV* represents vertical stress, *σe* stands for vertical effective stress, and α signifies Biot's effective stress coefficient.

Under conditions of normal pressure, the relationship between the normal vertical effective stress (*σn*) and the normal or hydrostatic pressure (*pn*) is given by equation 3.30:

$$\boldsymbol{\sigma_n = \sigma V - \alpha p_n} \quad (3.30)$$

*σn* denotes the normal vertical effective stress, while *σV* indicates overburden or vertical stress.

Figure 3.14 depicts the relationship between hydrostatic pressure, formation *PP*, overburden stress, and vertical effective stress in relation to the true vertical depth (*TVD*) within a standard oil and gas exploration well.

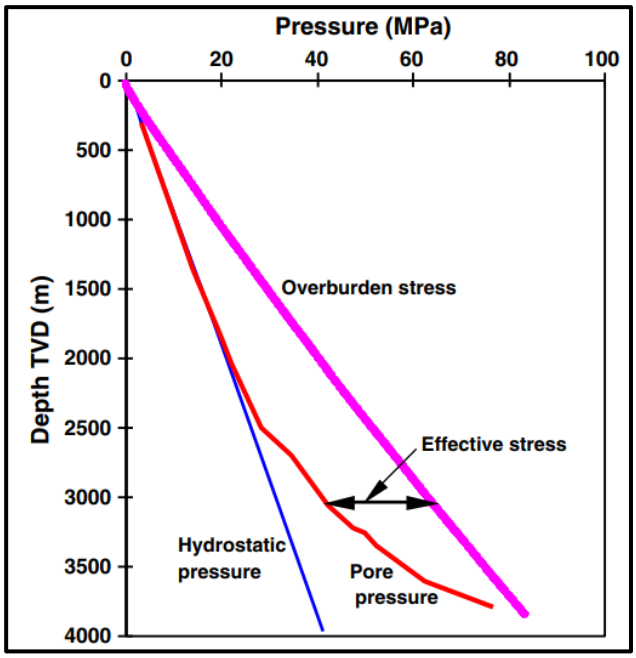



**Figure 3.14** A graphical representation of Hydrostatic pressure, *PP*, overburden stress, and effective stress in a borehole (Zhang, 2011).

### 3.1.5.9 Eaton's Method

Eaton (1975) introduced an empirical formula for forecasting *PP* gradient using sonic compressional transit time. Both of these factors necessitate a normal compaction trend line (*NCTL*) and overburden pressure (*OBP*), which can either be well-specific or derived from local or regional models. The equation 3.31 is given as:

$$\boldsymbol{P_{pg} = OBG - (OBG - P_{ng})(\frac{\Delta t_n}{\Delta t})^3} \qquad (3.31)$$

Where $P_{pg}$ signifies the *PP* gradient, *OBG* denotes the overburden pressure gradient, and $P_{ng}$ represents the hydrostatic *PP* gradient. $\Delta t_n$ denotes the sonic transit time derived from the normal compaction trend at the investigated depth, and $\Delta t$ signifies the sonic transit time acquired from well logging or seismic interval velocity. The traditional procedure for predicting *PP* via Eaton's method follows the workflow shown in Figure 3.15.

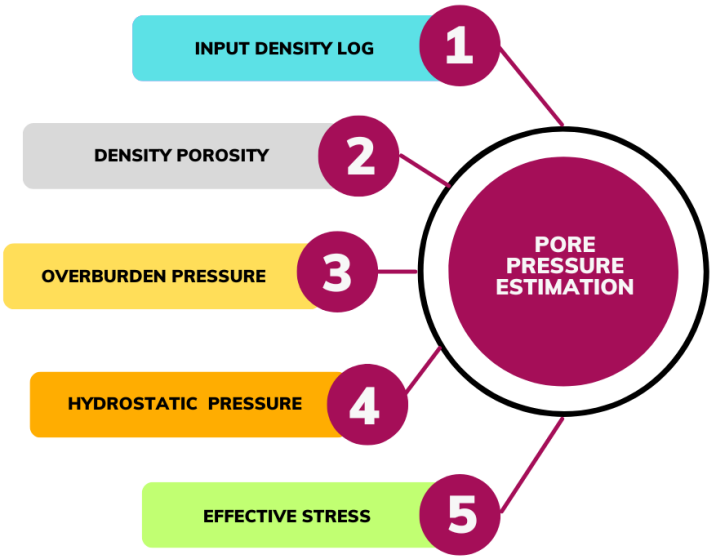



**Figure 3.15** *PP* estimation through conventional workflow. The parameters are required accordingly to generate the *PP* on wells.

## 3.2 Summary

The chapter comprises many interesting algorithms both conventional and *ML* for the assessment of below resolution varied reservoir sands. The logs processing is done for outliers and washout zone corrections using ML techniques, including One Class SVM and supervised ML. The supervised ML algorithm, i.e., *ETR*, is observed as the best one for washout zone correction after various technique evaluations through *MAE*, *MSE*, *RMSE*, *R2*, *RMSLE*, and *MAPE*. The *ML* algorithm is developed to approximate

petrophysics by keeping wells blind, and a decent match is observed with outcomes obtained through conventional techniques. A workflow is selected for detailed seismic structural interpretation and the generation of time and depth structural maps for the reservoir Ranikot/Khadro Formation. The inverted properties approximation is made via the simultaneous pre-stack seismic inversion technique using the available 3D seismic and well data. A comprehensive rock physics modeling based on *ML* techniques is devised for the estimation of elastic responses that tend to separate the facies in the petro-elastic domain. High frequency synthetic seismic volume *(HFSSV*) from *CWT* and *DNNs* is generated to produce high frequency attributes and facies that are validated at well locations. A novel approach to an advanced *ML* model is developed for *PP* predictions, integrating *CWT* and *RESNET* techniques.

# CHAPTER 4

# RESULTS AND DISCUSSION

## 4.1 Petrophysics

Petrophysics provides the depiction of physical reservoir parameters including porosity, volumetrics, fluids, etc. This research outcomes makes it possible to document the geological formations in a borehole and makes it easier to measure and identify the fluids and lithologies present in a reservoir.

### 4.1.1 Facies Analysis

Reservoir facies study is a critical step in the oil and gas development process. Conventionally the reservoir facies have been interpreted based on cut-off defined in Table 4.1.

**Table 4.1:** Petrophysical cut-off ranges employed for reservoir facies (Khan et al., 2022).

| Facies | Volume of Clay | $S_w$ |
|---|---|---|
| **Shale** | >0.35 | -Nil- |
| **Gas Sand** | <=0.35 | <=0.45 |
| **Wet Sand** | <=0.35 | >0.45 |

### 4.1.2 Outlier Removal of Zamzama Gas Field

The outlier removal is the key step for attaining consistency in the dataset. The exceptional or out-of-range values are identified in each well and then removed so that the data is cleaned for further processing.

#### 4.1.2.1 Outlier Removal of Zamzama-02

The pair-plot is drawn between the different well logs (*GR*, *RHOB*, *NPHI*, *DT*, *RES*) to visualize the presence of outliers. The red circle in Figure 4.1 shows the presence of outliers which can affect the petrophysical interpretation of the well logs. The red circled outliers can be removed through *ML* methods or traditional method.

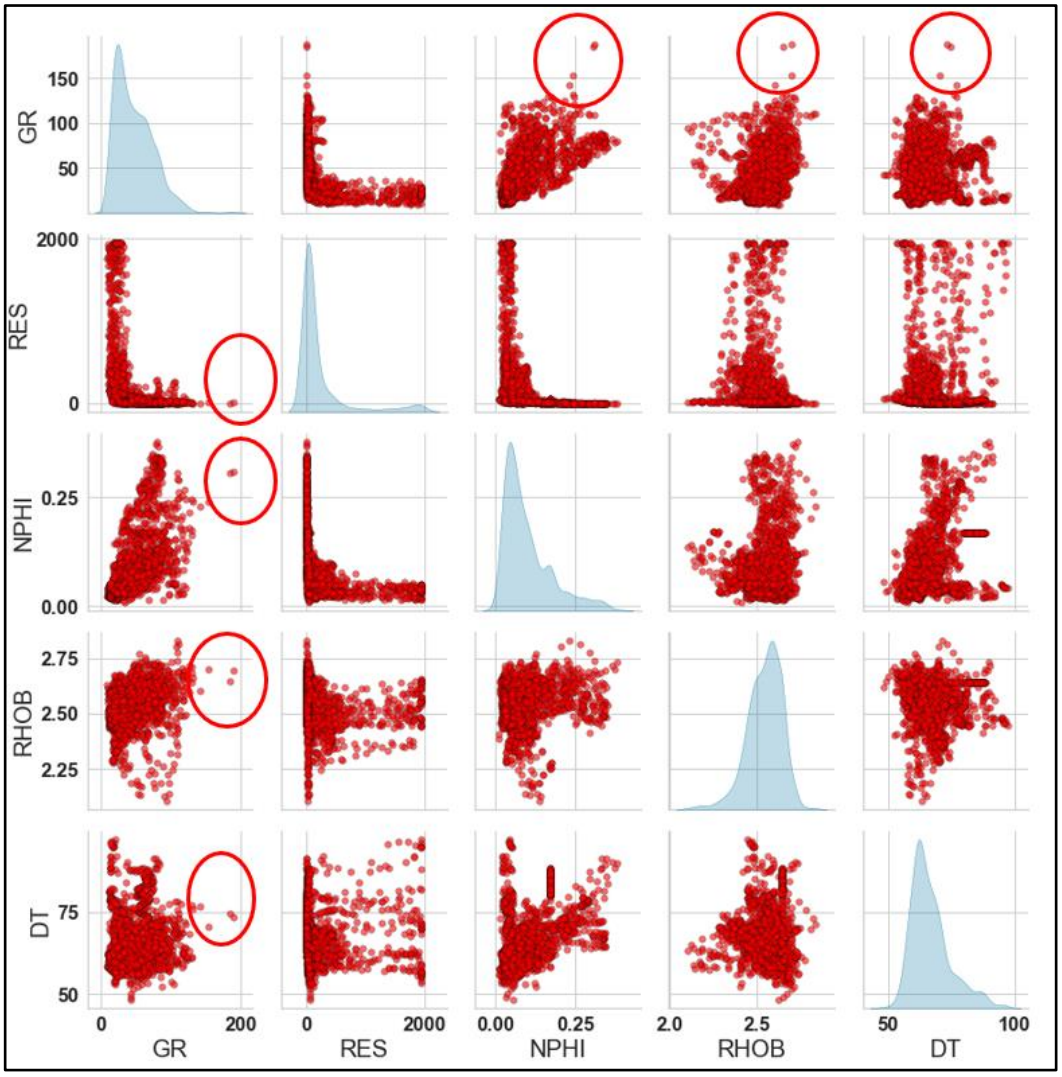



**Figure 4.1** Pair-plots between different well logs representing the appropriate outliers. Outliers are highlighted with red circles in the zone of interest.

In quest to identify and remove these outliers, algorithms such as isolation forest, Minimum Covariance, One-class-*SVM*, etc. were employed on the input logs. Furthermore, box plots were used to clearly visualize the data distribution through their quartiles as illustrated in Figure 4.2. The boxes depict the interquartile range (*IQR*) of the data and extend from the first quartile (*Q1*) to the third quartile (*Q3*). The midpoint of the box represents the median of the data. The quantity of outliers varies from algorithm to algorithm as mentioned in the Table 4.2. Basically, the Heat maps are used to visualize patterns or correlations of data points in a matrix (McDonald et al., 2021). One-Class-*SVM* outdated all other algorithms as it provided an appropriate number of outliers as delineated in the Table 4.2.

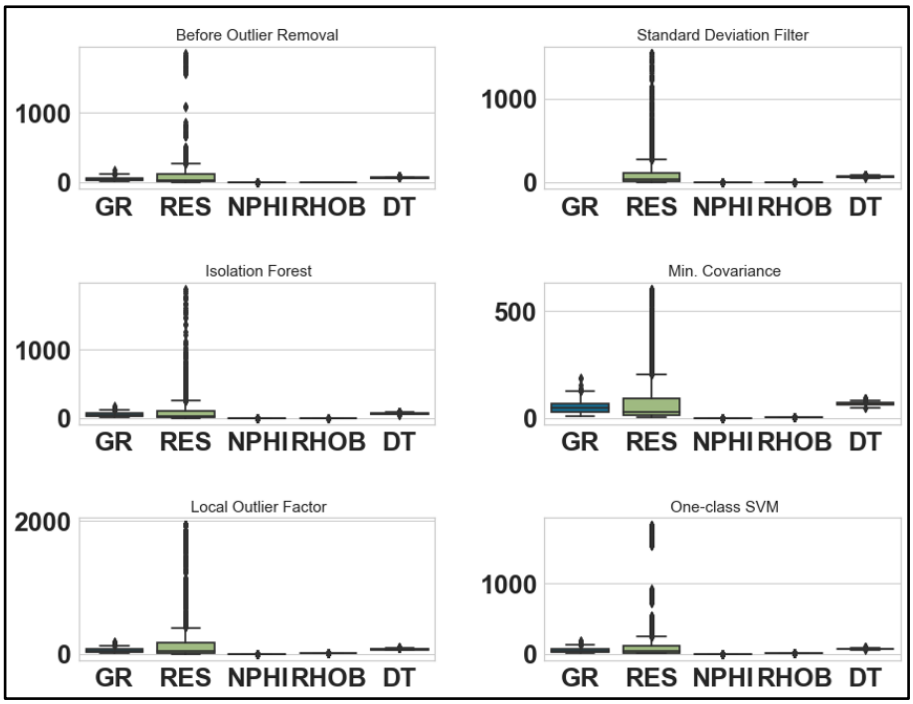



**Figure 4.2** Box Plots for *RHOB*, *NPHI*, *DT*, *RES* depicting outliers with data distribution in quartiles using several unsupervised algorithms i.e., Isolation Forest, Local outlier factor, One-class *SVM*, Minimum covariance and standard deviation filter.

The overall outliers by the application of various *ML* algorithms are summarized in the Table 4.2.

**Table 4.2:** Summary of outliers and inliers of the utilized well logs dataset identified by various employed *ML* algorithms.

| Machine Learning Algorithms | Inliers | Outliers |
|---|---|---|
| Original | 1798 | 0 |
| Standard Deviation | 1798 | 0 |
| Isolation Forest | 1618 | 180 |
| Min. Covariance | 1618 | 180 |
| Outlier Factor | 1618 | 180 |
| One-class SVM | 1618 | 180 |

After the removal of the outliers from measured logs, the pair-plots were re-evaluated, which clearly exhibited that the outliers have successfully been removed through One-Class *SVM* as given in Figure 4.3.

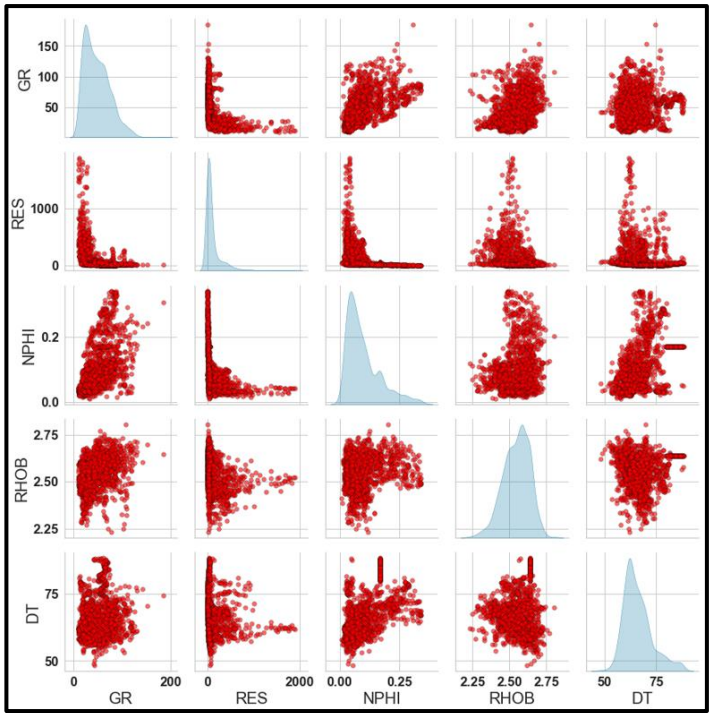



**Figure 4.3** Pair-plots between well logs to visualize the data distribution after successful outlier removal through One-Class-*SVM* approach.

The ID view of logs after the identification, removal, and optimization of logs via *SVM* technique is illustrated in Figure 4.4.

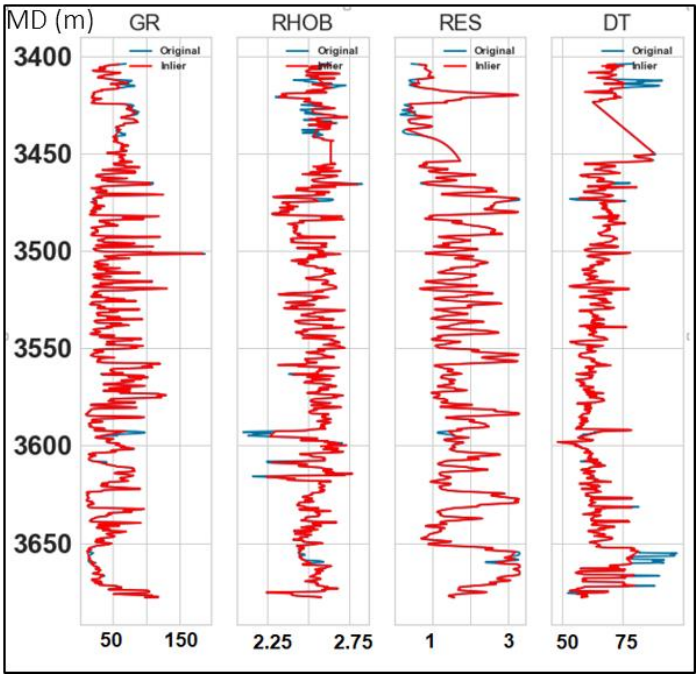



**Figure 4.4** The logs depict the removal of outliers while following the similar trends in comparison to original logs.

#### 4.1.2.2 Outlier Removal of Mehar-01

Following the similar process, the outliers of Mehar-01 is assessed and removed as illustrated in Figure 4.5. Green and blue scatter points indicate the zone of interest points of Mehar-01 respectively. The abnormality of data points is encircled that need to be removed and optimized before evaluating the petrophysical interpretation (Figure 4.5).

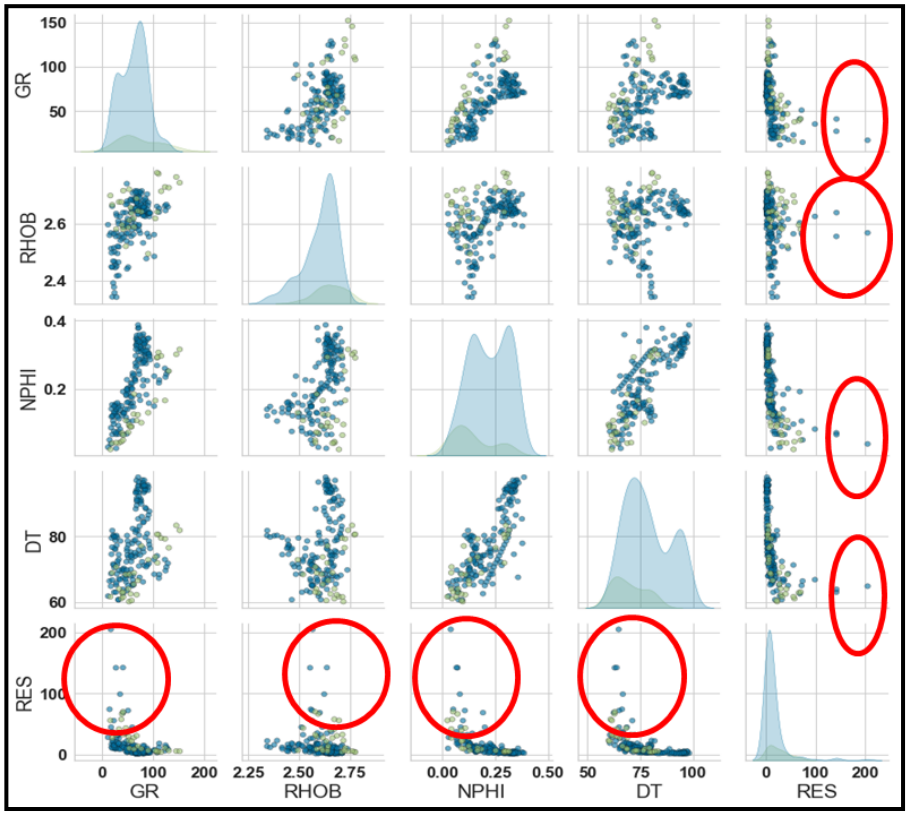
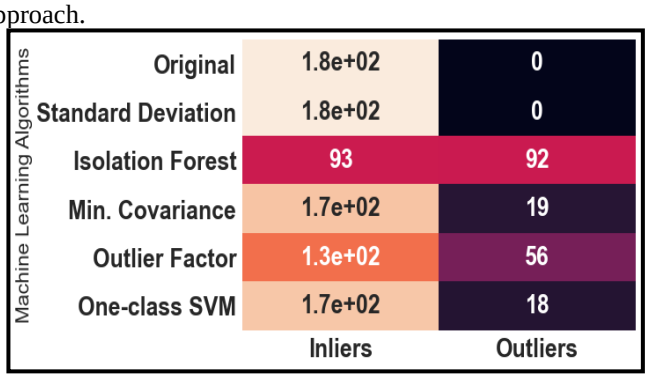



**Figure 4.5** Pair-plots between different well logs representing the appropriate outliers. Outliers are highlighted with red circles in the zone of interest.

A comparison is made at the box plot to view the data distribution through their quartiles in Figure 4.6, while a brief summary of the outliers and inliers using various *ML* algorithms are presented in Table 4.3. The outliers are removed after following the similar procedure and depicted in the pair-plot (Figure 4.7). The similar approach is implemented for all of the available wells of Zamzama along with the Mehar.

**Table 4.3:** Pair-plots between well logs to visualize the data distribution after successful outlier removal through One-Class-*SVM* approach.



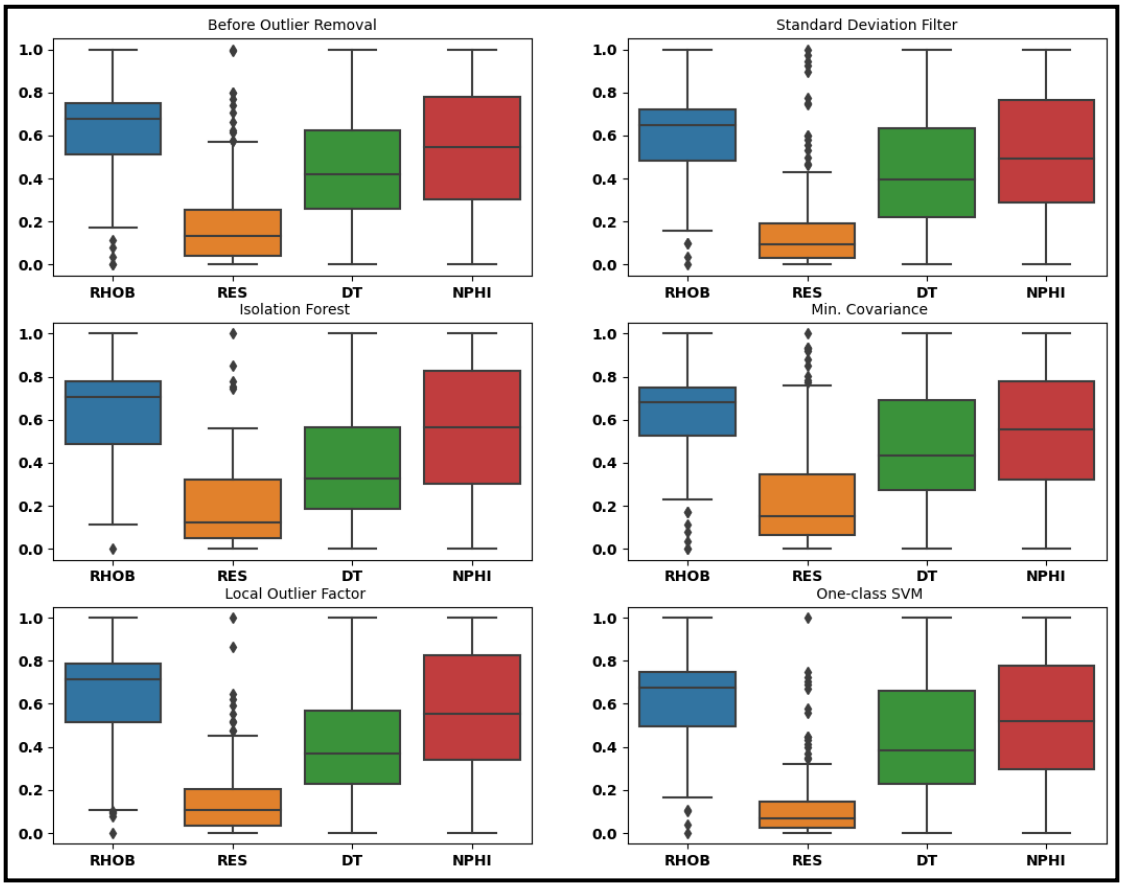




**Figure 4.6** Box Plots for *RHOB*, *NPHI*, *DT*, *RES* depicting outliers with data distribution in quartiles using several unsupervised algorithms i.e., Isolation Forest, Local outlier factor, One-class *SVM*, Minimum covariance and standard deviation filter.

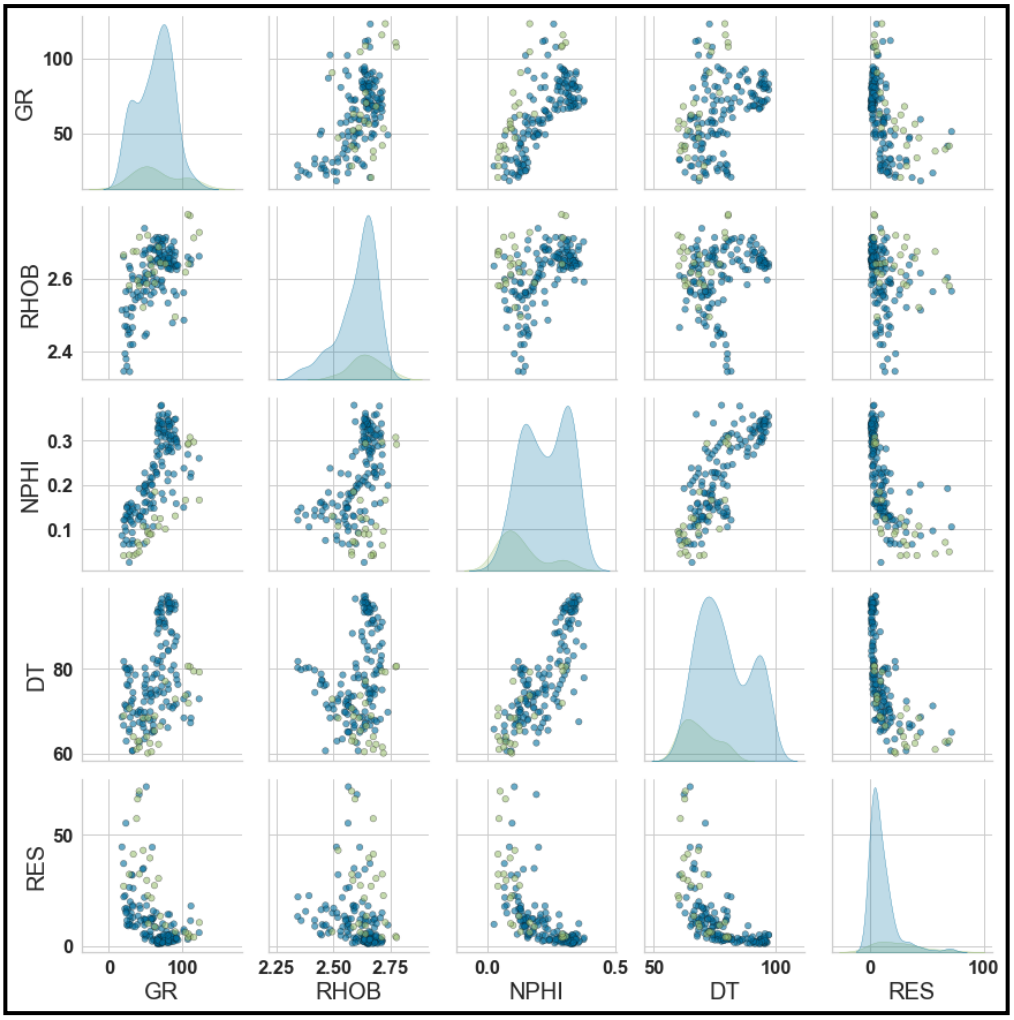



**Figure 4.7** Pair-plots between well logs to visualize the data distribution after successful outlier removal through One-Class-*SVM* approach.

### 4.1.3 Splice Zone Identification and Removal through Supervised *ML* Algorithms of Zamzama Field

The splice zones are the missing zone that are needed to be predicted via interpolation method or advanced approach (Figure 4.8). The interpolation method does not reflect the prediction of true data. Alternative of this approach, the advanced method is the supervised *ML* algorithms. After the outlier's removal, the splice zones on the well logs are removed through the visual inspections. The splice zones are predicted by using the associated non splice well logs through supervised *ML* algorithms. The density log associated with washout zones is removed and predicted by associated available well logs through supervised *ML* algorithms.

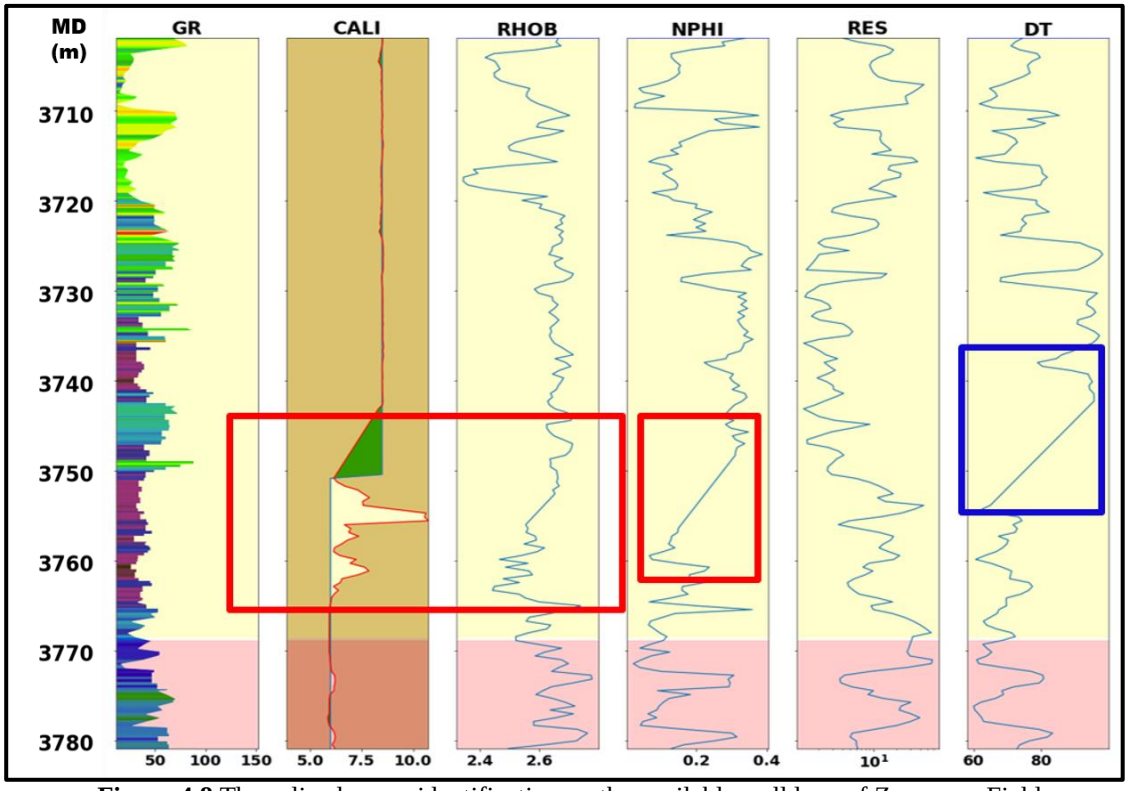



**Figure 4.8** The spliced zones identification on the available well logs of Zamzama Field.

The best *ML* algorithm is identified through evaluation metrics applied on the different sample datasets extracted from training datasets.

#### 4.1.3.1 Splice Zones approximation of Zamzama-02

The splice zone affected raw well logs have been corrected via supervised *ML* algorithms. All the other *ML* algorithms have been outdated by the *RFR* and *ETR*. The first track shows the zone of Ranikot/Khadro Formation and Pab Formation, second track shows the *GR*, third track consists of corrected density and measured density, fourth track is the corrected *NPHI* and measured *NPHI*, the fifth track is the crossover between corrected *NPHI* and corrected density, sixth track is the deep resistivity and last track is the corrected p-sonic (*DT*) and measured sonic (Figure 4.9).

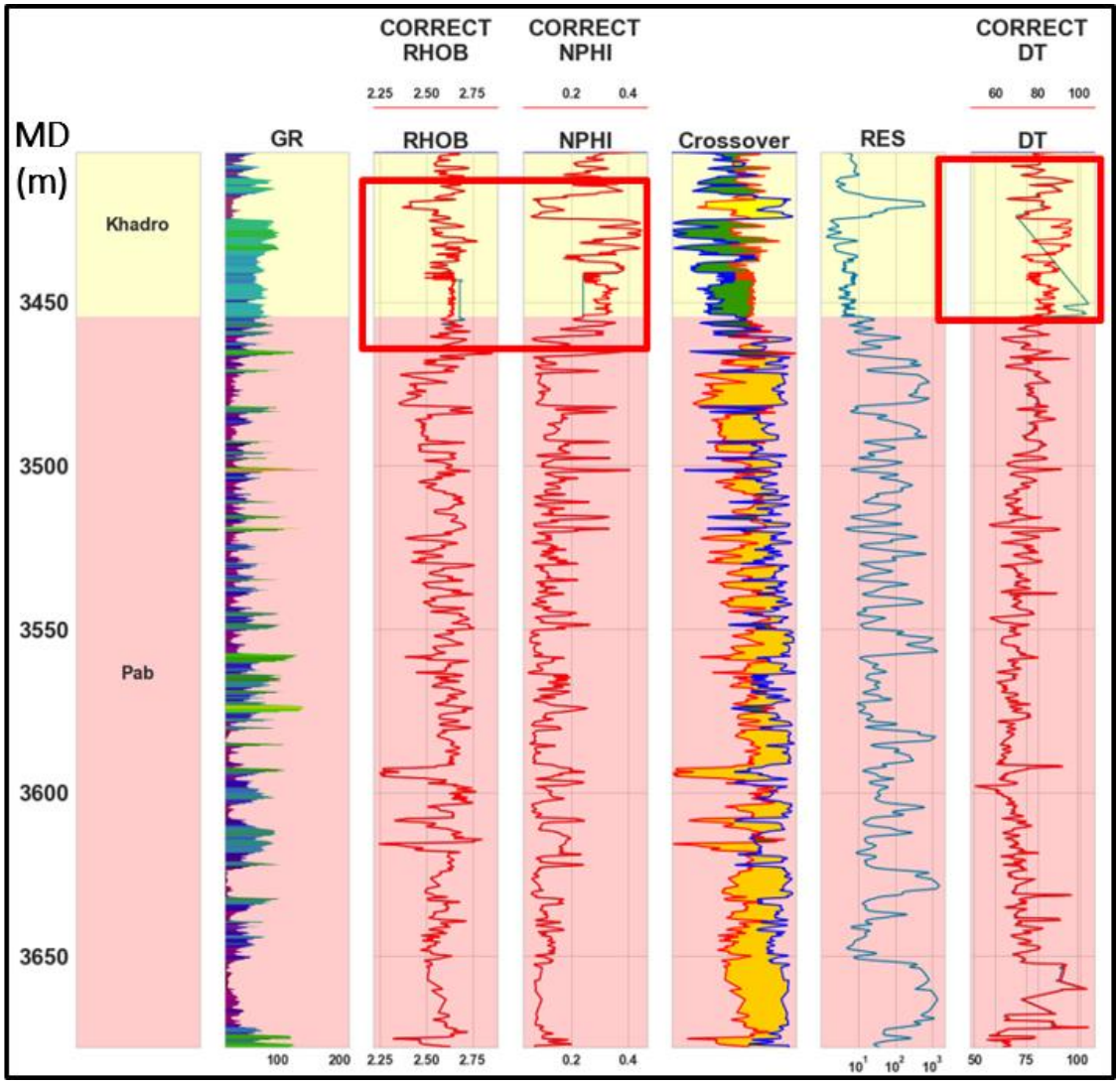



**Figure 4.9** The corrected spliced zone is bounded by the red rectangle while the red logs are the predicted through supervised *ML* techniques.

The best *ML* algorithm is identified through evaluation metrics applied on the different sample datasets extracted from training datasets. In the case of the splice zone affected density log, *ETR* predict the density log (Figure 4.10). In the case of the splice zone affected *NPHI* log, *RFR* predict the missing *NPHI* log quit reliably (Figure 4.11). In the case of the splice zone affected *DT* log, *ETR* predict the missing *DT* log (Figure 4.12).

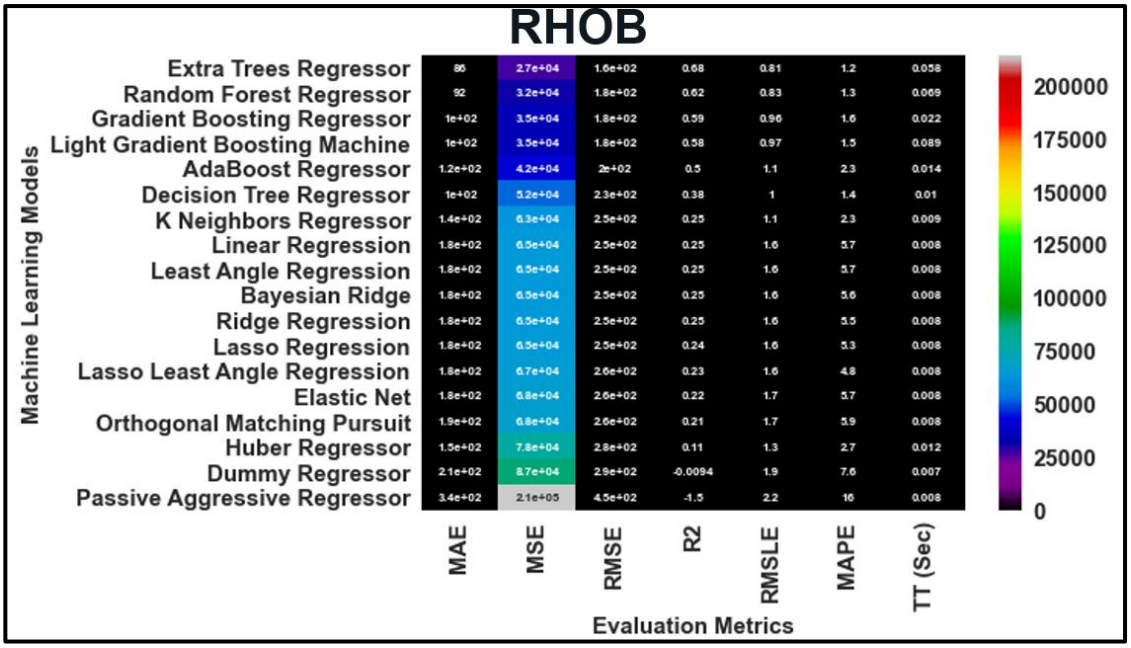



**Figure 4.10** Various *ML* algorithms and their evaluation metrics indicating best technique of *ETR* for the prediction of *RHOB* log.

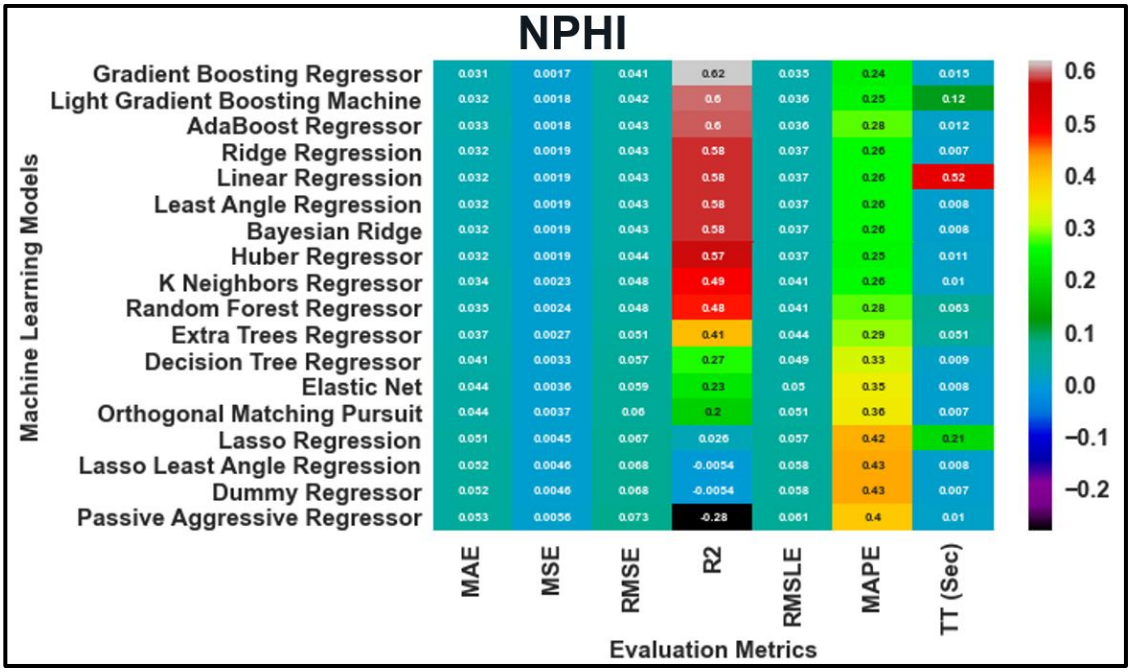



**Figure 4.11** Implemented *ML* algorithms with best technique for prediction of *NPHI* log i.e., *RFR*.

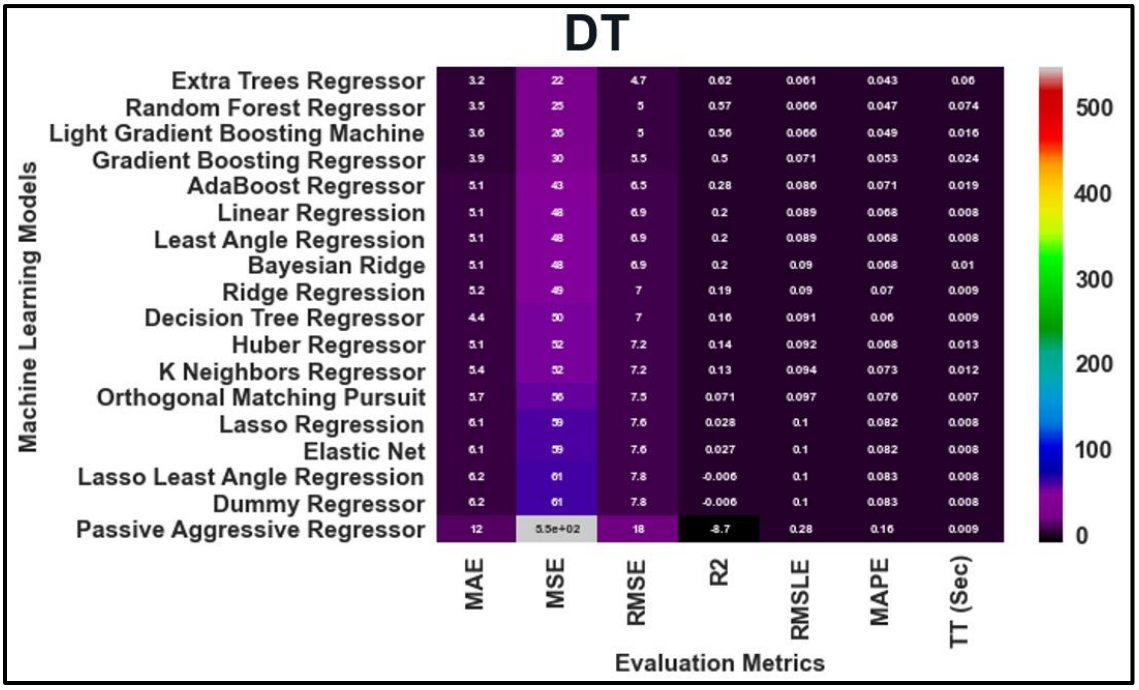



**Figure 4.12** *ML* algorithms with best technique for prediction of *DT* log such as *ETR*.

### 4.1.3.2 Splice Zones approximation of Mehar-01

The splice zone of Mehar-01 is done by adopting a similar approach. The models were trained using 60% of the data from Mehar-01 well, and the performance was validated using the remaining 40% of the data. The predicted *DT* (compressional wave velocity) and *RHOB* values are shown in red, while the measured logs are shown in blue in Figure 4.13. The trained models were then applied to the blind well, Mehar-02, to predict the *DT* and *RHOB* values, as shown in Figure 4.14. The first track shows the Ranikot/Khadro Formation zone and the gamma-ray, *CALI*, and bitesize measurements. The corrected density values can be compared to the measured density values in the *RHOB* (fourth) track, and the corrected and measured sonic values are depicted in the *DT* (sixth) track.

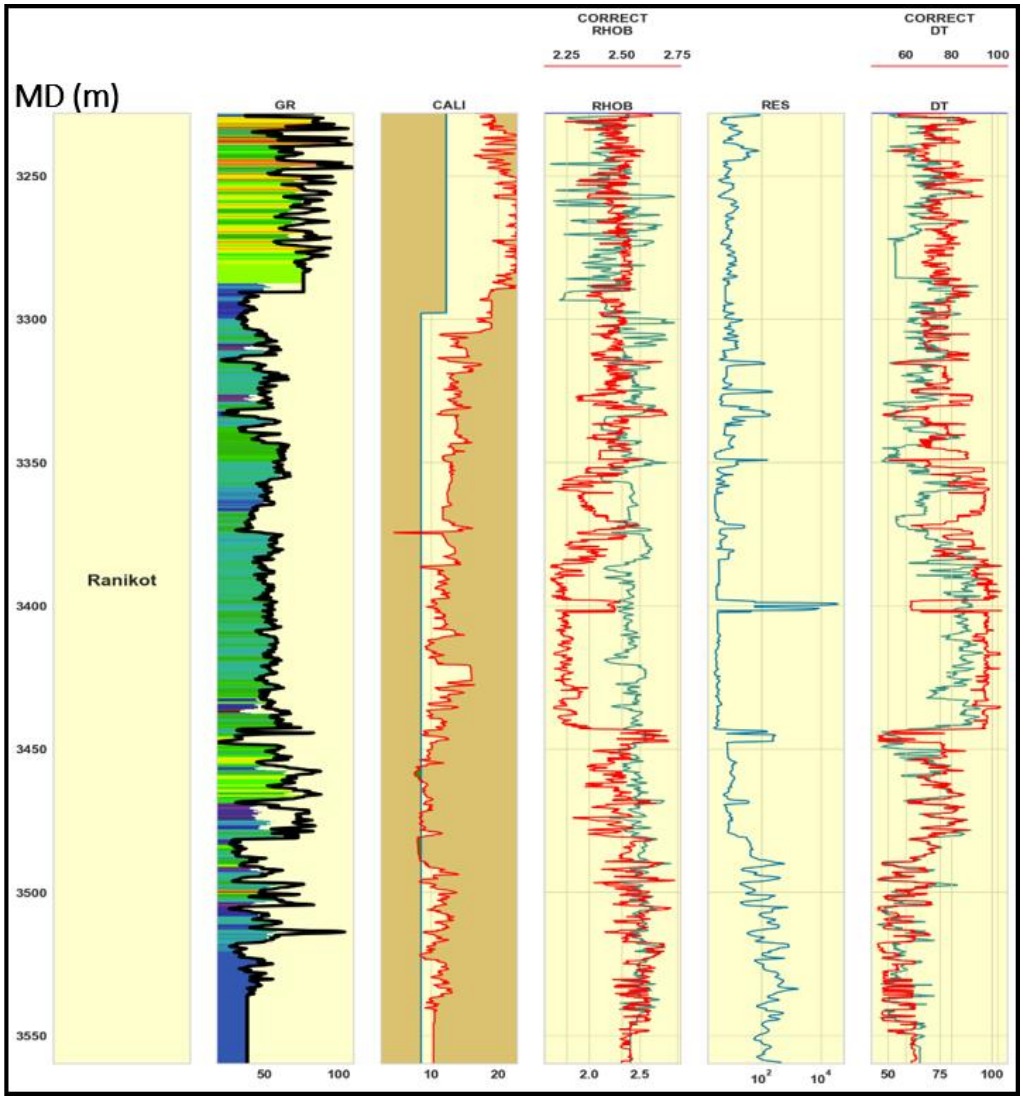



**Figure 4.13** Modeled and measured well log curves in training well i.e., Mehar-01. Blue color shows the measured curves while red color depicts corrected *RHOB* and *DT* in fourth and sixth track.

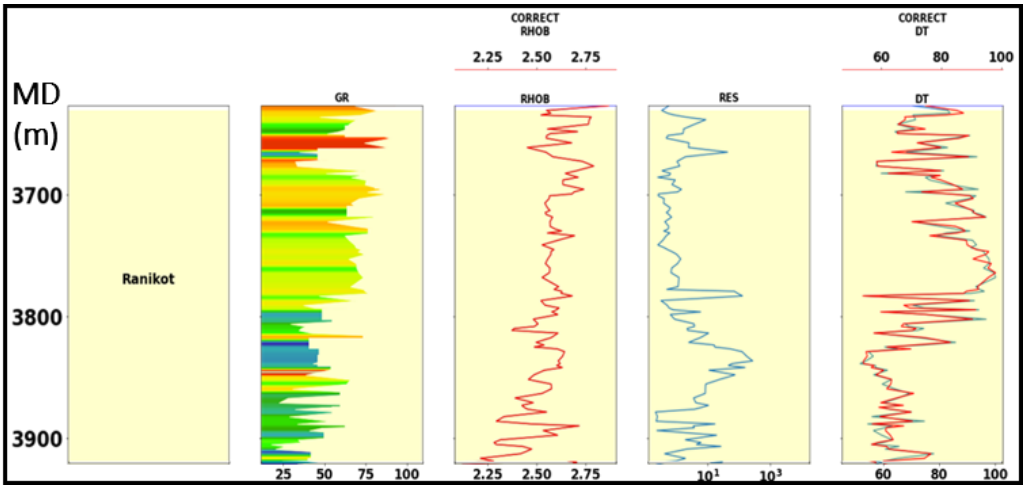



**Figure 4.14** Modeled and measured Well log curves in testing well i.e., Mehar-02. Blue color shows the measured curves while red color depicts corrected *RHOB* and *DT* in third and fifth track.

### 4.1.4 Washout Zones Correction through Supervised *ML* Algorithms of Zamzama Field

The Zamzama-02 density log is corrected via various algorithms, but best suited *ML* technique is selected for final approximation i.e., *ETR* Figure 4.15. The first track shows the zones of Ranikot/Khadro Formation and Pab Formation, second track shows the *GR*, third track is *CALI* and bitesize, fourth track consists of corrected density and measured density (mainly at the bounded area by the red box). The Fifth track is *NPHI*, sixth track is the deep resistivity and last track is the corrected p-sonic log. The *NPHI*, *RES*, and DT are employed to estimate the washout zones in the supervised *ML* technique.

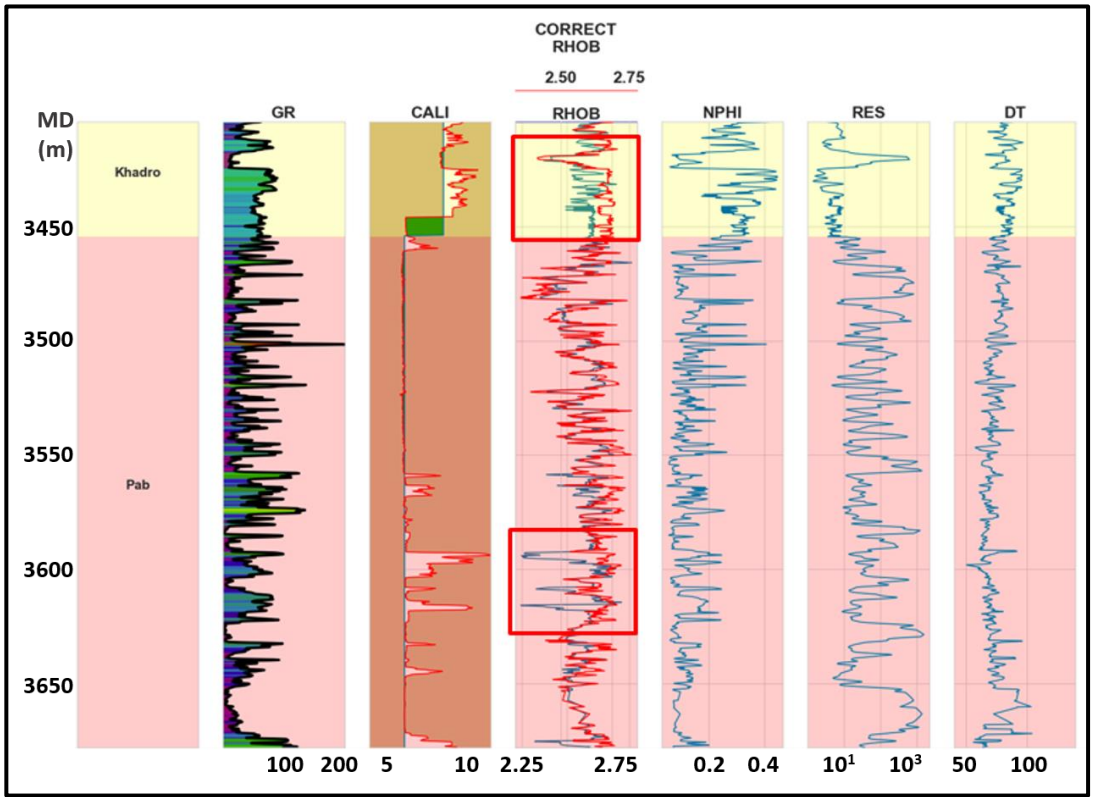



**Figure 4.15** Washout zones are corrected mainly at the red box bounding the corrected logs.

The selection of best method, i.e., *ETR* is made after detail evaluation of various *ML* technique as shown in the Figure 4.16. For all of the available wells, the technique is employed that effectively enhances the raw logs and brings them to their normal ranges.

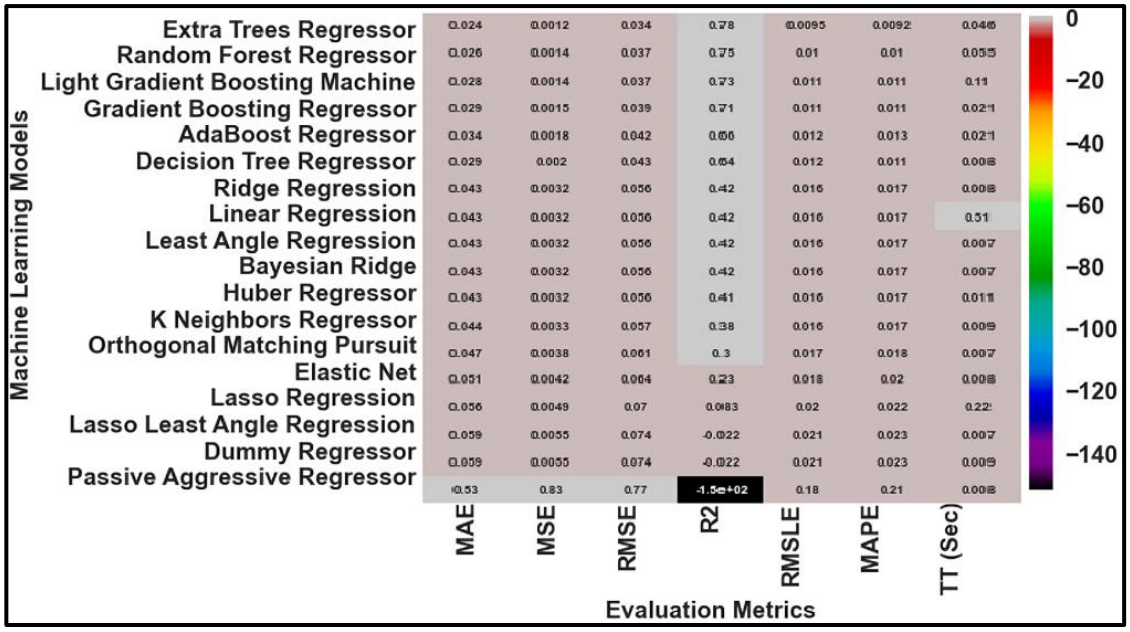



**Figure 4.16** Selection of best *ML* algorithm of *ETR* after evaluating various techniques.

### 4.1.4.1 Petrophysical interpretation of Zamzama-02 using conventional approach

The Ranikot/Khadro Formation at Zamzama-02 observed plausible gas sand body at 3420 m. The logs i.e., *LLD*, along with cross-over supported the interpreted potential sand body. According to Figure 4.17, the first track shows the zone of Ranikot/Khadro Formation and Pab Formation, second track shows the *GR*, third track consists of *RHOB*, forth track consists of the p-sonic (*DT*) while fifth trach comprises of deep resistivity (*RES*). The approximated clay volumetrics depicted low clay contents at the Ranikot/Khadro Formation level, hence provide a clue of good sand. Similarly, a cross-over is also observed while good *PHIE* is present (about 15%). In this sand interval, low $S_w$ is estimated while the facies are identified as gas sands in the last track.

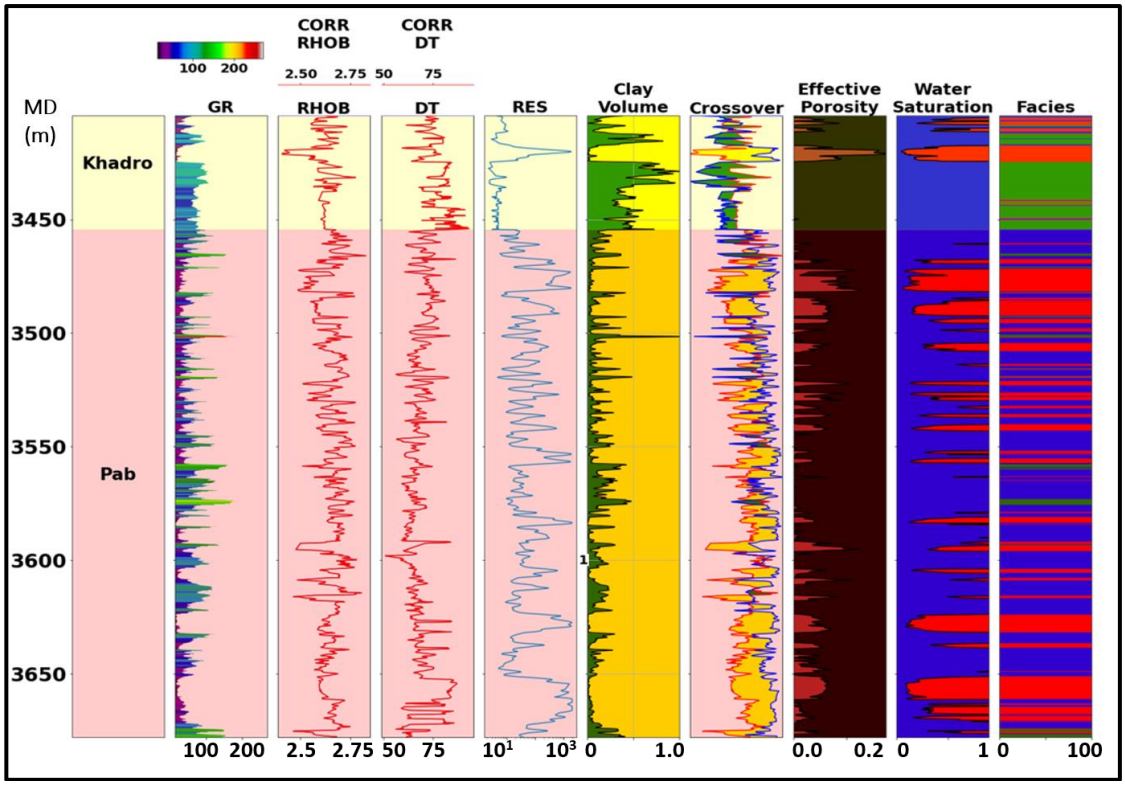



**Figure 4.17** Petrophysical interpretation of Zamzama-02 well indicated a potential sand body in the Ranikot/Khadro Formation.

### 4.1.4.2 Petrophysical interpretation of Zamzama-03 using conventional approach

Complete log suits are available for Zamzama-03 well. Well, Zamzama-03 is drilled at the depth of 4032 m, the last formation that is encounter in well is Fort Munro. As the Ranikot/Khadro Formation is our interested formation, therefore the petrophysical analysis is performed focusing on it (Figure 4.18). The similar arrangement of tracks is kept for Zamzama-03 as for Zamzama-02.

The petrophysical interpretation showed the hydrocarbon is present at the Pab Formation at some location and Ranikot/Khadro Formation shows two gas bearing zones at top of the formation named as Sand-01 and Sand-02. The result of two intervals is shown in the Table 4.4. A decent amount of gas is present at the Ranikot/Khadro Formation with good porosities (15-18%) (Khan et al., 2023).

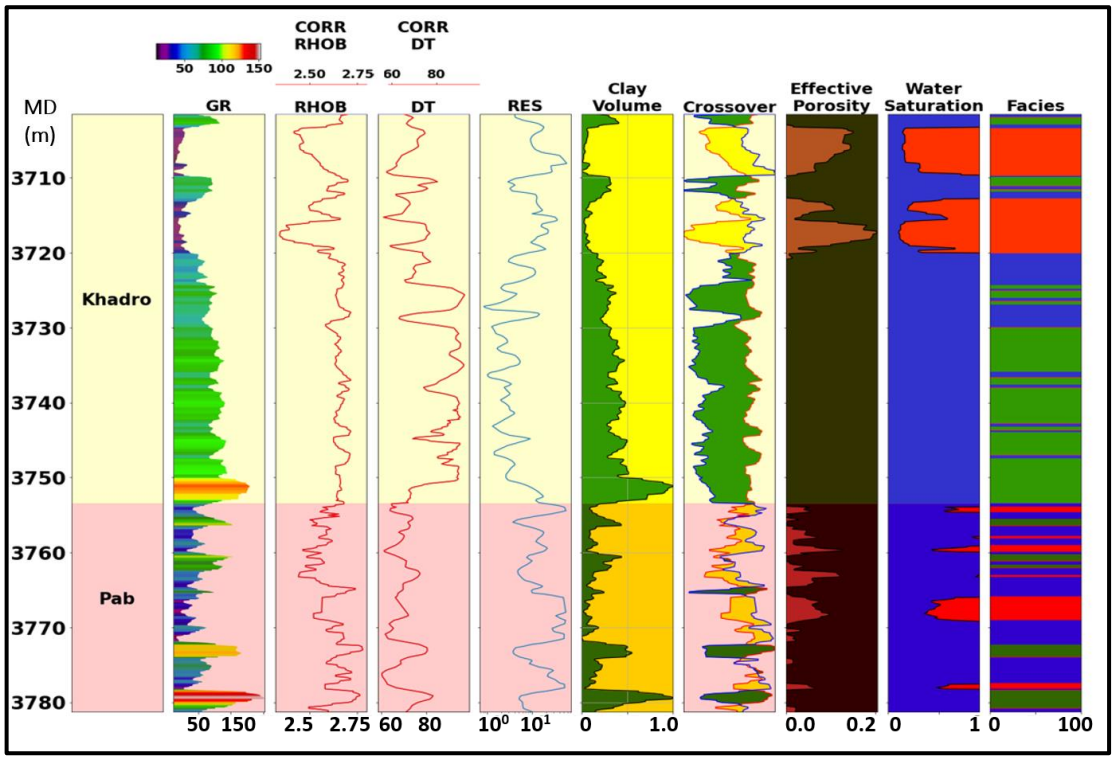



**Figure 4.18** Petrophysical analysis of Ranikot/Khadro Formation on Zamzama-03 along with the two interested sand of Ranikot/Khadro Sand-01 and Sand-02 at the top of Ranikot/Khadro Formation.

### 4.1.4.3 Petrophysical interpretation of Zamzama-04 using conventional approach

The petrophysical interpretation of Zamzama-04 delineated a few minor streaks of prospective sands that is evident by the low *GR* values with high deep resistivity within Ranikot/Khadro Formation. The effective porosities are also very low with minor streak of low $S_w$ i.e., depicted in the "$S_w$" track. Majorly, the Ranikot/Khadro Formation is significantly dominated by high content of shale (Figure 4.19).

### 4.1.4.4 Petrophysical Interpretation of Zamzama-05 using Conventional Approach

The logs available in Zamzama-05 are stable and the logs quality is good in the Ranikot/Khadro Formation. There is a considerable potential sand body at the top part of Ranikot/Khadro Formation that is interpreted with low *GR* values while contains less clay volumetrics. The *PHIE* range upto 13% in the potential sand body where low $S_w$ is observed (Figure 4.20).

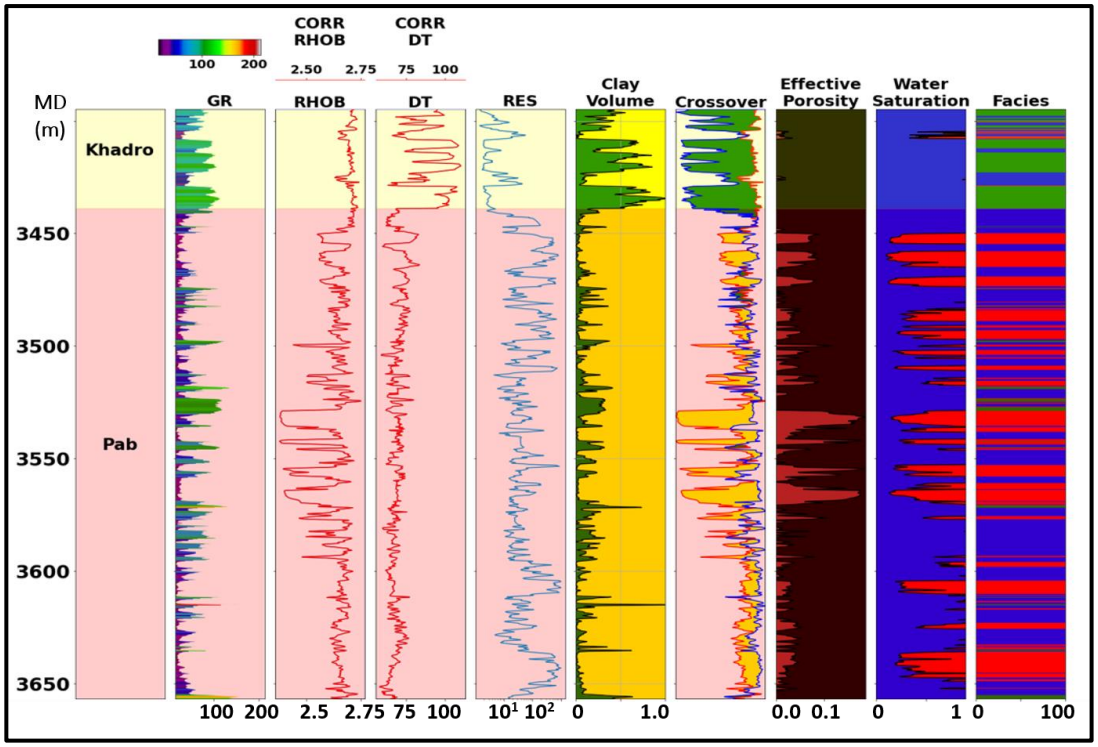



**Figure 4.19** Minor potential sands with low porosity and comparatively high-$S_w$ in Ranikot/Khadro Formation. The Ranikot/Khadro Formation is dominated by shales majority.

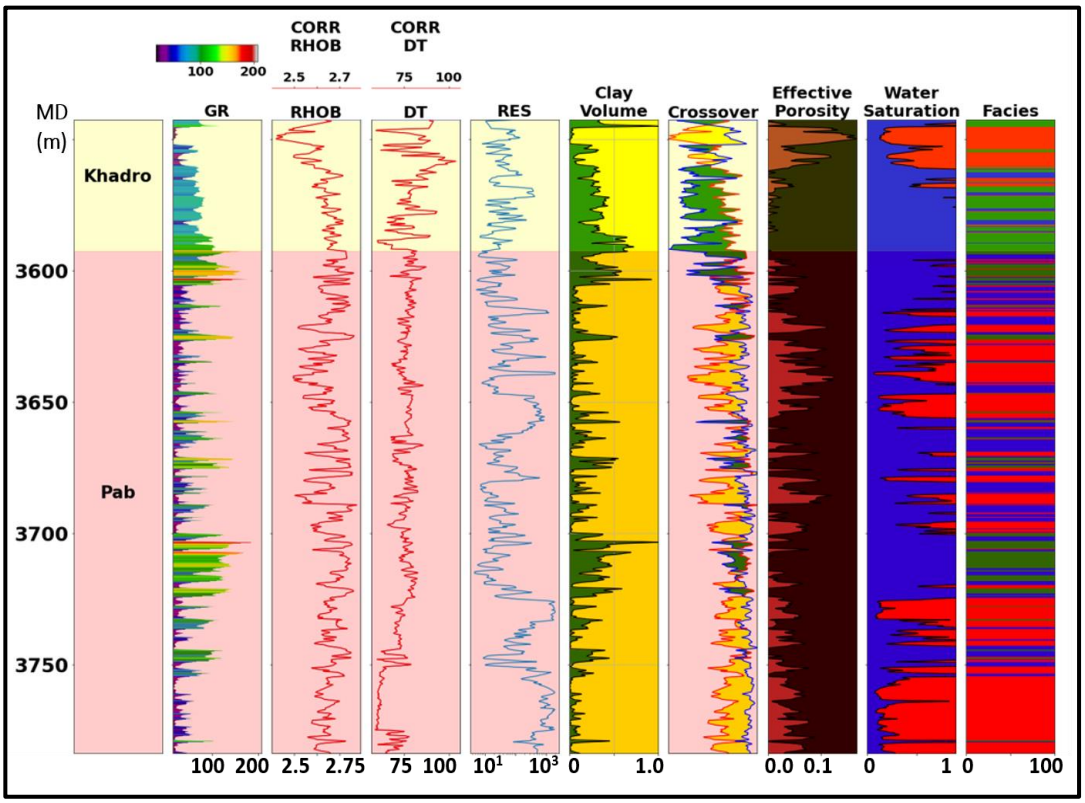



**Figure 4.20** The Ranikot/Khadro Formation bears plausible sand body where high *PHIE* with low $S_w$ is observed.

### 4.1.4.5 Petrophysical interpretation of Zamzama-06 using conventional approach

The Zamzama-06 well is not significant regarding the potential sand body. Less sand's concentration is evaluated at the Ranikot/Khadro Formation level as depicted by the high *GR* values. Effective porosities are fine but these zones are majorly comprises of water, therefore the gas saturations are negligible (Figure 4.21).

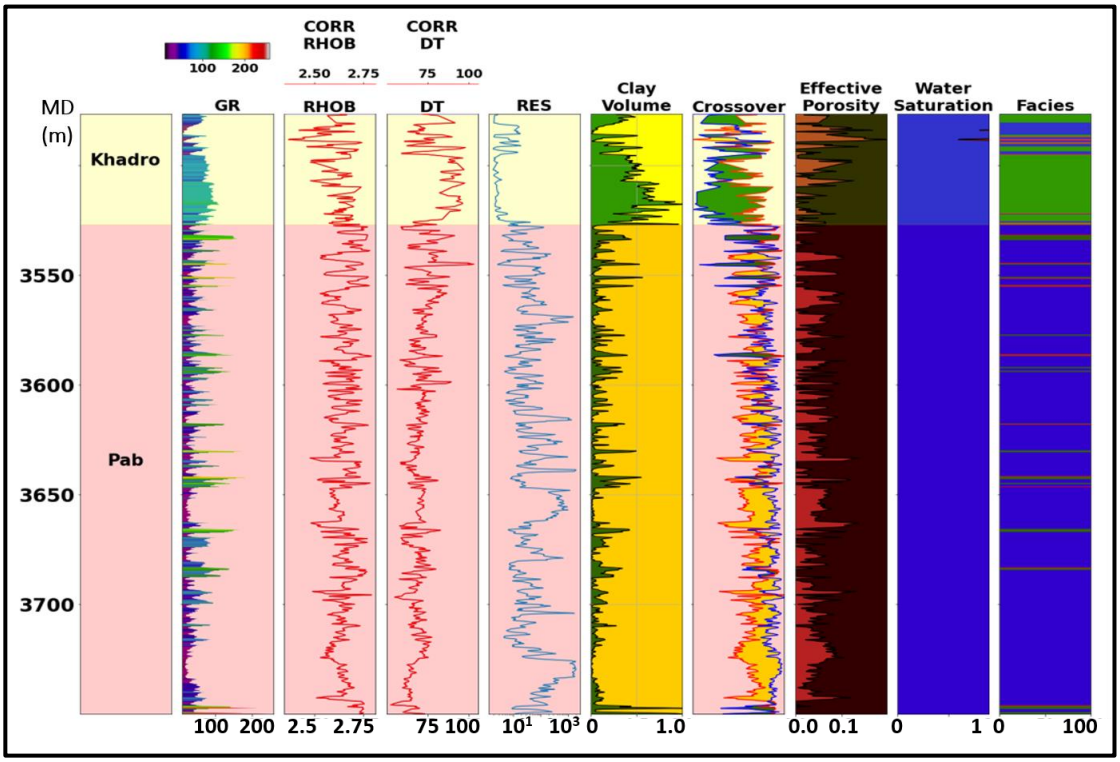



**Figure 4.21** Zamzama-06 well contains no significant potential sands with negligible gas saturations.

### 4.1.4.6 Petrophysical interpretation of Zamzama-07 using conventional approach

There are few streaks of sands in the Ranikot/Khadro Formation that are 2 to 3 m thick, however theses are filled with water with less concentration of gas. The *PHIE* is good upto 10% in the sands interval while the logs quality is fine (Figure 4.22).

### 4.1.4.7 Petrophysical Interpretation of Zamzama-08 using Conventional Approach

The Ranikot/Khadro Formation of Zamzama-08 is also potentially insignificant as depicted by the interpreted $S_w$ curve. Overall, the formation is filled with shales with certain streaks of sandstones. The porosity is fair to good in the sand streaks while the $S_w$ in the Ranikot/Khadro Formation is about 50% (Figure 4.23).

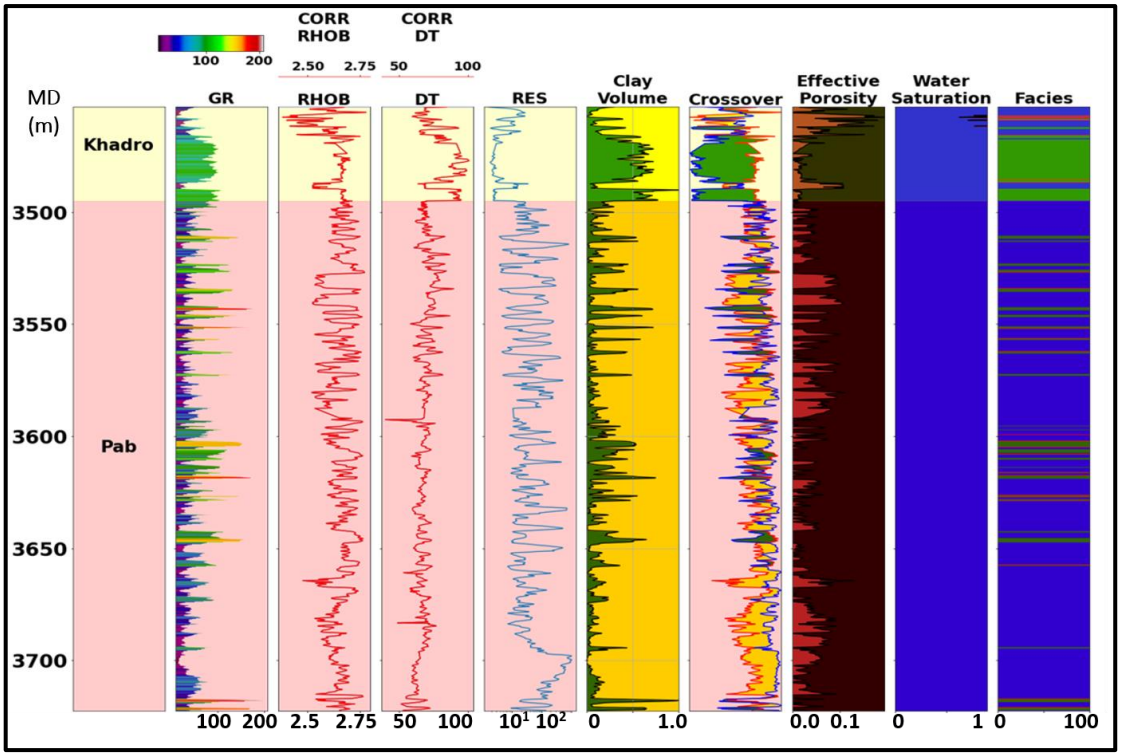



**Figure 4.22** Few sand streaks are depicted within Ranikot/Khadro Formation with low $V_{CL}$ while the $S_w$ curve showed no significance regarding the potential within sands.

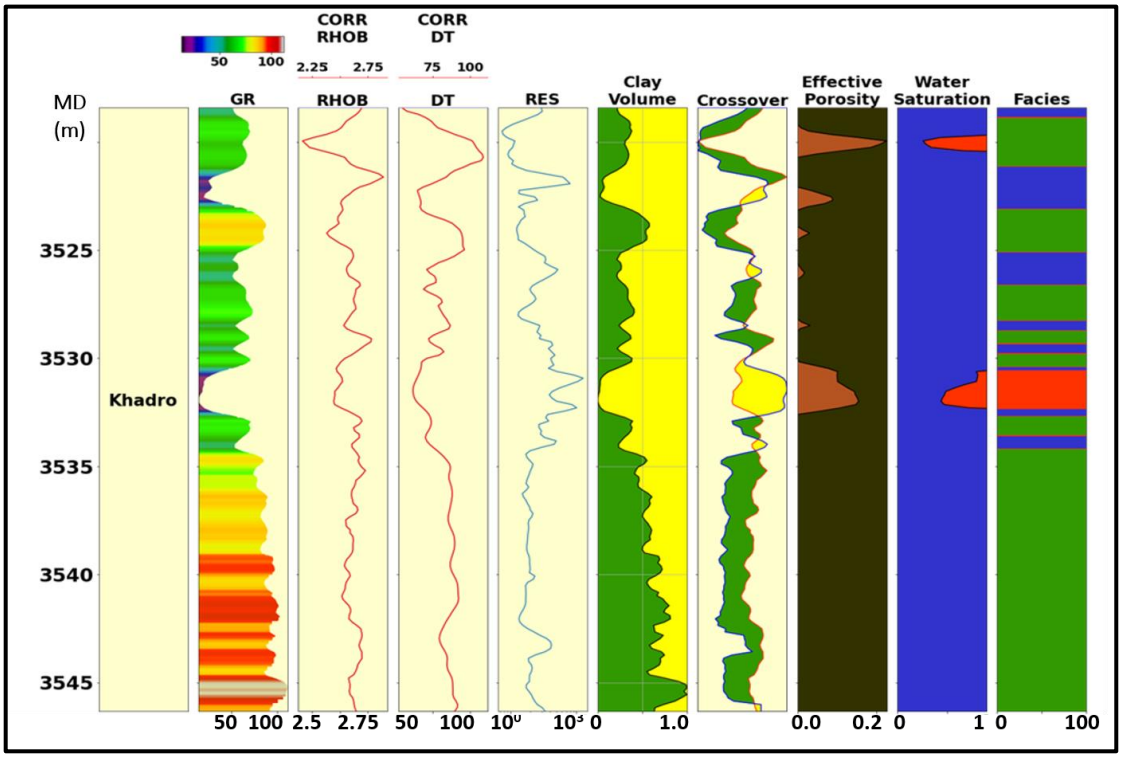



**Figure 4.23** Zamzama-08 depicts fair to good porosities at certain locations with low clay volumetrics but the $S_w$ is high in these streaks, i.e., above 50%.

A brief summary of the interpreted Ranikot/Khadro Formation is enlisted in Table 4.4 as follows. Basically, the Ranikot/Khadro Formation sands are divided into three sand types as Sand-01, 02 and 03.

**Table 4.4:** Petrophysical results of Ranikot/Khadro Formation.

| Wells | Sands | Avg. $S_W$ (fraction) | Avg. Clay (fraction) | Avg. *PHIE* (fraction) | Avg. Thickness (m) |
|---|---|---|---|---|---|
| **Zamzama-02** | Sand-01 | 0.37 | 0.22 | 0.08 | 1.7 |
| | Sand-02 | 0.16 | 0.055 | 0.15 | 6.1 |
| | Sand-03 | X | X | X | X |
| **Zamzama-03** | Sand-01 | 0.281 | 0.177 | 0.16 | 6 |
| | Sand-02 | 0.272 | 0.17 | 0.19 | 6 |
| | Sand-03 | X | X | X | X |
| **Zamzama-04** | Sand-01 | X | X | X | X |
| | Sand-02 | 0.39 | 0.19 | 0.05 | 2 |
| | Sand-03 | X | X | X | X |
| **Zamzama-05** | Sand-01 | 0.244 | 0.147 | 0.18 | 6 |
| | Sand-02 | 0.40 | 0.27 | 0.03 | 1 |
| | Sand-03 | X | X | X | X |
| **Zamzama-06** | Sand-01 | 0.90 | 0.23 | 0.12 | 2 |
| | Sand-02 | X | X | X | X |
| | Sand-03 | X | X | X | X |
| **Zamzama-07** | Sand-01 | 0.80 | 0.24 | 0.0973 | 2 |
| | Sand-02 | X | X | X | X |
| | Sand-03 | Bad bore-hole condition | | | |
| **Zamzama-08** | Sand-01 | 0.50 | 0.24 | 0.18 | 1 |
| | Sand-02 | 0.60 | 0.18 | 0.17 | 2 |
| | Sand-03 | X | X | X | X |

### 4.1.4.8 Petrophysical Interpretation of Mehar-01 using Conventional Approach

Complete log suits are available for Mehar 01 well. Well, Mehar-01 is drilled at the depth of 3999 m, the last formation that is encounter is Mehar-01 is Mughalkot. The formation tops of Mehar 01 are show in Table 4.5. The complete log data is only available

till Ranikot/Khadro Formation, so petrophysical analysis is only done on Ranikot/Khadro Formation in Mehar-01 the depth of formation is starting from 3228m and end on 3560m.

**Table 4.5:** Borehole Stratigraphy of Mehar-01.

| Formations | Formation Age | Formation Tops (m) | Thickness (m) |
|---|---|---|---|
| Siwalik | Miocene/ Pleistocene | 17 | 813 |
| Gaj | Miocene | 830 | 445 |
| Upper Nari | Oligocene | 1275 | 576 |
| Lower Nari | Oligocene | 1851 | 152 |
| Kirthar Limestone | Eocene | 2003 | 354 |
| Ghazij | Eocene | 2357 | 87 |
| Sui main Limestone | Eocene | 2444 | 618 |
| Dunghan | Paleocene | 3062 | 166 |
| Ranikot/ Khadro | Paleocene | 3228 | 332 |
| Pab Formation | Late Cretaceous | 3560 | 113 |
| Mughalkot | Late Cretaceous | 3673 | 132 |

In Mehar-01 only Ranikot/Khadro Formation is marked as shown in Figure 4.24. In Mehar-01, the hydrocarbon is saturated all over the Ranikot/Khadro Formation in various layers. Reservoir Facies have been interpreted in the facies column after applying Table 4.1 petrophysical cut-offs systematically. The petrophysical properties summary of Mehar-01 well is presented in Table 4.6.

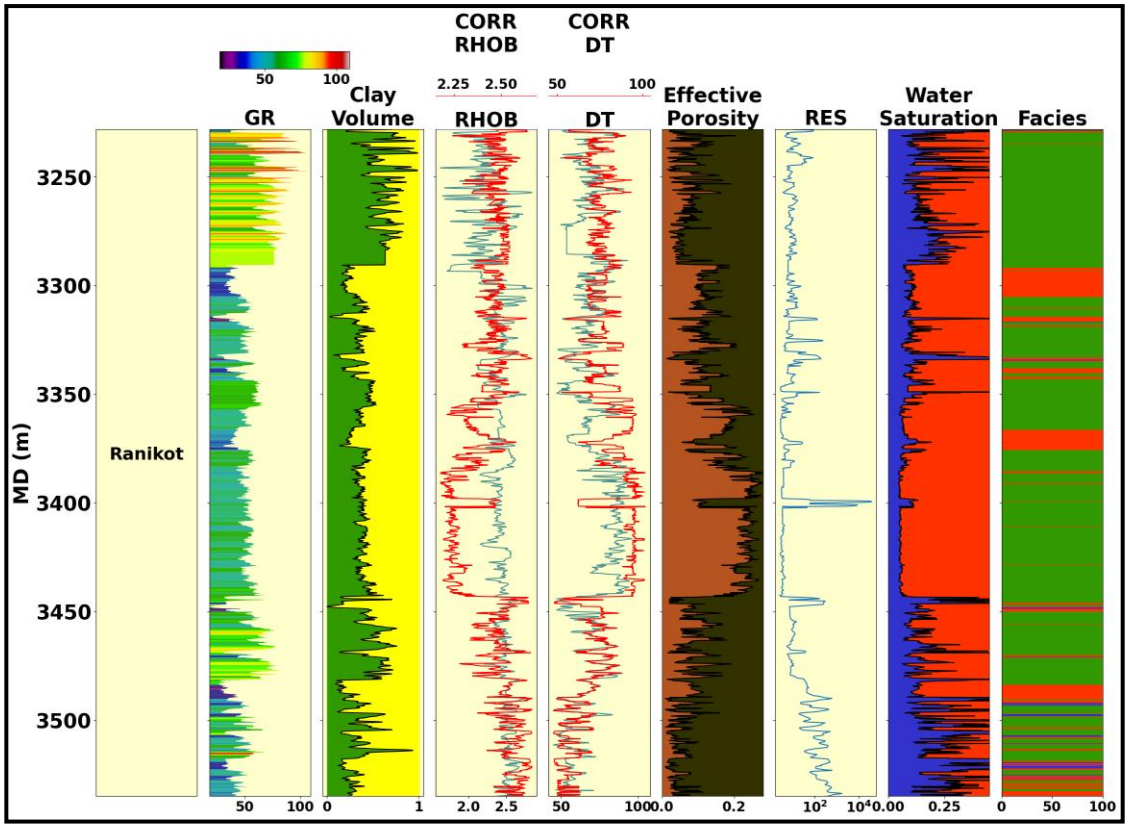



**Figure 4.24** Petrophysical interpretation of Mehar-01 depicting interpreted petrophysical properties (clay/shale volume, *PHIE*, and $S_w$) in their respective tracks while 2nd, 4th, 5th, and 6th tracks show the measured *GR*, *RHOB*, *DT* and *RES* logs respectively as well. Modeled *RHOB* and *DT* are represented by red color curves whereas blue represents measured logs. In the *PHIE* track light brown color reflects *PHIE* whereas in *Sw* track blue color shows the water. Facies track reflects the respective facies after applying petrophysical cut off ranges.

**Table 4.6:** Result of reservoir zones marked on Ranikot/Khadro Formation based on Mehar-01 well.

| **Ranikot/Khadro Formation** | | | | | |
|---|---|---|---|---|---|
| **Zone 1 (3239m- 3247m)** | | | | | |
| **Thickness(m)** | ***V*$_{SH}$ (%)** | ***PHIA* (%)** | ***PHIE* (%)** | ***S*$_W$ (%)** | ***S*$_H$ (%)** |
| 8 | 61 | 13 | 10 | 51 | 49 |
| **Zone 2 (3518m-3538m)** | | | | | |
| **Thickness(m)** | ***V*$_{SH}$ (%)** | ***PHIA* (%)** | ***PHIE* (%)** | ***S*$_W$ (%)** | ***S*$_H$ (%)** |
| 20 | 23 | 14 | 9 | 38.60 | 62.40 |

### 4.1.4.9 Petrophysical Interpretation of Mehar-02 using Conventional Approach

Complete log suits are available for Mehar-02. Mehar-02 is drilled at the depth of 4300 m, the last formation that is encounter is Mehar-02 is Pab Formation. The formation

tops of Mehar-02 are shown on the Table 4.7. The complete log data is only available till Pab Formation, so petrophysical analysis is done on Ranikot/Khadro Formation and Pab Formation in Mehar-02.

**Table 4.7:** Borehole Stratigraphy of Mehar-02.

| Formations | Formation Age | Formation Tops (m) | Thickness (m) |
|---|---|---|---|
| Siwalik | Miocene/ Pleistocene | 0 | 895 |
| Gaj | Miocene | 895 | 375 |
| Upper Nari | Oligocene | 1270 | 593 |
| Lower Nari | Oligocene | 1863 | 154 |
| Kirthar | Eocene | 2017 | 353 |
| Ghazij Formation | Eocene | 2370 | 71 |
| Sui Main Limestone | Eocene | 2441 | 602 |
| Dunghan | Paleocene | 3043 | 581 |
| Ranikot/Khadro | Paleocene | 3624 | 298 |
| Pab Formation | Late Cretaceous | 3922 | 316 |

Ranikot/Khadro Formation is lies at the depth of 3622 to 3922 m and the hydrocarbon is saturated at the bottom of Ranikot/Khadro Formation (Figure 4.25). The summary of petrophysical properties within identified zone in Mehar-02 is presented in Table 4.8.

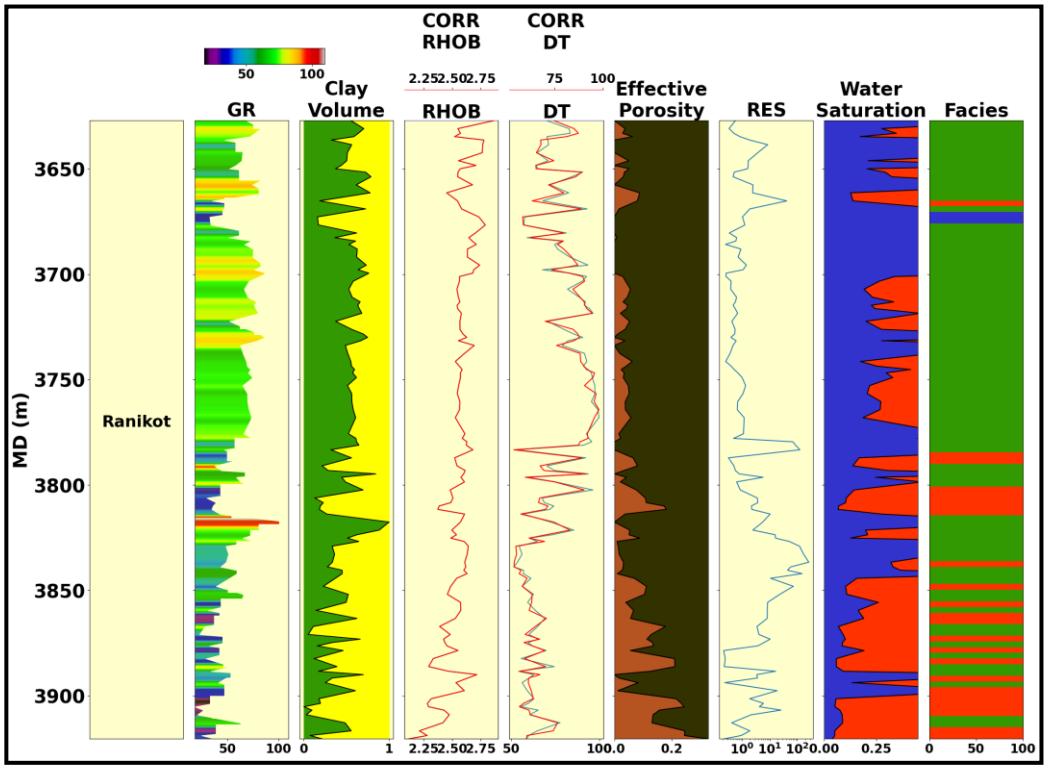



**Figure 4.25** Petrophysical interpretation of Mehar-02 is shown in $V_{CL}$, *PHIE*, and $S_w$ tracks while 2nd, 4th, 5th, and 6th tracks show the measured *GR*, *RHOB*, *DT* and *RES* logs respectively. Modeled *RHOB* and DT are represented by red color curves whereas blues represents measured logs. In the *PHIE* track light brown color reflects *PHIE* whereas in $S_w$ track blue color shows the water. Facies track reflects the respective facies after applying petrophysical cut off ranges.

**Table 4.8:** Result of reservoir zones marked on Ranikot/Khadro Formation based on Mehar-02.

| **Ranikot/Khadro Formation** | | | | | |
|---|---|---|---|---|---|
| **Zone 1 (3842m-3857m)** | | | | | |
| Thickness (m) | $V_{SH}$ (%) | *PHIA* (%) | *PHIE* (%) | $S_W$ (%) | $S_H$ (%) |
| 17 | 38 | 13.5 | 9.5 | 43.79 | 56.21 |
| **Zone 2 (3860m -3890m)** | | | | | |
| Thickness (m) | $V_{SH}$ (%) | *PHIA* (%) | *PHIE* (%) | $S_W$ (%) | $S_H$ (%) |
| 30 | 22.59 | 14 | 10.3 | 38.60 | 62.40 |
| **Zone 3 (3900m- 3917m)** | | | | | |
| Thickness(m) | $V_{SH}$ (%) | *PHIA* (%) | *PHIE* (%) | $S_W$ (%) | $S_H$ (%) |
| 17 | 35 | 12.5 | 8.3 | 54 | 46 |

### 4.1.5 Advanced *ML* Methods

Many of the advance novel *ML* techniques have been applied for evaluating the key petrophysical properties including porosities, volumetric, and saturations of wells.

#### 4.1.5.1 Petrophysical Properties Estimation on Blind Well Mehar-02

Among the use of traditional approaches, powerful *ML* is applied to identify petrophysical properties. The optimal model for a precise interpretation of well logs at various well locations is determined using a variety of *ML* models. The best model for accurate prediction of blind wells has been found using the heatmap of *ML* algorithm performance. To test the effectiveness of *ML* methods, dataset of Mehar-01 is trained, whereas Mehar-02 is kept blind to validate the results. After quality control using supervised and unsupervised learning algorithms, the petrophysical parameters such as the amount of shaly content, effective porosities, and $S_w$ is predicted using raw logs. Many algorithms are used to test the Mehar-01, whereas, the Extra Trees Regressor has outdated all the other algorithms for the predictions of volume of shale and effective porosities at the blind well Mehar-02. Whereas the *RFR* has best performance among all the other algorithms for the predictions of $S_w$.

The Extra Trees Regressor is an ensemble supervised *ML* approach that employs decision trees. This class implements a meta estimator that uses averaging to improve

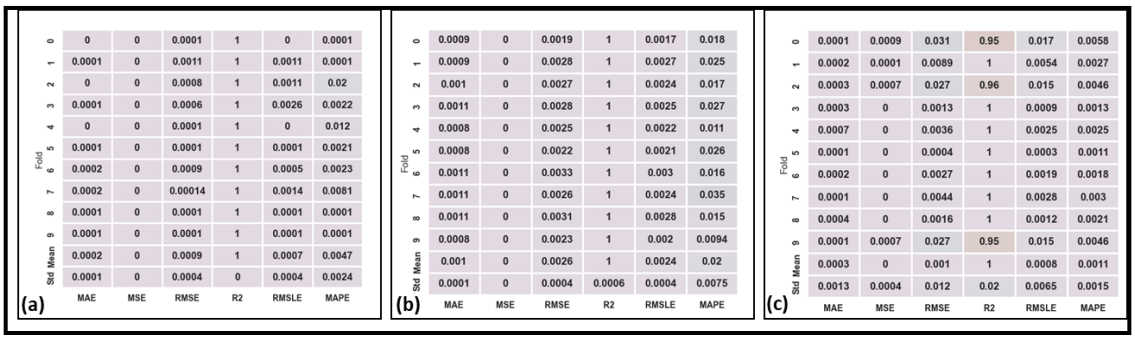
predicted accuracy and minimize overfitting by fitting several randomized decision trees (also known as extra-trees) on different sub-samples of the dataset. Extra trees is an ensemble-based technique which uses a cluster of decision trees to generate an average output (Hui et al., 2021). It operates by taking a vast number of decision trees to predict an average of the parameter of interest (Genuer et al., 2017; Geurts, et al., 2006; Breiman, 2001). Based on these groups, the prediction models are run by performing the computational algorithm. Finally, the accuracy of the results is analyzed using a coefficient of determination and mean squared error (Hui et al., 2021).

When it comes to predicting $S_w$, the *RFR* has outperformed all previous algorithms. The *RF* algorithm utilizes many decision trees to predict the average values of a parameter of interest by training the decision trees on random data sets (Otchere et al., 2022). *RF* is a commonly used algorithm as it is robust, minimizes the overfitting risk, and is easily interpretable (Breiman, 2001). This is due to its ability to estimate the relationship between the input variables with the target and subsequently prioritize the correlated features (Otchere et al., 2022).

#### 4.1.5.1.1 Cross Validation

K-fold cross-validation is a popular *ML* technique for model evaluation and performance estimates. The study employs a random subsampling cross-validation strategy because the available dataset is limited and must be efficiently utilized for both model training and testing. The initial dataset is partitioned into K subsets or folds of roughly equal size at random. The model is trained for each fold using training data from the remaining K-1 folds. After training with the training set, the model is tested on the held-out fold. The performance metrics are calculated in order to assess the model's performance on this fold. K folds, each serving as the testing set once, iteratively. At each fold, the testing subset represented by (K=i), i ranges from 0 to 9. The procedure is repeated K times, with each fold serving as a testing set only once. The model's overall performance is estimated by averaging the performance metrics from each iteration. Their means reflect the score of algorithms as mentioned before. The K Fold cross validation helps in avoiding overfitting and estimate good results. Figure 4.26 reflects the highest

$R^2$ value and least MSE for optimized 10 folds authenticating the outputs and validate the opted workflow performance.



(a)

| Fold | MAE | MSE | RMSE | R2 | RMSLE | MAPE |
|---|---|---|---|---|---|---|
| 0 | 0 | 0 | 0.0001 | 1 | 0 | 0.0001 |
| 1 | 0.0001 | 0 | 0.0011 | 1 | 0.0011 | 0.0001 |
| 2 | 0 | 0 | 0.0008 | 1 | 0.0011 | 0.02 |
| 3 | 0.0001 | 0 | 0.0006 | 1 | 0.0026 | 0.0022 |
| 4 | 0 | 0 | 0.0001 | 1 | 0 | 0.012 |
| 5 | 0.0001 | 0 | 0.0001 | 1 | 0.0001 | 0.0021 |
| 6 | 0.0002 | 0 | 0.0009 | 1 | 0.0005 | 0.0023 |
| 7 | 0.0002 | 0 | 0.00014 | 1 | 0.0014 | 0.0081 |
| 8 | 0.0001 | 0 | 0.0001 | 1 | 0.0001 | 0.0001 |
| 9 | 0.0001 | 0 | 0.0001 | 1 | 0.0001 | 0.0001 |
| Mean | 0.0002 | 0 | 0.0009 | 1 | 0.0007 | 0.0047 |
| Std | 0.0001 | 0 | 0.0004 | 0 | 0.0004 | 0.0024 |

(b)

| Fold | MAE | MSE | RMSE | R2 | RMSLE | MAPE |
|---|---|---|---|---|---|---|
| 0 | 0.0009 | 0 | 0.0019 | 1 | 0.0017 | 0.018 |
| 1 | 0.0009 | 0 | 0.0028 | 1 | 0.0027 | 0.025 |
| 2 | 0.001 | 0 | 0.0027 | 1 | 0.0024 | 0.017 |
| 3 | 0.0011 | 0 | 0.0028 | 1 | 0.0025 | 0.027 |
| 4 | 0.0008 | 0 | 0.0025 | 1 | 0.0022 | 0.011 |
| 5 | 0.0008 | 0 | 0.0022 | 1 | 0.0021 | 0.026 |
| 6 | 0.0011 | 0 | 0.0033 | 1 | 0.003 | 0.016 |
| 7 | 0.0011 | 0 | 0.0026 | 1 | 0.0024 | 0.035 |
| 8 | 0.0011 | 0 | 0.0031 | 1 | 0.0028 | 0.015 |
| 9 | 0.0008 | 0 | 0.0023 | 1 | 0.002 | 0.0094 |
| Mean | 0.001 | 0 | 0.0026 | 1 | 0.0024 | 0.02 |
| Std | 0.0001 | 0 | 0.0004 | 0.0006 | 0.0004 | 0.0075 |

(c)

| Fold | MAE | MSE | RMSE | R2 | RMSLE | MAPE |
|---|---|---|---|---|---|---|
| 0 | 0.0001 | 0.0009 | 0.031 | 0.95 | 0.017 | 0.0058 |
| 1 | 0.0002 | 0.0001 | 0.0089 | 1 | 0.0054 | 0.0027 |
| 2 | 0.0003 | 0.0007 | 0.027 | 0.96 | 0.015 | 0.0046 |
| 3 | 0.0003 | 0 | 0.0013 | 1 | 0.0009 | 0.0013 |
| 4 | 0.0007 | 0 | 0.0036 | 1 | 0.0025 | 0.0025 |
| 5 | 0.0001 | 0 | 0.0004 | 1 | 0.0003 | 0.0011 |
| 6 | 0.0002 | 0 | 0.0027 | 1 | 0.0019 | 0.0018 |
| 7 | 0.0001 | 0 | 0.0044 | 1 | 0.0028 | 0.003 |
| 8 | 0.0004 | 0 | 0.0016 | 1 | 0.0012 | 0.0021 |
| 9 | 0.0001 | 0.0007 | 0.027 | 0.95 | 0.015 | 0.0046 |
| Mean | 0.0003 | 0 | 0.001 | 1 | 0.0008 | 0.0011 |
| Std | 0.0013 | 0.0004 | 0.012 | 0.02 | 0.0065 | 0.0015 |

**Figure 4.26** Expressed K-fold Cross validation against each fold from 0-9 folds respectively for (a) Shale volume, (b) *PHIE* and (c) $S_w$. It represents statistical measures at each fold for the respective petrophysical property along with mean and standard deviation of K-folds.

The performance of various algorithms was evaluated using statistical metrics such as mean absolute error (*MAE*), root mean squared error (*RMSE*), and $R^2$, etc. as demonstrated in Figure 4.27, and Figure 4.28, and revealed that the *ETR* performed the best for predicting volume of shale and effective porosities with a maximum correlation coefficient of 1. Similarly, the evaluation metrics for predicting the $S_w$ has also been calculated, as shown in Figure 4.29, indicating the correlation coefficients and errors for each model, and declared *RFR* as the best model.

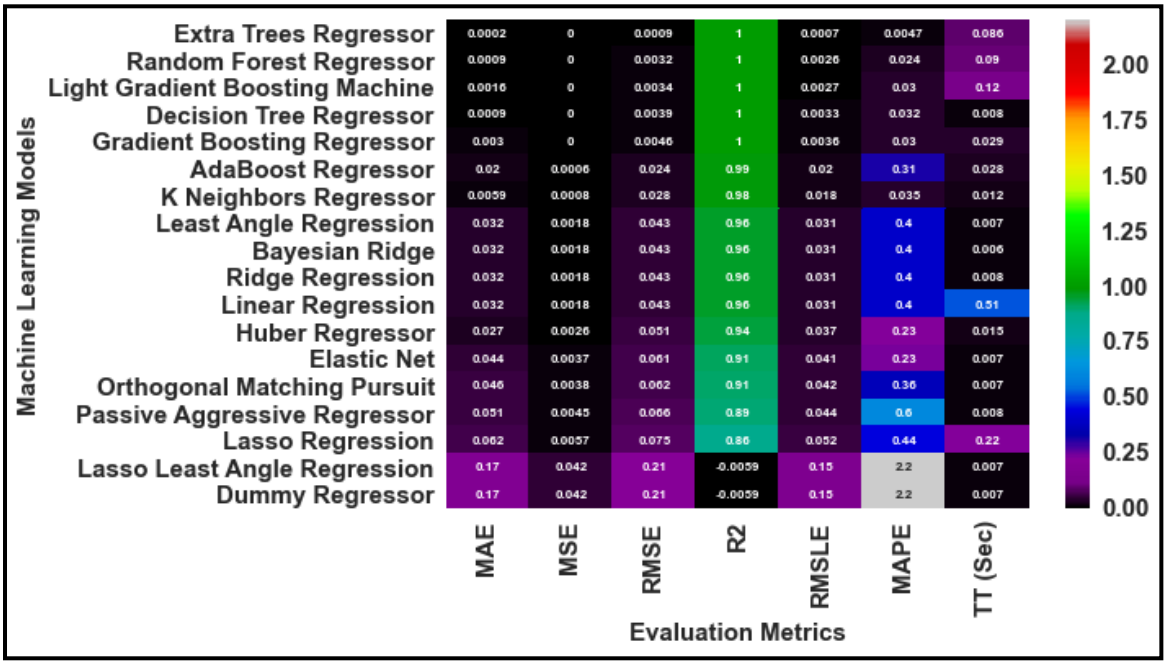



**Figure 4.27** Evaluation metrics for volume of shale reflecting the performance of algorithms (Akram et al., 2024).

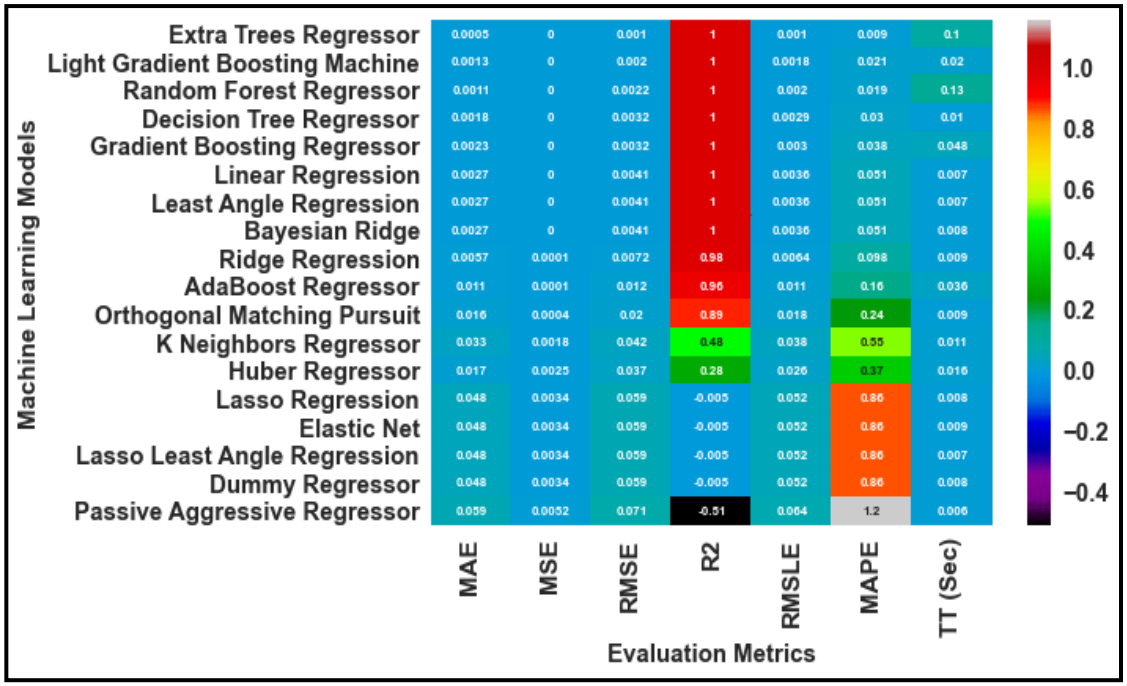



**Figure 4.28** Evaluation metrics for *PHIE* highlighting the statistical measures.

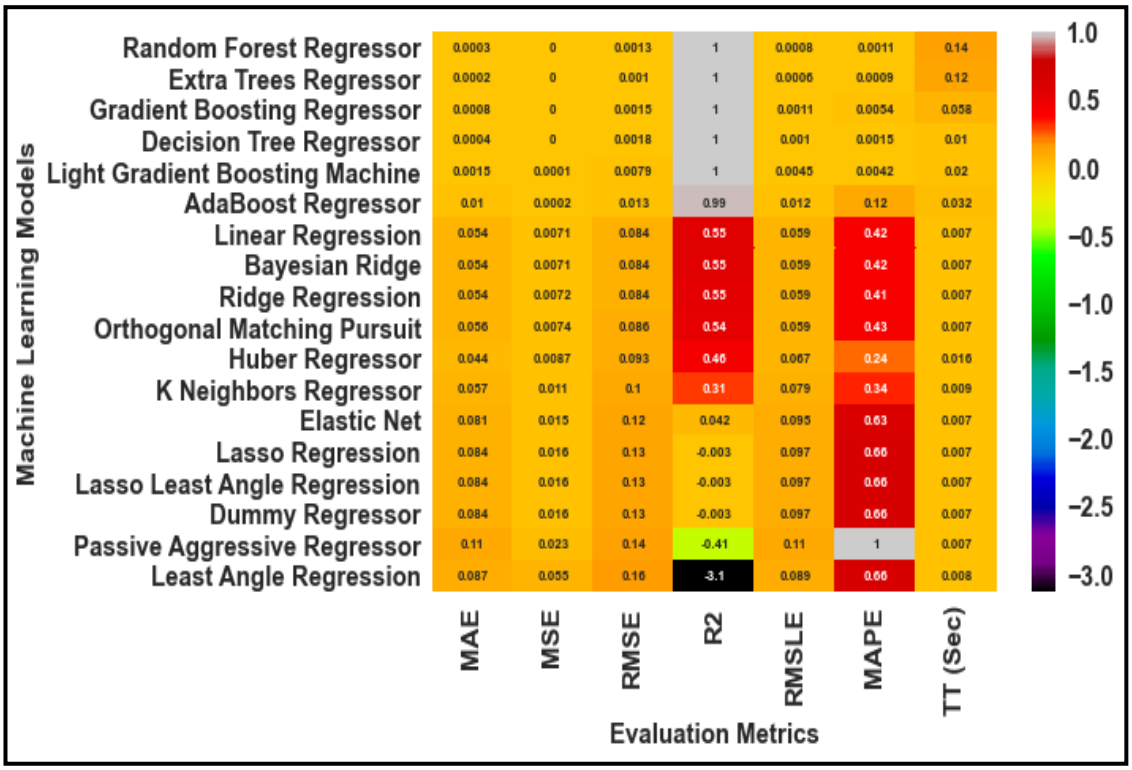



**Figure 4.29** Evaluation metrics of $S_w$ indicating performance of deployed algorithms.

#### 4.1.5.1.2 Feature Importance

The feature importance describes the input features that have a significant impact on the model's predictions. High-significance characteristics have substantial

relationships with the target variable. *GR* reflects the strong characteristics that provide useful insights into the volume of clay. Because the other parameters have 0% feature relevance, they are useless for predicting $V_{CL}$ (Figure 4.30a). In Figure 4.30b *GR*, *DPHI*, density represents the prominent characteristics that provide useful information regarding PHI. $V_{CL}$ -Linear, *GR*, and the other features have zero feature importance, indicating that they are unhelpful for predicting *PHIE*. Density and *PHIE* display the strong features that provide useful information regarding $S_w$ in Figure 4.30c.

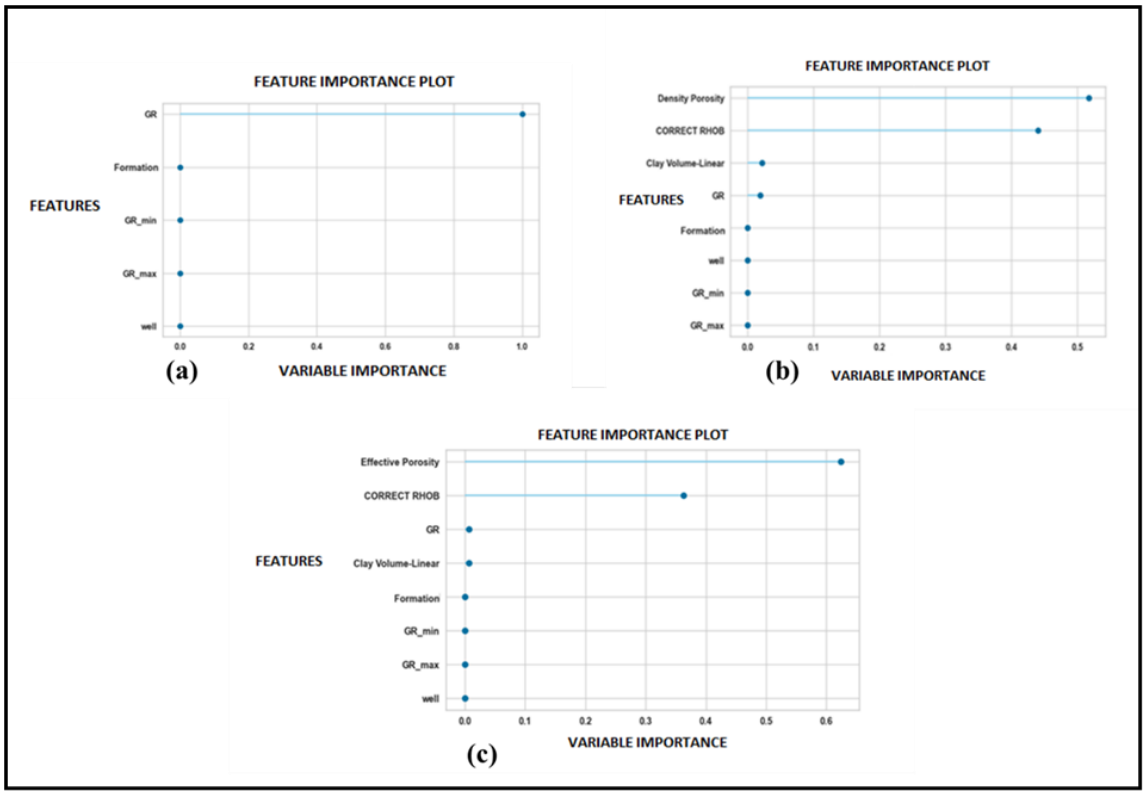



**Figure 4.30** (a) Feature Importance Plot: *ETR* for volume of clay predictions, (b) *PHIE* predictions (c) *RFR* for $S_w$ predictions.

The petrophysical properties estimated in Mehar-01 were further validated through the implementation of models that showed the best performance in Mehar-02 and after the feature importance study. The volume of shale, $S_w$, and effective porosities was predicted in Mehar-02 as a blind well. The results exhibit a strong correlation between the predicted and measured properties, as seen in Figure 4.31 (a, b, c), with the predicted curves shown in red and the conventional petrophysical properties shown in blue for comparison in Mehar-02.

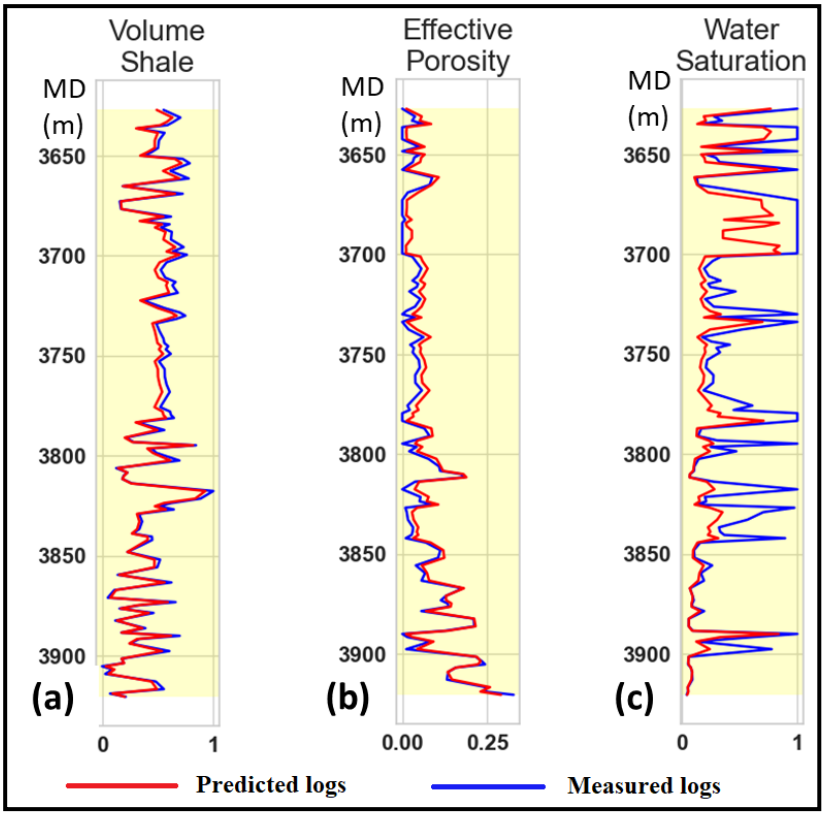



**Figure 4.31** The predicted and conventially interpreted petrophysical properties pf Mehar-02 (a). Predicted and conventially interpreted volume of shale (b). Predicted and conventially interpreted effective properties (c). Predicted and conventially interpreted $S_w$. Blue color shows the conventially interpreted curves while red color depicts predicted properties.

To assess the model's generalizability to new data and avoid overfitting, cross plots were created to evaluate the performance between the estimated petrophysical logs and predicted petrophysical logs. The plot in Figure 4.32 demonstrated that the $R^2$ score of the interpreted properties, between conventional petrophysics and *ML* approach for volume of shale, *PHIE*, and $S_w$ i.e., 0.99, 0.913 & 0.928 respectively.

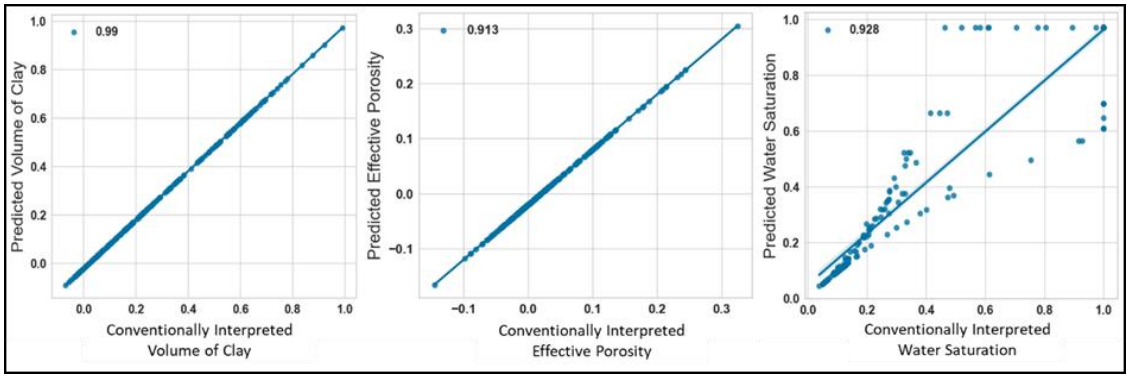



**Figure 4.32** Regression plots comparing the predicted and measured petrophysical properties and illustrate the high correlation amongst modeled and conventionally calculated properties, hence validating the results.

### 4.1.5.2 Petrophysical Properties Estimation on Blind well Zamzama-08

A similar way is adopted for predicting the various petrophysical properties at blind well location of Zamzama-08 by considering the relevant petrophysical properties of Zamzama-04 and Zamzama-05. The Figure 4.33 shows the estimation of shale volume on Zamzama-08 (blind well) through training of Shale volumetrics present in Zamzama-04 and Zamzama-05 (training well). A good match is observed at the blind well location between the shale volume estimated using conventional approaches with the *ML* technique predicted blindly.

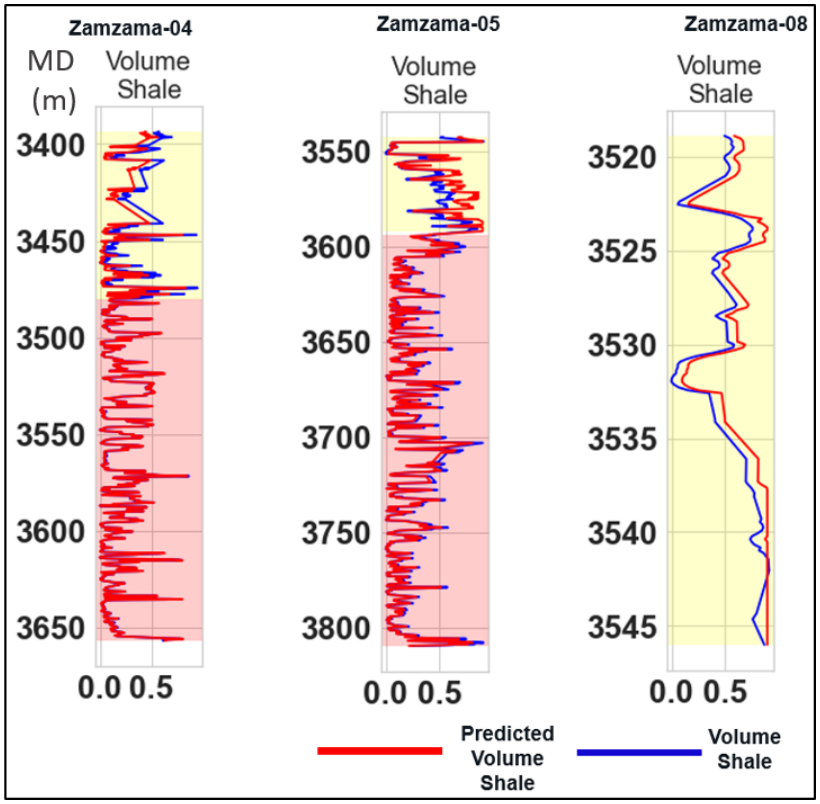



**Figure 4.33** Volume of Shale is predicted on blind well Zamzama-08 by training Zamzama-04 and Zamzama-05.

Similarly, *PHIE* and $S_w$ are modeled in blind well Zamzama-08 as depicted in Figure 4.34 and Figure 4.35. The evaluation metrics are created to identify the performance between the calculated petrophysical logs and predicted petrophysical logs (Figure 4.36). The plots show the $R^2$ score between the volume of shale and predicted volume of shale is 0.92, the $R^2$ score between the *PHIE* and predicted *PHIE* is 0.89 and the $R^2$ score between the $S_w$ and predicted $S_w$ is 0.775.

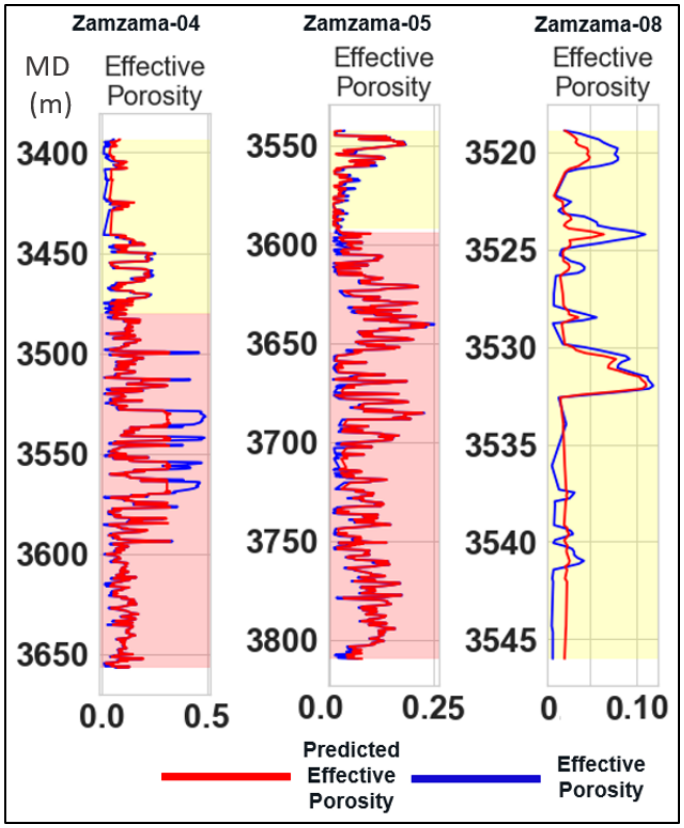



**Figure 4.34** *PHIE* is predicted on blind well Zamzama-08 by training Zamzama-04 and Zamzama-05.

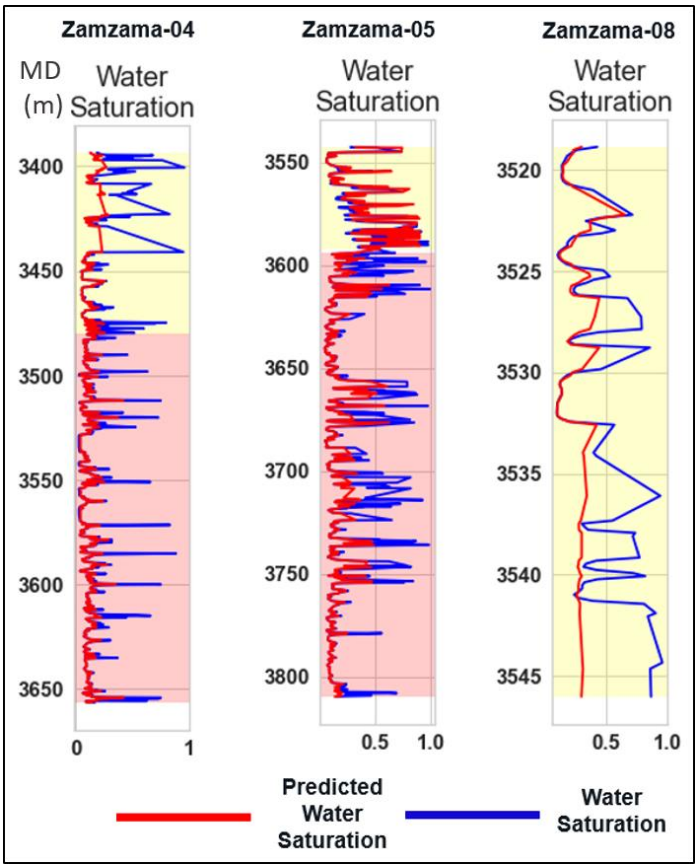



**Figure 4.35** *PHIE* is predicted on blind well Zamzama-08 by training Zamzama-04 and Zamzama-05.

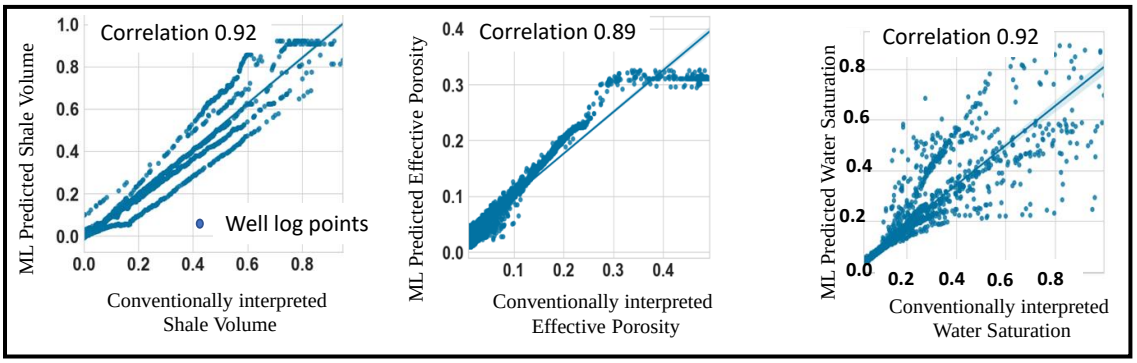



**Figure 4.36** Cross plots comparing the predicted and measured petrophysical properties, representing high correlation between them, hence validating the results.

### 4.1.6 Facies Analysis in Mehar-01

Using conventional and *ML* approaches, researchers investigated numerous methods for improving the accuracy and validation of this analysis. *ML* methods are increasingly being used in reservoir analysis. The goal was to discover and rank the best *ML* tools for estimating facies. Comparing conventional methods to *ML* techniques can improve reservoir facies analysis. The study represented the usage of techniques such as *DTC*, as well as the use of *ML* tools for predictive modeling. To assess the accuracy and performance of these algorithms in blind wells or unseen data, after 10 K-fold evaluation and statistical measurements (Figure 4.37).

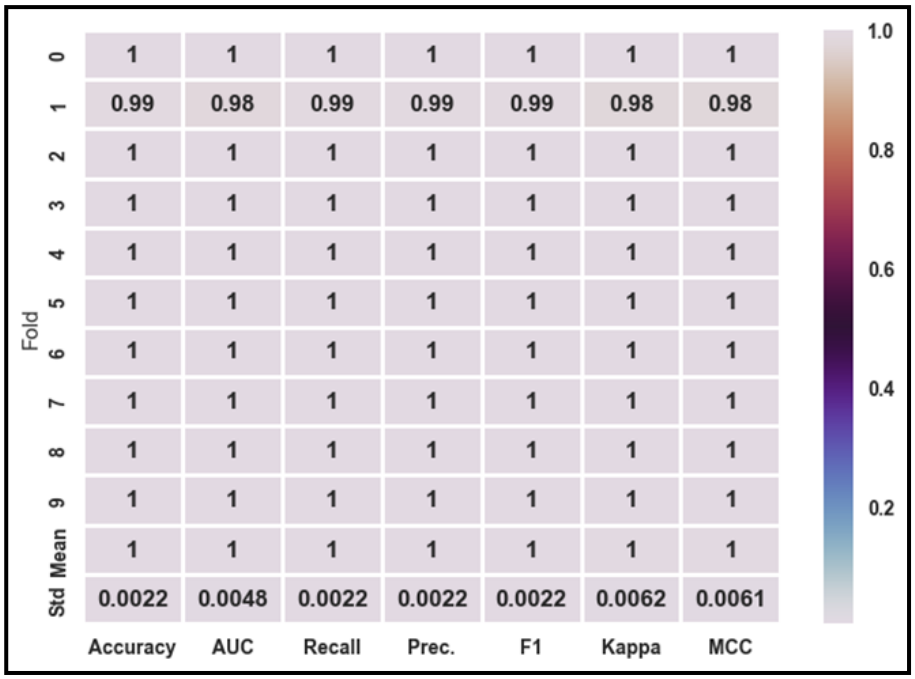



**Figure 4.37** K-fold cross validation highlight the measures at each fold for facies prediction.

*DTC* is the best algorithm for facies classification in which overall score is 1 at training well Mehar-01. The model performances are measured through Accuracy, area under curve (*AUC*), Recall, Precision, F1, Kappa, and Matthew's correlation coefficient (*MCC*) shown in the Figure 4.38.

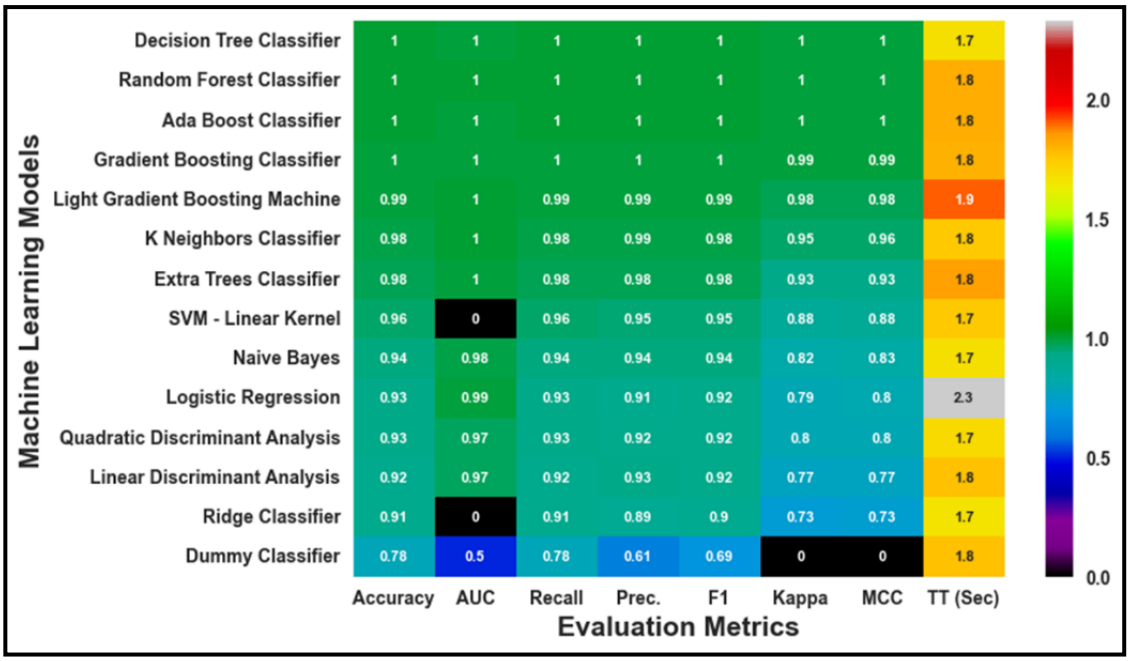



**Figure 4.38** Evaluation metrics represent the measures at each fold for facies prediction.

### 4.1.6.1 Confusion Matrix

The confusion matrix is drawn between each type of facies. Here 462 values of shale facies (class 0), 118 values of gas sand (class 1), and 12 values of wet sand (class 2) are accurately predicted at training well (Figure 4.39a). Whereas Confusion matrix shows accurately predicted 72 values of shale facies, 24 values of gas sand and 2 values of wet sand along with 100 % accuracy while only 10 values were falsely predicted for wet sand (Figure 4.39b) at blind well.

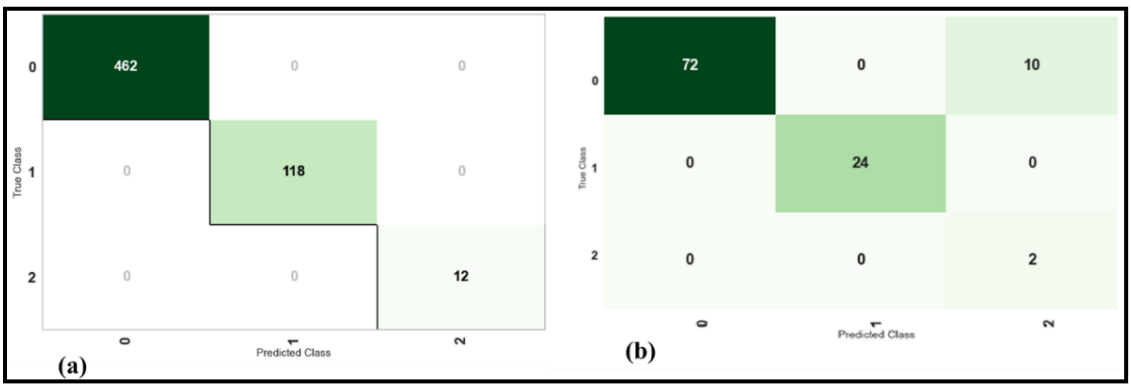



**Figure 4.39** Confusion metrics highlight the measures at each fold for facies prediction in (a) training well and (b) blind well.

*DTC* appeared as the best algorithm for predicating reservoir facies, and it has been further validating its results at the blind well Mehar-02. In Figure 4.40 is quite evident that the modeled reservoir facies are in quite align with the conventionally interpreted facies, hence validating the results.

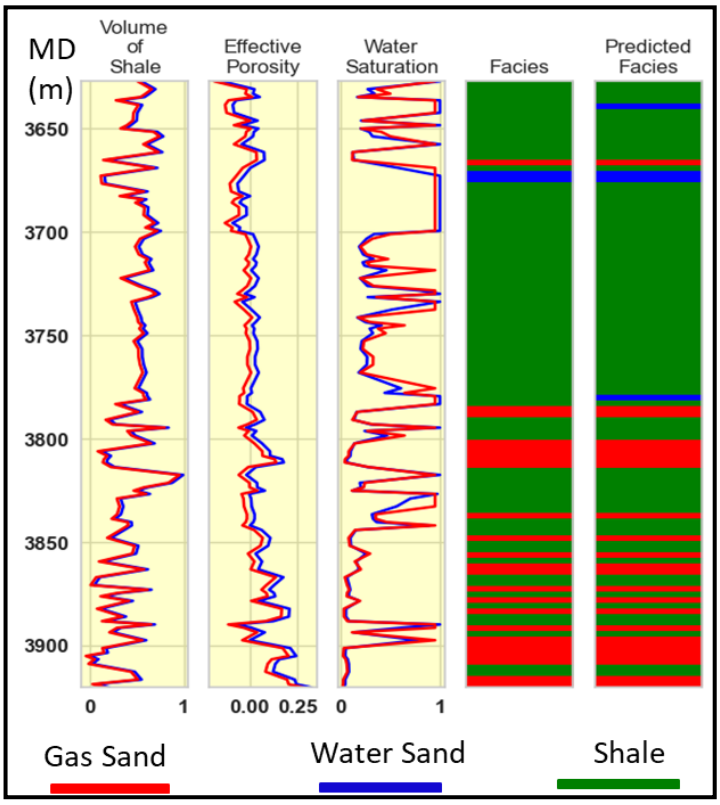



**Figure 4.40** Comparison of the facies predictions after *ML* with conventionally interpreted facies (4th Track) at the well showing a high correlation amongst the predicted results (5th Track).

Al-Mudhafar (2020) integrated *ML* and data analytics for clastic reservoir facies and discovered that LogitBoost is the most accurate algorithm, with 100% accuracy in total correct facies prediction. However, the total correct percentages for Multinom and XGBoost were 80.24% and 70.83%, respectively. Rafik and Kamel, (2017) employed a combination of *PCA*, model-based cluster analysis, and discriminant analysis to estimate permeability and porosity from well-log data. Three non-parametric techniques are investigated: alternating conditional expectations (*ACE*), generalized additive models (*GAM*), and *NNET* to predict permeability. The strategy of *ACE* strategy demonstrates superior performance as compared to the other two methods. Ahmadi et al., (2014a) conducted research to forecast the permeability and porosity of petroleum reservoirs in northern Persian Gulf oil fields using petrophysical logs. To overcome the

aforementioned issue, various *AI* approaches, such as fuzzy logic (*FL*) and *LSSVM*, were used. It is suggested that *LSSVM* and *FL* be used in conjunction with a genetic algorithm (*GA*). It has been observed that the correlation coefficient between model estimates and relevant real data is greater than 0.96 for the *GA-FL* technique and 0.97 for the *GA-LSSVM* approach and their result is more trustworthy in the case of porosity and permeability predictions. Ahmadi and Chen, (2019) evaluated various *ML* methods and suggested that the hybridized technique could predict the reservoir's petrophysical parameters with high accuracy. In the Central Indus Basin, Pakistan, few studies evaluated *ML* methods. Ali et al., (2021) used the *RFR* to forecast facies with an accuracy of 83.85%, and Ahmed et al., (2022) used a stacking method to combine the outputs of numerous models, including the *ETR* reservoir, with an accuracy of 87.23%.

In nutshell, this research highlights effectiveness of *ML* models in removing outliers from raw well logs and modeling the missing and bad well logs in an optimized and systematic ways. The *ML* has the ability to provide more accurate and reliable predictions compared to traditional techniques, making it a promising tool for the industry after testing and validation with excellent results. However, *ML*-based methods require large amounts of high-quality training data, and the accuracy of predictions depends on the quality of the training data and the complexity of the model.

For improving accuracy, efficiency, and decision-making, the use of *ML* techniques in petrophysical analysis has major implications. *ML* offers a data-driven strategy, the integration of many data sources, improved interpretation, and potential automation by utilizing cutting-edge algorithms and analyzing enormous volumes of well-structured data. These ramifications help petrophysicists comprehend subsurface formations more thoroughly, characterize reservoirs more accurately, and make wiser decisions.

## 4.2 Seismic Interpretation

The entire *3D* seismic cube has been interpreted, including the marking, and picking of the Ranikot/Khadro Formation, Pab Formation, and Fort Munro horizons, as well as fault marking. For the stratigraphical variation of formations throughout the field, an arbitrary line is displayed in the seismic section. Figure 4.41 represents the arbitrary line that is passed from all the well locations of the Zamzama Gas Field along with their

*GR* curve and with the interpreted horizons of the Ranikot/Khadro Formation, Pab Formation, and Fort Munro Formations.

Figure 4.42 shows an inline 401 in which all the three horizons of Ranikot/Khadro Formation, Fort Munro, and Pab Formation, along with the major faults, have been illustrated. On the east side of this data cube, there is a significant north-south trending reverse fault that was considerably younger and disturbs formations up to Cretaceous age. The one major thrust fault shown in black and two minor thrusts that terminated in the major fault have been marked. On the east side, there is another major thrust fault that has been shown in light blue. The direction of the line is west to east, while the direction of the major fault is dipping from northeast to southwest.

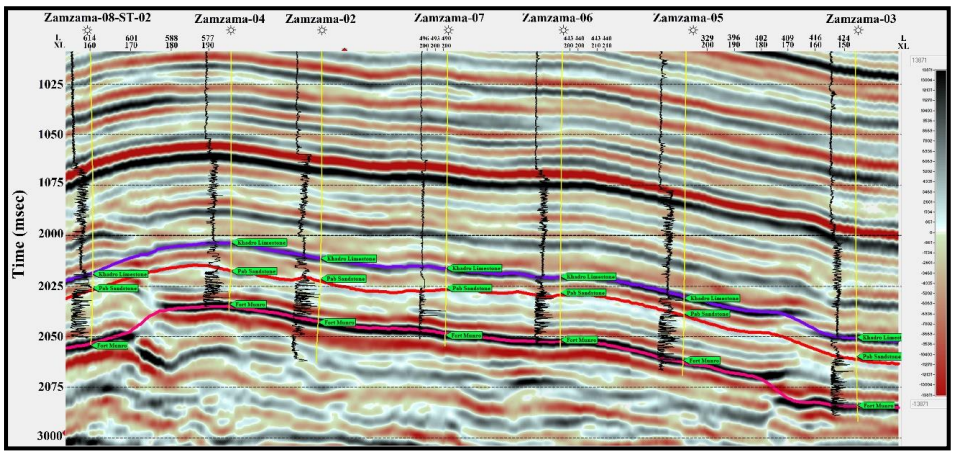



**Figure 4.41** Display of arbitrary Seismic section along with showing of all well of Zamzama Gas Field.

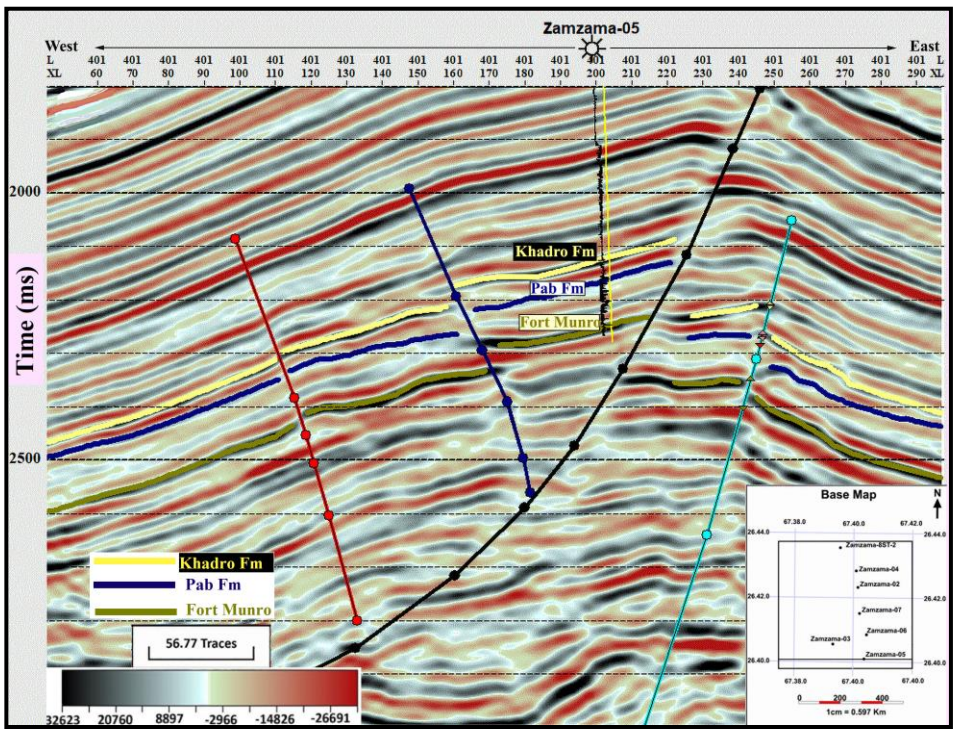



**Figure 4.42** Seismic section along with synthetic traces based on Zamzama-05 and inline is 401.

The Zamzama Field is a north-south oriented, elongated thrust-related anticline that slopes gently towards the west. Based on vintage *3D* seismic data interpretation, the field structure is divided into four compartments, such as Zamzama Far North, Core North, Core South, and East, with varying degrees of pressure compartmentalization. The major thrust bounds the anticline to the east, separating the hanging wall from the footwall structure. The thrust was previously thought to extend for around 40 km, dividing the main part of the field (hanging wall structure) from a less developed footwall anticline. Both the hanging wall and footwall structures contain two separate culminations. This interpretation suggests that the main bounding thrust connects several earlier thrust segments and separates the Zamzama-08 lateral ramp from the footwall structure. Smaller thrusts on the crestal hanging wall are considered to define individual compartments.

The Zamzama-08 well was initially drilled to evaluate the sub-thrust structure's potential. However, a pressure survey conducted in 2003 showed that the well was already depleted, indicating partial communication with the crestal hanging wall wells (Jackson et al., 2004). This suggested that the footwall was partially in pressure communication with the hanging wall and that the main thrust was leaky.

The new subsurface imaging has drastically altered the structural interpretation of the Zamzama Field. It suggests that the main thrusts defining the field are much less continuous than previously believed and generally do not connect directly. The main thrust is split into two segments by a significant lateral ramp located in the Zamzama East-1 well site, which connects the hanging wall with the footwall. An arbitrary seismic line from the crestal part of the Core South region of the field northwards along the Zamzama relay ramp displays continuous reflectors, and pressure data confirms structural continuity (Khan et al., 2022). The small en-echelon thrusts in the crestal hanging wall are short (<5km) and have an oblique-slip component to their thrusting. In the previous interpretation, these thrusts connected with the main thrust at their northern end but still allowed for tortuous connectivity across the Pab Formation reservoir via twisted pathways around the end of the crestal thrusts across the Core South region of the field.

### 4.2.1 Computation of Time Contour Maps

After the marking of faults and horizons of interested formations, the next step is to compute the time contour maps. Basically, fault polygons are usually digitized after plotting our interpretation on the base map, and then time contour maps are prepared, as seismic data is typically obtained in the time-domain. The time contour maps of Ranikot/Khadro Formation are shown in Figure 4.43. There are a couple of major thrust faults, depicted by the black color, dipping toward the western side, along with the minor faults that dip oppositely, resultant of a compressional regime and enclosing the hanging wall. The contour interval of the time map is kept at 0.025 sec. All the wells penetrated in the core area where the major faults separated the hanging wall from the footwall.

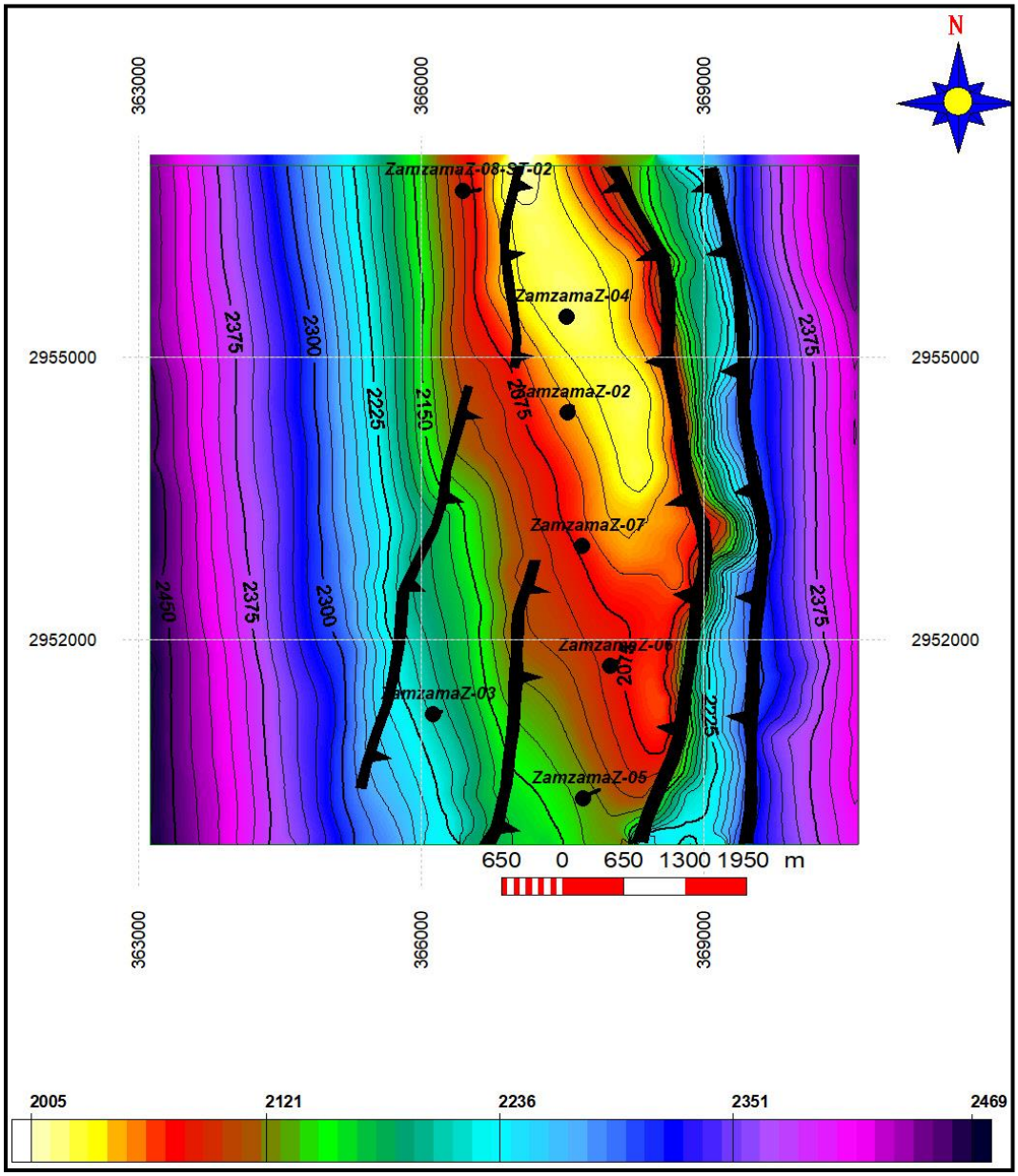



**Figure 4.43** Two-way travel time map of Ranikot/Khadro Formation along with major and minor faults. All wells are drilled in the hanging wall of the anticline.

Time contour map on Figure 4.43 shows that low time values are present where the wells penetrated in the crestal portion of the anticline. These low time values depict the shallow portion of anticline i.e., the hinge zone while the time contour values increase towards east-west direction (limbs of anticline). The increase in time values is observed in western side where a gentle dipping of the anticlinal is observed.

### 4.2.2 Computation of Depth Contour Maps

After making the time contour map the next step is to know the exact velocity of the formations. The exact velocity of formations is picked from check shot data that is basically a time-depth relationship. Velocity of interested formations Ranikot/Kadro Formation, Pab Formation and Mughalkot were assessed through their time and depth relationships from well ties.

After knowing the velocity next step is generating depth contouring map. The depth is calculated with the help of simple formula ($S = VT/2$), where $S$ = depth of the surface, $V$ is the velocity in m/sec, while $T$ signifies the 2-way travel time in ms.

The structure of the Ranikot/Khadro Formation is represented in true vertical depth sub-sea (*TVDSS*) and depicts the range, i.e., 3305 to 3660 m, in the main thrusted region of the core area (Figure 4.44). The arrangement of faults surrounding the hinge zone delineated the intensity of the tectonic activity in a compressional way and created major thrusts along with minor faults. The anticlinal structure elongated in the direction from north toward south and terminated in the direction of east-west. The depth counter map of the Ranikot/Khadro Formation is prepared with a contour interval of 25 m. The shallowest part (crestal zone) of the anticline, i.e., zones where wells were drilled, is depicted with low values that gradually increase in the east-west direction.

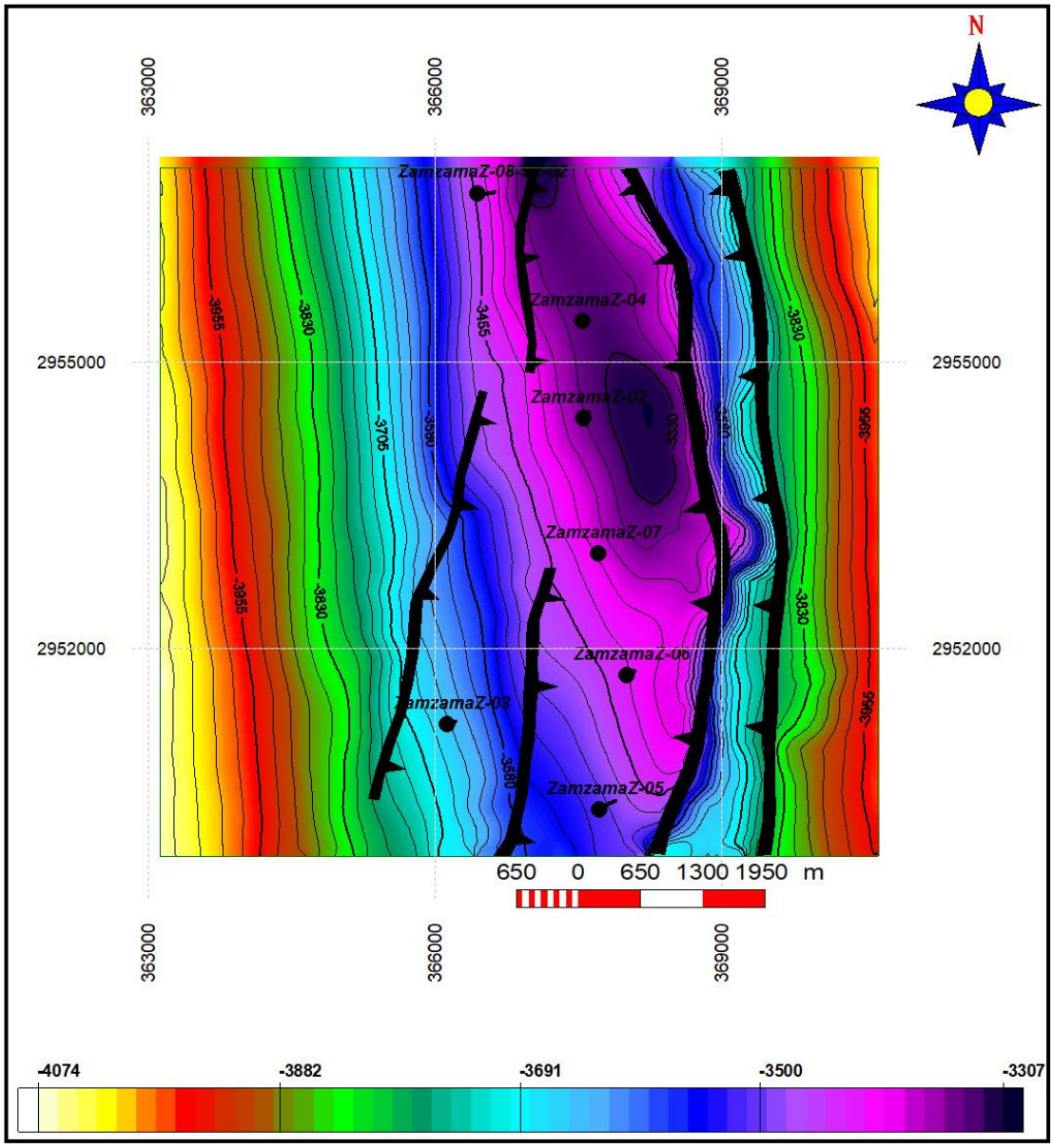



**Figure 4.44** Depth contour map of Ranikot/Khadro Formation clarify the wells penetration at the crest of anticline with the illustration of fault polygons.

## 4.3 Seismic Attribute

The seismic section of arbitrary line covering all the available wells i.e., Zamzama-02, 03, 04, 05, 06, 07, and 08 is displayed in Figure 4.45. The seismic data resolution is coarser regarding the gas-bearing sand facies within Ranikot/Khadro Formation along with the obscure reflection at various levels.

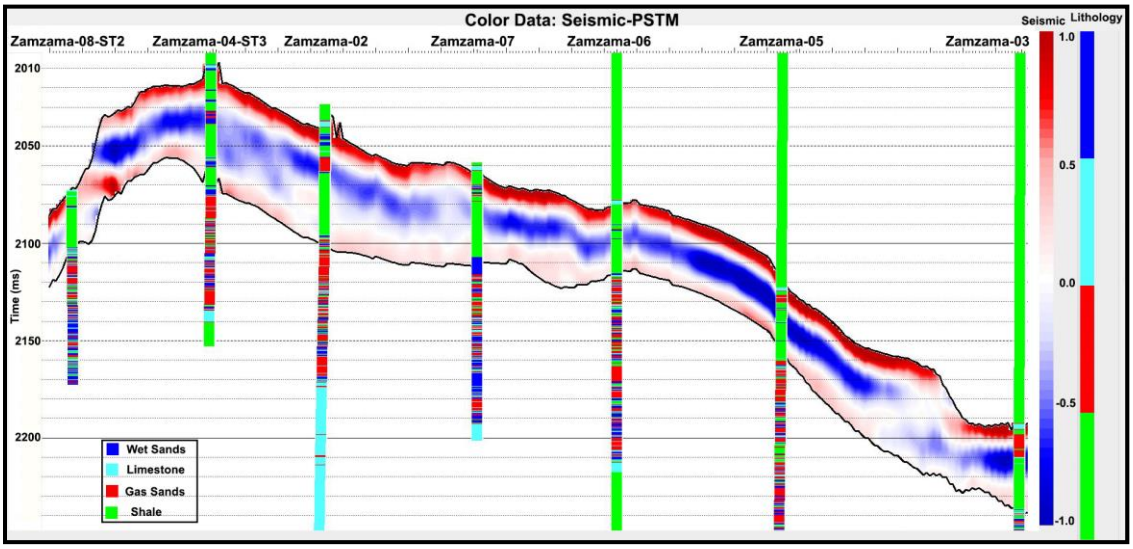



**Figure 4.45** Field seismic data shows coarser, smudge and obscure reflection at certain levels.

The extracted average energy attribute at Ranikot/Khadro Formation is represented in Figure 4.46. The attribute is coarser, and the gas sands are below seismic tuning thickness; thereby, the gas sands are not resolved. Such limitations restricted the distinction of various lithofacies, i.e., gas sands, wet sands, and shale.

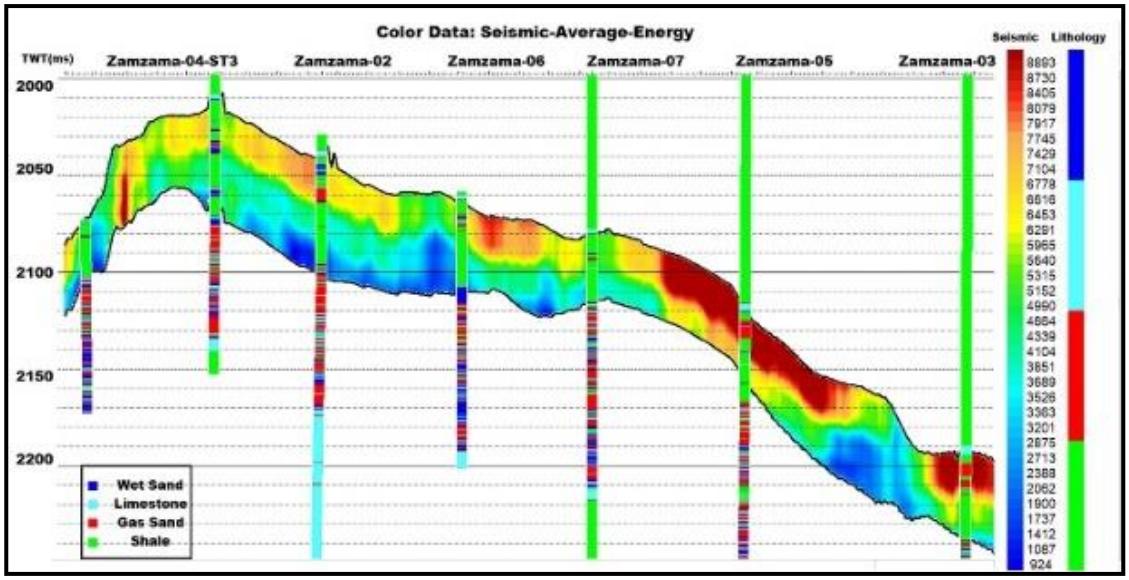



**Figure 4.46** Average energy seismic attribute on seismic data set with blocky & coarser resolution.

Similar to the Average energy attribute, RMS-amplitude extracted property doesn't clarify the gas sand facies, and the attribute is blocky with no obvious identification of thin potential sands Figure 4.47.

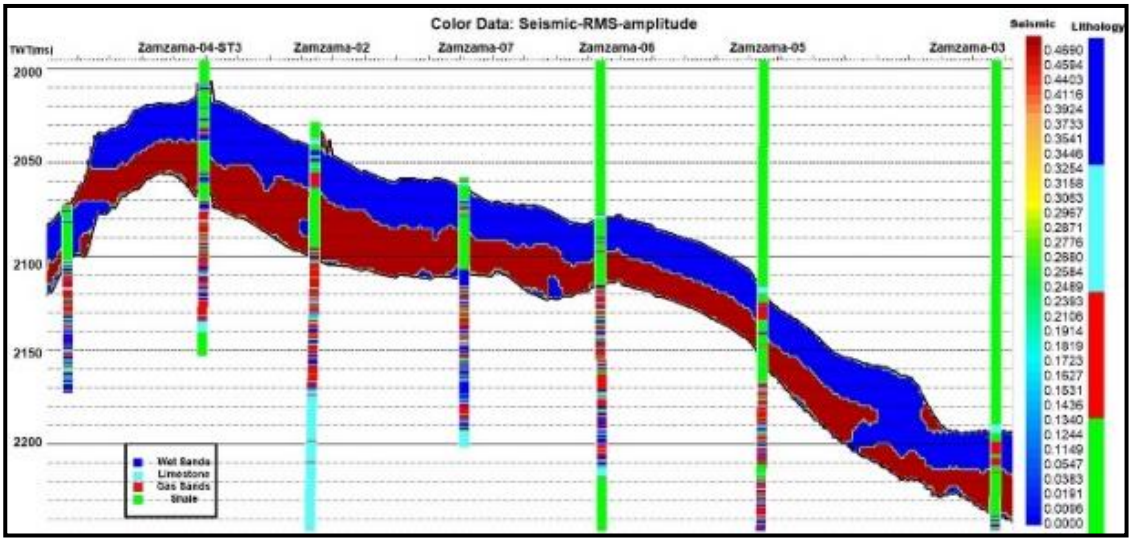



**Figure 4.47** *RMS* seismic attribute on seismic data set unable to distinguish fine reservoir layering.

The normalized seismic amplitude within the Ranikot/Khadro Formation is still unable to recognize the presence of plausible sands as the lower frequency content of the seismic data is not capable of resolving it (Figure 4.48).

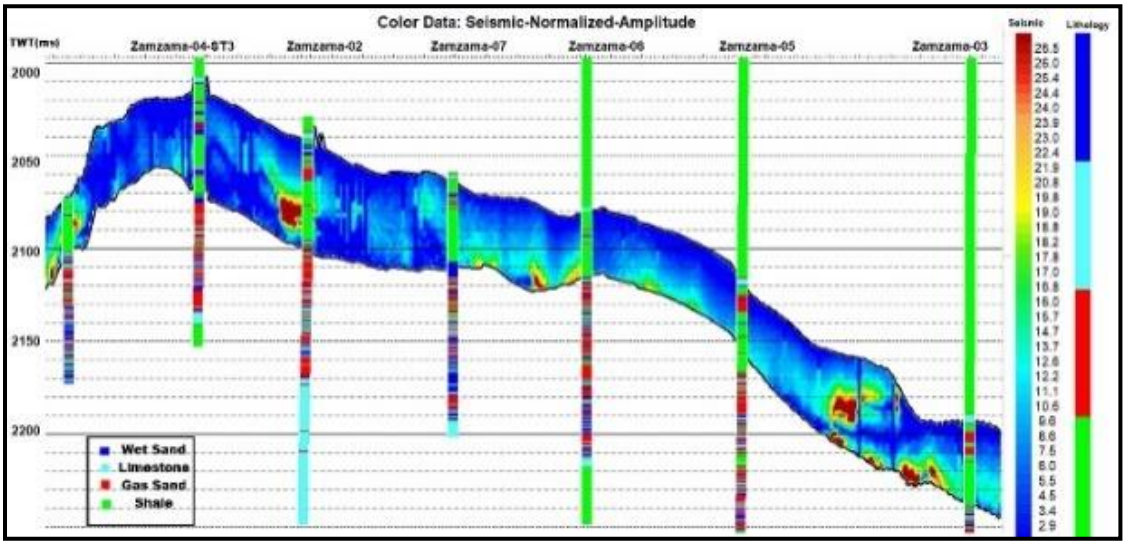



**Figure 4.48** Normalized seismic amplitude attribute extracted on field seismic data set.

Overall, the Ranikot/Khadro Formation is thin enough with the sand's intervals of few meters therefore it is not visible in conventional seismic data interpretation i.e.,

sweetness (Figure 4.49). For enhanced resolution the *ML* algorithm is developed that is comprehensively elaborated in discussion section.

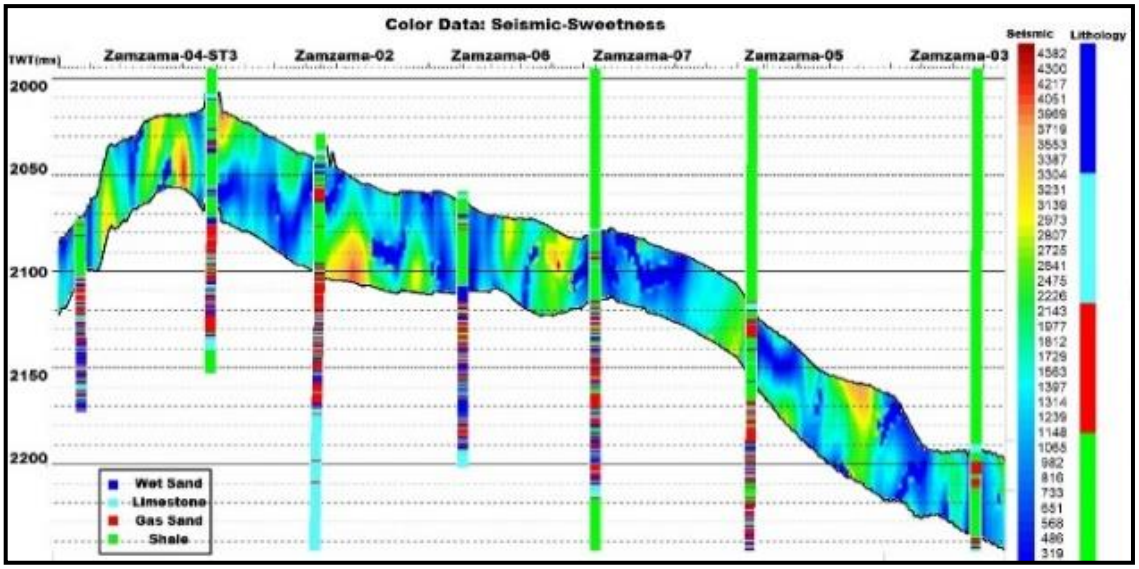



**Figure 4.49** Sweetness seismic attribute on seismic data set showing random results.

## 4.4 *RPM* and Seismic Inversion

The RPM plays a vital role in the identification and prediction of elastic behavior of reservoir properties. It generates a relationship between petrophysical and elastic attributes and assess the reservoir characteristics in the petro-elastic domain. The approximated elastic properties are employed in the inverted elastic properties so that a more reliable picture is attained that better understand the pay reservoir facies.

### 4.4.1 *ML* Based *RPM*

For the successful prediction of the missing logs, i.e., S-wave and optimization of density logs, in order to compensate for bad borehole conditions, a petro-elastic definition is reliably established over the cross plot between elastic properties, i.e., $Z_p$ plotted along the x-axis and $V_p/V_s$ ratio on the y-axis. The plotted elastic properties are colored with identified lithofacies that were recognized through conclusive cut-offs on petrophysical properties like $S_W$ and Vclay ($V_{CL}$) (Hussain et al., 2022; Durrani et al., 2022; Khan et al., 2022; Shakir et al., 2021) (Figure 4.50). Petrophysical cutoff ranges for all reservoir lithofacies can be defined using detailed petrophysical analysis. Whereas hydrocarbon bearing sands with low values of $S_W \leq 0.45$ and $V_{CL} \leq 0.30$ reflect the corresponding elastic ranges of $Z_p$ (7500-10500 g/cc*m/s) and $V_p/V_s$ ratio (1.45-1.6).

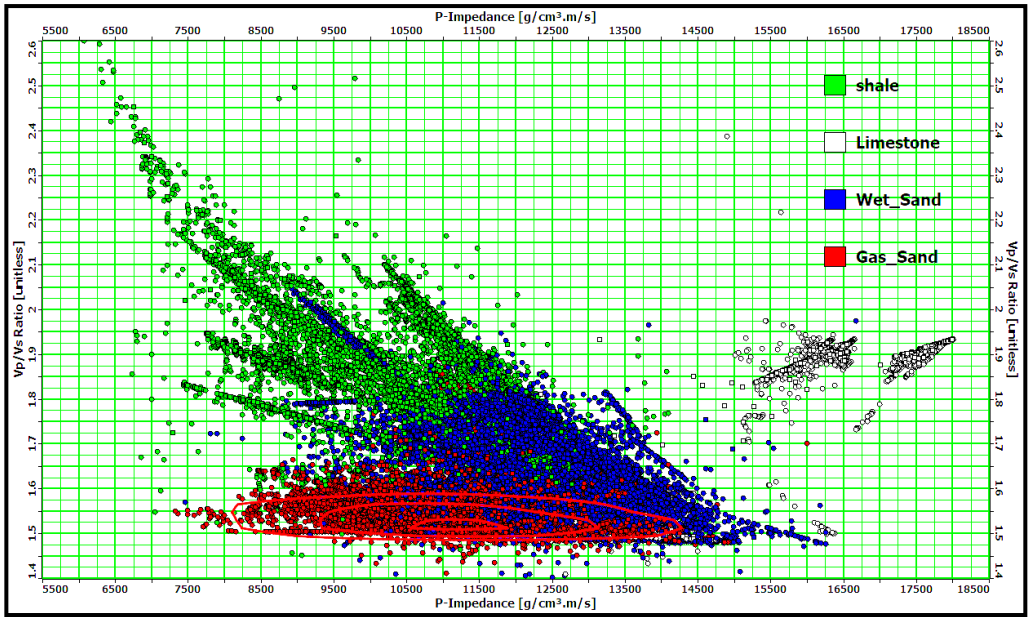



**Figure 4.50** *RPM* based cross-plot between elastic attributes such as $Z_p$ along the x-axis and $V_p/V_s$ ratio on the y-axis, colored with lithofacies (gas-sand, wet-sand, limestone, and shale) present within Ranikot/Khadro Formation, depicts their specific elastic ranges.

*SVM*, *DTR*, *RF*, and *GBR ML* algorithms were used to predict the missing logs and compensate for the rock physics modelling in an efficient and effective manner in the current research work. Each algorithm showed good results with the measured logs. However, *GBR* shows consistently good results in all carried out predictions of the $Z_p$ and $Z_s$. The $R^2$ score against each prediction made using the corresponding algorithm depicts the quality of the match attained. The higher the $R^2$ score reflects the good correlation with the measured log.

#### 4.4.1.1 One Dimensional (*1D*) elastic Properties Prediction using *ML*

By incorporating the *ML* workflow of *SVM*, *DTR*, *RF*, and *GBR*, the predictions were carried out for each elastic property by training the algorithm with the available dataset. Results are shown individually for each well, reflecting each elastic property predicted using different algorithms to evaluate the $R^2$ score.

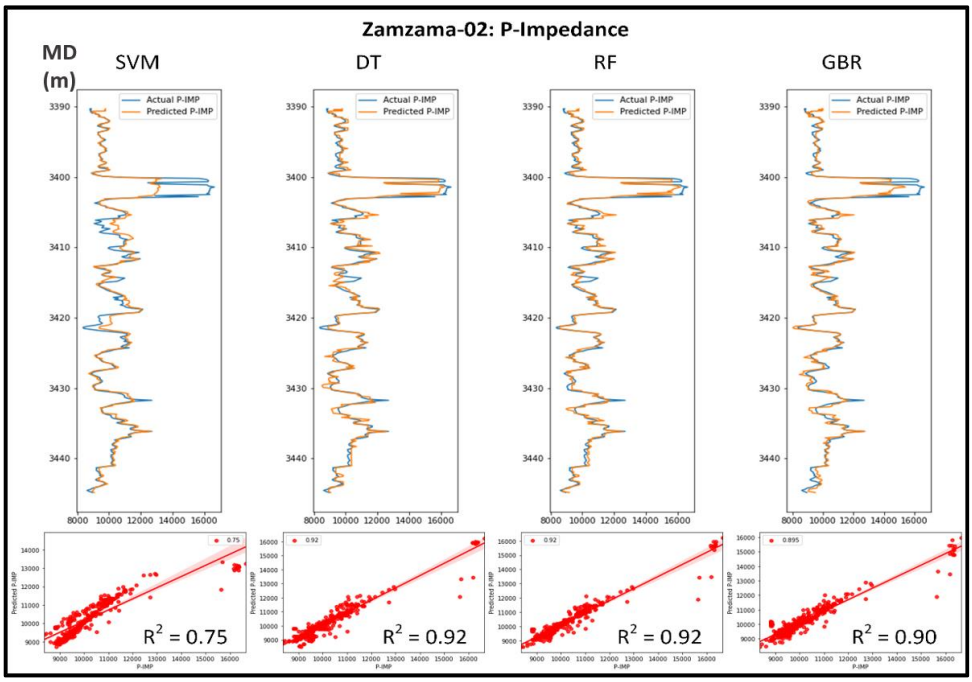



**Figure 4.51** Prediction of $Z_p$ at well location of Zamzama-02 using different *ML* algorithms by training from Zamzama-04 & 07. Blue is measured while orange is the predicted log. High $R^2$ score depicts the accuracy of between measured and predicted logs.

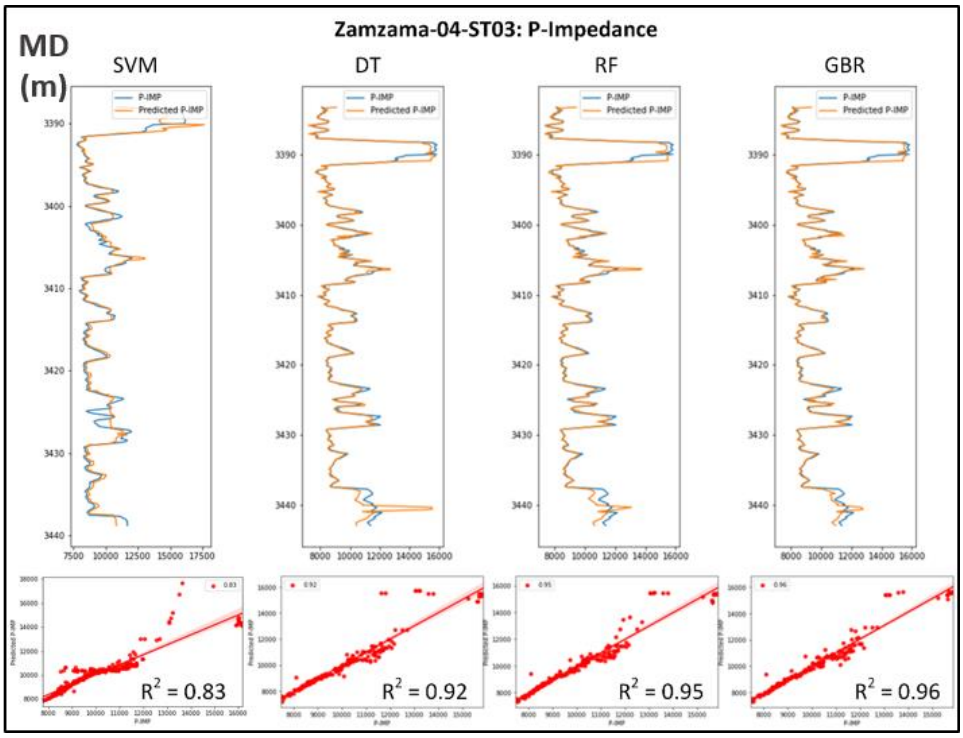



**Figure 4.52** Prediction of $Z_p$ at Zamzama-04-ST3 by employing different *ML* algorithms trained from Zamzama-02 & Zamzama-05. Blue is measured while orange is the predicted log. A decent match, i.e., modeled (orange) and measured (blue) logs, is attained.

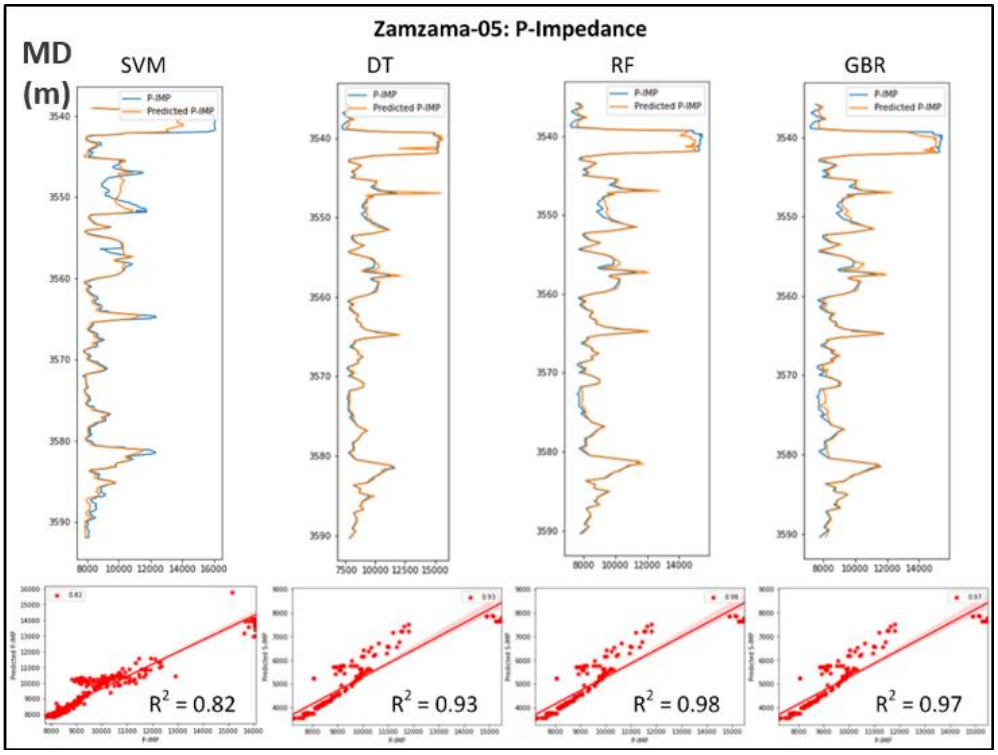



**Figure 4.53** Prediction of $Z_p$ at Zamzama-05 trained from Zamzama-02 & Zamzama-04. A reliable trend is accomplished amongst measured (blue) and predicted (orange) logs.

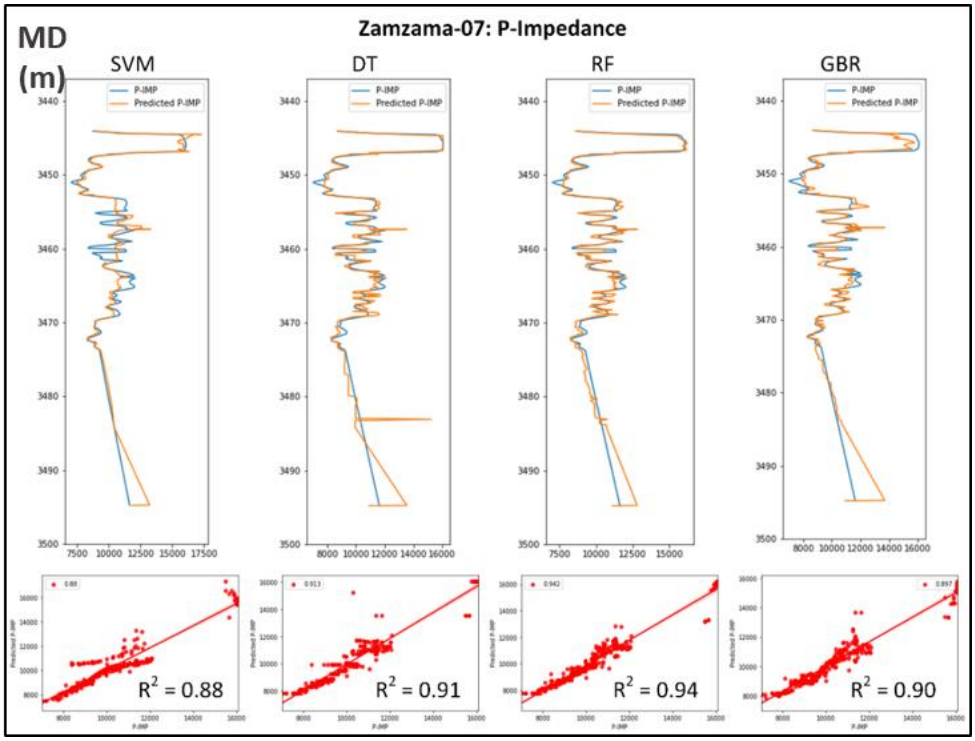



**Figure 4.54** Prediction of $Z_p$ at well location of Zamzama-07 trained from Zamzama-02, Zamzama-04-ST3 & Zamzama-05. A consistent match is observed for *DTR*, *RF*, and *GBR* with above 0.9 $R^2$ score.

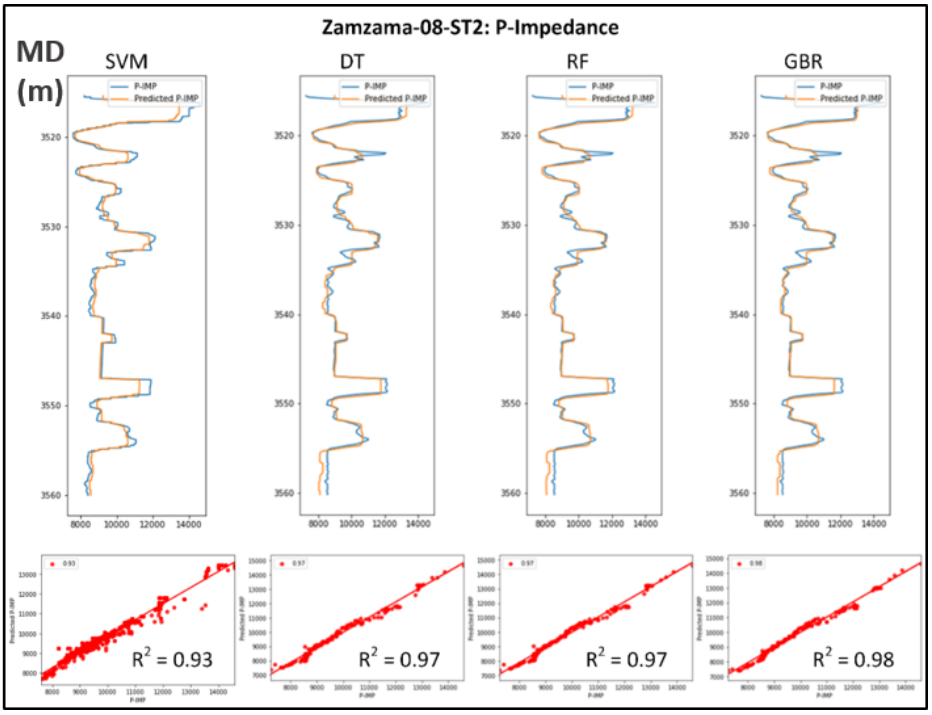



**Figure 4.55** Prediction of $Z_p$ at well location of Zamzama-08-ST2 training from Zamzama-05, Zamzama-06 & Zamzama-07. *GBR* shows a very good correlation along with trend matching between measured (blue) and predicted log (orange).

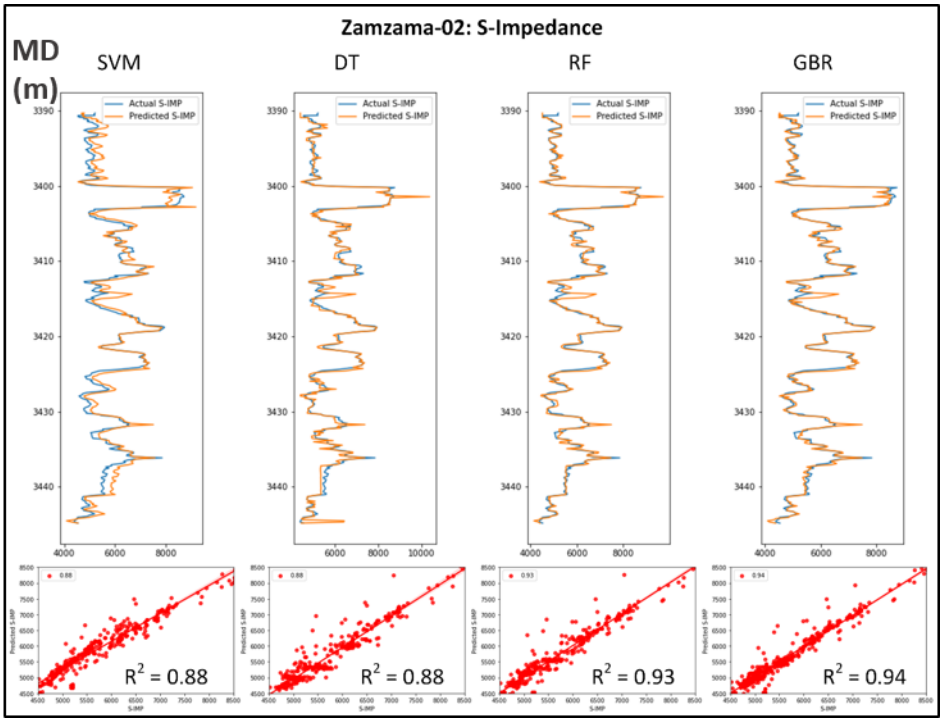



**Figure 4.56** Prediction of $Z_s$ Zamzama-02 using similar *ML* algorithms of $Z_p$ and trained from Zamzama-04-ST3, Zamzama-06 & Zamzama-07. *GBR* showed the highest correlation of measured (blue) and predicted (orange) log.

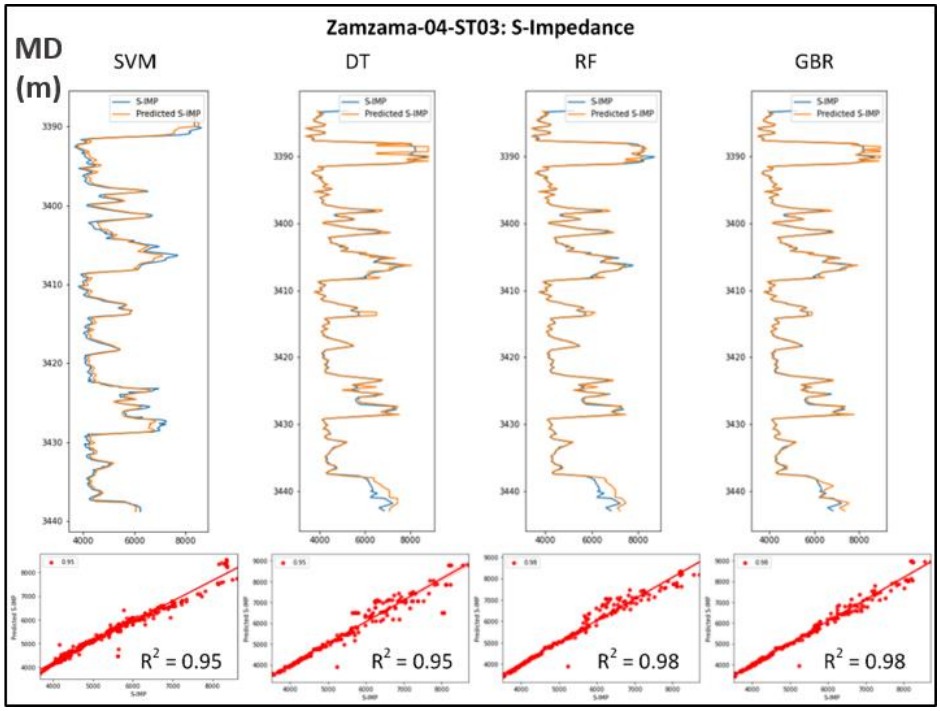



**Figure 4.57** Prediction of $Z_s$ Zamzama-04-ST3 using *ML* algorithms trained from Zamzama-02 & Zamzama-05. High correlation between measured (blue) and predicted (orange) logs.

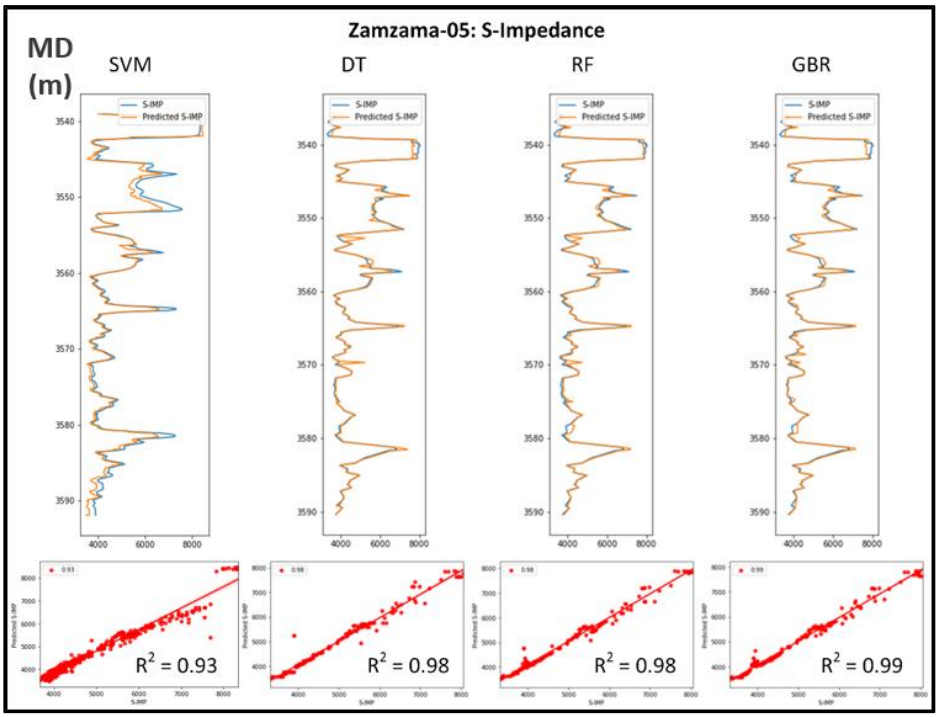



**Figure 4.58** Prediction of $Z_s$ Zamzama-05 using *ML* algorithms trained from Zamzama-02 & Zamzama-04. *GBR* showed high correlation between measured (blue) and modeled (orange) logs.

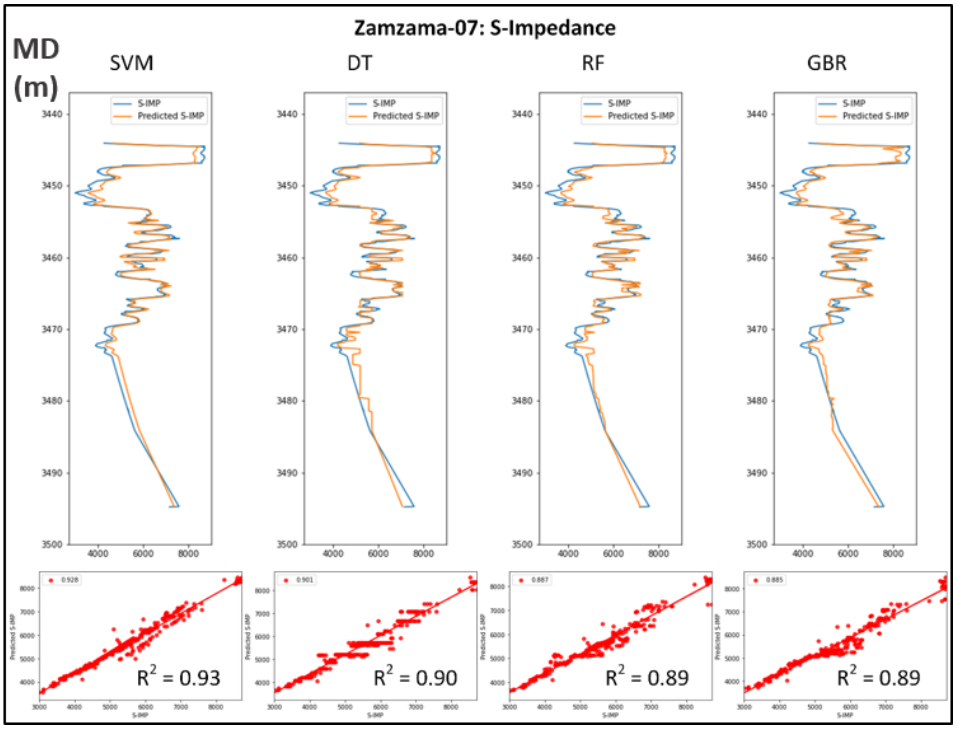



**Figure 4.59** Prediction of $Z_s$ Zamzama-07 using *ML* algorithms trained from Zamzama-02, Zamzama-04-ST3 & Zamzama-05. Measured (blue) and predicted (orange) showed good correlation.

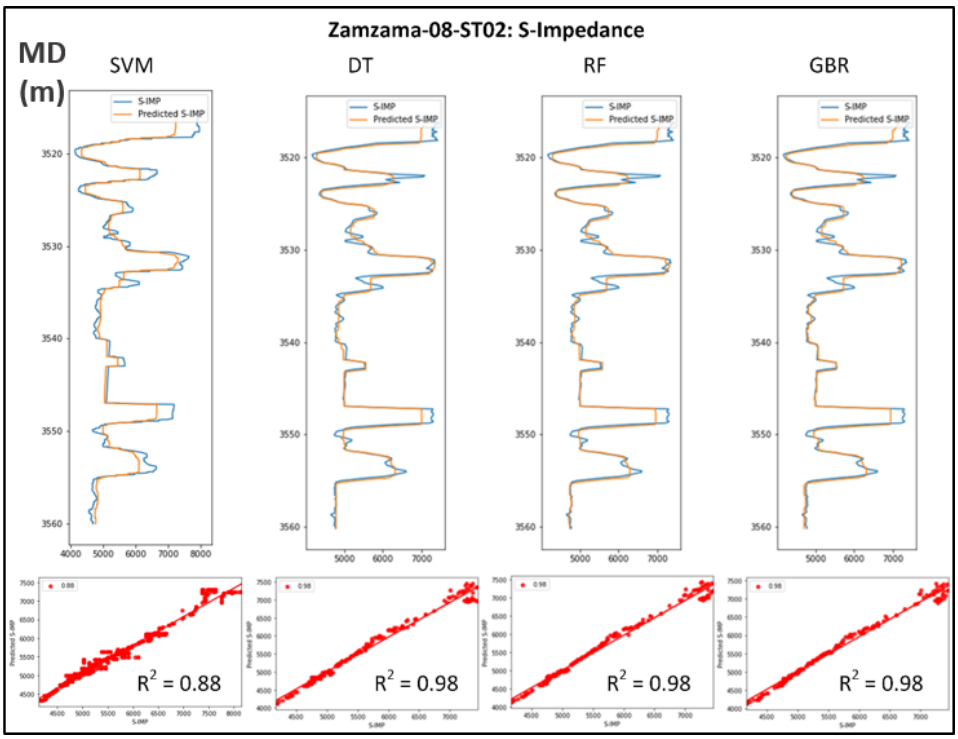



**Figure 4.60** Prediction of $Z_s$ well location Zamzama-08-ST2 using *ML* algorithms trained from Zamzama-05, Zamzama-06 & Zamzama-07. Measured (blue) and predicted (orange) logs showed reliable correlation.

#### 4.4.1.2 *ML* Algorithms Accuracy Measures

Based on the highest correlation obtained against the largest number of wells, the *GBR* was chosen as the final algorithm to characterize the thin sands of the Zamzama Gas Field (Table 4.9). Hence, the properties including $Z_p$, $Z_s$, $V_p/V_s$ ratio, *PHIE*, $V_{CL}$, and $S_W$ were trained and made comparison with actual values as follows (Figure 4.61).

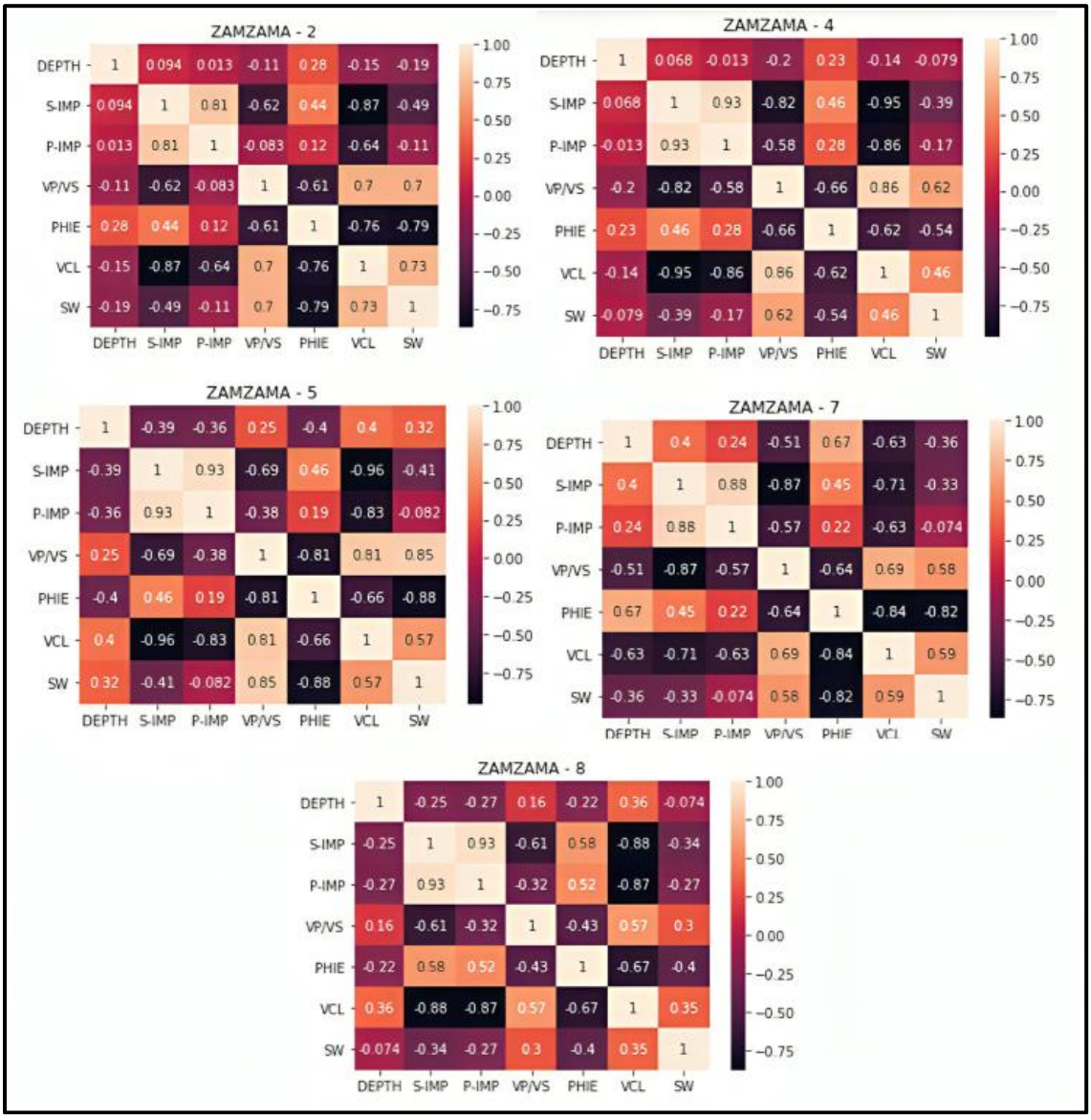



**Figure 4.61** Correlation matrix showing the actual vs predicted logs against key wells used for training the *ML* algorithm.

1. Zamzama-02 is trained using Zamzama-04 and Zamzama-05
2. Zamzama-04-ST3 is trained using Zamzama-05 and Zamzama-02
3. Zamzama-05 was trained using Zamzama-02 and Zamzama-04

4. Zamzama-07 is trained using Zamzama-02, Zamzama-04, and Zamzama-05
5. Zamzama-08-ST2 is trained using Zamzama-05, Zamzama-06, and Zamzama-07

An initial correlation matrix is observed to evaluate the variations in the plot. Results from Chicco et al. (2021) showed that the correlation matrix or coefficient of determination ($R^2$), which does not suffer from the interpretability issues of *MSE*, *RMSE*, *MAE*, and Mean Absolute Percentage Error *(MAPE)*, is more accurate and informative. Each log type is added in order to estimate the good quality match; hence the diagonal of each plot depicts the correlation match as shown in Figure 4.61. So, the correlation matrix plot is generated by evaluating actual versus predicted logs using *ML*. Each log type is added in order to estimate the good quality match; hence the diagonal of each plot depicts the correlation match.

**Table 4.9:** Different *ML* algorithms used to get the best correlation in order to characterize the thin sands of Zamzama Ranikot/Khadro Formation secondary reservoir. Overall, the *GBR* provided the best correlation values at each well.

| Well | Elastic Property | *ML* Algorithms $R^2$ Score | | | |
|---|---|---|---|---|---|
| | | ***SVM*** | ***DTR*** | ***RF*** | ***GBR*** |
| **Zamzama-02** | $Z_p$ | 0.75 | 0.92 | 0.92 | 0.90 |
| | $Z_s$ | 0.88 | 0.88 | 0.93 | 0.94 |
| **Zamzama-04-ST3** | $Z_p$ | 0.83 | 0.92 | 0.95 | 0.96 |
| | $Z_s$ | 0.95 | 0.95 | 0.98 | 0.98 |
| **Zamzama-05** | $Z_p$ | 0.82 | 0.93 | 0.98 | 0.97 |
| | $Z_s$ | 0.93 | 0.98 | 0.98 | 0.99 |
| **Zamzama-07** | $Z_p$ | 0.88 | 0.91 | 0.89 | 0.90 |
| | $Z_s$ | 0.93 | 0.90 | 0.89 | 0.89 |
| **Zamzama-08-ST2** | $Z_p$ | 0.93 | 0.97 | 0.97 | 0.98 |
| | $Z_s$ | 0.88 | 0.98 | 0.98 | 0.98 |

Therefore, it was concluded that the overall Gradient Booster Regressor provided the best prediction against each elastic property log at every given well location.

The thin secondary reservoir sands of the Ranikot/Khadro Formation (Khan et al., 2022) exhibit extreme heterogeneous behavior within the Zamzama Gas Field. The thin heterogeneous sands are assessed with by synthetic wedge model as detailed elaborated and depicted in Figure 4.76. The petrophysical analysis depicted a 7 m thick gas sand in the Zamzama-02 well at Ranikot/Khadro Formation level (Figure 4.76a), while the wedge model reflected the seismic resolving power of bed depicted thickness of around 60-65m. The major components employed into the synthetic wedge model including frequency,

velocity, and wavelet from the reservoir zone Figure 4.76b, c and d). The thin sands are below the seismic resolution limit, however seismic inversion is employed to observe the overall behavior of elastic properties. These thin sands (4-7 m thick) were successfully characterized across the study area with the implementation of advanced *ML* algorithms and techniques. Therefore, despite the excellent results obtained by all algorithms, *GBR* was chosen for final characterization based on the comparatively maximum correlation attained.

### 4.4.2 *AVO* Modeling and Sand Class Identification

By utilizing $V_p$, $V_s$, and $\rho$ (measured logs) and statistical wavelet, the *AVO* modeling (process of forward modeling) is performed using the Aki-Richards approximation (Aki, 1980) to calculate the reflection coefficients $R_p(\theta)$ as a function of incident angles according to the following equations 4.1-4.2.

$$\boldsymbol{R_p(\theta) = A\frac{\Delta V_P}{V_P} + B\frac{\Delta \rho}{\rho} + C\frac{\Delta V_s}{V_s}} \qquad (4.1)$$

Where,

$$\boldsymbol{A = \frac{1}{2cos^2\theta};\ B = 0.5 - \left[2\left(\frac{V_s}{V_p}\right)^2 sin^2\theta\right];\ C = -4\left(\frac{V_s}{V_p}\right)^2 sin^2\theta} \quad (4.2)$$

$A$ = Intercept, $B$ = gradient, $C$ = curvature

The output modeled *AVO* synthetic traces with coverage of angle ranges of about 30˚ is depicted in Figure 4.62 along with key logs (*GR*, *P-wave*, *RHOB*, and *S-wave*) and identified lithofacies.

The lithofacies identified two bodies of sands while Sand-01 and Sand-01bottom is selected for *AVO* intercept gradient analysis to assess the class of sand.

The *AVO* characteristics of seismic reflections from the interface between shales and gas sands is initially studied by (Rutherford and Williams, 1989). The Event 1 is placed at the top of the identified sand while Event 2 is kept at sand base (Figure 4.63). The *AVO* responses of the events are assessed at the amplitude vs. angle plot.

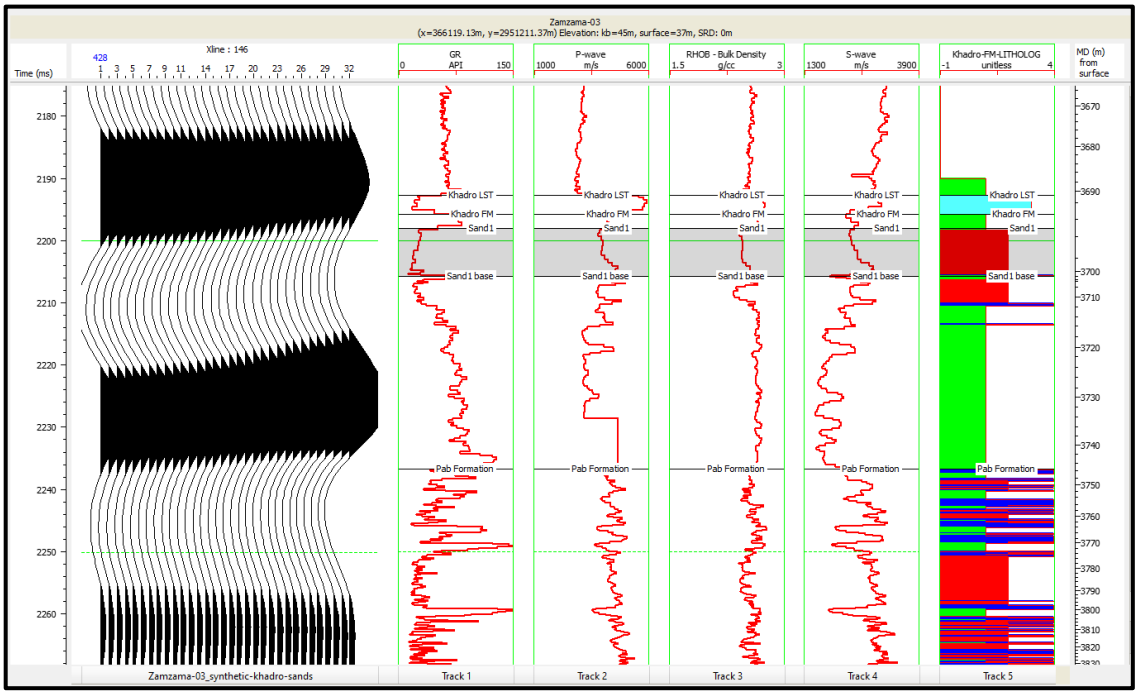



**Figure 4.62** The modeled *AVO* synthetic traces along with key well logs employed for *AVO* modeling and reservoir characterization while lithofacies log identified two zones of gas sands in the Ranikot/Khadro Formation.

The analysis of amplitude variations with angles is industrial standard for sand classification scheme. There are four sands classes identified using *AVO* analysis as follows:

Class 1: The impedance of sandstone is greater than that of shale. The interface between shale and this type of sandstone will provide a high *RC* and a positive zero offset, but the magnitude of the amplitude will decrease in order to offset.

Class 2: Sandstone has about identical *AI* with its cover (seal rock) and an amplitude that increases proportionately to offset. Based on the *RC* at zero offset, class 2 sandstone is separated into two classes: class 2a and class 2b.

Class 3: Sandstone has lower *AI* than its cover.

Class 4: Sandstone has negative *RC* at zero offset and lower impedance with amplitude that is decreasing against the offset. There is a change in polarity at a certain angle and then amplitude will be increasing proportionally to the offset.

According to the observed amplitude responses the top of the sand depicted the Class 1.

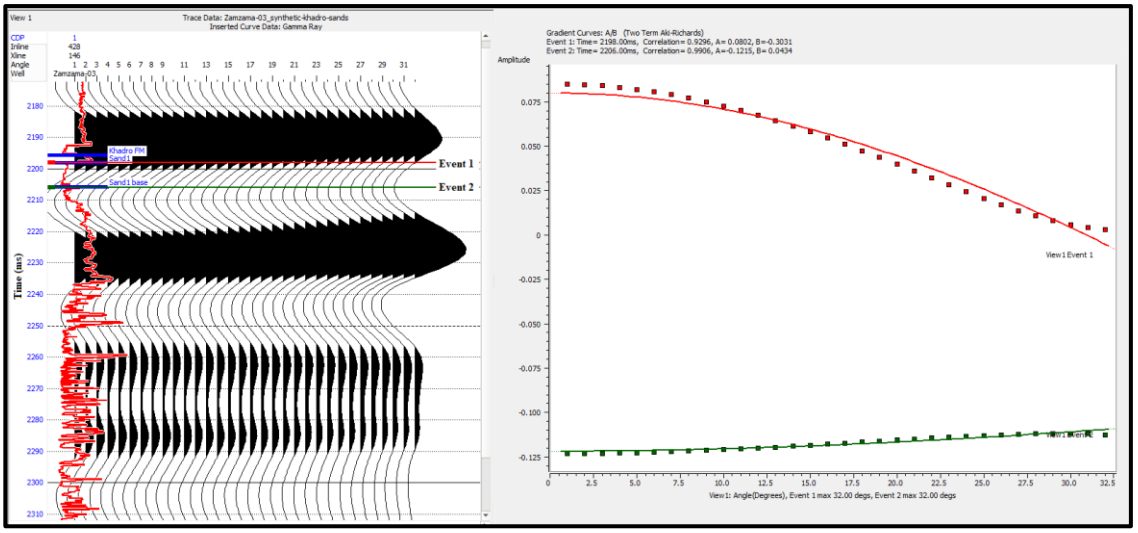



**Figure 4.63** The amplitude responses of the identified event observed at the amplitude vs. angles of incident crossplot.

Castagna and Swan (1997) proposed *AVO* crossplotting wherein an estimate of the normal-incidence reflectivity is plotted against a measure of the offset dependent reflectivity. The complete classification is represented in Figure 4.64.

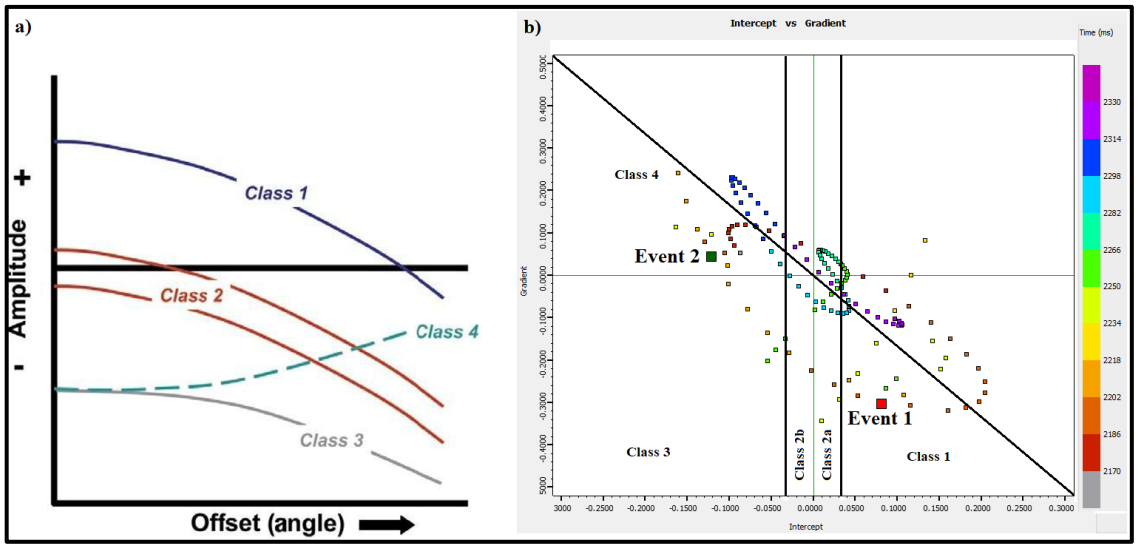



**Figure 4.64** a) Sand classification scheme. b) The *AVO* responses of Event 1 and Event 2 is plotted that depicted Event 1 (top sand) as Class 1 while the Event 2 as class 4.

### 4.4.3 Low Frequency Model

The LFM of $Z_p$ of Zamzama Gas Field passing all available wells is illustrated in Figure 4.65. The $Z_p$ of well logs having band-pass frequencies upto 15 Hz depicted good correlation with the LFM model covering Ranikot/Khadro Formation and Pab Formation for clear understating. The Pab Formation depicted low impedances as it is completely

comprising of sandstone in comparison with mixed lithologies of Ranikot/Khadro Formation.

Similarly, the LFM model of $Z_s$ depicted in Figure 4.66 among the corresponding well logs that are filtered according to the developed model (upto 15 Hz). A good accordance of well property with approximated model is observed with smooth trend of value ranges covering both of the reservoirs of Zamzama Gas Field.

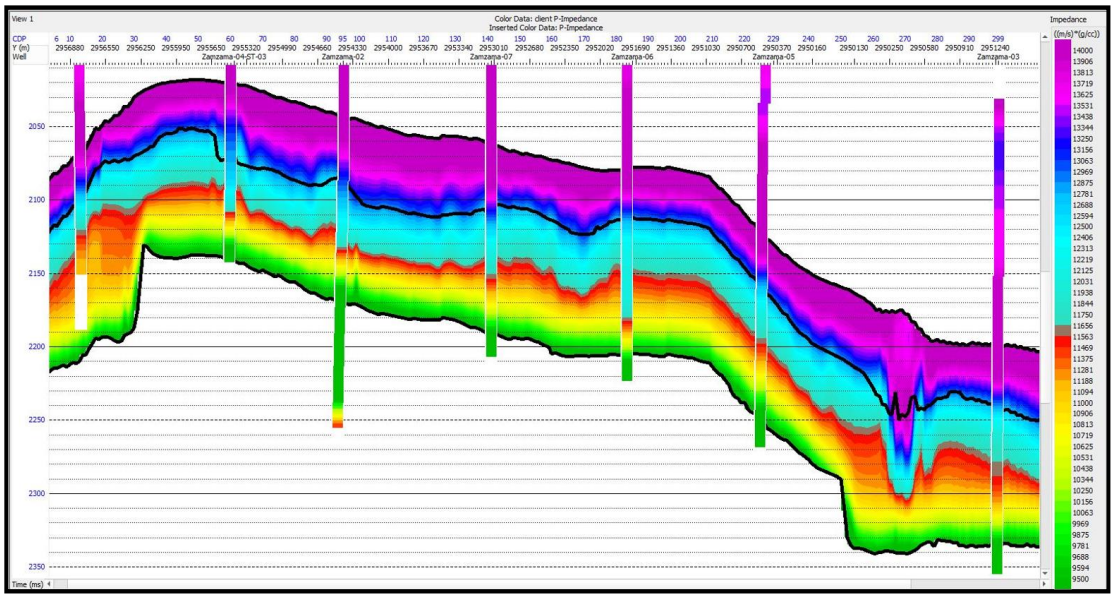



**Figure 4.65** Arbitrary line passing through the wells showing the low-frequency model of $Z_p$ used for Pre-Stack Inversion study overlain by the Zamzama Field wells with similar property of impedances filtered at 15 Hz.

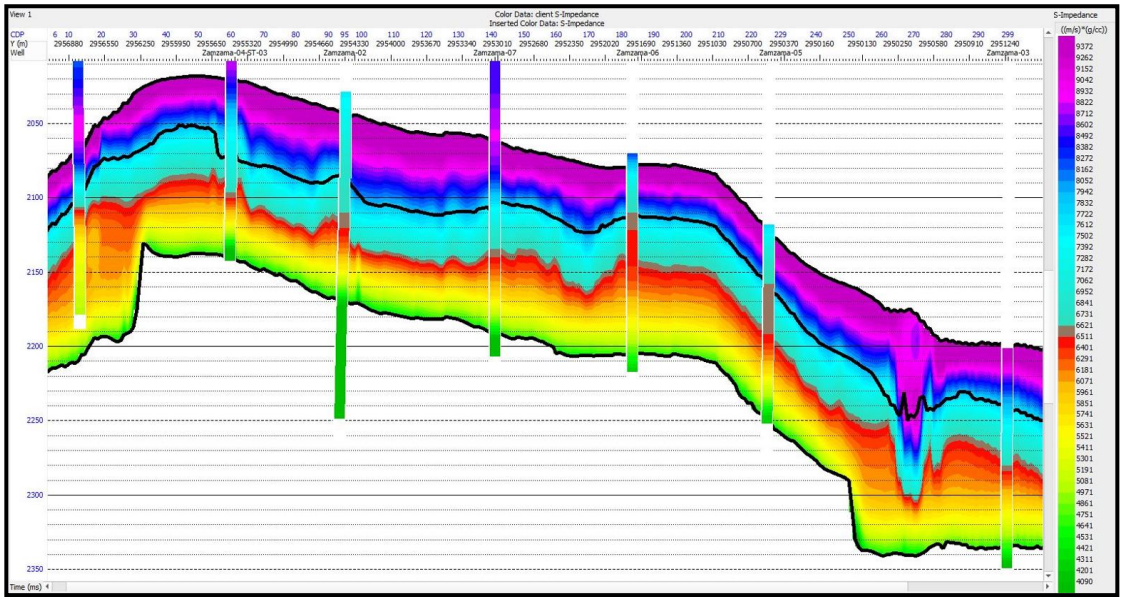



**Figure 4.66** Arbitrary line passing through the wells showing the low-frequency model of $Z_s$ overlain by the Zamzama Field wells having $Z_s$ log.

### 4.4.4 Pre-Stack Simultaneous Seismic Inversion (*PSSI*)

*PSSI* Figure 4.67 depicts a comprehensive evaluation of the measured impedances ($Z_p$, $Z_s$), density ($\rho$), $V_p/V_s$ ratio (blue), and their inverted properties (red), along with LFM (background trend with black color) at the well location of the Zamzama-05. The number of iterations is kept at 25, which iteratively lessens the error and brings the synthetic (blue) traces in accordance with field seismic (red). The synthetic traces are produced by taking inverted curves as input and using a wavelet extracted from the reservoir zone for the convolution purpose; hence, analyze the synthetics for error analysis through a forward model. The assessment of synthetics with field seismic depicted a correlation coefficient of 96.19% and minimum error values of about 0.2559. The coefficients of *k*, *kc*, *m*, and *mc* are calculated from the regression analysis on the crossplotting of well data along with the calculation of covariances from the relationship between $Z_p$, $Z_s$, and, $\rho$ as depicted in Figure 4.68 and Figure 4.69.

The output of Pre-Stack Inversion was represented as inverted $Z_p$ and $Z_s$ cross-sections. Calibration of inverted results with well log data showed very good matching within the zone of interest (Ranikot/Khadro Formation) at the well locations. Low impedance resides between high impedance at the reservoir level passing through wells fairly linked with the amount of potential fluid content present in that particular zone.

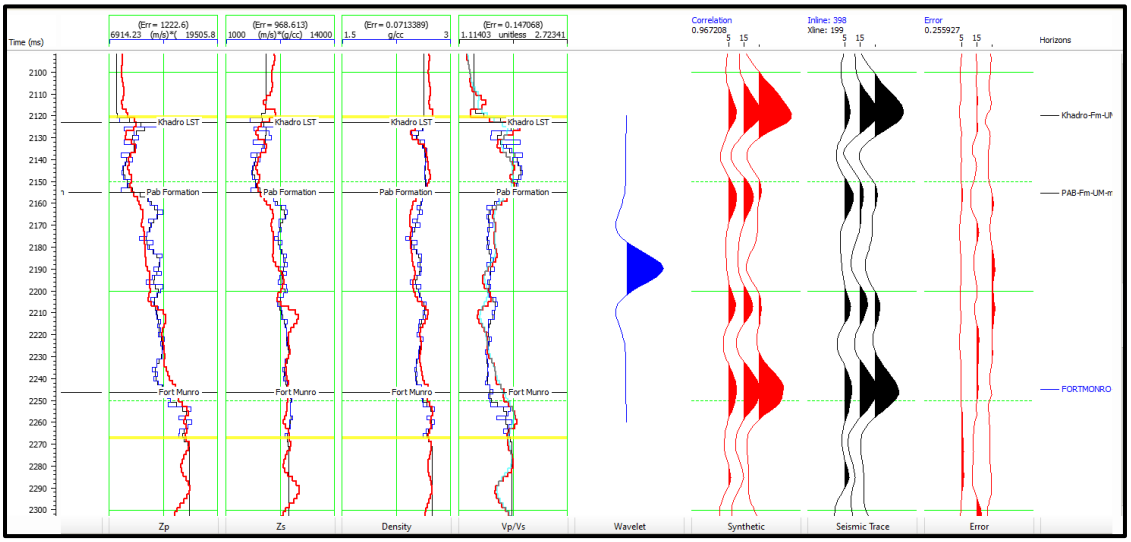



**Figure 4.67** Inversion analysis via comparison of inverted (red), measured (blue), and *LFM* (black) logs along with synthetic (red) generated from inverted logs by convolving with wavelet extracted reservoir zones correlation with seismic (black).

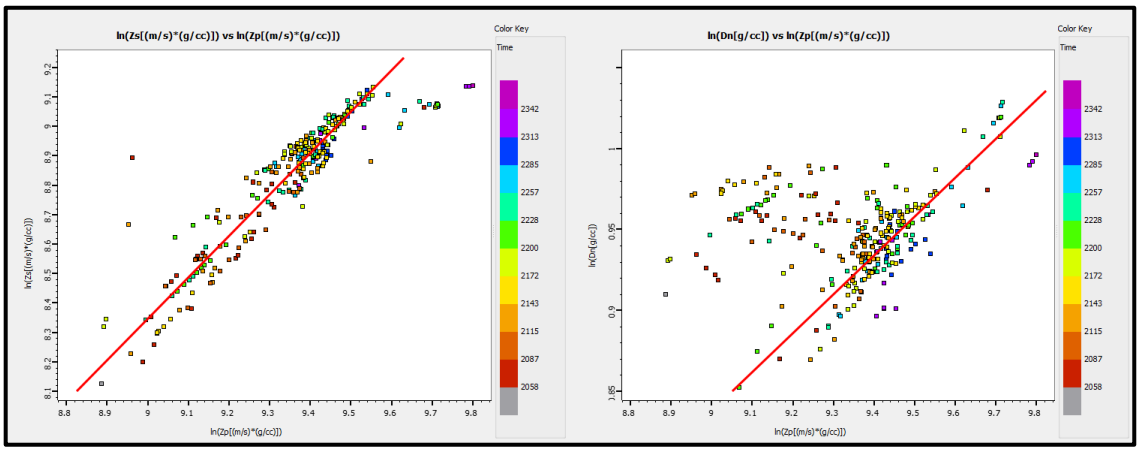



**Figure 4.68** Regression analysis based on natural logs on $Z_p$, $Z_s$, and $\rho$, i.e., $L_n$ ($Z_p$), $Ln$ ($Z_s$) and $L_n$ ($D_n$).

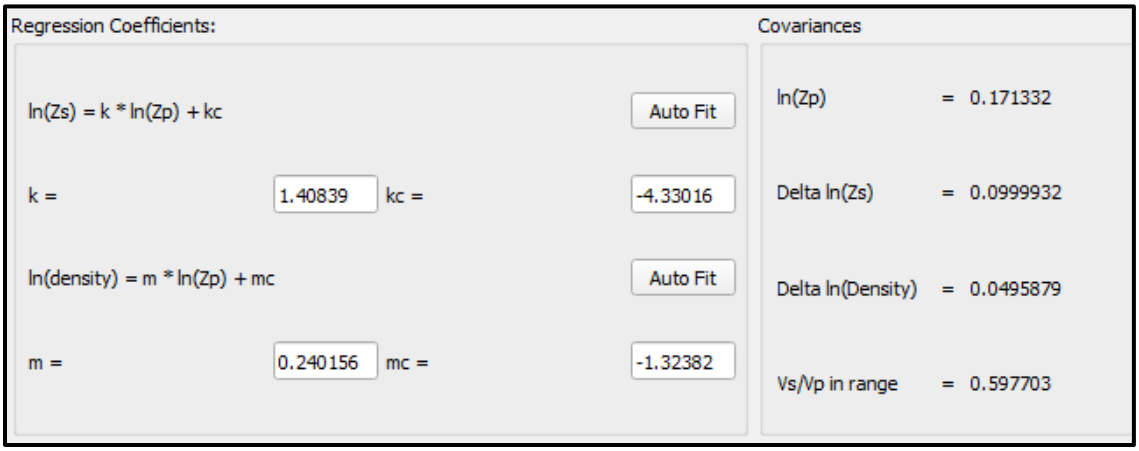



**Figure 4.69** The coefficients of *k*, *kc*, *m*, and *mc* are estimated from the regression analysis along with covariances.

In cross-section of $Z_p$ passing through all wells, the low values entrapped within high values possibly the potential sands representation (Figure 4.70). Furthermore, the inverted attribute of $Z_s$ also correspond to the outcomes attained from $Z_p$ showing lower values bounded between relatively higher $Z_s$ can be the plausible sands delineation (Figure 4.71). Overall, the seismic resolution is not enough to fully characterize the thin Ranikot/Khadro Formation (4-7 m), therefore the high resolution seismic is attained via advance *ML* algorithms that effectively trace out the potential sand's bodies on the inverted maps.

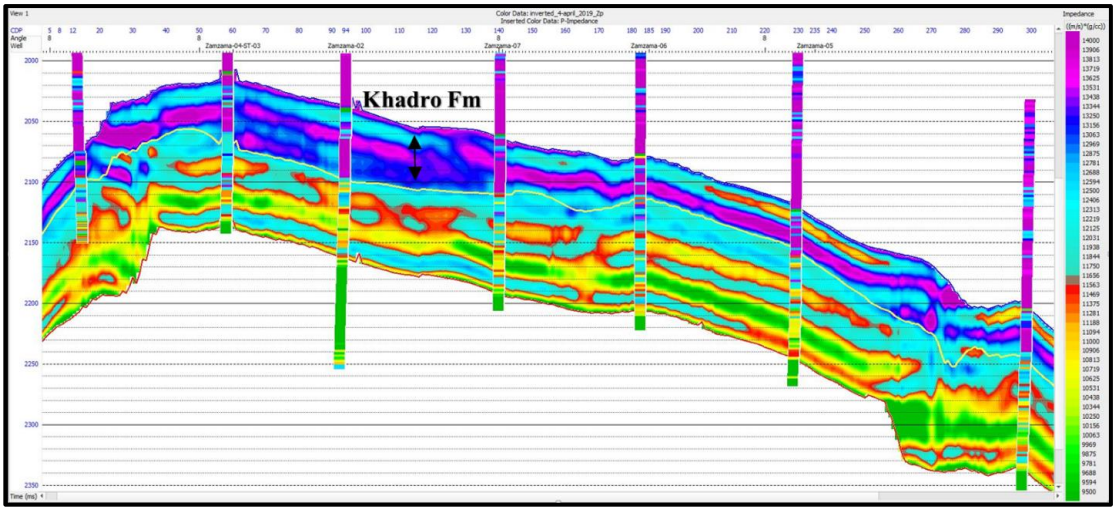



**Figure 4.70** The inverted impedance ($Z_p$) in a section view bisecting all Zamzama wells with inserted $Z_p$ curve in displayed wells.

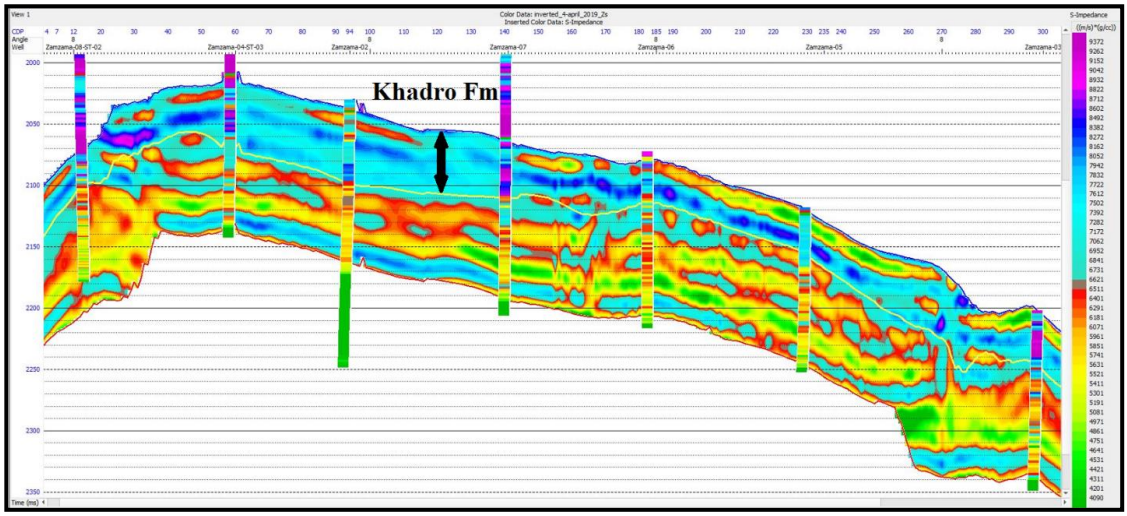



**Figure 4.71** The inverted impedance ($Z_s$) in a section view bisecting all Zamzama wells with inserted $Z_p$ curve in displayed wells.

### 4.4.5 Slices within Ranikot/Khadro Formation

Within the Ranikot/Khadro Formation, three sand bodies were identified based on petrophysics and were mainly characterized and entitled as Sand-01, Sand-02, and Sand-03, correspondingly marked at the well location (Figure 4.72) used within the study.

The inverted $Z_p$ slice of bottom part, comprising the Sand-03 interval (15 ms above Pab Formation) of the Ranikot/Khadro Formation. The stratigraphic slice shows that form north to south there are high values of impedance at the crestal portions of the anticline (core area where wells drilled) instead of few locations i.e., Zamzama-05 observed low impedance values (Figure 4.73).

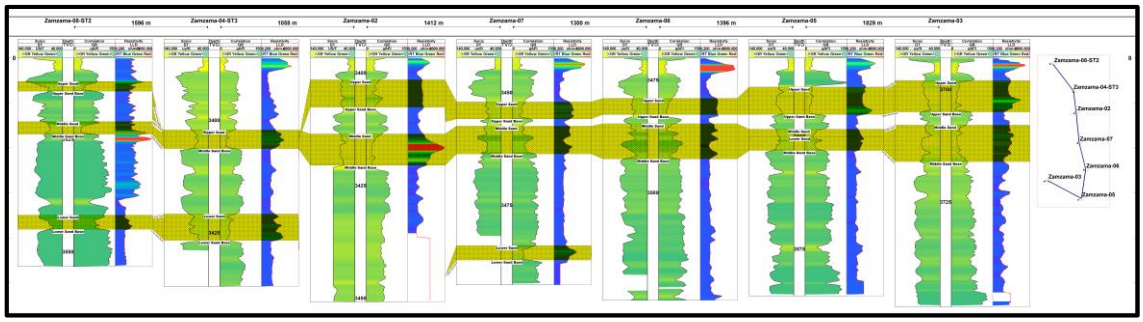


**Figure 4.72** Cross-correlation illustrating three sand layers distribution based on *GR* and P-sonic on all wells. The three sands are present in the wells that are further assessed by the petrophysical and cut-off values.

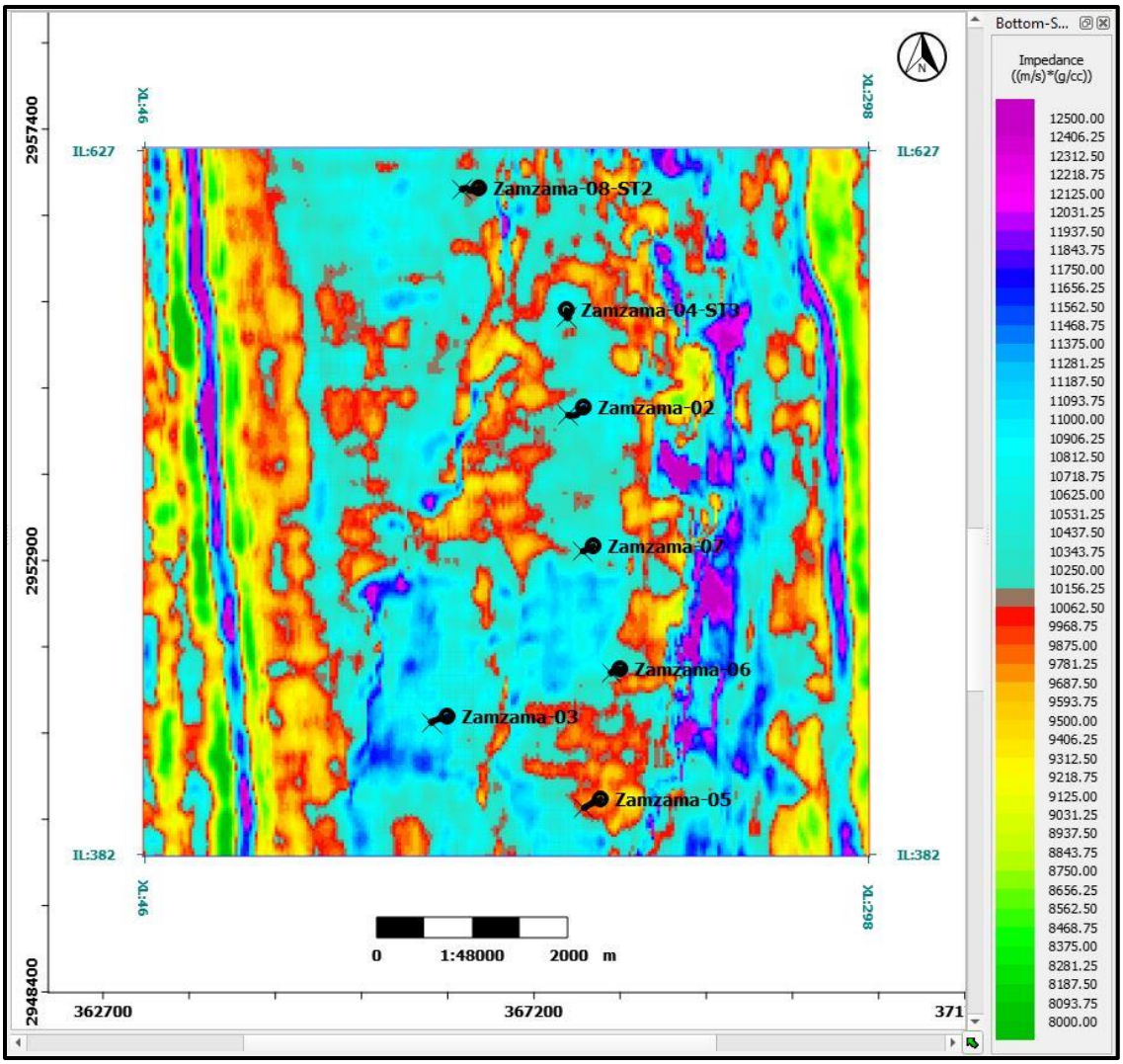



**Figure 4.73** Inverted $Z_p$ slice of Bottom sand of Ranikot/Khadro Formation.

The inverted $Z_p$ slice of the mid sands, i.e., Sand-02, is made by first shifting the Ranikot/Khadro Formation's horizon by 15 ms, and then mapping the mean of a 10 ms

window within the Ranikot/Khadro Formation, which shows no prominent clue for the potential sands. This is possibly due to the resolution limitation of seismic data regarding the thin gas sand beds Figure 4.74).

Similarly, the inverted $Z_p$ slice of top sands, i.e., Sand-01 (top of Ranikot/Khadro Formation's horizon up to a 15 ms window below), delineated low impedance values near Zamzama-04 and Zamzama-02 (Figure 4.75).

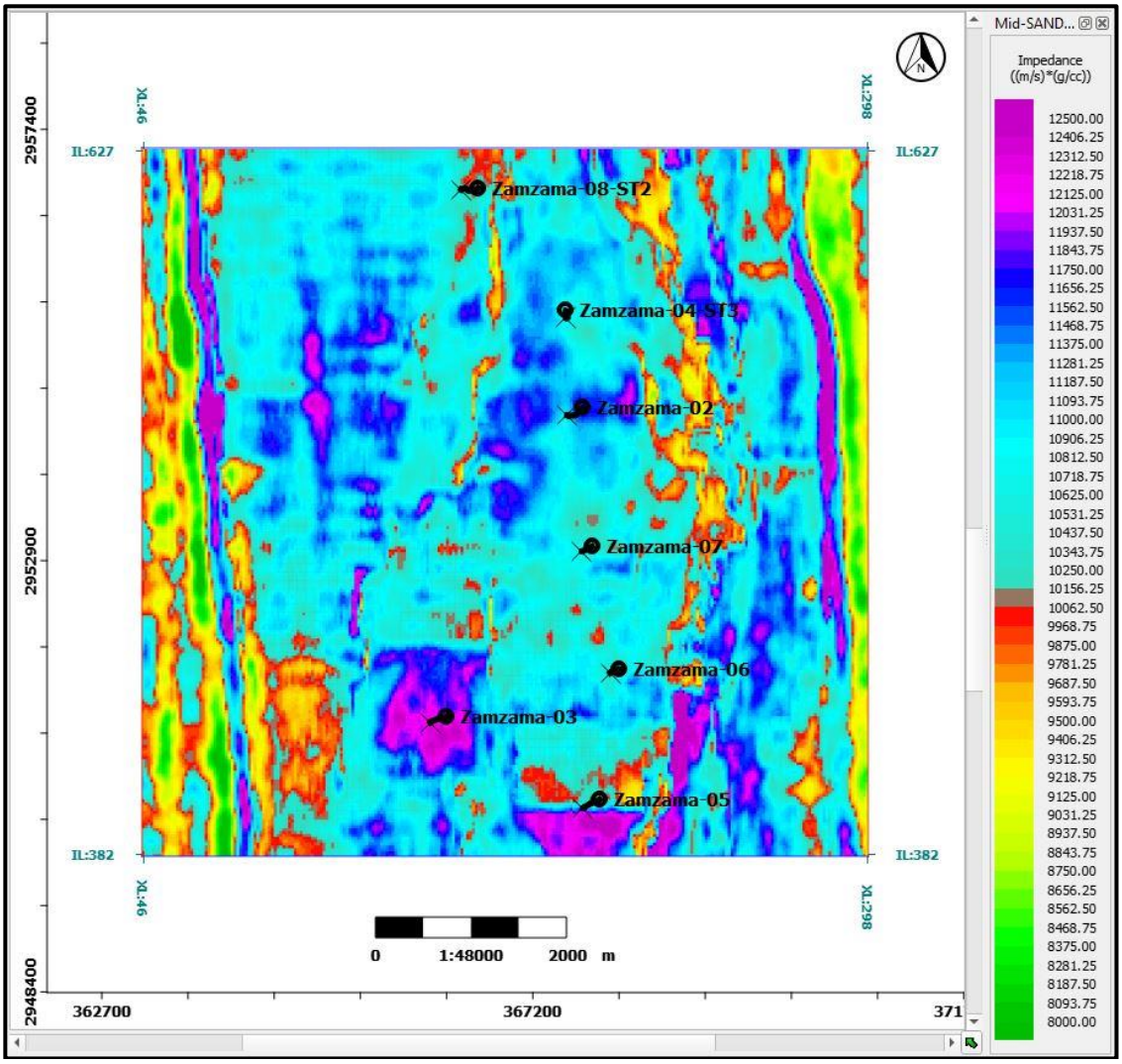



**Figure 4.74** Inverted $Z_p$ slice of Mid sand of Kadro Formation.

Conclusively, the thin sands of having maximum thickness of 7 m present Ranikot/Khadro Formation are not able to clearly resolve on the limited frequency spectrum that contain dominant frequency of about 20 Hz. In Figure, the potential gas bearing sands are identified with integrated relationship of elastic and petrophysical

properties. The gas-bearing sands have comparatively large porosities (up to 8%) with low values of $S_W \leq 0.45$ and $V_{CL} \leq 0.30$ reflecting correspondingly low elastic ranges in $Z_p$ (7500-9500 g/cc*m/s) and $V_p/V_s$ ratio (1.4-1.6). The potential sand is analyzed based on the wedge model so that resolving power of seismic data is assessed. The wedge model describes the shift in the amplitude of reflection events when the bed thickness decreases below the dominating wavelength of the seismic data (Saeed et al., 2020; Simm, 2009), represented by equation 4.3:

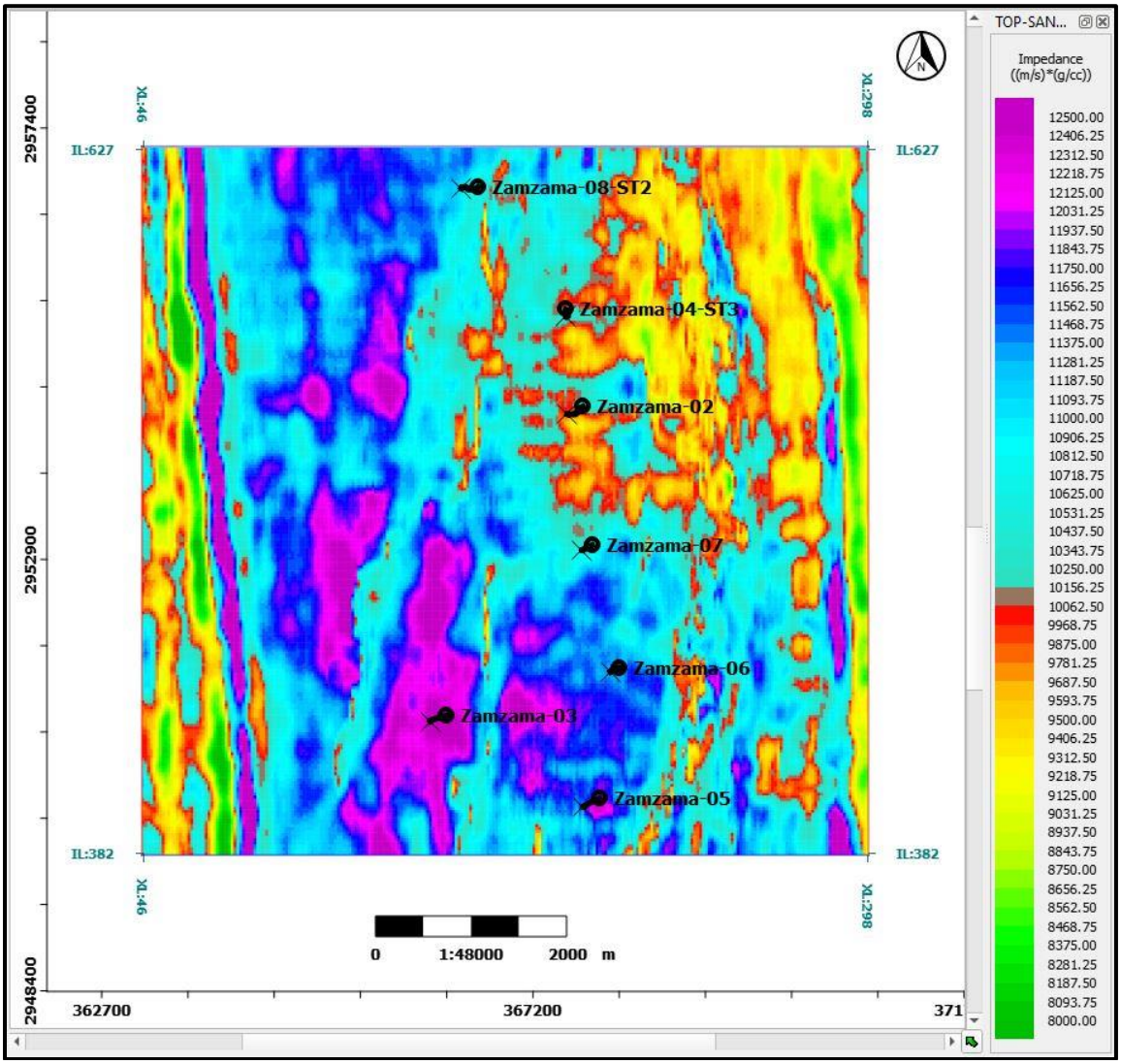



**Figure 4.75** Inverted $Z_p$ slice of Top sand of Kadro Formation.

$$\boldsymbol{Z = Vi/2.8f_d} \quad (4.3)$$

$Z$ represents the tuning thickness of bed, $Vi$ is the dominant velocity i.e., approximately 4000 m/s in the potential sand layers, while $f_d$ is the dominant frequency

within the reservoir zone that is about 20 Hz. The double-interference-wedge modeling concluded that constructive interference of waves having opposite polarity starts at 60 m of thickness, while maximum tuning thickness is observed around 40 m, having the highest amplitude (Figure 4.76). The synthetic seismic response is observed on the basis of interpreted horizons on zero-crossings at the top and base of thin Ranikot/Khadro Formation gas sand. Hence, such an observation depends on the shape and frequency of the wavelet and helps to notify the limitation of the resolving power of a seismic dataset, such as about 6 m of thin gas sand beds of the Ranikot/Khadro Formation. This technique helped to identify that seismic resolution is not much greater for Ranikot/Khadro Formation sands while it can resolve much thicker Pab Formation. Therefore, *ML* needs to be implemented for enhanced properties of elastic and petrophysical properties estimation within thin and heterogeneous reservoir Ranikot/Khadro Formation.

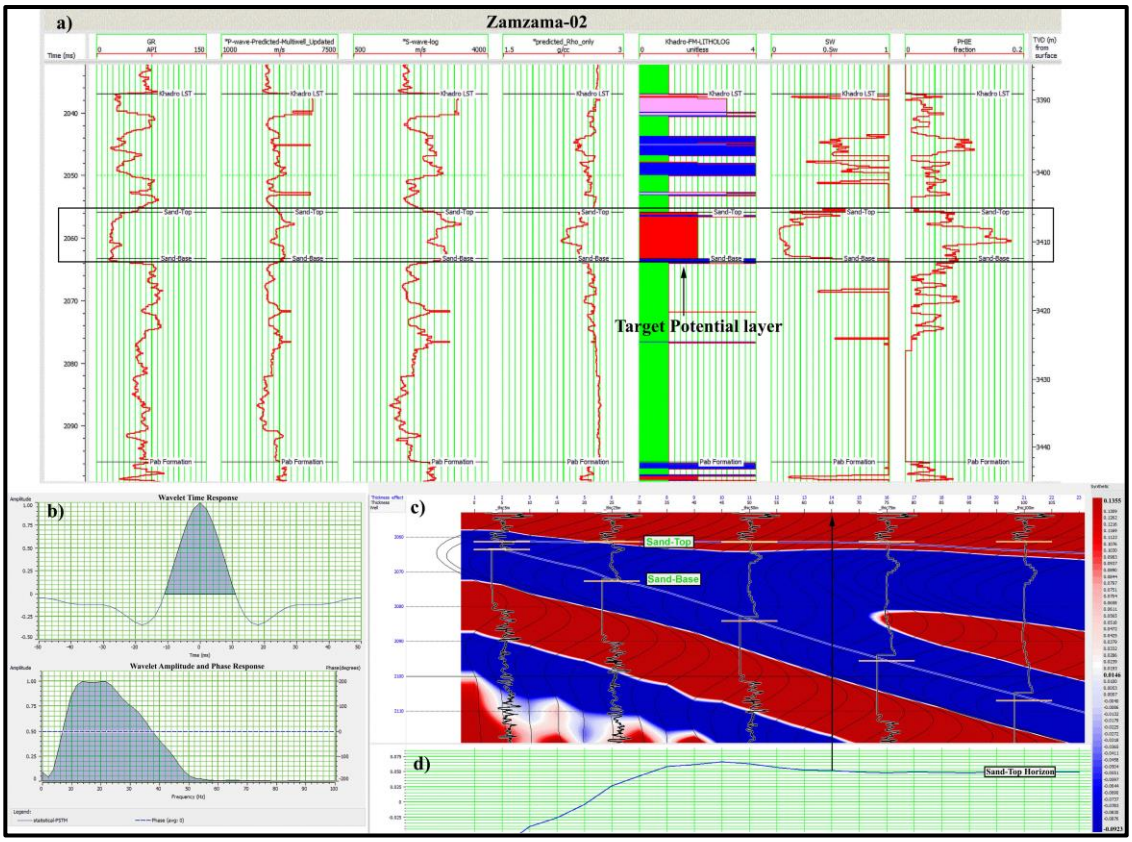



**Figure 4.76** a) A gas sand body with 7 m thickness encountered at Zamzama-02 well after petrophysical analysis. b) Statistical wavelet used to generate synthetic wedge model. c) Synthetic wedge models with variable thickness, starting from zero till 100 m reflects seismic tuning thickness for sand body. d) Around 60 m is the resolution limits for the seismic dataset under study.

### 4.4.6 *ML* Based Elastic, Petrophysical and Facies Prediction

Petrophysical interpretation confirmed the presence of the potential gas sands within the Zamzama-02, Zamzama-03, Zamzama-04-ST3, Zamzama-05, and Zamzama-07 wells, whereas Zamzama-06 and Zamzama-08-ST-2 did not encounter these sands. The Drill Stem Test (*DST*) reports of Zamzama-03 and Zamzama-05 confirmed that these sands were successfully perforated to obtain the gas production from Ranikot/Khadro Formation. Henceforth, these potential gas sand bodies are assessed with *ML* based elastic ($Z_p$, $Z_s$) (Figure 4.77, Figure 4.78) and petrophysical volumetrics, i.e., *PHIE* (Figure 4.79), to achieve high-resolution volumetrics. These volumes have a comparatively higher resolution due to the increased frequency spectrum. The *ML* technique that is best employed for *RPM* of elastic properties is employed for the prediction of volumetrics. The high resolution of volumetrics beyond the known well location across the field is one of the primary objectives of the study.

The best relationship for identifying the gas sand facies using the elastic logs is to have low *GR* along with low $S_W$ while having high values of *PHIE* (Durrani et al., 2022; Durrani et al., 2022a). Based on this defined criterion and with the help of key *ML* based elastic properties, the properties of gas sand facies were predicted across the whole volume, hence separating from the non-reservoir facies as well.

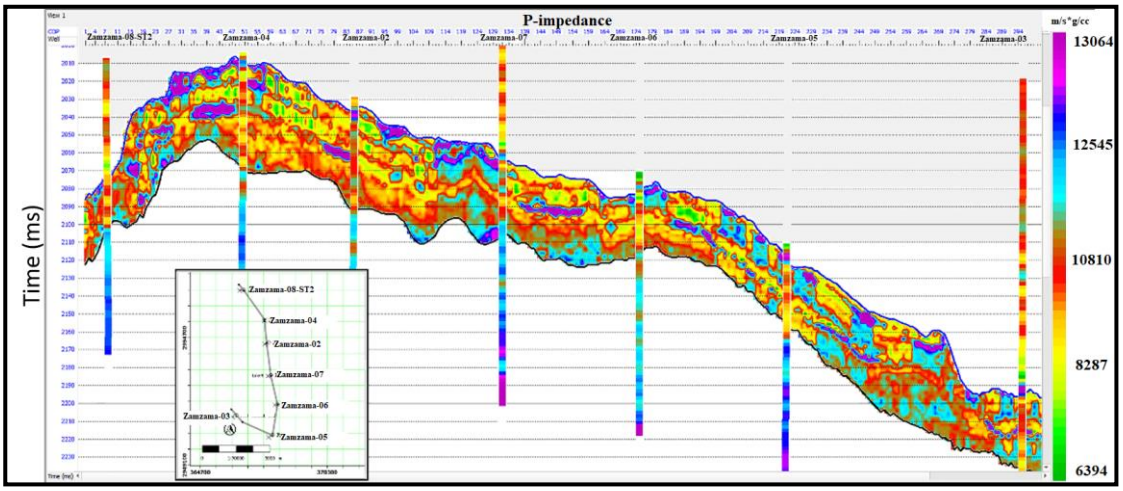



**Figure 4.77** *ML* based predicted Zp shown along an arbitrary line passing through all wells. The corresponding property log of Zp shown at each well location that shows good match with high resolution compared to Pre-stack seismic inversion.

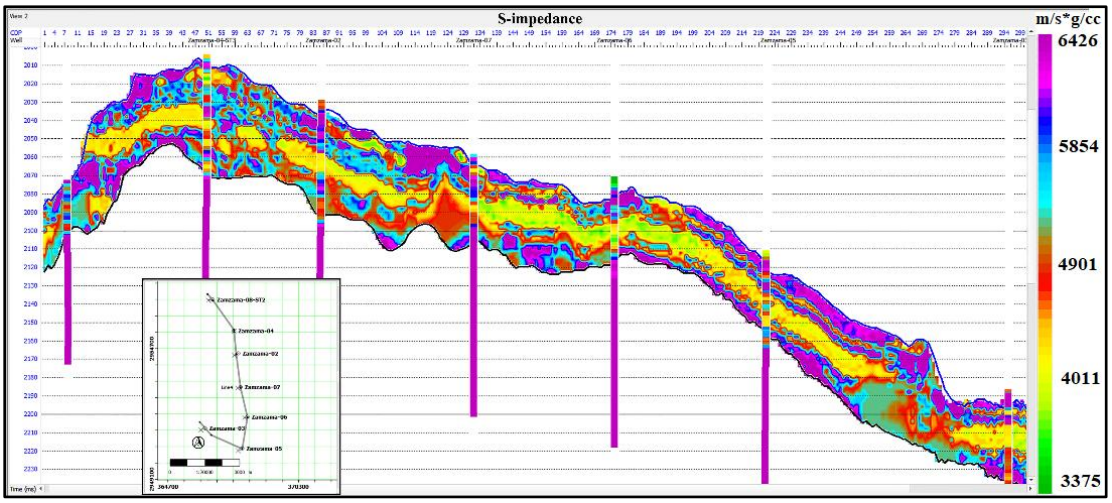



**Figure 4.78** *ML* based predicted $Z_s$ shown along an arbitrary line passing through all wells. The corresponding property log of $Z_s$ shown at each well location that reflect good match, authenticated the results.

The *PHIE* volumetric is depicted on all of the wells while the wells are colored with same log property (Figure 4.79). A reliable match of the modeled petrophysical property is observed on similar properties present within the wells. The three sands i.e., Sand-01, Sand-02, and Sand-03 (Top, middle, and Bottom) are mapped that showed good porosities distribution (about 9%) throughout the field (Figure 4.80).

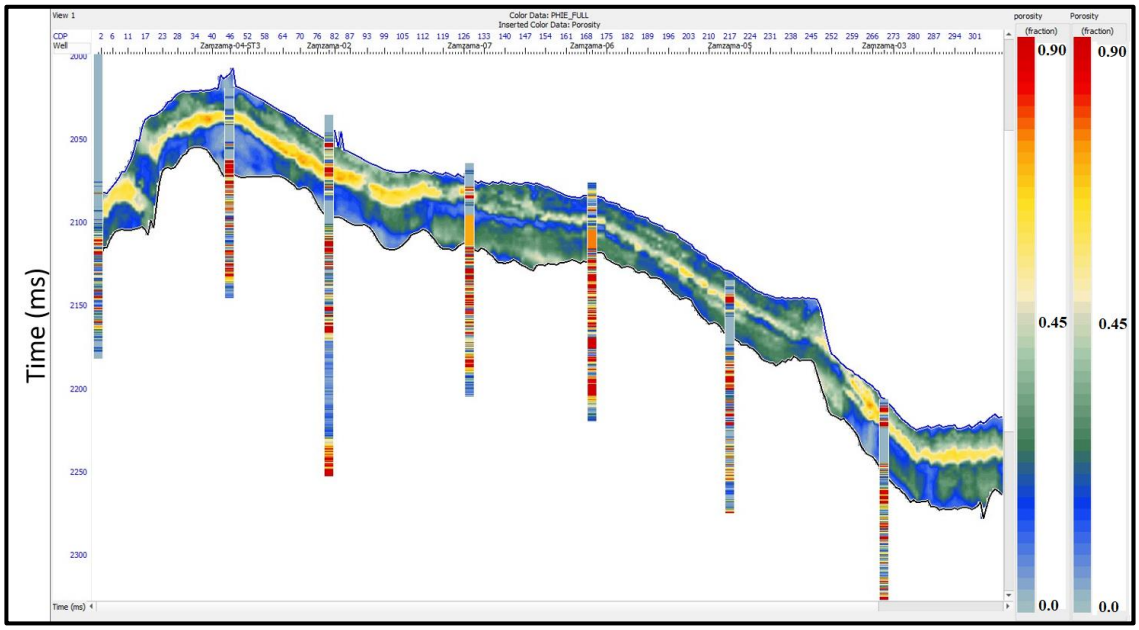



**Figure 4.79** *ML* based *PHIE* volumetrics passing through all wells delineates a reliable match with similar property of well that authenticated the *ML* results.

Similarly, the gas bearing sands are approximated passing through all wells (Figure 4.81). The net thickness of producing sands with thickness of about 7 m is

enhanced by the usage of *ML*. The detail mapping of sands facies distribution is depicted in Figures 4.82-4.84.

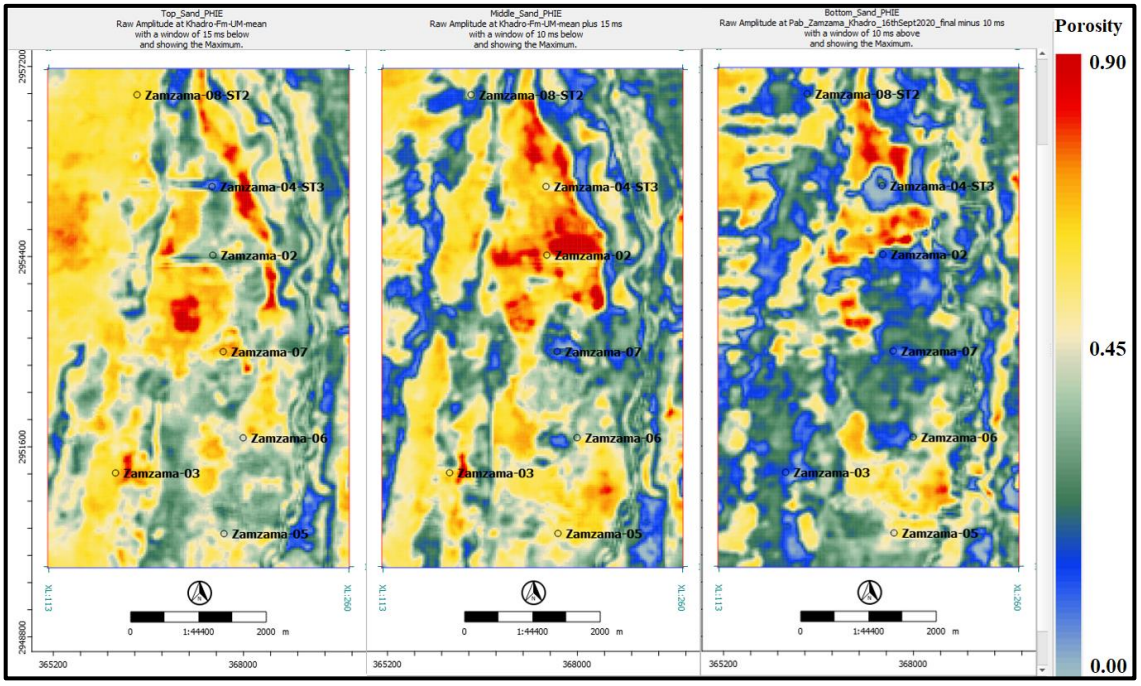



**Figure 4.80** Three Sands-01, 02, and 03 intervals (top, middle, and bottom) delineated a good distribution of porosities.

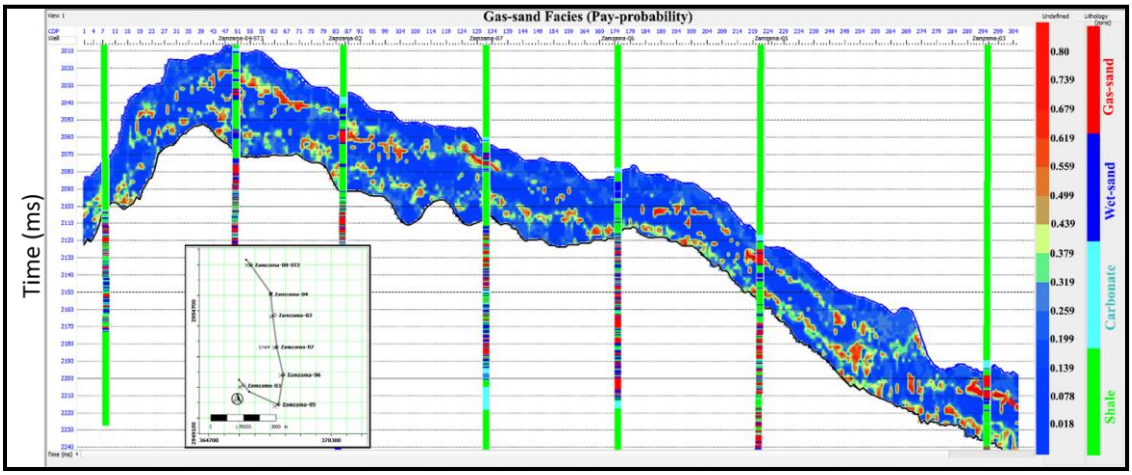



**Figure 4.81** *ML* based gas bearing sands passing through all wells delineates a reliable match with the identified gas sands facies of the wells authenticated the *ML* results.

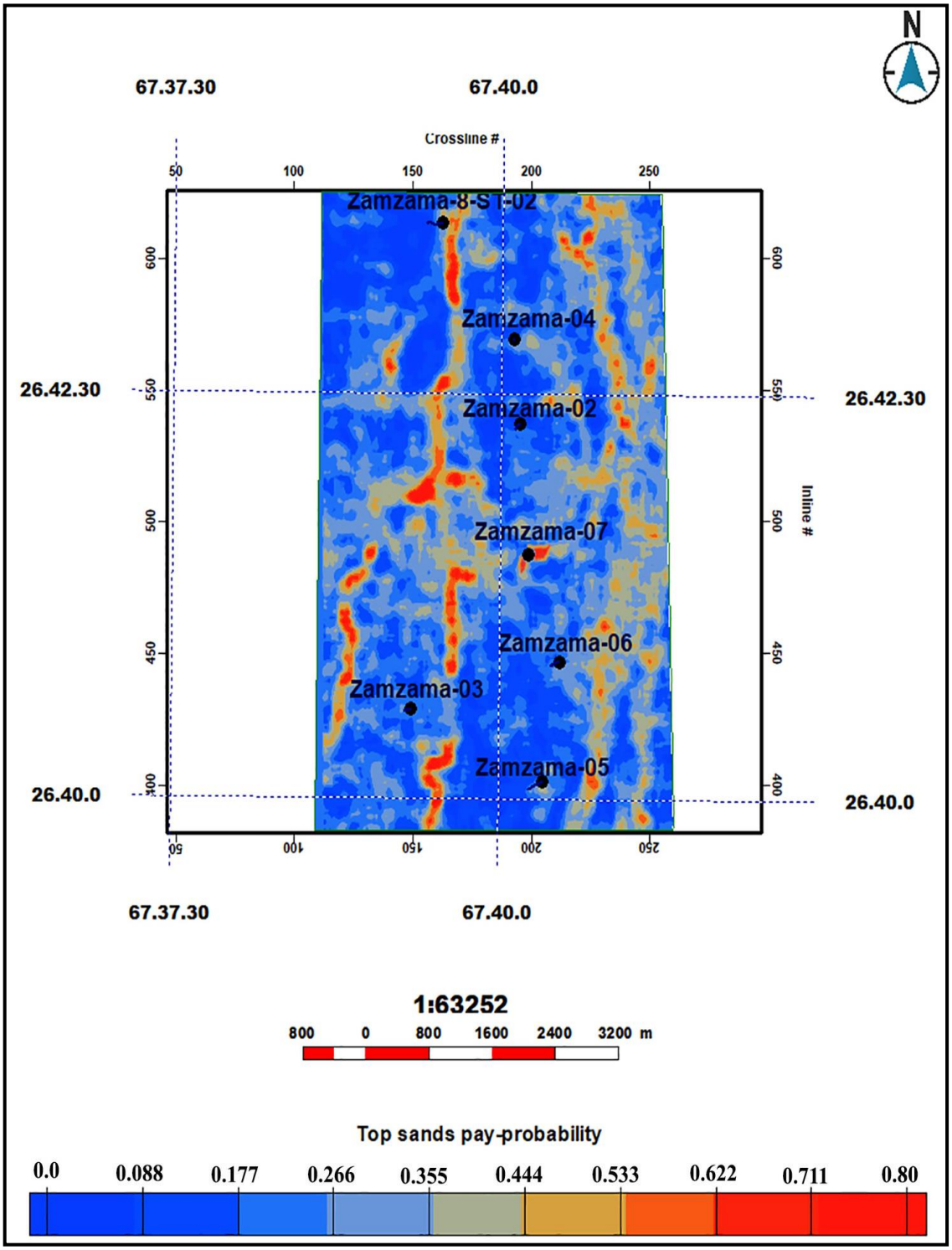



**Figure 4.82** Map illustrating the probability distribution to encounter the hydrocarbon bearing sand facies for the Sand-01 within Zamzama Gas Field. Red color depicts the highest probability to encounter the hydrocarbon bearing sand facies.

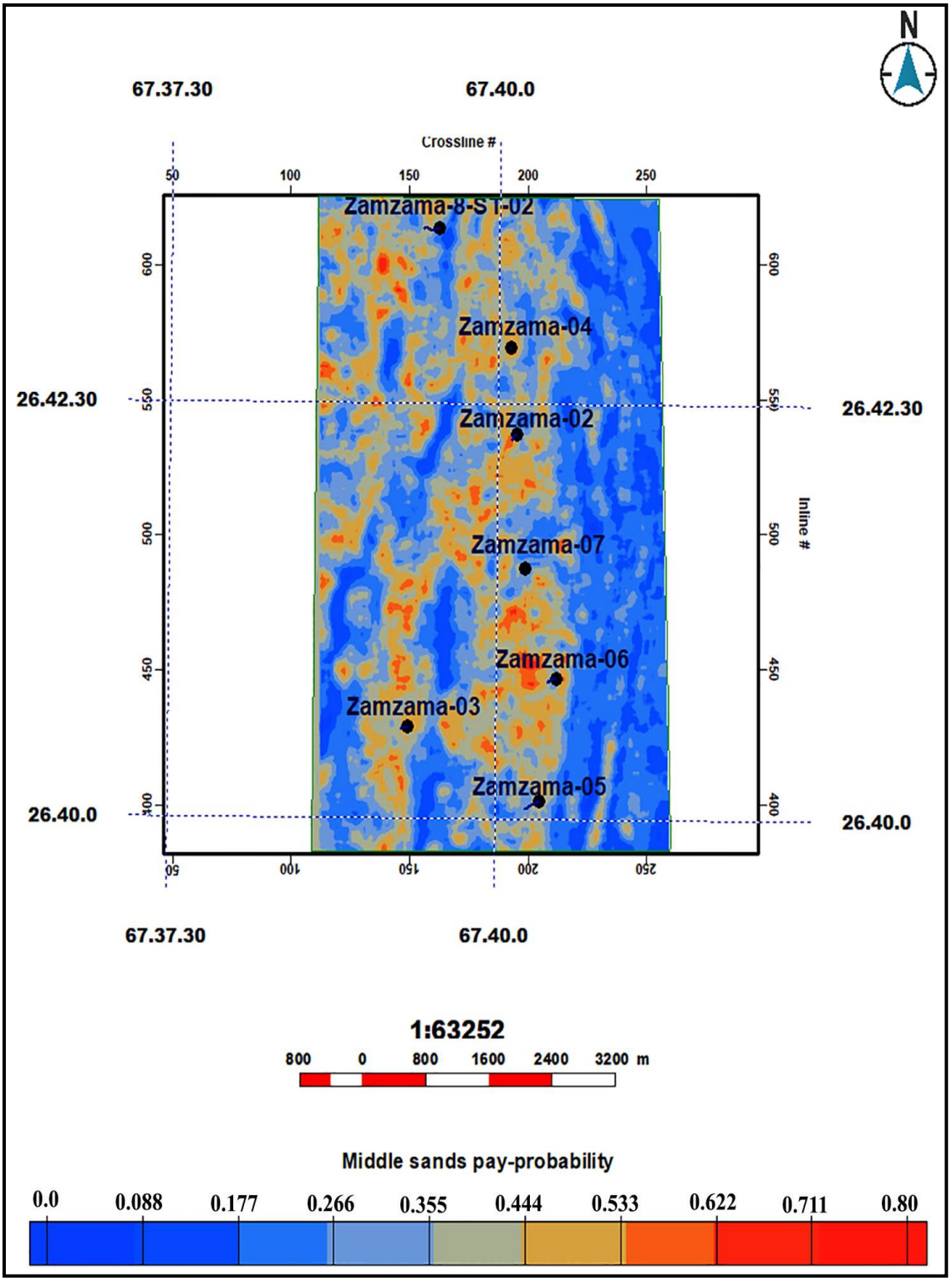



**Figure 4.83** Map illustrating the probability distribution to encounter the hydrocarbon bearing sand facies for the Sand-02 within Zamzama Gas Field. Red color depicts the highest probability to encounter the hydrocarbon bearing sand facies.

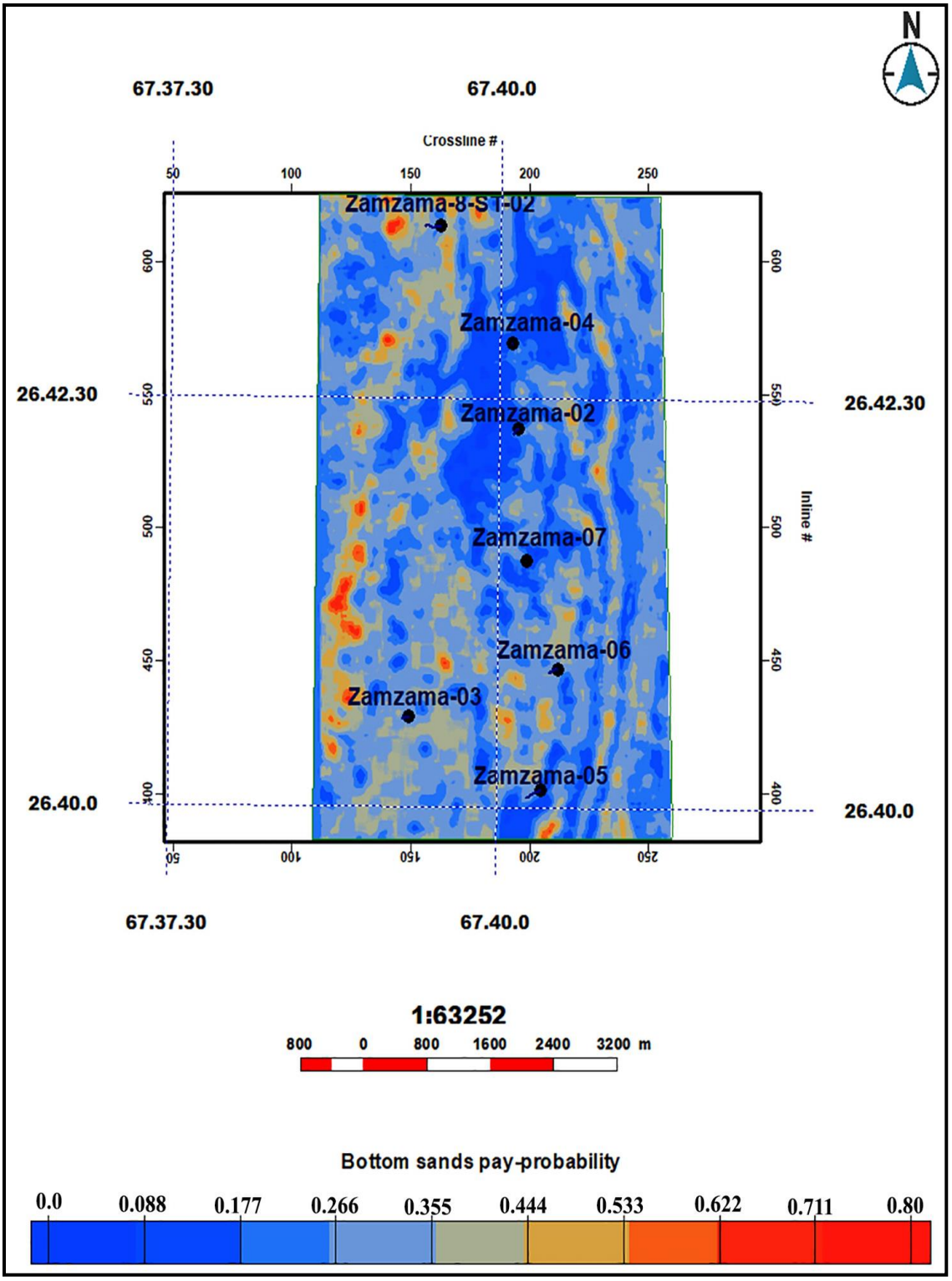



**Figure 4.84** Map illustrating the probability distribution to encounter the hydrocarbon bearing sand facies for the Sand Body 3 within Zamzama Gas Field. Red color depicts the highest probability to encounter the hydrocarbon bearing sand facies.

### 4.4.7 Continuous Wavelet Transform

The primary target is to resolve the thin sands by employing the Post-Stack seismic data through injecting high frequencies. The seismic frequency band ranges were

optimized from (0-5-30-40 Hz), with 20 Hz being the dominating frequency (Figure 4.76) to (0-5-60-100 Hz), while after injecting high frequencies peak frequency improved to 25 Hz, as evident by the synthetic wedge model of Zamzama-03 well (Figure 4.85). By applying the *CWT-DNN* approach after injecting high frequencies, a *HFSSV* has been generated. The results of output seismic with a sampling rate of 1 ms that drastically improved seismic resolution. The results of output seismic with a sampling rate of 1 ms drastically improved seismic resolution.

It has been validated by the lithology log that can be distinguished between sand or shales as thin as 2 m. The resolution of the lithology log is substantially higher than that of the typical seismic data.

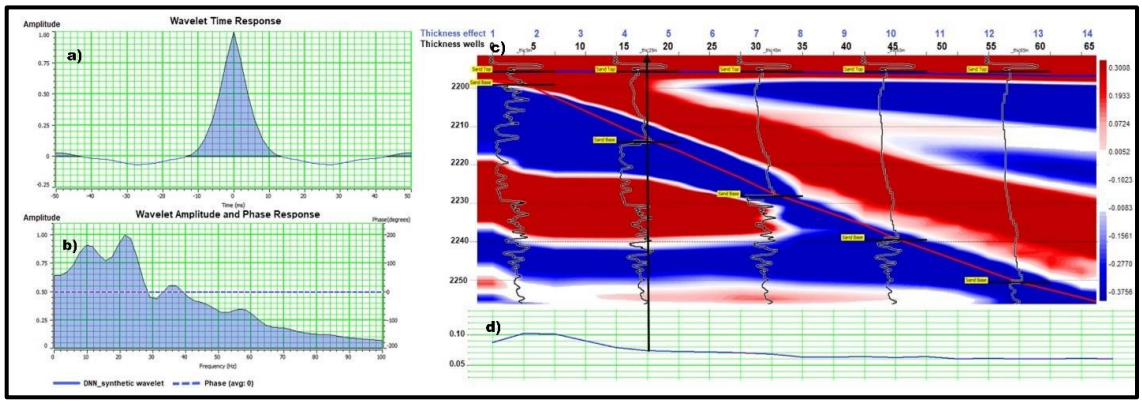



**Figure 4.85** a) The synthetic wedge model has improved resolution employing the high-frequency (0-5-60-100 Hz) wavelet b) with 25 Hz peak frequency and used for *CWT-DNNs* based high-frequency seismic synthetics volume. c) High resolution synthetic wedge model. d) Resolution limit around 18 m.

Different methods have been employed to organize Post-Stack seismic trace data into vectors or matrices for *NNET*. *CWT* is used on the original Post-Stack Seismic data to obtain seismic spectra (Figure 4.86). Before wavelet transformation, several parameters to the *CWT* about the Post-Stack seismic data are fed such as:

1. Sampling rate of Post-Stack Seismic data: If the Post-Stack seismic data sampling rate is 2.0 ms, then the frequency sampling rate which is the required parameter for *CWT* will be 200 s.
2. Frequency band of Post-Stack seismic data: The amplitude spectrum provides information about the frequency bands (0-5-30-40 Hz) for field seismic data, and vice versa.
3. Number of scales: The number of scales is dependent upon the quantity of data. If the rows of data for the *DL* model are fewer, the number of scales should be lower to avoid overfitting data. Constraining the number of scales will give less

certain results. In this case, the optimal number of scales is considered as 100 scales.

A non-stationary complex Morlet wavelet function has been employed after analysis to carry out *CWT* for Post-Stack seismic data. The *CWT* spectra reveal the hidden geological information in the Post-Stack seismic data by enhancing the signal's energy and suppressing noise in time and frequency non-stationary mode (Ali et al., 2020). *CWT* disintegrates seismic data in the form of an imaginary part, the real part and their corresponding magnitude. The frequency band of the initial Post-Stack seismic data can be fully utilized in this technique.

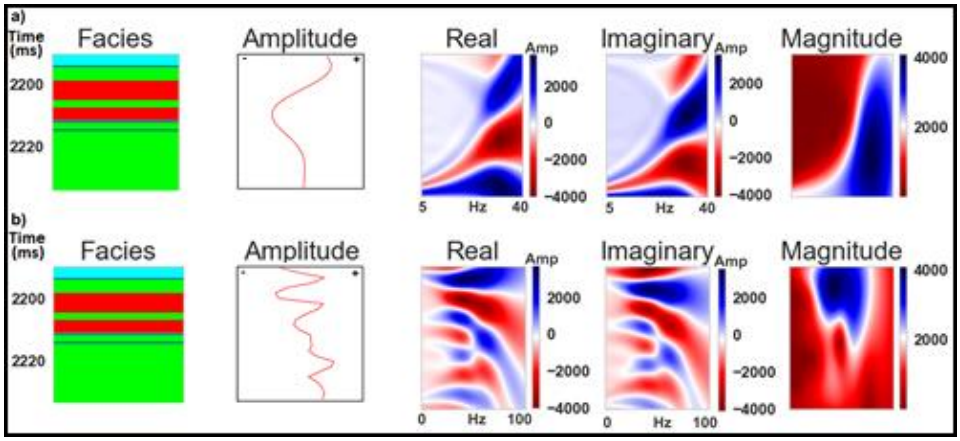



**Figure 4.86** *CWT* transformation seismic into real, imaginary, and magnitude components with facies log in the same interval at well location from a) seismic trace & b) high-frequency synthetic seismic.

### 4.4.8 Deep Neural Networks (*DNNs*) Technique

*NNET* are based on a neuron's computation and entail a weighted matrix of the input values. The synapses' value scaling and the neuron's combination of these values are symbolized by the weighted sums. A cascade of neurons would then perform a straightforward linear algebra process, outputting not only the weighted sum but rather the combined inputs that are subjected to a functional activity within the neuron. This technique seems to be a nonlinear function that only triggers a neuron to produce an output when the inputs exceed a certain threshold. *NNET*, by analogy, pertains to a non-linear function to the input values sum.

The following equation 4.4 is used for computation at each layer (Sze et al., 2017), as shown in Figure 4.87.

$$\boldsymbol{y_1} = \sum_{\boldsymbol{i=1}}^{\boldsymbol{3}} \boldsymbol{W_{ij}} * \boldsymbol{x_i} + \boldsymbol{b} \quad (4.4)$$

Where, $W_{ij}$ are weights, and $x_i$ , $y_j$ represents input and output activations.

The neural network comprises more than two layers, i.e., one or more hidden layers. Nowadays days, *DL* typically uses layers ranging from five to over a thousand.

Compared to shallower *NNET*, *DNNs* can learn high-level characteristics with greater complexity and abstraction. It employed a backpropagation approach for minimizing the difference between actual and predicted output by updating the weights of input features. Using *DNNs* to process visual data is one illustration of this application.

The study's first goal is to produce high-resolution fine-layering seismic by injecting high frequencies in field seismic. It has been achieved by using the impedance log, the *RC* is generated at wells, and a synthetic is generated after convolving the *RC* with a high-frequency wavelet. (Choi and Alkhalifah, 2011) has adhered to the fundamental idea that if the source wavelet's bandwidth is increased, the seismic trace will have a band-extended spectrum and obtain a better resolution from the same reflectivity as shown in Figure 4.86.

Alongside, *CWT* decomposes Post-Stack seismic field data into its real, imaginary, and magnitude parts (Figure 4.86a). By keeping the *CWT* output, as an input feature layer and earlier generated high-frequency synthetic as a target, the *DNN* approach is utilized to populate *HFSSV*. *DNN* employed 4 layers, including three hidden layers with 60, 40, and 20 neurons, respectively.

Further study’s key objective is to utilize high-resolution seismic to populate fine reservoir facies volume by resolving thin gas-sand facies. In this *HFSSV*, after validation at blind wells, is disintegrated into parts in the frequency domain by using the complex Morlet Wavelet function (Figure 4.86b).

In this step for *DNN* training, these three components are taken as input features/input layer by targeting high-resolution facies extracted at the well locations using petrophysical cutoffs. The retained *DNN* included three hidden layers and one input/features layer with 80, 60, and 40 neurons, respectively, using softmax as an activation function. As this activation function works best for multi-classification problems with *DNN* (Figure 4.87). The *DNN* generates results by keeping 70 % as training and the remaining 30% as testing dataset to ensure the model accuracy.

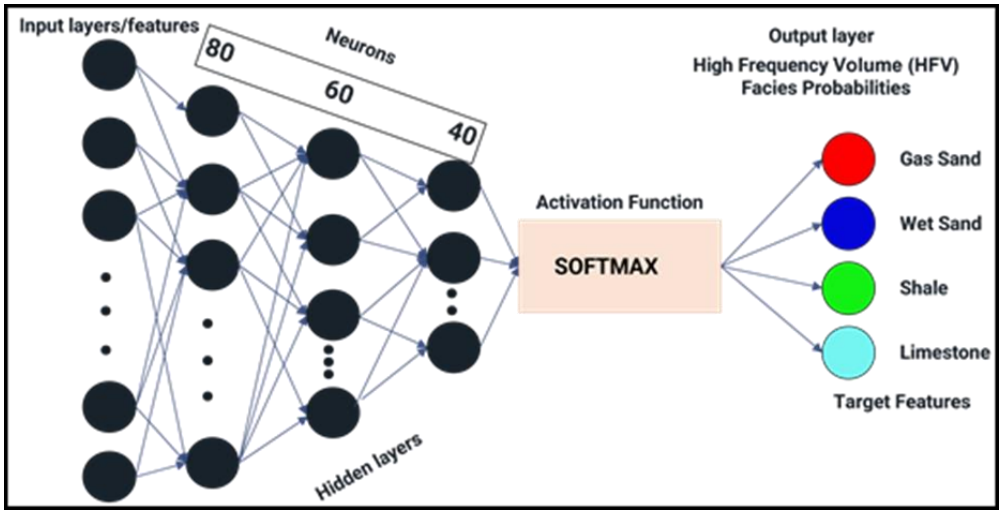



**Figure 4.87** *DNNs* schematic diagram taking input layers as neurons with activation function (softmax) to get output layer.

The *CWT-DNNs* model gives a higher resolution to capture fine reservoir gas-bearing facies more distinctly (Figure 4.88).

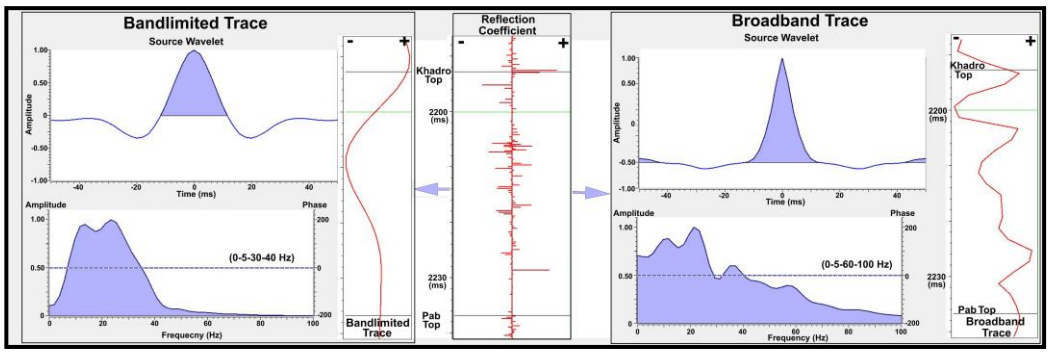



**Figure 4.88** Spectral enrichment from bandlimited to broadband in the convolutional model. Bandlimited trace exhibits coarser resolution while broadband trace reveals fine layers enhancing resolution.

The frequency band optimization process of spectral enhancement, which uses wider band wavelets to transform a narrowband trace into a broadband one. An end-to-end *ML* algorithm could carry out the spectral enhancement if enough narrow band-broadband pairs of traces are available. Comprehensive fine layers have been added to the enhanced frequency broadband trace, aligning the real geology found in the well *RC* with the results' veracity while bandlimited has missed this important fine. Therefore, after injecting high frequencies, seismic attributes are extracted, i.e., Average Energy, Normalized Amplitude, *RMS* amplitude, and Sweetness, extracted also using

conventional seismic data, depicted extraordinarily improved results for the *CWT-DNN* based high frequency synthetic volume (*HFSSV*). The identified facies of shale, limestone, wet and gas sands are also depicted on the wells.

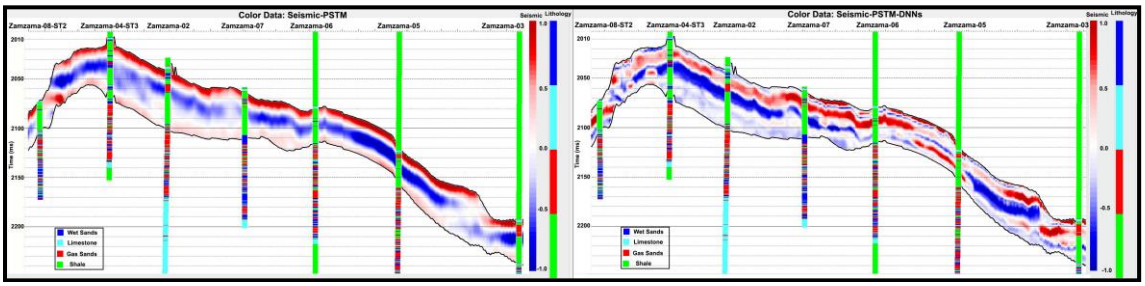



**Figure 4.89** a) Field seismic data shows coarser, smudge and obscure reflection at certain levels b) *CWT-DNNs* high-frequency volume of seismic synthetics overlain by high-frequency facies log reflects fine, coherent, and crisp reflections.

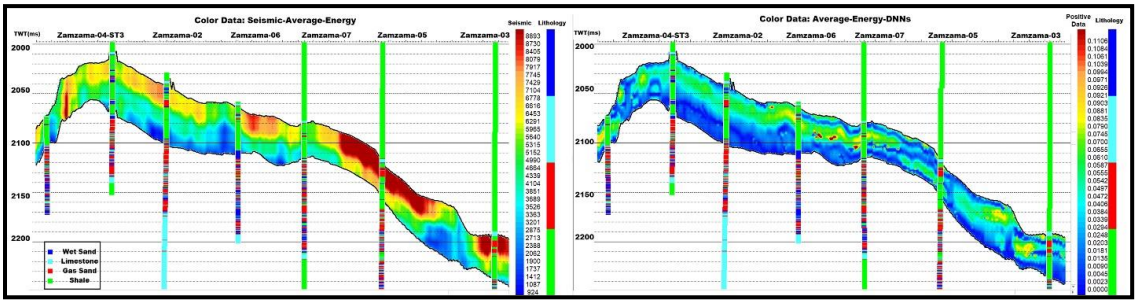



**Figure 4.90** Average energy seismic attribute a) on seismic data set with blocky & coarser resolution, b) *CWT-DNNs* seismic synthetics with enhanced, fine demarcation of beds overlain by facies log.

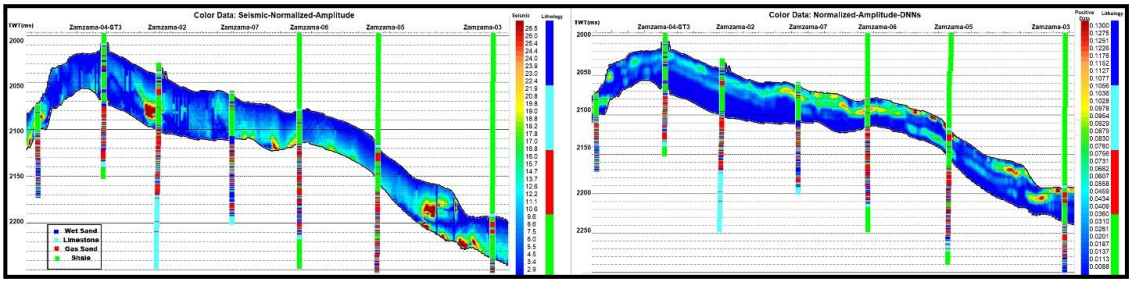



**Figure 4.91** Normalized seismic amplitude attribute a) field seismic data set, b) *CWT-DNNs* based *HFSSV* derived attribute with optimized layering following high-frequency facies layers present at wells.

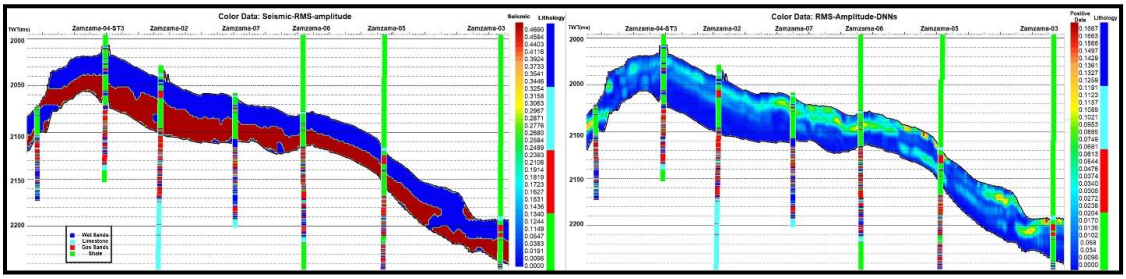



**Figure 4.92** *RMS* seismic attribute a) on seismic data set unable to distinguish fine reservoir layering, b) *CWT-DNNs* derived seismic synthetics attribute introduces high resolution layering match able at wells.

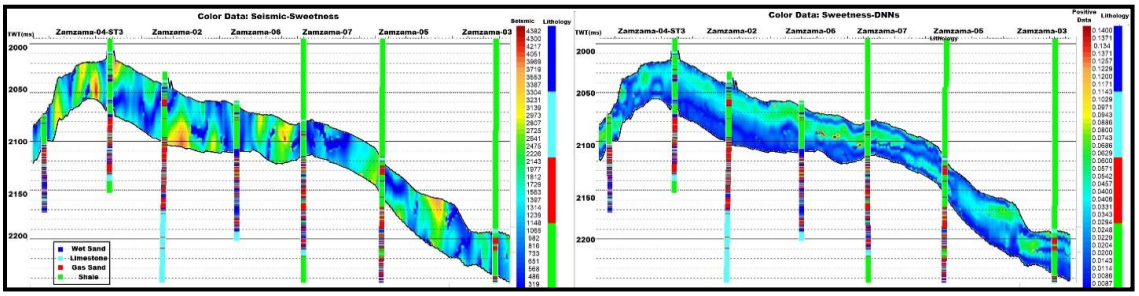



**Figure 4.93** Sweetness seismic attribute a) on seismic data set showing random results, b) *CWT-DNNs* based seismic synthetics attribute marks fine potential layers around the wells.

## 4.5 *PP* Prediction based on Conventional well logs using Advanced *ML* Approach

Pore pressure approximation is considered as a vital property for the effective modeling of reservoir and risk management. Its estimation provides an idea regarding the overpressure formation that hinders the drilling with severe outcomes. Many of the approaches (both conventional and ML based) have been employed to evaluate the pressures scenarios in the Ranikot/Khadro Formation.

### 4.5.1 Conventional Method for *PP* Estimation

*PP* has a crucial significance in drilling safety, efficiency, and cost-effectiveness. It significantly affects the behavior of subsurface materials and has important implications for the stability of rock formations, reservoir management, and geotechnical engineering. In this research, *PP* along with the geo-mechanical properties i.e., overburden pressure, horizontal minimum stress and horizontal maximum stress is estimated in Zamzama-06 using conventional logs i.e., *GR*, *DT* as given in Figure 4.94. These geo-mechanical properties lead to designation of the rock properties like their strength, compactness and thickness of the bed which help to optimize the drilling efficiency by preventing the collapse of wellbore and enlargement of hole.

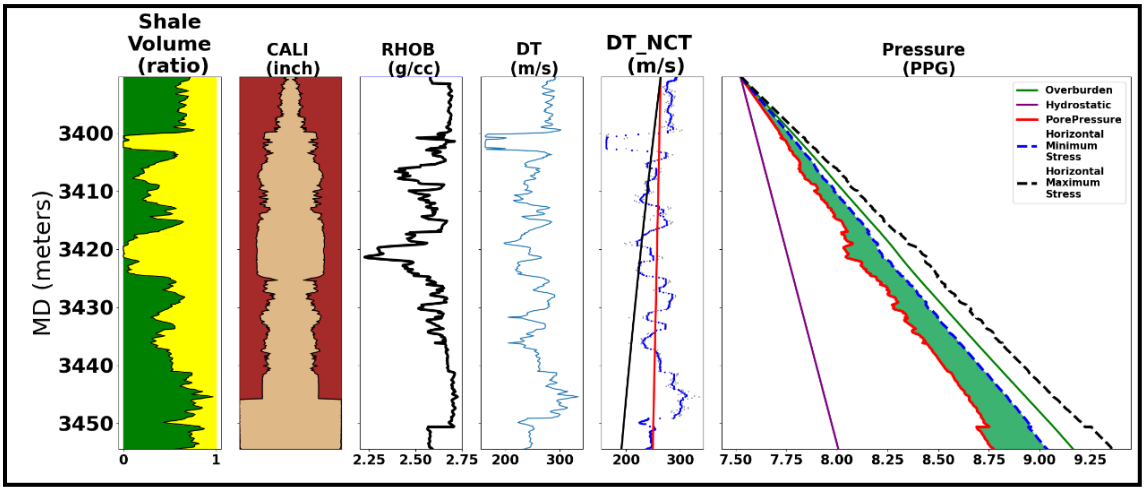



**Figure 4.94** *PP* calculated by *ML* predicted method in Ranikot/Khadro Formation. The pink color curve shows the overburden gradient (*OBG*) while blue curve represents *PP* gradient (*Ppg*) calculated from *DT* using Eq. (31). Maximum and minimum horizontal stress is also estimated, indicated by red and green curve.

### 4.5.2 *1D PP ML* Predictions

Conventional *PP* prediction and analysis techniques, while well-established and industry-proven for decades, suffer from time and labor intensiveness, necessitating specialized skills and equipment. In contrast, *ML*-driven interpretation capitalizes on *AI* and *ML* algorithms to scrutinize large datasets, revealing hidden patterns and relationships that may elude conventional approaches.

*ML* techniques replace traditional approaches to predict geo-mechanical properties with precision. Various *ML* models are tested to identify the optimal one for interpreting well logs at different locations. Employing a heat map helps in the identification of the optimal model for precise predictions on unseen wells. To assess the proficiency of *ML* techniques, the datasets from all seven wells are utilized. One well function as a "blind" well for validation purposes, while the remaining wells are employed for model training. After quality check, raw logs are used to predict geo-mechanical parameters, including *PP*, vertical stress, and horizontal maximum and minimum stress. Among the many algorithms used for training and testing, *GBR* outperforms all others for predicting *PP* and geo-mechanical properties in all blind wells. The different wells of Zamzama Gas Field taken into account for training and testing the model is mentioned in the Table 4.10.

*GBR* represents an ensemble technique comprised of a sequential arrangement of tree models, where each subsequent model homes in on the mistakes made by its forerunner. This method in *ML* employs a process called "boosting" to amalgamate less

potent prediction models, usually in the form of decision trees, into a more resilient and precise model (Rao et al., 2019). By employing the gradient descent loss function, it systematically reduces errors by iteratively updating initial estimations with refined ones. This iterative process culminates in the development of a final model that combines all preliminary estimations with appropriate weights for improved accuracy (Otchere et al., 2022).

**Table 4.10:** List of seven wells used to train and test the data. *ML* algorithm i.e., *GBR* has performed exceptionally well in predicting geo-mechanical properties.

| **Training Wells** | **Blind Well** | **Top *ML* Algorithm** |
|---|---|---|
| Zamzama-03,04,05,06,07,08 | ZZ-02 | *GBR* |
| Zamzama-02,04,05,06,07,08 | ZZ -03 | *GBR* |
| Zamzama-02,03,05,06,07,08 | ZZ -04 | *GBR* |
| Zamzama-02,03,04,06,07,08 | ZZ -05 | *GBR* |
| Zamzama-02,03,04,05,07,08 | ZZ -06 | *GBR* |
| Zamzama-02,03,04,05,06,08 | ZZ -07 | *GBR* |
| Zamzama-02,04,04,05,06,07 | ZZ -08 | *GBR* |

### 4.5.3 Cross Validation

K-fold cross-validation has been extensively implemented practice for evaluating the efficiency of ML models. Here, an automated workflow of random subsampling cross-validation is employed to estimate the competence of the prominent ML algorithm across diverse samples. The K-fold cross-validation divides the initial dataset's random division into K subsets/folds of approximately equivalent size. The K-1 folds are then trained by the model and evaluated on the remaining fold. This step works K times iteratively, with each fold being engaged as the testing dataset once randomly, denoted by (K=i), where i changes from 0 to 9. The computed performance metrics against each iteration are weighed and an overall model performance is calculated by averaging the results across all iterations. This methodology helps prevent overfitting and provides a reliable estimation of the algorithm's performance. In this research, GBR consistently delivered the best results across different samples by testing K=05, 10, and 20, while 10 K Folds produced the optimized results after validation. As 10-fold cross-validation offers a good balance between bias and variance in the performance estimate.

The effectiveness of algorithms was assessed through statistical measures including *MAE*, *RMSE*, $R^2$, among others. The evaluation metrics for all *ML* algorithms portray correlation coefficients and errors for each model, as depicted in Figure 4.95. The analysis highlights that among them, the *GBR* exhibited superior performance in predicting geo-mechanical properties with a correlation coefficient of 0.9.

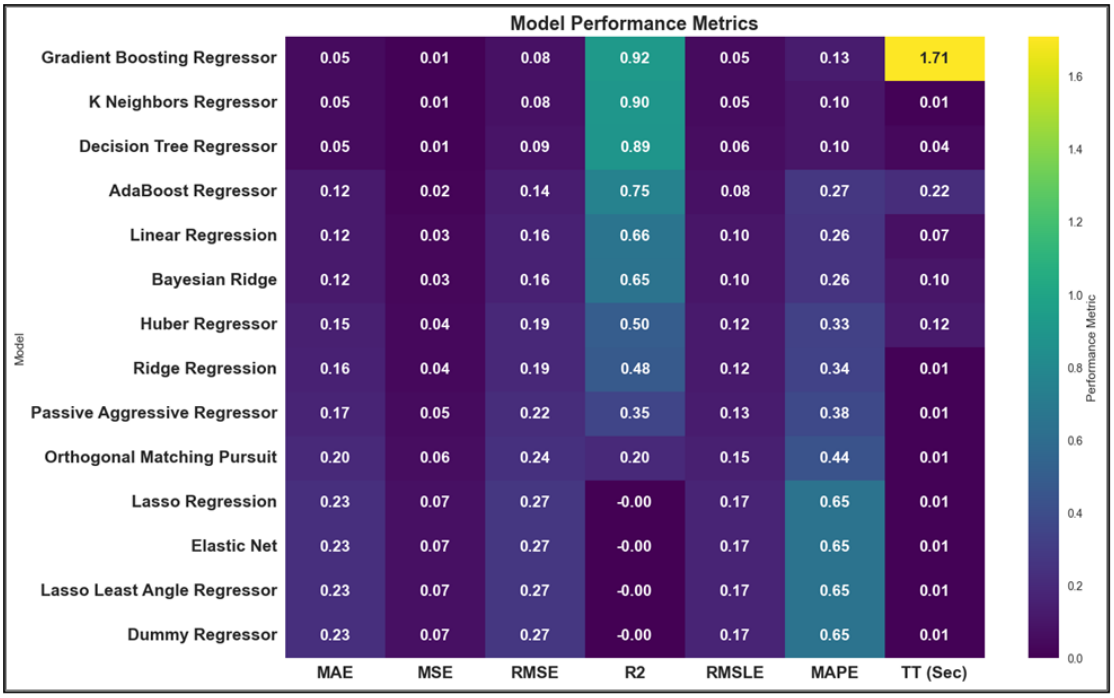



**Figure 4.95** Evaluation metrics for *PP* reflecting the performance of algorithms. Among all the algorithms, *GBR* performed well with correlation of 0.92.

The geo-mechanical properties estimated in all six wells, namely Zamzama-03, Zamzama-04, Zamzama-05, Zamzama-06, Zamzama-07, and Zamzama-08, are subsequently verified using models that exhibited optimal performance on Zamzama-02, a well-considered as a blind test. These six wells were employed for training purposes. The outcomes demonstrate a strong alignment between the projected and actual properties, as illustrated in Figure 4.96, Figure 4.97 and Figure 4.98. The anticipated curves are displayed in black, while the traditional *PP* is shown in dashed black for comparison within the context of Zamzama-02. Additionally, figures also include alternative algorithms like *RF* and light gradient boosting machine (*LGBM*). *RF* predicted curve is depicted with blue color, while predicted curve via *LGBM* algorithm is shown in red color. Among these algorithms, *GBR* shows a pronounced correlation between the forecasted and measured *PP* curve. $R^2$ score of the interpreted *PP* i.e., the correlation

between conventional and *ML* based approach is 0.41 in Zamzama-03 and 0.42 in Zamzama-05 respectively.

Zamzama-03 and Zamzama-05 have two top sand gas facies marked in the red color into which the *PP* has shown the high significance which indicates the potential fluid flow of gas in subsurface formations, suggesting that fluids are either entering or exiting the pores.

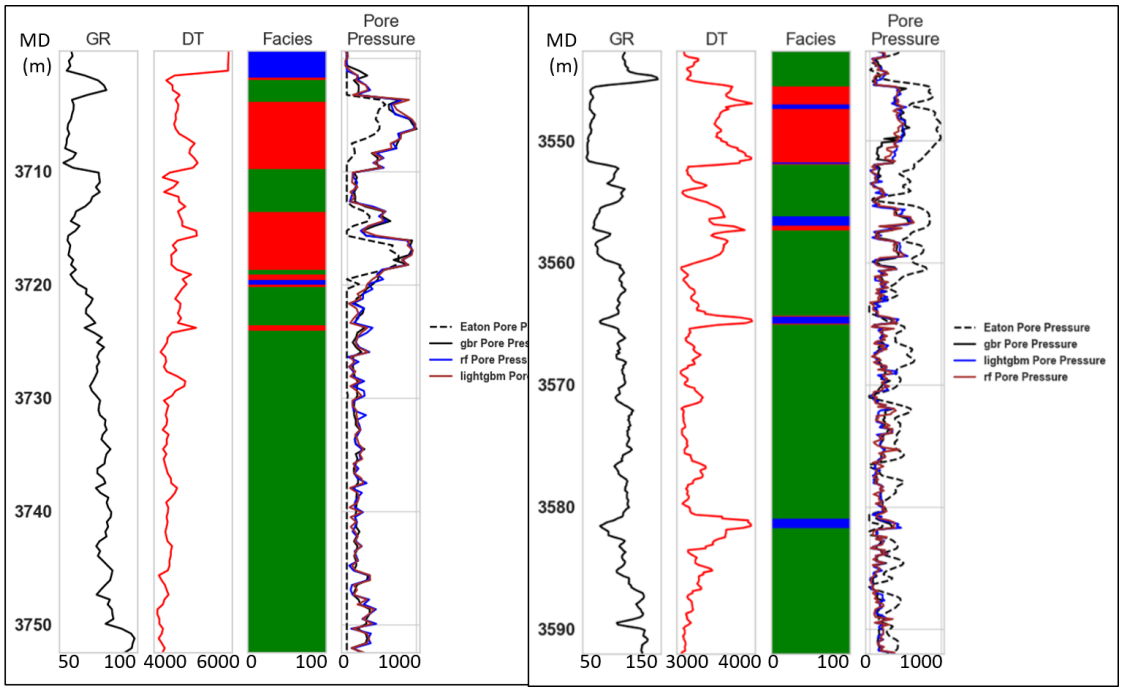



**Figure 4.96** *PP* predictions in Zamzama-03 and Zamzama-05. First two tracks show well log curves i.e., *GR* and *DT*. Third Track shows Facies and Last Track shows the predicted and conventionally interpreted *PP*. Black dotted curve shows the conventionally interpreted *PP* while black solid color depicts predicted *PP* via *GBR*, blue curve via *RF* and red curve predicted using *LGBM* algorithm. Sand gas facies marked in the red color into which the *PP* has shown the high significance.

Zamzama-04 has sand gas facies marked in the red color into which the *PP* has shown the high significance, indicating a potential fluid flow of gas in subsurface formations, as it suggests that fluids are either entering or exiting the pores. Among several algorithms, *GBR* exhibits a significant correlation between the predicted and actual *PP* curve. $R^2$ score of the interpreted *PP* i.e., the correlation between conventional and *ML* based approach is 0.41 in Zamzama-04 and 0.56 in Zamzama-06 respectively.

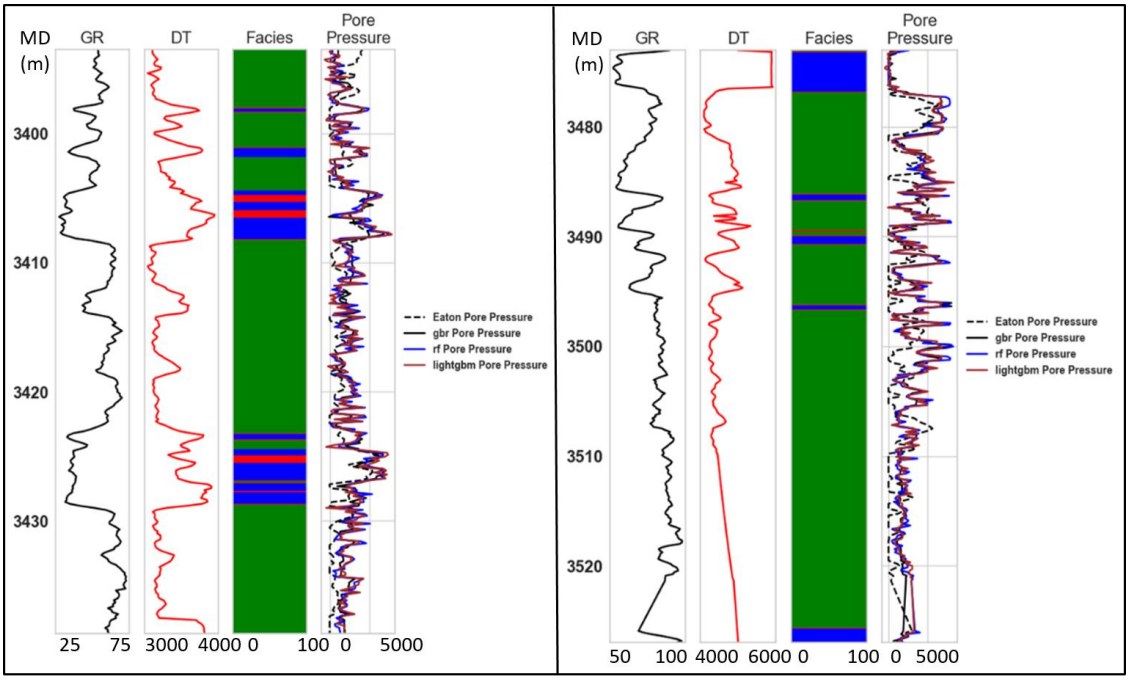



**Figure 4.97** *PP* predictions in Zamzama-04 and Zamzama-06. First two tracks show well log curves i.e., *GR* and *DT*. Third Track shows Facies and Last Track shows the predicted and conventionally interpreted *PP*. Black dotted curve shows the conventionally interpreted *PP* while black solid color depicts predicted *PP* via *GBR*, blue curve via *RF* and red curve predicted using *LGBM* algorithm. Sand gas facies marked in the red color into which the *PP* has shown the high significance.

Zamzama-07 has sand gas facies marked in the red color into which the *PP* has shown the high significance, indicating a potential fluid flow of gas in subsurface formations, as it suggests that fluids are either entering or exiting the pores. Among several algorithms, *GBR* demonstrates a robust correlation between the predicted and observed *PP* curve. $R^2$ score of the interpreted *PP* i.e., the correlation between conventional and *ML* based approach is 0.47 in Zamzama-07 and 0.37 in Zamzama-08 respectively.

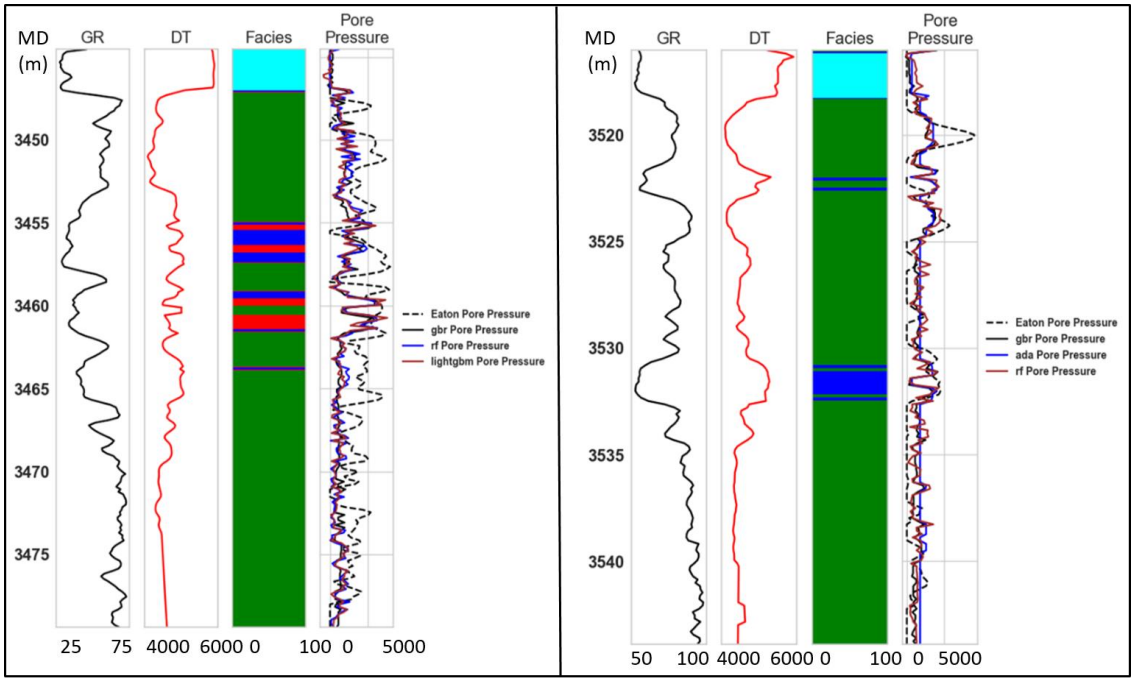



**Figure 4.98** *PP* predictions in Zamzama-07 and Zamzama-08. First two tracks show well log curves i.e., *GR* and *DT*. Third Track shows Facies and Last Track shows the predicted and conventionally interpreted *PP*. Black dotted curve shows the conventionally interpreted *PP* while black solid color depicts predicted *PP* via *GBR*, blue curve via *RF* and red curve predicted using *LGBM* algorithm. Sand gas facies marked in the red color into which the *PP* has shown the high significance.

### 4.5.4 *3D PP ML* Predictions

The spatial variations of *PP* in the thin Ranikot/Khadro Formation across the entire subsurface area, a three dimensional *CWT* is employed for the *3D PP* prediction i.e., The Ranikot/Khadro Formation is thin enough to be visible in conventional seismic data, therefore *CWT-DNNs* model gives higher resolution to capture reservoir gas-bearing facies more distinctly. (Zhang and Zhan, 2017) The relationship is established between *CWT* resultant traces and predicted *PP* on well location through *DNN* and the *PP* or other geo-mechanical properties are stimulated through the *3D* seismic data accurately. The workflow for predicting the uncertainties of *PP* throughout the seismic volume is given in Figure 4.99.

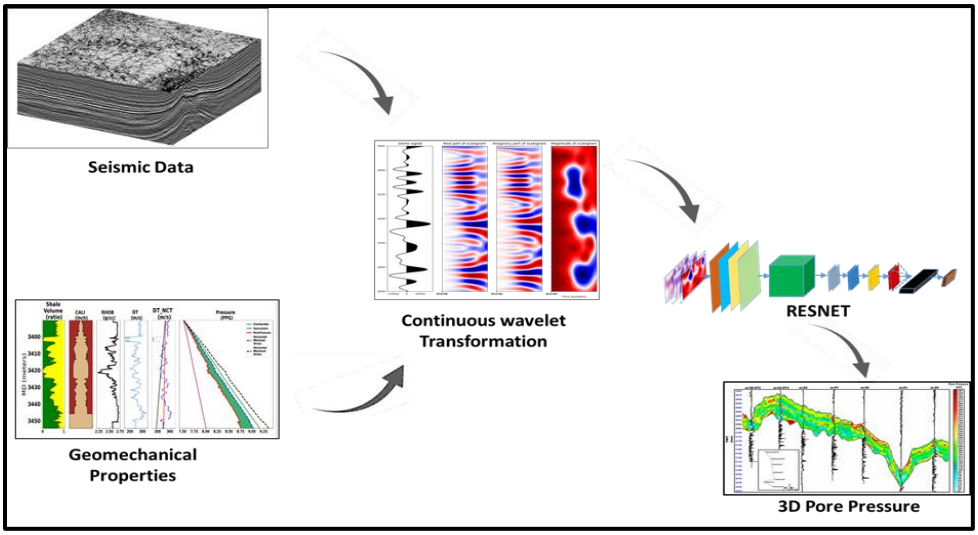



**Figure 4.99** Workflow of the *3D PP* predictions through advanced *ML* model.

In this study, seismic data is transformed into *CWT* features that highlight the hidden information from seismic data by converting it into the frequency domain i.e., real, and imaginary part, which further optimizes the prediction of *PP* throughout the area as shown in Figure 4.100. Likewise, the other geo-mechanical properties like poison ratio, horizontal stress etc. can also be predicted.

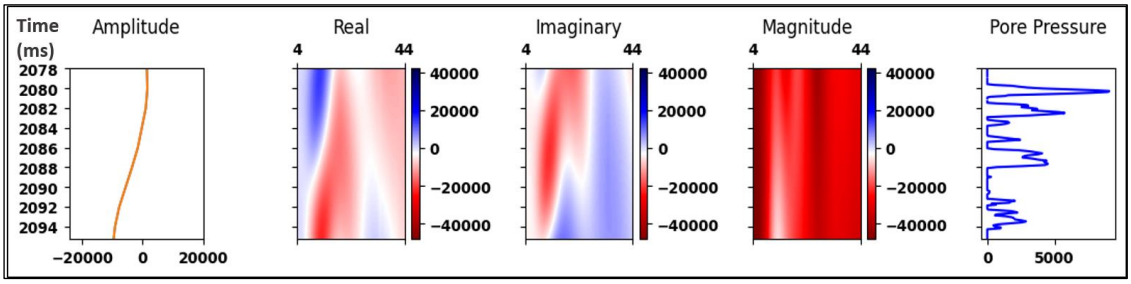



**Figure 4.100** *CWT* of the seismic traces (real, imaginary, and magnitude) in relation with predicted *PP*.

The real part of the *CWT* is fed into the input layer, which is defined by the specified input shape of *CWT* organized in three dimensions of data. Following that, a two dimensional (*2D*) convolutional layer with 64 filters and a 7x7 kernel size is used with padding to keep the spatial dimensions. Each convolutional layer's output undergoes batch normalization and rely activation before being down-sampled using a *2D* max pooling layer with a pool size of 3x3 with strides of 2.

Three residual blocks with 64 filters, a 3x3 kernel without a max-pooling layer and its output is followed by two residual blocks with 128 filters, a 3x3 kernel with strides of 2 with max-pooling layer and its output followed by six residual blocks with 256

filters, a 3x3 kernel with strides of 2 with max pooling layer and its output followed by three residual blocks with 512 filters, a 3x3 kernel with strides of 2 with max-pooling and global average *2D* pooling is applied on the output generated by last residual block to convert the spatial variations in one vector that is fed into output layer having one neuron with no activation function is added to produce the final output as shown in Figure 4.101. Colors in this figure show the convolution layers like red color, cyan and blue color, black hollow box comprises of bunch of different convolutional layers.

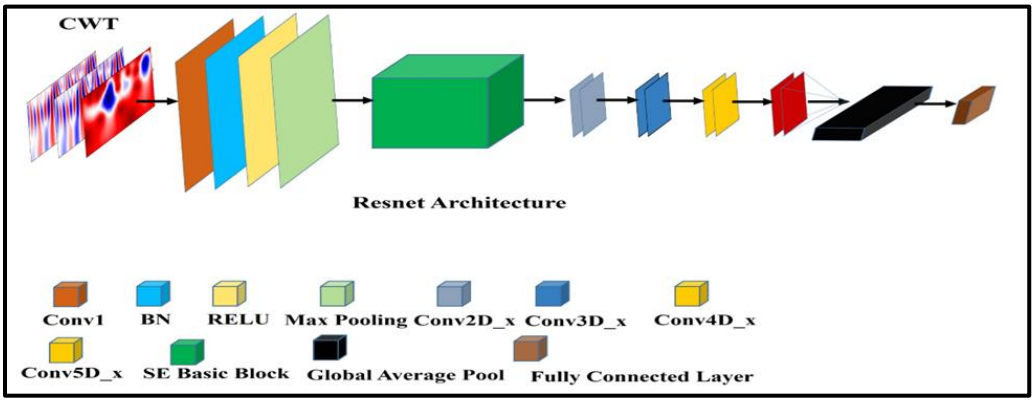



**Figure 4.101** *ResNET* Architecture to predict the *PP*. The multiple colors depict the convolutional layers and black box showing a bunch of different convolutional layers.

Figure 4.102 delineates the *PP* between the wells spatially via *DNN-CWT* relationship, which can also be affirmed through comparison between the predicted and estimated *PP*.

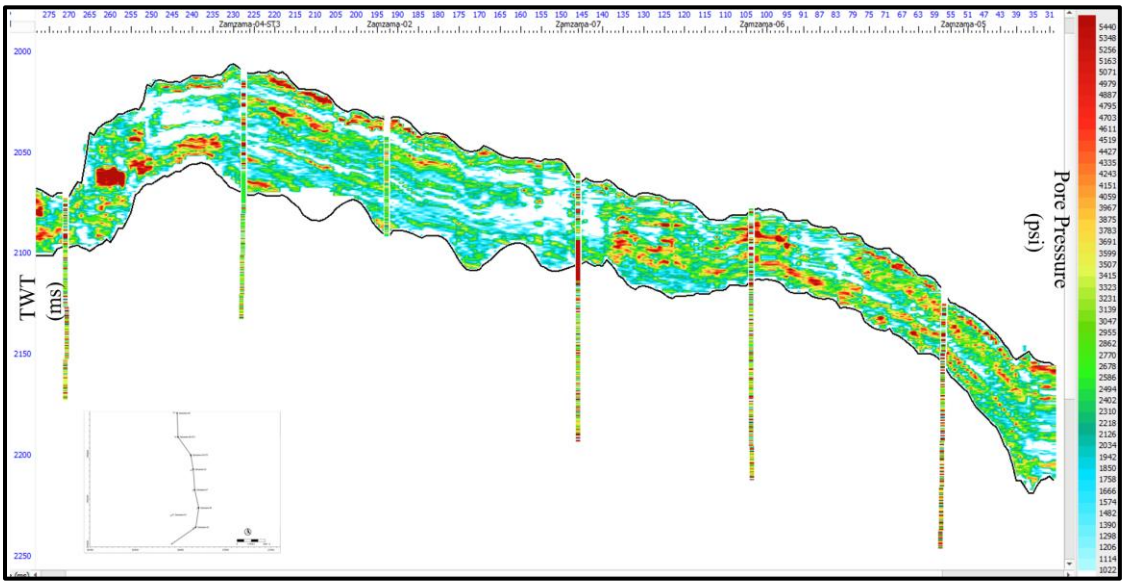



**Figure 4.102** Uncertainty of *PP* predicted between the wells spatially via *DNN-CWT* relationship on arbitrary seismic line passing through the employed wells.

## 4.6 Discussions

The area of research (Zamzama & Mehar Gas Field) lies toward the western section of the Kirthar Foredeep which comprises southern portion of the larger Lower Indus Basin (Ashraf et al., 2020; Ehsan et al., 2018). The Southern Indus Basin of Pakistan is a prolific hydrocarbon producer in Pakistan, bearing complete petroleum elements (Khan et al., 2022; Hussain et al., 2022; Ehsan et al., 2021; Ehsan and Gu, 2020; Ehsan et al., 2019; Mahmood, Ahmad and Ehsan, 2018).The presence of anticline structure provides suitable trap for placement of hydrocarbon in the hanging wall cut through high angle faults (Liu, 2017). A comprehensive petroleum system is available in the *KFTB* comprising Sembar and Goru shales as source and Girdo (Ranikot/Khadro) provides the seal in the region (Abbasi et al., 2016). Due to depositional variability and structural complexity in this basin, understanding reservoir rock, especially Ranikot/Khadro Formation is a significant problem. The Pab Formation is the primary reservoir and a plentiful gas producer in the Zamzama Gas Field (produced 1.7 *TCF* gas and ~9.6 *MMbbls* of condensate) but now in the decline phase. This field also produces from the Lower Ranikot/Khadro Formation, which is in stratigraphic pressure isolation from the Pab Formation. However, most of the production comes from Pab Formation. Moreover, the lateral ramps provided hanging-footwall a connection causing differential encroachment of saline water and created pressure depletion among the producing wells. Gas-Water-Contact (*GWC*) observed at deeper levels of the anticline at crestal portions, create an opportunity to drilled new wells in this field (Zafar et al., 2018). So, the secondary Ranikot/Khadro Formation grabs the attention here, but it is more complex in nature with multiple lithofacies and thin gas-baring sand layers. Although some wells have produced hydrocarbons, most wells did not encounter potential facies due to geological complexity (Munir et al., 2022; Khan et al., 2022). Previously, Jackson et al (2004) worked on the Ranikot/Khadro Formation with limited scope, focusing on a conventional approach that did not incorporate any advanced solutions. Khan et al (2022) and Shakir et al (2021) used Bayesian stochastic seismic inversion (BSSI) and *PSSI* to investigate the Ranikot/Khadro Formation sand layers, but they did not use advanced *ML* approaches.

For complete reservoir characterization of thin sands of Ranikot/Khadro Formations, seven wells, i.e., Zamzama-02, Zamzama-03, Zamzama-04, Zamzama-05, Zamzama-06, Zamzama-07, and Zamzama-08-ST-2, present in 3-Dimentional cube of

the Zamzam gas field, were employed along with three Mehar gas field's well, which are Mehar-01, Mehar-02, and Mehar-03. The well consisted of necessary information regarding formation tops, well logs, mineralogical literature, etc. that was further incorporated for well-to-seismic ties and petrophysical assessment. The analytical comparison of rock physics, petrophysics, seismic attributes, seismic inversion, petrophysical volumetrics and *PP* modeling with several multivariate prediction methodologies based on various *ML* algorithms was made regarding robustness, effectiveness, reliability, and resolution improvement. Many of the advanced *ML* approaches, such as linear regressions, *RFR*, *DTC*, *GBR*, *ETR*, *DTC*, and *DL* (*NNET*), were employed in the estimation of petrophysical and elastic properties (Priezzhev et al., 2019). Seismic and well data sets were integrated into conventional and *ML* procedures, to accurately approximate petrophysical and petro-elastic parameters. For maximum of the approximated properties, it was feasible to estimate the inverted petrophysical reservoir parameters more accurately and with less interpretational bias using the nonlinear predictive operators that *ML* techniques extract. According to Otchere et al (2022), it is an essential part of the *ML* algorithms to have the ability to preprocess the data with the aid of feature selection, along with removing extraneous factors in the trainings and predictions. Otchere et al (2022) preferred to use *GBR* and *RF* on eight feature selection techniques to model and characterize the shallow marine reservoir with better accuracy.

Many of the recent examples are presented based on *ML* techniques i.e., (Yasin et al., 2021a) consolidated rock physics modelling with Post-Stack seismic data, as well as the use of a joint inversion *ML* strategy that combines *SVM* and *PSO*. This technique assisted the author in capturing the thin clastic layers/channels of sand potential bodies while providing a reliable match of sand-shale packages using an arrangement of low-high *AI*. In contrast, the use of the Gaussian simulation technique allowed for the analysis of spatial changes in *3D* porosity models displaying consistency at wells. Finally, the author summarized the findings by predicting sweet spots identified with high porosity, permeability, and gas sands in direct comparison to low elastic property values, i.e., *P*-wave, *S*-wave, and $V_p/V_s$ ratio; inside Lower Goru Formation C-sands. Also, advance ML techniques are employed by recent researcher for reservoir characteristics based on petro-elastic attributes approximation in the Indus Basin (Hussain et al., 2022; Ahmed et al., 2022, Khan et al., 2022; Shakir et al., 2021).

Numerous approaches used in this study concentrates on the field's secondary reservoir, which is the primary exploration target at the present time in the area and nearby blocks, using unsupervised *ML* (One-Class Support Vector Machine) and supervised *ML* algorithms (*ETR*, *DTC*, *RFR*, and *GBR*) for *RPM* and predicting missing/poor logs i.e., density, *S-wave* etc., to properly highlighted Ranikot/Khadro Formation heterogeneous thin gas sands facies. The *RPM* approximated the missing *S-wave* in Zamzama-04, Zamzama-06, and Zamzama-07 based on *ML* techniques and more accurately estimated $Z_s$ using *GBR*. The *GBR* provided accuracy of above 90% and severely minimizing uncertainty along with the allied costs of the procedure. The main principle underlying *GBR* success is creating new base learners that are maximally correlated with the ensemble's negative gradient of the loss function, which is achieved by a sequential error fitting strategy. The accurate prediction of $V_s$ and hence $V_p/V_s$ ratio delineated a reliable relation on petro-elastic crossplot, i.e., x= $Z_p$, y=$V_p/V_s$ ratio, and z=identified lithofacies. The gas-bearing sands having $Z_p$ values of 7500 to 9500 (m/sec*g/cc) along with $V_p/V_s$ ratio of about 1.4 to 1.6 expressed in the crossplot relationship as decisive *ML* characteristics. The same algorithm of GRB is employed for the estimation of elastic and petrophysical volumes that reflected the heterogeneous Ranikot/Khadro Formation more prominently regarding thin potential sands. On the basis of the adopted workflow assisted with *ML* techniques, three units of reservoir sands were allocated regarding Paleocene gas-bearing sand intervals, and gas-bearing sand's probability distribution maps were prepared for each unit of reservoir sand body. Bringing into line the technique of seismic inversion with *ML* cutting-edge algorithms speeded the elastic property approximation regarding facies characterization but also increased the accuracy with reduced uncertainty of the process.

Petrophysics is an important discipline for assessing reservoir characteristics and creating new fields. The application of *ML* has improved drilling efficiency, data correction, reservoir property prediction, reservoir rock type, and other petrophysics-related jobs (Banas et al., 2021; Akkurt et al., 2018) . As an alternative of conventional approaches, advance *ML* i.e., one-class-*SVM*, *RFR*, *ETR*, *GBR*, *DTC* etc. are adopted for the identification of petrophysical properties. The optimal model is selected after assessing various *ML* techniques for a precise interpretation of well logs. Using a heatmap of *ML* algorithm performance, the best model for accurate prediction of blind wells was identified. The formation evaluation process through petrophysics reveals that One-class-*SVM* helped to reduce outliers with great certainty while the missing portions in P-sonic

(*DT*) and *RHOB* logs, employed in the petrophysical estimation, are precisely predicted via *GBR* and *ETR* with 0.66 and 0.88 $R^2$ respectively. Mehar-01's dataset is trained to assess the efficiency of *ML* approaches, while Mehar-02 is kept blind to validate the results. After quality control with supervised and unsupervised learning algorithms, raw logs are used to forecast petrophysical characteristics such as shaly content, *PHIE*, and $S_w$. Several approaches are utilized to train and test the Mehar-01, however the *ETR* has rendered all other algorithms obsolete for predicting shale volume and effective porosities at the blind well Mehar-02. Whereas the *RFR* outperforms all other algorithms in terms of $S_w$ predictions.

The *ETR* operates by taking a vast number of decision trees to predict an average of the parameter of interest (Geurts et al., 2006; Breiman, 2001). Based on these groups, the prediction models are run by performing the computational algorithm. In conclusion, the accuracy of the results is analyzed using a coefficient of determination and mean squared error (Hui et al., 2021). The *RFR* algorithm utilizes a large number of decision trees to predict the average values of a parameter of interest by training the decision trees on random data sets. It is one of the most commonly used algorithms due to its robustness, minimum overfitting risk, and is easily interpretable. This is due to its ability to estimate the relationship between the input variables with the target and subsequently prioritize the correlated features (Otchere et al., 2022). Hence providing reliable and optimized quality logs suitable for *ML* based petrophysics.

The *ML* worked on these augmented logs by dividing the data into 60% training and 40% testing. The *ETR* outperformed rest of models with a correlation of 0.99 and 0.91 among conventional and *ML* results. Likewise, *RFR* performed exceptionally well for $S_w$ modeling expressing the highest 0.93 correlation. Finally, *DTC* modeled reservoir facies with the best 91% accuracy and 0.935 *F1* measures at the blind well. Excellent calibration of >85% is met with the estimates obtained by the predictive model compared to conventional methods. The result of any modelling depends on the purity of the input data. More noise-free data helps produce more optimized results. In petrophysics, conventional methods were also employed so that results can be compared and later remove the noisy data through *ML* approaches. The K Fold cross validation helps in avoiding overfitting and estimate good results (Al-Mudhafar, 2016). The training through *ML* reflects the highest $R^2$ value and least mean square error (*MSE*) for optimized 10 folds authenticating the outputs and validate the opted workflow performance. The feature importance of *ML* describes the input features that have a significant impact on the

model's predictions i.e., parameters having 0% feature relevance are useless for predicting $V_{CL}$. Similarly, the confusion matrix is drawn between each type of facies. In petrophysical estimation 462 values of shale facies (class 0), 118 values of gas sand (class 1), and 12 values of wet sand (class 2) are accurately predicted at training well. Whereas Confusion matrix shows accurately predicted 72 values of shale facies, 24 values of gas sand and 2 values of wet sand along with 100 % accuracy while only 10 values were falsely predicted for wet sand at blind well. Similarly, missing sonic logs in the splice zone and density logs due to bad holes cause misleading results (Shakir et al., 2021; Hussain et al., 2017).

According to the previous results of *ML* attained by various researchers among the current study, it can be used as an analogue for the regional development of thin-bedded sandstone systems in the vast basins globally. The current research may serve as a vital example for the exploration of the remaining gas-bearing stratigraphic systems within the studied basin and similar surrounding basinal settings. As a result, investigating the reservoir facies by the integration of seismic signature and well log's high frequency content through *ML* technique highlighted the Paleocene Ranikot/Khadro Formation thin-bedded secondary reservoir for effective reservoir quality prediction. The *ML* based elastic attributes, i.e., $Z_p$, $Z_s$, and petrophysical properties (*PHIE*) along with the gas bearing sand probability distribution contributed effectively to the overall understanding of the Paleocene hydrocarbon plays.

The results of the facies classification aided by the high-resolution elastic properties approximated by *ML* technique are highly effective and promising to exploit the untapped and unexplored regions of the Indus Basin. The generated probability distribution maps can spot the light and direct the exploitation of these reservoirs on a regional and subsurface scale. Therefore, the identification of the best fit *ML* could improve reservoir prediction and simulation for reservoir management and recovery in such heterogeneous thin-bedded sandstone reservoirs. In terms of physical property prediction, our analysis indicates that the sandstone reservoirs can be classified into 3 units that can be traced along the studied area. These three levels, illuminated by the generated maps of elastic, petrophysical, and facies, can be the basis for field development and better exploitation of these thin-bedded secondary reservoirs that was beyond the seismic resolution of conventional seismic data. The class of sands is initially assessed through the *AVO* analysis of high frequency well data that demarcated the sands as class-2 revealing its exploration importance.

In nutshell, this research highlights effectiveness of *ML* models in removing outliers from raw well logs and modeling the missing and bad well logs in an optimized and systematic ways. The *ML* has the ability to provide more accurate and reliable predictions compared to traditional techniques, making it a promising tool for the industry after testing and validation with excellent results. However, *ML*-based methods require large amounts of high-quality training data, and the accuracy of predictions depends on the quality of the training data and the complexity of the model.

For evaluating seismic resolution limitation, the synthetics wedge model was constructed at available wells to understand seismic resolution limits and resulted that field seismic fails to resolute thin ~ 6 m sands. Therefore, the aim was to increase the field seismic data itself by increasing its frequency band width through *DL* along with the integration of *CWT*. High frequencies have been injected carefully for attaining high resolution seismic which is then populated throughout the *3D* cube and a high-frequency volume is obtained. This high frequency synthetic seismic volume (*HFSSV*) is obtained by the *CWT-DNN* integrated approach and thereby the enhanced attribute analysis was carried out. It further validated that the developed algorithm by correlating the original field results exhibited improved results confirmed by the production at wells.

The field seismic and its extracted attributes reveal blocky, smeared, and ambiguous images at certain levels, which mostly failed to highlight thin gas sand layers as in this study. The geological information extracted from these extracted seismic attributes has not matching the overlain well information and did not produce efficient results. Whereas, seismic and its corresponding attributes after injecting high-frequency content uncover fine layers by exhibiting logical, consistent, and sharp reflections matching the high-resolution facies log.

The *HFSSV* attribute results are matching at the producing and non-producing facies at well locations passing through the arbitrary line. It comprehensively shows a clear difference and optimization of the results produced with application of this techniques. Likewise, incorporation of the more available well data, core data, and other supportive studies data would be valuable and enhance the subsurface geological perspective in a more vibrant way. As a result, incorporating these enhanced and additional datasets with in advance integrated techniques would be the study's future recommendations for further research.

Finally, the *ML* technique assessed thin sand beds to induce high frequencies from well data into the band-limited seismic data that increased the vertical resolution and

hence became able to resolve these thin sands. Innovative *ML* techniques proved its capability of providing broadband seismic attributes therefore minimized the requirement of introducing high frequencies. *ML* identified features by employing well-data, trained the property, and approximate the consistent features using the traces of seismic data, hence successfully permitting to assess the beds present below seismic resolving power among structurally complicated area.

*PP* is one of the significant properties of the reservoir in terms of safe drilling, efficiency of the recovery, and well bore stability. It significantly affects the behavior of subsurface materials and has important implications for the stability of rock formations, reservoir management, and geotechnical engineering (Zhang and Zhan, 2017). In the study area, *PP* along with the geo-mechanical properties i.e., overburden pressure, horizontal minimum stress and horizontal maximum stress is estimated using conventional logs i.e., *GR*, *DT*. These geo-mechanical properties lead to designation of the rock properties like their strength, compactness and thickness of the bed which help to optimize the drilling efficiency by preventing the collapse of wellbore and enlargement of hole. *ML* techniques replace traditional approaches to predict geo-mechanical properties with precision. Various *ML* models were tested to identify the optimal one for interpreting well logs at different locations. The heat map aided in the selection of best model at blind wells. In every process, six wells out of seven were selected in the training while keeping one as blind. Among many algorithms used for training and testing, *GBR* outperforms all others for predicting *PP* and geo-mechanical properties especially at blind well location.

The execution of K-fold cross-validation (process iterates K times) entails the initial dataset's random division into K subsets or folds of approximately equivalent size. Performance metrics are computed for each iteration, and an overall assessment of model performance is derived by averaging the outcomes across all iterations. This approach helps avoid overfitting and provides a reliable estimation of the algorithm's performance. In this research, *GBR* consistently delivered the best results across different samples, validated by 10 K Folds. The effectiveness of various applied *ML* algorithms was assessed through statistical measures including *MAE*, *RMSE*, $R^2$ that suggested *GBR* as one of the best approaches for *PP* and its mechanical properties estimation. The outcomes demonstrated a strong alignment between projected and actual properties using *GBR* in comparison with *RF*, *LGBM*, etc. For the *3D PP* prediction i.e., the spatial variations of *PP* in the thin Ranikot/Khadro Formation across the entire subsurface area, a *3D CWT* is

employed. The Ranikot/Khadro Formation is thin enough to be visible in conventional seismic data, therefore *CWT-DNNs* model gives higher resolution to capture reservoir gas-bearing facies more distinctly (Zhang and Zhan, 2017). The relationship is established between *CWT* resultant traces and predicted *PP* on well locations through *DNN* and the *PP* or other geo-mechanical properties are stimulated through the *3D* seismic data accurately. The real part of the *CWT* is fed into the input layer, which is defined by the specified input shape of *CWT* organized in three dimensions of data. Following that, a *2D* convolutional layer with 64 filters and a 7x7 kernel size is used with padding to keep the spatial dimensions. Each convolutional layer's output undergoes batch normalization and rely activation before being down-sampled using a *2D* max pooling layer with a pool size of 3x3 with strides of 2. Three residual blocks with 64 filters, a 3x3 kernel without a max-pooling layer and its output is followed by two residual blocks with 128 filters, a 3x3 kernel with strides of 2 with max-pooling layer and its output followed by six residual blocks with 256 filters, a 3x3 kernel with strides of 2 with max pooling layer and its output followed by three residual blocks with 512 filters, a 3x3 kernel with strides of 2 with max-pooling and global average *2D* pooling is applied on the output generated by last residual block to convert the spatial variations in one vector that is fed into output layer having one neuron with no activation function is added to produce the final output. The convolution layers are presented with various colors like red, cyan and blue color, black hollow box etc. in this research work.

The results indicate that the data-driven *ML* approach for estimating *PP* is both reliable and efficient compared to conventional methods. This approach only requires conditioned well logs and *ML* algorithms to accurately determine *PP* at well locations. By utilizing *CWT* of Post-Stack seismic data, geologists and geophysicists can obtain a *3D PP* volume through this data-driven approach, facilitated by *ML* algorithms. In contrast, the conventional methods for *PP* estimation have been challenging.

## 4.7 Summary

The integrated approach optimized the geological outputs and clarified the heterogeneous thin potential reservoir sands. The petrophysical and elastic properties of the plausible sands are summarized in Table 4.10.

**Table 4.10:** Key reservoir properties assessed through employed techniques.

| **Property** | **Value of reservoir sands** |
|---|---|
| Petrophysics (Mehar gas field) | $V_{SH}$ (<25%), $S_W$ (<40%), *PHIE* (10%) |
| Petrophysics (Zamzama gas field) | $V_{SH}$ (<20%), $S_W$ (<0.35%), *PHIE* (10%) |
| *ML* based P-impedance | <10000 m/s*g/cc |
| *ML* based S-impedance | <4500 /s*g/cc |
| *ML* based *PHIE* | ≈9% |
| Gas sand Facies probability | >75% |
| *ML* based *PP* prediction | ≈4500-5000 psi |

# CHAPTER 5

# CONCLUSIONS AND RECOMENDATIONS

## 5.1 Conclusion

The thesis significantly demonstrated the advanced novel approaches of *ML* implemented successfully for the characterization of thin heterogeneous Ranikot/Khadro reservoir sands. The key outcomes of the employed techniques are described as follows:

(1) Petro-elastic relationship has been comprehensively developed through *ML* for optimized understanding of the thin heterogeneous sands. Four key facies have been demarcated successfully including shales, wet sands, gas-sands, and limestone within the Ranikot/Khadro Formation where the thin reservoir sands (4 –7 m in thickness) were par below seismic resolution but have the potential for hydrocarbon production. The high probabilities gas-sands zones have been demarcated for further field development.

(2) The reservoir sand layers are classified into Sand-01, Sand-02, and Sand-03 according to petrophysical characteristics and these were penetrated in all studied wells except Zamzama-06 and Zamzama-08-ST-02. These potential gas sand bodies were extended and evaluated beyond the known well location and across the field, marking them as the primary objective of the research with the implementation of advanced *ML* algorithms and techniques. To resolve the thin-bed reservoirs and obtain the best possible results, we have applied four *ML* approaches including *SVM, DTR, RF,* and *GBR*. All aforementioned algorithms have excellent results, but the *GBR* outperforms the other algorithms based on the comparative maximum correlation and $R^2$ attained i.e., on average above 90%.

(3) The study presents a competent approach that highlights the effectiveness of using *ML* techniques for well logs outlier removal, prediction of missing logs, and petrophysical interpretation. In the context of well log data conditioning, the one Class-*SVM* algorithm emerged as the most effective, including various algorithms employed in handling outliers. The *GBR* performed best with $R^2$ 0.66 and MSE 0.0034 to predict the missing splice zone *DT* log, while the *ETR* had $R^2$ 0.88 and MSE 0.0005 to optimize the bad hole *RHOB* log, which are essential inputs for accurate *ML* results. Among the multiple algorithms considered, the *ETR* algorithm demonstrated excellent performance with $R^2$ > 99.9% and 0 MSE in modelling shale volume and *PHIE,* while the *RFR* algorithm had $R^2$ =99.9% with

0 MSE and remarkably predicted $S_w$ at the training well. While the strong correlations of 0.99, 0.91, and 0.93 were observed between predicted and conventionally measured shale volume, *PHIE*, and $S_w$ respectively, after thorough testing and validation, using Mehar-02 as a blind well. Similarly, key reservoir facies have been predicted with high correlation at blind well along with excellent statistical measures. The results mark the effectiveness of *ML* models in petrophysical interpretation and data quality control, surpassing conventional approaches.

(4) Likewise, the *DNNs* proved to be a flexible and more accurate *AI* approach. The workflow built an effective relationship with *CWT*-derived seismic properties for obtaining high-frequency volumes (*HFV*). The *HFV* based on *CWT-DNNs* extremely improves the seismic resolution evident by the comparison of synthetic wedge models i.e., conventional seismic and *CWT-DNNs* volumes. The attributes obtained from HFV are more pronounced for the thin Khadro Formation's reservoir sands. This method produced high-resolution elastic ($Z_p$, $Z_s$, $V_p/V_s$ ratio) along with petrophysical properties ($V_{CL}$, *PHIE*, gas-sand facies) for the whole volume after validation at the blind wells. The *CWT-DNNs* approach is typically useful for areas where acquiring high-resolution seismic is challenging or cost-prone. The *AI*-guided *DNN* exploited the untapped potential in acquired seismic volumes for improved reservoir characterization with low uncertainty.

(5) This study concludes an efficient approach that demonstrates the efficacy of employing *ML* techniques for predicting pore-pressure. Among the various models developed, the *GBR* algorithm proved to be the most effective, achieving an $R^2$ of 0.92 for *PP* prediction using conditioned well log data. Furthermore, these predictions were utilized as input, and the entire seismic volume was populated using the *CWT-DNN* relationship to accurately predict the spatial distribution of *PP*. The results indicate that the data-driven *ML* approach for precise estimating *PP* is both reliable and efficient compared to conventional methods.

## 5.2 Recommendations

The study results are highly promising, and the blind wells corroborate the findings. As a result, a complete dataset should be used, and the study's findings should be expanded to the regional level for optimal well development.

*ML* demonstrated an optimal integrated method for improved results. The results indicate that the high probability zones should be investigated for enhancement of hydrocarbon exploitation in a heterogeneous, thin reservoir.

The Signal-to-noise ratio of the available seismic dataset need to be optimized by reprocessing the seismic data with the latest technology and optimized parameters.

The shear wave component in well logging is very important log for fluid demarcation along with others logs that need to acquire in all future wells.

In the future, seismic and well-log data should be collected by focusing on thin beds with high resolution. As a result, the issues associated with this complicated reservoir must be addressed in order to resolve thin potential layers.

# REFERENCES


Abad, A.R.B. et al. (2021) 'Predicting oil flow rate through orifice plate with robust machine learning algorithms', *Flow Measurement and Instrumentation*, 81, p. 102047.

Abbasi, S.A. et al. (2016) 'Study of fault configuration related mysteries through multi seismic attribute analysis technique in Zamzama gas field area, southern Indus Basin, Pakistan', *Geodesy and Geodynamics*, 7(2), pp. 132–142.

Abdelaal, A., Elkatatny, S. and Abdulraheem, A. (2021) 'Data-driven modeling approach for pore pressure gradient prediction while drilling from drilling parameters', *ACS omega*, 6(21), pp. 13807–13816.

Aghli, G. et al. (2016) 'Fractured zones detection using conventional petrophysical logs by differentiation method and its correlation with image logs', *Journal of Petroleum Science and Engineering*, 142, pp. 152–162.

Ahmad, N. et al. (2013) 'Shale gas potential of lower Cretaceous Sembar Formation in middle and lower Indus basin, Pakistan', *Pakistan Journal of Hydrocarbon Research*, 23, pp. 51–62.

Ahmadi, M.-A. et al. (2014a) 'Connectionist model predicts the porosity and permeability of petroleum reservoirs by means of petro-physical logs: application of artificial intelligence', *Journal of Petroleum Science and Engineering*, 123, pp. 183–200.

Ahmadi, M.-A. et al. (2014b) 'Connectionist model predicts the porosity and permeability of petroleum reservoirs by means of petro-physical logs: application of artificial intelligence', *Journal of Petroleum Science and Engineering*, 123, pp. 183–200.

Ahmadi, M.A. and Chen, Z. (2019) 'Comparison of machine learning methods for estimating permeability and porosity of oil reservoirs via petro-physical logs', *Petroleum*, 5(3), pp. 271–284.

Ahmed, R. and Ali, S.M. (1991) 'Tectonic and structural development of the eastern part of Kirthar Fold Belt and its hydrocarbon prospects', *Pakistan Journal of Hydrocarbon Research*, 3(2), pp. 19–31.

Ahmed, S.A., Hussain, M. and Khan, Z.U. (2022) 'Supervised machine learning for predicting shear sonic log (DTS) and volumes of petrophysical and elastic attributes, Kadanwari Gas Field, Pakistan', *Frontiers in Earth Science*, 10, p. 919130.

Aki, K. (1980) 'Quantitative seismology', *Theory and models.*, 1, p. 932p.

Akkurt, R. et al. (2018) ‘Accelerating and enhancing petrophysical analysis with machine learning: a case study of an automated system for well log outlier detection and reconstruction’, in *SPWLA Annual Logging Symposium*. SPWLA, p. D043S008R004.

Akram, S. et al. (2024) 'ResNet and CWT Fusion: A New Paradigm for Optimized Heterogeneous Thin Reservoir Evaluation', ACS omega, 9, pp. 4775-4791.

Al Khalifah, H., Glover, P.W.J. and Lorinczi, P. (2020) ‘Permeability prediction and diagenesis in tight carbonates using machine learning techniques’, *Marine and Petroleum Geology*, 112, p. 104096.

Alam, M. et al. (2003) ‘An overview of the sedimentary geology of the Bengal Basin in relation to the regional tectonic framework and basin-fill history’, *Sedimentary geology*, 155(3–4), pp. 179–208.

Al-Anazi, A. and Gates, I.D. (2010) ‘A support vector machine algorithm to classify lithofacies and model permeability in heterogeneous reservoirs’, *Engineering Geology*, 114(3–4), pp. 267–277.

Ali, A. et al. (2018) ‘Resource potential of gas reservoirs in South Pakistan and adjacent Indian subcontinent revealed by post-stack inversion techniques’, *Journal of Natural Gas Science and Engineering*, 49, pp. 41–55.

Ali, M. et al. (2021) ‘Machine learning-A novel approach of well logs similarity based on synchronization measures to predict shear sonic logs’, *Journal of Petroleum Science and Engineering*, 203, p. 108602.

Ali, Muhammad et al. (2019) ‘Petrophysical analysis of well logs for reservoir evaluation: A case study of “Kadanwari” gas field, middle Indus basin, Pakistan’, *Arabian Journal of Geosciences*, 12, pp. 1–12.

Alloghani, M. et al. (2020) ‘A systematic review on supervised and unsupervised machine learning algorithms for data science’, *Supervised and unsupervised learning for data science*, pp. 3–21.

Al-Mudhafar, W.J. (2016) ‘Incorporation of bootstrapping and cross-validation for efficient multivariate facies and petrophysical modeling’, in *SPE Rocky Mountain Petroleum Technology Conference/Low-Permeability Reservoirs Symposium*. SPE, p. SPE-180277-MS.

Al-Mudhafar, W.J. (2020) ‘Integrating machine learning and data analytics for geostatistical characterization of clastic reservoirs’, *Journal of Petroleum Science and Engineering*, 195, p. 107837.

Amigun, J.O., Olisa, B. and 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), pp. 181–187.

Anifowose, F.A., Labadin, J. and Abdulraheem, A. (2017) 'Ensemble machine learning: An untapped modeling paradigm for petroleum reservoir characterization', *Journal of Petroleum Science and Engineering*, 151, pp. 480–487.

Ashraf, U. et al. (2020) 'Controls on reservoir heterogeneity of a shallow-marine reservoir in Sawan Gas Field, SE Pakistan: Implications for reservoir quality prediction using acoustic impedance inversion', *Water*, 12(11), p. 2972.

Ashraf, U. et al. (2021) 'A core logging, machine learning and geostatistical modeling interactive approach for subsurface imaging of lenticular geobodies in a clastic depositional system, SE Pakistan', *Natural Resources Research*, 30, pp. 2807–2830.

Asim, S. et al. (2014) 'Structural and Stratigraphical correlation of seismic profiles between Drigri anticline and Bahawalpur high in central Indus Basin of Pakistan', *International Journal of Geosciences*, 5(11), p. 1231.

Avseth, P. et al. (2010) 'Rock-physics diagnostics of depositional texture, diagenetic alterations, and reservoir heterogeneity in high-porosity siliciclastic sediments and rocks—A review of selected models and suggested work flows', *Geophysics*, 75(5), pp. 75A31-75A47.

Avseth, P. and Skjei, N. (2011) 'Rock physics modeling of static and dynamic reservoir properties—A heuristic approach for cemented sandstone reservoirs', *The Leading Edge*, 30(1), pp. 90–96.

Avseth, P.A. and Odegaard, E. (2004) 'Well log and seismic data analysis using rock physics templates', *First break*, 22(10).

Ayodele, T.O. (2010) 'Types of machine learning algorithms', *New advances in machine learning*, 3, pp. 19–48.

Azeem, T. et al. (2017) 'An integrated petrophysical and rock physics analysis to improve reservoir characterization of Cretaceous sand intervals in Middle Indus Basin, Pakistan', *Journal of Geophysics and Engineering*, 14(2), pp. 212–225.

Babasafari, A.A. et al. (2020) 'A new approach to petroelastic modeling of carbonate rocks using an extended pore-space stiffness method, with application to a carbonate reservoir in Central Luconia, Sarawak, Malaysia', *The Leading Edge*, 39(8), pp. 592a1-592a10.

Bahmaei, Z. and Hosseini, E. (2020) ‘Pore pressure prediction using seismic velocity modeling: case study, Sefid-Zakhor gas field in Southern Iran’, *Journal of Petroleum Exploration and Production Technology*, 10, pp. 1051–1062.

Banas, R., McDonald, A. and Perkins, T.J. (2021) ‘Novel methodology for automation of bad well log data identification and repair’, in *SPWLA Annual Logging Symposium*. SPWLA, p. D031S022R003.

Banks, C.J. and Warburton, J. (1986) ‘‘Passive-roof’duplex geometry in the frontal structures of the Kirthar and Sulaiman mountain belts, Pakistan’, *Journal of structural Geology*, 8(3–4), pp. 229–237.

Bannert, D. and Raza, H.A. (1992) ‘The segmentation of the Indo-Pakistan plate’, *Pakistan Journal of Hydrocarbon Research*, 4(2), pp. 5–18.

Barclay, F. et al. (2008) ‘Seismic inversion: Reading between the lines’, *Oilfield Review*, 20(1), pp. 42–63.

Batta, I. et al. (2022) ‘SVR-based Multimodal Active Subspace Analysis for the Brain using Neuroimaging Data’, *bioRxiv*, p. 2022.07. 28.501879.

Bender, F. and Raza, H.A. (1995) ‘Geology of Pakistan’.

Besse, J. and Courtillot, V. (1988) ‘Paleogeographic maps of the continents bordering the Indian Ocean since the Early Jurassic’, *Journal of Geophysical Research: Solid Earth*, 93(B10), pp. 11791–11808.

Biot, M.A. (1941) ‘General theory of three-dimensional consolidation’, *Journal of applied physics*, 12(2), pp. 155–164.

Bisht, B.S. et al. (2013) ‘Integration of petrophysics & rock-physics modeling in single workflow reduces uncertainty in seismic reservoir characterization: a case study’, *Geohorizons*, pp. 44–47.

Björk, G. et al. (2010) ‘Flow of Canadian basin deep water in the Western Eurasian Basin of the Arctic Ocean’, *Deep Sea Research Part I: Oceanographic Research Papers*, 57(4), pp. 577–586.

Breiman, L. (2001) ‘Random forests’, *Machine learning*, 45, pp. 5–32.

Busby, C. and Pérez, A.A. (2011) *Recent Advances in the: Tectonics of Sedimentary Basins*. Wiley-Blackwell.

Carrazzone, J.J. et al. (1996) *Method for deriving reservoir lithology and fluid content from pre-stack inversion of seismic data*. Google Patents.

Caruana, R. and Niculescu-Mizil, A. (2006) 'An empirical comparison of supervised learning algorithms', in *Proceedings of the 23rd international conference on Machine learning*, pp. 161–168.

Castagna, J.P., Batzle, M.L. and Eastwood, R.L. (1985) 'Relationships between compressional-wave and shear-wave velocities in clastic silicate rocks', *geophysics*, 50(4), pp. 571–581.

Castagna, J.P. and Swan, H.W. (1997) 'Principles of AVO crossplotting', *The leading edge*, 16(4), pp. 337–344.

Chapelle, O., Schölkopf, B., & Zien, A. (Eds.) (2006). *Semi-supervised learning*. MIT press.

Chen, Q. and Sidney, S. (1997) 'Seismic attribute technology for reservoir forecasting and monitoring', *The Leading Edge*, 16(5), pp. 445–448.

Chen, J., Li, X., Lei, Y., & Jiang, J. (2020). Random forest for mineral prospectivity mapping using geochemical data and remote sensing imagery. *Remote Sensing*, 12(13), 2112. https://link.springer.com/article/10.1007/s11053-019-09510-8

Choi, Y. and Alkhalifah, T. (2011) 'Source-independent time-domain waveform inversion using convolved wavefields: Application to the encoded multisource waveform inversion', *Geophysics*, 76(5), pp. R125–R134.

Chopra, S. and Marfurt, K.J. (2005) 'Seismic attributes—A historical perspective', *Geophysics*, 70(5), pp. 3SO-28SO.

Clochard, V. et al. (2009) 'Post-stack versus pre-stack stratigraphic inversion for CO 2 monitoring purposes: a case study for the saline aquifer of the Sleipner field', in *SEG Technical Program Expanded Abstracts 2009*. Society of Exploration Geophysicists, pp. 2417–2421.

Darvishpour, A. et al. (2019) 'Wellbore stability analysis to determine the safe mud weight window for sandstone layers', *Petroleum Exploration and Development*, 46(5), pp. 1031–1038.

Dolan, P. (1990) 'Pakistan: a history of petroleum exploration and future potential', *Geological Society, London, Special Publications*, 50(1), pp. 503–524.

Dou, Q., Sun, Y. and Sullivan, C. (2009) 'Rock-physics-based heterogeneity characterization of a carbonate reservoir in the Permian Basin', in *SEG International Exposition and Annual Meeting*. SEG, p. SEG-2009-1945.

Downton, J.E. (2005) *Seismic parameter estimation from AVO inversion*. University of Calgary, Department of Geology and Geophysics.

Dreiseitl, S. et al. (2010) ‘Outlier detection with one-class SVMs: an application to melanoma prognosis’, in *AMIA annual symposium proceedings*. American Medical Informatics Association, p. 172.

Durrani, M.Z.A. et al. (2022) ‘Characterization of seismic anisotropy using azimuthal AVO analysis (AVAz)-An application case study in the deep and tight carbonate reservoirs from Potwar Basin onshore Pakistan’, *Journal of Applied Geophysics*, 205, p. 104767.

Eaton, B.A. (1975) ‘The equation for geopressure prediction from well logs’, in *SPE Annual Technical Conference and Exhibition?* SPE, p. SPE-5544-MS.

Ehsan, M. et al. (2018) ‘A geological study of reservoir formations and exploratory well depths statistical analysis in Sindh Province, Southern Lower Indus Basin, Pakistan’, *Kuwait Journal of Science*, 45(2).

Ehsan, M. et al. (2019) ‘A modified approach for volumetric evaluation of shaly sand formations from conventional well logs: A case study from the talhar shale, Pakistan’, *Arabian Journal for Science and Engineering*, 44, pp. 417–428.

Ehsan, M. et al. (2021) ‘An integrated approach to evaluate the unconventional hydrocarbon generation potential of the Lower Goru Formation (Cretaceous) in Southern Lower Indus basin, Pakistan’, *Journal of Earth System Science*, 130(2), p. 90.

Ehsan, M. and Gu, H. (2020) ‘An integrated approach for the identification of lithofacies and clay mineralogy through Neuro-Fuzzy, cross plot, and statistical analyses, from well log data’, *Journal of Earth System Science*, 129, pp. 1–13.

Erofeev, A. et al. (2019) ‘Prediction of porosity and permeability alteration based on machine learning algorithms’, *Transport in Porous Media*, 128, pp. 677–700.

Farsi, M. et al. (2021) ‘Predicting formation pore-pressure from well-log data with hybrid machine-learning optimization algorithms’, *Natural Resources Research*, 30, pp. 3455–3481.

Fattahi, H. and Karimpouli, S. (2016) ‘Prediction of porosity and water saturation using pre-stack seismic attributes: a comparison of Bayesian inversion and computational intelligence methods’, *Computational Geosciences*, 20, pp. 1075–1094.

Fatti, J.L. et al. (1994) ‘Detection of gas in sandstone reservoirs using AVO analysis: A 3-D seismic case history using the Geostack technique’, *Geophysics*, 59(9), pp. 1362–1376.

Fowler, J.N. et al. (2004) ‘Two-dimensional kinematic modeling of the southern Kirthar fold belt, Pakistan’.

Gardner, G.H.F., Gardner, L.W. and Gregory, Ar. (1974) ‘Formation velocity and density—The diagnostic basics for stratigraphic traps’, *Geophysics*, 39(6), pp. 770–780.

Gassmann, F. (1951) ‘Elastic waves through a packing of spheres’, *Geophysics*, 16(4), pp. 673–685.

Genuer, R. et al. (2017) ‘Random forests for big data’, *Big Data Research*, 9, pp. 28–46.

Géron, A. (2022) *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow*. O’Reilly Media, Inc.

Geurts, P., Ernst, D. and Wehenkel, L. (2006) ‘Extremely randomized trees’, *Machine learning*, 63, pp. 3–42.

Ghahramani, Z. (2003) ‘Unsupervised learning’, in *Summer school on machine learning*. Springer, pp. 72–112.

Gholami, A. (2016) ‘A fast automatic multichannel blind seismic inversion for high-resolution impedance recovery’, *Geophysics*, 81(5), pp. V357–V364.

Ghosh, D. et al. (2018) ‘New workflow in reservoir modelling-incorporating high resolution seismic and rock physics’, in *Offshore Technology Conference Asia*. OTC, p. D032S004R002.

Gnos, E. (1998) ‘Peak metamorphic conditions of garnet amphibolites beneath the Semail ophiolite: implications for an inverted pressure gradient’, *International Geology Review*, 40(4), pp. 281–304.

Goodway, B. (2001) ‘AVO and Lamé constants for rock parameterization and fluid detection’, *CSEG recorder*, 26(6), pp. 39–60.

Gorain, S. and Thakur, P. (2015) ‘“Attribute based Inversion” a tool for reservoir characterization: a case study-Kalol Field, Cambay Basin, India’, *Acta Geodaetica et Geophysica*, 50(3), pp. 321–338.

Grana, D., Pirrone, M. and Mukerji, T. (2012) ‘Quantitative log interpretation and uncertainty propagation of petrophysical properties and facies classification from rock-physics modeling and formation evaluation analysis’, *Geophysics*, 77(3), pp. WA45–WA63.

Gültekin, T. B., Demir, Ö., & Sertkaya, C. (2019). A GIS-based spatial data mining approach for landslide susceptibility mapping using machine learning algorithms. *Environmental Earth Sciences*, 78(16), 413.

Hampson, D.P., Russell, B.H. and Bankhead, B. (2005) 'Simultaneous inversion of pre-stack seismic data', in *SEG Technical Program Expanded Abstracts 2005*. Society of Exploration Geophysicists, pp. 1633–1637.

Hampson, D.P., Schuelke, J.S. and Quirein, J.A. (2001) 'Use of multiattribute transforms to predict log properties from seismic data', *Geophysics*, 66(1), pp. 220–236.

Hart, B.S. (2008) 'Channel detection in 3-D seismic data using sweetness', *AAPG bulletin*, 92(6), pp. 733–742.

Hastie, T. et al. (2009) *The elements of statistical learning: data mining, inference, and prediction*. Springer.

Hazbeh, O. et al. (2021) 'Hybrid computing models to predict oil formation volume factor using multilayer perceptron algorithm', *Journal of Petroleum and Mining Engineering*, 23(1), pp. 17–30.

Helle, H.B., Bhatt, A. and Ursin, B. (2001a) 'Porosity and permeability prediction from wireline logs using artificial neural networks: a North Sea case study', *Geophysical Prospecting*, 49(4), pp. 431–444.

Helle, H.B., Bhatt, A. and Ursin, B. (2001b) 'Porosity and permeability prediction from wireline logs using artificial neural networks: a North Sea case study', *Geophysical Prospecting*, 49(4), pp. 431–444.

Helmy, T., Fatai, A. and Faisal, K. (2010) 'Hybrid computational models for the characterization of oil and gas reservoirs', *Expert Systems with Applications*, 37(7), pp. 5353–5363.

Herron, D.A. (2011) *First steps in seismic interpretation*. Society of Exploration Geophysicists.

Hu, L. et al. (2013) 'A new pore pressure prediction method-back propagation artificial neural network', *Electron. J. Geotech. Eng*, 18, pp. 4093–4107.

Huang, L., Dong, X. and Clee, T.E. (2017) 'A scalable deep learning platform for identifying geologic features from seismic attributes', *The Leading Edge*, 36(3), pp. 249–256.

Huang, Z. et al. (2011) 'Numerical study on two-phase flow through fractured porous media', *Science China Technological Sciences*, 54, pp. 2412–2420.

Hui, G. et al. (2021) 'Machine learning-based production forecast for shale gas in unconventional reservoirs via integration of geological and operational factors', *Journal of Natural Gas Science and Engineering*, 94, p. 104045.

Hussain, M. et al. (2017) ‘Reservoir characterization of basal sand zone of lower Goru Formation by petrophysical studies of geophysical logs’, *Journal of the Geological Society of India*, 89, pp. 331–338.

Hussain, M. et al. (2021) ‘Optimized rock physics model implementation to characterize the heterogeneous B-interval reservoir facies in the petro-elastic domain, Mubarak Block, Lower Indus Basin, Pakistan.’, *Kuwait Journal of Science* [Preprint].

Hussain, M. et al. (2022) ‘Application of machine learning for lithofacies prediction and cluster analysis approach to identify rock type’, *Energies*, 15(12), p. 4501.

Hussein, M., El-Ata, A.S.A. and El-Behiry, M.G. (2019) ‘Prospects delineation and evaluation using the CDP gathers characterization, onshore Nile Delta, Egypt’, *Petroleum Research*, 4(4), pp. 365–375.

Ismail, A. et al. (2020) ‘Identification of gas zones and chimneys using seismic attributes analysis at the Scarab field, offshore, Nile Delta, Egypt’, *Petroleum Research*, 5(1), pp. 59–69.

Jadoon, I., Lawrence, R.D. and Hassan, K.S. (1994) ‘Mari-Bugti pop-up zone in the central Sulaiman fold belt, Pakistan’, *Journal of Structural Geology*, 16(2), pp. 147–158.

Jadoon, S.-U.-R.K. et al. (2020) ‘Structural styles, hydrocarbon prospects, and potential of Miano and Kadanwari fields, Central Indus Basin, Pakistan’, *Arabian Journal of Geosciences*, 13, pp. 1–13.

Jensen, E.H. et al. (2016) ‘Quantitative interpretation using inverse rock-physics modeling on AVO data’, *The Leading Edge*, 35(8), pp. 677–683.

Jiang, L. et al. (2020) ‘Rock physics modeling using machine learning’, in *SEG International Exposition and Annual Meeting*. SEG, p. D031S042R006.

Kadri, I.B. (1995) ‘Petroleum geology of Pakistan: Pakistan Petroleum Limited’, *Karachi, Pakistan* [Preprint].

Kamel, M.H. and Mohamed, M.M. (2006) ‘Effective porosity determination in clean/shaly formations from acoustic logs with applications’, *Journal of Petroleum Science and Engineering*, 51(3–4), pp. 267–274.

Karimpouli, S. et al. (2018) ‘Application of alternative digital rock physics methods in a real case study: a challenge between clean and cemented samples’, *Geophysical Prospecting*, 66(4), pp. 767–783.

Kazmi, A.H. (1988) ‘Stratigraphy of the dungan group in Kach-Ziarat Area, NE Balochistan’, *Journal of Himalayan Earth Sciences*, 21.

Kazmi, A.H. and Abbasi, I.A. (2008) ‘Stratigraphy & historical geology of Pakistan’.

Kazmi, A.H. and Jan, M.Q. (1997) ‘Geology and tectonics of Pakistan’, *(No Title)* [Preprint].

Kemal, A. (1991) ‘Hydrocarbon potential in Pakistan-A geological perspective’, *AAPG Bulletin (American Association of Petroleum Geologists);(United States)*, 75(CONF-910403-).

Khan, A. et al. (2011) ‘Structural styles and hydrocarbon potential of western Kirthar Fold Belt’, in *PAPG/SPE Annual Technical Conference*, pp. 147–174.

Khan, N. et al. (2016) ‘Sequence stratigraphic analysis of Eocene Rock Strata, Offshore Indus, southwest Pakistan’, *Marine Geophysical Research*, 37, pp. 207–228.

Khan, Z.U. et al. (2022) ‘Gas-bearing sands appraisal for Zamzama Gas field in Pakistan through inverted elastic attributes assisted with PNN approximation of petrophysical properties’, *Kuwait Journal of Science*, 49(4).

Khan, Z.U. et al. (2023) ‘Bayesian stochastic inversion with petro-elastic relation to quantify thin gas sands of Khadro Formation, Zazmama gas field’, *Episodes Journal of International Geoscience*, 46(3).

Korjani, M. et al. (2016) ‘A new approach to reservoir characterization using deep learning neural networks’, in *SPE Western Regional Meeting*. SPE, p. SPE-180359-MS.

Kotsiantis, S.B., Zaharakis, I. and Pintelas, P. (2007) ‘Supervised machine learning: A review of classification techniques’, *Emerging artificial intelligence applications in computer engineering*, 160(1), pp. 3–24.

Krebs, J.R. et al. (2009) ‘Fast full-wavefield seismic inversion using encoded sources’, *Geophysics*, 74(6), pp. WCC177–WCC188.

Küçükdemirci, M. and Sarris, A. (2022) ‘GPR data processing and interpretation based on artificial intelligence approaches: Future perspectives for archaeological prospection’, *Remote Sensing*, 14(14), p. 3377.

Kuhn, M. and Johnson, K. (2013) *Applied predictive modeling*. Springer.

Lantz, B. (2019) *Machine learning with R: expert techniques for predictive modeling*. Packt publishing ltd.

Liu, Y. (2017) ‘Applications of machine learning for seismic quantitative interpretation’, in *GeoConvention*, p. 5.

Liu, Y., Sun, J., & Li, Q. (2019). Geochemical anomaly identification using hierarchical clustering and self-organizing maps: a case study in the Grasset Lake area, Canada. *Minerals*, 9(7), 380.

MacGregor, L. and Andreis, D. (2012) 'Invited Organization: Integrating Seismic, CSEM, and Well Log Data for Reservoir Characterization', in *Offshore Technology Conference*. OnePetro.

Maddahi, I., Moradzadeh, A. and Nejati Kalateh, A. (2020) 'Comparison of pore pressure prediction using conventional seismic velocity and acoustic impedance-based methods', *Journal of Petroleum Research*, 29(109), pp. 96–107.

Mahetaji, M., Brahma, J. and Sircar, A. (2020) 'Pre-drill pore pressure prediction and safe well design on the top of Tulamura anticline, Tripura, India: a comparative study', *Journal of Petroleum Exploration and Production Technology*, 10, pp. 1021–1049.

Mahgoub, M.I., Padmanabhan, E. and Abdullatif, O.M. (2018) 'Facies and porosity 3D models constrained by stochastic seismic inversion to delineate Paleocene fluvial/lacustrine reservoirs in Melut Rift Basin, Sudan', *Marine and Petroleum Geology*, 98, pp. 79–96.

Mahmood, M.F., Ahmad, Z. and Ehsan, M. (2018) 'Total organic carbon content and total porosity estimation in unconventional resource play using integrated approach through seismic inversion and well logs analysis within the Talhar Shale, Pakistan', *Journal of Natural Gas Science and Engineering*, 52, pp. 13–24.

Male, F., Jensen, J.L. and Lake, L.W. (2020) 'Comparison of permeability predictions on cemented sandstones with physics-based and machine learning approaches', *Journal of Natural Gas Science and Engineering*, 77, p. 103244.

Manzoor, U. et al. (2023) 'Harnessing Advanced Machine-Learning Algorithms for Optimized Data Conditioning and Petrophysical Analysis of Heterogeneous, Thin Reservoirs', *Energy & Fuels*, 37(14), pp. 10218–10234.

Marfurt, K.J. et al. (1998) '3-D seismic attributes using a semblance-based coherency algorithm', *Geophysics*, 63(4), pp. 1150–1165.

Maurya, S.P., Singh, N.P. and Singh, K.H. (2020) *Seismic inversion methods: a practical approach*. Springer.

Mavko, G., Mukerji, T. and Dvorkin, J. (2020) *The rock physics handbook*. Cambridge university press.

Miah, M.I., Zendehboudi, S. and Ahmed, S. (2020) ‘Log data-driven model and feature ranking for water saturation prediction using machine learning approach’, *Journal of Petroleum Science and Engineering*, 194, p. 107291.

Moncayo, E., Tchegliakova, N. and Montes, L. (2012) ‘Pre-stack seismic inversion based on a genetic algorithm: A case from the Llanos Basin (Colombia) in the absence of well information’, *CT&F-Ciencia, Tecnología y Futuro*, 4(5), pp. 5–20.

Morozov, I.B. and Ma, J. (2009) ‘Accurate poststack acoustic-impedance inversion by well-log calibration’, *Geophysics*, 74(5), pp. R59–R67.

Müller, A.C. and Guido, S. (2016) *Introduction to machine learning with Python: a guide for data scientists*. O’Reilly Media, Inc.

Munir, M.N., Zafar, M. and Ehsan, M. (2022) ‘Comparative and Statistical Analysis of Core-Calibrated Porosity with Log-Derived Porosity for Reservoir Parameters Estimation of the Zamzama GAS Field, Southern Indus Basin, Pakistan’, *Arabian Journal for Science and Engineering*, pp. 1–16.

Mustafa, A. et al. (2022) ‘Data-driven machine learning approach to predict mineralogy of organic-rich shales: An example from Qusaiba Shale, Rub’al Khali Basin, Saudi Arabia’, *Marine and Petroleum Geology*, 137, p. 105495.

Muther, T. et al. (2022) ‘Unconventional hydrocarbon resources: geological statistics, petrophysical characterization, and field development strategies’, *Journal of Petroleum Exploration and Production Technology*, 12(6), pp. 1463–1488.

Nabi, A. et al. (2019) ‘Seismotectonic analyses of Karachi Arc, Southern Kirthar Fold Belt, Pakistan’, *International Journal of Economic and Environmental Geology*, 10(2), pp. 9–18.

Nanda, N.C. (2021) *Seismic data interpretation and evaluation for hydrocarbon exploration and production*. Springer.

Naseer, M.T. (2020a) ‘Seismic attributes and reservoir simulation’application to image the shallow-marine reservoirs of Middle-Eocene carbonates, SW Pakistan’, *Journal of Petroleum Science and Engineering*, 195, p. 107711.

Naseer, M.T. (2020b) ‘Seismic attributes and reservoir simulation’application to image the shallow-marine reservoirs of Middle-Eocene carbonates, SW Pakistan’, *Journal of Petroleum Science and Engineering*, 195, p. 107711.

Naveshki, M. et al. (2021) ‘Prediction of Bubble Point Pressure Using New Hybrid Computational Intelligence Models.’, *Journal of Chemical & Petroleum Engineering*, 55(2).

Novak Zelenika, K., Novak Mavar, K. and Brnada, S. (2018) ‘Comparison of the sweetness seismic attribute and porosity–thickness maps, sava depression, Croatia’, *Geosciences*, 8(11), p. 426.

Nwachukwu, A. et al. (2018) ‘Machine learning-based optimization of well locations and WAG parameters under geologic uncertainty’, in *SPE improved oil recovery conference*. OnePetro.

Ooilithica Geoscience Ltd (2004) ‘Zamzama Field: Reservoir Sedimentology Synthesis, Dadu Block, Pakistan’.

Otchere, D.A. et al. (2022) ‘Application of gradient boosting regression model for the evaluation of feature selection techniques in improving reservoir characterisation predictions’, *Journal of Petroleum Science and Engineering*, 208, p. 109244.

Patriat, P. and Achache, J. (1984) ‘India–Eurasia collision chronology has implications for crustal shortening and driving mechanism of plates’, *Nature*, 311(5987), pp. 615–621.

Pendrel, J. (2006) ‘Seismic inversion–still the best tool for reservoir characterization’, *CSEG Recorder*, 31(1), pp. 5–12.

Poulton, S.W. and Raiswell, R. (2002) ‘The low-temperature geochemical cycle of iron: from continental fluxes to marine sediment deposition’, *American journal of science*, 302(9), pp. 774–805.

Priezzhev, I.I. et al. (2019) ‘Direct prediction of petrophysical and petroelastic reservoir properties from seismic and well-log data using nonlinear machine learning algorithms’, *The Leading Edge*, 38(12), pp. 949–958.

Qayyum, F. et al. (2016) ‘Evaluation of source rocks using one dimensional maturity modeling in Lower Indus Basin, Pakistan’, *Arabian Journal of Geosciences*, 9, pp. 1–22.

Quadri, V.-N. and Shuaib, S.M. (1986) ‘Hydrocarbon prospects of southern Indus basin, Pakistan’, *AAPG bulletin*, 70(6), pp. 730–747.

Qureshi, M.A. et al. (2021) ‘Geo-seismic model for petroleum plays an assessment of the Zamzama area, Southern Indus Basin, Pakistan’, *Journal of Petroleum Exploration and Production*, 11, pp. 33–44.

Radovich, B.J. and Oliveros, R.B. (1998) ‘3-D sequence interpretation of seismic instantaneous attributes from the Gorgon Field’, *The Leading Edge*, 17(9), pp. 1286–1293.

Radwan, A.E., Wood, D.A. and Radwan, A.A. (2022) 'Machine learning and data-driven prediction of pore pressure from geophysical logs: A case study for the Mangahewa gas field, New Zealand', *Journal of Rock Mechanics and Geotechnical Engineering*, 14(6), pp. 1799–1809.

Raef, A.E. et al. (2015) '3D seismic attributes and well-log facies analysis for prospect identification and evaluation: Interpreted palaeoshoreline implications, Weirman Field, Kansas, USA', *Journal of Petroleum science and Engineering*, 133, pp. 40–51.

Rafik, B. and Kamel, B. (2017) 'Prediction of permeability and porosity from well log data using the nonparametric regression with multivariate analysis and neural network, Hassi R'Mel Field, Algeria', *Egyptian journal of petroleum*, 26(3), pp. 763–778.

Rajabi, M. et al. (2021) 'Novel hybrid machine learning optimizer algorithms to prediction of fracture density by petrophysical data', *Journal of Petroleum Exploration and Production Technology*, 11, pp. 4375–4397.

Ray, A.K. and Chopra, S. (2016) 'Building more robust low-frequency models for seismic impedance inversion', *First Break*, 34(5).

Reading, A.M. et al. (2015) 'Combining machine learning and geophysical inversion for applied geophysics', *ASEG Extended Abstracts*, 2015(1), pp. 1–5.

Reine, C. (2017) 'Discovering a supermodel-A rock-Physics tutorial', *CSEG Recorder*, 42(2), pp. 20–25.

Richards, G. et al. (2020) 'Pore pressure prediction based on the Full Effective Stress (FES) method', in *Third EAGE Workshop on Pore Pressure Prediction*. European Association of Geoscientists & Engineers, pp. 1–5.

Richards, J.P., Boyce, A.J. and Pringle, M.S. (2001) 'Geologic evolution of the Escondida area, northern Chile: A model for spatial and temporal localization of porphyry Cu mineralization', *Economic Geology*, 96(2), pp. 271–305.

Rider, M.H. (1986) 'The geological interpretation of well logs'.

Rizwan, M. et al. (2018) 'Amplitude versus offset (AVO) modelling and analysis for quantitative interpretation of porosity and saturation: A case study for Sawan gas field, middle Indus basin, Pakistan', *Geofísica internacional*, 57(2), pp. 151–160.

Russell, B. and Hampson, D. (1991) 'Comparison of poststack seismic inversion methods', in *SEG Technical Program Expanded Abstracts 1991*. Society of Exploration Geophysicists, pp. 876–878.

Russell, B. and Hampson, D. (1999) *Comparison of post-stack seismic inversion methods SEG Technical Program Expanded Abstracts*. Society of Exploration Geophysicists.

Rutherford, S.R. and Williams, R.H. (1989) 'Amplitude-versus-offset variations in gas sands', *Geophysics*, 54(6), pp. 680–688.

Saeed, W. et al. (2020) 'An integrated petrophysical-based wedge modeling and thin bed AVO analysis for improved reservoir characterization of Zhujiang Formation, Huizhou sub-basin, China: A case study', *Open Geosciences*, 12(1), pp. 256–274.

Sams, M. and Carter, D. (2017) 'Stuck between a rock and a reflection: A tutorial on low-frequency models for seismic inversion', *Interpretation*, 5(2), pp. B17–B27.

Sayers, C.M. and den Boer, L.D. (2011) 'Rock physics-based relations for density and S-velocity versus P-velocity in deepwater subsalt Gulf of Mexico shales', *The Leading Edge*, 30(12), pp. 1376–1381.

Schlumberger (1974) *Log interpretation—applications*. New York, Schlumberger Limited.

Searle, M.P. et al. (1997) 'Shisha Pangma leucogranite, south Tibetan Himalaya: Field relations, geochemistry, age, origin, and emplacement', *The Journal of Geology*, 105(3), pp. 295–318.

Shah, S.M.I. (2009) 'Stratigraphy of Pakistan (memoirs of the geological survey of Pakistan)', *The Geological Survey of Pakistan*, 22, pp. 93–114.

Shakir, U. et al. (2021) 'Improved gas sand facies classification and enhanced reservoir description based on calibrated rock physics modelling: A case study', *Open Geosciences*, 13(1), pp. 1476–1493.

Shen, K. et al. (2017) 'Numerical investigations on pore-pressure response of suction anchors under cyclic tensile loadings', *Engineering Geology*, 227, pp. 108–120.

Shuaib, S.M. (1973) 'Subsurface Petrographic Study of Joints in Variegated Siltstone-Sandstone and Khairabad Limestone, Pakistan', *AAPG Bulletin*, 57(9), pp. 1775–1778.

Siddiqui, N.K. (2012) 'A prospective Neoproterozoic-Cambrian hydrocarbon/exploration play in the Kirthar Fold Belt, Pakistan', *Geological Society, London, Special Publications*, 366(1), pp. 123–130.

Simm, R. (2009) 'Simple net pay estimation from seismic: A modelling study', *first break*, 27(9).

Smewing, J.D. et al. (2002) 'Sequence stratigraphy of the southern Kirthar fold belt and middle Indus basin, Pakistan'.

Sohail, G.M. and Hawkes, C.D. (2020) ‘An evaluation of empirical and rock physics models to estimate shear wave velocity in a potential shale gas reservoir using wireline logs’, *Journal of petroleum science and engineering*, 185, p. 106666.

Sultan, M. and Gipson Jr, M. (1995) ‘Reservoir Potential of the Maastrichtian Pab sandstone in the Eastern Sulaiman Fold-Belt, Pakistan’, *Journal of Petroleum Geology*, 18(3), pp. 309–328.

Taner, M.T. et al. (1994) ‘Seismic attributes revisited’, in *SEG technical program expanded abstracts 1994*. Society of Exploration Geophysicists, pp. 1104–1106.

Taner, M.T. and Images, R.S. (2001) ‘Seismic attributes: CSEG Recorder’, *Rock Solid Images, Houston, USA* [Preprint].

Teh, W.J., Willhite, G.P. and Doveton, J.H. (2012) ‘Improved reservoir characterization using petrophysical classifiers within electrofacies’, in *SPE Improved Oil Recovery Conference?* SPE, p. SPE-154341-MS.

Terzaghi, K., Peck, R.B. and Mesri, G. (1996) *Soil mechanics in engineering practice*. John wiley & sons.

Williams, M.D. (1959) ‘19. Stratigraphy of the Lower Indus Basin, West Pakistan’, in *World Petroleum Congress*. WPC, p. WPC-8019.

Wuest, T. et al. (2016) ‘Machine learning in manufacturing: advantages, challenges, and applications’, *Production & Manufacturing Research*, 4(1), pp. 23–45.

Xu, C. et al. (2019) ‘When petrophysics meets big data: What can machine do?’, in *SPE Middle East Oil and Gas Show and Conference*. OnePetro.

Yasin, Q. et al. (2021a) ‘Application of machine learning tool to predict the porosity of clastic depositional system, Indus Basin, Pakistan’, *Journal of Petroleum Science and Engineering*, 197, p. 107975.

Yasin, Q. et al. (2021b) ‘Application of machine learning tool to predict the porosity of clastic depositional system, Indus Basin, Pakistan’, *Journal of Petroleum Science and Engineering*, 197, p. 107975.

Yoshida, C., Ikeda, S. and Eaton, B.A. (1996) ‘An investigative study of recent technologies used for prediction, detection, and evaluation of abnormal formation pressure and fracture pressure in North and South America’, in *IADC/SPE Asia Pacific Drilling Technology Conference and Exhibition?* SPE, p. SPE-36381-MS.

Young, R.A. and LoPiccolo, R.D. (2003) ‘A comprehensive AVO classification’, *The Leading Edge*, 22(10), pp. 1030–1037.

Yu, H., Chen, G. and Gu, H. (2020) ‘A machine learning methodology for multivariate pore-pressure prediction’, *Computers & Geosciences*, 143, p. 104548.

Zafar, Z.A. et al. (2018) ‘A radical seismic interpretation re-think resolves the structural complexities of the Zamzama Field, Kirther Foredeep, Pakistan. PAPG’, in *SPE annual technical conference, December. Islamabad, Pakistan*, pp. 11–12.

Zeoppritz, K. (1919) ‘On the reflection and propagation of seismic waves’, *Erdbebenwellen VIIB, Gottiger Nachrichten.(1)*, pp. 66–84.

Zhang, G. et al. (2019) ‘Extended Hilbert transform and application for seismic attributes’, *Earth and Space Science*, 6(5), pp. 873–886.

Zhang, J. (2011) ‘Pore pressure prediction from well logs: Methods, modifications, and new approaches’, *Earth-Science Reviews*, 108(1–2), pp. 50–63.

Zhang, L. and Zhan, C. (2017) ‘Machine learning in rock facies classification: An application of XGBoost’, in *International Geophysical Conference, Qingdao, China, 17-20 April 2017*. Society of Exploration Geophysicists and Chinese Petroleum Society, pp. 1371–1374.

Zhang, Y., Li, S., Guo, L., & Wang, Z. (2020). Reservoir property prediction with transfer learning based on well-log data. *Journal of Petroleum Science and Engineering*, 187, 106818.https://repositories.lib.utexas.edu/items/f094e756-6ce7-48a6-b9d3-8a23b24c9efb.

Zhang, Y. et al. (2020) ‘Geological characteristics and abnormal pore pressure prediction in shale oil formations of the Dongying depression, China’, *Energy Science & Engineering*, 8(6), pp. 1962–1979.

Zhang, Z. et al. (2021) ‘Permeability and porosity prediction using logging data in a heterogeneous dolomite reservoir: An integrated approach’, *Journal of Natural Gas Science and Engineering*, 86, p. 103743.

Zhang, Z. and Alkhalifah, T. (2020) ‘High-resolution reservoir characterization using deep learning-aided elastic full-waveform inversion: The North Sea field data example’, *Geophysics*, 85(4), pp. WA137–WA146.

Zheng, A. and Casari, A. (2018) *Feature engineering for machine learning: principles and techniques for data scientists*. O’Reilly Media, Inc.

Zhong, Z. et al. (2019) ‘A deep learning approach to anomaly detection in geological carbon sequestration sites using pressure measurements’, *Journal of hydrology*, 573, pp. 885–894.

# APPENDICES-PUBLICATIONS

# APPENDICES-A

222 (2023) 211451

Contents lists available at ScienceDirect

Geoenergy Science and Engineering

journal homepage: www.sciencedirect.com/journal/geoenergy-science-and-engineering

## Seismic driven reservoir classification using advanced machine learning algorithms: A case study from the Lower Ranikot/Khadro sandstone gas reservoir, Kirthar Fold Belt, Lower Indus Basin, Pakistan

Umar Manzoor [a], Muhsan Ehsan [a,*], Ahmed E. Radwan [b], Muyyassar Hussain [c], Muhammad Khizer Iftikhar [c], Farooq Arshad [c]

[a] Department of Earth and Environmental Sciences, Bahria University, Islamabad, Pakistan
[b] Faculty of Geography and Geology, Institute of Geological Sciences, Jagiellonian University, Kraków, 30-387, Poland
[c] Reservoir Characterization Department, LMK Resources Pakistan (Private) Limited, Pakistan

ARTICLE INFO

*Keywords:*
Seismic inversion
Rock physics
Facie classification
Machine learning
Attributes
Thin bed
Clastic reservoir
Lower indus basin
Pakistan

ABSTRACT

Reservoir characterization of thin sand bed reservoirs has been a challenge for petroleum explorers across the globe. In this study, we have studied the heterogeneous Paleocene sandstone gas reservoir of Lower Ranikot/Khadro, Kirthar Fold Belt, Lower Indus Basin, Pakistan. The studied reservoir is par below seismic resolution, with an average thickness that varies from 4 to 7 m at places and has provided a good amount of prosecution in some producing wells. The solution to the above challenges has been put forth where an integrated approach of seismic attributes, petrophysical properties, well logs, and facies data sets are combined with advanced machine learning algorithms to get better resolution for the studied thin-bedded sandstone resource. In addition, the thin heterogeneous sands of the Ranikot/Khadro Formation were characterized by predicting the elastic properties of the reservoir and classifying the facies distribution of hydrocarbon-bearing sands within the Zamzama Gas Field. Several machine learning (ML) algorithms were chosen and implemented on pre-stack seismic data from seven wells, two of which were produced from the Paleocene reservoir. The Gradient Boosting Regressor (GBR) produced promising results with the demarcation of thin sands with high accuracy levels. Based on a successful petro-elastic relationship, thin gas sands were highlighted. The low VCL of 30% and $S_w$ of 45% cut-off reflect gas sands with P-impedances of 7500–9500 m/s × g/cc and a Vp/Vs ratio of 1.4–1.6 as decisive ML characteristics. Based on the applied workflow aided by ML, three reservoir units have been assigned for the Paleocene hydrocarbon-bearing sands, and gas sand probability distribution maps have been developed for each reservoir unit. Aligning the seismic inversion technique with machine learning advanced algorithms not only speeds up the prediction of elastic properties for facies classification but also increases accuracy to over 90% while drastically reducing uncertainty and associated costs in the process. Advanced machine learning algorithms have allowed us to cater to thin sand beds by eliminating the need for injecting frequencies. Instead, it recognizes features from the well data, trains itself, and predicts the corresponding features on the seismic trace, thus effectively allowing us to characterize the beds that are below seismic resolution in any given complex geology. The generated probability distribution maps of the three recognized thin-bedded sand units can be the basis for petroleum exploration and development in the study area. The discussed technique is successfully applied to complex folds and thrust belts, and hence can be implemented successfully in other heterogeneous thin-bed reservoirs for the successful classification of hydrocarbon-bearing sand distribution. It would optimize the results in a short period of time and with fewer economics.

### 1. Introduction

Energy demand in this technologically advanced world is increasing exponentially, and the oil and gas industries are extremely under pressure to be able to meet the supply with increasing demand day by day. Innovative and robust solutions are required to meet the challenges

* Corresponding author.
*E-mail address:* muhsanehsan98@hotmail.com (M. Ehsan).

https://doi.org/10.1016/j.geoen.2023.211451
Received 1 October 2022; Received in revised form 11 December 2022; Accepted 3 January 2023
Available online 9 January 2023

# APPENDICES-B

energy&fuels

pubs.acs.org/EF Article

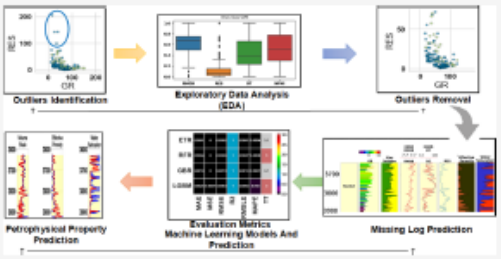
## Harnessing Advanced Machine-Learning Algorithms for Optimized Data Conditioning and Petrophysical Analysis of Heterogeneous, Thin Reservoirs

Umar Manzoor,[⊥] Muhsan Ehsan,[*,⊥] Muyyassar Hussain, Muhammad Khizer Iftikhar, Kamal Abdelrahman, S. M. Talha Qadri, Farooq Arshad, Komal Ashraf, and Mohammed S. Fnais

Cite This: https://doi.org/10.1021/acs.energyfuels.3c01293 Read Online

ACCESS | Metrics & More | Article Recommendations

**ABSTRACT:** Petrophysical analysis is an industry-standard practice for reservoir evaluation as it provides critical inputs for characterizing subsurface formations and estimating resource potential. Khadro/Ranikot Formation sands are proliferous producers in the Central Indus Basin, Pakistan. The demarcate potential in intercalated sand shale layers that are thin and heterogeneous makes it a challenging reservoir. Conventional petrophysical interpretation is laborious and does not produce up-to-mark results due to reservoir complexity, data limitations, and associated uncertainties. Hence, an emerging and delicate machine-learning (ML) approach has been comprehensively applied to analyze the potential and robustly interpret well log data while addressing the associated challenges. This case study entails a thorough evaluation of well log quality, assessing several algorithms such as least-squares support vector machines (one-class SVM), Random Forest Regressor (RFR), Extra Tree Regressor (ETR), Gradient Boosting Regressor (GBR), Decision Tree Classifier (DTC), etc. to compare their efficacy and reliability. One-class SVM helps to reduce outliers with great certainty, while the missing logs sonic (DT) and density (RHOB) are precisely predicted via GBR and ETR with 0.66 and 0.88 $R^2$, respectively. Hence, providing reliable and optimized quality logs suitable for ML-based petrophysics. ML worked on these augmented logs by dividing the data into 60% training and 40% testing. The ETR outperformed the rest of the models with a correlation of 0.99 and 0.91 among conventional and ML results. Likewise, RFR performed exceptionally well for water saturation modeling, expressing the highest 0.93 correlation. Finally, DTC modeled reservoir facies with the best 91% accuracy and 0.935 F1 measures at the blind well. Excellent calibration of >85% is met with the estimates obtained by the predictive model compared to conventional methods. This comprehensive approach offers cost-effective and robust workarounds for modern formation evaluation with minimal uncertainty and resource-efficient multiwell interpretation within complex reservoirs and sets the stage for further research in the machine-learning ecosystem.




## 1. INTRODUCTION

The reservoir understudy, the Ranikot/Khadro Formation, consists of three thin sand layers. The concerned reservoir consists of varying facies, including shale, wet sand, gas, and limestone, with varying petrophysical properties. The tricky reservoir behaves normally at certain intervals, while at other intervals, it behaves abnormally, making it challenging to resolve the net pay amongst the reservoir intervals. Therefore, estimating petrophysical characteristics from well logs through conventional approaches is crucial and laborious and may not provide a comprehensive understanding of the zone of interest. Along with reservoir complexities, data limitations, including the missing of important DT logs in the splice zone and the presence of a bad hole RHOB log add more difficulties to a complex problem. Thereby, a more robust approach like machine-learning (ML) is the need of the hour to address a nonstationary, nonlinear problem. So, ML algorithms not only help to mitigate the noise in the data but also help to predict the key missing logs, as poor data will bring poor results and miss important information. ML also helps in developing more accurate and efficient petrophysical prediction models to delineate the pay zones highlighted for perforation after testing and validation. The novice and comprehensive approach mark

Received: April 14, 2023
Revised: June 22, 2023