



PEOPLE'S DEMOCRATIC REPUBLIC OF ALGERIA MINISTRY OF HIGHER
EDUCATION AND SCIENTIFIC RESEARCH

KASDI MERBAH UNIVERSITY – OUARGLA
FACULTY OF HYDROCARBONS, RENEWABLE ENERGIES AND EARTH AND UNIVERSE
SCIENCES
DEPARTMENT OF EARTH AND UNIVERSE SCIENCES

THESIS

Doctorate - 3rd Cycle

Field: Geology

Speciality: Petroleum Geology

TOPIC

**Machine Learning Based Geomechanics Prediction While Drilling Using Mud
Logging Data: Tight Fractured Quartzites, Rhourde Nouss Field, Algeria**

Presented by: **KERMIA Hanane**

Publicly defended on 17/05/2026

Before the jury:

President

Thesis Supervisor

BELEKSIR Mohamed Saleh

Pr

Univ. Ouargla

Co-Supervisor

ZEDDOURI Azziez

Pr

Univ. Ouargla

Examiner

KECHICHED Rabah

Pr

Univ. Ouargla

Examiner

AMEUR-ZAIMECHE Ouafi

MCA

Univ. Ouargla

Examiner

HEDDAM Salim

Pr

Univ. Skikda

Examiner

MAZOUZ EL Hadi

MCA

Univ. Oum EL Bouaghi

Academic Year: 2025/2026

Acknowledgment

First, I would like to thank Allah for blessing me and guiding me in my journey to finish this work. I would like to thank my advisors **Professor BELKSEIR MOHAMED SALAH** and **Professor ZEDDOURI AZIEZ** from the University of Ouargla for their continuous support, for their patience, motivation and immense knowledge.

I would like to express my sincere gratitude and deep appreciation to **Doctor AMEUR-ZAIMECH OUAFI**, who played a pivotal role in guiding and supporting me throughout the completion of the scientific publication. He was not merely an academic supervisor, but also a devoted scientific mentor. He provided his guidance with utmost sincerity, sparing no effort or knowledge, and was a true source of support during all stages of this work until it was successfully completed. I extend to him my heartfelt thanks and profound appreciation for his continuous support and constant encouragement.

I extend my appreciation and thanks to **Doctor BOUTAGHANE Ayoub** for his efforts and valuable contribution to the completion of the scientific publication.

I would also like to extend my sincere gratitude and deep appreciation to **Professor KECHICHED RABAH** for his support, assistance, and valuable guidance throughout the preparation and writing of the scientific publication.

I would also like to express my sincere gratitude to the jury members for agreeing to evaluate this work and for their valuable feedback.

I would like to express my gratitude to **the personnel the Laboratory of Underground Reservoirs of Oil, gas, and Aquifers-** University of Ouargla for their hospitality and providing logistic assistance.

Dedication

I dedicate this work to my father, may Allah bless his soul, who were my source of strength and support throughout my life, never sparing any effort in encouraging and guiding me, and to my beloved mother, my primary support and guiding light.

I also dedicate this work to my dear sisters, to the wives of my brothers who have always supported me, and to my beloved brothers for their continuous support and encouragement throughout the completion of this work.

I dedicate this work especially to my sister, who was not born from my mother, my companion on this path, the keeper of my thoughts and writings, my beloved sister and friend, and constant supporter, *BLAHA Wahiba*, and to her children, who are a piece of the soul of my brother *Mourad*, may Allah have mercy on him.

For all those who supported me conditionally or unconditionally.

List of abbreviations

Δt	Transit time	NPHI	Neutron Porosity
AB	Adaptive Boosting	OGR	Oil gas reservoirs
AI	Artificial intelligence	PE	Photoelectric factor
ANFIS	Adaptive neuro-fuzzy inference system	POTA	Potassium
ANN	Artificial neural network	PR	Poisson's ratio
BBOE	Billion barrels of oil equivalent	Q	Drilling fluid flow Rate in gpm
CALI	Borehole diameter	QH	El hamra quartzite
CB	Categorical Boosting	R	Correlation coefficient
CBM	Coalbed methane	R ²	Coefficient of determination
CV	Cross validation	RF	Random foreste
DD	Drilling data	RMSE	Root mean square error
DEPTH	Depth	ROHB	Bulk density
DL	Deep learning	ROP	Rate of penetration. ft/hr
DPR	Dynamic Poisson's ratio	RPM	Rotation per minute
DTc	Compressional slowness. $\mu\text{s}/\text{ft}$	RT	Resistivity
DTs	Shear slowness. $\mu\text{s}/\text{ft}$	SGRs	Shale gas reservoirs
E	Dynamic Young's modulus	SPP	Standpipe pressure. psi
ELM	Extrem learning machine	Std Dev	Standar deviation
EOR	Enhanced Oil Recovery	SVM	Support vector machine
FN	Functional network	SVR	Support vector regression
GB	Gradient boosting	T	Torque. klbf.ft
GBDT	Gradient Boosted Decision Trees	THOR	Thorium
GBDT	Gradient-boosted decision tree	TOC	Total organic carbon
GR	Gamma ray (natural gamma-ray)	URAN	Uranium
GRNN	Generalized regression neural network	V _{dyn}	Dynamic Poisson's ratio
LD	Logging data	V _p	Compressional wave velocities
LGBM	Light Gradient Boosting	V _s	Shear wave velocities
MAE	Mean absolute error	V _{stat}	Static Poisson's ratio
ML	Machine learning	WOB	Weight on bit. klbm
MLPNN	Multilayer perceptron neural network	XGB	Extreme gradient boosting

Abstract

For accurate characterization of petroleum reservoirs and the optimization of drilling and production operations, a comprehensive understanding of the geomechanical properties of rock formations is required, particularly Poisson's ratio and Young's modulus, due to their fundamental role in wellbore stability assessment, in-situ stress estimation, drilling performance optimization, and hydraulic fracturing design. However, the estimation of these properties has traditionally relied on direct laboratory measurements on core samples or on sonic and density well logs. These methods are often unavailable or of limited accuracy in tight and naturally fractured reservoirs, owing to poor core recovery, data loss, and the degradation of log quality resulting from complex drilling conditions and the influence of drilling mud.

This thesis aims to develop intelligent machine learning-based frameworks for the continuous and real-time prediction of geomechanical properties, specifically dynamic Poisson's ratio (*DPR*) and dynamic Young's modulus (*DE*), using drilling data and geophysical well log data. The developed models were trained and validated using 247 data points for *DPR* prediction and 1894 data points for *DE* prediction. The tight and naturally fractured El Hamra Quartzite Formation, of Rhourde Nouss field in southeastern Algeria, was selected as a case study for this research.

The main findings of this thesis can be summarized as follows:

- Dynamic Poisson's ratio prediction is highly accurate using real-time drilling data, with Extreme Gradient Boosting (XGB) achieving ($R^2 \approx 0.856$, $RMSE \approx 0.025$, $MAE \approx 0.019$).
- Dynamic Young's modulus prediction reaches optimal performance with XGB applied to conventional well logs ($R^2 \approx 0.963$, $RMSE \approx 2.281$, $MAE \approx 1.355$), particularly after reservoir classification based on borehole diameter CALIPER
- The innovative CALIPER (borehole diameter)-based reservoir classification strategy improved model accuracy by reducing internal data variability ($R^2 \approx 0.963$, $RMSE \approx 2.281$, $MAE \approx 1.355$).
- Using real-time drilling data represents a technical and economic alternative to traditional methods, enhancing decision making and reducing field operation costs.

The results of this thesis confirm the effectiveness of artificial intelligence techniques in achieving accurate geomechanical characterization of complex petroleum reservoirs, including tight and fractured reservoirs, while saving time and effort and reducing costs associated with laboratory testing and conventional logging operations notably when losing of travel time (Δt). Furthermore, this study highlights the importance of integrating operational data with intelligent models as a modern and promising approach for developing more efficient and reliable reservoir evaluation strategies in contemporary petroleum engineering.

Keywords:

Poisson's ratio, Young's modulus, tight fractured reservoirs, artificial intelligence, Extreme Gradient Boosting, drilling data, geophysical well logs, Rhourde Nouss, Algeria.

Résumé

Pour une caractérisation précise des réservoirs pétroliers et l'optimisation des opérations de forage et de production, une connaissance approfondie des propriétés géomécaniques des formations rocheuses est indispensable, en particulier le coefficient de Poisson et le module de Young's, en raison de leur rôle fondamental dans l'évaluation de la stabilité du puits, l'estimation des contraintes in situ, l'optimisation des performances de forage et la conception des opérations de fracturation hydraulique. Cependant, l'estimation de ces propriétés repose traditionnellement sur des mesures directes en laboratoire réalisées sur des carottes, ou sur des diagraphies acoustiques. Ces méthodes sont souvent indisponibles ou d'une précision limitée dans les réservoirs compacté et naturellement fracturés, en raison du faible taux de récupération des carottes, de la perte de données et de la dégradation de la qualité des diagraphies liée aux conditions de forage complexes et à l'influence de la boue de forage.

Cette thèse vise à développer des cadres intelligents basés sur l'apprentissage automatique pour la prédiction continue et en temps réel des propriétés géomécaniques, en particulier le coefficient de Poisson dynamique (*DPR*) et le module de Young's dynamique (*DE*), à partir des données de forage et des données de diagraphies géophysiques. Les modèles ont été entraînés et validés sur 247 points de données pour le *DPR* et 1 894 points de données pour le *DE*. La formation de quartzite d'El Hamra, compacté et naturellement fracturée, de champ de Rhourde Nouss au sud-est de l'Algérie, a été choisie comme étude de cas pour cette recherche.

Les principaux résultats de cette thèse peuvent être résumés comme suit :

- La prédiction du coefficient de Poisson dynamique à partir des données de forage en temps réel présente une précision élevée, le modèle Extreme Gradient Boosting (XGB) obtenant ($R^2 \approx 0,856$, $RMSE \approx 0,025$ et $MAE \approx 0,019$).
- La prédiction du module de Young dynamique atteint une performance optimale avec XGB appliqué aux diagraphies conventionnelles, avec ($R^2 \approx 0,963$, $RMSE \approx 2,281$ et $MAE \approx 1,355$), particulièrement après la classification de réservoir basée sur le diamètre du puits (CALI).
- La classification innovante par CALI améliore la précision en réduisant la variabilité interne des données ($R^2 \approx 0.963$, $RMSE \approx 2.281$, $MAE \approx 1.355$).
- L'utilisation des données de forage en temps réel est une alternative technique et économique aux méthodes traditionnelles, améliorant la prise de décision et réduisant les coûts opérationnels.

Les résultats de cette thèse confirment l'efficacité des techniques d'intelligence artificielle pour une caractérisation géomécanique précise des réservoirs pétroliers complexes, y compris les réservoirs compacté et fracturés, tout en permettant un gain de temps et d'efforts et une réduction des coûts associés aux essais en laboratoire et aux opérations de diagraphie conventionnelles. En outre, cette étude met en évidence l'importance de l'intégration des données opérationnelles avec des modèles intelligents comme une approche moderne et prometteuse pour le développement de stratégies d'évaluation des réservoirs plus efficaces et plus fiables en ingénierie pétrolière contemporaine.

Mots-clés

Coefficient de Poisson, module de Young, réservoirs compacté et fracturés, intelligence artificielle, Extreme Gradient Boosting, données de forage, données de diagraphies, Rhourde Nouss, Algérie.

ملخص

من أجل التوصيف الدقيق للمكانم البترولية وتحسين عمليات الحفر والإنتاج، يتطلب الأمر معرفة شاملة بالخصائص الجيوميكانيكية للتكوينات الصخرية، ولا سيما معامل بواسون ومعامل يونغ، لما لهما من دور أساسي في تقييم استقرار البئر، وتقدير الإجهادات في الموقع، وتحسين أداء الحفر، وتصميم عمليات التكسير الهيدروليكي غير أن تقدير هذه الخصائص يعتمد تقليدياً على القياسات المخبرية المباشرة على العينات اللبّية أو على التسجيلات الصوتية والكثافية، وهي طرائق غالباً ما تكون غير متاحة أو محدودة الدقة في المكانم الضيقة والمتصدعة طبيعياً، بسبب ضعف استرجاع اللب، وفقدان البيانات، وتدهور جودة التسجيلات نتيجة ظروف الحفر المعقدة وتأثير طين الحفر تهدف هذه الأطروحة إلى تطوير أطر ذكية قائمة على التعلّم الآلي للتنبؤ المستمر وفي الزمن الحقيقي بالخصائص الجيوميكانيكية، بالخصوص نسبة بواسون الديناميكي (dynamic Poisson's ratio) ومعامل يونغ الديناميكي (dynamic Young's modulus)، اعتماداً على بيانات الحفر وبيانات التسجيلات الجيوفيزيائية. تم تدريب النماذج المطورة والتحقق من صحتها باستخدام 247 نقطة بيانات للتنبؤ بمعامل بواسون الديناميكي و1894 نقطة بيانات للتنبؤ بمعامل يونغ الديناميكي. وقد تم اختيار تكوين كوارتزيت الحمرة (Quartzite El Hamra) الضيق والمتصدع طبيعياً، بحقل غورد النص (Rhourde Nouss) جنوب شرق الجزائر، كحالة تطبيقية لهذه الدراسة.

يمكن تلخيص النتائج الرئيسية لهذه الأطروحة على النحو التالي

- تُظهر تنبؤ نسبة بواسون الديناميكية باستخدام بيانات الحفر في الوقت الفعلي دقة عالية، حيث حقّق نموذج (XGB) Extreme Gradient Boosting قيم $R^2 \approx 0.856$ ، $RMSE \approx 0.025$ ، و $MAE \approx 0.019$.
 - يصل تنبؤ معامل يونغ الديناميكي إلى أداء أمثل عند تطبيق XGB على تسجيلات الآبار التقليدية، مع قيم $R^2 \approx 0.963$ ، $RMSE \approx 2.281$ ، و $MAE \approx 1.355$ ، خاصة بعد تصنيف المكانم بناءً على قطر البئر. (CALI)
 - أسهمت الاستراتيجية المبتكرة لتصنيف المكانم المبنية على CALI في تحسين دقة النماذج من خلال تقليل التباين الداخلي للبيانات، مما عزز موثوقية التنبؤ ($R^2 \approx 0.963$, $RMSE \approx 2.281$, $MAE \approx 1.355$)
 - يعد استخدام بيانات الحفر في الزمن الحقيقي بديلاً تقنياً واقتصادياً للطرق التقليدية، مع تحسين اتخاذ القرار وتقليل تكاليف العمليات الميدانية.
- تؤكد نتائج هذه الأطروحة فعالية تقنيات الذكاء الاصطناعي في التوصيف الجيوميكانيكي الدقيق للمكانم البترولية المعقدة، بما في ذلك المكانم الضيقة والمتصدعة، مع توفير الوقت والجهد وتقليل التكاليف المرتبطة بالاختبارات المخبرية وعمليات التسجيل التقليدية .

كما تُبرز هذه الدراسة أهمية دمج البيانات التشغيلية مع النماذج الذكية كاتجاه حديث وواعد لتطوير استراتيجيات تقييم مكانم أكثر كفاءة وموثوقية في هندسة البترول الحديثة.

الكلمات المفتاحية:

معامل بواسون، معامل يونغ، المكانم الضيقة والمتصدعة، الذكاء الاصطناعي، التعزيز التدرّجي المتطّرف، بيانات الحفر، التسجيلات الجيوفيزيائية، غورد النص، الجزائر

TABLE OF CONTENTS

DEDICATION	I
ACKNOWLEDGMENT	II
ABSTRACT	III
RESUME	IV
المُلخَص	V
LIST OF FIGURES	IV
LIST OF TABLES	IX
GENERAL INTRODUCTION	1
Chapter I Overview of Unconventional Reservoirs	4
Introduction to unconventional reservoirs	4
1.1. Type of unconventional reservoirs	5
1.1.1. Unconventional oil reservoirs.....	5
1.1.1.1. Shale oil reservoirs	5
1.1.1.2. Tight oil reservoirs.....	6
1.1.2. Unconventional gas reservoirs.....	6
1.1.2.1. Coal-bed reservoirs	6
1.1.2.2. Shale gas reservoirs	7
1.1.2.3. Tight gas Reservoirs	8
1.2. Unconventional Reservoirs in The World	10
1.2.1. Factors Controlling the Global Distribution of Unconventional Reservoirs	13
1.2.1.1. Tectonic controls	13
1.2.1.2. Paleogeographic Evolution and Continental Drift.....	13
1.2.1.3. Depositional Environment and Organic Matter Preservation	14
1.3. Unconventional Reservoirs in Algeria	15
1.4. Natural Fractures	17
1.4.1. Type of Natural Fractures in reservoirs	18
1.4.1.1. Structural classification.....	18
1.4.1.2. Hydrological Classification	19
1.4.2. General Importance of Natural Fractures in reservoirs	20
1.4.2.1. Positive Effects of Natural Fractures	20
1.4.2.2. Negative Effects of Natural Fractures	21
Chapter II Regional and Local Geology of the Study Area	22
2.1. General Overview of the Berkine Basin.....	22
2.2. Geographic framework of the Berkin basin and location of the Rhourde Nouss field	23
2.3. Geological Context.....	24

2.3.1. Tectonic Evolution	24
2.3.1.1. The Pan-African Phase	24
2.3.1.2. Cambro-Ordovician Distension	25
2.3.1.3. Taconic compression (Caradoc- Ashgill)	25
2.3.1.4. Silurian Distension.....	25
2.3.1.5. The Siluro-Devonian or Caledonian compression.....	25
2.3.1.6. Frasnian distention.....	26
2.3.1.7. Hercynian Unconformity	26
2.3.1.8. Austrian phase	26
2.3.1.9. Pyrenean Phase	26
2.3.2. Stratigraphic Aspect	27
2.4. Petroleum Aspect of the Rhourde Nouss field	33
2.4.1. Source Rocks	33
2.4.2. Sealing Rocks	33
2.4.3. Traps	33
2.4.4. Reservoir Rocks.....	34
2.4.4.1. The Upper Argillaceous–Sandstone Triassic formation (TAGS).....	34
2.4.4.2. The infra-TAGS reservoirs	34
2.4.4.3. The deep reservoir of the Cambro-Ordovician Hamra Quartzites.....	34
Conclusion.....	36
Chapter III Artificial Intelligence and Machine Learning Techniques	37
Introduction	36
3.1. Intelligence Artificial (AI).....	36
3.1.1. Definition and conception of intelligence artificial	36
3.2. Machine Learning (ML)	37
3.2.1. Definition and principals of machine learning	37
3.2.2. Types of machine learning	38
3.2.2.1. Supervised learning	38
3.2.2.2. Unsupervised learning	39
3.2.2.3. Reinforcement learning.....	40
3.2.2.4. Deep Learning	41
3.3. Employed Algorithms for Predictive Modeling	42
3.3.1. Random Forest.....	42
3.3.1.1. Random Forest: Principles and Methodology	42
3.3.1.2. Splitting Criteria	43
3.3.2. Prediction Aggregation.....	43

3.3.3. Gradient Boosting	44
3.3.3.1. Gradient Boosting: Principles and Methodology	44
3.3.3.2. Algorithmic Procedure	45
3.3.4. Extreme gradient boosting	45
3.3.4.1. Extreme gradient boosting: principles and methodology	45
3.3.4.2. Additive Boosting Strategy	46
3.3.4.3. Regression Tree Ensemble and Prediction Formulation	46
3.3.5. Hierarchical Tree Structure	46
3.3.5.1. Regularized Objective Function and Overfitting Prevention	47
3.3.5.2. Regularization Mechanics and Generalization Enhancement	47
3.3.6. Categorical Boosting Algorithm	47
3.3.6.1. Categorical Boosting: principles and methodology	47
3.3.6.2. Additive Model Formulation	48
3.3.6.3. Ordered Boosting and Regularization	48
3.3.7. Adaptive Boosting algorithm	48
3.3.7.1. Adaptive Boosting: principles and methodology	48
3.3.7.2. Objective Function in AdaBoost	48
3.3.8. Light Gradient Boosting	49
3.3.8.1. Light Gradient Boosting: principles and methodology	49
3.3.8.2. Objective Function optimization	49
3.3.8.3. Robustness and Generalization Capability	50
3.4. Model interpretability and explainability tools	50
3.4.1. Feature importance	51
3.4.2. SHapley Additive explanations	51
3.5. Application of intelligence artificial in petroleum engineering	52
3.6. Application of Artificial Intelligence in Geomechanical Property Estimation	52
Chapter IV Practical Application of Artificial Intelligence in Predicting Geomechanical Properties of Tight Fractured Reservoir Quartzite El Hamra	36
Introduction	54
Part I. Real time dynamic Poisson's ratio prediction	56
Introduction	56
4.1.1. Material and methods	57
4.1.1.1. Data collection and description	57
4.1.1.2. Compiled Dataset Characterization	57
4.1.1.3. Implementation of Machine Learning Techniques	60
4.1.2. Results	63

4.1.3. Discussion.....	67
4.1.3.1. Evaluation of ML in predicting dynamic Poisson's in tight fractured reservoir	67
4.1.3.2 Feature importance significance and implications.....	69
4.1.3.3. Evaluation of XGBoost Performance for Real-Time Prediction of Dynamic Poisson's Ratio Using Drilling Data	71
4.1.3.4. Advantages and recommendations for future research to improve the technique	74
Conclusion.....	75
Part II. Dynamic Young's Modulus Prediction.....	76
Introduction	76
4.2.1. Material and methods	77
4.2.1.1. Data collection and description.....	77
4.2.1.2. Compiled Dataset Characterization	78
4.2.1.3. Developing and Assessing Machine Learning Models.....	81
4.2.1.4. Implementation of Machine Learning Techniques	83
4.2.3. Results	85
4.2.4. Discussion.....	95
4.2.4.1. Evaluation of ML in predicting dynamic Young's modulus in tight fractured reservoir	95
4.2.4.2. SHAP Interpretability analysis for the XGB for Class 1 Reservoir.....	100
4.2.4.3. Proposed Avenues for Refining the Technique in Subsequent research	101
Conclusion.....	103
Part III. Scientific and practical significance of the study.....	105
4.3.1. Scientific implications	105
4.3.1.1. Evaluation of machine learning geomechanical model accuracy	105
4.3.1.2. Beyond evaluation metrics: the role of data type and reservoir complexity in model behavior	107
4.3.2. Operational and economic Implications of Real-Time Geomechanical Characterization Based on Drilling.....	112
4.3.2.1. Enabling Real-Time Geomechanical Characterization During Drilling.....	112
4.3.2.2. Implications for Hydraulic Fracturing and Uncertainty Reduction	113
4.3.2.3. Economic Implications: cost efficiency and resource optimization	113
4.4. Impact of the Study on the Rhourde Nouss Field.....	114
General Conclusion	117
Appendix	142

LIST OF FIGURES

Figure 1. Schematic Diagram Illustrating the Evolution from Conventional to Unconventional Reservoirs	5
Figure 2. Geological schematic showing the distribution of conventional and unconventional hydrocarbon accumulations (Ahmed and Meehan, 2016).	10
Figure 3. Evolution of Global Unconventional Reservoir Production: History and Projections, (a) unconventional gas reservoirs and (b) unconventional oil reservoirs.	12
Figure 4. Global distribution map of cumulative production and unconventional oil and gas reserves (Data compiled from; (World Energy Council, 2013; EIA, 2013).....	15
Figure 5. Geological Distribution of Unconventional Hydrocarbon Systems within major sedimentary basins in Algeria. (source: Webography).....	17
Figure 6. Types of natural fractures in reservoir cores: open (O.f), closed (C.f) and mineral-filled fracture (M-f.f).	20
Figure 7. Improved satellite image of the Maghreb region Northwest Africa, illustrating the primary tectonic domains of the area and the position of the study area Rhourde Nouss (Galeazzi et al., 2010).....	23
Figure 8. Location of the Rhourde Nouss field study area in Algeria.....	24
Figure 9. Stratigraphic summary column of the Rhourde Nouss field in Algeria, highlighting the El Hamra Quartzite formations (Beicip/SONATRACH, 2002).....	32
Figure 10. Petroleum Sources in the basin of Berkine (Boote et al., 1998)	34
Figure 11. Sequence stratigraphic framework and depositional architecture of the Mid-Ordovician Hamra Quartzites formation within the Saharan platform (Galeazzi et al., 2010).	35
Figure 12. Hierarchical Relationship Between Artificial Intelligence, Machine Learning, and Deep Learning. (source: Webography)	37
Figure 13. Main Paradigms of Machine Learning and Their Typical Applications (source: Webography).....	38
Figure 14. Workflow of Supervised Machine Learning: From Raw Data to Output Predictions (source: Webography)	39
Figure 15. Workflow of Unsupervised Machine Learning: From Raw Data to Output Predictions (source: Webography)	40
Figure 16. Workflow of Reinforcement Learning: From Raw Data to Output Predictions (source: Webography)	41
Figure 17. Deep Learning: Example Illustrating Deep Neural Network Functioning (source: Webography).....	42
Figure 18. Random forest workflow: From Raw Data to Output Predictions (source: Webography).....	43
Figure 19. Structural Design of Gradient Boosting Decision Trees (Deng et al., 2021).....	44
Figure 20. Structural Design of Extreme Gradient Boosting (Sahour et al., 2021).	46

Figure 21. Schematic comparison of tree structures in ensemble learning algorithms: AdaBoost, XGBoost, LightGBM, and CatBoost.....	50
Figure 22. Correlation matrix and correlation analysis between drilling parameters and the dynamic Poisson's ratio	60
Figure 23. Flowchart illustrating the ML model architecture developed to predict formation dynamic Poisson's ratio using drilling data tight reservoir.	63
Figure 24. Measured dynamic Poisson's ratio versus predicted dynamic Poisson's ratio for the best-performing ML for the testing subset 3 best scenarios (1, 2 and 3).	65
Figure 25. Learning curve of the mean squared error as a function of the training data size...	67
Figure 26. Loss curve behavior during model training and validation.....	67
Figure 27. Feature importance scores illustrating the contribution of each drilling parameter to the prediction of dynamic Poisson's ratio.....	70
Figure 28. Model interpretability illustrated through the SHapley Additive exPlanations (SHAP) summary plot.	71
Figure 29. Comparative visualization of dynamic Poisson's ratio predictions from three machine learning models using box, violin, radar plots, and a Taylor diagram for the QH-Full Reservoir.	72
Figure 30. Correlation matrix and correlation coefficient between the dynamic Young's modulus E and logging data for three reservoirs.....	80
Figure 31. Flowchart illustrating the ML model architecture developed to predict dynamic Young's modulus using logging data	83
Figure 32. Measured Young's modulus versus predicted Young's modulus for the best-performing ML for the testing subset for the QH-Class 1 Reservoir.	88
Figure 33. Measured Young's modulus versus predicted Young's modulus for the best-performing ML for the testing subset for the QH-Full Reservoir	91
Figure 34. Measured Young's modulus versus predicted Young's modulus for the best-performing ML for the testing subset for the QH-Class 2 Reservoir.	94
Figure 35. Comparative results (R^2) with published studies designed to predict Young's modulus.....	98
Figure 36. Comparative visualization of dynamic Young's modulus predictions from six machine learning models using box, violin, radar plots, and a Taylor diagram for the QH-Class 1 Reservoir.	99
Figure 37. Comparative visualization of dynamic Young's modulus predictions from six machine learning models using box, violin, radar plots, and a Taylor diagram for the QH-Full Reservoir.	100
Figure 38. Comparative visualization of dynamic Young's modulus predictions from six machine learning models using box, violin, radar plots, and a Taylor diagram for the QH-Class 2 Reservoir.	101
Figure 39. Model interpretability illustrated through the SHapley Additive exPlanations (SHAP) summary plot.	102
Figure 40. Comparison of Evaluation Metrics for Poisson's Ratio and Young's Modulus ...	106
Figure 41. comparative evaluation of model performance Metrics across reservoir class.....	111

LIST OF TABLES

Table 1. Summary of the Measured Variables Used as Input Parameters from Mud Logging Data.....	57
Table 2. Statistical Analysis of value distributions for measured Poisson's ratio and DD variables.	58
Table 3. Hyperparameter setting applied to the three configured ML models	Error!
Bookmark not defined.	
Table 4. Input variable scenarios of drilling data for dynamic Poisson's ratio prediction.	64
Table 5. Dynamic Poisson's ratio prediction performances using RF, GB, and XGB.	64
Table 6. Results from studies designed to predict Poisson's ratio from drilling and well logging variables inputs.....	73
Table 7. Summary of the Measured Variables Used as Input Parameters from Logging Data.	78
Table 8. Analyse statistique de la distribution des valeurs des LD mesurées et du module de Young's calculé.....	78
Table 9. Input Data Scenarios from Well Logging Variables for Dynamic Young's Modulus Prediction.	82
Table 10. Overview of Model Optimization Methods and Hyperparameter Tuning	143
Table 11. Dynamic Young's modulus E prediction performances of the six ML for the QH-Class 1 Reservoir.....	86
Table 12. Dynamic Young's modulus E prediction performances of the six ML for the QH-Full Reservoir.....	90
Table 13. Dynamic Young's modulus E prediction performances of the six ML for the QH-Class 2 Reservoir.....	93
Table 14. Comparison of results with published studies designed to predict Young's modulus from drilling and logging variables inputs.....	96

GENERAL INTRODUCTION

GENERAL INTRODUCTION

Amid the increasing demand for energy, the world has witnessed, over recent decades, a significant acceleration towards exploiting unconventional reservoirs as a primary energy source. It has become evident that these reservoirs hold strategic importance in securing future energy resources, especially in light of the declining reserves of conventional reservoirs. Unconventional reservoirs are considered promising sources (Holditch, 2003; Caineng et al., 2015), containing vast reserves of oil and gas that can significantly contribute to meeting the growing global energy needs (Bentley, 2002; Kjärstad and Johnsson, 2009). With this shift, these reservoirs have become a focal point of interest for many countries and global companies, as they are seen as a crucial part of the future of energy (Vedachalam et al., 2015).

However, the exploitation of unconventional reservoirs faces numerous challenges due to the unpredictability of their behavior during drilling and production (Zhang et al., 2025), especially those characterized by natural fracture networks (Dong et al., 2020). These natural fractures present two main scenarios. The first scenario benefits production and economics (Nelson, 1985; Gale et al., 2014), as the presence of these fractures facilitates the extraction process for petroleum engineers by enhancing the permeability of the reservoir (Bai et al., 2020), thereby improving fluid flow (Shang et al., 2024). On the other hand, the second scenario hinders production and exploitation, as these fractures pose significant challenges that cannot be ignored during drilling and production operations (Dusseault, 2011; Halim et al., 2021). They can also lead to a loss of control over the drilling path and fluid flow (Epelle and Gerogiorgis, 2020), increasing operational risks and negatively affecting the economic viability of the project.

In response to these challenges, a precise and comprehensive reservoir characterization is essential (Abdulhadi et al., 2025). This requires a thorough investigation of the petrophysical properties that govern the reservoir's capacity to store and transmit hydrocarbons, such as porosity and permeability (Tiab and Donaldson, 2024). Additionally, there must be a focus on the geomechanical properties such as Poisson's ratio, Young's modulus and in-situ stresses, which examine the behavior of rocks under varying pressures during drilling and production, especially under conditions of horizontal drilling and hydraulic fracturing (Nagel and Sanchez-Nagel, 2015; Zoback and Kohli, 2019). These studies help in understanding how the reservoir interacts with extraction processes.

Traditionally, the geomechanical properties of reservoirs, such as Poisson's ratio and Young's modulus, are measured from laboratory samples using triaxial experiments and compression tests (Fjaer et al., 2008; Liu, 2017). However, these methods face significant limitations in naturally fractured unconventional reservoirs, where the complexity of the rock and the density of fracture networks make it difficult to obtain accurate representative samples. Additionally, the use of rock wave behavior logs during drilling (log data) encounters technical issues (Barree et al., 2009; Ashraf et al., 2021), notably the loss of travel time (Δt) for shear and compressional waves affected by drilling fluid density, leading to inaccurate readings, as noted by Zerrouki et al. (2014), making traditional estimation of geomechanical properties challenging and unreliable in these reservoirs.

To address these challenges, many researchers have turned to artificial intelligence (AI) techniques due to their proven ability to efficiently solve complex problems (Dramschi, 2020; Popa and Cassidy, 2012). Most of these studies have focused on predicting geomechanical properties in conventional reservoirs, particularly Poisson's ratio and Young's modulus, using drilling or logging data without relying entirely on traditional methods (Tariq et al., 2017; Elkatatny, 2018; Ahmed et al., 2021; Mahmoud et al., 2022; Siddig and Elkatatny, 2021; Sanei et al., 2023; Alakbari et al., 2024). AI-based models have demonstrated the ability to handle incomplete or information-loss-affected data, improving prediction accuracy and providing reliable estimates of geomechanical properties. This capability significantly enhances reservoir management and supports the optimal planning of drilling and hydraulic fracturing operations.

Unlike previous studies that focused on predicting geomechanical properties in conventional reservoirs, this research aims to address an important scientific gap in the literature related to the development of tight fractured reservoirs. The main contribution of this work lies in the first-time application of machine learning tools to predict geomechanical properties, particularly Poisson's ratio and Young's modulus, in tight fractured reservoirs. This is achieved by employing widely used supervised machine learning algorithms known for their high accuracy and robustness in reservoir parameter prediction, including Random Forest from the bagging and Boosting family; Gradient boosting, Extreme gradient boosting, Light gradient boosting machine, Categorical boosting and Adaptive boosting.

Drilling and logging data from wells penetrating the Cambro Ordovician Quartzite El Hamra Formation, located in the Rhourde Nouss field in southeastern Algeria, were used as input variables. These reservoirs face significant challenges, including poor core recovery and low quality logging data affected by drilling fluids and adverse borehole conditions, which make the accurate measurement or estimation of geomechanical properties difficult. The

proposed machine learning models provide flexible and reliable predictions, enabling the generation of a continuous profile of Poisson's ratio and Young's modulus.

This approach enhances drilling efficiency, reduces operational costs, and offers a practical tool for geomechanical analysis and reservoir management, serving as a complementary methodology to conventional approaches for the exploitation of unconventional reservoirs.

Thesis Structure and Organization

This thesis is organized into four main chapters, thoughtfully structured to examine the geological framework, theoretical principles, and practical application of artificial intelligence for predicting the geomechanical properties of tight fractured reservoirs.

Chapter I This chapter provides a comprehensive overview of unconventional reservoirs, addressing their main types (oil and gas, including tight, shale, and coal-bed), the factors controlling their global distribution, their spatial distribution within Algeria, and the overall significance of natural fractures in reservoir behavior and performance.

Chapter II Presents the regional and local geology of the study area, including a general overview of the Berkine Basin, its geographic framework and the location of the Rhourde Nouss field, as well as the geological context covering tectonic evolution, stratigraphy, and the petroleum characteristics of the field.

Chapter III Introduces artificial intelligence and machine learning, explaining their concepts and applications in petroleum engineering. It reviews machine learning principles, types, deep learning, and the algorithms used in this study for predictive modeling. The chapter also covers model interpretability using tools like Feature Importance and SHAP.

Chapter IV Presents the practical AI application for predicting geomechanical properties in a tight fractured reservoir Quartzite El Hamra Formation:

- Part I uses drilling data (DD) for real-time prediction of dynamic Poisson's ratio.
- Part II uses logging data (LD) for predicting dynamic Young's modulus, with a CALIPER based reservoir classification.
- Part III addresses the key methodological, operational and economic implication of the study.

A general conclusion summarizing all the results obtained in this study, along with the perspectives and potential directions for future research, is presented at the end of this thesis.

Chapter I Overview of Unconventional Reservoirs

Introduction to unconventional reservoirs

Unconventional reservoirs (including coal bed gas, shale oil/gas, and tight oil/gas) are a distinct type of rock formations characterized by their ability to generate and/or trapping hydrocarbons (Glorioso and Rattia, 2012), such as oil and gas, in substantial quantities (Kundu, 2023). Their distinctiveness lies in their geological composition, petrophysical properties, and the specialized techniques required for hydrocarbon extraction.

The first distinguishing feature of unconventional reservoirs compared to conventional reservoirs; the petroleum system. Unlike trap reservoirs (conventional reservoirs), the petroleum system in unconventional reservoirs does not rely on the traditional integration of the essential elements of a conventional petroleum system: source rock, reservoir rock, cap rock or seal, trap and migration. Instead, these elements are either integrated or modified in a way that makes the petroleum system highly complex (Chengzao, 2017).

Geophysically, source reservoirs (unconventional reservoirs) are defined as sedimentary layers stacked at great depths (Fangzheng, 2019; Gong et al., 2023), where their porosity and permeability are significantly reduced compared to trap reservoirs (Sander et al., 2017), leading to extremely slow, and even impossible fluid flow mechanisms within the rocks (Malozyomov et al., 2023), even under high pressure (Zhang et al., 2024).

Geophysicists face significant challenges in detecting this type of reservoir, as well as in defining its boundaries and the distribution of hydrocarbons within it, due to the often unclear or weak seismic response (Kaczmarczyk and Slota-Valim, 2020). As a result, more advanced techniques are used, such as 3D seismic analysis and geophysical models that integrate petrophysical data (Yang et al., 2013).

On the other hand, these complexities present major technical challenges during drilling and production stimulation, leading petroleum engineers to employ specialized techniques such as horizontal drilling and hydraulic fracturing (Epelle and Gerogiorgis., 2020), which, in turn, increase production costs, making it economically unfeasible (Gregory et al., 2011; King, 2012). However, despite this, and due to the continuous increase in energy demand driven by predictions of depletion of conventional hydrocarbon resources, most countries have turned to exploit this type of reservoir, provided that specific studies are conducted to enhance its performance alongside economic production.

Figure 1 is a conceptual triangular diagram highlights the differences between conventional and unconventional gas reservoirs in terms of increasing reservoir volume,

decreasing permeability and quality, and the need for advanced extraction technologies in unconventional reservoirs. This explains the growing global interest in their development amid the rising demand for energy resources.

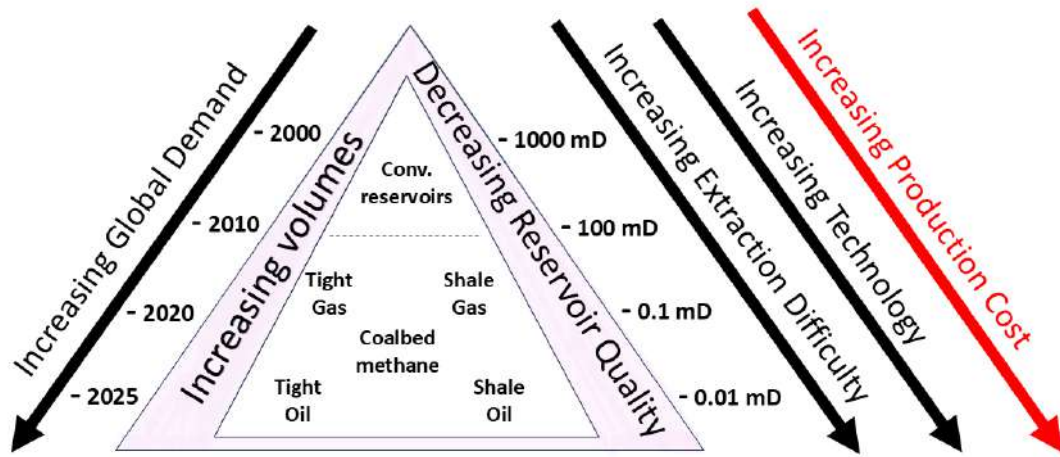


Figure 1. Schematic Diagram Illustrating the Evolution from Conventional to Unconventional Reservoirs

1.1. Type of unconventional reservoirs

1.1.1. Unconventional oil reservoirs

1.1.1.1. Shale oil reservoirs

Shale oil reservoirs are self-sourced reservoirs, where oil is both generated and trapped within the same formation. This means the oil does not migrate long distances to form separate reservoirs (conventional reservoirs) (Taghavinejad et al., 2021). Instead, it remains confined within a dual-porosity rock system comprising nanoscale pores (Guangming, 2008), natural fracture networks, and tight matrices, which impede fluid movement (Clarkson et al., 2016). In this scenario, the source rock generating the oil simultaneously serves as the reservoir rock that stores it.

This type of reservoir forms in marine and lacustrine depositional environments, where fine clay particles and organic matter accumulate. Over millions of years, with increasing sedimentation and absence of oxygen, temperature and pressure rise to between 60–120°C, triggering thermal maturation. This process produces Type I kerogen, rich in algae, and Type II kerogen, rich in marine organic matter, both of which favor oil generation (Kundu, 2023). In contrast, continental shale oil is characterized by Type III kerogen, which ranges from low to high maturity and predominantly produces gas and coal.

1.1.1.2. Tight oil reservoirs

Tight oil is found in low-permeability source rocks where it has migrated only a short distance, forming tight oil reservoirs within porous media (Zeng et al., 2024). These reservoirs are typically categorized as tight sandstones or low-permeability limestones based on their lithology (Fan et al., 2024). The oil in such formations generally has an API gravity above 40°, with porosity and permeability values usually below 12% and 0.1 millidarcy (mD), respectively (Zou, 2013).

Production from tight oil reservoirs requires significant pressure drawdowns, which can reduce reservoir pressure below the bubble point, resulting in two-phase flow (Tabatabaie and Pooladi-Darvish, 2017). Due to the low recovery rates of these reservoirs, a substantial portion of the oil remains trapped, necessitating the use of enhanced oil recovery (EOR) techniques (Muggeridge et al., 2016). Compared to conventional sandstone reservoirs, tight sandstones exhibit higher water saturation, greater rock density, and increased capillary pressure, while also being highly stress-sensitive, meaning their pore sizes fluctuate with changes in net stress (Zou, 2013). Additionally, tight sandstones frequently exhibit abnormally high formation pressures, whereas conventional reservoirs typically operate under normal or subnormal pressure conditions (Zeng et al., 2024). Extracting tight oil requires advanced recovery techniques such as hydraulic fracturing and horizontal drilling to achieve commercial viability (Fan et al., 2024).

Unlike shale oil, which originates from organic-rich source rocks, or oil-shale oil, which is produced through high-temperature processing of immature source rocks, tight oil is located in non-source rock formations and must be stimulated for economic production (Wenzhi et al., 2020). While shale oil extraction is primarily achieved in fractured mudshale, large-scale production from unfractured mudstones or shales remains unproven (Jarvie, 2010).

1.1.2. Unconventional gas reservoirs

1.1.2.1. Coal-bed reservoirs

Coalbed methane (CBM) is considered an unconventional gas resource due to its unique storage and production mechanisms. Unlike conventional gas reservoirs, where hydrocarbons accumulate in porous and permeable rocks, CBM is stored within coal layers, where the coal itself acts as both the source and reservoir rock, this means that little to no migration occurs from the source rock (Moore, 2012; Zou, 2017).

Coalbed methane reservoirs typically contain over 95% methane (CH₄), along with small amounts of heavier hydrocarbons such as ethane (C₂H₆) and propane (C₃H₈), as well as

non-hydrocarbon gases like nitrogen (N_2) and carbon dioxide (CO_2). The porosity and permeability of coal layers are usually below 10% and 0.1 mD, respectively, making gas extraction challenging (Flores, 2013). Unlike conventional reservoirs, where gas is primarily stored in pore spaces, CBM is held within the coal matrix through adsorption, alongside free gas present in fractures (cleats). The adsorption mechanism dominates gas storage in CBM, whereas in shale gas reservoirs (SGRs), free gas is more prevalent. Both CBM and SGRs rely on natural fractures for gas flow; however, the contribution of each storage and flow mechanism differs between the two (Sharma et al., 2019). This fundamental distinction affects exploration, reservoir characterization, well completion, and production strategies. Given these factors, flow modeling approaches used for SGRs can be applied to CBM reservoirs, albeit with modifications to account for differences in gas desorption and transport behavior (Erdenetsogt et al., 2010).

1.1.2.2. Shale gas reservoirs

Shale gas reservoirs are considered a type of unconventional hydrocarbon reservoir, where the shale layer simultaneously acts as a source rock, a reservoir, and a seal for hydrocarbons. These fine-grained sedimentary rocks contain a high proportion of organic matter and are characterized by extremely low permeability, often within the micro- to nano-Darcy range (Jarvie, 2007; Jarvie, 2012; Syed et al., 2022). Due to this very low permeability, natural gas cannot flow freely within the shale rock, necessitating the use of advanced stimulation techniques such as horizontal drilling and hydraulic fracturing to enhance permeability and economically produce gas (Gregory et al., 2011; Zhang and Tingyun, 2015; Zou et al., 2017).

Gas is stored within shale through several mechanisms: as free gas in pores and fractures, adsorbed gas on the surface of organic matter and clay minerals, and sometimes as dissolved gas within the organic matter. Adsorbed gas can constitute up to 85% of the total gas present in thermally mature, organic-rich rocks (Zhang, 2012; Mavko et al., 2020). The pore structure in shale is highly complex, ranging from ultra-micropores (<0.7 nm) to macropores (>50 nm), which significantly affects gas transport mechanisms within the rock, such as diffusion, slip flow, and Knudsen flow (Ross and Bustin, 2009).

Shale formations often exhibit a dual-porosity system, where both the rock matrix and natural fractures contribute to fluid storage and flow; however, fractures often dominate the effective permeability and producibility (Zou et al., 2017; Syed et al., 2022). The efficiency of shale gas reservoirs depends on several critical petrophysical and geomechanical factors,

including total organic carbon content (TOC), thermal maturity (commonly measured by vitrinite reflectance), porosity, brittleness (related to quartz and carbonate content), and clay mineral composition (Jarvie, 2007; Zhang, 2012; Chen and Xiao, 2014; Mavko et al., 2020). Mechanical properties of the rocks, such as Young's modulus and hardness, also play an important role in designing drilling and hydraulic fracturing operations to ensure economically viable and effective reservoir exploitation (Syed et al., 2022).

Hydraulic fracturing technology, which involves injecting a mixture of water, sand, and chemical additives at high pressure, is used to open and extend natural fractures or create new fractures within the shale, thereby enhancing permeability and facilitating gas flow to the wellbore (Zhang and Tingyun, 2015). Optimal site selection for drilling and fracturing is closely linked to a detailed understanding of the reservoir's geomechanical and petrophysical characteristics, including fracture distribution, organic content, and thermal maturity (Zhang, 2012).

1.1.2.3. Tight gas Reservoirs

Tight reservoirs are geological formations characterized by very low porosity, often less than 10%, and extremely low permeability, generally below 0.1 millidarcy, which significantly limits the ability of the rocks to transmit fluids naturally (Ahr, 2011; Tiab and Donaldson, 2024). These reservoirs include various rock types such as tight sandstone, siltstone, limestone, and shale, which share physical and geomechanical properties that impede hydrocarbon flow (Worden et al., 2011; Zou et al., 2012; Ilgen et al., 2017). Producing hydrocarbons from these reservoirs presents significant challenges due to their low permeability, necessitating advanced techniques such as horizontal drilling and hydraulic fracturing to stimulate the rock and create artificial flow pathways (Milad et al., 2021; Qun et al., 2022; Wu et al., 2022).

Multiple geological factors contribute to the compaction in these reservoirs, including the long-term overburden pressure from sediment accumulation (Settari, 2002), pressure-induced recrystallization, pore pressure depletion due to production (Swarbrick and Osborne, 1998), as well as thermal and geomechanical effects that lead to pore volume reduction and decreased porosity and permeability (Settari, 2002; Ghassemi, 2012). Studies have shown that processes such as cementation, clay diagenesis, and chemical alterations play key roles in modifying pore structures and affecting reservoir quality (Lai et al., 2016; Lai et al., 2018; Wang et al., 2020).

Tight reservoirs heavily rely on networks of natural or induced fractures to improve connectivity between isolated pores and provide alternative fluid flow paths, thereby enhancing

effective permeability and productivity (Gale et al., 2007; Warpinski et al., 2009; Gupta et al., 2021). The characteristics of these fractures, including orientation, connectivity, aperture, as well as mineral composition and mechanical properties of the rock, influence fluid flow efficiency and reservoir stimulation response (Manzocchi, 2002; Ghassemi, 2012; Gale et al., 2014). Therefore, precise geomechanical characterization of the tight reservoir is essential for optimizing well design and stimulation strategies (Dusseault, 2011; Ghassemi, 2012).

Fractured tight reservoirs form an important subclass where natural fractures play a critical role in controlling hydrocarbon storage and transport properties (Questiaux et al., 2010; Xu et al., 2016). These fractures may be open, partially mineral filled (e.g., calcite or quartz), or completely sealed, and their connectivity and spatial distribution directly impact production performance (Gale et al., 2014). Understanding the interaction between fracture networks and the rock matrix is vital for building accurate reservoir models that enable improved recovery strategies and increased extraction efficiency (Viswanathan et al., 2022; Yu et al., 2023).

Natural fractures in these reservoirs develop due to various factors including reactivation of ancient faults under ongoing tectonic stresses (Roure et al., 2005; , changes in pore pressure from production or environmental shifts, thermal contraction of rocks, and organic matter decomposition within source rocks (Warren and Root, 1963; Segall and Fitzgerald, 1998; Ghassemi, 2012), all contributing to the formation of complex fracture networks that serve as critical fluid flow pathways (Berkowitz, 2002; Gale et al., 2007).

Settari and Walters, 2001; Minkoff et al. 2003; Dusseault. 2011; Hu et al. 2013 indicate that the coupled interaction between fluid flow and geomechanics, especially under compaction and shear conditions, causes dynamic changes in permeability and porosity, affecting pressure distribution within the reservoir and potentially leading to geological and secondary fracture failures that can enhance permeability and improve production.

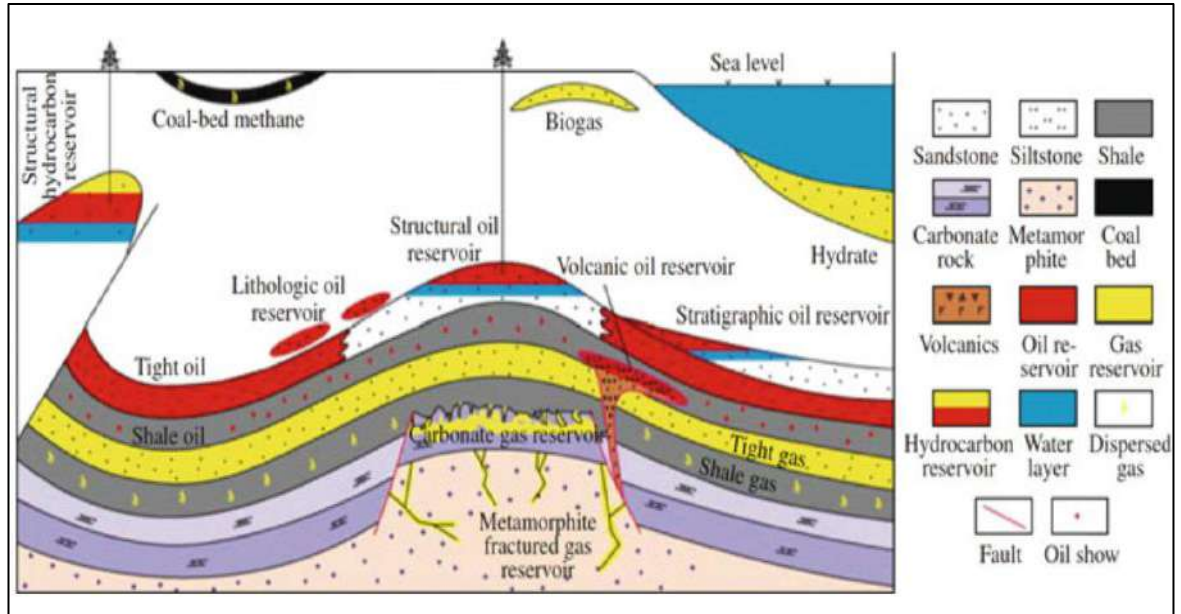


Figure 2. Geological schematic showing the distribution of conventional and unconventional hydrocarbon accumulations (Ahmed and Meehan, 2016).

1.2. Unconventional Reservoirs in The World

Amid the increasing global demand for energy, the past few decades have seen a rapid expansion in the development and utilization of unconventional hydrocarbon reservoirs, which have transitioned from being considered supplementary sources to becoming integral parts of the global energy strategy. This paradigm shift is largely driven by the progressive depletion of conventional oil and gas reserves and the imperative to establish sustainable and secure long-term energy supplies. Unconventional resources including shale oil and gas, tight oil and gas, and coalbed methane are increasingly recognized for their substantial reserves (Hongjun et al., 2016), broad geographic availability, and their critical role in compensating for the declining output from conventional fields (Vedachalam et al., 2015; Caineng et al., 2018). According to the International Energy Agency (IEA), worldwide production from unconventional hydrocarbons exceeded 25 million barrels of oil equivalent per day in 2023, marking a significant contribution to the global energy supply (Papon, 2025).

North America remains at the forefront of unconventional hydrocarbon exploitation, with the United States leading the charge through prolific shale plays such as the Permian Basin, Bakken, Marcellus, Eagle Ford, and Barnett formations, collectively producing over 9 million barrels per day (Soeder, 2018; Schultz et al., 2020). Canada has established itself as a significant force in the coalbed methane (CBM) sector, particularly within the Alberta Basin (Gunter et al., 1997). As of 2005, there were over 7,764 CBM wells operating in Alberta (Mayer et al., 2007), producing approximately 150 million cubic feet per day (MMcf/d) (Hein et al., 2012).

Production continued to increase significantly, with daily output approaching 500 million cubic feet in subsequent years (Hacquebard, 2002). This rapid growth reflects technological advancements, such as the use of nitrogen-based hydraulic fracturing, along with strong infrastructure and favorable regulatory environments (Orujov et al., 2023).

In South America, Argentina has emerged as a key shale oil and gas producer, largely due to the Vaca Muerta Formation, which ranks among the largest shale resources globally outside North America, with estimated recoverable reserves exceeding 16 billion barrels of oil (Rosa and D'Odorico, 2019; Minisini et al., 2020; Collins et al., 2021).

Europe possesses significant geological potential for the development of unconventional hydrocarbon resources, particularly shale gas and tight gas (Caineng et al., 2015), in countries such as the United Kingdom, Poland, and Ukraine. Key formations in this context include the Bowland Shale in the UK, with estimated reserves ranging from approximately 822 to 2,281 trillion cubic feet of gas (Clarke et al., 2018), the Baltic Basin in Poland, and the Yuzivska and Oleska formations in Ukraine, which contain over 3.5 trillion cubic meters of gas (Stulberg, 2015; Lazaruk and Karabyn, 2020). Despite these substantial reserves, commercial-scale production has yet to be realized, with activities remaining limited to exploration and appraisal phases (Syed et al., 2022). This delay is primarily attributed to stringent environmental regulations imposed on extraction technologies especially hydraulic fracturing as well as widespread public opposition to these techniques in several European countries (Gény, 2010; Tolmachev et al., 2020).

Asia holds extensive unconventional reserves, particularly in China and India. China has made significant strides in the Sichuan and Tarim Basins, achieving shale gas production exceeding 20 billion cubic meters in 2022 (Fengyin et al., 2023; Chen et al., 2024). India has prioritized the development of coalbed methane (CBM) resources, particularly in the Damodar Valley, which forms part of the Gondwana coal belt and is characterized by substantial Permian-age coal seams (Singh and Hajra, 2018). The nation possesses more than 2.6 trillion cubic meters of CBM reserves (Joshi et al., 2013). In addition, India is actively investigating shale gas potential in sedimentary basins including Cambay, Krishna-Godavari, Cauvery, and Assam-Arakan (DWIVEDI, 2016). Additionally, Southeast Asian countries, such as Indonesia and Pakistan, are currently assessing their unconventional hydrocarbon resource potential. In Indonesia, the focus is primarily on coalbed methane (CBM) in the Barito and South Sumatra basins (Kim et al., 2016), where the coal-bearing formations date back to the Eocene and Miocene epochs, with estimated resources exceeding 450 trillion cubic feet (Acquah-Andoh et al., 2019). In Pakistan, promising shale gas opportunities exist within mature organic-rich

Jurassic and Cretaceous shale formations in the Indus Basin and Sulaiman ranges, with recoverable gas estimates exceeding 100 trillion cubic feet (Muther et al., 2022).

Africa, while still underexplored, possesses noteworthy unconventional hydrocarbon potential. Algeria ranks among the countries with the largest technically recoverable shale gas resources estimated at over 700 trillion cubic feet primarily within the Illizi, Ahnet and berkine basins (Galeazzi et al., 2010; Layachi, 2013; Smith et al., 2017). South Africa's Karoo Basin, especially the Permian-age Ecca Group shales, along with Egypt's Western Desert highlighting the Jurassic-Cretaceous Khatatba and Alam El Bueib formations represent significant prospects for the development of unconventional hydrocarbon resources (Belica et al., 2017; Bosworth and Tari, 2020). In Egypt, exploration efforts in the Western Desert have identified recoverable shale oil reserves exceeding 4.6 billion barrels (Dolson et al., 2001).

In Australia and Oceania, unconventional energy development, particularly coalbed methane (CBM), is well advanced, especially within the Bowen Basin, which dates back to the Permian period, and the Surat Basin, Jurassic to Cretaceous periods (Ham and Kantzas, 2008; Freij-Ayoub, 2012). Australia annually produces over 1,000 petajoules of CBM. The Cooper Basin, from the Permian to Triassic periods, and the Canning Basin, Devonian to the Permian, further exhibit strong potential for shale and tight gas exploitation (Mastalerz and Drobnik, 2020).

Figure 3 illustrates a conceptual diagram depicting the rising production of global unconventional reservoirs, (a). unconventional gas reservoirs (U.S. Energy Information Administration, 2016; (P.E.S.A. Staff, 2024), (b). unconventional oil reservoirs (Gould and McGlade, 2019; Raimi et al., 2023), highlighting their increasing importance in the global energy landscape and emphasizing the need for their continued development and utilization.

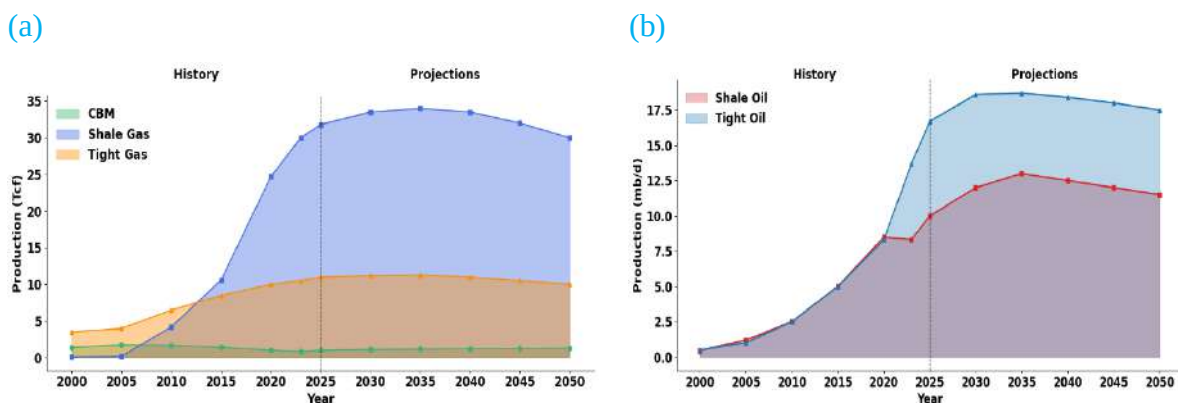


Figure 3. Evolution of Global Unconventional Reservoir Production: History and Projections, (a) unconventional gas reservoirs and (b) unconventional oil reservoirs.

1.2.1. Factors Controlling the Global Distribution of Unconventional Reservoirs

1.2.1.1. Tectonic controls

A key geological factor that controls the distribution and development of shale and tight reservoirs worldwide is each region's tectonic history. This history determines the structural framework, sedimentary settings, and thermal maturity needed to form and preserve unconventional hydrocarbon systems (Bois et al., 1982; Shuichang et al., 2011). The most productive shale and tight gas plays globally are typically located within stable cratonic interiors (Mann et al., 2003) or foreland basins areas characterized by long term tectonic stability or consistent compressional forces during mountain building events (Reading, 2009; Allen and Allen, 2013). Such environments favor the deposition of thick, organic rich shales (DeCelles and Giles, 1996), often formed under oxygen poor (anoxic) marine conditions during periods of rising sea levels that flooded continental margins and low-lying areas (Beaumont, 1981; Lash and Engelder, 2011).

One of the most prominent geological models illustrating the influence of tectonic history on the distribution of shale and tight reservoirs is found in north America, shale formations like the Marcellus, Barnett, and Eagle Ford developed within foreland basins influenced by the Appalachian orogeny, providing ideal conditions for organic rich shale accumulation (Lash and Engelder, 2011). In South America, Argentina's Vaca Muerta Formation formed in a back-arc basin associated with Andean subduction, boasting high total organic carbon (TOC) and well-preserved organic material (Tenaglia et al., 2020). Likewise, China's Sichuan Basin, an intracratonic basin noted for its tectonic stability (Liu et al., 2021), underwent several marine transgression events that promoted the development of prolific shale units such as the Longmaxi Shale (Beagle, 2025). Despite their diverse locations, these basins share tectonic characteristics that have enabled the formation of abundant unconventional hydrocarbon deposits.

1.2.1.2. Paleogeographic Evolution and Continental Drift

The paleogeographic evolution of the continents has had a fundamental impact on the formation and distribution of unconventional hydrocarbon resources (Rogner, 1997). During the Palaeozoic and Mesozoic eras, the Earth's landmasses were joined together in a massive supercontinent Pangea (Evans, 2013). This configuration created extensive inland seas and marine environments where large quantities of organic-rich shales were deposited over millions of years (DiMichele, 2014; Ogg, 2015). These shale formations later became important source rocks for hydrocarbons (Olsen, 1997; Lučić and Bosworth, 2018). As Pangea began to break

apart due to plate tectonic activity, the once-continuous sedimentary basins were split into separate continental blocks (Stampfli et al., 2013). This continental drift led to the current distribution of basins across different continents, many of which still contain the same types of sedimentary rocks and organic-rich layers that are now considered targets for unconventional resource development (Caineng et al., 2015). For instance, the Karoo Basin in South Africa and the Paraná Basin in Brazil were once connected as part of the ancient supercontinent Gondwana (Milani and De Wit, 2008). This common geological background has resulted in similar stratigraphic layers and depositional settings in both basins, including substantial deposits of organic-rich shale (Milani and De Wit, 2008; Linol et al., 2014). The Karoo Basin features extensive rock sequences from the Paleozoic to Mesozoic eras, with notable coal-bearing layers, glacial sediments, and thick black shales within the Ecca Group, making it a prime area for shale gas exploration (KING, 1958; Paul et al., 2025). Similarly, the Paraná Basin is one of South America's largest intracratonic basins, containing Paleozoic and Mesozoic deposits such as the organically rich Irati Formation, regarded as one of Brazil's key potential shale gas reservoirs (Milani and Zalán, 1999; Rocha et al., 2020; Martins et al., 2023). These common geological traits have contributed to ongoing exploration efforts for shale gas and unconventional hydrocarbons in both regions.

1.2.1.3. Depositional Environment and Organic Matter Preservation

Depositional environment and sedimentation rate strongly control the spatial distribution of unconventional reservoirs (Caineng et al., 2015). Shale formations with high total organic carbon (TOC), like the Barnett Shale in the U.S. and the Longmaxi Shale in China, were deposited in deep marine settings characterized by low oxygen levels and minimal energy, which favored the preservation and buildup of organic material (Jarvie et al., 2007; Borjigin et al., 2017). The slow sedimentation in these environments contributed to increased organic carbon concentration (Canfield, 1994; Burdige, 2007). Additionally, factors such as thermal maturity and burial history play a crucial role in the generation and release of hydrocarbons from these source rocks (McCarthy et al., 2011; Al-Areeq, 2018). Adequate burial depth combined with proper geothermal gradients (Kolawole and Evenick, 2023), as seen in the Alberta Basin in Canada and the Ordos Basin in China, facilitates the conversion of kerogen into extractable hydrocarbons (Li et al., 2012; Ardakani et al., 2017).

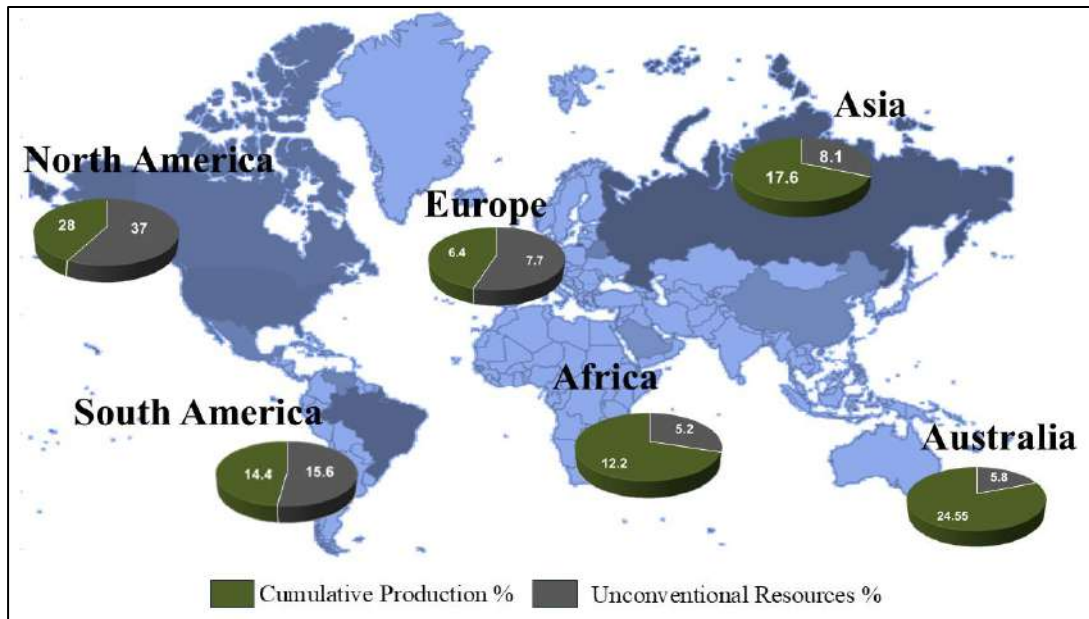


Figure 4. Global distribution map of cumulative production and unconventional oil and gas reserves (Data compiled from; [World Energy Council, 2013](#); [EIA, 2013](#))

1.3. Unconventional Reservoirs in Algeria

Unconventional hydrocarbon reservoirs in Algeria, such as tight gas, shale gas, shale oil, and coalbed methane, have gained significant interest recently due to rising global energy demands and the gradual decline of conventional reserves ([Abdulhadi et al., 2025](#)). Algeria possesses extensive sedimentary basins including Berkine, Illizi, Ahnet, Mouydir, Timimoun, Reggane, and Tindouf, many of which contain promising unconventional resources ([Galeazzi et al., 2010](#); [Kaced and Arab, 2012](#)). These basins feature source rocks and reservoirs characterized by low permeability, high organic content, and the need for advanced extraction methods like horizontal drilling and hydraulic fracturing ([Lučić and Bosworth, 2018](#); [Baouche et al., 2021](#)).

The Berkine Basin, situated in the Central Sahara, is a key area for tight and shale gas studies. Its Cambro-Ordovician sandstone formations, notably Rhourd Enouss, the Tadrart (Upper Cambrian) and In Salah (Lower Ordovician) ([Askri et al., 1995](#); [Djouder, 2019](#); [Fello et al., 2024](#)), have deep, quartz-rich reservoirs with very low porosity (3–8%) and extremely low permeability, classifying them as tight gas reservoirs ([Benayad et al., 2013](#); [Troudi et al., 2018](#)). These formations are typically overpressured and thermally mature, requiring hydraulic stimulation to be economically productive ([Economides and Nolte, 2000](#)). Shale gas has also been discovered in this basin, mainly in the Silurian Tanezzuft Formation and Devonian

Frasnian shales, which are thick, rich in total organic carbon, and thermally mature, favoring gas generation and storage (Soua, 2014; Fello et al., 2024).

Shale gas potential is also significant in the Ahnet and Mouydir basins, where thick and overmature Paleozoic organic shales lie at depths over 3500 meters (Kaced and Arab, 2012). These basins have compact source reservoir systems with minimal hydrocarbon migration, necessitating stimulation to release gas (Yahi et al., 2001; Rossi et al., 2002). They are regarded as some of the most promising yet underexplored formations in North Africa (Radwan, 2024).

The Timimoun and Reggane basins in southwestern Algeria hold substantial tight gas accumulations within Jurassic and Early Cretaceous sandstone layers (Bekkouche et al., 2011; Arab and Djezzar, 2011). These reservoirs are deeply buried (3000-4200 meters), poorly permeable, and overpressured, preserving dry gas but requiring stimulation techniques (Djezzar et al., 2022). Despite being sometimes informally mentioned as chalk reservoirs, these sandstones are siliciclastic and tight rather than carbonate chalks. (Beydoun, 1995; Radwan, 2024).

Coalbed methane is a less developed unconventional resource but shows potential in the western Saharan basins, mainly around Bechar and Tindouf, where Carboniferous and Permian coal-bearing sequences are present. Although data is limited, initial geochemical and stratigraphic evidence indicates viable coalbed methane plays (Baouche and Wood, 2020).

Exploiting these unconventional resources is technically and economically challenging. Their seismic signatures are often weak or ambiguous, demanding sophisticated 3D seismic and integrated petrophysical modeling (Bacon et al., 2007; Aminzadeh and Dasgupta, 2013). Development also requires costly drilling and stimulation efforts. Nonetheless, Algeria's strategic goal to diversify its energy portfolio and maintain its key role as a global gas supplier has led to increased focus on evaluating and developing these resources, often through partnerships with international companies and research groups (Stambouli, 2011; Abada and Bouharkat, 2018).

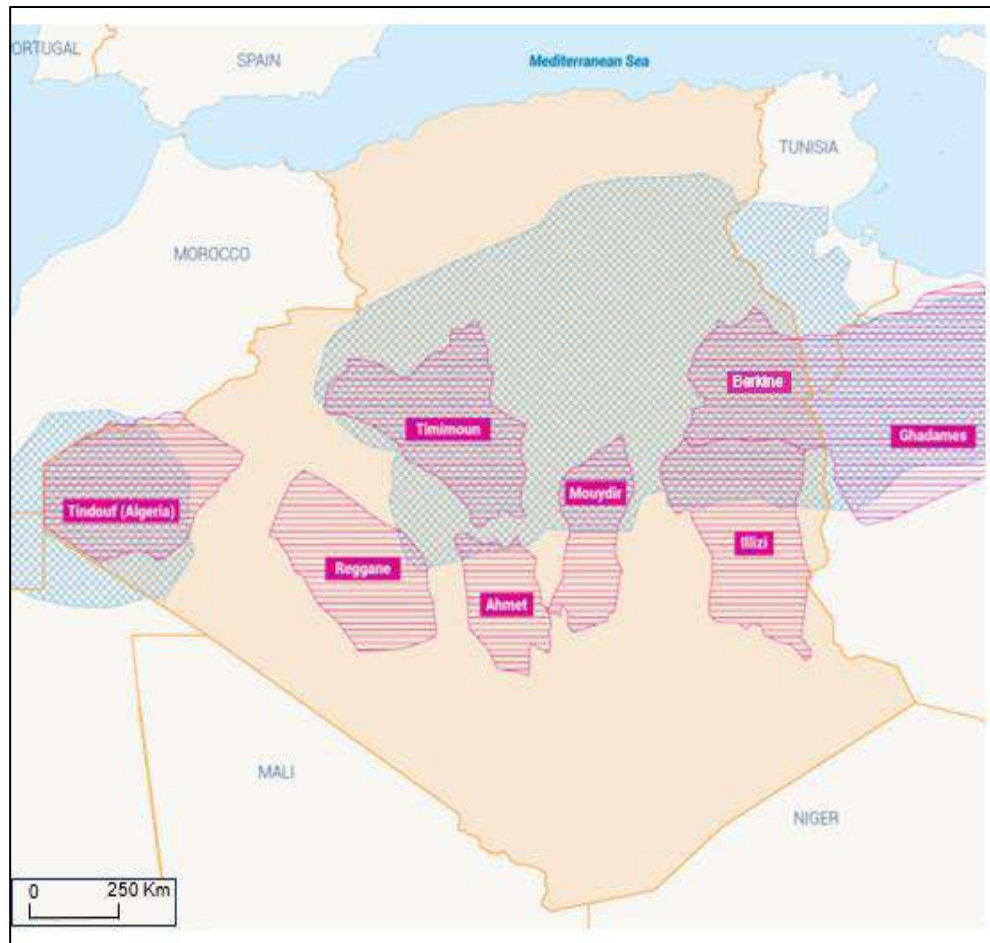


Figure 5. Geological Distribution of Unconventional Hydrocarbon Systems within major sedimentary basins in Algeria. (source: Webography)

1.4. Natural Fractures

Natural fractures are cracks or breaks in rocks that form naturally over geological time due to various geological processes and tectonic or mechanical stresses (Gudmundsson, 2011; Hu, 2022). Leads to a break within a rock that occurs when cohesion between two surfaces is lost (Van Golf-Racht, 1982; Zhang and Yin, 2017). In experimental structural geology, fractures are defined as deformations that occur when cohesion is temporarily lost or when resistance to differential stress fails, resulting in the release of stored energy (Neuendorf et al., 2005). These fractures develop from several causes, including tectonic forces related to plate movements, mountain formation, or continental rifting (Molnaret al., 2007); uneven compaction as sediment layers build up (Maltman, 2012); thermal expansion or contraction of rocks (Atkinson, 2015); and pressure reduction caused by erosion or the removal of overlying material (Gale et al., 2014). The fractures differ in size, spacing, and direction and may appear as isolated cracks or as complex interconnected networks (Gale et al., 2014; Wang, 2019). They

have a direct impact on the permeability and fluid storage capacity of the rock (Jones, 1975; Guerriero et al., 2013; Rashid et al., 2023). In rocks with low permeability, such as tight sandstones or oil shales, these natural fractures act as essential pathways that enable the movement of oil and gas, which otherwise cannot flow through the rock's matrix (Gale et al., 2014; Tiab and Donaldson, 2024). The importance of these fractures is influenced by the rock's mechanical properties, including stiffness and brittleness, the history of stress, and the structural setting (Paterson and Wong, 2005; Atkinson, 2015; Li et al., 2018). They are crucial for planning drilling and stimulation methods like hydraulic fracturing and for assessing the productivity potential of unconventional reservoirs (Cramer, 2008; Warpinski et al., 2009).

1.4.1. Type of Natural Fractures in reservoirs

The classification of natural fractures is considered a fundamental process for understanding their origin and role in petroleum reservoirs (Bratton et al., 2006). Many researchers have proposed various classifications based on structural, hydrodynamic, morphological, and even geochemical criteria. Morphological classification focuses on the geometric shape of fractures (length, width, and orientation) (Müller et al., 2012; Gale et al., 2014), while temporal classification distinguishes between primary fractures related to deposition and secondary fractures caused by subsequent tectonic movements (Bratton et al., 2006; Laubach et al., 2009). Additionally, some studies have proposed geochemical classifications concerned with the mineral composition filling the fractures, reflecting pressure and temperature conditions during geological evolution stages (Gudmundsson, 2011; Ukar et al., 2020). At the structural level, Gudmundsson (2011) suggested a classification based on the dominant stress type, dividing fractures into tensile, shear, and compressive types. From a hydrological and petroleum perspective, Van Golf-Racht (1982) and Nelson (2001) categorized fractures as open, semi-open, closed, or mineral-filled, depending on their permeability to fluids.

1.4.1.1. Structural classification

This classification is based on how fractures form under different tectonic stresses. According to Gudmundsson (2011), fractures fall into three key types:

a. Tensile fractures

Develop when rocks experience stress pulling them apart perpendicular to the fracture surface (Gudmundsson, 2011; Atkinson, 2015), causing distinct openings usually aligned perpendicular to the highest principal stress (Lang et al., 2018). These fractures effectively facilitate fluid storage and flow, especially in low-permeability reservoirs (Wang et al., 2024).

b. Shear fractures

Arise from shearing forces where rock blocks slide past each other along a fracture plane (Atkinson, 2015). They are generally inclined relative to tectonic stress and often associated with microfaults (Wibberley et al., 2000; Anders et al., 2014). Despite lower permeability than tensile fractures (Zhang et al., 2017), they connect fracture networks to maintain fluid flow continuity in reservoirs (Zhang et al., 2009).

c. Compression fractures

Form under strong compressive forces (Goldstein and Osipenko, 2019), fragmenting rocks into sets roughly parallel to the stress direction (Atkinson, 2015; Li et al., 2018). These fractures tend to be closed or have very narrow openings, limiting fluid flow (Snow, 1968; Zhang et al., 2009). Despite their limited hydraulic role, compression fractures hold significant geological value as they preserve evidence of past tectonic events and may enhance the mechanical integrity of the rock mass (Gudmundsson, 2011; Fossen, 2016).

1.4.1.2. Hydrological Classification

This classification focuses on how natural fractures influence fluid passage, a critical factor for reservoir performance. Van Golf-Racht (1982) and Nelson (2001) identify four main categories:

a. Open fractures

Have empty internal spaces that markedly increase permeability and act as primary channels for petroleum, gas, and groundwater flow, often corresponding to tensile fractures (Stearns and Friedman, 1972; Aydin, 2000).

b. Partially open fractures

Contain some mineral deposits but still permit reduced fluid flow, common in reservoirs subjected to repeated tectonic events leaving fractures partially open (Wennberg et al., 2016; Ruse, 2024).

c. Closed fractures

Form through mineral recementation (Kenzie, 2024) or tectonic compression that closes fracture openings (Germanovich and Dyskin, 2000). These are hydraulically inactive but reflect previous tectonic phases that modified the fracture system (Dahi, 2009; Gale et al., 2014).

d. Mineral-filled fractures

Are fully occupied by minerals like calcite, quartz, or barite, making them impermeable (Farag et al., 2010; Wright et al., 2023). Though they do not conduct fluids (Bratton et al.,

2006), they provide valuable geological insights into mineralization and tectonic evolution for reservoir history reconstruction (Bratton et al., 2006; Bons et al., 2022).

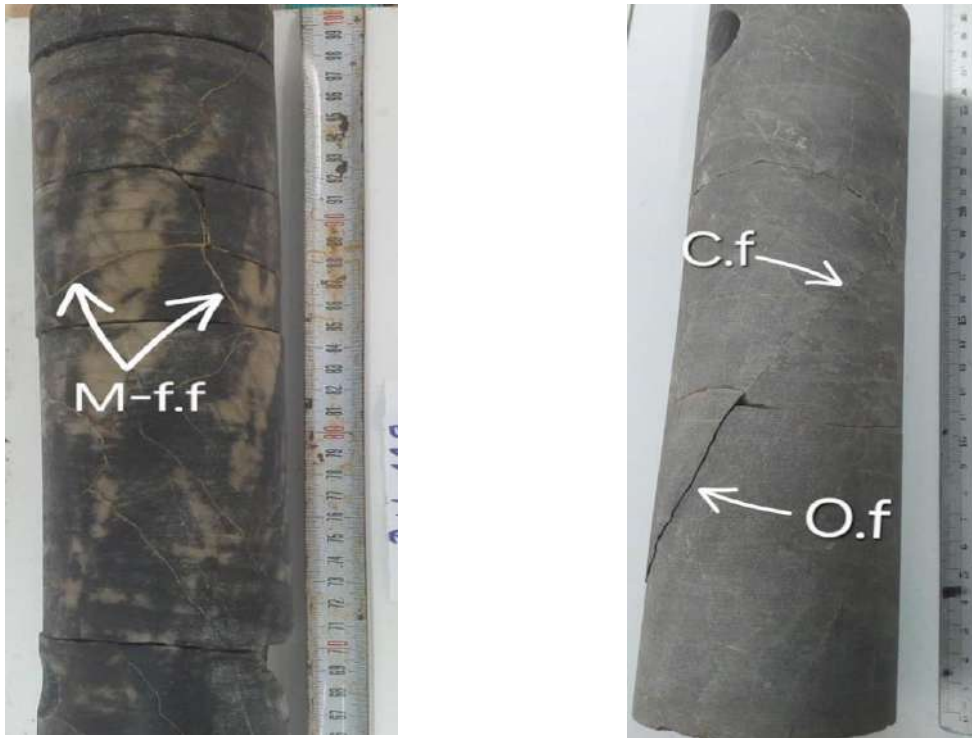


Figure 6. Types of natural fractures in reservoir cores: open (O.f), closed (C.f) and mineral-filled fracture (M-f.f).

1.4.2. General Importance of Natural Fractures in reservoirs

The impact of natural fractures on petroleum reservoirs is a critical topic requiring a thorough understanding of their multiple roles in either enhancing or complicating reservoir exploitation.

1.4.2.1. Positive Effects of Natural Fractures

a. Enhancement of Permeability and Productivity

Open, interconnected fractures make it easier for fluids to flow freely through otherwise tight, dense rocks (Sahimi, 2011). These fractures play a crucial role in boosting oil and gas production in low-permeability reservoirs by effectively replacing poor natural rock permeability (Kharrat and Ott, 2023; Yin et al., 2024). For example, in shale gas reservoirs, fractures form the main channels allowing gas flow to wells, making extraction feasible and economically viable (Qun et al., 2022).

b. Improved Hydraulic Fracturing Efficiency

The flexibility and correct orientation of natural fractures facilitate the propagation of induced fractures during hydraulic fracturing operations (Gale et al., 2007; Tang et al., 2022). This process opens and extends the fracture network, increasing the surface area of rock-fluid contact, which enhances well productivity (Zhang et al., 2009).

c. Role in Geomechanical Equilibrium

Natural fractures help release localized stresses within the reservoir, particularly in deep drilling zones, preventing unexpected collapses or cracks that could halt production or damage equipment (Ahmed and Salehi, 2021). This stress redistribution maintains wellbore stability, contributing to safer operational conditions (Jalali and Dusseault, 2012).

1.4.2.2. Negative Effects of Natural Fractures

a. Loss of Drilling Fluids

When open fractures connect to the reservoir surface or extensive fracture networks, they create pathways for drilling fluids to escape from wells, leading to technical challenges such as fluid loss (Epelle and Gerogiorgis, 2020; Yang et al., 2022), increased drilling time, and higher costs (Taleghani et al., 2016).

b. Heterogeneity and Impact on Production

Heterogeneous fractures especially those partially closed or mineral-filled create uneven fluid flow paths (Kharrat and Ott, 2023), producing non-uniform reservoir drainage (Gonzalez and Aguilera, 2013). This phenomenon results in some rock areas being bypassed, reducing overall production efficiency (Khan and Mandal, 2022).

c. Wellbore Instability

Misalignment of fracture orientations with prevailing stresses may cause rock around wells to expand or slip (Al-Bazali et al., 2009), triggering localized collapses or leaks that threaten well integrity and reservoir stability, particularly in deep or geologically complex reservoirs (Ahmed and Salehi, 2021).

Chapter II Regional and Local Geology of the Study Area

2.1. General Overview of the Berkine Basin

The study area, Rhourde Nouss field, belongs to the Berkine basin (the formerly giant Ghadames basin), which is a tectonic depression dating from the Paleozoic to Early Mesozoic periods, situated on the Saharan Platform just south of the Atlas Orogenic Belt (see Figure 7). It originated during the Paleozoic as an intracratonic depocenter characterized by thick sedimentary layers cover of more than 7,000 m, deposited atop a Proterozoic basement made up of metamorphic and igneous rocks. A notable Late Paleozoic unconformity separates these early sediments from the younger Mesozoic sediment wedge, which thickens toward the north and forms part of the passive margin of northwestern Africa related to the Tethys realm (Dardour et al., 2004; Galeazzi et al., 2010; Aissaoui et al., 2016; Fello et al., 2024).

The Proterozoic basement formed during the Pan-African Orogenic event, while the subsequent tectono-sedimentary cycles tied to Gondwana, the Tethyan Ocean, and Alpine movements have shaped the basin's structural and stratigraphic framework (Bouabdellah and Slack, 2016). The sedimentary record throughout the Phanerozoic, along with associated unconformities and structural elements, encapsulates the basin's geological evolution (Galeazzi et al., 2010).

Alongside nearby structures such as the Tihemboka, Amguid–El Biod, and Hassi Messaoud highs, the Berkine Basin constitutes a key petroleum province of global significance (Eschard et al., 2010; Fello et al., 2024). The estimated ultimate recovery is around 39 billion barrels of oil equivalent (BBOE). In particular, the Berkine Basin together with the El Biod high contains proven reserves exceeding 20 billion barrels of oil and about 43 trillion cubic feet of natural gas (IHS Database, 2008) (Galeazzi et al., 2010).

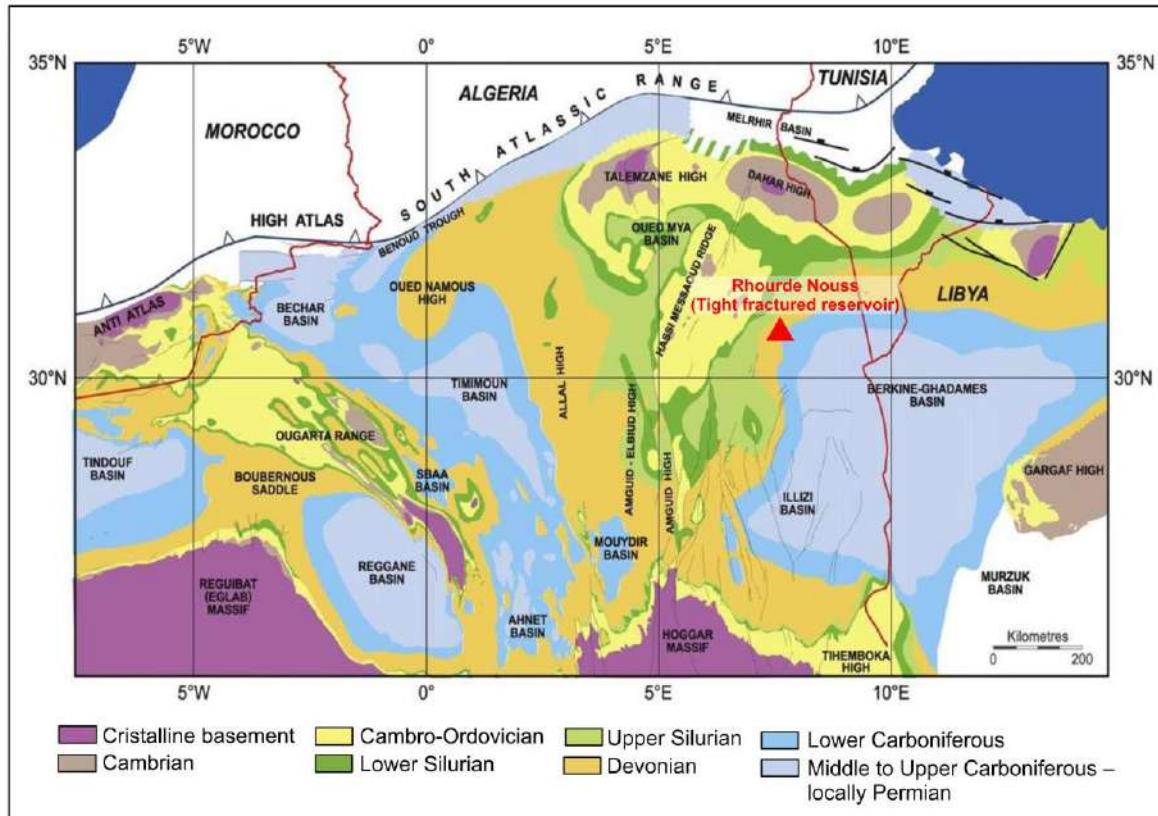


Figure 7. Improved satellite image of the Maghreb region Northwest Africa, illustrating the primary tectonic domains of the area and the position of the study area Rhourde Nouss (Galeazzi et al., 2010).

2.2. Geographic framework of the Berkin basin and location of the Rhourde Nouss field

The Berkine Basin represents a broad geological extension within the Eastern Saharan Platform, it's located between latitudes 29° and 34° North and longitudes 5° and 10° East. Its geographical boundaries are defined as follows:

- To the east: the Sirte Basin, extending between Tunisia and Libya.
- To the south: the Illizi Basin.
- To the west: the Amguid–El Biod–Hassi Messaoud structural high.
- To the north: the Ain-Roumana ridge and the Dahar dome.

The Rhourde Nouss field extends between longitudes 06°24' E and 07°00' E and latitudes 29°16' N and 30°00' N, is part of the Illizi Province, this area covers an estimated 10,000 km². and is located approximately 850 km south of Algiers, 230 km southeast of the Hassi Messaoud field, and 350 km northwest of Aïn Aménas (BeicipFranlab, 2002; WEC, 2007).

- To the north, it borders the Gassi Touil region,
- To the south, it adjoins the Tin Fouyé Tabenkourt (TFT) area.

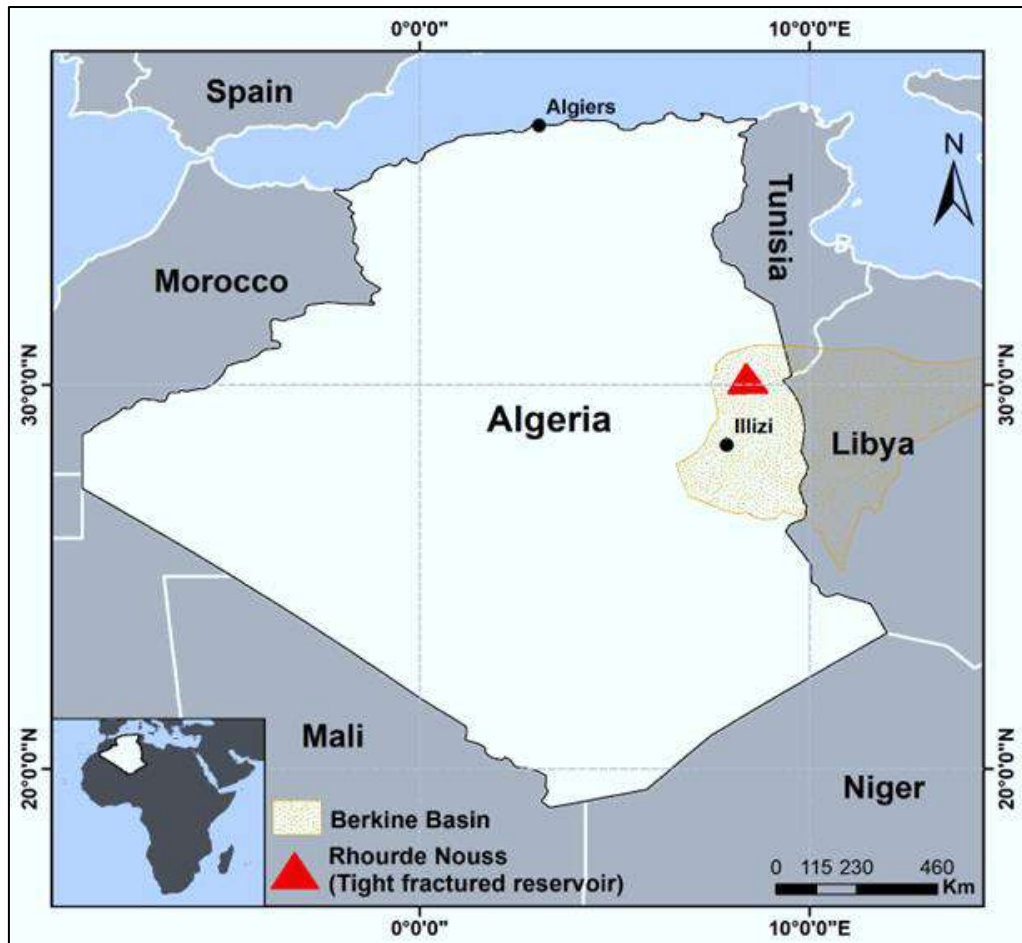


Figure 8. Location of the Rhourde Nouss field study area in Algeria.

2.3. Geological Context

2.3.1. Tectonic Evolution

The region's tectonic history is part of the broader evolution of the Saharan Platform, shaped by ancient pre-Paleozoic structures and major tectonic events such as the Pan-African, Hercynian, and Alpine orogenies (Garfunkel, 2002). It belongs to the Triassic Province, an intraplate epicontinental basin (Perrodon, 1985), and is closely linked to the Amguid El-Biod uplift, which has experienced several tectonic phases throughout geological time (Boudjemaa, 1987). This complex evolution reflects the interaction of tectonic forces that have influenced the region's structural framework and sedimentation patterns.

2.3.1.1. The Pan-African Phase

was marked by the collision between the rigid West African Craton and the more deformable East African Craton, leading to the formation of significant vertical faults oriented approximately north-south across the Precambrian basement (Liégeois et al., 2003; Maddaloni, 2021). These fault systems are believed to have experienced substantial horizontal shifts, with

displacements possibly reaching up to 100 kilometers. Within these prominent fault zones, networks of faults trending northeast-southwest (NE-SW) and northwest-southeast (NW-SE) intersect, reflecting a tectonic regime in which the principal stresses were oriented predominantly east-west (E–W) (Basheer, 2024). This complex fault pattern illustrates the intense tectonic deformation and reactivation that shaped the region's early geological framework (Fowler and Hamimi, 2024).

2.3.1.2. Cambro-Ordovician Distension

A significant erosion event flattened the landscape and structural features, resulting in variations in sediment thickness and facies on either side of the NS fault (Ghienne et al., 2023). This phenomenon is observable within the Cambro-Ordovician stratigraphy and suggests tectonic instability between the more rigid western African block and the comparatively less rigid eastern block, a process that coincided with ongoing sediment deposition (Boudjemaa, 1987; Djezzar and Boualam, 2025).

2.3.1.3. Taconic compression (Caradoc- Ashgill)

Based on Issawi and Farouk (2023), the compression phase, known as Taconic compression, mainly caused considerable uplift of the ancient shield regions, specifically the Reguibat and Targui areas, along with reverse fault activity on faults oriented roughly north-south. Additionally, Zazoun and Mahdjoub (2003) identified a transtensional tectonic phase oriented northwest-southeast within the glacial deposits of Tamadjert (Unit IV) in the Tassili N'Ajjer and Illizi Basin, which coincided in time with the Ashgillian glacial tectonic episode during the Hirnantian stage.

2.3.1.4. Silurian Distension

During the Silurian period, the region experienced tectonic extension following the Ordovician glaciation (Kogbe and Burollet, 1990). The retreat of the ice sheets led to a global rise in sea levels and the development of tranquil, reducing marine settings. These environmental conditions promoted the accumulation of black, organic-rich, radioactive shales, characterized by elevated uranium content. Known as Silurian graptolitic shales, these sediments serve as important stratigraphic indicators and are considered potential source rocks within North African basins such as Berkine and Illizi (Derder et al., 2018).

2.3.1.5. The Siluro-Devonian or Caledonian compression

The sandstone layers found in the southeastern region indicate a transition towards a compressional tectonic setting, marking the initiation of the Caledonian phase. This phase represents a time of tectonic reactivation that affected both the Precambrian basement and the

overlying Paleozoic formations. It is characterized by tectonic uplift, folding, and erosion, reflecting the extensive impact of the Caledonian orogeny along the northern margin of Gondwana. The tectonic influences of this phase are clearly observed in the Saharan basins of Berkine, Illizi, and Ahnet-Mouydir ([Galeazzi et al., 2010](#)).

2.3.1.6. Frasnian distention

Corresponds to the deposition of the Hassi Tabenkort Formation in the Illizi Basin. This period is marked by significant tectono-sedimentary activity, including: thickness and facies variations mainly along major NE–SW faults; areas of non-deposition and localized erosion in the Ahara High; and intense volcanic activity characterized by basic rock flows in the eastern Hoggar region. The extensional phase that influenced the Lower Devonian also led to the formation of two major unconformities: one following the Emsian and another known as the intra-Devonian unconformity ([Galeazzi et al., 2010](#)).

2.3.1.7. Hercynian Unconformity

The Hercynian cycle, encompasses two major tectonic phases. The first, known as the early Hercynian phase, is characterized by a compressive regime-oriented NE/SW, inferred from the observation of striations in the upper Devonian Khenig sandstones. The second, referred to as the major Hercynian phase, is distinguished by compressive stresses oriented NW/SE, which led to the development of parallel fault systems that structured the landscape into horsts and grabens, reflecting a differential deformation of the Earth's crust ([Galeazzi et al., 2010](#); [Ghalgaoui et al., 2023](#)).

2.3.1.8. Austrian phase

The Austrian phase, was triggered by the opening of the South Atlantic Ocean between South America and Africa ([Scotese, 2021](#)). This event caused an east-west compressional stress that reactivated existing faults, creating an echelon fault systems-oriented NE–SW. These structural developments played a key role in facilitating the migration and entrapment of hydrocarbons ([Macgregor and Reeves, 2025](#)).

2.3.1.9. Pyrenean Phase

The Pyrenean Phase, occurring from the Late Cretaceous to the Eocene, represents a NW-SE oriented compressional tectonic event linked to the convergence between the African and Eurasian plates ([Ford et al., 2022](#)). This phase resulted in the development of anticlines and reverse faults, reactivation of pre-existing fault zones, and the creation of structural traps that are conducive to hydrocarbon accumulation. In basins such as Berkine and Illizi, the Pyrenean

Phase played a crucial role in shaping the tectono-sedimentary framework and advancing the petroleum system (Galeazzi et al., 2010; Vergés and Fernández, 2012).

2.3.2. Stratigraphic Aspect

According to Beicip/SONATRACH (2002), the stratigraphic succession in the Rhourde Nouss area is categorized into three primary units that correspond to the major geological eras. The Paleozoic unit is mainly composed of siliciclastic sediments that have been influenced by Hercynian erosion. The Mesozoic unit features alternating layers of clayey sandstone, evaporite, and carbonate formations. Lastly, the Cenozoic unit consists of unconformably deposited Miocene to Pliocene sediments that overlie the older rock layers.

- **The Basement**

Consists of a granitic and/or metamorphic complex that exhibits structural features such as folds and faults. Understanding of the basement in the Rhourde Nouss area is limited due to the few wells that have actually penetrated this

- **The Paleozoic**

The oldest Paleozoic sediments, dating back to the Silurian period, comprise a sequence of clayey sandstone deposited within the context of the Hercynian unconformity. The formations from the Ordovician and Silurian are relatively well-preserved, whereas the Devonian and Carboniferous deposits have experienced substantial erosion.

- **The Cambrian**

The formation mainly consists of Hassi Leïla sandstones, featuring a lower part made up of poorly sorted, fine- to coarse-grained quartzitic sands, and an upper part composed of fine quartzites, with an average thickness of approximately 104 meters.

- **The Ordovician**

This stratigraphic interval is notably complex, comprising alternating layers of quartzitic sandstones and clays. Its thickness typically varies between 750 and 1,250 meters and consists of multiple stacked units, arranged sequentially from the bottom to the top as follows:

- **Miribel Sandstones**

These sandstones, dating back to the Lower Tremadocian, have an approximate thickness of 180 meters. They primarily consist of fine to very fine quartzitic sandstone, interbedded with gray to black clay layers and clayey-silty sandstones.

- **El-Gassi Clays**

from the Middle Tremadocian, are about 268 meters thick and feature alternating layers of quartzitic sandstones and dark gray, occasionally plastic, silty clays.

- **El-Attchane Sandstones**

comprise finely grained, silty, well-sorted quartzitic sandstones interspersed with clay layers.

- **Hamra Quartzites**

dated to the Middle Arenig, vary between 80 and 232 meters thick and consist of gray-brown fine- to medium-grained quartzites with interspersed black clay layers.

- **Ouargla Sandstones**

from the Upper Arenig, range from 91 to 100 meters thick and are composed mainly of gray-brown, fine- to medium-grained quartzitic sandstones with black clay interbeds.

- **Azzel Clays**

dated to the Llanvirn, are between 64 and 117 meters thick and consist of dark gray, silty clays that are variably hardened.

- **Oued Saret Sandstones**

from the Llandeilian, have a thickness of 103 to 137 meters and are fine-grained, sometimes quartzitic.

- **Microconglomeratic Clays**

dating to the Caradoc, are 66 to 119 meters thick and exclusively feature microconglomeratic clays containing quartz grains.

- **M'kratta Bed**

from the Ashgill, is 6 to 10 meters thick and consists of quartzitic sandstones that range from very fine to medium and coarse grains.

- **The Silurian**

The Silurian period in the Rhourde Nouss area is composed of two primary lithological units: a thick, basal clay-rich layer, followed by a more varied clayey-sandstone layer above it.

- **The clayey Silurian**

dated from the Llandovery to Wenlock stages (Llandoveryan-Taranian-Wenlock), have a thickness between 232 and 341 meters. These deposits primarily consist of micaceous black clays that are rich in fossilized fauna such as brachiopods, gastropods, and graptolites. They correspond to the well-known "graptolitic shales," which indicate relatively deep and tranquil marine environments.

- **Clayey-Sandstone Silurian**

Assigned to the Ludlow, this series represents the F6 reservoir and is composed of six superimposed units, numbered from M1 to B2:

- **Unit M1:** Indurated silty clays with a few carbonate intercalations at the base; thickness ranges from 32 to 37 meters.
- **Unit M2:** Indurated silty clays, sometimes interbedded with Mederba sandstones; thickness ranges from 61 to 153 meters.
- **Unit A1:** Fine- to medium-grained sandstones, well consolidated with siliceous cement, interbedded with thin layers of gray to black clays; thickness between 13 and 18 meters.
- **Unit A2:** Fine silty sandstones with clayey interbeds at the base; thickness ranges from 98 to 126 meters.
- **Unit B1:** A succession of compact quartzitic sandstones and dark gray, micaceous, indurated clays. This unit was affected by Hercynian erosion, particularly on the southeastern flank of the Rhourde Nouss structure, and shows a variable thickness from 0 to 149 meters.
- **Unit B2:** Alternation of quartzitic sandstones and micaceous clays, also truncated by Hercynian erosion, mainly in the southeastern and southwestern parts of the studied area. Its thickness varies from 0 to 111 meters.

- **The Devonian, Carboniferous, and Permian**

Are entirely missing in the Rhourde Nouss area due to severe Hercynian erosion that removed these layers. Consequently, the Triassic deposits lie directly on the Silurian, which itself has been locally eroded and truncated by this erosional event.

- **The Mesozoic**

In the Rhourde Nouss area, the Mesozoic is represented by three principal periods: the Triassic, Jurassic, and Cretaceous. This era signifies a phase of renewed sedimentation that followed the Hercynian erosion, which had extensively stripped away the Paleozoic formations.

- **The Triassic**

The Triassic lies unconformably on the clayey-sandstone Silurian, following the Hercynian erosion. It consists of five well-differentiated stratigraphic units, arranged from base to top:

- **Lower Clayey-Sandstone Triassic (TAGI):** Carnian in age, with a thickness ranging from 15 to 57 m. It consists of fine-grained sandstones with siliceous–clayey cement, interbedded with silty clay layers and microconglomeratic beds at the base.
- **Intermediate Triassic I:** Lower Norian in age, with a thickness between 29 and 70 m. It is composed of fine, clayey sandstones at the base, gradually grading upward into a quartzitic texture.

- **Middle Clayey Triassic:** Assigned to the Middle Norian, with a thickness of 33 to 53 m. This unit consists of brown to red, silty and dolomitic clays, sometimes indurated.
- **Intermediate Triassic II:** Upper Norian in age, 53 to 77 m thick, characterized by an alternation of fine sandstones and red to green silty clays.
- **Upper Clayey-Sandstone Triassic (TAGS):** Rhaetian in age, with a thickness ranging from 100 to 198 m. It begins with gray-white, fine-grained sandstones, sometimes microconglomeratic, and is overlain by brown to red, silty, slightly dolomitic clays. This unit forms the main reservoir of the Rhourde Nouss region.

- **The Jurassic:**

The Jurassic period is prominently represented in the Rhourde Nouss region by a thick and varied sedimentary sequence, which records changing depositional environments from evaporitic lagoonal settings to more restricted marine conditions. This sequence is divided into three main stratigraphic units: the Lias, the Dogger, and the Malm.

- **The Lias**

The Lias is divided into five stratigraphic units, ordered from the bottom to the top as follows:

- **The Clayey Lias (S1):** with a thickness of 39 to 62 meters, consists of varicolored clays mixed with silts.
- **The Saliferous Lias (S2):** is a thick layer ranging from 202 to 443 meters and is characterized by massive white to pinkish salt deposits interspersed with brown clay beds, with anhydrite levels at the base serving as markers.
- **Horizon B:** varying from 14 to 42 meters thick, includes gravelly, vacuolar, or silty dolomites with slight anhydrite content.
- **Horizon H:** has a thickness between 13 and 41 meters and consists of compact white anhydrite with a chalky texture and some pink selenite gypsum layers, serving as an excellent regional seismic marker.
- **The Clayey-Dolomitic Lias:** with thickness from 140 to 220 meters, is composed of silty dolomitic clays with some anhydrite interbeds.

- **The Dogger**

The Dogger is made up of two primary units:

- **The Lagoonal Dogger:** which ranges from 0 to 322 meters thick and includes alternating silty clays and fine- to medium-grained sandstones.
- **The Clayey Dogger:** with a thickness of 55 to 228 meters, consists of varicolored clays interspersed with fine sandstone layers.

- **The Malm:** The Malm which varies from 0 to 232 meters thick, mainly consists of soft silty clays, sometimes dolomitic, and represents the Jurassic period's end in the study area.

- **The Cretaceous:**

In the Rhourde Nouss region, the Cretaceous is represented by a thick sedimentary sequence, divided into two main units: the Lower Cretaceous and the Upper Cretaceous. This succession reflects a progressive evolution of depositional environments, from deltaic and coastal settings to deep marine, and sometimes evaporitic environments.

- **Lower Cretaceous:**

- **Neocomian:** Thickness 46-345 m, composed of dolomitic silty clays alternating with fine- to coarse-grained sands and fine to very fine sandstones, occasionally enriched in lignite or pyrite.
- **Barremian:** Average thickness 529 m, characterized by an alternation of sands, fine sandstones, and reddish sandy clays, slightly dolomitic. The upper part marks an Austrian unconformity.
- **Aptian:** Poorly developed (0-22 m), composed of calcareous-dolomitic marls.
- **Albian:** Thickness 30-216 m, composed of fine- to coarse-grained sands and clay-dolomitic sandstones.

- **Upper Cretaceous:**

- **Cenomanian:** Thickness 62-111 m, composed of carbonate clays with anhydrite layers.
- **Turonian:** Thickness 69-113 m, formed of alternating plastic varicolored clays, white to beige limestones (sometimes clayey), and calcareous marls.
- **Senonian:** Subdivided into three subunits:
- **Saliferous Senonian:** Thickness 13-80 m, composed of massive white salt, occasionally interbedded with clays.
- **Anhydritic Senonian:** Thickness 111-233 m, consisting of massive beige anhydrite alternating with carbonate clays.
- **Carbonate Senonian:** Thickness 0-194 m, composed of microcrystalline dolomitic limestones.

- **The Cenozoic**

The Cenozoic is represented only by the Miocene-Pliocene. Older formations Paleocene, Eocene, and Oligocene are absent, likely due to erosion or non-deposition related to Pyrenean tectonics.

• Miocene-Pliocene

With a thickness ranging from 154 to 677 m, it rests unconformably on the Senonian carbonate formations. It is composed of fine to very coarse-grained sands, sometimes sandstone-rich, with interbeds of sandy, slightly dolomitic clays and layers of sandy limestone.

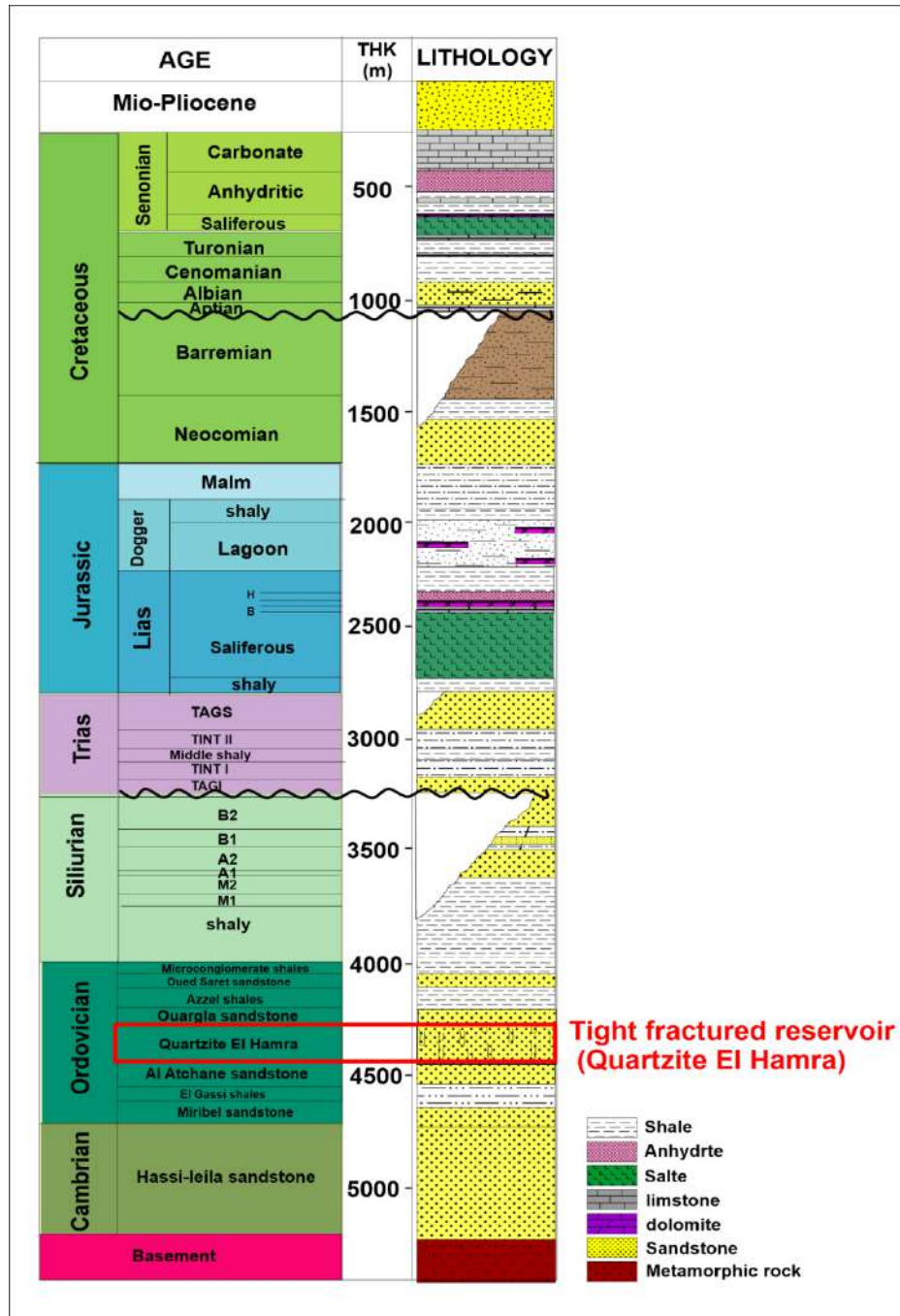


Figure 9. Stratigraphic summary column of the Rhourde Nouss field in Algeria, highlighting the El Hamra Quartzite formations (Beicip/SONATRACH, 2002).

2.4. Petroleum Aspect of the Rhourde Nouss field

The petroleum system in the Rhourde Nouss region stems from an extended and intricate geological history that promoted hydrocarbon generation, migration, accumulation, and trapping. It depends on the interaction of superior source rocks, effective migration routes, robust reservoir layers, and reliable top seals. Analyzing these components and their interconnections proves crucial for assessing the area's hydrocarbon prospects, especially amid coexisting conventional and tight reservoirs.

2.4.1. Source Rocks

The primary petroleum system in the Rhourde Nouss field is tied to the basal Silurian source rocks, while a secondary system relates to the deeper Hamra Quartzite reservoir. In practice, the Silurian shale stands as the sole significant source rock locally, exhibiting marked variations in generative capacity: the basal section is organically rich and aligns with the radioactive peak but remains quite thin, whereas the middle and upper sections are substantially thicker yet hold only mediocre to poor organic content ([Beicip Report, 1995](#)). For the Hamra Quartzite, potential input from other Ordovician clay-rich units like the El Gassi Shales cannot be discounted ([WEC, 2007](#)).

2.4.2. Sealing Rocks

The sealing formations within the petroleum system of the study area are highly efficient. The Liassic evaporites create a continuous, robust regional barrier that overlays all the structural highs in the area. For the Hamra Quartzite reservoir, the trapping and preservation of hydrocarbons are ensured by the Azzel Shales, which act as a reliable seal.

2.4.3. Traps

The primary hydrocarbon traps in the region developed during the Austrian tectonic phase, leading to substantial structural closures impacting both the Triassic reservoirs and the Silurian clayey-sandstone layers. It is thought that these traps were predominantly filled by hydrocarbons migrating laterally over long distances from the west and north, with possible contributions from vertical migration along fault networks ([Beicip Report, 1995](#)).

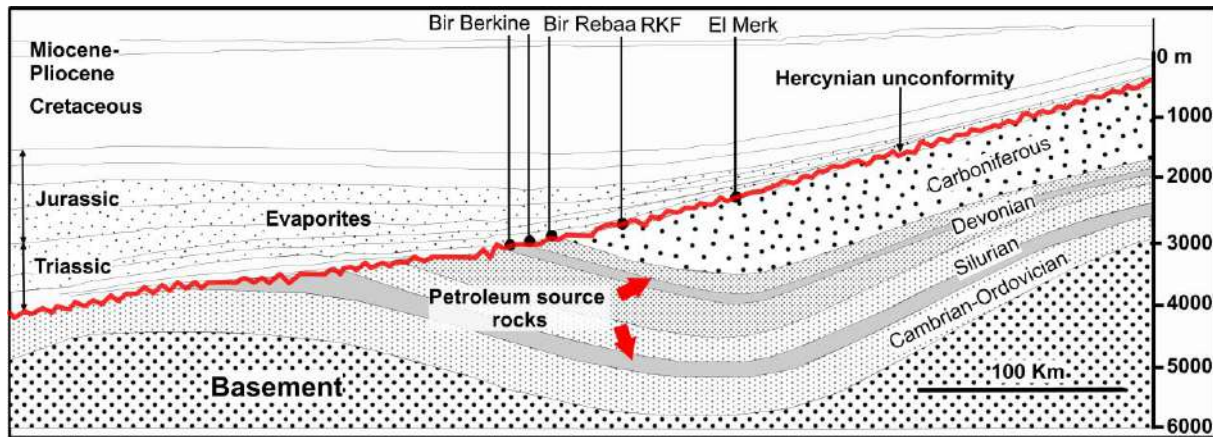


Figure 10. Petroleum Sources in the basin of Berkine (Boote et al., 1998)

2.4.4. Reservoir Rocks

In the Rhourde Nouss field, hydrocarbon deposits yield oil and condensate-rich gas from multiple reservoir intervals, classified into three main reservoir systems:

2.4.4.1. The Upper Argillaceous–Sandstone Triassic formation (TAGS)

The foremost reservoir, which varies in thickness from 100 to 200 meters throughout the field. This formation is divided into two parts: a lower segment consisting of coarse, laterally extensive sandstones interpreted as braided river channel deposits, and an upper segment characterized by greater heterogeneity and higher clay content, associated with a meandering river system. Porosity within these reservoirs ranges from 15% to 20%, with permeability commonly reaching up to 1 Darcy (Beicip, 2002; WEC, 2007).

2.4.4.2. The infra-TAGS reservoirs

Display more heterogeneous properties and reduced quality, primarily owing to fluvial channel networks. They include the Intermediate Triassic and TAGI reservoirs, featuring porosities typically between 5% and 15%, and permeabilities under 10 mD. These units overlie the Hercynian unconformity in a transgressive manner and can be absent locally due to non-deposition, such as in the Rhourde Adra structure. Additionally, the Hercynian unconformity facilitates hydraulic connectivity between these infra-TAGS Triassic reservoirs and the Silurian units (B2, B1, A2, A1, M2, M1). Optimal reservoir quality is found in the A unit, formed by marine platform sandstone bars, with net pay up to 50 meters thick and porosity from 10% to 15%.

2.4.4.3. The deep reservoir of the Cambro-Ordovician Hamra Quartzites

The Quartzites El Hamra formation (QH) is one of the unconventional tight fractured reservoirs. Composed of mature quartz arenates typical of those deposited in a shallow marine

environment (subtidal sandflat), and is situated between two regional marine shale formations: the Gassi shales (lower Ordovician) and the Azzel shales (middle Ordovician), intercalates with the El Atchane sandstone and the Ouargla sandstone.

This formation is divided into four units (QH1, QH2, QH3 and QH4) from top to base according to the variation in the inflection of gamma rays (GR) and the variation of resistivity.

- **QH-1:** A clear basal marker corresponding to the lowest gamma-ray value at the base of the Ouargla Sandstone, marking the true top of the Hamra Quartzites.
- **QH-2:** A major marker of the Rhourde Nouss complex, identified in all drilled wells.
- **QH-3 and QH-4:** Well identified locally, although some uncertainties remain in certain wells where the reservoir was not fully drilled.

This reservoir is economically promising according to its significant thickness (200 to 250 meters), an extension of approximately 10,000 km², and its intense natural fracturing resulting from the global tectonic history of the Saharan platform (Boudjema, 1987). Despite its thickness and horizontal continuity, this QH formation yields very poor matrix petrophysical characteristics (average porosity of 3.5 % and matrix permeability of 0.1 mD), which only the presence of fractures improves productivity. (Beicipe, 2002).

The depositional and stratigraphic evolution of the Hamra Quartzite formation within the Saharan platform is illustrated in figure 10, which shows a schematic sequence stratigraphic framework of the Mid-Ordovician interval, highlighting the lateral facies variations and their relationship with sea-level fluctuation (after Galeazzi et al., 2010)

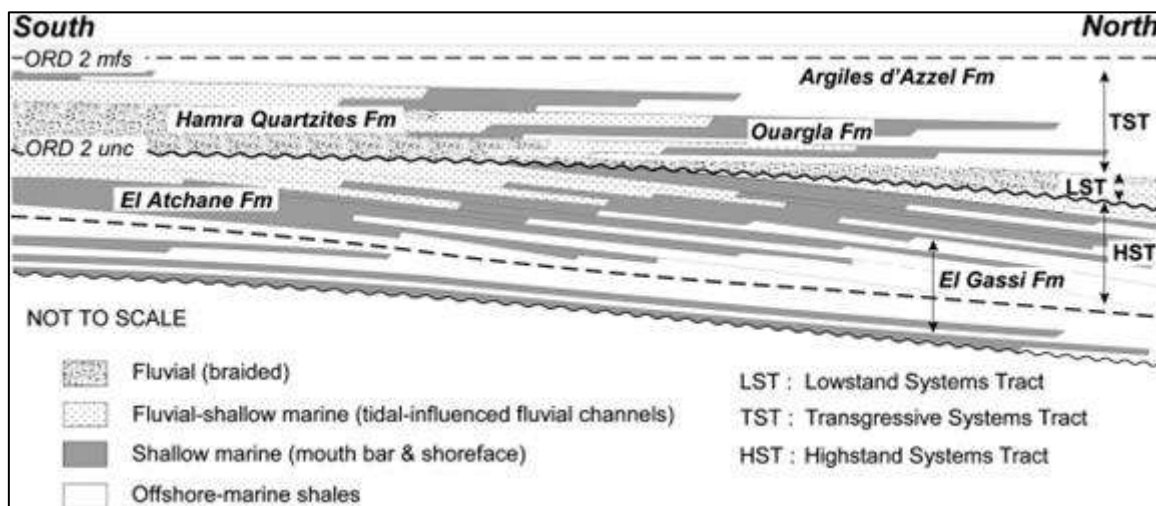


Figure 11. Sequence stratigraphic framework and depositional architecture of the Mid-Ordovician Hamra Quartzites formation within the Saharan platform (Galeazzi et al., 2010).

Conclusion

The study area, the Rhourde Nouss field, is located within the Berkine Basin on the Saharan Platform, south of the Atlas Orogenic Belt. This basin represents a major intracratonic depression that developed from the Paleozoic to the Early Mesozoic and is characterized by a thick sedimentary succession overlying a Precambrian basement formed during the Pan-African orogeny. Its geological evolution reflects a complex tectonic history marked by successive Pan-African, Hercynian, and Alpine phases, which controlled both the structural framework and sedimentary architecture.

The Rhourde Nouss field constitutes one of the most significant hydrocarbon provinces in Algeria. Stratigraphically, the region comprises three main units: a predominantly siliciclastic Paleozoic sequence, a Mesozoic succession of sandstones, evaporites, and carbonates, and a Cenozoic cover.

The petroleum system primarily derives from organic-rich Silurian shales, while Liassic evaporites provide an effective seal for hydrocarbon preservation. Structural traps, mainly formed during the Austrian tectonic phase, have played a key role in hydrocarbon accumulation. The main reservoir systems include the Triassic TAGS formation, the infra-TAGS units, and the deeply buried Cambro-Ordovician Hamra Quartzites (HQ), which represent the main target of this study.

This formation QH is characterized by significant thickness and horizontal continuity, but exhibits very low matrix porosity and permeability (average porosity of 3.5 % and matrix permeability of 0.1 mD), making reservoir productivity largely dependent on natural fractures. These fractures, developed throughout the basin's complex tectonic history, play a critical role in enhancing fluid flow and hydrocarbon recovery, rendering the QH a strategically important target for exploration and production.

Chapter III Artificial Intelligence and Machine Learning Techniques

Introduction

Artificial intelligence (AI) has emerged as one of the fundamental pillars of science and engineering, due to its ability to address intricate systems with multiple variables, uncertainty, and challenges in exact mathematical modeling (Li and Du, 2017; Xu et al., 1021). In contrast to conventional methods dependent on simplified physical or statistical models, AI offers a versatile computational approach that extracts insights directly from data, bypassing the need for assumed variable interrelationships (Bishop, 2006; Norvig, 2016). This strength positions AI as a powerful solution for challenges where traditional techniques prove inadequate or imprecise (Marr, 2019).

AI's advancements stem from algorithms incorporating machine learning principles, hybrid intelligent frameworks, and sophisticated optimization methods like evolutionary and genetic algorithms (Hastie et al., 2009; Lu, 2019; Kaur et al., 2025). These have proven highly capable in processing complex datasets, revealing concealed patterns, and aiding predictions and decisions amid noisy data and evolving system dynamics (Haykin, 2009;).

Fueled by surging computational capabilities and abundant high-quality datasets, machine learning (ML) and deep learning (DL) have surged in engineering contexts. They excel at capturing intricate variable interactions and high-dimensional nonlinear dynamics, ideal for systems resistant to deterministic formulations (LeCun et al., 2015; Goodfellow et al., 2016). Additionally, these methods facilitate robust, generalizable predictive models derived purely from data.

3.1. Intelligence Artificial (AI)

3.1.1. Definition and conception of intelligence artificial

Artificial intelligence (AI) serves as the overarching conceptual framework for computational methods designed to enable machines to demonstrate intelligent behavior. The term was coined by McCarthy (1956), who defined it as the science and engineering of creating intelligent machines. Machine learning (ML), a key subfield within AI, was formalized by Samuel (1959) as computers' capacity to learn from experience without explicit programming, emphasizing algorithms that detect patterns and relationships directly from data.

Deep learning (DL), an advanced extension of ML, traces its origins to early neural network studies, including Rosenblatt's perceptron (Rosenblatt, 1958), and evolved through

multi-layer networks and backpropagation (Rumelhart et al., 1986). DL employs deep artificial neural networks to extract hierarchical feature representations from vast, complex datasets (Goodfellow et al., 2016). This progression from AI to ML to DL illustrates the historical and methodological development of intelligent systems, forming the basis for contemporary data driven technologies (LeCun et al., 2015).

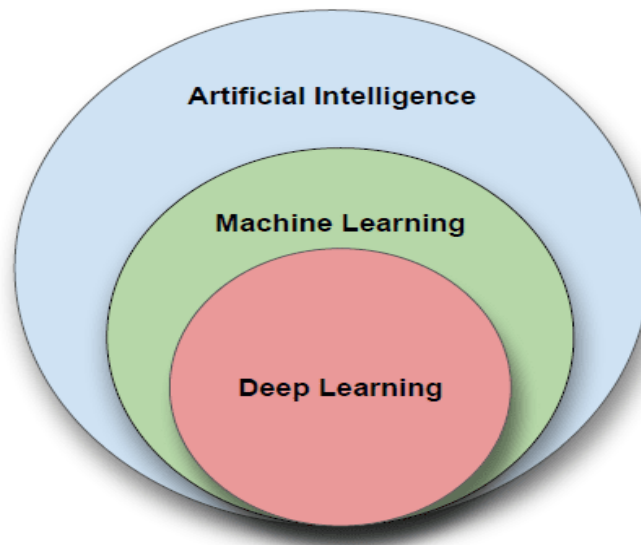


Figure 12. Hierarchical Relationship Between Artificial Intelligence, Machine Learning, and Deep Learning. (source: Webography)

3.2. Machine Learning (ML)

3.2.1. Definition and principals of machine learning

Machine Learning is a subfield of artificial intelligence aimed at developing algorithms that automatically acquire knowledge from data, extract key patterns, and generalize them to untrained new data for prediction or decision-making without the need for detailed programming of each step (Dramsche, 2020; Samuel, 2000). The origin of the term traces back to Arthur Samuel in 1959, who described the goal of machine learning as enabling the computer to learn and improve automatically through experience (Samuel, 2000).

From a theoretical perspective, machine learning is classified within Statistical Learning Theory, which provides a mathematical foundation for explaining how models learn functions that link inputs to outputs while effectively generalizing to unfamiliar data (Vapnik, 2013). Machine learning relies on creating an adjustable statistical model or function, which is optimized through a training process focused on minimizing the loss function (Loss Function) using optimization techniques such as gradient descent (Gradient Descent) (Goodfellow et al., 2016).

The core foundation of machine learning lies in its ability to generalize lessons learned from training data to new scenarios. During training, the model's parameters are adjusted to minimize the loss function, which measures prediction deviations, while maintaining a balance between the model's complexity and flexibility to avoid overfitting or underfitting since excessive simplification leads to poor performance on both training and test data (Hastie et al., 2005). Additionally, machine learning models excel at capturing complex nonlinear interactions between variables, making them ideal for diverse scientific and engineering applications (Bishop, 2006). Accurate performance evaluation through validation and testing stages remains essential before deploying the model in real-world applications, to ensure its reliability and effectiveness (Goodfellow et al., 2016).

3.2.2. Types of machine learning

Machine learning is classified into multiple primary categories based on the nature of available data, learning approaches, and interaction mechanisms with the environment. However, the most prevalent classification in scientific literature divides it into three fundamental types: supervised learning, unsupervised learning, and reinforcement learning (Bishop, 2006; Hastie et al., 2005; Goodfellow et al., 2016).

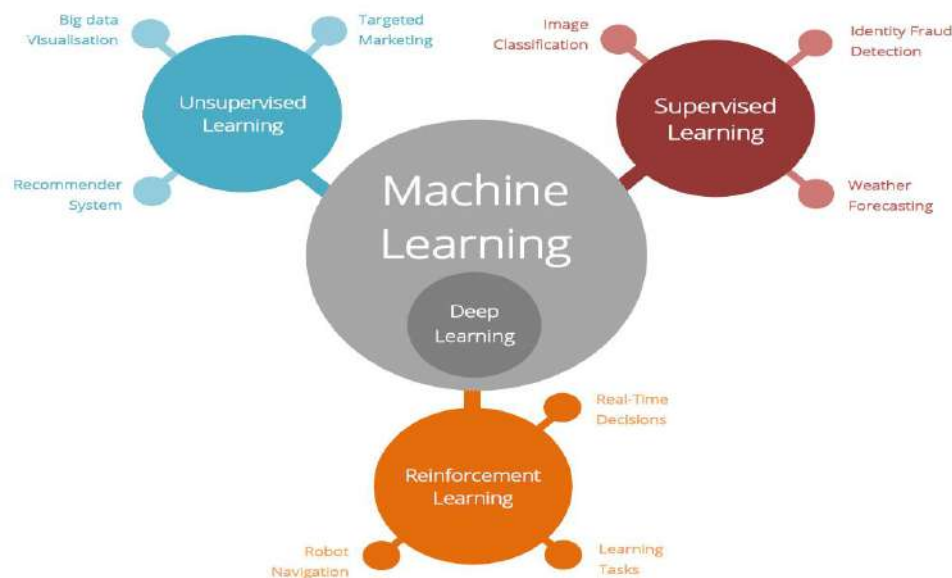


Figure 13. Main Paradigms of Machine Learning and Their Typical Applications (source: Webography)

3.2.2.1. Supervised learning

Supervised learning represents the most widely adopted and applied type across scientific and engineering domains. This approach trains models using a labeled dataset comprising input-output pairs X and Y . The primary objective is to derive an approximating

function that captures the relationship between inputs and outputs, enabling predictions of unknown values for new data (Hastie et al., 2005).

Supervised learning problems are generally classified into two main categories:

- Regression problems, where outputs are continuous, such as predicting petrophysical properties or mechanical parameters.
- Classification problems, where outputs are categorical or discrete, such as identifying rock types or reservoir conditions (Bishop, 2006).

Its algorithms minimize a loss function that quantifies discrepancies between predictions and actual values through optimization techniques like gradient descent. Prominent examples include linear and logistic regression, support vector machines (SVM), decision trees, random forests (RF), and deep learning algorithms (Goodfellow et al., 2016).

Supervised learning finds extensive application in petroleum engineering, geophysics, and well log data analysis due to the availability of measured and labeled datasets (Dramsch, 2020).

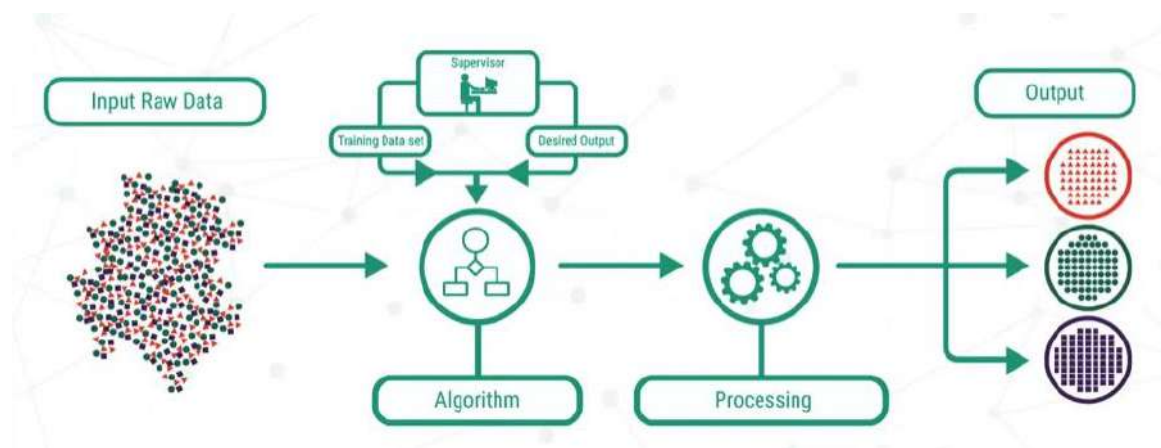


Figure 14. Workflow of Supervised Machine Learning: From Raw Data to Output Predictions (source: Webography)

3.2.2.2. Unsupervised learning

Unlike supervised learning, unsupervised learning processes unlabeled data without predefined output values. Its core focus lies in revealing internal structures, hidden relationships, and latent patterns within the data without direct supervision (Bishop, 2006). Key applications encompass:

- Clustering to partition data into homogeneous groups.
- Dimensionality reduction to simplify high-dimensional data while preserving maximum information content (Hastie et al., 2005).

Prominent algorithms include k-means, hierarchical clustering, DBSCAN, and principal component analysis (PCA). It serves as a critical tool for exploratory data analysis, particularly

in early stages of geological and reservoir studies where relationships remain unknown (Jain, 2010).

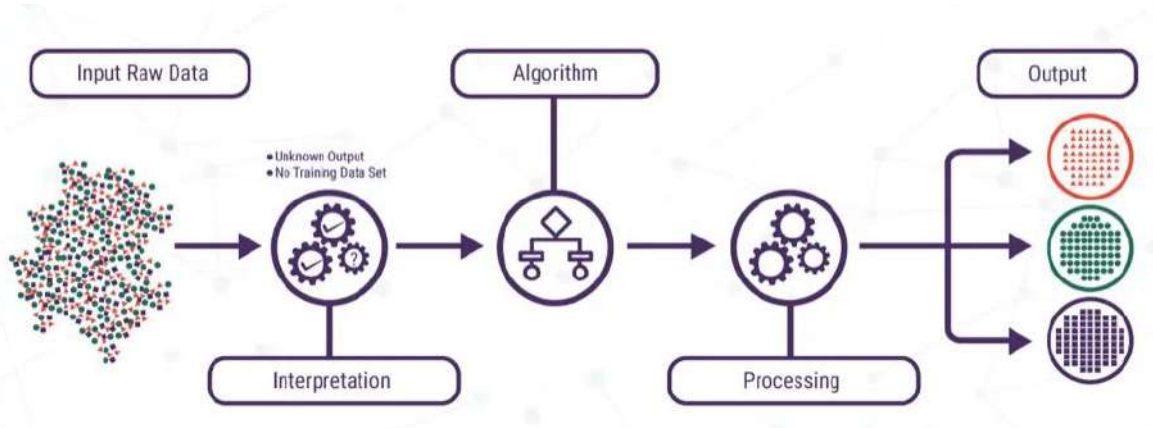


Figure 15. Workflow of Unsupervised Machine Learning: From Raw Data to Output Predictions (source: Webography)

3.2.2.3. Reinforcement learning

Reinforcement learning provides a distinct framework wherein an intelligent agent interacts with a specified environment through actions, receiving rewards or penalties, to maximize long-term cumulative returns (Sutton and Barto, 2018).

Rather than relying on labeled data, it depends on trial-and-error and suits dynamic, complex systems challenging to model analytically. Notable applications include robotics, gaming, optimal control, and industrial system management (Kaelbling et al., 1996; Ding et al., 2020). Recent years have witnessed substantial advances through its integration with deep neural networks, termed deep reinforcement learning, enabling effective solutions for highly complex problems (Goodfellow et al., 2016).

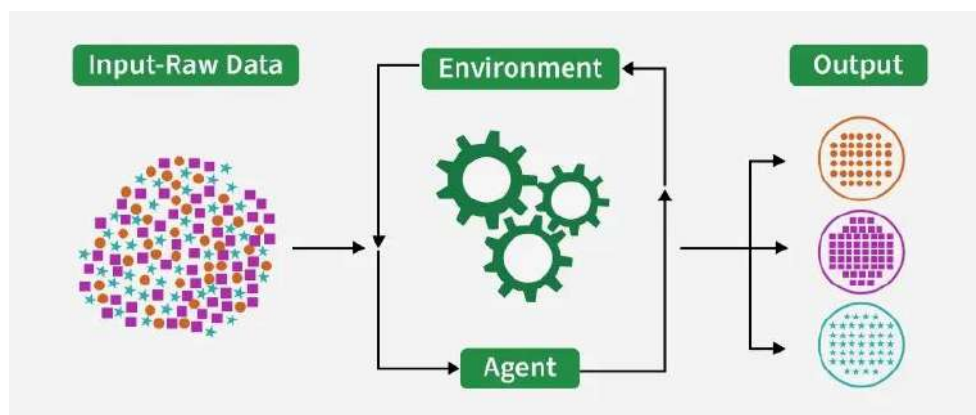


Figure 16. Workflow of Reinforcement Learning: From Raw Data to Output Predictions (source: Webography)

3.2.2.4. Deep Learning

Deep learning (DL) constitutes an advanced branch of artificial intelligence within the machine learning framework, relying on multi layered deep artificial neural networks to model complex, nonlinear relationships between inputs and outputs through entirely data driven approaches (LeCun et al., 2015; Goodfellow et al., 2016). This paradigm is grounded in hierarchical representation learning, wherein each layer of the neural network captures distinct levels of abstraction. Shallow layers process low level features, while deeper layers extract increasingly abstract and complex patterns, enabling efficient representation of high dimensional data structures (Bengio et al., 2013).

Deep learning models distinguish themselves through automatic feature extraction from raw data, obviating the need for intensive manual feature engineering a key differentiator from traditional machine learning algorithms. This approach minimizes reliance on domain expertise while enhancing model generalization and adaptability to complex, unstructured datasets (Schmidhuber, 2015).

Training deep learning models employs backpropagation integrated with numerical optimization algorithms such as Stochastic Gradient Descent, Adam, and RMSProp, iteratively adjusting network weights to minimize the loss function and improve prediction accuracy (Rumelhart et al., 1986; Kingma & Ba, 2015).

Given their high parameter counts and susceptibility to overfitting, advanced regularization techniques including Dropout, Batch Normalization, Early Stopping, and Weight Regularization are essential to ensure computational stability and balance model complexity with generalization capability (Srivastava et al., 2014; Ioffe and Szegedy, 2015). This theoretical and technical framework has enabled deep learning to achieve remarkable success in handling high-dimensional, noisy, and nonlinear data, establishing it as a cornerstone tool across computer vision, natural language processing, signal analysis, and geosciences including petroleum engineering applications and well log analysis (Haykin, 2009; Dramschi, 2020).

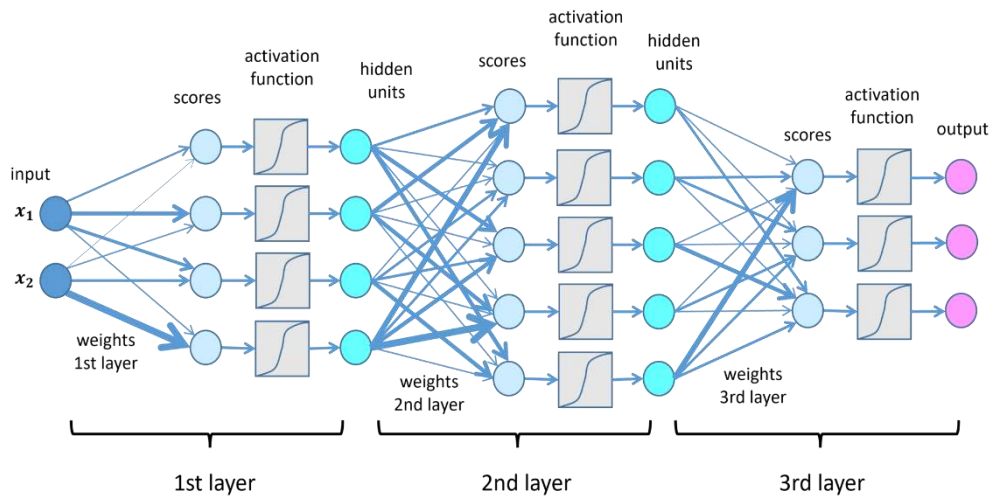


Figure 17. Deep Learning: Example Illustrating Deep Neural Network Functioning (source: Webography)

3.3. Employed Algorithms for Predictive Modeling

3.3.1. Random Forest

3.3.1.1. Random Forest: Principles and Methodology

The Random Forest (RF) algorithm represents a supervised ensemble learning technique within the tree-based methods family. Developed by [Breiman \(2001\)](#), it addresses the overfitting tendency of standalone decision trees by integrating numerous independent trees into a unified predictive framework, thereby improving generalization performance ([Costa and Pedreira, 2023](#)).

At its foundation, Random Forest constructs a multitude of decision trees, each trained on bootstrap samples extracted from the original dataset a process termed Bootstrap Aggregating (Bagging) ([Altman and Krzywinski, 2017](#)). This methodology effectively lowers variance while maintaining the capacity to capture intricate nonlinear data relationships ([Breiman, 2001](#); [Liaw and Wiener, 2002](#)).

Individual trees emerge through recursive partitioning of the feature space, where each internal node selects a splitting rule from a randomized subset of predictors. This randomization fosters tree diversity, constituting a cornerstone of the Random Forest methodology ([Breiman, 2001](#)).

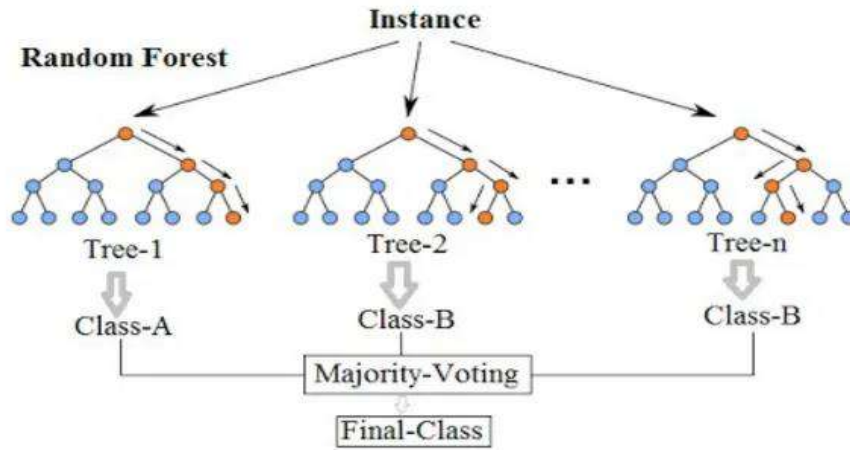


Figure 18. Random forest workflow: From Raw Data to Output Predictions (source: Webography)

3.3.1.2. Splitting Criteria

a. Classification Problems

In classification scenarios, entropy rooted in Shannon's information theory serves as a primary criterion:

$$E = - \sum_{i=1}^c p_i \log(p_i) \quad (1)$$

Where: c signifies class count and p_i denotes class i 's prior probability.

Optimal splits maximize information gain, measured as entropy reduction from parent to child nodes (Sarica et al., 2017; Schonlau and Zou, 2020).

b. Regression Problems

In regression, Mean Squared Error (MSE) prevails:

$$\text{MSE} = \frac{1}{n} \sum_{j=1}^n (y_j - \hat{y}_j)^2 \quad (2)$$

Where: y_j and $\hat{y}_j$ represent observed and predicted values, respectively. The best split minimizes MSE at each node (Liaw and Wiener, 2002; Jaiswal and Samikannu, 2017).

3.3.2. Prediction Aggregation

Final predictions aggregate individual tree outputs: predictions are obtained through majority voting for classification and averaging for regression:

$$\hat{y} = \text{mode}(\hat{y}^{(1)}, \dots, \hat{y}^{(T)}), \hat{y} = \frac{1}{T} \sum_{t=1}^T \hat{y}^{(t)} \quad (3)$$

Where: T is the total number of trees in the forest (Breiman, 2001; Salman et al., 2024).

3.3.3. Gradient Boosting

3.3.3.1. Gradient Boosting: Principles and Methodology

Gradient Boosting (GB) represents a supervised ensemble learning method distinct from parallel tree construction in Random Forests. Instead, it builds decision trees sequentially, with each subsequent tree designed to compensate for the residual errors of its predecessors (Natekin and Knoll, 2013; Bahad and Saxena, 2019; Bentéjac et al., 2021). This iterative refinement prioritizes challenging cases, yielding superior predictive performance (Aziz et al., 2020; Sahin, 2022).

Gradient Boosting optimizes a loss function through stage-wise tree addition. At iteration m , the predictive function updates as:

$$F_m = F_{m-1}(x) + \gamma_m h_m(x) \quad (4)$$

Where: $F_{m-1}(x)$ denotes the prior stage's prediction, $h_m(x)$ constitutes the new weak learner (typically a decision tree), and γ_m represents the learning rate optimized to minimize overall loss (Natekin and Knoll, 2013).

Loss functions adapt to task requirements: mean squared error for regression, logistic loss for classification, or concordance index for survival modeling (Hosen and Amin, 2021).

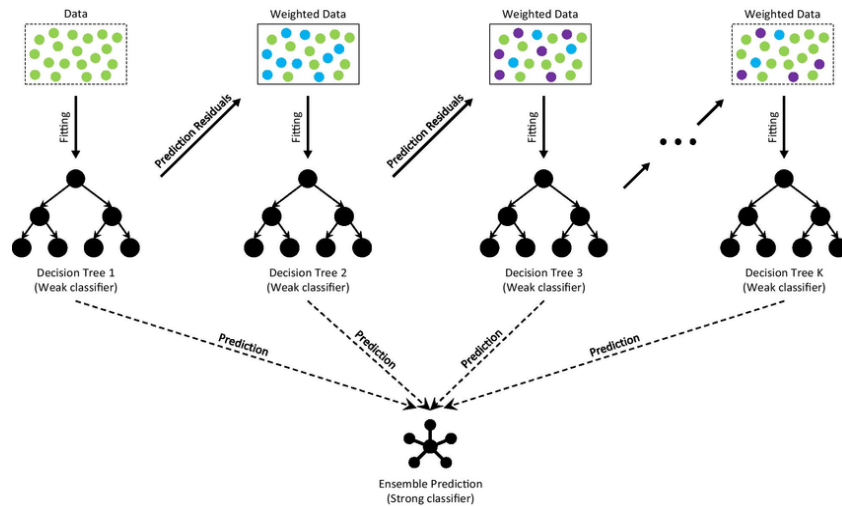


Figure 19. Structural Design of Gradient Boosting Decision Trees (Deng et al., 2021).

3.3.3.2. Algorithmic Procedure

1- Initialize the model with a constant value, e.g., the mean of the target variable:

$$F_0(x) = \arg \min_{\gamma} \sum_{i=1}^n L(y_i, \gamma) \quad (5)$$

2- For iterations $m = 1$ to M :

- Calculate negative gradients (residuals):

$$r_{im} = - \left[\frac{\partial L(y_i, F(x_i))}{\partial F(x_i)} \right]_{f=f_{m-1}} \quad (6)$$

- Train weak learner $h_m(x)$ on residuals r_{im}
- Determine optimal step size:

$$\gamma_m = \arg \min_{\gamma} \sum_{i=1}^n L(y_i, F_{m-1}(x_i) + \gamma h_m(x_i)) \quad (7)$$

- Update model:

$$F_m(x) = F_{m-1}(x) + \gamma_m h_m(x) \quad (8)$$

The final model yields $F_M(x)$ as the prediction ([Bentéjac et al., 2021](#); [Natekin and Knoll, 2013](#)).

3.3.4. Extreme gradient boosting

3.3.4.1. Extreme gradient boosting: principles and methodology

The Extreme Gradient Boosting (XGBoost) algorithm constitutes a sophisticated evolution of the Gradient Boosted Decision Tree (GBDT) framework, engineered to simultaneously elevate predictive precision and computational performance ([Liang et al., 2020](#); [Chi, 2024](#)). Classified among ensemble learning methodologies, XGBoost integrates multiple weak learners to forge a robust predictive system ([Naghbi et al., 2020](#)). Specifically, it employs shallow regression trees trained sequentially, with each successive tree calibrated to rectify residual errors persisting from the antecedent ensemble ([Friedman, 2002](#); [Chen et al., 2015](#)).

3.3.4.2. Additive Boosting Strategy

XGBoost's learning paradigm adheres to an additive boosting methodology, wherein each iteration introduces a new tree designed to approximate the negative gradient of the loss function relative to prevailing model predictions (Zhang et al., 2023). This gradient directed optimization strategically emphasizes challenging samples, systematically diminishing cumulative error rates (Natekin and Knoll, 2013). Distinguished from conventional gradient boosting, XGBoost incorporates second-order Taylor expansion of the loss function, facilitating superior optimization accuracy and accelerated convergence (Chen et al., 2015).

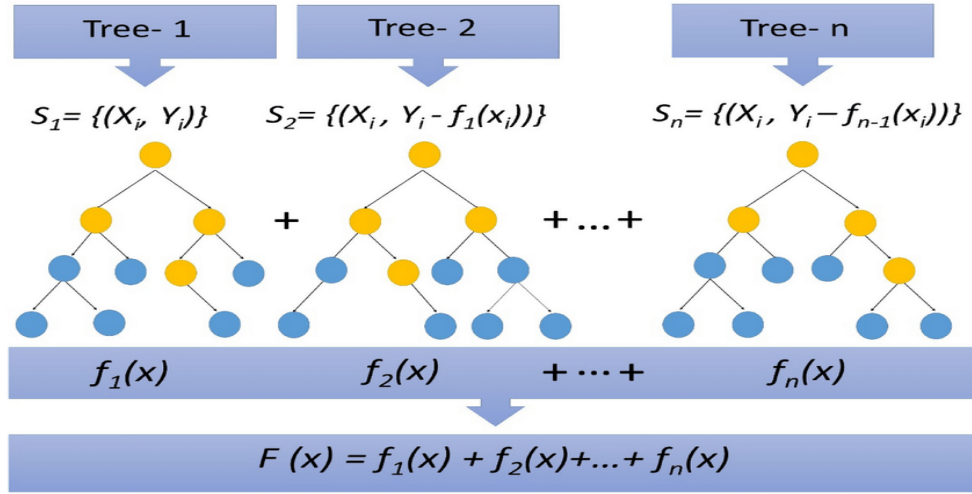


Figure 20. Structural Design of Extreme Gradient Boosting (Sahour et al., 2021).

3.3.4.3. Regression Tree Ensemble and Prediction Formulation

Predictions emerge from an ensemble of regression trees, wherein each constituent tree $f_k(x)$ contributes additively to the ultimate output. The comprehensive model prediction manifests as:

$$\hat{y} = \sum_{k=1}^K f_k(x_i) \quad (9)$$

Where $\hat{y}$ is the predicted output, K is the total number of trees, and f_k embodies the k^{th} regression tree function (Qiu et al., 2022; Xu et al., 2023).

This cumulative architecture empowers XGBoost to adeptly discern intricate nonlinear relationships and higher-order feature interactions.

3.3.5. Hierarchical Tree Structure

Individual regression trees within XGBoost comprise internal decision nodes coupled with terminal leaf nodes. Internal nodes execute data partitioning via feature threshold splits,

while leaf nodes house continuous output scores (Zhang et al., 2021). This stratified tree topology progressively subdivides the feature space into increasingly homogeneous regions, conferring exceptional flexibility for modeling complex patterns and variable interdependencies (Chen and Guestrin, 2016).

3.3.5.1. Regularized Objective Function and Overfitting Prevention

XGBoost's hallmark feature resides in its explicit regularization framework, which governs model complexity and counters overfitting tendencies (Khodabakhshi and Bijani, 2024). The optimized objective function amalgamates two principal elements: a differentiable loss function quantifying disparities between observed and predicted values, alongside a regularization term penalizing excessive tree complexity (Asselman et al., 2023). During iteration g , the objective formulates as:

$$K^{(g)} = \sum_{i=0}^n m(y_i + \hat{y}_i) + \sum_{n=1}^p \alpha(t_p) \quad (10)$$

Where: m the loss function, (y_i) the actual value and $(\hat{y}_i)$ the predicted or estimated value by the model. (α) Regularization function, (t_p) the function of tree P.

3.3.5.2. Regularization Mechanics and Generalization Enhancement

The regularization component conventionally penalizes both leaf node proliferation and leaf weight magnitudes, thereby attenuating variance and bolstering generalization efficacy (Ali et al., 2023). This deliberate regularization paradigm differentiates XGBoost from traditional gradient boosting approaches, substantially augmenting its resilience and predictive superiority particularly within noisy or high-dimensional data environments (Chen et al., 2015; Ali et al., 2023).

3.3.6. Categorical Boosting Algorithm

3.3.6.1. Categorical Boosting: principles and methodology

Categorical Boosting (CatBoost) constitutes an ensemble learning methodology grounded in Gradient Boosted Decision Trees (GBDT), engineered for seamless handling of categorical features while curbing overfitting via ordered boosting (Chen et al., 2024). This technique employs random permutation of the dataset, followed by model training that restricts each sample's exposure exclusively to preceding samples in the sequence excluding the sample itself and all subsequent instances (Prokhorenkova et al., 2018; Hancock and Khoshgoftaar,

2020). Consequently, predictions for individual data points emerge without exposure to their actual target values during training.

3.3.6.2. Additive Model Formulation

Consistent with other boosting algorithms, CatBoost assembles the ultimate predictive model through an additive combination of decision trees: Eq (9), this cumulative structure empowers CatBoost to iteratively rectify prediction discrepancies while adeptly modeling nonlinear dependencies and feature interactions (Zhong et al., 2021; Ding et al., 2025).

3.3.6.3. Ordered Boosting and Regularization

CatBoost refines a regularized objective function by sequentially adapting each tree to residuals from the prior ensemble, while maintaining unbiased gradient estimates through systematic permutations of training data. This ordered boosting paradigm substantially alleviates overfitting risks and bolsters model generalization performance (Chen et al., 2024; Cai et al., 2024).

$$\mathcal{J}_{CB} = \sum_{i=1}^n \mathcal{L}(y_i, \hat{y}_i) + \sum_{k=1}^K \Omega(f_k) \quad (11)$$

Where: y_i is the true target value, $\hat{y}_i$ the predicted output, $\mathcal{L}(\cdot)$ a differentiable loss function, f_k represents the k -th oblivious decision tree and $\Omega(f_k)$ the regularization term.

3.3.7. Adaptive Boosting algorithm

3.3.7.1. Adaptive Boosting: principles and methodology

Adaptive Boosting (AdaBoost) algorithm ranks among the foundational algorithms in the field of ensemble learning based on boosting, aiming to construct a robust predictive model through the sequential integration of a series of weak learners, which are typically simple decision trees (Ying et al., 2013; Schapire, 2013). The algorithm relies on the principle of adaptive learning, whereby training sample weights are redistributed in each iteration to assign greater weights to samples that were misclassified by previous learners, compelling the subsequent model to concentrate on the most challenging cases (Wang, 2012).

3.3.7.2. Objective Function in AdaBoost

The adaptive behavior translates mathematically through the objective function, which AdaBoost seeks to minimize across all iterations (Zhang and Yang, 2016). In classification problems, the algorithm focuses on reducing exponential loss, which severely penalizes prediction errors, thereby amplifying the influence of difficult samples in subsequent learning

processes (Schapire, 2013). The final model comprises a weighted additive combination of weak learners, where their weights reflect their capacity to reduce the overall loss function (Hu et al., 2008).

$$L = i = 1 \sum_{i=1}^n \exp(-y_i F(x_i)) \quad (12)$$

Where: n is the number of training samples, x_i the i -th input sample, $y_i \in \{-1, +1\}$: true class label.

Where the additive model is defined as:

$$F(x) = \sum_{m=1}^M \alpha_m h_m(x) \quad (13)$$

Where: h_m : the m -th weak learner, α_m : weight assigned to the weak learner, M : number of boosting iterations

This framework, predicated on sequential optimization of the objective function, achieves effective balance between minimizing predictive error and enhancing the model's generalization capability a factor explaining AdaBoost's superiority across diverse classification and regression applications (Solomatine and Shrestha, 2004; Ding et al., 2022).

3.3.8. Light Gradient Boosting

3.3.8.1. Light Gradient Boosting: principles and methodology

The Light Gradient Boosting Machine (LightGBM) represents one of the leading machine learning models based on gradient boosting for decision trees, distinguished by its training speed, computational efficiency, and capacity to handle large, high dimensional datasets (Ju et al., 2019; Liu et al., 2023). LightGBM employs a leaf-wise tree growth strategy (best-first), where the leaf yielding the greatest loss reduction is split first, enabling the model to represent complex, nonlinear relationships more efficiently than traditional approaches (Tang et al., 2020; Gan et al., 2021).

3.3.8.2. Objective Function optimization

LightGBM optimizes the objective function through gradient boosting techniques, with each subsequent tree correcting residual errors from the previous model (Al-Kasassbeh et al., 2020; Wang et al., 2023). The general objective function formulates as:

$$L(\theta) = i = 1 \sum_{i=1}^n \ell(y_i, \hat{y}_i) + \sum_{i=1}^n \Omega(fk) \quad (14)$$

Where: y_i : true sample values, $\hat{y}_i$: current model predictions, $\ell(y_i, \hat{y}_i)$: loss function, $\Omega(f_k)$: regularization for tree k , K : number of trees in the model, θ : all model parameters.

3.3.8.3. Robustness and Generalization Capability

Light Gradient Boosting machine excels at processing diverse, multidimensional, noisy data while preserving prediction accuracy and performance stability (Zheng et al., 2025), where generalization improves through hyperparameter tuning (learning rate, tree count, max depth) (Zhang et al., 2021), reducing overfitting risk and establishing LightGBM as a robust, reliable algorithm for complex scientific and engineering applications (Van, 2023).

Figure. 21 schematic comparison of tree structures in ensemble learning algorithms: AdaBoost, XGBoost, LightGBM, and CatBoost. This diagram illustrates the sequential structure and tree types used in each ensemble algorithm, highlighting differences in growth strategy and symmetry for clear comparison.

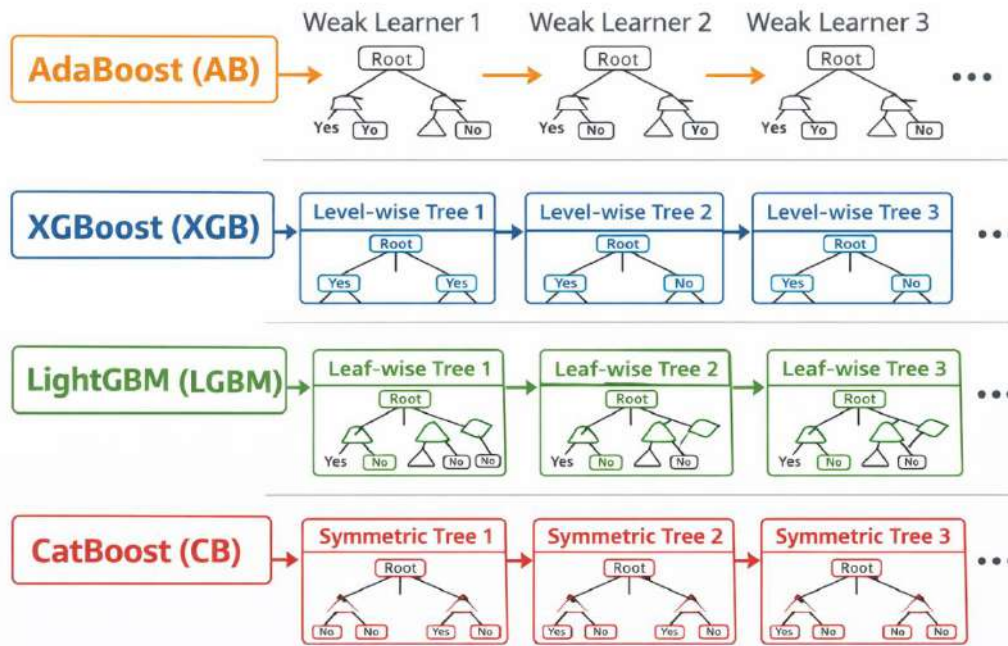


Figure 21. Schematic comparison of tree structures in ensemble learning algorithms: AdaBoost, XGBoost, LightGBM, and CatBoost

3.4. Model interpretability and explainability tools

In contemporary machine learning applications particularly within scientific and engineering domains comprehending the decision making mechanisms of models has become imperative. Model interpretability and comprehension are fundamental for ensuring transparency, trust, and accountability in predictive analyses (Gilpin et al., 2018; Hassija et al.,

2024). These concepts enable researchers and practitioners to assess prediction reliability and identify key influencing factors. Without interpretation, complex models such as ensemble methods or gradient boosting algorithms function as black boxes, limiting their utility in safety-critical or sensitive applications (Carvalho et al., 2019; Linardatos et al., 2020).

Numerous tools and techniques have been developed to enhance model interpretability. Among the most widely adopted are Feature Importance and SHapley Additive explanations (SHAP), which provide clear insights into how input variables impact model predictions (Hassija et al., 2024). Utilizing these tools bridges the gap between model complexity and practical result comprehension, facilitating informed decision making and increasing confidence in machine learning outcomes (Carvalho et al., 2019; ElShawi et al., 2021).

3.4.1. Feature importance

Feature importance measures the contribution of each input variable to a model's predictions, helping interpret complex machine learning algorithms and enhance transparency (Musolf et al., 2022). It identifies which features most influence outputs, guiding decision-making and improving reliability. Common methods include permutation importance and tree-based metrics, which evaluate how individual features affect model performance (Huynh-Thu et al., 2012).

Advanced techniques, such as fuzzy information fusion, integrate multiple importance measures for robust assessments in high-dimensional or complex datasets (Rengasamy et al., 2022). Feature importance also supports feature selection, reduces model complexity, and prevents overfitting (Kaneko, 2023). By highlighting the most influential predictors, it provides insights into underlying patterns, validates hypotheses, and ensures actionable, trustworthy results in applied domains.

3.4.2. SHapley Additive explanations

SHAP is an advanced tool for interpreting complex machine learning models, allowing the analysis of how each feature contributes to the model's final prediction (Ekanayake et al., 2022). The core idea is to assign each feature a contribution value that reflects its relative impact on the output, improving the transparency of black-box models (Nohara et al., 2019). SHAP works by comparing the model's output with and without the feature, evaluating all possible feature combinations to precisely measure the effect of each feature (Mangalathu et al., 2020).

This approach allows understanding feature interactions and their influence on predictions, enabling a comprehensive interpretation of models (Gebreyesus et al., 2023). SHAP can identify the most influential features, analyze unexpected results, and improve

decision making based on the models. The technique can also handle incomplete or noisy data, increasing prediction reliability. SHAP is compatible with various models, including ensemble methods, random forests, and gradient boosting algorithms, making it a flexible tool for evaluating model performance and enhancing understanding of how models operate.

3.5. Application of intelligence artificial in petroleum engineering

Artificial intelligence (AI) has found extensive application in petroleum engineering over recent decades, benefiting from more than seventy years of machine learning (ML) development in Earth sciences modeling ([Dramschi, 2020](#)). Numerous studies have demonstrated the capability of ML techniques to predict reservoir and drilling-related properties with high accuracy. Consequently, AI-based approaches have been widely adopted for petroleum reservoir characterization and modeling, production optimization and monitoring, as well as the prediction of key petrophysical properties.

For instance, [Zerrouki et al. \(2014\)](#) employed well log data to estimate natural fracture porosity using fuzzy ranking theory combined with artificial neural networks (ANN). [Ameur-Zaimeche et al. \(2022\)](#) developed real-time porosity prediction models based on drilling data by applying several ML algorithms, including extreme learning machine (ELM), support vector regression (SVR), random forest regression (RFR), multilayer perceptron neural networks (MLPNN), and generalized regression neural networks (GRNN). Similarly, [Zhao et al. \(2020\)](#) utilized the random forest algorithm to predict relative permeability.

AI techniques have also been extensively applied to pressure-related evaluations. Pore pressure estimation has been successfully performed using artificial neural networks ([Abdelaal et al., 2022](#); [Kianoush et al., 2023](#)), while the formation pressure gradient has been predicted in real time using ANN and adaptive neuro-fuzzy inference systems (ANFIS). In the context of drilling fluids, [Alsabaa et al. \(2021\)](#) proposed new ANN-based correlations to improve the monitoring of all-oil mud rheological properties.

Moreover, AI methodologies have been increasingly used to optimize drilling operations. These applications include the prediction and optimization of the rate of penetration (ROP) using ANN, random forest, and ANFIS models ([Elkatatny et al., 2017](#); [Alsaihati et al., 2022](#)), as well as the evaluation and prediction of equivalent circulation density using various AI techniques ([Abdelgawad et al., 2019](#); [Abdelaal et al., 2023](#)).

3.6. Application of Artificial Intelligence in Geomechanical Property Estimation

Recent studies have demonstrated that artificial intelligence (AI) represents an effective tool for characterizing the geomechanical properties of reservoirs, such as dynamic Poisson's

ratio, Young's modulus, compressive strength, and in-situ stresses, which are critical factors for evaluating wellbore stability and rock behavior during drilling and production (Zoback, 2010).

For instance, Siddig et al. (2021) showed that using artificial neural networks (ANN) combined with adaptive neuro-fuzzy inference systems (ANFIS) can accurately predict Poisson's ratio in carbonate, sandstone, and shale reservoirs, achieving high correlation coefficients up to $R = 0.99$. Ahmed et al. (2021) confirmed that integrating ANN, ANFIS, and SVM allows precise Poisson's ratio predictions in geologically complex reservoirs, reflecting the capability of machine learning algorithms to handle geological variability.

Regarding dynamic Young's modulus estimation, studies such as Davoodi et al. (2024), Wei et al. (2024), and Yue et al. (2023) demonstrated that using mud-logging or LWD data with ANN, RF, and SVR enables high-accuracy predictions of Young's modulus while reducing reliance on intensive laboratory measurements. Additionally, Siddig et al. (2023) and Mahmoud et al. (2022) showed that combining algorithms like SVR, RF, and ANFIS enhances real-time predictive capabilities, enabling faster and more reliable engineering decisions.

For pore pressure and in-situ stress estimation, studies by Yao et al. (2023) and Bruna Teixeira Silveira et al. (2023) employed advanced neural networks (BPNN, LSTM, NN) to deliver accurate real-time predictions during drilling, allowing continuous monitoring and improved operational performance. Meanwhile, Ibrahim et al. (2022) demonstrated that integrating RF, FN, and ANFIS provides accurate modeling of in-situ stresses and their impact on reservoir stability.

These studies illustrate that AI has evolved from simply predicting numerical geomechanical values to becoming an integrated and advanced approach over the years, combining multi-source data and providing adaptive estimates suitable for heterogeneous reservoirs. This evolution has improved the practical feasibility of AI applications in reservoir engineering by saving time and cost compared to conventional methods while maintaining the high accuracy required for engineering and operational design (Davoodi et al., 2024; Siddig et al., 2021; Mahmoud et al., 2022).

Chapter IV Practical Application of Artificial Intelligence in Predicting Geomechanical Properties of Tight Fractured Reservoir Quartzite El Hamra

Introduction

Geomechanics plays a crucial role in petroleum reservoir engineering (Zoback, 2010), as it provides insights into the stress and strain behavior of reservoir rocks under various operational conditions (Sone and Zoback, 2014). Understanding geomechanical properties such as Poisson's ratio, Young's modulus and the in-situ stresses is essential for optimizing reservoir management practices and ensuring safe and efficient operations (Yin et al., 2017). These properties influence key engineering decisions, including hydraulic fracturing, wellbore stability, and drilling performance (King, 2012; Maleki et al., 2014).

Accurate knowledge of geomechanical behavior helps predict how rocks will respond during drilling, production, and hydraulic fracturing operations (Mohan et al., 2017). For example, variations in rock strength and elastic moduli can impact fracture propagation, reservoir compaction, and subsidence (Jing, 2003; Zoback and Kohli, 2019). Additionally, geomechanical studies assist in designing safe well trajectories and preventing operational issues such as wellbore collapse, lost circulation, and induced seismicity (Haimson and Fairhurst, 1967; Moulin et al., 2019).

In unconventional reservoirs, especially tight and naturally fractured reservoirs, geomechanical characterization is even more critical due to the high complexity and heterogeneity of the rock (Dusseault, 2011; Dong et al., 2020). Proper geomechanical understanding ensures better planning for hydraulic fracturing, enhanced production, and reduced operational risks (Zhang et al., 2025).

In this context, predicting geomechanical properties using data-driven approaches has received increasing attention in recent years, as artificial intelligence techniques provide effective tools for modeling the complex nonlinear relationships between operational parameters and rock properties. Although Poisson's ratio and Young's modulus are both fundamental geomechanical properties, each responds differently to rock behavior during drilling operations.

Accordingly, drilling parameters were employed to estimate the dynamic Poisson's ratio in order to capture the instantaneous response of the rock under drilling conditions, while well log data were used to estimate the dynamic Young's modulus, as they provide a more stable representation of rock stiffness and elastic properties. This approach also enables a systematic comparison between the two data sources to evaluate their ability to predict these properties and to determine the most suitable option from scientific, technical, time, and cost perspectives.

This chapter is therefore organized into three main sections. The first section focuses on predicting the dynamic Poisson's ratio using drilling parameters, while the second section addresses the prediction of dynamic Young's modulus using well log data. The third section is dedicated to discussing the scientific implications of the study by analyzing the results and evaluating their methodological and practical significance for geomechanical characterization in unconventional reservoirs.

Part I. Real time dynamic Poisson's ratio prediction

Introduction

Poisson's ratio (PR) is a fundamental geomechanical property that quantifies how materials deform under applied stress, specifically describing the ratio of lateral (perpendicular) strain to axial strain (Gercek, 2007; Belyadi et al., 2019; Ashraf et al., 2021). This parameter is particularly significant in petroleum engineering operations, including hydraulic fracturing, where it directly influences fracture geometry, propagation, and containment within the reservoir (Pickett, 1963; Najibi and Asef, 2014; Onalo et al., 2018). Additionally, PR plays a vital role in drilling activities, as it helps engineers make informed real-time decisions regarding operational challenges such as stuck bottom-hole assemblies, casing pack-offs, or wellbore stability issues (Al-Zankawi et al., 2017). By accurately quantifying Poisson's ratio, hydrocarbon production can be optimized while maintaining the integrity and stability of the wellbore (Ingraffea et al., 2014).

Traditionally, Poisson's ratio is measured in the laboratory using triaxial compression tests on core plug samples (Fjaer et al., 2008; Liu, 2017). However, the reliability of this approach is often limited by the availability and quality of core samples, which are expensive and sometimes inaccessible due to complex geological conditions (Al-Mudhafar, 2017; Ameer-Zaimeche et al., 2020; Abbas & Al-Mudhafar, 2021; Ashraf et al., 2021; Al-Mudhafar et al., 2022; Sanei et al., 2023; Abdulkhaleq et al., 2024). Furthermore, natural fractures and microcracks in tight reservoirs can alter the measured PR values, as demonstrated by Walsh (1965) and more recently by Narimani et al. (2024), who highlighted the significant influence of fracture density and rock mass heterogeneity on the estimation of PR.

As an alternative to laboratory measurements, Poisson's ratio can also be estimated using acoustic log data, such as compressional (P-wave) and shear (S-wave) velocities, available in most wells. This method offers a cost-effective approach while enabling continuous profiling along the wellbore (Barree et al., 2009; Ashraf et al., 2021). Nevertheless, logging-derived estimations can be affected by missing or noisy data, leading to uncertainties, especially in unconventional tight fractured reservoirs with dense natural fracture networks. These challenges underscore the importance of combining multiple methods and carefully interpreting PR values for accurate geomechanical characterization and safe reservoir exploitation.

4.1.1. Material and methods

4.1.1.1. Data collection and description

The dataset used in this study was acquired from a single horizontal well drilled in a tight, naturally fractured reservoir within the Quartzite El Hamra formation, in 2011. The well reached a total vertical depth exceeding 807,086.64 feet, targeting a highly compact and mechanically complex lithology. The collected data include the dynamic Poisson's ratio (DPR) and a comprehensive set of drilling parameters (DD), all of which were continuously recorded at the surface in real time during the drilling of the 6-inch hole section.

The dataset comprises 247 measurement points, sampled at a consistent resolution of one measurement per meter along the drilled interval. Each recorded value of the dynamic Poisson's ratio corresponds exactly to the same depth point as the associated drilling parameters, ensuring full data alignment and consistency for subsequent analysis and modeling. The drilling parameters considered in this study include rotary speed, weight on bit, mud flow rate, standpipe pressure, torque, and rate of penetration. These parameters collectively reflect the mechanical interaction between the drilling system and the rock formation and are known to be sensitive to variations in rock mechanical behavior.

Table 1. Summary of the Measured Variables Used as Input Parameters from Mud Logging Data.

variable	Unit	Measurement Device	Abbreviation
Weight on Bit	Tons	Weight Sensor	WOB
Rate of Penetration	m/h	Drill bit performance meter	ROP
Torque	ft.lbs	Torque gauge	T
Standpipe Pressure	psi	Pressure gauge	SPP
Flow	L/min	Flow meter	Q
Drill-string Rotary Speed	min ⁻¹	RPM gauge	RPM

4.1.1.2. Compiled Dataset Characterization

As in any study based on machine learning tools, the well data were organized into Excel files (.xlsx format), with each row representing individual measurements, including logging data and the dynamic Young's modulus. During the data acquisition process, certain operational

conditions and recording tool errors may have affected the accuracy of the raw data collected by the logging instruments.

To ensure the reliability of the dataset used as input for machine learning models, it was necessary to subject the raw data to a preprocessing stage. This stage involved removing missing values, noisy data, and outliers. In this context, univariate statistics were employed, focusing on the analysis of each variable individually by summarizing its fundamental characteristics, such as measures of central tendency and variability.

a. Univariate statistics

This approach enabled the evaluation of data distributions, the identification of patterns, and the detection of outliers, thereby allowing problematic records to be identified and eliminated. Consequently, this process provided a reliable foundation for subsequent multivariate analyses (Cooksey et al., 2020).

Table 2. Statistical Analysis of value distributions for measured Poisson's ratio and DD variables.

DD	Unit	Mean	Min	Max	Std.Dev.	Skew
ROP	m/hr	1.270	0.270	3.420	0.659	0.997
WOB	t	4.373	1.200	13.200	2.100	1.459
RPM	min-1	549.982	454.000	605.000	22.398	-0.835
T	klb*f	1.649	0.000	3.650	1.045	-0.630
SPP	psi	3711.049	3057.000	4117.000	204.812	-0.063
FLW	l/mn	856.615	683.000	928.000	25.330	-2.541
DPR	/	0.078	-0.102	0.248	0.062	-0.107

Table 2. summarizes the descriptive statistics of the dataset used in this study, including the mean, minimum, maximum, standard deviation (Std.Dev), and skewness. Pearson correlation coefficient values between the dynamic Poisson's ratio and each of the input variables (Drilling Data) were also computed. This data was randomly divided into two groups; among the 247 records derived to build the model (100%). 80% of the data points were used for training, for testing purposes, the remaining 20% of the data was set aside purely for testing the model, ensuring it remained separate from the training phase to prevent any data leakage into the models (Bishop et al., 2006; Pedregosa et al., 2011; Yehia et al., 2024). In addition to dynamic Poisson's ratio, where derived from shear and compressional acoustic slowness logs, which was considered the target output. To estimate the dynamic Poisson's ratio, two equations (15) and (16) have been used; (Gassmann., 1951; Seddig et al., 2021).

$$v_{dyn} = \frac{v_p^2 - 2v_s^2}{2(v_p^2 - v_s^2)} \quad (15)$$

Where, V_{dyn} : dynamic Poisson's ratio, V_p : compressional wave velocities and V_s : shear wave velocities.

$$v_{dyn} = \frac{\frac{1}{2}(\Delta t_{shear}/\Delta t_{comp})^2 - 1}{(\Delta t_{shear}/\Delta t_{comp})^2 - 1} \quad (16)$$

Where, V_{dyn} : dynamic Poisson's ratio, Δt_{comp} : compressional slowness ($\mu s/ft$) and Δt_{shear} : shear slowness ($\mu s/ft$).

b. Bivariate statistics

Bivariate statistics, examines relationships between two variables, helping identify patterns, dependencies, and multicollinearity. It aids feature selection by highlighting variables with the strongest impact on the target and supports model evaluation by comparing predicted and actual values. Visualizations like heatmaps simplify interpretation and decision-making, enhancing the effectiveness of machine learning (James et al., 2013; Guyon and Elisseeff, 2003).

Correlation analysis was used to examine the relationship between the drilling data input variables of the model and the dynamic Poisson's ratio as the output. Figure 22 presents the Correlation matrix and correlation coefficients between the measured dynamic Poisson's ratio and the various drilling data variables. The analysis reveals that the dynamic Poisson's ratio exhibits a weak positive correlation with the torque (T: 0.12; $p = 0.068$), and standpipe pressure (SPP: 0.063; $p = 0.328$), suggesting that increases in these parameters slightly coincide with increases in Poisson's ratio. Conversely, negative correlations were observed with the rate of penetration (ROP: -0.122; $p = 0.055$), weight on bit (WOB: -0.253; $p < 0.001$), rotary speed (RPM: -0.493; $p < 0.001$), and mud flow rate (Q: -0.10; $p = 0.119$), indicating that higher values of these parameters tend to associate with a decrease in the dynamic Poisson's ratio. These results provide a preliminary understanding of how drilling mechanics may influence geomechanical behavior in tight fractured reservoirs and inform subsequent modeling and feature selection processes.

A correlation analysis was performed using a correlation matrix to evaluate the relationships among the input variables and to detect potential collinearity before training the machine learning models. The results show that most variables present low to moderate correlation levels, with correlation coefficients (r) varying between positive and negative values, indicating the absence of severe multicollinearity among the predictors used in the

models. A noticeable relationship was observed between the standpipe pressure (SPP) and the flow rate (Q), with a correlation coefficient of $r = 0.71$. This relationship can be explained by the direct physical link between the circulating drilling fluid flow rate and the standpipe pressure during drilling operations. However, this correlation remains below the critical threshold commonly associated with severe multicollinearity ($|r| \geq 0.8$) (Dormann et al., 2013), suggesting that these variables can be used simultaneously in the machine learning models without significantly affecting their stability or interpretability.

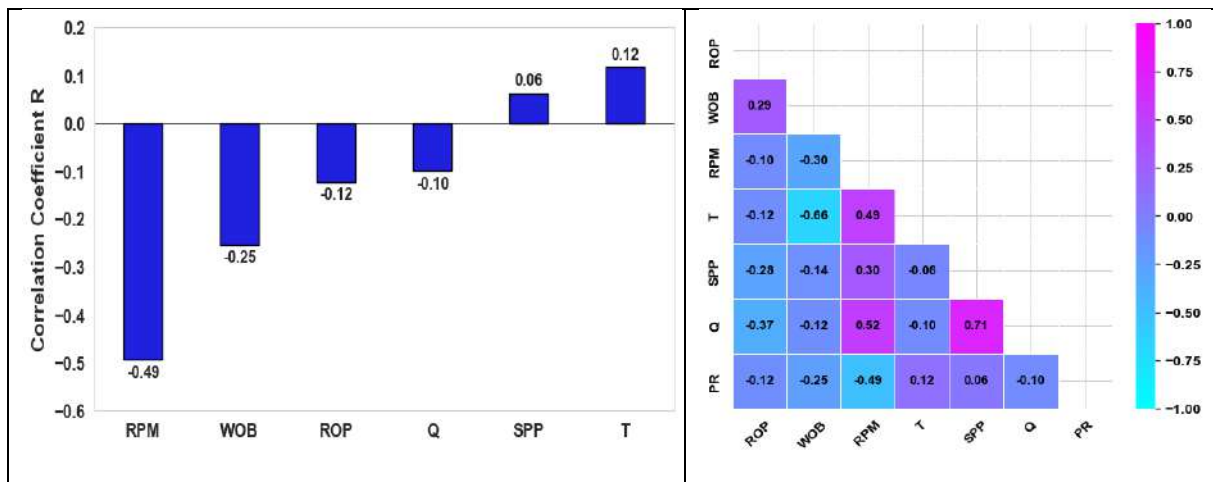


Figure 22. Correlation matrix and correlation analysis between drilling parameters and the dynamic Poisson's ratio

4.1.1.3. Implementation of Machine Learning Techniques

Three machine learning algorithms were implemented using well established models from the scikit-learn Python library, which is one of the most widely used libraries in the field of machine learning and scientific data analysis. This library provides an integrated environment for developing supervised learning models, including tools for data preprocessing, model construction, hyperparameter tuning, and statistical performance evaluation. Its use enabled the efficient and systematic implementation of the selected algorithms in a reproducible manner, thereby enhancing the reliability of the results and ensuring accurate comparison of model performance (Pedregosa et al., 2011). The adopted algorithms include Random Forest (RF), Gradient Boosting (GB), and Extreme Gradient Boosting (XGB), all of which were trained within a supervised learning framework using a subset of the available data.

a. Tuning Machine Learning Algorithm Hyperparameters

Table 3 (in the Appendix) presents the optimal hyperparameter settings selected for the three configured machine learning models, namely Random Forest (RF), Gradient Boosting

(GB), and Extreme Gradient Boosting (XGB). These parameters were determined through a tuning process aimed at improving predictive accuracy while enhancing model generalization and reducing the risk of overfitting.

For the Gradient Boosting (GB) model, a low learning rate (0.01) was adopted to ensure gradual and stable performance improvement. The maximum tree depth was limited to 3, thereby controlling model complexity, while min samples leaf = 4 and min samples split = 2 contributed to stabilizing the learning process. The number of estimators was set to 50, with a subsample value of 1.0, indicating that the entire dataset was used at each boosting iteration.

Regarding the XGBoost (XGB) model, a colsample_bytree value of 0.3 was selected to reduce tree correlation and improve model robustness. The model was configured with 100 trees, a maximum depth of 5, and a learning rate of 0.1, providing a balanced trade-off between convergence speed and predictive accuracy.

For the Random Forest (RF) model, 100 estimators were used to ensure sufficient tree diversity, while a maximum tree depth of 10 allowed the model to capture nonlinear relationships without inducing excessive overfitting.

The hyperparameter configurations summarized in this table reflect a carefully balanced compromise between predictive performance, model stability, and generalization capability, ensuring a fair and reliable comparison of the three machine learning algorithms.

b. Prediction Performance Evaluation Metrics

To evaluate and compare the performance of the models, several evaluation metrics from the sklearn.metrics module of the Python library were employed (Buitinck et al., 2013), including the coefficient of determination (R^2), root mean squared error (RMSE), and mean absolute error (MAE) as defined by Eq: (17) to (19).

These metrics played a critical role in validating prediction accuracy and guiding the hyperparameter optimization process.

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (X_i - Y_i)^2}, (0 \leq RMSE < +\infty) \quad (17)$$

$$MAE = \frac{1}{N} \sum_{i=1}^N |Y_i - X_i|, (0 \leq MAE < +\infty) \quad (18)$$

$$R^2 = \left[\frac{\frac{1}{N} \sum_{i=1}^N (XI) \cdot (\overline{XI})}{\sqrt{(XI - \overline{XI})^2} \cdot \sqrt{(YI - \overline{YI})^2}} \right]^2, (0 < R^2 < +1) \quad (19)$$

Where: (N) is the number of observations, (XI) is the true value for observation, (YI) is the predicted value for observation, $(\overline{XI})$ and $(\overline{YI})$ are the means of variables X and Y

c. Developing and Assessing Machine Learning Models

To analyze the possibility of excluding certain drilling data variables without affecting the accuracy of dynamic Poisson's ratio (DPR) predictions, a feature selection analysis was conducted based on the methodology of [Bermingham et al. \(2015\)](#). This involved a series of trial-and-error experiments to test multiple combinations of input variables and select the most effective ones, for detailed analysis to identify the optimal set of features. Each machine learning model was trained and tested on six different sets of drilling parameters, providing a systematic and transparent framework to generate a continuous profile of DPR and accurately assess the contribution of each input.

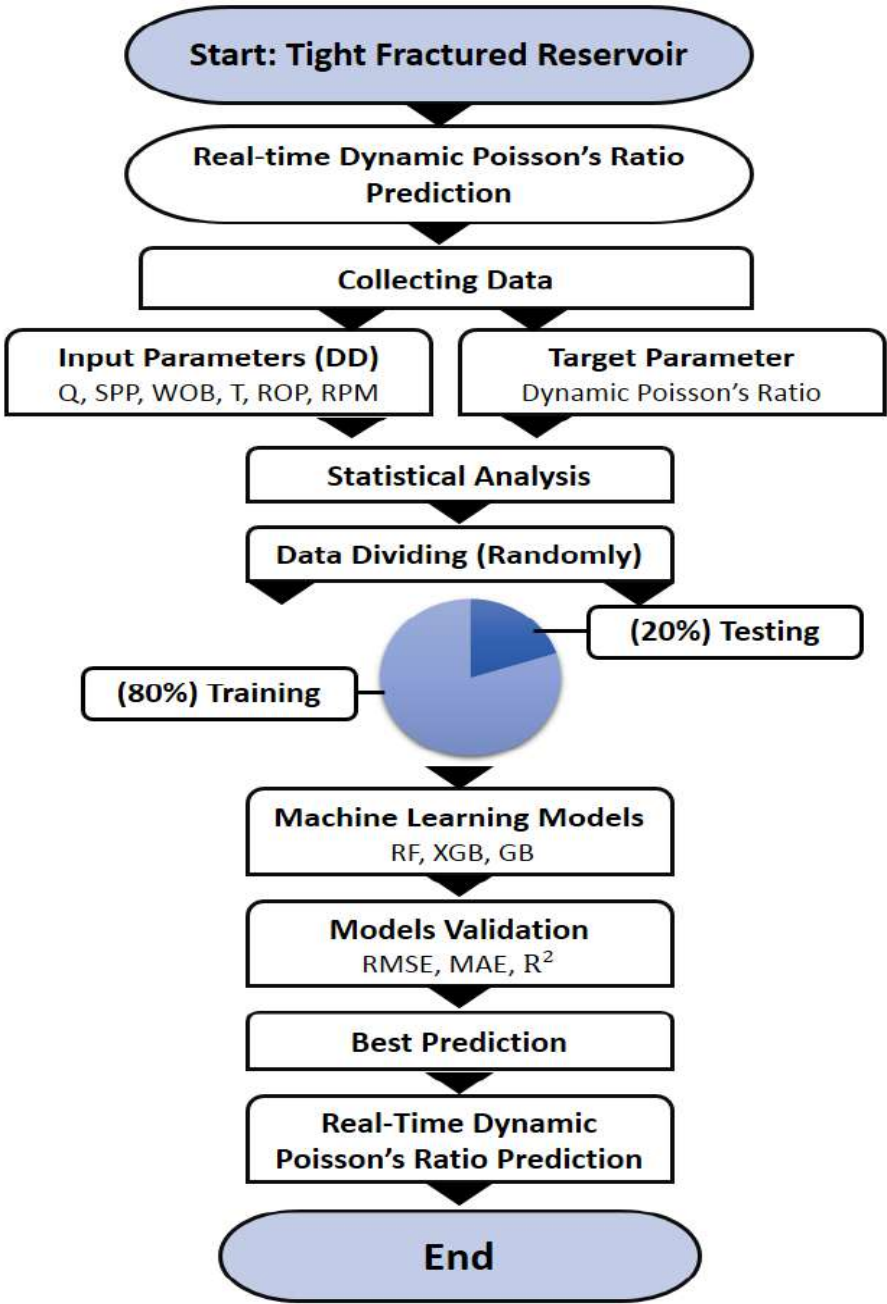


Figure 23. Flowchart illustrating the ML model architecture developed to predict formation dynamic Poisson's ratio using drilling data tight reservoir.

4.1.2. Results

Table 4 follows a sequential reduction, presenting a logical progression from a complete set of variables in (Scenario 1) to a single variable (Scenario 6), which consists only of ROP, ROP is considered as a consistent variable across all scenarios since it is a fundamental indicator of drilling efficiency and is directly influenced by formation characteristics. This reduction

allows for the observation of how each parameter's inclusion affects models performance, providing insights into the relative importance of each variable.

Table 3. Input variable scenarios of drilling data for dynamic Poisson's ratio prediction.

Combination Reference	Input Variables	Output
1	ROP, RPM, SPP, T, WOB, Q	DPR
2	ROP, RPM, SPP, T, WOB	DPR
3	ROP, RPM, SPP, WOB	DPR
4	ROP, RPM, SPP	DPR
5	ROP, RPM	DPR
6	ROP	DPR

Table 5 and Figure 24 present the comprehensive results of the prediction performance for dynamic Poisson's ratio (DPR) across six different scenarios. Out of these six scenarios, only three scenarios 1, 2, and 3 produced promising predictive outcomes. In the first scenario, the Random Forest (RF) and Extreme Gradient Boosting (XGB) models recorded lower prediction errors compared to the Gradient Boosting (GB) model. Specifically, the RMSE values were 0.027 for RF and 0.025 for XGB, while both models had MAE values of 0.020. In terms of the coefficient of determination (R^2), XGB achieved the highest value of 0.852, surpassing both RF and GB models.

For the second and third scenarios, the XGB model consistently generated lower RMSE values of 0.025 and 0.026, and MAE values of 0.019 and 0.021, respectively, when compared to other models. Additionally, XGB exhibited the best alignment with measured dynamic Poisson's ratio, achieving R^2 values of 0.856 and 0.840 in the second and third scenarios, respectively.

The RF model followed closely, with error metrics and R^2 values showing relatively high agreement with the observed data. Its highest R^2 was 0.829, reflecting close matching to XGB in terms of prediction accuracy.

In contrast, the GB model showed the highest RMSE values across the three scenarios, recorded as 0.031, 0.034, and 0.035, and the highest MAE values of 0.024, 0.024, and 0.027. Furthermore, its predictive accuracy was the lowest among the models, with R^2 values of 0.0778, 0.0745, and 0.0725 across scenarios 1, 2, and 3, respectively.

Table 4. Dynamic Poisson's ratio prediction performances using RF, GB, and XGB.

ML Scenarios	Training Set			Testing Set		
	RMSE	MAE	R-squared	RMSE	MAE	R-squared
RF 1	0.013	0.009	0.936	0.027	0.020	0.829

XGB 1	0.001	0.001	0.999	0.025	0.020	0.852
GB 1	0.010	0.010	0.936	0.031	0.024	0.778
RF 2	0.013	0.009	0.940	0.029	0.022	0.812
XGB 2	0.001	0.001	0.999	0.025	0.019	0.856
GB 2	0.013	0.010	0.934	0.034	0.024	0.745
RF 3	0.013	0.009	0.938	0.031	0.023	0.785
XGB 3	0.001	0.001	0.999	0.026	0.021	0.840
GB 3	0.015	0.011	0.920	0.035	0.027	0.725
RF 4	0.014	0.009	0.938	0.035	0.025	0.730
XGB 4	0.002	0.001	0.998	0.039	0.027	0.651
GB 4	0.015	0.012	0.916	0.038	0.027	0.682
XGB 5	0.017	0.012	0.892	0.035	0.024	0.722
GB 5	0.013	0.006	0.941	0.039	0.024	0.655
RF 5	0.018	0.014	0.877	0.035	0.023	0.719
RF 6	0.036	0.025	0.558	0.057	0.043	0.265
XGB 6	0.035	0.023	0.578	0.060	0.043	0.205
GB 6	0.036	0.024	0.554	0.056	0.042	0.292

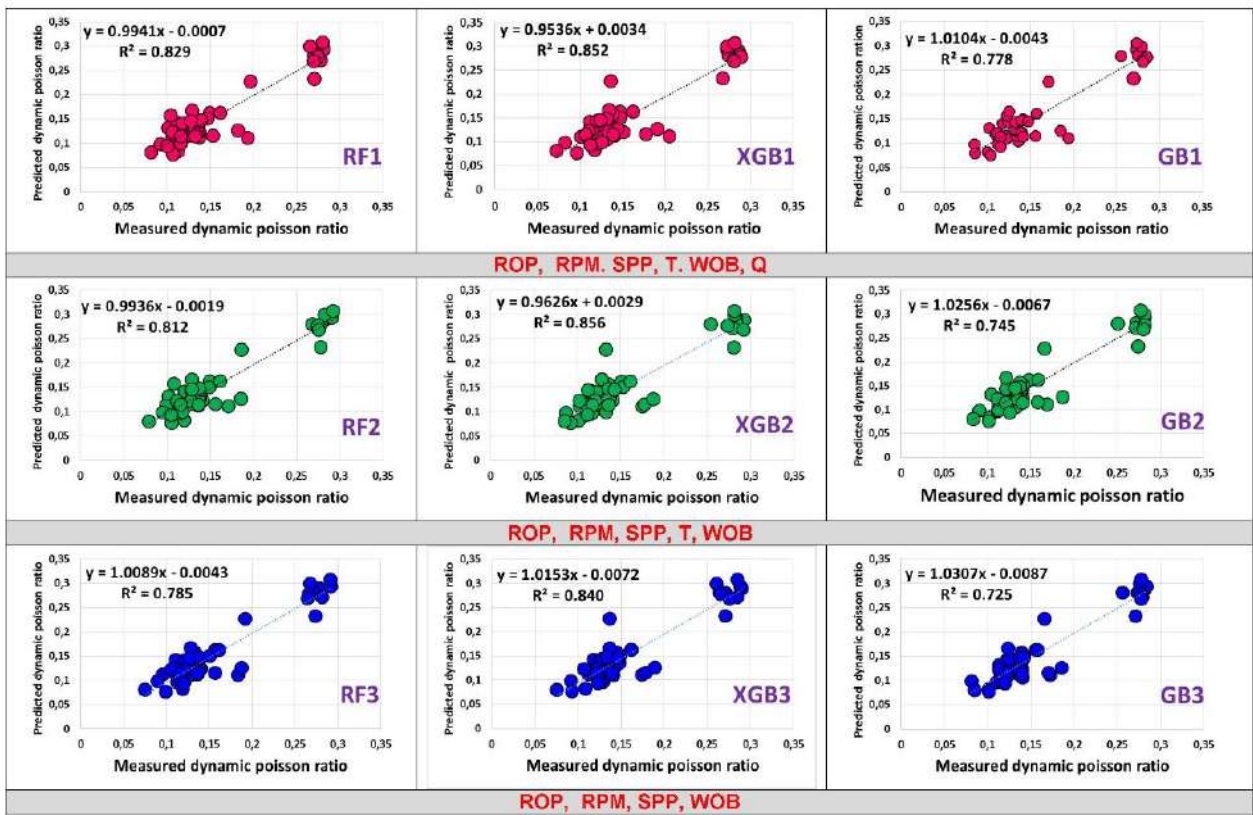


Figure 24. Measured dynamic Poisson's ratio versus predicted dynamic Poisson's ratio for the best-performing ML for the testing subset 3 best scenarios (1, 2 and 3).

The learning curve shown in [Figure 25](#) reflects the dynamic relationship between the amount of data used for training and the model's performance on both training and testing datasets. In the initial stages of adding data, the learning curve shows a continuous increase in both training and testing errors, indicating that the model temporarily struggles to adapt to new information. However, it is observed that the training error begins to decrease gradually compared to the testing error, with both remaining close to each other, indicating an initial alignment between the model's performance on training and testing data. As more data points are added, a divergence between training and testing errors begins to appear, serving as an early indicator of potential overfitting, meaning the model has started to adapt excessively to the training data at the expense of generalizing to new data. Subsequently, the training error continues to decrease gradually and converges again with the testing error as more data is incorporated, suggesting that the model is gaining better learning capability and becomes more accurate with an increasing size of the training dataset ([Goodfellow, 2016](#)).

On the other hand, the loss curve shown in [Figure 26](#) provides a detailed view of how performance loss changes during training over multiple epochs. The curve exhibits a sharp and rapid decline during the initial epochs, reflecting the model's ability to quickly capture the primary patterns in the data. After this rapid decrease, the loss curve stabilizes at a relatively low level around epochs 20, 40, and 60, and this stability continues until approximately epoch 80, indicating that the model has achieved stable and effective learning. During this phase, a slight increase in testing loss can be observed, representing a brief period of overfitting where the model is sensitive to minor variations in the testing data compared to the training data. Following this temporary increase, the loss curve returns to a low and stable level until the end of the training process, indicating the overall stability of the model's performance and its ability to learn efficiently with minimal overfitting issues ([Bishop et al., 2006](#); [Smith, 2017](#); [Thongsuwan, 2021](#)).

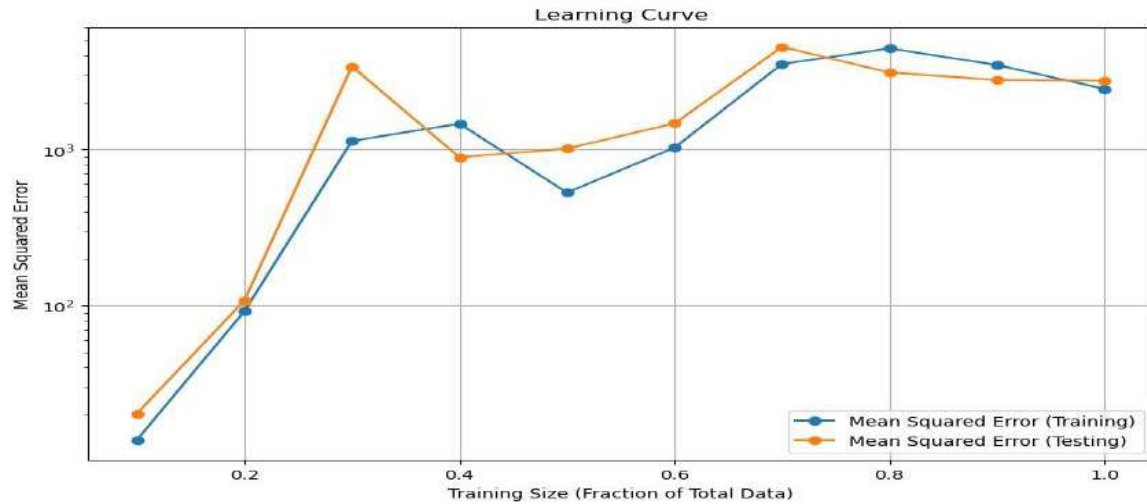


Figure 25. Learning curve of the mean squared error as a function of the training data size

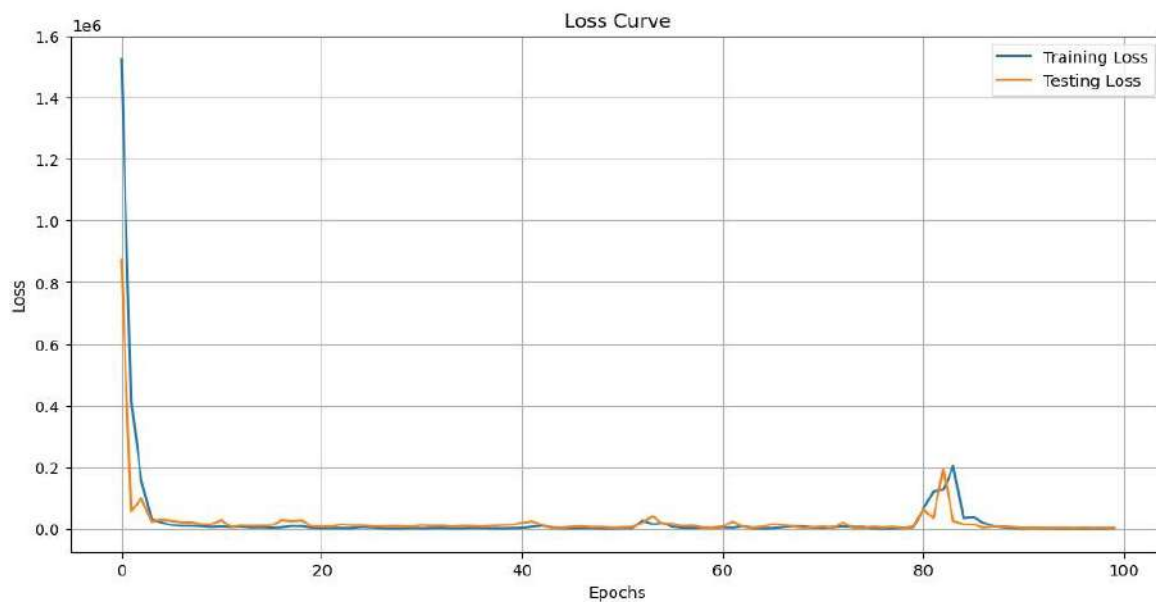


Figure 26. Loss curve behavior during model training and validation

4.1.3. Discussion

4.1.3.1. Evaluation of ML in predicting dynamic Poisson's in tight fractured reservoir

Table 6 presents a comprehensive summary of recent studies that have addressed the prediction of Poisson's ratio in various types of hydrocarbon reservoirs, including carbonate, sandstone, and shale reservoirs, using a wide range of machine learning algorithms across different regions worldwide. These studies highlight the growing trend toward applying artificial intelligence techniques to predict geomechanical properties, relying on different types of data particularly well logging data and, in some cases, drilling parameters as primary inputs.

This comparative overview allows the results of the present study to be positioned within a broader scientific context by relating them to previously published work.

In this context, [Alakbari et al. \(2024\)](#) employed a deep learning-based algorithm to predict Poisson's ratio using only well logging data, specifically bulk density (RHOB), compressional transit time (DTc), and shear transit time (DTs). Their results demonstrated a high level of agreement between measured and predicted values, with correlation coefficient values reaching 0.967. Similarly, [Elkatatny \(2018\)](#) investigated the prediction of static Poisson's ratio in a carbonate reservoir using the same set of logging inputs (RHOB, DTc, and DTs), but through a comparative evaluation of several machine learning algorithms, including Artificial Neural Networks (ANN), Support Vector Machines (SVM), and the Adaptive Neuro-Fuzzy Inference System (ANFIS). Among these models, ANFIS achieved the best performance, with correlation coefficients as high as 0.985. These values are very close to those reported by [Alakbari et al. \(2024\)](#).

On the other hand, [Tariq et al. \(2017\)](#) focused on a reduced input set, applying an Artificial Neural Network model to predict Poisson's ratio in a carbonate reservoir using only compressional and shear wave velocities (V_p and V_s). Despite the limited number of input variables, the model exhibited excellent performance, with a correlation coefficient of 0.985. These findings were supported by a subsequent study by [Tariq et al. \(2018\)](#), in which the input dataset was expanded to include gamma ray (GR), bulk density (RHOB), along with V_p and V_s , using a Functional Network algorithm. The model maintained a similarly high level of predictive accuracy.

Overall, the studies summarized in [Table 6](#) demonstrate that machine learning-based approaches are effective for predicting Poisson's ratio across different reservoir types, particularly when well logging data are used as inputs. Moreover, these works show that the developed models are capable of producing smooth and reliable Poisson's ratio predictions with high statistical accuracy, highlighting the strong potential of machine learning techniques for geomechanical reservoir characterization.

Unlike some authors (e.g., [Alakbari et al., 2024](#); [Elkatatny, 2018](#); [Tariq et al., 2018](#); [Tariq et al., 2017](#)), transit time (Δt) is a vital parameter for predicting the Poisson's ratio. Therefore, it represents the time taken by acoustic waves to travel through the rock, providing essential information on rock mechanical properties. Missing or inaccurate transit time readings, often affected by changes in mud properties, can compromise predictions. This issue is more significant in heterogeneous reservoirs and tight formations with extensive natural fracture networks and high pressure, where accurate prediction during drilling is challenging.

To prevent damage to rock properties, measures such as determining optimal mud weight are applied, especially when using dense oil-based mud. Maintaining hydraulic column pressure equal to or slightly above reservoir pressure ensures wellbore stability and minimizes formation risks. However, increasing mud density can reduce sound wave propagation, leading to inaccurate or lost transit time readings (Zerrouki et al., 2014).

4.1.3.2 Feature importance significance and implications

In our case study of the tight fractured Quartzite El Hamra reservoir, three prominent machine learning algorithms were evaluated: Random Forest (RF), Extreme Gradient Boosting (XGB), and Gradient Boosting (GB). Drilling parameters, such as rate of penetration (ROP), weight on bit (WOB), rotary speed (RPM), mud flow rate (Q), standpipe pressure (SPP), and torque (T), served as input variables for the models. These parameters are readily available in all drilled wells and can be automatically recorded in real time during drilling, making data collection efficient in terms of both time and cost without requiring additional equipment (Epelle et al., 2020).

The machine learning models underwent rigorous training and validation for each algorithm using a carefully divided dataset. Six different feature-selection scenarios were tested, encompassing various combinations of input variables, as detailed in Table 3.

According to the performance evaluation of the examined models, the Extreme Gradient Boosting (XGB) algorithm in the second scenario emerged as the super-performing model, demonstrating the highest alignment with the dynamic Poisson's ratio and the lowest prediction errors ($R^2 \approx 0.856$, $RMSE \approx 0.025\%$, and $MAE \approx 0.019\%$). Figure 27 and 28 presents the relative importance of input features and the SHapley Additive exPlanations (SHAP) summary plot for the XGB model, emphasizing the dominant role of rotary speed (RPM), which ranks as the most influential variable. Notably, RPM is a mechanical drilling parameter that directly impacts formation deformation (Wei et al., 2020). Considering the surface rotational speed, the drill bit can rotate at a rate up to three times higher within the formation (Arjun and Teodoriu, 2013), thereby enhancing the processes of formation deformation and disruption.

Furthermore, the Extreme Gradient Boosting (XGB) model shows a notable relative importance for rotational torque (T), as depicted in Figure 27 and 28. The inclusion of this parameter improves model performance, highlighting a flexible behaviour relationship between the algorithm and torque, which contributes to an increased drilling penetration rate in the reservoir formation (Payne et al., 1997; Xie et al., 2018). As the torque applied to the drill bit rises, the rotational force increases (Alam et al., 2011), thereby facilitating deformation and

breakdown of the formation. Considering that this study is based on a horizontal well in the tight fractured Quartzite El Hamra reservoir, the extensive natural fracture network leads to higher fluid losses compared to other reservoir types (Li et al., 2016; Wang et al., 2018). Consequently, standpipe pressure (SPP) tends to decrease with increasing drilling fluid loss (Labenski et al., 2003), which explains the relatively low importance assigned to SPP by the XGB model.

Moreover, maintaining wellbore stability during horizontal drilling presents specific challenges due to the rising lateral pressure as drilling advances, along with the requirement to preserve directional control (Wang and Sterling, 2007). The impact of weight on bit (WOB) becomes more intricate in horizontal wells compared to vertical wells, being affected by factors such as stabilizer diameter, their placement, and borehole diameter (Lian et al., 2015). As a result, the optimal WOB in horizontal drilling tends to be lower than in vertical drilling to compensate for the increasing frictional resistance encountered during the process (Chen et al., 2018). Consequently, the XGB algorithm assigns limited relative importance to WOB, which slightly reduces the prominence of the rate of penetration (ROP) in the model's feature ranking.

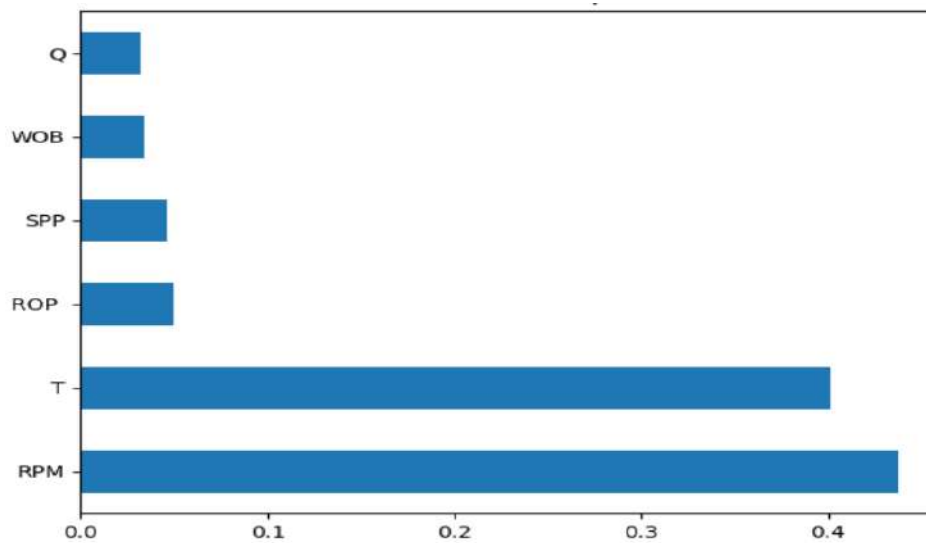


Figure 27. Feature importance scores illustrating the contribution of each drilling parameter to the prediction of dynamic Poisson's ratio.

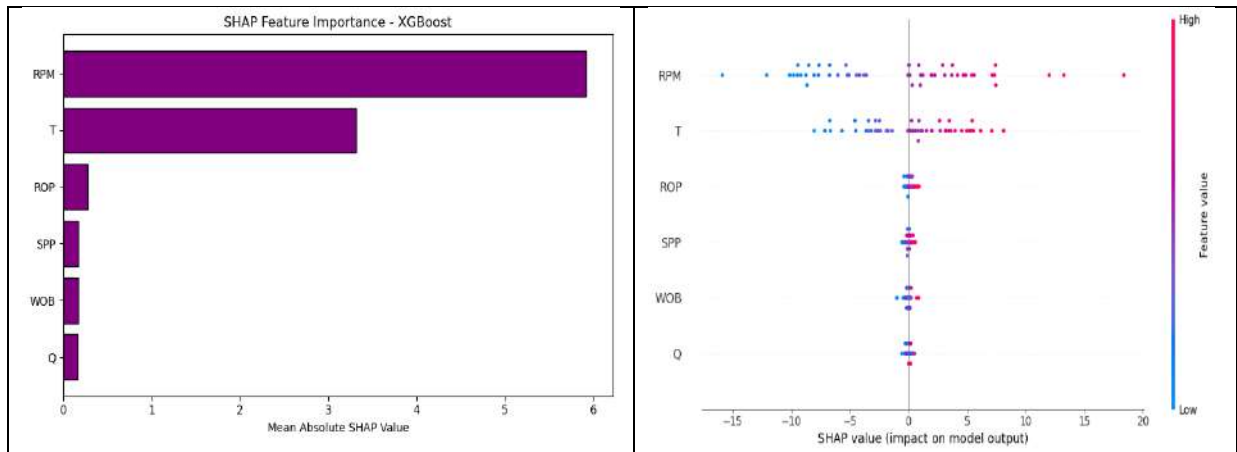
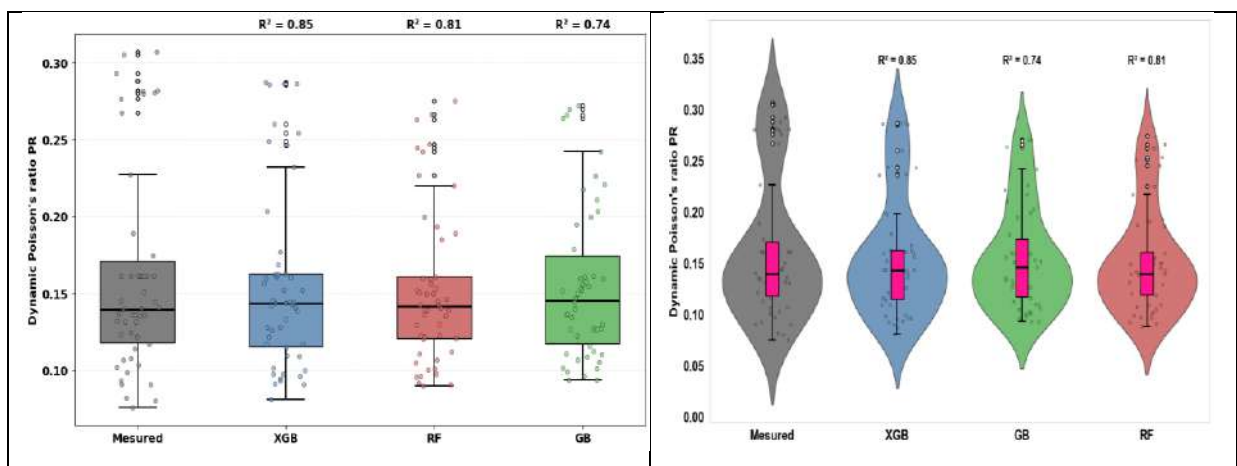


Figure 28. Model interpretability illustrated through the SHapley Additive exPlanations (SHAP) summary plot.

4.1.3.3. Evaluation of XGBoost Performance for Real-Time Prediction of Dynamic Poisson's Ratio Using Drilling Data

This research demonstrates the adaptability of the Extreme Gradient Boosting Algorithm, applied for the first time in this study to predict dynamic Poisson's ratio in tight fractured reservoirs. To the authors' knowledge, no prior study has specifically investigated dynamic Poisson's ratio prediction in such reservoirs using XGB. Nevertheless, the algorithm exhibited exceptional outstanding superiority compared to other machine learning methods [Figure 29](#), Random Forest and Gradient Boosting when employing drilling parameters as inputs, even under the challenging conditions inherent to horizontal drilling in tight naturally fractured and reservoir. These findings confirm XGB as the most responsive and flexible algorithm for real-time predictions during drilling operations.



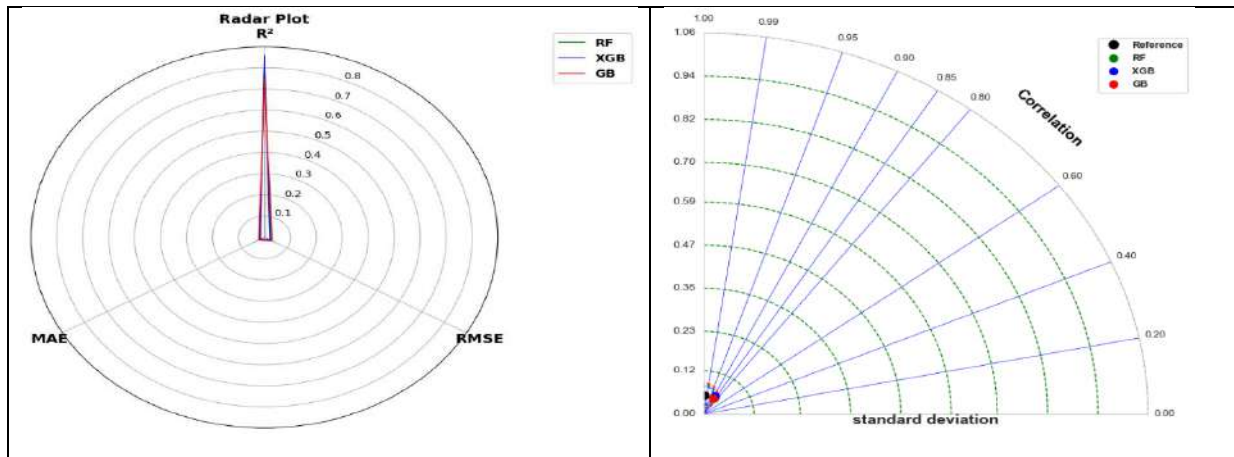


Figure 29. Comparative visualization of dynamic Poisson's ratio predictions from three machine learning models using box, violin, radar plots, and a Taylor diagram for the QH-Full Reservoir.

Furthermore, as summarized in Table 6, several key studies employed real-time drilling parameters combined with different machine learning algorithms to predict Poisson's ratio. Siddig et al. (2021) developed models based on Artificial Neural Networks (ANN) and Adaptive Neuro-Fuzzy Inference Systems (ANFIS) for three reservoir types carbonate, sandstone, and shale using WOB, T, SPP, RPM, ROP, and Q as input parameters. These models exhibited high predictive performance with R-values ranging from 0.98 to 0.99. Ahmed et al. (2021) applied drilling data along with ANN, ANFIS, and Support Vector Machine (SVM) algorithms in a complex lithology for predicting static Poisson's ratio, achieving R-values between 0.90 and 0.91. It is noteworthy that the performance of ANN and ANFIS models reported by Siddig et al. (2021) decreased when applied to reservoirs of different characteristics, indicating that reservoir type and properties can significantly affect the influence of drilling parameters. Later, Siddig and Elkatatny (2021) successfully developed ANN, ANFIS, and Random Forest models using drilling data inputs, reaching excellent performance comparable to models using logging data, with R-values of 0.99. When comparing these results, particularly those obtained with XGB in our study, with previous literature using either drilling data (Siddig et al., 2021; Ahmed et al., 2022; Siddig and Elkatatny, 2021) or logging data (Alakbari et al., 2024; Elkatatny, 2018; Tariq et al., 2017, 2018), it is evident that our predictions are more accurate.

Table 5. Results from studies designed to predict Poisson's ratio from drilling and well logging variables inputs.

	In put	Type Reservoirs	Algorithms	Performances	References
Logging data	RHOB, DTc and DTs	/	-Deep learning	R= 0.967 AAPRE=3.228, APRE=1.054 RMSE =4.389	Alakbari et al., 2024
	RHOB, DTc and DTs	Carbonate	- ANN - ANFIS - SVM	R = 0.97 AAPRE=1.13	Elkatatny., 2018
	GR, RHOB, Vp and Vs	Carbonate	- FN	R = 0.985	Tariq et al., 2018
	Vp and Vs	Carbonate	- ANN	R = 0.985	Tariq et al., 2017
Drilling data	WOB, SPP, RPM, ROP, T, and Q	/	- ANN - SVM - RF	R= 0.99	Siddig and Elkatatny., 2021
	WOB, SPP, RPM, ROP, T, and Q	Complex lithology	- ANN - ANFIS - SVM	-ANN (R =0.90, AAPE=8.12) -ANFIS (R=0.91, AAPE=7.25) -SVM (R=0.90, AAPE=6.57)	Ahmed et al., 2021
	WOB, T, SPP, RPM, ROP and Q	Carbonate, sandstone, and shale.	- ANN - ANFIS	- ANN R =0.99 - ANFIS R=0.98	Siddig et al., 2021
	WOB, T, SPP, RPM, ROP and Q	Tight Fractured Reservoir	- RF - XGB - GB	- XGB (R≈0.92, RMSE≈0.025, MAE≈0.019) - RF (R≈0.91, RMSE≈0.027, MAE≈0.020) - GB (R≈0.88, RMSE≈0.031%, MAE≈0.024)	This study

NB: To ensure a consistent comparison, the R^2 values were converted into R values. The evaluation metrics considered in the studies included the correlation coefficient (R) between predicted and measured Poisson's ratio, along with the reported error percentages. However, some studies did not provide error percentages and only reported the correlation coefficient (R). Consequently, due to the absence of complete error data, the following comparative analysis is based solely on the available R -values for consistency.

Therefore, employing the Extreme Gradient Boosting (XGB) algorithm in combination with drilling data for real time prediction is regarded as a highly effective strategy to tackle the challenges associated with predicting the dynamic Poisson's ratio in tight fractured reservoirs. These reservoirs are characterized by high heterogeneity and exhibit unpredictable behavior during drilling, particularly in horizontal wells. Key difficulties include missing transit time measurements, which are affected by drilling fluid properties and downhole conditions, the expensive nature of core sampling, and issues arising from fractures that impede core recovery.

It should be emphasized that this methodology can also be assessed and potentially extended to other types of tight reservoirs and varying depths, ensuring broader applicability and reliability.

4.1.3.4. Advantages and recommendations for future research to improve the technique

Unlike well logs and core samples, drilling parameters are continuously available and systematically recorded by mud logging services during drilling operations in every well. It is important to highlight the critical role of mud logging, which provides real-time data that can serve as input for predicting and modeling missing reservoir parameters using historical data and machine learning techniques. Despite their availability in all wells, the use of these parameters can be limited by drilling conditions, operational strategies, and the direct impact of reservoir petrophysical properties. Therefore, it is essential to recognize the value of drilling parameters, ensure their careful monitoring, and improve the accuracy of sensing devices to enable a robust approach based on drilling data and machine learning models. Additionally, incorporating more parameters and expanding the dataset such as pore pressure, mud properties, and real-time gas measurements during drilling can further enhance model performance and training.

Surface drilling parameters are adjusted according to the prevailing drilling conditions, particularly the borehole diameter. This adjustment provides petroleum engineers with valuable information to tackle the various challenges associated with predicting the dynamic Poisson's ratio in such reservoirs. Accurate control and understanding of this critical mechanical property are essential in petroleum engineering, especially during drilling operations. It aids in forecasting the risk of borehole collapse and fracture development by estimating in-situ stresses and determining the optimal mud weight window (Gholami et al., 2014). Furthermore, it supports the real-time identification of deformation zones, enabling the calculation of the appropriate density and volume of cement slurry required to resist deformation, prevent cracking, and avoid debonding from the formation or casing (Ravi et al., 2002; Xi et al., 2021). This approach ultimately ensures the long-term stability and operational integrity of the well.

Poisson's ratio is a key parameter in hydraulic fracturing, as it helps estimate the minimum horizontal stress necessary to determine the fracture initiation pressure window. It also plays a vital role in optimizing fracture fluid properties and injection rates and has a direct impact on fracture geometry, including fracture length and height. Incorporating Poisson's ratio into predictive models is crucial for accurately forecasting fracture propagation, which in turn improves fracture treatment design, particularly in naturally fractured reservoirs (Nagel et al., 2012; Takbiri et al., 2019).

This methodology not only advances future research but also enhances the exploitation and performance management of tight fractured reservoirs. Moreover, it contributes to more efficient drilling operations by reducing additional costs and saving time.

Conclusion

This study applied Gradient Boosting, Random Forest, and Extreme Gradient Boosting machine learning algorithms to develop an advanced approach for real-time, continuous prediction of dynamic Poisson's ratio in a tight fractured reservoir. The predictions were based on commonly recorded drilling parameters including ROP, RPM, SPP, WOB, T, and Q collected from a horizontal well in the Quartzite El Hamra formation within the Rhourde Nouss field. A thorough comparison of the three models across multiple input scenarios showed that the XGB model achieved the highest predictive accuracy, with a correlation coefficient of approximately $R^2 \approx 0.856$ and the lowest error rates (RMSE ≈ 0.025 and MAE ≈ 0.019), outperforming both GB and RF.

The findings demonstrate that using drilling parameters as input for predicting dynamic Poisson's ratio can achieve results comparable to those obtained with well logging data across different machine learning models. Additionally, this approach is particularly valuable when transit time data is unavailable, eliminating the need for laboratory measurements on core samples and reducing both costs and time. Continuous real-time monitoring and practical application of drilling parameters during operations provide a reliable understanding of formation characteristics, with the XGB model showing high predictive capability due to its computational flexibility in handling multiple drilling inputs.

This approach advances future research, enhances the exploitation of tight fractured reservoirs, and contributes economically by optimizing drilling operations and minimizing additional costs. By focusing on predictive modeling of mechanical properties in complex, tight, and highly fractured reservoirs, this work addresses a critical gap in the literature, illustrating how real time data can improve drilling efficiency without extra expenses. The results are particularly useful for petroleum engineers in optimizing wellbore stability, forecasting fracture initiation, refining well design, and ensuring continuous monitoring and control of reservoir performance in similar tight reservoir environments.

Part II. Dynamic Young's Modulus Prediction

Introduction

Young's modulus is one of the most critical parameters for evaluating the mechanical behavior of rocks, as it quantifies the stiffness of a material under axial stress. It is defined as the ratio of stress (σ) to strain (ϵ) within the elastic range of deformation, according to Hooke's law as Eq (20):

$$\sigma = E\epsilon \quad (20)$$

This modulus is fundamental in geomechanics and rock physics, where it serves as a key indicator of a rock's ability to deform elastically and recover its shape once the applied stress is removed (Sanei et al., 2023). Accurate knowledge of Young's modulus is vital for applications such as wellbore stability analysis, hydraulic fracturing design, subsidence prediction, and in-situ stress estimation (Jaeger et al., 2007; Fjaer et al., 2008; Zoback, 2010).

Young's modulus can be measured either statically or dynamically. Static measurements are typically obtained from laboratory experiments on core samples subjected to uniaxial or triaxial compression, while dynamic measurements are derived from compressional (V_p) and shear (V_s) wave velocities recorded through acoustic well logging (Siddig and Elkatatny, 2021). Dynamic values often exceed static ones due to differences in testing conditions: dynamic measurements involve high-frequency elastic waves that do not damage the rock, whereas static loading can induce microcracking, leading to lower observed stiffness (Onalo et al., 2018; Elkatatny, 2022). To bridge this gap, numerous empirical correlations have been developed to estimate static Young's modulus from dynamic data (Eissa and Kazi, 1988; Chang et al., 2006).

Understanding and accurately determining Young's modulus is therefore essential in constructing reliable geomechanical models, especially for unconventional reservoir specially in fractured tight reservoirs where variations in petrophysical parameters can significantly affect elastic behavior (Castagna et al., 1985; Mavko et al., 2009).

This approach highlights the importance of Young's modulus as a key parameter in characterizing the geomechanical performance of rocks and in enabling more accurate and effective solutions in unconventional reservoirs.

4.2.1. Material and methods

4.2.1.1. Data collection and description

The dataset employed in this study originates from a single vertical well drilled within a tight, naturally fractured reservoir referred to as Quartzite El Hamra, in 2015. This well reached a total vertical depth of over 200.8 m. The collected data comprises both the dynamic Young's modulus (E) and several logging parameters (LD), which are essential in hydrocarbon exploration and reserve evaluation (Szabo., 2011). Geomechanical and Petrophysical properties were extracted from well log data using various empirical equations and theoretical frameworks (Chitale et al., 2003; Asquith et al., 2004). These parameters, extensively referenced in geological literature, also function as key inputs for various studies (Ellis et al., 1988).

The dataset includes 1894 measurements; each recorded at one-centimetre intervals. Each value of the dynamic Young's modulus (E) corresponds to associated logging parameter values, including depth (DEPTH) in meters, records borehole diameter (CALI) in inches, Resistivity measurements at various depths of investigation (RT01, RT02, RT03, RT06, RT09, RT12) in ohm·m, neutron porosity (NPHI) in porosity units, photoelectric factor (PE) in cm²/g, and natural gamma-ray (GR) in API units, the elemental concentrations of potassium (POTA) in %, thorium (THOR) and uranium (URAN) in ppm.

The reservoir was divided into two reservoirs based on wellbore diameter (CALI): the first reservoir (QH-Class 1), with a diameter less than or equal to 7", contains 1828 measurements, while the second reservoir (QH-Class 2), with a diameter greater than 7", contains 181 measurements. Additionally, the entire reservoir (QH-Full Reservoir) was considered as a third reservoir.

The dataset of the three reservoirs were split randomly into two portions: 80% of the data was allocated for training the machine learning model, while the remaining 20% was kept for testing. This approach ensured that the test data remained entirely separate from the training process, thus avoiding any risk of data leakage. The dynamic Young's modulus, which served as the predicted variable, was calculated using shear and compressional acoustic slowness logs, applying Gassmann's (1951) Eq (21).

$$E = \frac{\rho \cdot V_s^2 (3v_p^2 - 4v_s^2)}{v_p^2 - v_s^2} \quad (21)$$

Where, E : dynamic Young's modulus, ρ : density (kg/m^3), V_p : compressional wave velocities (m/s) and V_s : shear wave velocities (m/s).

Table 6. Summary of the Measured Variables Used as Input Parameters from Logging Data.

variable	Unit	Measurement Device	Abbreviation
Borehole Diameter	inches	CALI log tool	CALI
Resistivity; RT01, RT02, RT03, RT06, RT09, RT12	ohm·m	Electrical Resistivity Logging Tool	RT
Neutron Porosity	porosity units	Neutron Density Logging Tool	NPHI
Photoelectric Factor	cm ² /g	Density/Photoelectric Logging Tool	PE
Natural Gamma-Ray	API units	Gamma Ray Spectrometer	GR
Potassium	%	Spectrometry Logging Tool	POTA
Thorium	ppm	Spectrometry Logging Tool	THOR
Uranium	ppm	Spectrometry Logging Tool	URAN

4.2.1.2. Compiled Dataset Characterization

In this section, the same data processing steps applied in the dynamic Poisson's ratio study were followed. Well log data for the dynamic Young's modulus were organized and subjected to the same preprocessing procedures, including data cleaning and quality checks, to ensure the accuracy and reliability of the inputs used for the machine learning models.

Table.8 provides the descriptive statistics of the dataset, including the mean, minimum, maximum, standard deviation, and skewness. Furthermore, Pearson correlation coefficients were determined to evaluate the relationships between the dynamic Young's modulus and the various logging parameters.

Table 7. Analyse statistique de la distribution des valeurs des LD mesurées et du module de Young's calculé.

QH-Full Reservoir								
variable	mean	min	25	50	75	max	std	Skewness
DEPT	3501.287	3402.100	3449.725	3497.850	3555.575	3602.900	59.257	0.059
CALI	6.423	6.087	6.189	6.265	6.565	8.299	0.366	1.767
RT01	303.054	8.595	20.914	41.336	107.972	2000.000	597.291	2.045
GR	97.018	12.309	40.566	71.783	163.665	238.779	60.554	0.338
RT02	76.990	8.610	20.288	36.918	77.562	755.507	93.772	2.112
NPHI	0.077	0.000	0.019	0.038	0.154	0.225	0.070	0.640
RT03	56.230	8.619	20.108	35.401	68.492	608.788	52.540	2.036
PE	3.322	1.836	2.488	3.602	3.984	6.464	0.854	-0.170
RT06	45.356	8.627	19.613	33.350	60.050	265.462	33.552	1.218
POTA	2.165	0.131	0.755	1.051	4.313	5.479	1.775	0.614

RT09	43.011	8.628	19.527	32.868	58.038	213.074	30.047	1.089
THOR	11.951	1.708	5.109	8.791	18.853	34.875	7.293	0.334
RT12	42.294	8.628	19.452	32.594	57.179	206.500	29.115	1.077
URAN	2.153	0.264	1.117	1.885	3.062	7.330	1.236	0.895
E	58.188	33.645	47.719	59.674	66.877	94.860	11.878	-0.110
QH-Class 1 Reservoir								
DEPT	3506.153	3402.100	3456.275	3501.950	3559.725	3602.900	59.870	0.040
CALI	6.343	6.087	6.186	6.260	6.415	6.999	0.231	1.316
RT01	297.569	8.595	20.605	40.250	98.717	2000.000	596.020	2.070
GR	96.626	12.309	40.130	65.081	165.035	238.779	61.364	0.355
RT02	75.266	8.610	19.996	36.045	72.140	755.507	93.962	2.163
NPHI	0.079	0.000	0.018	0.038	0.156	0.225	0.071	0.593
RT03	54.857	8.619	19.742	34.468	64.367	608.788	52.579	2.128
PE	3.234	1.836	2.337	3.402	3.956	6.464	0.871	0.043
RT06	44.231	8.627	19.300	32.336	56.786	265.462	33.454	1.314
POTA	2.213	0.131	0.765	1.098	4.362	5.479	1.787	0.561
RT09	41.948	8.628	19.216	31.854	54.955	213.074	29.926	1.187
THOR	11.729	1.708	5.099	7.997	18.817	34.875	7.274	0.390
RT12	41.253	8.628	19.189	31.688	53.999	206.500	28.989	1.176
URAN	2.093	0.264	1.068	1.786	3.017	7.330	1.223	0.943
E	57.988	33.645	47.209	59.506	66.787	94.860	11.972	-0.084
QH-Class 2 Reservoir								
DEPT	3511.931	3406.300	3417.900	3559.600	3570.200	3601.800	76.549	-0.325
CALI	7.294	7.000	7.090	7.197	7.444	8.299	0.277	1.320
RT01	194.633	17.672	32.061	57.884	127.801	1964.569	420.544	3.280
GR	65.694	27.999	38.896	45.050	93.046	154.934	36.662	0.934
RT02	69.269	17.474	30.277	49.735	88.423	367.946	62.351	2.513
NPHI	0.030	0.006	0.020	0.026	0.038	0.059	0.014	0.540
RT03	56.364	17.327	29.049	44.754	76.502	192.199	36.377	1.455
PE	3.779	2.756	3.610	3.814	3.990	4.439	0.311	-0.711
RT06	48.165	17.327	28.139	39.221	65.656	123.962	24.953	0.922
POTA	0.917	0.540	0.751	0.850	0.981	2.156	0.311	2.175
RT09	46.181	17.327	27.835	38.000	62.473	111.106	22.667	0.822
THOR	9.938	3.465	4.494	6.330	15.746	26.843	6.852	0.898
RT12	45.566	17.327	27.828	37.310	61.723	107.130	22.035	0.794
URAN	2.013	0.414	1.066	1.414	2.941	5.606	1.325	1.128
E	63.337	53.816	59.367	62.450	66.674	74.535	5.389	0.468

NB: QH-Class 1 Reservoir: $CALI < 7$, QH-Class 2 Reservoir: $CALI > 7$.

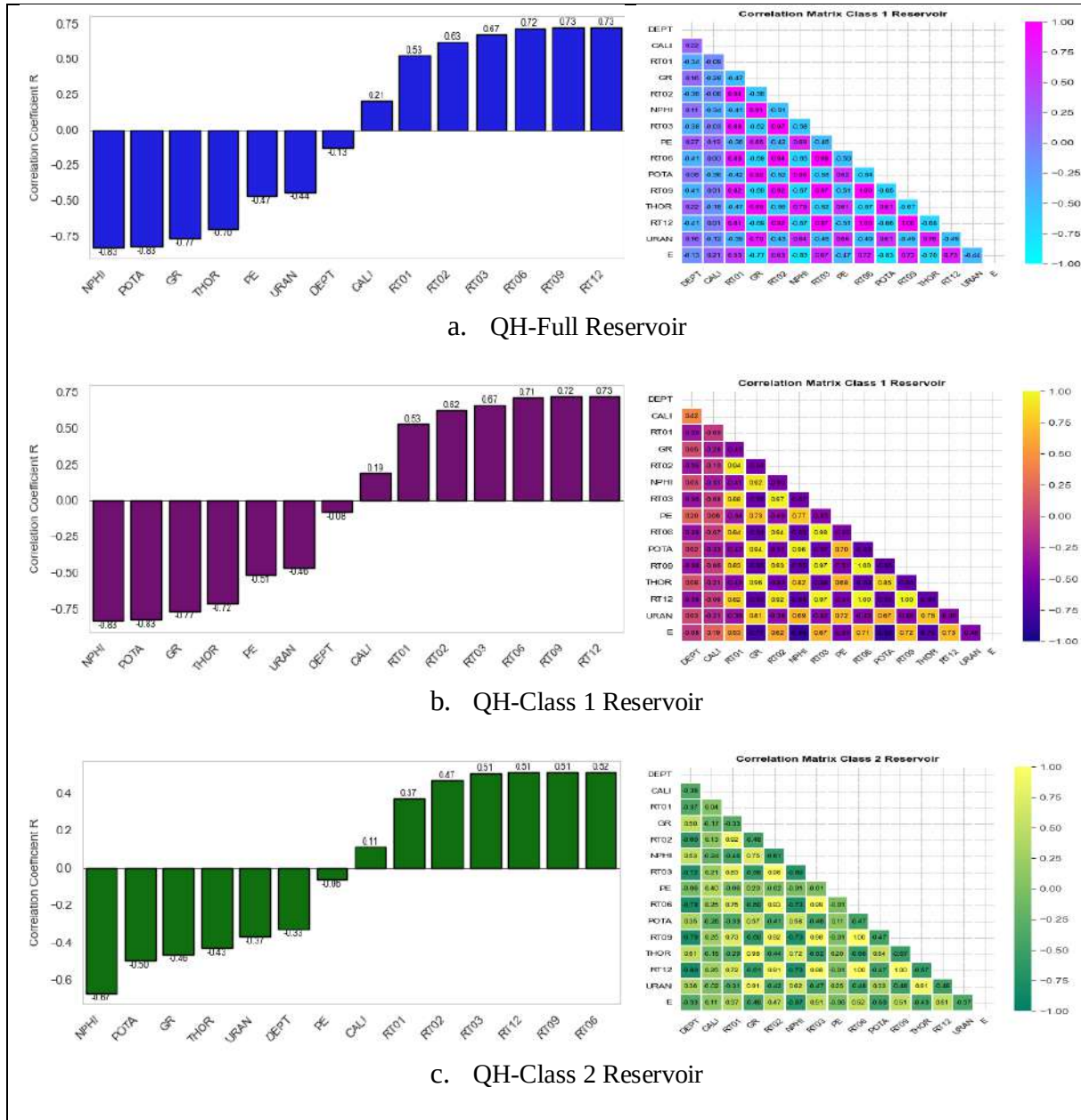


Figure 30. Correlation matrix and correlation coefficient between the dynamic Young's modulus E and logging data for three reservoirs.

A statistical evaluation was conducted to investigate how the drilling parameters (model inputs) are associated with the dynamic Young's modulus (model output) in the QH-Full Reservoir. The analysis reveals positive correlations between E and variables such as CALI, RT01, RT02, RT03, RT06, RT09, and RT12, with correlation coefficients (R) of (R = 0.208; $p < 0.001$), (R = 0.528; $p < 0.001$), (R = 0.626; $p < 0.001$), (R = 0.671; $p < 0.001$), (R = 0.720; $p < 0.001$), (R = 0.730; $p < 0.001$), and (R = 0.730; $p < 0.001$), respectively. Conversely, E is negatively correlated with DEPT, GR, NPHI, PE, POTA, THOR, and URAN, exhibiting R values of (R = -0.132; $p < 0.001$), (R = -0.768; $p < 0.001$), (R = -0.832; $p < 0.001$), (R = -0.471;

$p < 0.001$), ($R = -0.826$; $p < 0.001$), ($R = -0.698$; $p < 0.001$), and ($R = -0.440$; $p < 0.001$), respectively, as illustrated in [Figure 30.a](#).

In the QH-Class 1 Reservoir, statistical correlation analysis revealed that the dynamic Young's modulus (E) exhibited positive correlations with CALI ($R = 0.188$; $p < 0.001$), RT01 ($R = 0.529$; $p < 0.001$), RT02 ($R = 0.623$; $p < 0.001$), RT03 ($R = 0.666$; $p < 0.001$), RT06 ($R = 0.715$; $p < 0.001$), RT09 ($R = 0.725$; $p < 0.001$), and RT12 ($R = 0.725$; $p < 0.001$). Conversely, E showed negative correlations with DEPT ($R = -0.080$; $p < 0.001$), GR ($R = -0.770$; $p < 0.001$), NPHI ($R = -0.831$; $p < 0.001$), PE ($R = -0.514$; $p < 0.001$), POTA ($R = -0.828$; $p < 0.001$), THOR ($R = -0.715$; $p < 0.001$), and URAN ($R = -0.462$; $p < 0.001$), as illustrated in [Figure 30.b](#). These results indicate that resistivity-related parameters are strongly and positively associated with E , whereas radioactive and porosity-related logs tend to exhibit strong negative correlations.

In the QH-Class 2 reservoir, statistical correlation analysis showed that the dynamic Young's modulus (E) has significant positive correlations with RT01 ($R = 0.373$; $p < 0.001$), RT02 ($R = 0.467$; $p < 0.001$), RT03 ($R = 0.507$; $p < 0.001$), RT06 ($R = 0.515$; $p < 0.001$), RT09 ($R = 0.513$; $p < 0.001$), and RT12 ($R = 0.512$; $p < 0.001$). Conversely, E displayed significant negative correlations with DEPT ($R = -0.331$; $p < 0.001$), GR ($R = -0.462$; $p < 0.001$), NPHI ($R = -0.670$; $p < 0.001$), POTA ($R = -0.498$; $p < 0.001$), THOR ($R = -0.427$; $p < 0.001$), and URAN ($R = -0.370$; $p < 0.001$). No significant correlations were observed with CALI ($R = 0.114$; $p = 0.127$) and PE ($R = -0.062$; $p = 0.405$). These results suggest that resistivity-related logs are positively associated with E , while radioactive and porosity-related parameters tend to be negatively correlated. As illustrated in [Figure 30.c](#).

4.2.1.3. Developing and Assessing Machine Learning Models

To evaluate whether certain logging variables could be excluded without compromising the accuracy of dynamic Young's modulus (E) predictions, a feature selection analysis was conducted, following the approach of ([Bermingham et al., 2015](#)). Through a series of trial-and-error experiments, multiple combinations of input variables were tested. The most effective combinations, as shown in [Table 9](#), were selected for detailed analysis to identify the optimal set of features for predicting dynamic Young's modulus. Each machine learning model was trained and tested on six distinct sets of logging parameters, providing a systematic and transparent framework to generate a continuous profile of E and to assess the relevance of each input.

Table 8. Input Data Scenarios from Well Logging Variables for Dynamic Young's Modulus Prediction.

Combination Reference	Input Variables	Output
1	Dept, NPHI, POTA, GR, PE, Cali, URAN, RT10, THOR, RT20, RT30, RT60, RT90	E
2	Dept, NPHI, POTA, GR, PE, Cali, URAN, RT10, THOR, RT20, RT30	E
3	Dept, NPHI, POTA, GR, PE, Cali, URAN, RT10, THOR	E
4	Dept, NPHI, POTA, GR, PE, Cali, URAN	E
5	Dept, NPHI, POTA, GR, PE	E
6	Dept, NPHI, POTA	E

The table above [Table 9](#). illustrates a structured reduction in the number of input variables, starting with a comprehensive set in Combination 1 (scenario 1) and narrowing down progressively to just three variables in Combination 6 (scenario 6). The initial configuration includes a full suite of petrophysical, elemental, and resistivity-based measurements such as Dept, NPHI, GR, PE, Cali, POTA, URAN, THOR, and RT10 to RT90. As the combinations progress, the resistivity and elemental inputs are gradually removed. The final combination retains only Dept, NPHI, and POTA, representing a minimal yet potentially informative feature set. This systematic reduction enables clear observation of how the exclusion of certain parameters impacts the model's performance, revealing the relative importance of each variable in predicting dynamic Young's modulus.

[Figure 31](#). illustrates the workflow adopted for the machine learning modeling process. It begins with the collection of logging data, and dynamic Young's modulus derived from shear and compressional acoustic slowness logs, The process then proceeds through data classification into three reservoir types based on wellbore diameter, followed by data processing, model development, and performance evaluation

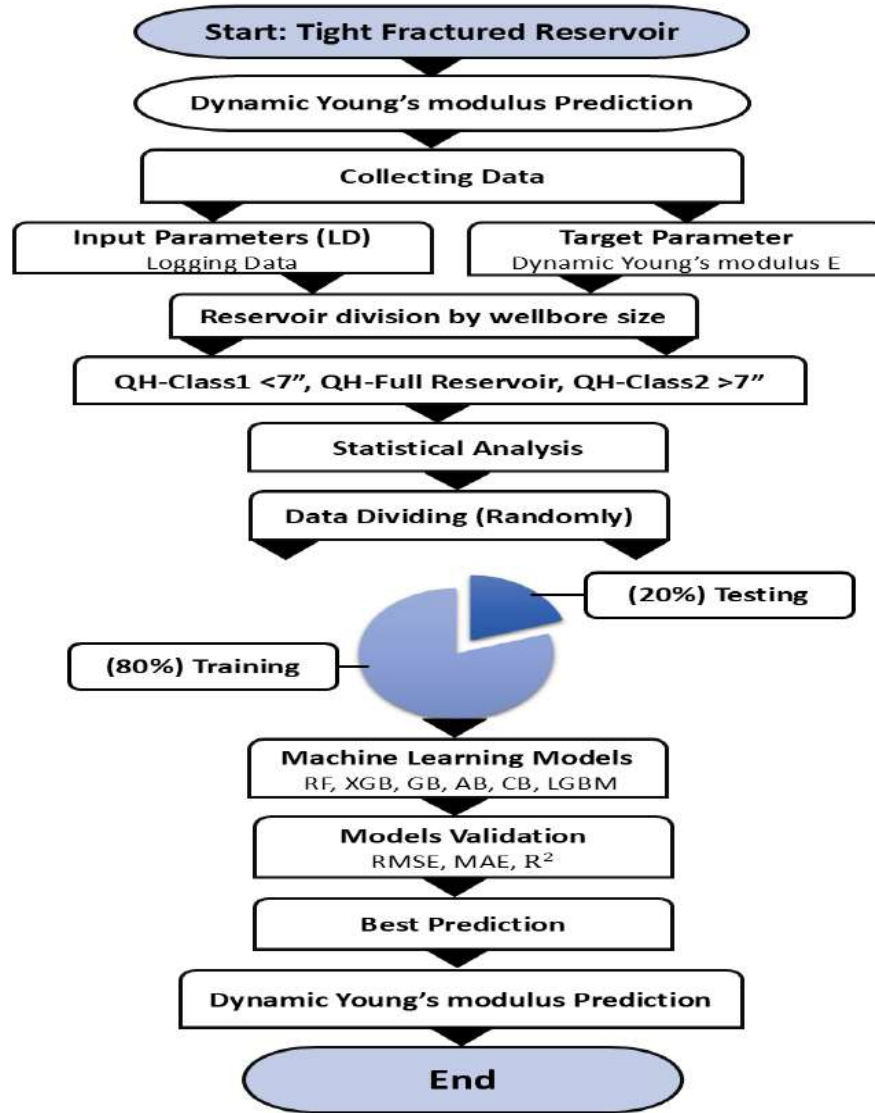


Figure 31. Flowchart illustrating the ML model architecture developed to predict dynamic Young's modulus using logging data

4.2.1.4. Implementation of Machine Learning Techniques

To predict dynamic Young's modulus, six machine learning (ML) algorithms were implemented using models from the scikit-learn Python library (Pedregosa et al., 2011; Varoquaux et al., 2015). These algorithms Random Forest (RF), Gradient Boosting (GB), Extreme Gradient Boosting (XGB), AdaBoost (AB), CatBoost (CB), and LightGBM (LGBM) were trained within a supervised learning framework using a subset of the available data. To assess and compare model performance, several evaluation metrics from the sklearn.metrics Python library (Buitinck et al., 2013) module were employed, including the coefficient of determination (R^2), root mean squared error (RMSE), and mean absolute error (MAE). These

metrics played a vital role in validating predictive accuracy and guiding hyperparameter optimization.

a. Hyperparameters Tuning

Tuning hyperparameters plays a crucial role in effectively training machine learning models. In this study, optimal hyperparameter configurations for six algorithms Random Forest (RF), Gradient Boosting (GB), Extreme Gradient Boosting (XGB), AdaBoost (AB), CatBoost (CB), and LightGBM (LGBM) were identified through a structured and systematic approach. This process relied exclusively on the training dataset to prevent data leakage, thereby ensuring that test results accurately reflect each model's ability to generalize to new, unseen data. The selected hyperparameter values were chosen to strike a balance between maximizing predictive performance and minimizing overfitting, promoting robust and reliable generalization across various data scenarios.

The optimal values of set of hyperparameters were determined using a grid search strategy and are summarized in [Table .10 \(in the Appendix\)](#)

The RF model was configured with 100 decision trees (`n_estimators = 100`), which is sufficient to ensure model stability without overfitting. The maximum tree depth (`max_depth`) was limited to 10, preventing overly complex trees. The number of features considered when looking for the best split (`max_features`) was set to 1.0, meaning all features were used. The minimum number of samples required to split an internal node (`min_samples_split`) and to be at a leaf node (`min_samples_leaf`) were set to 2 and 1, respectively. A `random_state` of 42 was used for reproducibility of results. ([Pedregosa et al., 2011](#)).

XGBoost was tuned to use 100 trees (`n_estimators = 100`) with a relatively shallow depth (`max_depth = 5`) to control overfitting. A learning rate (`learning_rate = 0.1`) was applied to scale the contribution of each tree. The subsample parameter was set to 1.0, meaning all training data were used in each boosting round. A `random_state` of 42 was used for consistent results, and verbosity was set to 0 to suppress output messages during training. ([Chen and Guestrin, 2016](#))

The GB model shared several hyperparameters with XGB. It used 100 trees (`n_estimators = 100`), and a learning rate of 0.1 was adopted. The maximum depth of trees was slightly lower (`max_depth = 3`), helping to keep the model simpler. The minimum samples required to split (`min_samples_split = 2`) and for leaf nodes (`min_samples_leaf = 1`) were the same as in RF. The subsample was set to 1.0, and `random_state = 42` ensured replicability. ([Pedregosa et al., 2011](#)).

The AdaBoost model used 100 estimators ($n_estimators = 100$) and a learning rate of 1.0, which is the default value in many implementations and controls the weight of each estimator. While no depth parameter is explicitly set for AB (as it often uses decision stumps), $random_state = 42$ was included for reproducibility. (Freund and Schapire, 1997; Pedregosa et al., 2011).

CatBoost was configured with 100 trees ($n_estimators = 100$) and a moderate learning rate of 0.1. The max_depth was set to -1, meaning the depth was not restricted and could be optimized automatically by the algorithm. $Colsample_bytree$ was set to 1.0, indicating that all features were considered when building each tree. The model used a $random_state$ of 42 for consistency (Prokhorenkova et al., 2018).

LGBM used 100 boosting rounds ($n_estimators = 100$) with a learning rate of 0.1. It included several key hyperparameters for regularization: $num_leaves = 31$, which controls the complexity of each tree; $min_child_samples = 20$ to avoid overfitting small datasets; and $min_child_weight = 0.001$, a parameter that controls minimum sum of instance weight needed in a child. The subsample rate was set to 1.0 to use the full dataset in each iteration, and $random_state = 42$ was applied (Ke et al., 2017).

This thorough hyperparameter tuning ensured that each algorithm was optimized for the dataset in question, improving both the predictive accuracy and the generalizability of the models.

4.2.3. Results

Table 11 and Figure 32 summarize the prediction performance of six machine learning models XGBoost (XGB), Random Forest (RF), Gradient Boosting (GB), AdaBoost (AB), CatBoost (CB), and LightGBM (LGBM) for estimating the dynamic Young's modulus in the QH-Class 1 Reservoir. Among all models and scenarios, XGB1 stands out as the best performer. It yields the best match with the measured dynamic Young's modulus in terms of R^2 , achieving 0.9638 on the testing set, along with the lowest RMSE (2.2813) and MAE (1.3551), indicating strong accuracy and generalization. The other XGB configurations (XGB2 and XGB3) also demonstrated robust results, with R^2 values of 0.9621 and 0.9596, respectively, maintaining low RMSEs (2.335 and 2.4116) and MAEs (1.3674 and 1.3872).

The RF models ranked second in performance, with RF1 delivering a testing R^2 of 0.9474, RMSE of 2.7508, and MAE of 1.5442, closely following the XGB models. RF2 and RF3 maintained similar prediction quality, confirming RF's reliability as an alternative.

In contrast, the GB models showed relatively weaker performance. Even the best GB scenario, GB2, only reached a testing R^2 of 0.9086, with higher RMSE (3.6268) and MAE (2.4314). The worst performance among them was seen in GB6, which yielded a lower R^2 (0.8566) and the highest errors among GB models (RMSE $\approx$ 4.5423, MAE $\approx$ 2.9520). These results indicate GB's limited ability to accurately capture the nonlinear relationships in the data for this task.

The AdaBoost (AB) models delivered the least satisfactory results overall. Their testing R^2 values ranged from 0.8209 (AB1) down to 0.7611 (AB6), with considerably high RMSEs (5.0752–5.8617) and MAEs (4.0674–4.5890), indicating poor fitting and high prediction errors.

Mean while, CatBoost (CB) models showed moderate performance. CB1 provided a decent testing R^2 of 0.9486, which represents the best among all CB scenarios, accompanied by an RMSE of 2.7196 and MAE of 1.7894. However, later scenarios like CB6 exhibited reduced accuracy ($R^2 \approx$ 0.9151, RMSE $\approx$ 3.4949, MAE $\approx$ 2.4317), reflecting some instability in performance across scenarios.

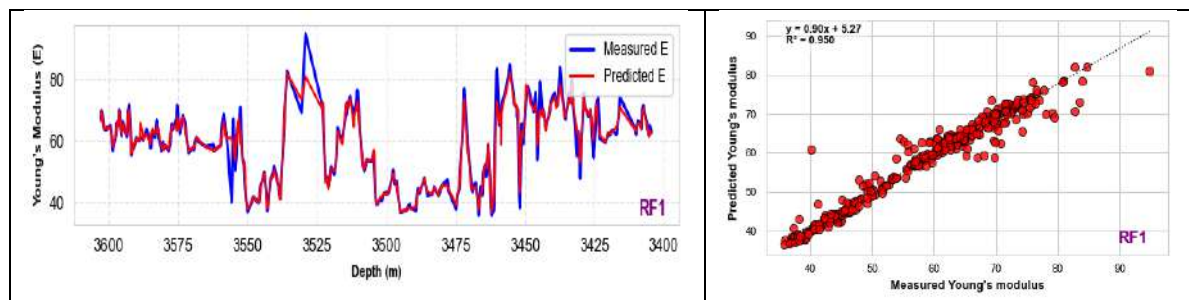
LightGBM (LGBM) models followed a similar trend to CB, with LGBM1 and LGBM2 showing strong results. LGBM1 achieved a testing R^2 of 0.9519, RMSE of 2.6307, and MAE of 1.6922, which places it among the top performing models just behind XGB. Later LGBM scenarios, such as LGBM6, displayed lower performance ($R^2 \approx$ 0.9290, RMSE $\approx$ 3.1959, MAE $\approx$ 2.1301), yet still performed better than GB and AB models in most cases.

These results clearly highlight the superiority of the XGBoost algorithm, particularly the XGB1 model, which yields the best match with the measured dynamic Young's modulus in terms of R^2 , while also delivering the lowest prediction errors (RMSE and MAE). The RF and LGBM models also demonstrate strong performance, making them competitive options for this regression task. In contrast, the GB and especially the AB models show more limited accuracy, suggesting they are less suitable for reliably predicting the dynamic Young's modulus in this context.

Table 9. Dynamic Young's modulus E prediction performances of the six ML for the QH-Class 1 Reservoir.

ML MODELS	Training Set			Testing Set		
	R^2	RMSE	MAE	R^2	RMSE	MAE
XGB1	0.9996	0.2403	0.1688	0.9638	2.2813	1.3551
XGB2	0.9995	0.2558	0.1822	0.9621	2.335	1.3674
XGB3	0.9996	0.225	0.1595	0.9596	2.4116	1.3872
XGB4	0.9993	0.3129	0.2185	0.9395	2.9498	1.7623
XGB5	0.9986	0.444	0.308	0.9412	2.9076	1.8092
XGB6	0.9959	0.7626	0.4823	0.9329	3.106	1.9449

RF1	0.9924	1.0404	0.5767	0.9474	2.7508	1.5442
RF2	0.9923	1.0468	0.5729	0.9458	2.7909	1.5387
RF3	0.9925	1.0331	0.5646	0.947	2.7621	1.5479
RF4	0.9908	1.1464	0.6654	0.9399	2.9413	1.713
RF5	0.9904	1.1714	0.6826	0.9429	2.8652	1.7037
RF6	0.9908	1.1494	0.6585	0.9429	2.8661	1.6855
GB1	0.9506	2.6581	1.9118	0.9063	3.6707	2.4785
GB2	0.9509	2.6500	1.8991	0.9086	3.6268	2.4314
GB3	0.9473	2.7462	1.9437	0.9016	3.7631	2.5097
GB4	0.929	3.1870	2.3186	0.8702	4.3216	2.9244
GB5	0.9255	3.2638	2.3183	0.8798	4.1577	2.8012
GB6	0.9039	3.7067	2.5169	0.8566	4.5423	2.952
AB1	0.8534	4.5786	3.7568	0.8209	5.0752	4.0674
AB2	0.8622	4.4388	3.5197	0.8308	4.9338	3.7525
AB3	0.8509	4.618	3.7277	0.8113	5.2099	4.1068
AB4	0.8139	5.1579	4.1662	0.7833	5.5837	4.3566
AB5	0.8102	5.2097	4.1249	0.7775	5.6572	4.2901
AB6	0.7894	5.4879	4.3974	0.7611	5.8617	4.589
CB1	0.9883	1.2929	0.9942	0.9486	2.7196	1.7894
CB2	0.9879	1.3127	0.9881	0.9457	2.7938	1.8413
CB3	0.9873	1.3481	1.0132	0.9461	2.7836	1.8541
CB4	0.9803	1.679	1.2715	0.9268	3.2454	2.2164
CB5	0.9747	1.9007	1.4076	0.9296	3.1826	2.2351
CB6	0.963	2.3012	1.7271	0.9151	3.4949	2.4317
LGBM1	0.9926	1.0257	0.7046	0.9519	2.6307	1.6922
LGBM2	0.9919	1.0792	0.7339	0.9519	2.6314	1.6691
LGBM3	0.9919	1.0789	0.7314	0.9486	2.7186	1.7034
LGBM4	0.9877	1.3254	0.9179	0.9294	3.1876	2.0601
LGBM5	0.9793	1.7221	1.1798	0.9272	3.2368	2.157
LGBM6	0.9667	2.1809	1.4917	0.929	3.1959	2.1301



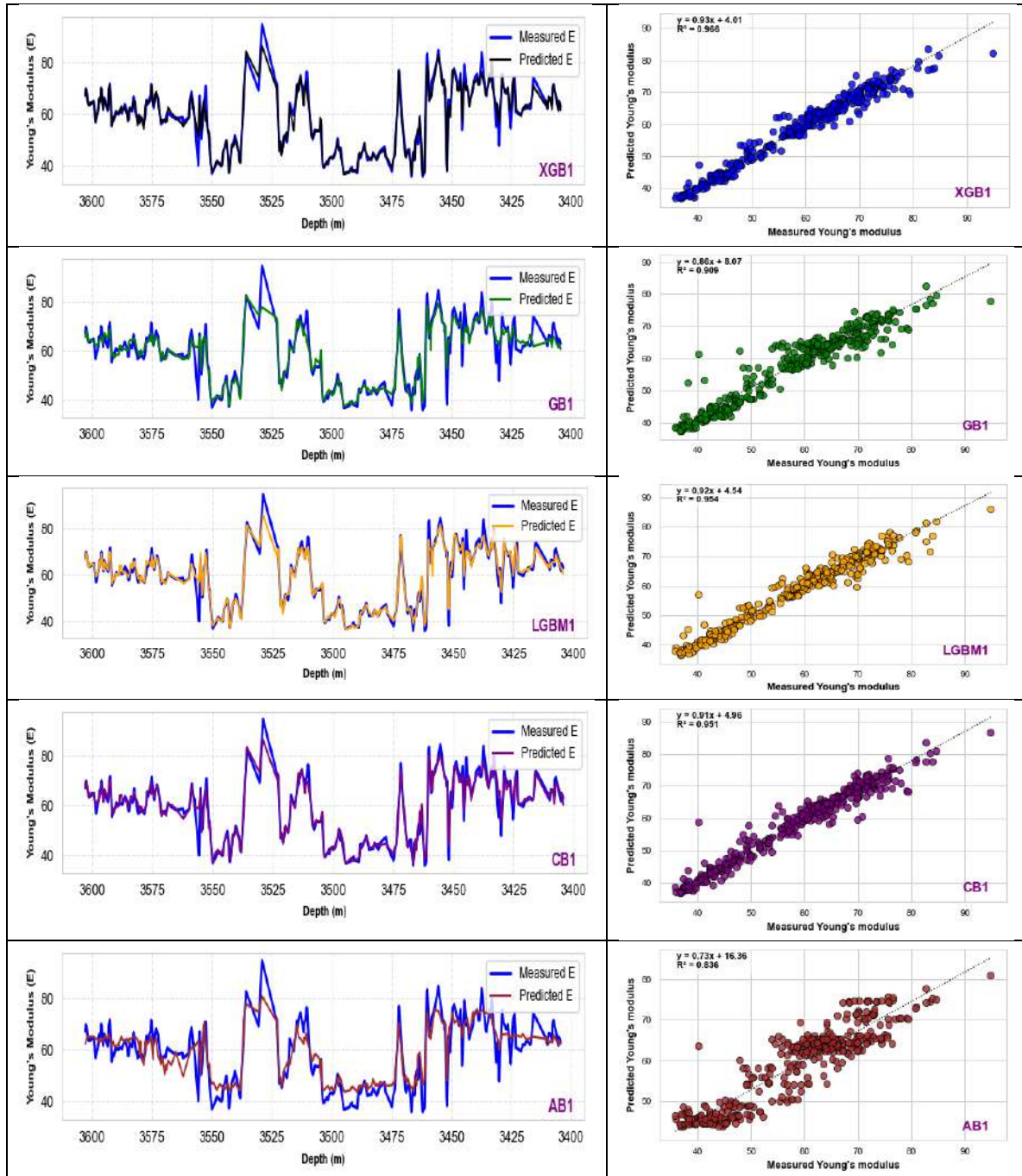


Figure 32. Measured Young's modulus versus predicted Young's modulus for the best-performing ML for the testing subset for the QH-Class 1 Reservoir.

The evaluation of the six machine learning algorithms for predicting dynamic Young's modulus on the QH-Full Table 12 and Figure 33 reservoir test dataset reveals clear differences in their predictive performance. Among these, the XGBoost (XGB) models consistently demonstrate the strongest correspondence with the measured values. Specifically, XGB1 attains the highest R^2 of 0.9529, with RMSE and MAE of 2.6546 and 1.5392, respectively, representing the best overall fit. Other XGB variants (XGB2 to XGB6) also show excellent predictive power,

with R^2 values gradually declining from 0.9507 to 0.9310, and corresponding increases in RMSE and MAE, indicating a slight decrease in precision but still maintaining superior accuracy relative to other methods.

Random Forest (RF) models follow closely behind, offering a solid match to the observed dynamic Young's modulus. Their R^2 values range from 0.9449 (RF1) to 0.9301 (RF6), paired with RMSE values between 2.8726 and 3.2343 and MAE values from 1.6884 to 1.8382. This performance suggests RF models reliably replicate the dynamic modulus trends, though with marginally higher errors compared to the XGB models.

CatBoost (CB) models also provide strong predictive capabilities, with R^2 values spanning from 0.9474 down to 0.9183 across the six variants. Their RMSE ranges from 2.8073 to 3.4985, and MAE from 1.9106 to 2.3778, indicating a robust yet slightly less precise fit than XGB and RF. CatBoost thus proves to be a competitive algorithm for this task, offering a good balance of accuracy and stability.

LightGBM (LGBM) models exhibit consistent predictive power as well, with R^2 values decreasing from 0.9445 to 0.9201, RMSE increasing from 2.884 to 3.4589, and MAE rising from 1.8695 to 2.3021 across the model variations. These metrics illustrate that LGBM models maintain strong correlations with the dynamic Young's modulus but do not surpass the leading XGB models.

Gradient Boosting (GB) models show moderate predictive strength, with R^2 values falling between 0.9066 and 0.8876, RMSE values increasing from 3.7407 to 4.1032, and MAE from 2.6437 to 2.767. While still producing reasonable predictions, their fit to the measured data is clearly less accurate than the previously mentioned ensemble methods.

Lastly, AdaBoost (AB) models provide the weakest match to the dynamic Young's modulus data. R^2 values for AB range from 0.8145 down to 0.7793, while RMSE and MAE reach the highest observed errors, up to 5.7495 and 4.5309 respectively. This indicates that AdaBoost is less suited for accurately modeling the dynamic Young's modulus in this reservoir compared to the other tested algorithms.

Overall, the XGB models, particularly XGB1, generate the best match with the measured dynamic Young's modulus in terms of R^2 and error metrics. RF and CB models also show strong performance and could serve as reliable alternatives. LGBM offers consistent, though slightly lower accuracy, while GB and AB models display progressively reduced predictive effectiveness for this application.

Table 10. Dynamic Young's modulus E prediction performances of the six ML for the QH-Full Reservoir.

ML MODELS	Training Set			Testing Set		
	R ²	RMSE	MAE	R ²	RMSE	MAE
XGB1	0.9996	0.2238	0.1629	0.9529	2.6546	1.5392
XGB2	0.9995	0.2591	0.1848	0.9507	2.7172	1.5499
XGB3	0.9994	0.2937	0.2063	0.946	2.8431	1.5719
XGB4	0.9992	0.3414	0.246	0.9423	2.9407	1.7568
XGB5	0.9987	0.4252	0.3019	0.9367	3.0785	1.7876
XGB6	0.9962	0.7231	0.4968	0.931	3.214	1.8846
RF1	0.9914	1.0953	0.6236	0.9449	2.8726	1.7395
RF2	0.992	1.0523	0.6017	0.9458	2.8494	1.6884
RF3	0.9919	1.0621	0.5982	0.9448	2.8755	1.6877
RF4	0.99	1.1787	0.6976	0.9349	3.1221	1.8293
RF5	0.9903	1.1604	0.691	0.934	3.1441	1.8179
RF6	0.9908	1.1274	0.6689	0.9301	3.2343	1.8382
GB1	0.9417	2.8442	2.0019	0.9066	3.7407	2.6437
GB2	0.9364	2.9717	2.1101	0.9037	3.7974	2.6779
GB3	0.939	2.9097	2.0673	0.9052	3.7673	2.6377
GB4	0.9217	3.2969	2.4113	0.8853	4.1438	2.8856
GB5	0.9199	3.334	2.406	0.889	4.0763	2.7812
GB6	0.9049	3.6333	2.5319	0.8876	4.1032	2.767
AB1	0.8529	4.5196	3.6764	0.8145	5.2701	4.2341
AB2	0.8482	4.5902	3.7421	0.814	5.2774	4.2126
AB3	0.843	4.6682	3.7214	0.8027	5.4359	4.3063
AB4	0.8143	5.0771	4.066	0.7961	5.5262	4.2945
AB5	0.8104	5.1305	3.9366	0.7923	5.5772	4.1768
AB6	0.7908	5.3886	4.2594	0.7793	5.7495	4.5309
CB1	0.9872	1.3319	1.0075	0.9474	2.8073	1.9106
CB2	0.9863	1.3805	1.0553	0.9467	2.8255	1.9011
CB3	0.9855	1.4176	1.0717	0.9466	2.8284	1.9917
CB4	0.9761	1.82	1.3718	0.9249	3.3535	2.2024
CB5	0.9704	2.0267	1.5076	0.9183	3.4972	2.3459
CB6	0.9628	2.2712	1.6986	0.9183	3.4985	2.3778
LGBM1	0.9922	1.0392	0.714	0.9445	2.884	1.8695
LGBM2	0.9917	1.0758	0.7398	0.9423	2.9384	1.8747
LGBM3	0.9908	1.1277	0.7656	0.942	2.9478	1.9304
LGBM4	0.9864	1.3723	0.9667	0.9323	3.1831	2.0227
LGBM5	0.9789	1.7132	1.207	0.9296	3.247	2.0606
LGBM6	0.9673	2.1306	1.5002	0.9201	3.4589	2.3021

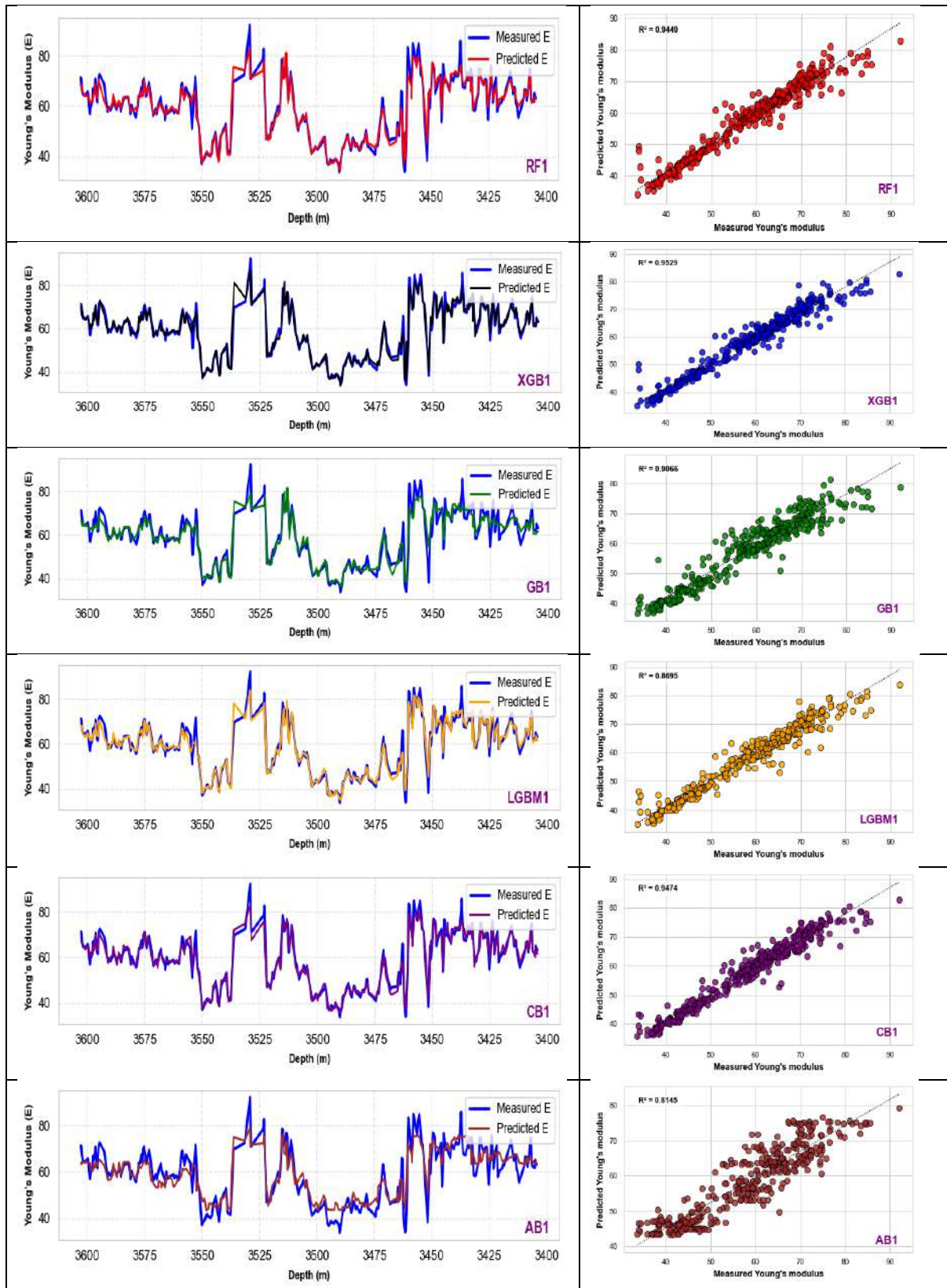


Figure 33. Measured Young's modulus versus predicted Young's modulus for the best-performing ML for the testing subset for the QH-Full Reservoir

In the prediction of the dynamic Young's modulus for the QH-Class 2 Reservoir [Table 13](#) and [Figure 34](#), XGBoost (XGB) models demonstrate strong performance across all six scenarios. Among them, XGB6 yields the best match with the measured values, achieving $R^2 \approx 0.9326$, $RMSE \approx 1.3766$, and $MAE \approx 0.9595$. This is closely followed by XGB5 ($R^2 \approx 0.9053$) and XGB1 ($R^2 \approx 0.8922$), both providing reliable generalization. XGB2 also offers good predictive ability ($R^2 \approx 0.8888$), while XGB3 slightly underperforms with $R^2 \approx 0.8595$. The weakest scenario, XGB4, shows a lower fit ($R^2 \approx 0.7659$) and the highest RMSE and MAE among XGB models (2.5653 and 1.7704, respectively), indicating reduced robustness in that case.

The Random Forest (RF) algorithm shows consistent and strong performance, particularly with RF6, which delivers the best match ($R^2 \approx 0.9387$, $RMSE \approx 1.3122$, $MAE \approx 0.9114$) among all RF scenarios. RF1 and RF5 also yield high R^2 values of 0.8998 and 0.9006, respectively, with acceptable error levels. RF2 and RF3 exhibit moderate performance ($R^2 \approx 0.8756$ and 0.8658), while RF4 presents the lowest R^2 of 0.8157, accompanied by a relatively high RMSE (2.2757), reflecting weaker generalization in that scenario.

As for the Gradient Boosting (GB) models, GB6 offers the best match with the dynamic modulus ($R^2 \approx 0.9317$, $RMSE \approx 1.3859$, $MAE \approx 1.0184$), slightly outperforming GB5 ($R^2 \approx 0.9197$) and GB1 ($R^2 \approx 0.9152$). GB2 and GB3 also yield solid results with R^2 values of 0.9134 and 0.9112, respectively. The lowest-performing scenario is GB4, with $R^2 \approx 0.8508$, higher RMSE (2.0477), and MAE (1.4345), indicating decreased predictive accuracy compared to the other GB configurations.

On the other hand, AdaBoost (AB) shows the weakest performance overall. Its best result is observed in AB1, which yields a match of $R^2 \approx 0.7908$, $RMSE \approx 2.425$, and $MAE \approx 2.1601$. AB2 and AB3 show similar behavior ($R^2 \approx 0.77$ – 0.79), while the lowest performance appears in AB6, which records $R^2 \approx 0.7242$, the highest RMSE (2.7841), and MAE (2.3121), indicating substantial prediction error across this algorithm.

The CatBoost (CB) algorithm proves to be the most robust overall, with CB5 producing the best match among all models tested: $R^2 \approx 0.9538$, $RMSE \approx 1.139$, and $MAE \approx 0.914$. CB6 also performs exceptionally well ($R^2 \approx 0.9509$), followed by CB2 ($R^2 \approx 0.9303$) and CB1 ($R^2 \approx 0.921$). Even the lowest performing CB scenario, CB4, still maintains a respectable $R^2 \approx 0.852$, confirming the algorithm's stability and reliability.

Finally, LightGBM (LGBM) models provide moderate prediction accuracy. The best match is achieved by LGBM1 ($R^2 \approx 0.8505$, $RMSE \approx 2.0496$, $MAE \approx 1.5107$), while LGBM2 and LGBM3 follow with R^2 values of 0.842 and 0.8163, respectively. The lowest predictive

performance appears in LGBM6 and LGBM5, with $R^2 \approx 0.7867$ and 0.7844 , respectively. LGBM4 also performs modestly ($R^2 \approx 0.7708$, $RMSE \approx 2.5381$), further confirming that LightGBM is less effective in replicating the measured dynamic Young's modulus compared to the boosting-based XGB and CB.

Overall, across the six machine learning models, CatBoost and XGBoost consistently yield the best matches with the measured dynamic Young's modulus, both in terms of R^2 and error metrics, while AdaBoost and LightGBM demonstrate weaker predictive capabilities, especially in the more challenging scenarios.

Table 11. Dynamic Young's modulus E prediction performances of the six ML for the QH-Class 2 Reservoir.

ML MODELS	Training Set			Testing Set		
	R^2	RMSE	MAE	R^2	RMSE	MAE
XGB1	1	0.001	0.0008	0.8922	1.7409	1.2813
XGB2	1	0.0008	0.0006	0.8888	1.7678	1.2925
XGB3	1	0.0008	0.0006	0.8595	1.9871	1.4356
XGB4	1	0.0012	0.0008	0.7659	2.5653	1.7704
XGB5	1	0.0011	0.0008	0.9053	1.6318	1.1697
XGB6	1	0.0039	0.0029	0.9326	1.3766	0.9595
RF1	0.9729	0.8879	0.6232	0.8998	1.6783	1.2571
RF2	0.972	0.9024	0.6269	0.8756	1.87	1.3902
RF3	0.9715	0.9109	0.6443	0.8658	1.9419	1.415
RF4	0.9655	1.0011	0.6923	0.8157	2.2757	1.6296
RF5	0.9665	0.9872	0.6937	0.9006	1.6715	1.3476
RF6	0.9743	0.8638	0.6324	0.9387	1.3122	0.9114
GB1	0.9934	0.4364	0.3383	0.9152	1.5441	1.0945
GB2	0.9927	0.4608	0.3474	0.9134	1.5603	1.0903
GB3	0.9913	0.5032	0.3763	0.9112	1.5795	1.1957
GB4	0.9889	0.569	0.4165	0.8508	2.0477	1.4345
GB5	0.9862	0.6345	0.4804	0.9197	1.5027	1.1451
GB6	0.9857	0.6448	0.4978	0.9317	1.3859	1.0184
AB1	0.8897	1.7905	1.5549	0.7908	2.425	2.1601
AB2	0.8807	1.8627	1.5574	0.7888	2.4365	2.0305
AB3	0.8505	2.085	1.7632	0.771	2.5368	2.2512
AB4	0.8729	1.9223	1.6791	0.7537	2.6309	2.2425
AB5	0.7903	2.4692	1.9554	0.7519	2.6404	2.2085
AB6	0.7771	2.5461	2.1186	0.7242	2.7841	2.3121
CB1	0.9965	0.3173	0.2589	0.921	1.4897	1.1242
CB2	0.9968	0.3049	0.2504	0.9303	1.3992	1.0427
CB3	0.9967	0.3115	0.2557	0.9077	1.6105	1.1731
CB4	0.9965	0.3194	0.2594	0.852	2.0396	1.4456
CB5	0.995	0.3828	0.3042	0.9538	1.139	0.914
CB6	0.9921	0.4801	0.3862	0.9509	1.1742	0.9328
LGBM1	0.9524	1.1761	0.8728	0.8505	2.0496	1.5107
LGBM2	0.9489	1.2188	0.894	0.842	2.1074	1.5458
LGBM3	0.9453	1.261	0.9131	0.8163	2.272	1.7158

LGBM4	0.9311	1.4158	1.0646	0.7708	2.5381	1.9484
LGBM5	0.8898	1.7902	1.3464	0.7844	2.4617	1.9704
LGBM6	0.8432	2.1354	1.635	0.7867	2.4484	1.9833

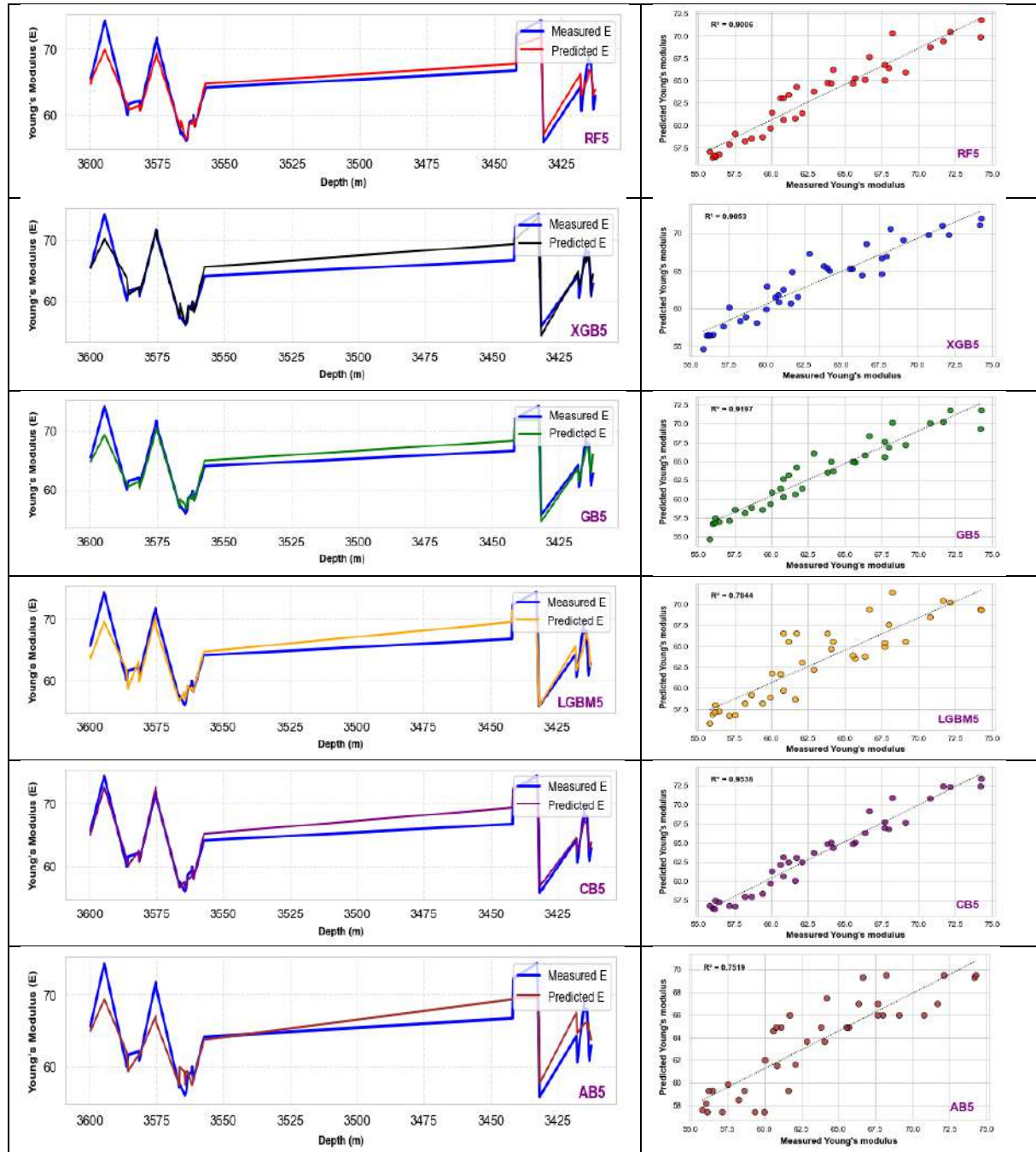


Figure 34. Measured Young's modulus versus predicted Young's modulus for the best-performing ML for the testing subset for the QH-Class 2 Reservoir.

4.2.4. Discussion

4.2.4.1. Evaluation of ML in predicting dynamic Young's modulus in tight fractured reservoir

The results of this study reveal notable differences in the performance of the applied machine learning algorithms, both when applied to the full dataset (QH-Full Reservoir) and when applied after reservoir classification based on CALI data (QH-Class 1 Reservoir), (QH-Class 2 Reservoir). Where it is clearly evident the superiority of the XGB algorithm over the other five algorithms RF, GB, AB, CB, and LGBM when applied to the two reservoirs (QH-Full) and (QH-Class 1), as it produced the best match with the measured dynamic Young's modulus (E) and the lowest errors XGB1; R^2 of (0.952, 0.963), RMSE of (2.654, 2.281) and MAE of (1.539, 1.355) respectively. It was followed by the LGBM algorithm, which generated matching values with the dynamic Young's modulus LGBM1; R^2 of (0.951, 0.944), RMSE (2.630, 2.884) and MAE (1.692, 1.869) respectively, when applied to the same reservoirs. followed by the algorithms CB, RF, GB, AB.

However, the performance of both algorithms declined when applied to the third reservoir (QH-Class 2) containing a data volume more than ten times smaller than that of the QH-Full reservoir and QH-Class 1, where the CB algorithm outperformed the others, achieving the best match with the measured dynamic Young's modulus and the lowest errors CB5; R^2 of (0.953), RMSE of (1.139) and MAE of (0.914) followed by GB, XGB, RF, LGBM, AB algorithms.

The AB algorithms showed the lowest accuracy in predicting the dynamic Young's modulus E across all three reservoirs (QH-Full, QH-Class 1, QH-Class 2) with R^2 (0.814, 0.820, 0.751) respectively, RMSE (5.270, 5.075, 2.640) respectively and MAE (4.234, 4.067, 2.208) respectively. Nevertheless, the R values greater than 0.70 between measured and predicted variables are generally considered acceptable for predictive modeling, as stated by [Legates and McCabe \(1999\)](#) and [Moriasi et al. \(2007\)](#). Based on this threshold, all six machine learning models demonstrated satisfactory performance, showing acceptable R values for both the training and validation datasets.

These differences can be attributed to several factors, primarily the algorithms' ability to handle complex geotechnical data and the benefits of classification in reducing internal data variability ([Zeng et al., 2024](#)). Unconventional reservoir data (tight fractured reservoirs) are typically characterized by high complexity, dispersion, and non-linearity, which pose significant challenges to accurate modeling especially when laboratory measurements and well

log data are unavailable specially acoustic data affected by the density of the drilling fluid used while drilling in fractured tight reservoirs (Zerrouki et al., 2014).

The six algorithms demonstrated varying behaviors when dealing with this data. XGBoost clearly outperformed the others on both the classified and full datasets, which can be attributed to its capacity for capturing nonlinear relationships between variables and its use of advanced optimization techniques that mitigate overfitting, such as regularized boosting and complexity control (Chen and Guestrin., 2016).

Random Forest also performed well, particularly when patterns were repeated in the dataset. However, the lack of fine-tuned optimization strategies in its structure makes it less effective than XGBoost in capturing complex interactions (Breiman, 2001).

In contrast, both Gradient Boosting and AdaBoost exhibited high sensitivity to internal noise in the data, which impacted their generalization ability especially when applied to the full especially the data of heterogeneous and complex reservoirs, such as fractured tight reservoirs with high pressure (Freund and Schapire., 1997).

CatBoost and LightGBM showed average performance, as they were somewhat affected by the distribution of variable values, which influenced the stability of their results across different scenarios a trend previously documented in similar geological applications (Prokhorenkova et al., 2018; Ke et al., 2017).

Table 12. Comparison of results with published studies designed to predict Young's modulus from drilling and logging variables inputs

Reference	R ²	ML model	Input Parameters	Reservoir Type
Seddig and Elkatatny., 2021	0.97	ANN, SVM	WOB, RPM, ROP, T, SPP, Q	Carbonate, sand and shale
Mahmoud et al., 2021	0.98	ANN	WOB, RPM, ROP, Flow	Carbonate, sand and shale
Mahmoud et al., 2022	0.99	SVR, RF, ANFIS-SC	WOB, T, SPP, ROP, RPM, Q	Carbonate, sand and shale
Elkatatny., 2022	0.98	ANN	WOB, T, SPP, ROP, RPM, Q	Carbonate, sand and shale
Sanei et al., 2023	0.99	A-MLP, GA-RBF	DT, LI, ROHB, X, Y, Z, MD, DTs, NPHI	Anisotropic reservoir
	0.96	XGB1	Dept, NPHI, POTA, GR, PE, Cali, URAN,	Tight fractured reservoir
This study	0.95	LGBM1	RT10, THOR, RT20, RT30, RT60, RT90	

As shown in Table 14 and Figure 35 reports a summary of some recent studies that discuss the prediction of Young's modulus utilizing drilling data or logging data to develop accurate machine learning models.

[Seddig and Elkatatny \(2021\)](#), [Mahmoud et al. \(2021, 2022\)](#), and [Elkatatny \(2022\)](#) relied on drilling variables as inputs. These studies employed a variety of algorithms, including Artificial Neural Networks (ANN), Support Vector Machines (SVM), hybrid models like ANFIS, and ensemble techniques such as Random Forests (RF) and Support Vector Regression (SVR). Despite the differences in modeling approaches and algorithms, all of these models achieved excellent results, with R^2 values ranging between 0.97 and 0.99, demonstrating the effectiveness of the ML algorithms to predicting Young's modulus, even across different reservoir types (carbonate, sandstone, and shale).

However, the methodology diverges in the study conducted by [Sanei et al. \(2023\)](#), which applied more advanced hybrid models (A-MLP, GA-RBF) and used both spatial and petrophysical inputs (such as DT, LI, ROHB, and NPHI) in an anisotropic reservoir setting. Their results also achieved a very high performance ($R^2 \approx 0.99$), emphasizing the role of complex spatial input in enhancing model precision.

Unlike the previous studies that primarily depended on drilling data, the present study follows an approach similar to that of [Sanei et al. \(2023\)](#), utilizing well log data including GR, NPHI, POTA, PE, CALI, URAN, THOR, and multiple deep resistivity measurements (RT10 to RT90) while deliberately excluding acoustic (DT) and density (RHOB) logs, which are commonly used for predicting dynamic Young's modulus. This methodology was applied to a tight fractured reservoir, characterized by complexity and heterogeneity, and low petrophysical properties. Despite these inherent challenges, the XGB1 and CB5 models demonstrated strong predictive performance, achieving R^2 values of 0.963 and 0.953 respectively.

The obtained results, specifically XGB when compared with the literature which used DD as inputs ([Seddig and Elkatatny., 2021](#); [Mahmoud et al., 2021](#); [Mahmoud et al., 2022](#)), and logging data ([Sanei et al., 2023](#)) appear closely aligned and accurate [Figure 33](#).

on the other hand, adopting a reservoir classification approach based on borehole diameter helped improve the performance of the algorithms a specially the XGB algorithm by reducing internal variability within each class, which allows algorithm to learn from more homogeneous data ([Zhao et al., 2022](#)).

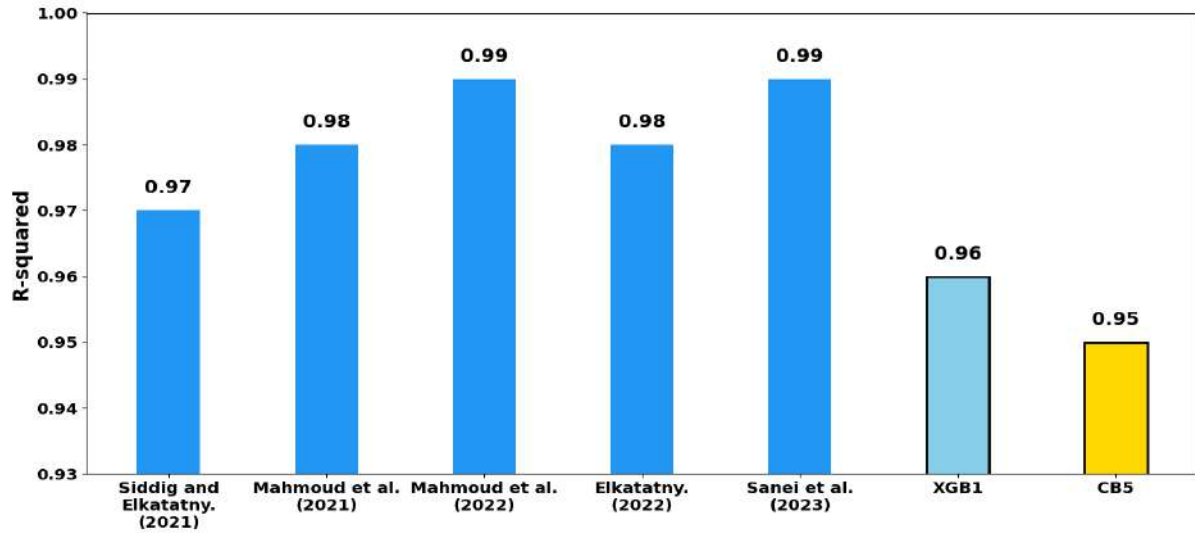


Figure 35. Comparative results (R^2) with published studies designed to predict Young's modulus

Figures 36, 37, and 38 include box plots, violin plots, radar plot and a Taylor diagram, comparing the distributions of dynamic Young's modulus predictions of the six ML models applied to each reservoir. The superior dynamic Young's modulus prediction performance of model XGB compared to the other ML models is evident from the Taylor diagram. This is clearly demonstrated by the notable increase in the performance metrics ($R^2 \approx 0.963$, $RMSE \approx 2.281$ and $MAE \approx 1.351$) in the QH-Class 1 reservoir compared to the QH-Full ($R^2 \approx 0.952$, $RMSE \approx 2.654$ and $MAE \approx 1.539$) reservoir. Likewise, it improves model generalization reducing noise caused by data from zones with contrasting characteristics, allowing the models to generate more stable and realistic relationships between the input features and the target variable (dynamic Young's modulus E) (Chang and Lippmann., 1990; Salih and Abdulazeez., 2021).

Additionally, a reservoir classification based on borehole diameter (CALI) provide an indirect representation of geomechanical properties and reflect zones of mechanical weakness, natural fractures, or changes in rock density, which means it serves as an important indicator of reservoir quality variation (Odling et al., 1999; Mohammadinia et al., 2023).

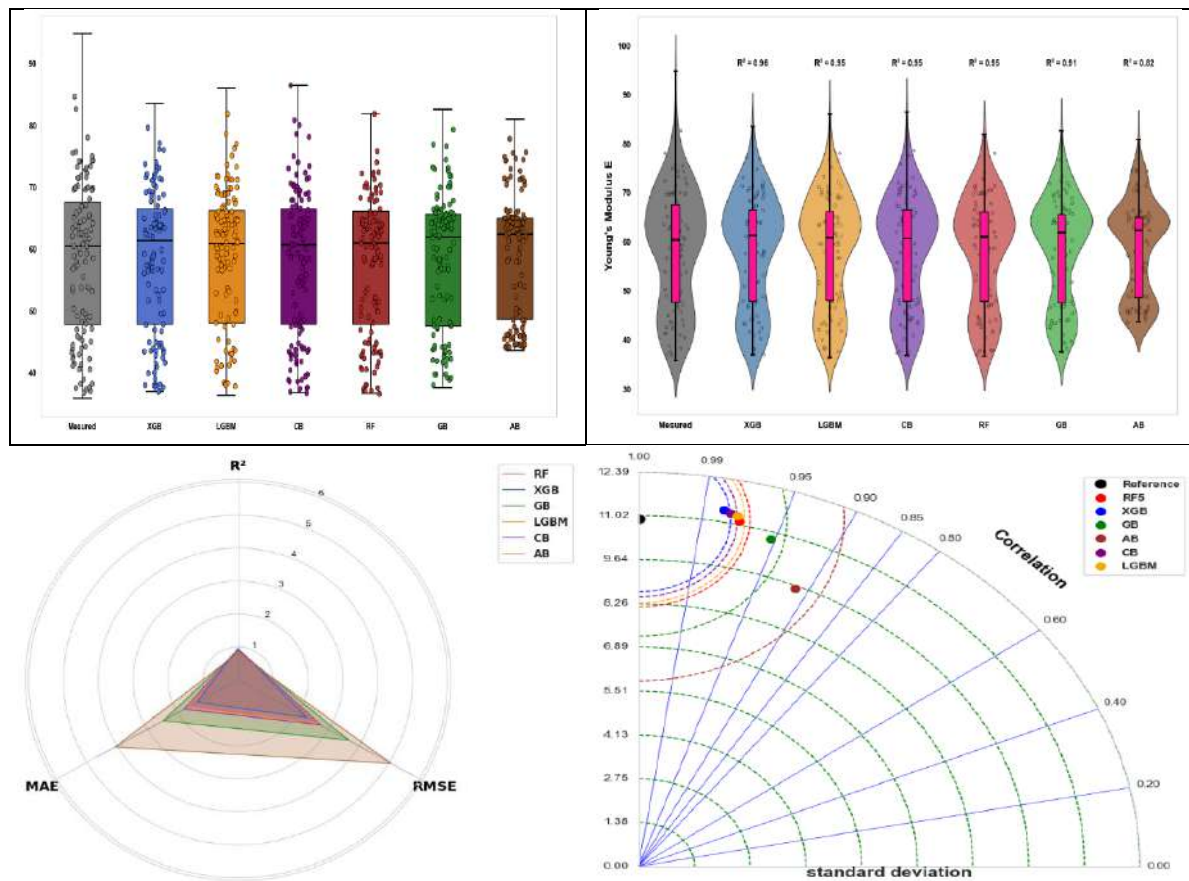
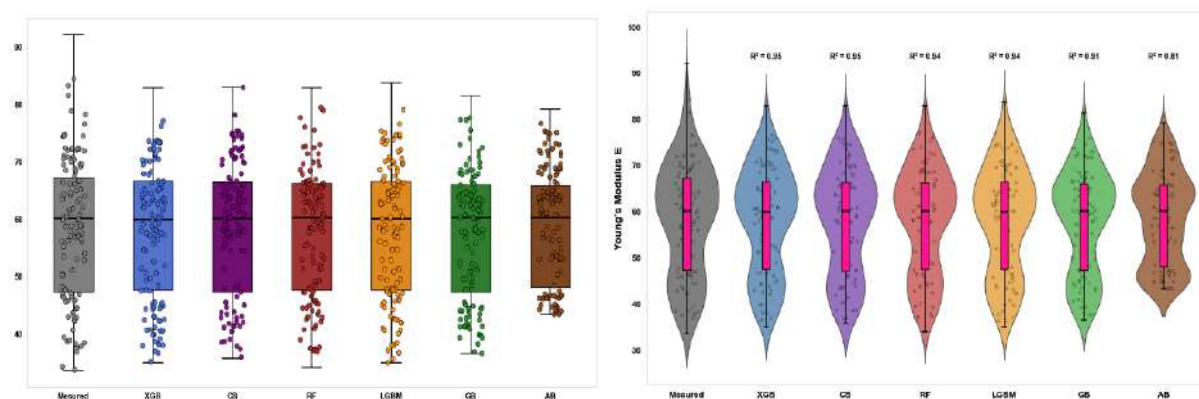


Figure 36. Comparative visualization of dynamic Young's modulus predictions from six machine learning models using box, violin, radar plots, and a Taylor diagram for the QH-Class 1 Reservoir.



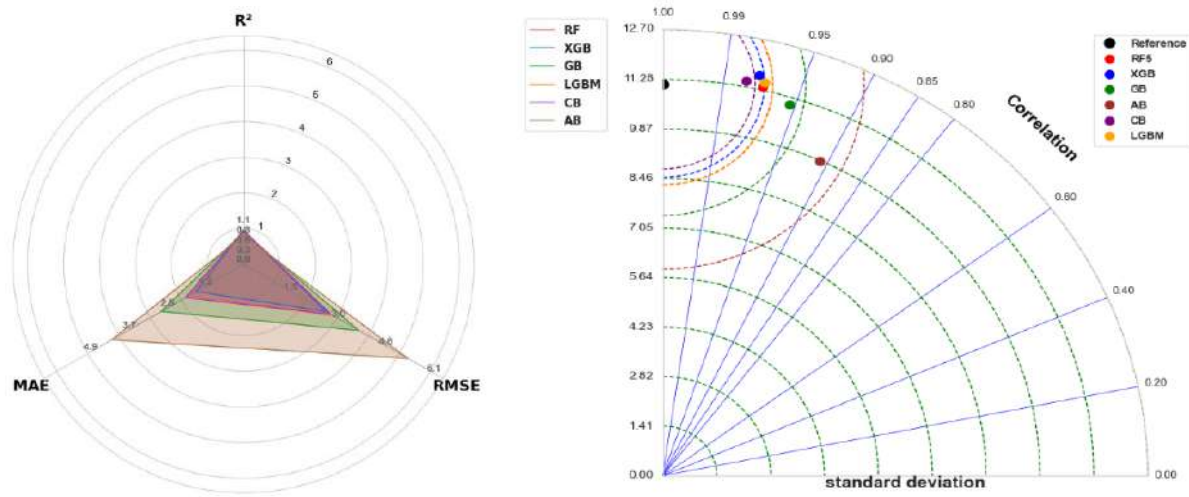
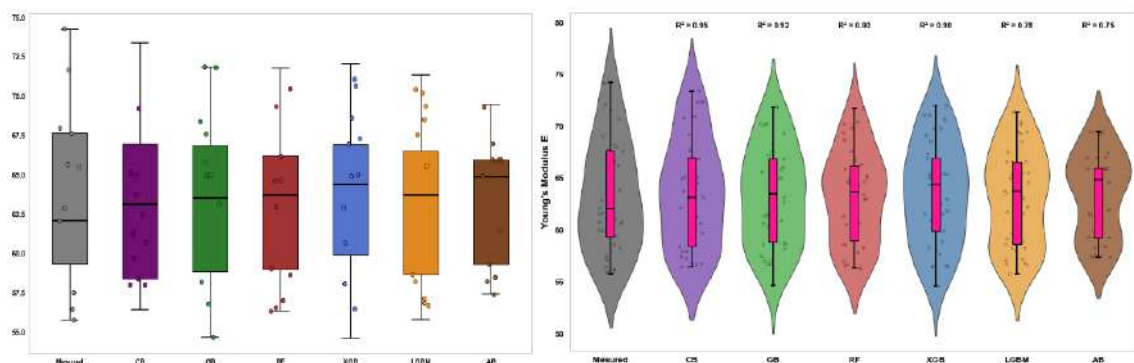


Figure 37. Comparative visualization of dynamic Young's modulus predictions from six machine learning models using box, violin, radar plots, and a Taylor diagram for the QH-Full Reservoir.

4.2.4.2. SHAP Interpretability analysis for the XGB for Class 1 Reservoir

The SHAP analysis [Figure 39](#) applied to the XGBoost model revealed that neutron porosity (NPHI) is the most influential factor in predicting the dynamic Young's modulus in tight fractured reservoir Quartzite El Hamra formation. From a geomechanical perspective, the inverse relationship between porosity and Young's modulus is well established, higher porosity implies a greater proportion of voids within the rock matrix, which reduces its overall stiffness and, consequently, lowers the Young's modulus ([Mavko et al., 2009](#); [Kahloul et al., 2019](#)).

The SHAP results indicate that high NPHI values have a negative impact on the output, meaning they contribute to a lower predicted Young's modulus. This aligns with the physical understanding of porous rock behavior. Conversely, [Harju et al. \(2023\)](#) emphasized that low NPHI values reflect denser and more compact rocks, which are mechanically stronger and associated with higher Young's modulus values.



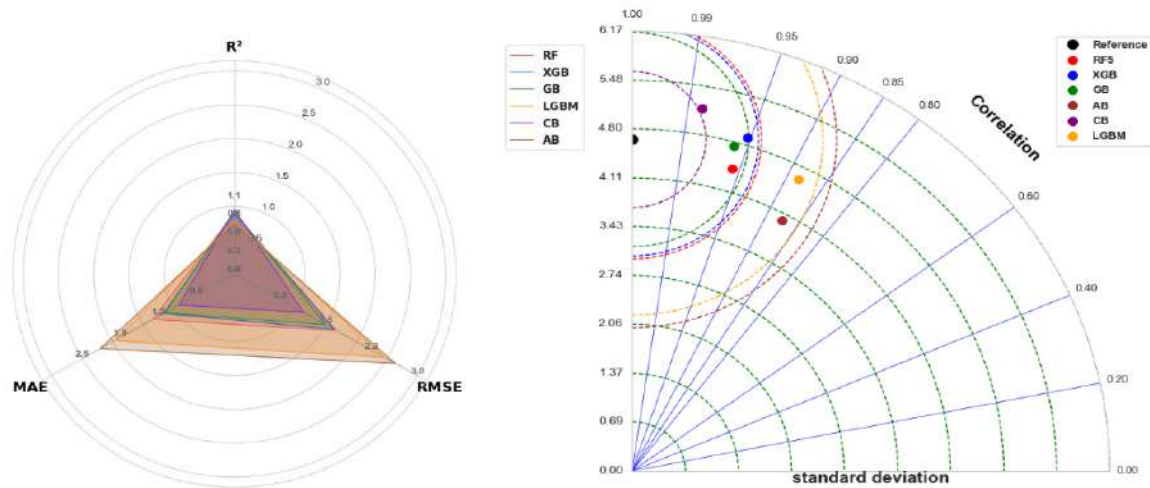


Figure 38. Comparative visualization of dynamic Young's modulus predictions from six machine learning models using box, violin, radar plots, and a Taylor diagram for the QH-Class 2 Reservoir.

These results are particularly logical in the context of fractured reservoirs, where porosity is sometimes linked to the distribution of natural fractures, which in turn affects the mechanical response of the formation. [Jiang and Yang \(2018\)](#) emphasized that the complex interactions between fracture networks and geomechanical properties particularly in fractured formations can lead to nonlinear mechanical responses. These include a noticeable reduction in rock stiffness due to the presence of discontinuities, natural fractures, and mechanically weak zones, which alter stress distribution and affect the elastic behavior of the reservoir.

Therefore, NPHI not only serves as a conventional porosity indicator but also reflects the geomechanical effects of rock structure and fracture distribution. This makes it a key variable for accurate modeling of properties such as Young's modulus in unconventional tight fractured reservoirs.

4.2.4.3. Proposed Avenues for Refining the Technique in Subsequent research

The approach of predicting dynamic Young's modulus using well log data has proven highly effective in complex reservoirs such as tight fractured reservoirs. Utilizing geophysical logging parameters like GR, NPHI, CALI, RT, URAN, THOR and POTAS without relying on acoustic logs, which can be compromised by high mud density resulting in signal loss or waveform distortion, enables enhancing prediction performance and accurate capture of dynamic Young's modulus along the wellbore.

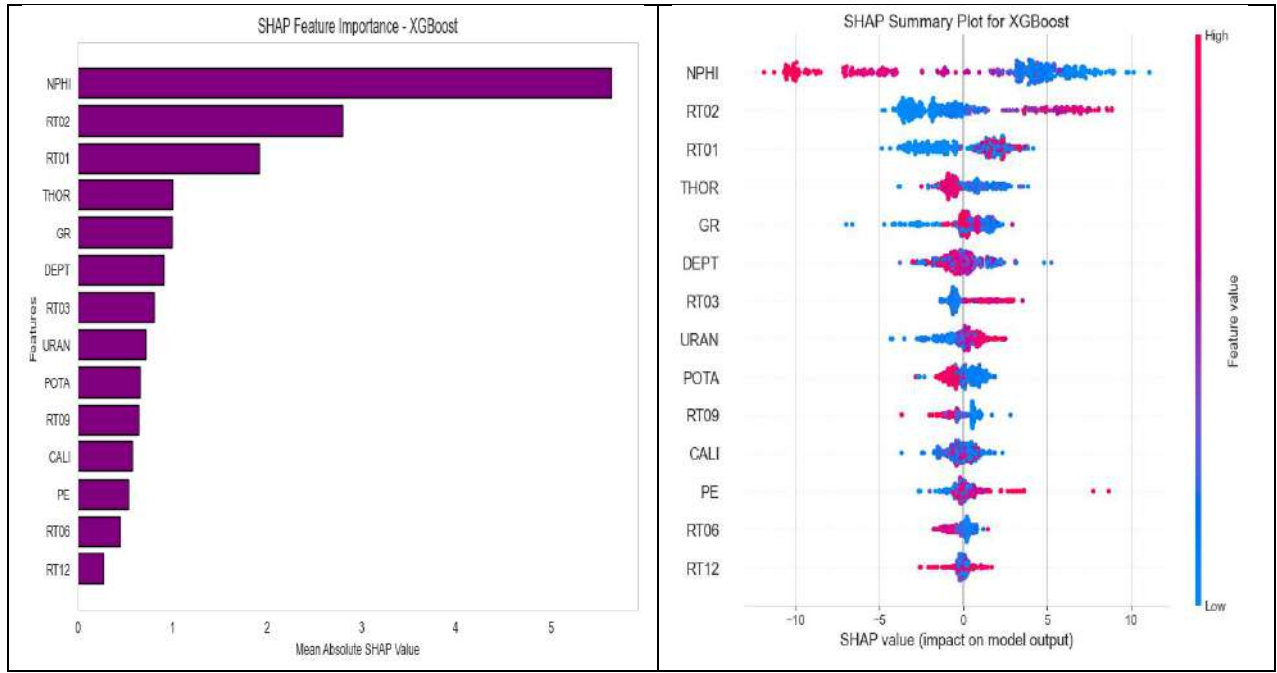


Figure 39. Model interpretability illustrated through the SHapley Additive exPlanations (SHAP) summary plot.

This method offers the key advantage of being applicable even when acoustic or density data are unavailable or of poor quality.

Moreover, integrating a reservoir classification based on borehole diameter (CALI) serves as a valuable strategy to improve model performance. CALI measurements act as indirect indicators of geomechanical weakness, natural fractures, or rock density variations. Classifying the reservoir based on this parameter reduces intra-class variability, allowing machine learning models to learn from more homogeneous subsets, thereby improving generalization and prediction stability.

To further strengthen this predictive approach, it is recommended to incorporate real-time drilling data such as RPM, T, SPP and WOB. These parameters provide a dynamic and continuous reflection of formation mechanical properties during drilling. Additionally, integrating real-time gas data (methane C1, ethane C2, total gas ...) would enhance the estimation of reservoir saturation conditions, which often correlate with rock stiffness.

Combining drilling data with real-time gas measurements would substantially improve the accuracy of real-time dynamic Young's modulus prediction. For even greater precision, additional inputs could be considered, including bulk density (RHOB) improves model accuracy and robustness especially when acoustic logs are compromised or absent, as is common in tight fractured reservoirs.

This methodology should be generalized to other reservoir types and extended to predict additional geomechanical parameters, such as shear modulus or compressive strength. This would allow for a more robust and comprehensive characterization of tight fractured reservoirs.

Conclusion

This study employed six machine learning algorithms to provide an advanced approach for continuous prediction of dynamic Young's modulus in a tight fractured reservoir Quartzite El Hamra formation. The study demonstrated the effectiveness of machine learning algorithms in predicting the dynamic Young's modulus (E_{dyn}) in the tight fractured reservoir. Among the six tested models (XGB, LGBM, CB, RF, GB, AB), the XGB algorithm showed the best performance on both the full and classified datasets, achieving the highest correlation values ($R^2 \approx 0.963$ for QH-Class 1) and the lowest errors ($RMSE \approx 2.281$, $MAE \approx 1.355$), confirming its ability to capture complex nonlinear relationships and efficiently handle data noise. LGBM showed similar but slightly lower performance, while CB, RF, GB, and AB exhibited variable results depending on data volume and homogeneity, with AB being the least accurate across all scenarios.

In this study, a novel reservoir classification approach based on borehole diameter (CALI) was also adopted. This classification significantly improved model performance by reducing internal data variability, allowing the algorithms to learn from more homogeneous subsets. This methodology is particularly suitable for tight and naturally fractured reservoirs such as Quartzite El Hamra, where acoustic measurements are often affected by high mud density during drilling or poor log quality. Using only well logging data (GR, NPHI, CALI, RT, URAN, THOR, POTA) without relying on acoustic or density logs highlights the importance of this approach, as it enables accurate and continuous prediction of formation mechanical properties even in the absence of conventional acoustic measurements.

In light of these findings, the proposed methodology represents a significant advancement in the geomechanical characterization of tight and naturally fractured reservoirs, as it provides a reliable alternative to laboratory measurements and acoustic logs, which are often unavailable or of poor quality in such reservoirs. Relying on geophysical well log data along the borehole enables continuous estimation of the dynamic Young's modulus, significantly reducing the time and costs associated with laboratory testing of rock samples.

Furthermore, this methodology contributes to improving well design and enhancing borehole stability, while providing effective support for decision-making during geomechanical evaluation and engineering planning stages. The results also demonstrate the potential for

extending this approach to predict additional geomechanical parameters and to apply it to different types of unconventional reservoirs.

Part III. Scientific and practical significance of the study

AI based geomechanical characterization studies currently represent a significant qualitative advancement in the field of petroleum engineering, particularly concerning unconventional reservoirs such as tight reservoirs and naturally fractured reservoirs. The implications of these studies extend across multiple levels, encompassing scientific, methodological, operational, economic, and technological aspects.

4.3.1. Scientific implications

4.3.1.1. Evaluation of machine learning geomechanical model accuracy

The results of this dissertation strongly support the use of real-time drilling data as a primary and effective approach for predicting the geomechanical properties of rocks, as an alternative to the exclusive reliance on conventional well logging data. This preference is justified by a set of operational and scientific advantages that make drilling data more suitable for modern engineering applications, particularly in the context of real time drilling operations. These data are continuously available during the drilling process in all wells, without the need for additional operations or specialized equipment. Moreover, they represent a cost-effective data source compared to well logs and laboratory experiments, which require longer acquisition times, advanced instrumentation, and higher operational costs (Ameur-Zaimeche et al. 2020). In addition, drilling data directly reflect the mechanical response of rocks during penetration (Hassaan et al., 2024), providing a high representative value of the actual geomechanical behavior of the subsurface formation.

Within this framework, the results obtained in this dissertation demonstrate that models based on real time drilling data achieved outstanding performance in predicting Poisson's ratio, with very low error values ($RMSE \approx 0.025$ and $MAE \approx 0.019$), indicating high numerical accuracy and strong suitability for direct engineering applications. In contrast, as illustrated in Figure 40, although well log-based models exhibited higher explanatory power in predicting Young's modulus, with an R^2 value of 0.963, their associated error metrics were higher than those obtained from drilling data-based models achieving ($RMSE \approx 2.281$, $MAE \approx 1.355$). This highlights a fundamental distinction between statistical explanatory capability and the actual numerical accuracy of predictions.

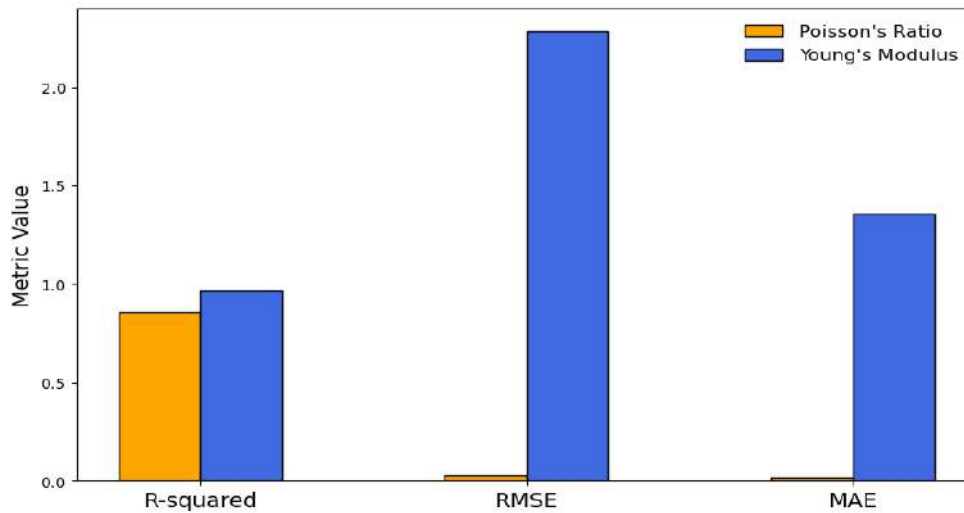


Figure 40. Comparison of Evaluation Metrics for Poisson's Ratio and Young's Modulus

These findings indicate that the coefficient of determination (R^2) is an important and essential metric for evaluating a model's ability to explain data variability and to understand the statistical relationships between variables (Chicco et al., 2021). Nevertheless, relying on R^2 alone may not be sufficient to assess the quality of predictive models in sensitive operational applications. A high R^2 value does not necessarily guarantee low prediction errors or an adequate level of engineering accuracy, as this metric does not directly quantify the numerical deviations between predicted and observed values (Alexander et al., 2015).

In this context, recent literature, such as the studies by (Willmott and Matsuura, 2005), (Spiess and Neumeier, 2010) and (Li, 2017), has emphasized that R^2 should be interpreted with caution, particularly when evaluating nonlinear models or comparing models with different levels of complexity. Consequently, it becomes essential to complement R^2 with error-based performance metrics that reflect the actual predictive behavior of the model. Absolute error metrics such as MAE and RMSE are therefore considered complementary to the coefficient of determination and more representative of true model performance from an engineering standpoint, as they directly quantify the deviation between predicted and actual values and provide indicators that are interpretable in practical and operational comparisons.

Based on the promising results demonstrated in this study and analyses, predicting geomechanical properties using real-time drilling data can be considered an effective and suitable approach for reservoir design applications and modern drilling operations management. This approach combines immediate data availability, reduced cost, and high numerical prediction accuracy, while allowing the coefficient of determination to be used as a supportive

explanatory indicator within an integrated evaluation framework that primarily relies on error metrics with clear engineering significance.

4.3.1.2. Beyond evaluation metrics: the role of data type and reservoir complexity in model behavior

The results of this study reveal a noticeable disparity in the performance of geomechanical prediction models when using drilling data compared to well logging data. These differences cannot be explained solely through conventional evaluation metrics, whether in terms of statistical explanatory power as reflected by the coefficient of determination (R^2) or in terms of actual numerical accuracy expressed by error indicators such as RMSE and MAE. Rather, the underlying causes extend to deeper factors related to the nature and quality of the data, as well as the type of input variables used in model training.

The fundamental differences between drilling data and well logging data in terms of their physical origin, statistical structure, and sensitivity to the geological and geomechanical characteristics of the reservoir lead to distinct learning behaviors within artificial intelligence-based algorithms. This effect becomes even more pronounced in unconventional and heterogeneous reservoirs, where high variability, discontinuities, and complex mechanical rock behavior significantly influence how algorithms learn from data. Consequently, understanding the interaction between evaluation metrics, data type, and reservoir characteristics represents a critical step in explaining the observed differences in model performance and provides an essential framework for analyzing algorithm behavior in subsequent sections.

a. Influence of data type on machine learning model accuracy

The results of this study demonstrate that the type of input variables used to train machine learning models is a critical factor in determining the accuracy and stability of predictions of geomechanical properties, such as Poisson's ratio and Young's modulus. Models based on real-time drilling data exhibited strong and consistent performance compared to those developed using well-log data, across the various geological scenarios investigated, including heterogeneous and naturally fractured reservoirs.

Drilling data include operational parameters such as rate of penetration (ROP), weight on bit (WOB), rotational speed (RPM), standpipe pressure (SPP), and torque (T). These variables indirectly yet instantaneously reflect the mechanical response of rocks during the drilling process. Several studies have demonstrated that such parameters are strongly correlated with rock mechanical properties, making drilling data a reliable and information-rich source for

geomechanical characterization, particularly in real-time applications (Elkatatny et al., 2018; Siddig and Elkatatny, 2021; Hassaan et al., 2024).

In contrast, well-log data rely on physical and geophysical measurements acquired after drilling operations and are often affected by several factors, including drilling fluid type, borehole conditions, geophysical noise, and subtle lithological variations. Although well logs may exhibit higher correlation coefficients in certain cases, their high sensitivity to noise and lack of temporal continuity can lead to increased numerical instability in predictions, especially in complex geological environments (Dramsch, 2020; Mahmoud et al., 2021).

These findings indicate that drilling data, owing to their continuous and time-dependent nature and their ability to capture the direct interaction between the rock and the drilling tool, provide more consistent and stable information for training artificial intelligence models, even in cases where the coefficient of determination is relatively lower. This explains why drilling-based models achieved lower prediction errors (RMSE and MAE) compared to well-log-based models, reinforcing their suitability for sensitive operational and engineering applications (Hastie et al., 2009; Dramsch, 2020).

Therefore, the observed differences in model performance are not solely attributable to the algorithms employed, but are closely linked to the nature of the input variables, their physical response to the geomechanical behavior of the rocks, and their ability to capture reservoir complexity and heterogeneity.

b. Influence of Reservoir properties on machine learning prediction accuracy

The geological and mechanical characteristics of a reservoir are key factors that govern the reliability and accuracy of machine learning models, particularly when predicting geomechanical properties in complex environments such as unconventional or naturally fractured reservoirs (Syed et al., 2022). The geomechanical behavior of rocks is not represented solely by fixed elastic properties but results from a complex interaction between mineral composition, cohesion patterns, fracture density, lithological heterogeneity, and in-situ stress distribution, all of which directly affect rock response during drilling and the geophysical measurements recorded in the well (Fjaer et al., 2008; Zoback, 2010; Askaripour et al., 2022).

In relatively homogeneous reservoirs, the relationships between input variables (whether real-time drilling data or well logs) and geomechanical properties are more stable, allowing machine learning models to capture physical patterns more effectively and achieve higher predictive accuracy (Syed et al., 2022; Isah et al., 2025). In contrast, in heterogeneous or naturally fractured reservoirs, rapid variations in mechanical properties with depth and the

presence of directional rock behavior (mechanical anisotropy) increase the nonlinearity of these relationships, posing additional challenges for statistical and learning algorithms (Warpinski and Teufel, 1987; Jaeger et al., 2009).

This issue becomes even clearer when using well logs, as the elastic measurements derived from them are directly linked to the physical properties of rocks, such as density, seismic velocities, and fabric, which in turn reflect the geomechanical behavior of formations under in-situ stress conditions. This close physical correlation enhances the ability of machine learning models to capture governing relationships between input variables and target geomechanical properties, especially when the data adequately represent lithological and structural variations along the well (Syed et al., 2022; Zhao et al., 2025). Moreover, the sensitivity of well logs to subtle lithological and formation changes allows models to detect mechanical responses related to stress variations and natural fractures, particularly in complex reservoirs. This contributes to improved characterization of rock geomechanical behavior, as the recorded measurements reflect an integrated physical response of the formation rather than isolated properties, as supported by recent studies confirming the effectiveness of integrating well logs into geomechanical parameter modeling using machine learning techniques (Alameedy et al., 2022; Sanei et al., 2023; Gabry et al., 2023).

In contrast, drilling data reflect an integrated mechanical response of rocks during penetration, where variables such as rate of penetration (ROP), torque (T), weight on bit (WOB), and standpipe pressure (SPP) are directly affected by rock resistance and changes in mechanical properties. This dynamic and integrated nature makes drilling data more capable of representing the overall geomechanical behavior of the reservoir, particularly in heterogeneous environments, explaining the stability of predictions and lower absolute errors reported in several recent studies (Elkatatny, 2018; Siddig and Elkatatny, 2021; Hassaan et al., 2024).

Furthermore, reservoir characteristics control the generalization capability of models. Models trained on reservoirs with specific mechanical behavior may perform differently when applied to reservoirs with varying fracture intensity or stress conditions. This highlights the importance of advanced strategies, such as Transfer Learning or multi-reservoir modeling, to enhance model performance across wide-ranging geomechanical variability (Pan and Yang, 2009; Bergen et al., 2019).

Overall, the accuracy of machine learning models in predicting geomechanical properties cannot be separated from the geological and mechanical context of the reservoir. The success of these models largely depends on the compatibility of input variables with the

physical behavior of rocks and the ability of the data to represent the true complexity of the reservoir, rather than solely on algorithm selection or hyperparameter tuning.

c. Influence of Algorithm Characteristics and Structure on Predictive Model Performance

The specific structure of gradient boosting family algorithms is based on the principle of sequential model construction, in which a set of weak learners is generated iteratively, and each new model is directed toward correcting the residual shortcomings of the previous ones through error-weighting mechanisms and weight updating strategies (Natekin and Knoll, 2013; Chen and Guestrin, 2016). This algorithmic structure enables the capture of nonlinear relationships and complex interactions among variables, particularly when dealing with multiple geological and geophysical inputs that represent the physical behavior of rock systems (Bentéjac et al., 2021).

This structural framework is clearly reflected in the study results when the data are classified according to borehole diameter Figure 41. The QH-Class 1 Reservoir, categorized by a borehole diameter of less than 7 inches, exhibits a high degree of structural homogeneity along with a large number of logging data points (1828 points). This context provided a favorable environment for extracting complex geomechanical relationships when adopting the first scenario (XGB1; $R^2 \approx 0.963$, RMSE ≈ 2.281 , MAE ≈ 1.355), which relies on a comprehensive set of geophysical inputs, namely: Dept, NPHI, POTA, GR, PE, Cali, URAN, THOR, and RT. The diversity of these inputs enabled the sequential structure of the algorithm to capture subtle patterns that accurately reflect the mechanical behavior of the reservoir (Friedman, 2002; Chen et al., 2015; Lawal et al., 2024; Isah et al., 2025), particularly in heterogeneous reservoirs with limited petrophysical properties, such as tight reservoirs characterized by a dense network of natural fractures, as is the case for the Red Quartzite reservoir investigated in this study.

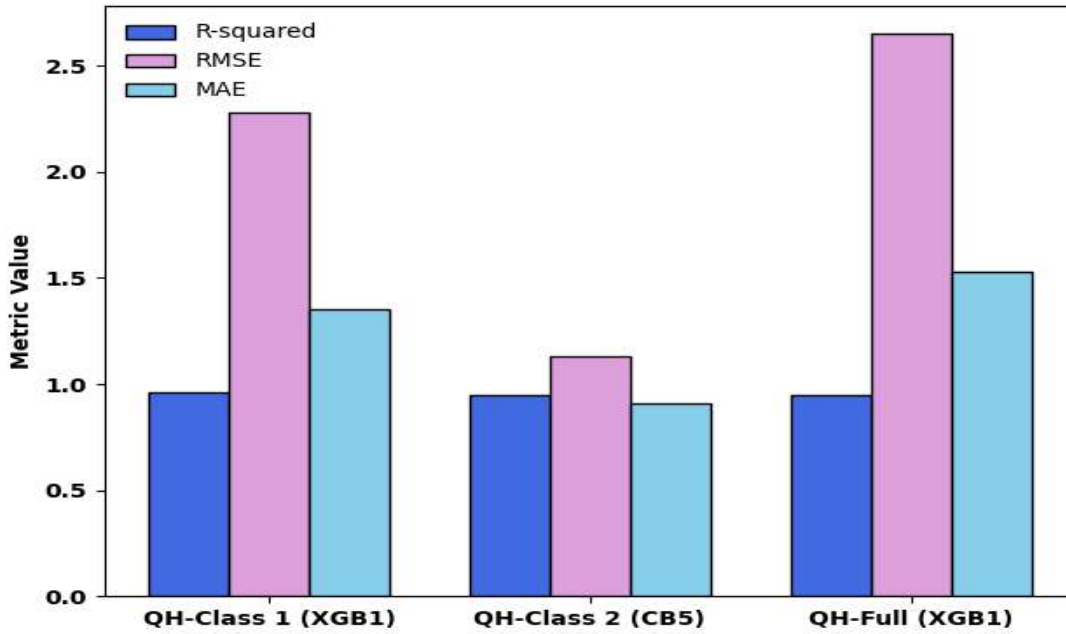


Figure 41. comparative evaluation of model performance Metrics across reservoir class

Similarly, although the complete dataset (QH-Full Reservoir), which includes all available data points (1894 points) without prior classification by borehole diameter, exhibits relatively limited structural variability due to the smaller proportion of data associated with boreholes less than 7 inches compared to those greater than or equal to 7 inches, XGBoost was still able to maintain its superior performance when using the full set of input variables (XGB1; $R^2 \approx 0.952$, $RMSE \approx 2.654$, $MAE \approx 1.539$). This confirms the capability of the algorithm's sequential architecture to handle large and diverse datasets without compromising the overall consistency of the predictive model (Friedman, 2002; Chen et al., 2015).

In contrast, the QH-Class 2 Reservoir represents a small dataset in terms of the number of data points (181 points), yet it is homogeneous with respect to borehole diameter (≥ 7 inches). In this context, the CatBoost algorithm demonstrated notable efficiency under the fifth scenario (CB5; $R^2 \approx 0.953$, $RMSE \approx 1.139$, $MAE \approx 0.914$), which relies on a reduced set of key input variables, namely: Dept, NPHI, POTA, GR, and PE. This selection allowed the model to focus on the most representative variables governing the geomechanical behavior of the reservoir under conditions of limited data availability. This behavior highlights the ability of CatBoost algorithm, through its symmetric tree structure and bias-reduction mechanisms, to achieve an effective balance between model simplicity and the richness of extracted information, making it particularly well suited for small datasets when structural homogeneity is preserved (Friedman, 2002; Prokhorenkova et al., 2018; Hancock and Khoshgoftaar, 2020; Chen et al., 2024).

These applied results demonstrate that variations in algorithm performance are not solely related to data volume, but are also strongly influenced by the degree of data homogeneity and the selection of appropriate input scenarios as emphasized in several geoscience-oriented machine learning studies (Chen and Guestrin, 2016; Prokhorenkova et al., 2018; Dramsch, 2020; Syed et al., 2022; Lawal et al., 2024). While powerful boosting algorithms such as XGBoost exhibit clear superiority when large or relatively homogeneous datasets with a high number of input variables are available, algorithms such as CatBoost emerge as effective solutions in cases of limited data, particularly when the number of input variables is reduced and attention is focused on those most representative of the geomechanical behavior of rocks.

Accordingly, this study confirms that the alignment between algorithm characteristics, data structure, and input scenarios is a critical factor in achieving reliable predictive performance (Dramsch, 2020). The selection of the optimal model should therefore be based on an integrated understanding of the geological and statistical context, rather than on the absolute superiority of a single algorithm (Syed et al., 2022; Lawal et al., 2024).

4.3.2. Operational and economic Implications of Real-Time Geomechanical Characterization Based on Drilling

4.3.2.1. Enabling Real-Time Geomechanical Characterization During Drilling

The results of this study highlight that predicting geomechanical properties based on drilling data represents one of the most valuable applied outcomes with high operational relevance, particularly in the context of real-time drilling operations (Elkatatny et al., 2017). Unlike conventional well logging, which requires additional time and significant costs and is typically available only after specific drilling stages are completed, drilling data provide a continuous stream of geomechanical information from the onset of operations throughout well advancement, and across all wells without exception (Ahmed et al., 2021; Siddig et al., 2021; Siddig and Elkatatny., 2023).

This real-time geomechanical characterization enables immediate adjustment of drilling parameters in response to changes in rock mechanical behavior, directly contributing to improved wellbore stability and mitigation of drilling-related issues such as mud losses and drill string sticking. This capability is especially critical in heterogeneous and naturally fractured reservoirs, where geomechanical properties may vary rapidly over small depth intervals, making rapid operational response a decisive factor in the success of field operations (Zoback, 2010; Fjær et al., 2008).

Recent studies indicate that integrating artificial intelligence models with real-time drilling data significantly enhances geomechanical risk management, reduces non-productive time, and improves overall operational efficiency. This integration leads to cost reduction and increased drilling reliability, particularly in complex environments characterized by high geological uncertainty ([Dramschi, 2020](#); [Ameur-Zaimeche et al., 2022](#)).

4.3.2.2. Implications for Hydraulic Fracturing and Uncertainty Reduction

Geomechanical properties particularly Young's modulus and Poisson's ratio play a central role in the design and execution of hydraulic fracturing operations, as they govern rock resistance to deformation and failure, fracture propagation behavior, and preferred fracture orientations under in-situ stress conditions ([Zoback, 2010](#); [Fjær et al., 2008](#)). Geomechanical models, especially those driven by drilling data, allow continuous updating of these properties, enabling more accurate identification of highly fracture-prone zones, improved estimation of optimal fracturing pressures, fluid volumes, and injection rates that minimize undesirable deformation or ineffective fracturing ([Economides and Nolte, 2000](#); [Warpinski et al., 2009](#)).

In tight fractured reservoirs, where geological and geomechanical uncertainty is inherently high, this approach contributes to uncertainty reduction by integrating real-time information rather than relying on post-drilling or sparsely representative datasets. It also enables real-time refinement of fracture propagation models, facilitating anticipation of interactions between natural and induced fractures and reducing operational risks associated with their impact on wellbore stability and equipment integrity ([Warpinski et al., 2009](#)).

Furthermore, the use of drilling data for geomechanical property prediction enhances the design of targeted hydraulic fracturing programs, including optimal zone selection for high-pressure fluid injection and the sequencing of fracturing stages along the wellbore based on geomechanical variability. This leads to more effective utilization of natural fracture networks, reduced energy and resource waste, and improved overall reservoir productivity ([Zoback, 2010](#); [Lake, 2014](#)).

4.3.2.3. Economic Implications: cost efficiency and resource optimization

This study confirms that drilling-data-based geomechanical characterization represents a cost-effective alternative to expensive well logging programs and laboratory experiments with limited representativeness. Accurate numerical prediction of Young's modulus and Poisson's ratio using data already available in all wells results in substantial operational cost reductions, both by minimizing reliance on off-site laboratory testing and by reducing the need for

additional logging runs in each well (Ahmed et al., 2021; Siddig et al., 2021; Siddig and Elkatatny., 2023; Kermia et al., 2025).

Real-time geomechanical characterization also contributes to reducing non-productive time associated with issues such as wellbore collapse or tool sticking, as drilling parameters can be adjusted instantaneously in response to evolving rock properties. This improves resource utilization efficiency, including drilling fluids, equipment, and personnel.

Moreover, this approach enhances production efficiency by supporting improved planning of hydraulic fracturing operations and optimal selection of perforation and stimulation zones, thereby reducing waste and increasing the economic return per well. This aligns with the global transition toward cost-efficient, high-performance digital solutions in petroleum engineering, particularly in the context of unconventional reservoir development challenges and the increasing pressure to reduce operational costs and environmental impacts (Dramschi, 2020; Lawal et al., 2024; Isah et al., 2025).

4.4. Impact of the Study on the Rhourde Nouss Field

The results of this study demonstrate significant scientific, operational, and economic implications for the Rhourde Nouss field, located within the Berkine Basin, particularly in the context of the Cambro-Ordovician Hamra Quartzite reservoir, which is characterized by its tight nature, intense natural fracturing, and primary reliance on hydraulic fracturing for production.

In this context, the conventional core-based approach faces major limitations in this field, as the fractured and weakly consolidated nature of the Hamra Quartzite reservoir often leads to difficulties in core recovery or to the retrieval of poor-quality samples due to disintegration along fracture planes. Moreover, the presence of a dense natural fracture network renders these samples poorly representative of the true geomechanical behavior of the reservoir and limits their suitability for reliable laboratory testing. Consequently, reliance on laboratory measurements in such a context becomes of limited effectiveness in terms of accuracy and representativeness.

In addition, the low reliability of sonic transit time (Δt) measurements constitutes one of the major challenges in reservoir characterization. The use of high-density drilling mud, which is necessary to ensure wellbore stability and prevent collapse due to pressure differentials, leads to attenuation or loss of the acoustic signal as it penetrates the drilling fluid, resulting in inaccurate or misleading measurements. This represents a major limitation for conventional geomechanical characterization methods based on sonic and density logs.

To overcome these limitations, the application of machine learning models based on real-time drilling data provides an effective and cost-efficient framework for continuous geomechanical characterization across the field, while bypassing the constraints associated with core quality and conventional logging data. In this context, geomechanical properties, particularly Young's modulus and Poisson's ratio, play a critical role in controlling hydraulic fracture propagation, rock brittleness, and fracture orientation under in-situ stress conditions.

Accurate real-time prediction of these parameters enables optimization of hydraulic fracturing design through the selection of optimal injection pressures, appropriate fluid volumes, and adjusted pumping rates in accordance with the mechanical response of the rock, thereby reducing the risk of treatment failure or ineffective fracture development. These models also contribute to identifying the most fracture-prone zones within the Hamra Quartzite reservoir, particularly in heterogeneous and naturally fractured environments, allowing optimal exploitation of natural fracture networks and improved connectivity between induced and natural fractures.

Within this framework, this study provides additional value by contributing to the mitigation of one of the major challenges affecting the Rhourde Nouss field, namely the progressive increase in water cut during production, resulting from water migration through natural fractures that act as preferential flow pathways. Accurate geomechanical characterization enables a better understanding of the interaction between hydraulic fractures and natural fractures, helping to control fracture propagation and avoid their connection with water-bearing zones. It also supports the selection of fracturing intervals with weak connectivity to water-saturated fractures and limits the activation or re-opening of water-conducting pathways through optimized control of injection pressures and fluid volumes. Consequently, this approach contributes to reducing early water breakthrough and improving hydrocarbon recovery efficiency.

Furthermore, real-time geomechanical characterization reduces uncertainties related to fracture behavior and enhances control over the interaction between hydraulic and natural fractures, which is a key factor in improving well productivity in the Rhourde Nouss field. From an operational perspective, this approach enhances wellbore stability and reduces non-productive time through immediate adaptation to changes in rock properties during drilling. From an economic standpoint, the use of continuously available drilling data reduces the need for costly logging operations and laboratory testing, making this approach highly suitable for large-scale field development. Overall, this study provides an integrated solution that supports the optimization of hydraulic fracturing design and execution, improves production efficiency,

mitigates water cut issues, and enhances the productivity of tight reservoirs in the Rhourde Nouss field.

General Conclusion

General Conclusion

Effective exploitation of petroleum reservoirs relies on a comprehensive understanding of reservoir properties before and during drilling, particularly geomechanical parameters such as Poisson's ratio and Young's modulus. These parameters play a critical role in addressing key drilling and production challenges, including wellbore stability, in-situ stress estimation, drilling performance optimization, and hydraulic fracturing design. Conventionally, Poisson's Ratio and Young's modulus are estimated through laboratory measurements on core samples or derived from acoustic and density well logs. However, in tight and naturally fractured reservoirs, these conventional approaches are often limited due to poor core recovery, missing and degraded acoustic log quality caused by drilling fluids and complex borehole conditions.

To overcome these limitations, this research proposes intelligent, data driven frameworks for the real-time and continuous prediction of geomechanical properties using machine learning techniques. In the first part of the study, an innovative model was developed to predict dynamic Poisson's ratio while drilling, relying solely on surface recorded drilling data acquired from a horizontal well in the Cambro-ordovician tight fractured Quartzite El Hamra formation in the Rhourde Nouss field, southeast Algeria. Three widely used machine learning algorithms Random Forest, Gradient Boosting, and Extreme Gradient Boosting were trained and evaluated using key drilling parameters, including (WOB, T, SPP, RPM, ROP, Q). Among these models, XGBoost demonstrated superior predictive performance, achieving a high correlation coefficient ($R^2 \approx 0.856$) and the lowest prediction errors (RMSE ≈ 0.025 , MAE ≈ 0.019), highlighting its robustness and ability to capture complex drilling formation interactions.

In parallel, the study further demonstrated the effectiveness of machine learning in predicting the dynamic Young's modulus (E_{dyn}) using geophysical well logging data acquired from a vertical well. Six algorithms (XGB, LGBM, CB, RF, GB, and AB) were evaluated using conventional petrophysical logs (GR, NPHI, CALI, RT, URAN, THOR, and POTA). The results consistently showed that XGBoost outperformed the other models on both full and classified datasets, reaching a high predictive accuracy ($R^2 \approx 0.963$ for QH-Class 1) with low error values (RMSE ≈ 2.281 , MAE ≈ 1.355). These results confirm the strong capability of boosting-based models to handle nonlinear relationships and noisy data commonly encountered in tight fractured reservoirs.

To further enhance prediction reliability, a novel reservoir classification strategy based on borehole diameter (CALI) was introduced. This classification significantly improved model

performance by reducing internal data variability and enabling the algorithms to learn from more homogeneous data subsets. Such an approach is particularly well suited for tight and naturally fractured formations like Quartzite El Hamra, where acoustic and density logs are frequently unreliable or unavailable due to high mud weight and adverse borehole conditions.

The proposed methodologies represent a significant advancement in characterizing geomechanical properties in tight fractured reservoirs, while this study also demonstrates the effectiveness of artificial intelligence techniques in predicting reservoir geomechanical properties, enhancing estimation accuracy and improving evaluation processes. By providing reliable alternatives to laboratory tests and conventional acoustic methods, these approaches enable continuous and cost-effective estimation of key mechanical properties using drilling data or well logs.

An additional advantage of this study lies in highlighting the significant benefits of using drilling data for the prediction of geomechanical compared to conventional logging data. Drilling data are continuously acquired in real time at the surface, without the need for additional downhole operations, making them inherently more cost-effective and operationally efficient.

Unlike well logging, which is often performed after drilling and may be limited by borehole conditions, tool availability, and high operational costs, drilling parameters provide uninterrupted measurements throughout the drilling process. This enables real-time estimation of critical reservoir and geomechanical properties, allowing timely decision-making, early risk mitigation, and optimization of drilling operations.

Moreover, in challenging environments such as tight fractured reservoirs, where logging data may be degraded or unavailable, drilling data offer a reliable alternative for continuous reservoir characterization. Consequently, integrating drilling data with intelligent models represents a strategic shift toward more economical, real-time, and data driven reservoir evaluation workflows in modern petroleum engineering.

References

- Abada, Z., & Bouharkat, M. (2018). Study of management strategy of energy resources in Algeria. *Energy Reports*, 4, 1-7.
- Abdelaal, A., Elkatatny, S., & Abdulraheem, A. (2022). Real-time prediction of formation pressure gradient while drilling. *Scientific Reports*, 12(1), 11318.
- Abdelaal, A., Elkatatny, S., Gamal, H., & Ziadat, W. (2023, June). Drilling Data Based Approach for Equivalent Circulation Density Prediction While Drilling. In *ARMA US Rock Mechanics/Geomechanics Symposium* (pp. ARMA-2023). ARMA.
- Abdelgawad, K. Z., Elzenary, M., Elkatatny, S., Mahmoud, M., Abdulraheem, A., & Patil, S. (2019). New approach to evaluate the equivalent circulating density (ECD) using artificial intelligence techniques. *Journal of Petroleum Exploration and Production Technology*, 9, 1569-1578.
- Abdulahadi, D., Ali, J. A., & Hama, S. M. (2025). Advanced Techniques for Improving the Production of Natural Resources from Unconventional Reservoirs: A State-of-the-Art Review. *Energy & Fuels*.
- Acquah-Andoh, E., Putra, H. A., Ifelebuegu, A. O., & Owusu, A. (2019). Coalbed methane development in Indonesia: Design and economic analysis of upstream petroleum fiscal policy. *Energy Policy*, 131, 155-167.
- Ahmed, A., Elkatatny, S., & Abdulraheem, A. (2021). Real-time static Poisson's ratio prediction of vertical complex lithology from drilling parameters using artificial intelligence models. *Arabian Journal of Geosciences*, 14, 1-13.
- Ahmed, S., & Salehi, S. (2021). Failure mechanisms of the wellbore mechanical barrier systems: Implications for well integrity. *Journal of Energy Resources Technology*, 143(7), 073007.
- Ahmed, S., & Salehi, S. (2021). Failure mechanisms of the wellbore mechanical barrier systems: Implications for well integrity. *Journal of Energy Resources Technology*, 143(7), 073007.
- Ahmed, U., & Meehan, D. N. (Eds.). (2016). *Unconventional oil and gas resources: exploitation and development*. CRC Press.
- Ahr, W. M. (2011). *Geology of carbonate reservoirs: the identification, description and characterization of hydrocarbon reservoirs in carbonate rocks*. John Wiley & Sons.
- Aissaoui, M. N., Bédir, M., & Gabtni, H. (2016). Petroleum assessment of berkine–Ghadames basin, southern Tunisia. *AAPG Bulletin*, 100(3), 445-476.
- Alakbari, F. S., Mohyaldinn, M. E., Ayoub, M. A., Hussein, I. A., Muhsan, A. S., Ridha, S., & Salih, A. A. (2024). A gated recurrent unit model to predict Poisson's ratio using deep learning. *Journal of Rock Mechanics and Geotechnical Engineering*, 16(1), 123-135.
- Alam, K., Mitrofanov, A. V., & Silberschmidt, V. V. (2011). Experimental investigations of forces and torque in conventional and ultrasonically-assisted drilling of cortical bone. *Medical engineering & physics*, 33(2), 234-239.
- Alameedy, U., Alhaleem, A. A., Isah, A., Al-Yaseri, A., El-Husseiny, A., & Mahmoud, M. (2022). Predicting dynamic shear wave slowness from well logs using machine learning methods in the Mishrif Reservoir, Iraq. *Journal of Applied Geophysics*, 205, 104760.
- Al-Areeq, N. M. (2018). Petroleum source rocks characterization and hydrocarbon generation. *Recent Insights in Petroleum Science and Engineering*, 7(1).
- Al-Bazali, T. M., Zhang, J., Wolfe, C., Chenevert, M. E., & Sharma, M. M. (2009). Wellbore instability of directional wells in laminated and naturally fractured shales. *Journal of Porous Media*, 12(2).
- Alexander, D. L., Tropsha, A., & Winkler, D. A. (2015). Beware of R²: simple, unambiguous assessment of the prediction accuracy of QSAR and QSPR models. *Journal of chemical information and modeling*, 55(7), 1316-1322.

- Ali, Z. A., Abduljabbar, Z. H., Tahir, H. A., Sallow, A. B., & Almufti, S. M. (2023). eXtreme gradient boosting algorithm with machine learning: A review. *Academic Journal of Nawroz University*, 12(2), 320-334.
- Al-Kasassbeh, M., Abbadi, M. A., & Al-Bustanji, A. M. (2020, July). LightGBM algorithm for malware detection. In *Science and Information Conference* (pp. 391-403). Cham: Springer International Publishing.
- Allen, P. A., & Allen, J. R. (2013). *Basin analysis: Principles and application to petroleum play assessment*. John Wiley & Sons.
- Alsabaa, A., Gamal, H., Elkatatny, S., & Abdulraheem, A. (2021). New correlations for better monitoring the all-oil mud rheology by employing artificial neural networks. *Flow Measurement and Instrumentation*, 78, 101914.
- Alsaihati, A., Elkatatny, S., & Gamal, H. (2022). Rate of penetration prediction while drilling vertical complex lithology using an ensemble learning model. *Journal of Petroleum Science and Engineering*, 208, 109335.
- Altman, N., & Krzywinski, M. (2017). Ensemble methods: bagging and random forests. *Nature Methods*, 14(10), 933-935.
- Ameur-Zaimeche, O., Kechiched, R., Heddam, S., & Wood, D. A. (2022). Real-time porosity prediction using gas-while-drilling data and machine learning with reservoir associated gas: Case study for Hassi Messaoud field, Algeria. *Marine and Petroleum Geology*, 140, 105631.
- Aminzadeh, F., & Dasgupta, S. N. (2013). *Geophysics for petroleum engineers* (Vol. 60). Newnes.
- Anders, M. H., Laubach, S. E., & Scholz, C. H. (2014). Microfractures: A review. *Journal of Structural Geology*, 69, 377-394.
- Arab, M., & Djezzar, S. (2011, March). Gas Potential Evaluation Of Reggane Basin, South-Western Algeria. In *Offshore Mediterranean Conference and Exhibition* (pp. OMC-2011). OMC.
- Ardakani, O. H., Sanei, H., Ghanizadeh, A., McMechan, M., Ferri, F., & Clarkson, C. R. (2017). Hydrocarbon potential and reservoir characteristics of lower cretaceous Garbutt Formation, Liard Basin Canada. *Fuel*, 209, 274-289.
- Arjun Patil, P., & Teodoriu, C. (2013). Model development of torsional drillstring and investigating parametrically the stick-slips influencing factors. *Journal of Energy Resources Technology*, 135(1), 013103.
- Ashraf Ahmed, Salaheldin Elkatatny, Ahmed Alsaihati, "Applications of Artificial Intelligence for Static Poisson's Ratio Prediction While Drilling", *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 9956128, 10 pages, 2021. <https://doi.org/10.1155/2021/9956128>
- Askaripour, M., Saeidi, A., Mercier-Langevin, P., & Rouleau, A. (2022). A review of relationship between texture characteristic and mechanical properties of rock. *Geotechnics*, 2(1), 262-296.
- Askri, H., Belmecheri, A., Benrabah, B., Boudjema, A., Boumendjel, K., Daoudi, M., ... & Terkmani, M. (1995, March). *Geology of Algeria*. In *Well evaluation conference Algeria* (pp. 1-93). Houston, TX, USA: Schlumberger.
- Asquith G B, Krygowski D and Gibson C R 2004 *Basic well log analysis* (2nd edn), American Association of Petroleum Geologists, Tulsa.
- Asselman, A., Khaldi, M., & Aammou, S. (2023). Enhancing the prediction of student performance based on the machine learning XGBoost algorithm. *Interactive Learning Environments*, 31(6), 3360-3379. 10.1080/10494820.2021.1928235
- Atkinson, B. K. (Ed.). (2015). *Fracture mechanics of rock*. Elsevier.
- Aydin, A. (2000). Fractures, faults, and hydrocarbon entrapment, migration and flow. *Marine and petroleum geology*, 17(7), 797-814.

- Aziz, N., Akhir, E. A. P., Aziz, I. A., Jaafar, J., Hasan, M. H., & Abas, A. N. C. (2020, October). A study on gradient boosting algorithms for development of AI monitoring and prediction systems. In 2020 International conference on computational intelligence (ICCI) (pp. 11-16). IEEE.
- Bacon, M., Simm, R., & Redshaw, T. (2007). 3-D seismic interpretation. Cambridge University Press.
- Bahad, P., & Saxena, P. (2019, December). Study of adaboost and gradient boosting algorithms for predictive analytics. In International Conference on Intelligent Computing and Smart Communication 2019: Proceedings of ICSC 2019 (pp. 235-244). Singapore: Springer Singapore.
- Bai, Y., Cao, G., Wei, G., Nan, X., Cheng, Q., Du, T., & An, H. (2020). Mechanism of permeability and oil recovery during fracturing in tight oil reservoirs. *Processes*, 8(8), 972.
- Baouche, R., & Wood, D. A. (2020). Characterization and estimation of gas-bearing properties of Devonian coals using well log data from five Illizi Basin wells (Algeria). *Advances in Geo-Energy Research*, 4(4), 356-371.
- Baouche, R., Sen, S., Ganguli, S. S., & Boutaleb, K. (2021). Petrophysical and geomechanical characterization of the Late Cretaceous limestone reservoirs from the Southeastern Constantine Basin, Algeria. *Interpretation*, 9(4), SH1-SH9.
- Barree, R. D., Gilbert, J. V., & Conway, M. W. (2009, January). Stress and rock property profiling for unconventional reservoir stimulation. In SPE Hydraulic Fracturing Technology Conference and Exhibition (pp. SPE-118703). SPE.
- Basheer, A. A. (2024). Deep Crustal and Upper Mantle Structures in North Africa: A Review. *The Geology of North Africa*, 21-45.
- Beagle, J. R. (2025). Tracing Gas Interaction and Mixing Processes in Natural Gases from the Sichuan and Bohai Bay Basins, China: Geochemical Insights from Noble Gas Isotopic Signatures (Doctoral dissertation, Lancaster University (United Kingdom)).
- Beaumont, C. (1981). Foreland basins. *Geophysical Journal International*, 65(2), 291-329.
- Beicip.Franlab, 1995 : révision du model géologique du champs de RhourdeNouss
- BeicipFranab ,2002 : étude sédimentologique du bassin de Berkin, rapport interne sonatrach.
- Bekkouche, D., & Meraghni, N. (2011, March). The Replacement Of Prospective Resources Of The Algerian Sedimentary Basins. In Offshore Mediterranean Conference and Exhibition (pp. OMC-2011). OMC.
- Belica, M. E., Tohver, E., Poyatos-Moré, M., Flint, S., Parra-Avila, L. A., Lanci, L., ... & Pisarevsky, S. A. (2017). Refining the chronostratigraphy of the Karoo Basin, South Africa: magnetostratigraphic constraints support an Early Permian age for the Eccu Group. *Geophysical Journal International*, 211(3), 1354-1374.
- Benayad, S., Park, Y. S., Chaouchi, R., & Kherfi, N. (2013). Unconventional resources in Algeria: appraisal result from the Hamra Quartzite reservoir. *Geosciences Journal*, 17(3), 313-327.
- Bentéjac, C., Csörgő, A., & Martínez-Muñoz, G. (2021). A comparative analysis of gradient boosting algorithms. *Artificial Intelligence Review*, 54(3), 1937-1967.
- Bentley, R. W. (2002). Global oil & gas depletion: an overview. *Energy policy*, 30(3), 189-205.
- Bergen, K. J., Johnson, P. A., de Hoop, M. V., & Beroza, G. C. (2019). Machine learning for data-driven discovery in solid Earth geoscience. *Science*, 363(6433), eaau0323.
- Berkowitz, B. (2002). Characterizing flow and transport in fractured geological media: A review. *Advances in water resources*, 25(8-12), 861-884.
- Birmingham, M. L., Pong-Wong, R., Spiliopoulou, A., Hayward, C., Rudan, I., Campbell, H., ... & Haley, C. S. (2015). Application of high-dimensional feature selection: evaluation for genomic prediction in man. *Scientific reports*, 5(1), 1-12. <https://doi.org/10.1038/srep10312>

- Beydoun, Z. R. (1995). Productive Middle East clastic oil and gas reservoirs: their depositional settings and origins of their hydrocarbons. *Sedimentary Facies Analysis: a Tribute to the Research and Teaching of Harold G. Reading*, 329-354.
- Bishop, C. M., & Nasrabadi, N. M. (2006). *Pattern recognition and machine learning* (Vol. 4, No. 4, p. 738). New York: springer.
- Bois, C., Bouche, P., & Pelet, R. (1982). Global geologic history and distribution of hydrocarbon reserves. *AAPG Bulletin*, 66(9), 1248-1270.
- Bons, P. D., Cao, D., De Riese, T., González-Esvertit, E., Koehn, D., Naaman, I., ... & Gomez-Rivas, E. (2022). A review of natural hydrofractures in rocks. *Geological Magazine*, 159(11-12), 1952-1977.
- Borjigin, T., Baojian, S. H. E. N., Lingjie, Y. U., Yunfeng, Y., Zhang, W., Cheng, T., ... & Jianzhong, Q. (2017). Mechanisms of shale gas generation and accumulation in the Ordovician Wufeng-Longmaxi Formation, Sichuan Basin, SW China. *Petroleum exploration and development*, 44(1), 69-78.
- Bosworth, W., & Tari, G. (2020). Hydrocarbon accumulation in basins with multiple phases of extension and inversion: examples from the Western Desert (Egypt) and the western Black Sea. *Solid Earth Discussions*, 2020, 1-31.
- Bouabdellah, M., & Slack, J. F. (2016). Geologic and metallogenic framework of North Africa. In *Mineral Deposits of North Africa* (pp. 3-81). Cham: Springer International Publishing.
- Boudjema, A. (1987). *Évolution structurale du bassin pétrolier "triasique" du Sahara Nord oriental (Algérie)* (Doctoral dissertation, Paris 11).
- Bratton, T., Canh, D. V., Van Que, N., Duc, N. V., Gillespie, P., Hunt, D., ... & Sonneland, L. (2006). The nature of naturally fractured reservoirs. *Oilfield Review*, 18(2), 4-23.
- Breiman, L. (2001). Random forests. *Machine learning*, 45, 5-32.
- Bruna Teixeira Silveira, B. S., Claudio Rabe, C. R., Cristian Mejia, C. M., & Deane Roehl, D. R. (2023, June). Neural network model to predict pore pressure and wellbore stability from LWD log and drilling parameter. In *ARMA US Rock Mechanics/Geomechanics Symposium* (pp. ARMA-2023). ARMA.
- Buitinck, L., Louppe, G., Blondel, M., Pedregosa, F., Mueller, A., Grisel, O., ... & Varoquaux, G. (2013). API design for machine learning software: experiences from the scikit-learn project. *arXiv preprint arXiv:1309.0238*.
- Buitinck, L., Louppe, G., Blondel, M., Pedregosa, F., Mueller, A., Grisel, O., ... & Varoquaux, G. (2013). API design for machine learning software: experiences from the scikit-learn project. *arXiv preprint arXiv:1309.0238*.
- Burdige, D. J. (2007). Preservation of organic matter in marine sediments: controls, mechanisms, and an imbalance in sediment organic carbon budgets?. *Chemical reviews*, 107(2), 467-485.
- Cai, Y., Yuan, Y., & Zhou, A. (2024). Predictive slope stability early warning model based on CatBoost. *Scientific Reports*, 14(1), 25727.
- Caineng, Z. O. U., Guangming, Z. H. A. I., Hongjun, W. A. N. G., Jianzhong, L. I., Zhaoming, W. A. N. G., Zhixin, W. E. N., ... & Xin, L. I. (2015). Formation, distribution, potential and prediction of global conventional and unconventional hydrocarbon resources. *Petroleum Exploration and Development*, 42(1), 14-28.
- Caineng, Z. O. U., Guangming, Z. H. A. I., Zhang, G., Hongjun, W., Zhang, G., Jianzhong, L., ... & Liang, K. (2015). Formation, distribution, potential and prediction of global conventional and unconventional hydrocarbon resources. *Petroleum Exploration and Development*, 42(1), 14-28.
- Caineng, Z. O. U., Zhi, Y. A. N. G., Dongbo, H. E., Yunsheng, W., Jian, L., Ailin, J., ... & Shen, Y. (2018). Theory, technology and prospects of conventional and unconventional natural gas. *Petroleum exploration and development*, 45(4), 604-618.

- Canfield, D. E. (1994). Factors influencing organic carbon preservation in marine sediments. *Chemical geology*, 114(3-4), 315-329.
- Carvalho, D. V., Pereira, E. M., & Cardoso, J. S. (2019). Machine learning interpretability: A survey on methods and metrics. *Electronics*, 8(8), 832.
- Castagna, J. P., Batzle, M. L., & Eastwood, R. L. (1985). Relationships between compressional-wave and shear-wave velocities in clastic silicate rocks. *geophysics*, 50(4), 571-581.
- Chang, C., Zoback, M. D., & Khaksar, A. (2006). Empirical relations between rock strength and physical properties in sedimentary rocks. *Journal of Petroleum Science and Engineering*, 51(3-4), 223-237.
- Chang, E., & Lippmann, R. P. (1990). Using genetic algorithms to improve pattern classification performance. *Advances in neural information processing systems*, 3.
- Chen, D., Zhang, C., Yang, M., Li, H., Wang, C., Diwu, P., ... & Wang, Y. (2024). Research on water invasion law and control measures for ultradeep, fractured, and low-porosity sandstone gas reservoirs: A case study of Kelasu gas reservoirs in Tarim Basin. *Processes*, 12(2), 310.
- Chen, J., & Xiao, X. (2014). Evolution of nanoporosity in organic-rich shales during thermal maturation. *Fuel*, 129, 173-181.
- Chen, T., & Guestrin, C. (2016, August). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining* (pp. 785-794). <https://dl.acm.org/doi/abs/10.1145/2939672.2939785>
- Chen, T., He, T., Benesty, M., Khotilovich, V., Tang, Y., Cho, H., ... & Zhou, T. (2015). Xgboost: extreme gradient boosting. *R package version 0.4-2*, 1(4), 1-4.
- Chen, X., Yang, J., & Gao, D. (2018). Drilling performance optimization based on mechanical specific energy technologies. *Drilling*, 1(1), 133-162.
- Chen, Y., Chen, B., & Shayilan, A. (2024). Combining categorical boosting and Shapley additive explanations for building an interpretable ensemble classifier for identifying mineralization-related geochemical anomalies. *Ore Geology Reviews*, 173, 106263.
- Chengzao, J. I. A. (2017). Breakthrough and significance of unconventional oil and gas to classical petroleum geology theory. *Petroleum Exploration and Development*, 44(1), 1-10.
- Chi, M. (2024, November). Optimization of Prediction Model Based on Gradient Boosting Decision Tree. In *2024 International Conference on Integrated Intelligence and Communication Systems (ICIICS)* (pp. 1-5). IEEE.
- Chicco, D., Warrens, M. J., & Jurman, G. (2021). The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *Peerj computer science*, 7, e623.
- Chitale D V, Coates G R, Day P I and Sigal R F 2003 System and method for clay typing using NMR-based porosity modeling; USOO6646437B1 Google Patents.
- Clarke, H., Turner, P., Bustin, R. M., Riley, N., & Besly, B. (2018). Shale gas resources of the Bowland Basin, NW England: a holistic study. *Petroleum Geoscience*, 24(3), 287-322.
- Clarkson, C. R., Haghsheenas, B., Ghanizadeh, A., Qanbari, F., Williams-Kovacs, J. D., Riazi, N., ... & Deglint, H. J. (2016). Nanopores to megafractures: Current challenges and methods for shale gas reservoir and hydraulic fracture characterization. *Journal of Natural Gas Science and Engineering*, 31, 612-657.
- Collins, G., Jones, M. P., Krane, J., Medlock, K., & Monaldi, F. (2021). Shale Renders the 'Obsolescing Bargain' Obsolete: Political risk and foreign investment in Argentina's Vaca Muerta. *Resources Policy*, 74, 102269.
- Cooksey, R. W., & Cooksey, R. W. (2020). Descriptive statistics for summarising data. *Illustrating statistical procedures: Finding meaning in quantitative data*, 61-139.

- Cramer, D. D. (2008, February). Stimulating unconventional reservoirs: lessons learned, successful practices, areas for improvement. In SPE Unconventional Resources Conference/Gas Technology Symposium (pp. SPE-114172). SPE.
- Dahi Taleghani, A. (2009). Analysis of hydraulic fracture propagation in fractured reservoirs: an improved model for the interaction between induced and natural fractures.
- Dardour, A. M., Boote, D. R. D., & Baird, A. W. (2004). Stratigraphic controls on Palaeozoic petroleum systems, Ghadames basin, Libya. *Journal of Petroleum Geology*, 27(2), 141-162.
- Davoodi, S., Mehrad, M., Wood, D. A., Al-Shargabi, M., Eremyan, G., & Shulgina, T. (2024). A novel data-driven model for real-time prediction of static Young's modulus applying mud-logging data. *Earth Science Informatics*, 17(6), 5771-5793.
- DeCelles, P. G., & Giles, K. A. (1996). Foreland basin systems. *Basin research*, 8(2), 105-123.
- Deng, H., Zhou, Y., Wang, L., & Zhang, C. (2021). Ensemble learning for the early prediction of neonatal jaundice with genetic features. *BMC medical informatics and decision making*, 21(1), 338.
- Derder, M. E. M., Henry, B., Maouche, S., Merabet, N. E., Amenna, M., & Bayou, B. (2018). Paleomagnetism of the Western Saharan basins: an overview. *The Geology of the Arab World---An Overview*, 291-318.
- DiMichele, W. A. (2014). Wetland-dryland vegetational dynamics in the Pennsylvanian ice age tropics. *International Journal of Plant Sciences*, 175(2), 123-164.
- Ding, X., Hasanipanah, M., Rouhani, M. M., & Nguyen, T. (2025). Hybrid catboost models optimized with metaheuristics for predicting shear strength in rock joints. *Bulletin of Engineering Geology and the Environment*, 84(3), 150.
- Ding, Y., Zhu, H., Chen, R., & Li, R. (2022). An efficient AdaBoost algorithm with the multiple thresholds classification. *Applied sciences*, 12(12), 5872.
- Ding, Z., Huang, Y., Yuan, H., & Dong, H. (2020). Introduction to reinforcement learning. In *Deep reinforcement learning: fundamentals, research and applications* (pp. 47-123). Singapore: Springer Singapore.
- Djezzar, S., & Boualam, A. (2025). Structural Analysis and Fracture Kinematics Using Seismic 2D and Geological Maps. *Geophysical Exploration for Hydrocarbon Reservoirs, Geothermal Energy, and Carbon Storage: New Technologies and AI-based Approaches*, 205-220.
- Djezzar, S., Boualam, A., Ouadi, H., Laalam, A., Mouedden, N., Merzoug, A., & Chemmakh, A. (2022, September). Geological characterization of lower devonian reservoirs in Reggane Basin, Algeria. In *SPE Annual Technical Conference and Exhibition?* (p. D011S011R001). SPE.
- Djouder, H. (2019). Silurian Succession from North Africa: Sedimentology, Ichnology and Thermal History for a New Era of Hydrocarbon Exploration. *Universite de Liege (Belgium)*.
- Dolson, J. C., Shann, M. V., Matbouly, S. I., Hammouda, H., & Rashed, R. M. (2001). Egypt in the twenty-first century: petroleum potential in offshore trends. *GeoArabia*, 6(2), 211-230.
- Dong, K., Jiang, M., Li, J., & Zhang, D. (2020). Research progresses in formation mechanism of complex fracture network for unconventional reservoir. *Arabian Journal of Geosciences*, 13(15), 750.
- Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., ... & Lautenbach, S. (2013). Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography*, 36(1), 27-46.
- Dramsch, J. S. (2020). 70 years of machine learning in geoscience in review. *Advances in geophysics*, 61, 1-55.
- Dusseault, M. B. (2011). Geomechanical challenges in petroleum reservoir exploitation. *KSCE Journal of Civil Engineering*, 15(4), 669-678.
- DWIVEDI, A. K. (2016). Petroleum exploration in India-a perspective and endeavours.
- Economides, M. J., & Nolte, K. G. (2000). *Reservoir stimulation* (Vol. 18). New York: Wiley.

- Eissa, E. A., & Kazi, A. (1988). Relation between static and dynamic Young's moduli of rocks. *International Journal of Rock Mechanics and Mining & Geomechanics Abstracts*, 25(6).
- Ekanayake, I. U., Meddage, D. P. P., & Rathnayake, U. (2022). A novel approach to explain the black-box nature of machine learning in compressive strength predictions of concrete using Shapley additive explanations (SHAP). *Case Studies in Construction Materials*, 16, e01059.
- Elkhatatny, S. (2018). Application of artificial intelligence techniques to estimate the static Poisson's ratio based on wireline log data. *Journal of Energy Resources Technology*, 140(7), 072905.
- Elkhatatny, S. (2022). Real-time prediction of the dynamic Young's modulus from the drilling parameters using the artificial neural networks. *Arabian Journal for Science and Engineering*, 47(9), 10933-10942.
- Elkhatatny, S. M., Tariq, Z., Mahmoud, M. A., & Al-AbdulJabbar, A. (2017, June). Optimization of rate of penetration using artificial intelligent techniques. In *ARMA US Rock Mechanics/Geomechanics Symposium* (pp. ARMA-2017). ARMA.
- Ellis D, Howard J, Flaum C, McKeon D, Scott H, Serra O and Simmons G 1988 Mineral logging parameters: Nuclear and acoustic; *Tech. Rev.* 36(1) 38.
- ElShawi, R., Sherif, Y., Al-Mallah, M., & Sakr, S. (2021). Interpretability in healthcare: A comparative study of local machine learning interpretability techniques. *Computational Intelligence*, 37(4), 1633-1650.
- Epelle, E. I., & Gerogiorgis, D. I. (2020). A review of technological advances and open challenges for oil and gas drilling systems engineering. *AIChE Journal*, 66(4), e16842.
- Erdenetsogt, B. O., Lee, I., Lee, S. K., Ko, Y. J., & Bat-Erdene, D. (2010). Solid-state C-13 CP/MAS NMR study of Baganuur coal, Mongolia: Oxygen-loss during coalification from lignite to subbituminous rank. *International Journal of Coal Geology*, 82(1-2), 37-44.
- Eschard, R., Braik, F., Bekkouche, D., Rahuma, M. B., Desaubliaux, G., Deschamps, R., & Proust, J. N. (2010). Palaeohighs: their influence on the North African Palaeozoic petroleum systems.
- Evans, D. A. (2013). Reconstructing pre-Pangean supercontinents. *Bulletin*, 125(11-12), 1735-1751.
- Fan, P., Liu, Y., Lin, Z., Guo, H., & Li, P. (2024). Experimental Study on the Efficiency of Fracturing Integrated with Flooding by Slickwater in Tight Sandstone Reservoirs. *Processes*, 12(11), 2529.
- Fangzheng, J. I. A. O. (2019). Re-recognition of "unconventional" in unconventional oil and gas. *Petroleum Exploration and Development*, 46(5), 847-855.
- Farag, S., Mas, C., Maizeret, P. D., Li, B., & Van Hung, L. (2010). An integrated workflow for granitic-basement-reservoir evaluation. *SPE Reservoir Evaluation & Engineering*, 13(06), 893-905.
- Fello, N. M., Deaf, A. S., & Leila, M. (2024). Petroleum Geology of North Africa. In *The Geology of North Africa* (pp. 265-303). Cham: Springer International Publishing.
- Fengyin, X. U., Wei, H. O. U., Xiong, X., Borui, X. U., Peng, W. U., Hongya, W. A. N. G., ... & Delei, M. A. O. (2023). The status and development strategy of coalbed methane industry in China. *Petroleum Exploration and Development*, 50(4), 765-783.
- Fjaer, E., Holt, R. M., Horsrud, P., & Raaen, A. M. (2008). *Petroleum related rock mechanics*. Elsevier.
- Flores, R. M. (2013). *Coal and coalbed gas: fueling the future*. Newnes.
- Ford, M., Masini, E., Vergés, J., Pik, R., Ternois, S., Léger, J., ... & Calassou, S. (2022). Evolution of a low convergence collisional orogen: a review of Pyrenean orogenesis. *BSGF-Earth Sciences Bulletin*, 193(1), 19.
- Fossen, H. (2016). *Structural geology*. Cambridge university press.
- Fowler, A. R., & Hamimi, Z. (2024). The Tuareg Shield and Trans-Saharan Orogenic Belt of Central Part of Northwest Africa: An Overview. *The Geology of North Africa*, 69-108.
- Freund, Y., & Schapire, R. E. (1997). A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of computer and system sciences*, 55(1), 119-139.

- Friedman, J. H. (2002). Stochastic gradient boosting. *Computational statistics & data analysis*, 38(4), 367-378.
- Gabry, M. A., Ali, A. G., & Elsayy, M. S. (2023, April). Application of Machine Learning Model for Estimating the Geomechanical Rock Properties Using Conventional Well Logging Data. In *Offshore Technology Conference* (p. D021S028R004). OTC.
- Gale, J. F., Laubach, S. E., Olson, J. E., Eichhubl, P., & Fall, A. (2014). Natural fractures in shale: A review and new observations. *AAPG bulletin*, 98(11), 2165-2216.
- Gale, J. F., Reed, R. M., & Holder, J. (2007). Natural fractures in the Barnett Shale and their importance for hydraulic fracture treatments. *AAPG bulletin*, 91(4), 603-622.
- Galeazzi, S., Point, O., Haddadi, N., Mather, J., & Druesne, D. (2010). Regional geology and petroleum systems of the Illizi–Berkine area of the Algerian Saharan Platform: An overview. *Marine and Petroleum Geology*, 27(1), 143-178.
- Gan, M., Pan, S., Chen, Y., Cheng, C., Pan, H., & Zhu, X. (2021). Application of the machine learning lightgbm model to the prediction of the water levels of the lower columbia river. *Journal of Marine Science and Engineering*, 9(5), 496.
- Garfunkel, Z. (2002). Early Paleozoic sediments of NE Africa and Arabia: Products of continental-scale erosion, sediment transport, and deposition. *Israel Journal of Earth Sciences*, 51.
- Gebreyesus, Y., Dalton, D., Nixon, S., De Chiara, D., & Chinnici, M. (2023). Machine learning for data center optimizations: feature selection using Shapley additive exPlanation (SHAP). *Future Internet*, 15(3), 88.
- Gény, F. (2010). Can unconventional gas be a game changer in European gas markets?. Oxford Institute for Energy Studies.
- Germanovich, L. N., & Dyskin, A. V. (2000). Fracture mechanisms and instability of openings in compression. *International Journal of Rock Mechanics and Mining Sciences*, 37(1-2), 263-284.
- Ghalgaoui, M., Sebei, K., Manai, D., Boussigua, H., Dkhaili, N., Kharbachi, S., & Inoubli, M. H. (2023). Paleozoic geological evolution of Southeastern Tunisia, North Africa: New insights from subsurface data. *Marine and Petroleum Geology*, 151, 106185.
- Ghassemi, A. (2012). A review of some rock mechanics issues in geothermal reservoir development. *Geotechnical and Geological Engineering*, 30, 647-664.
- Ghienne, J. F., Abdallah, H., Deschamps, R., Guiraud, M., Gutiérrez-Marco, J. C., Konaté, M., ... & Rubino, J. L. (2023). The Ordovician record of North and West Africa: unravelling sea-level variations, Gondwana tectonics, and the glacial impact.
- Gholami, R., Moradzadeh, A., Rasouli, V., & Hanachi, J. (2014). Practical application of failure criteria in determining safe mud weight windows in drilling operations. *Journal of Rock Mechanics and Geotechnical Engineering*, 6(1), 13-25.
- Gilpin, L. H., Bau, D., Yuan, B. Z., Bajwa, A., Specter, M., & Kagal, L. (2018, October). Explaining explanations: An overview of interpretability of machine learning. In *2018 IEEE 5th International Conference on data science and advanced analytics (DSAA)* (pp. 80-89). IEEE.
- Glorioso, J. C., & Rattia, A. (2012, March). Unconventional reservoirs: basic petrophysical concepts for shale gas. In *SPE/EAGE European Unconventional Resources Conference and Exhibition* (pp. SPE-153004). Spe.
- Goldstein, R. V., & Osipenko, N. M. (2019). About compression fracture. *Physical Mesomechanics*, 22(6), 439-455.
- Gómez-Navarro, T., & Ribó-Pérez, D. (2018). Assessing the obstacles to the participation of renewable energy sources in the electricity market of Colombia. *Renewable and Sustainable Energy Reviews*, 90, 131-141.
- Gong, L., Liu, K., & Ju, W. (2023). Advances in the study of natural fractures in deep and unconventional reservoirs. *Frontiers in Earth Science*, 10, 1096643.

- Gonzalez, L., & Aguilera, R. (2013). Effect of Natural-Fracture Density on Production Variability of Individual Wells in the Nikanassin Tight Gas Formation. *Journal of Canadian Petroleum Technology*, 52(02), 131-143.
- Goodfellow, I., Bengio, Y., Courville, A., & Bengio, Y. (2016). *Deep learning* (Vol. 1, No. 2). Cambridge: MIT press.
- Gould, T., & McGlade, C. (2019). Could tight oil go global?
- Gregory, K. B., Vidic, R. D., & Dzombak, D. A. (2011). Water management challenges associated with the production of shale gas by hydraulic fracturing. *Elements*, 7(3), 181-186.
- Guangming, Z. H. A. I. (2008). Speculations on the exploration and development of unconventional hydrocarbon resources. *Natural Gas Industry*, 28(12), 1-3.
- Gudmundsson, A. (2011). *Rock fractures in geological processes*. Cambridge University Press.
- Guerriero, V., Mazzoli, S., Iannace, A., Vitale, S., Carravetta, A., & Strauss, C. (2013). A permeability model for naturally fractured carbonate reservoirs. *Marine and petroleum geology*, 40, 115-134.
- Gunter, W. D., Gentzis, T., Rottenfusser, B. A., & Richardson, R. J. H. (1997). Deep coalbed methane in Alberta, Canada: a fuel resource with the potential of zero greenhouse gas emissions. *Energy Conversion and Management*, 38, S217-S222.
- Gupta, I., Rai, C., Devegowda, D., & Sondergeld, C. H. (2021). Fracture hits in unconventional reservoirs: A critical review. *SPE Journal*, 26(01), 412-434.
- Guyon, I., & Elisseeff, A. (2003). An introduction to variable and feature selection. *Journal of Machine Learning Research*, 3, 1157-1182.
- Hacquebard, P. A. (2002). Potential coalbed methane resources in Atlantic Canada. *International Journal of Coal Geology*, 52(1-4), 3-28.
- Halim, M. C., Hamidi, H., & Akisanya, A. R. (2021). Minimizing formation damage in drilling operations: a critical point for optimizing productivity in sandstone reservoirs intercalated with clay. *Energies*, 15(1), 162.
- Ham, Y., & Kantzas, A. (2008, June). Development of coalbed methane in Australia: unique approaches and tools. In *SPE Unconventional Resources Conference/Gas Technology Symposium* (pp. SPE-114992). SPE.
- Hancock, J. T., & Khoshgoftaar, T. M. (2020). CatBoost for big data: an interdisciplinary review. *Journal of big data*, 7(1), 94.
- Harju, J. A., Malki, M. L., Rabiei, M., & Rasouli, V. (2023). Log scale rock physics modelling to study the effect of mineralogy on elastic properties of the Bakken Formation. *Geoenergy Science and Engineering*, 223, 211432.
- Hassaan, S., Mohamed, A., Ibrahim, A. F., & Elkatatny, S. (2024). Real-Time Prediction of Petrophysical Properties Using Machine Learning Based on Drilling Parameters. *ACS omega*, 9(15), 17066-17075.
- Hassija, V., Chamola, V., Mahapatra, A., Singal, A., Goel, D., Huang, K., ... & Hussain, A. (2024). Interpreting black-box models: a review on explainable artificial intelligence. *Cognitive Computation*, 16(1), 45-74.
- Hastie, T., Tibshirani, R., Friedman, J., & Franklin, J. (2005). The elements of statistical learning: data mining, inference and prediction. *The Mathematical Intelligencer*, 27(2), 83-85.
- Hastie, T., Tibshirani, R., Friedman, J., & Franklin, J. (2005). The elements of statistical learning: data mining, inference and prediction. *The Mathematical Intelligencer*, 27(2), 83-85.
- Haykin, S. (2009). *Neural networks and learning machines*, 3/E. Pearson Education India.
- He, Q., Wang, Z., Liu, C., & Wu, H. (2024). Identifying nonuniform distributions of rock properties and hydraulic fracture trajectories through deep learning in unconventional reservoirs. *Energy*, 291, 130329.

- Hein, F., Geol, P., Rokosh, D., Tan, L., Alpak, F., Ameen, M. S., ... & Asia, A. A. P. G. Tight-Gas Sands Committee Report EMD Annual Meeting, April 2012.
- Holditch, S. A. (2003). The increasing role of unconventional reservoirs in the future of the oil and gas business. *Journal of petroleum technology*, 55(11), 34-79.
- Hongjun, W. A. N. G., Feng, M. A., Xiaoguang, T. O. N. G., Zuodong, L., Zhenzhen, W. U., Denghua, L. I., ... & Liuyan, Y. A. N. G. (2016). Assessment of global unconventional oil and gas resources. *Petroleum Exploration and Development*, 43(6), 925-940.
- Hosen, M. S., & Amin, R. (2021). Significant of gradient boosting algorithm in data management system. *Eng. Int*, 9(2), 85-100.
- Hu, L. (2022). A review of mechanical mechanism and prediction of natural fracture in shale. *Arabian Journal of Geosciences*, 15(6), 474.
- Hu, L., Winterfeld, P. H., Fakcharoenphol, P., & Wu, Y. S. (2013). A novel fully-coupled flow and geomechanics model in enhanced geothermal reservoirs. *Journal of Petroleum Science and Engineering*, 107, 1-11.
- Hu, W., Hu, W., & Maybank, S. (2008). Adaboost-based algorithm for network intrusion detection. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 38(2), 577-583.
- Huynh-Thu, V. A., Saeys, Y., Wehenkel, L., & Geurts, P. (2012). Statistical interpretation of machine learning-based feature importance scores for biomarker discovery. *Bioinformatics*, 28(13), 1766-1774.
- Ilgen, A. G., Heath, J. E., Akkutlu, I. Y., Bryndzia, L. T., Cole, D. R., Kharaka, Y. K., ... & Suarez-Rivera, R. (2017). Shales at all scales: Exploring coupled processes in mudrocks. *Earth-Science Reviews*, 166, 132-152.
- Isah, A., Tariq, Z., Mustafa, A., Mahmoud, M., & Okoroafor, E. R. (2025). A Review of Data-Driven Machine Learning Applications in Reservoir Petrophysics. *Arabian Journal for Science and Engineering*, 1-35.
- Isah, A., Tariq, Z., Mustafa, A., Mahmoud, M., & Okoroafor, E. R. (2025). A Review of Data-Driven Machine Learning Applications in Reservoir Petrophysics. *Arabian Journal for Science and Engineering*, 1-35.
- Issawi, B., & Farouk, S. (2023). An overview of the phanerozoic geology in Egypt. *The Phanerozoic Geology and Natural Resources of Egypt*, 3-26.
- Jaeger, J. C., Cook, N. G. W., & Zimmerman, R. (2007). *Fundamentals of Rock Mechanics*. Wiley-Blackwell.
- Jaeger, J. C., Cook, N. G., & Zimmerman, R. (2009). *Fundamentals of rock mechanics*. John Wiley & Sons.
- Jaiswal, J. K., & Samikannu, R. (2017, February). Application of random forest algorithm on feature subset selection and classification and regression. In *2017 world congress on computing and communication technologies (WCCCT)* (pp. 65-68).Ieee.
- Jalali, M. R., & Dusseault, M. B. (2012). Coupling geomechanics and transport in naturally fractured reservoirs. *International Journal of Mining and Geo-Engineering*, 46(2), 105-131.
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning* (Vol. 112, p. 18). New York: springer.
- Jarvie, D. M. (2010, September). Unconventional oil petroleum systems: shales and shale hybrids. In *AAPG International & Conference Exchange*, Calgary, Alberta, Canada (pp. 1-21).
- Jarvie, D. M. (2012). Shale Resource Systems for Oil and Gas: Part 2  Shale-oil Resource Systems.
- Jarvie, D. M., Hill, R. J., Ruble, T. E., & Pollastro, R. M. (2007). Unconventional shale-gas systems: The Mississippian Barnett Shale of north-central Texas as one model for thermogenic shale-gas assessment. *AAPG bulletin*, 91(4), 475-499.

- Jarvie, D. M., Hill, R. J., Ruble, T. E., & Pollastro, R. M. (2007). Unconventional shale-gas systems: The Mississippian Barnett Shale of north-central Texas as one model for thermogenic shale-gas assessment. *AAPG bulletin*, 91(4), 475-499.
- Jiang, J., & Yang, J. (2018). Coupled fluid flow and geomechanics modeling of stress-sensitive production behavior in fractured shale gas reservoirs. *International Journal of Rock Mechanics and Mining Sciences*, 101, 1-12.
- Jones Jr, F. O. (1975). A laboratory study of the effects of confining pressure on fracture flow and storage capacity in carbonate rocks. *Journal of Petroleum Technology*, 27(01), 21-27.
- Joshi, D., Prajapati, P., Sharma, P., & Sharma, A. (2023). Past, present and future of Coal Bed Methane (CBM): a review with special focus on the Indian scenario. *International Journal of Coal Preparation and Utilization*, 43(2), 377-402.
- Ju, Y., Sun, G., Chen, Q., Zhang, M., Zhu, H., & Rehman, M. U. (2019). A model combining convolutional neural network and LightGBM algorithm for ultra-short-term wind power forecasting. *Ieee Access*, 7, 28309-28318.
- Kaced, M., & Arab, M. (2012). The potential of shale gas plays in Algeria. In *25th World Gas Conference*, Kuala Lumpur.
- Kaczmarczyk, W., & Słota-Valim, M. (2020). Multidisciplinary Characterization of Unconventional Reservoirs Based on Correlation of Well and Seismic Data. *Energies*, 13(17), 4413.
- Kaelbling, L. P., Littman, M. L., & Moore, A. W. (1996). Reinforcement learning: A survey. *Journal of artificial intelligence research*, 4, 237-285.
- Kahloul, L., Chadli, H., Boudebane, S., & Iost, A. (2019). Performances tribologique et électrochimique de l'alliage TiNi poreux. *Synthèse: Revue des Sciences et de la Technologie*, 25(1), 109-124.
- Kaneko, H. (2023). Interpretation of machine learning models for data sets with many features using feature importance. *ACS omega*, 8(25), 23218-23225.
- Kaur, R., Kumar, R., & Aggarwal, H. (2025). Systematic Review of Artificial Intelligence, Machine Learning, and Deep Learning in Machining Operations: Advancements, Challenges, and Future Directions. *Archives of Computational Methods in Engineering*, 1-54.
- Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., ... & Liu, T. Y. (2017). Lightgbm: A highly efficient gradient boosting decision tree. *Advances in neural information processing systems*, 30.
- Kenzie, T. A. (2024). Natural Fracture Systems in Outcrop of the Lower Triassic Sulphur Mountain Formation: Investigation of Fracture Timing Using Structural, Petrographic, and Geochemical Datasets.
- Khan, M. Y., & Mandal, A. (2022). The impact of permeability heterogeneity on water-alternating-gas displacement in highly stratified heterogeneous reservoirs. *Journal of Petroleum Exploration and Production Technology*, 12(3), 871-897.
- Kharrat, R., & Ott, H. (2023). A comprehensive review of fracture characterization and its impact on oil production in naturally fractured reservoirs. *Energies*, 16(8), 3437.
- Khodabakhshi, M. J., & Bijani, M. (2024). Predicting scale deposition in oil reservoirs using machine learning optimization algorithms. *Results in Engineering*, 22, 102263.
- Kianoush, P., Mohammadi, G., Hosseini, S. A., Keshavarz Faraj Khah, N., & Afzal, P. (2023). ANN-based estimation of pore pressure of hydrocarbon reservoirs—a case study. *Arabian Journal of Geosciences*, 16(5), 302.
- Kim, Y. M., Jang, H. C., & Lee, J. H. (2016). Application of probabilistic approach to evaluate coalbed methane resources using geological data of coal basin in Indonesia. *Geosciences Journal*, 20, 229-238.
- King, G. E. (2012, February). Hydraulic fracturing 101: What every representative, environmentalist, regulator, reporter, investor, university researcher, neighbor and engineer should know about

- estimating frac risk and improving frac performance in unconventional gas and oil wells. In SPE Hydraulic Fracturing Technology Conference and Exhibition (pp. SPE-152596). SPE.
- KING, P. L. C. (1958). Basic palaeogeography of Gondwanaland during the late Palaeozoic and Mesozoic eras. *Quarterly Journal of the Geological Society*, 114(1-4), 47-70.
- Kjärstad, J., & Johnsson, F. (2009). Resources and future supply of oil. *Energy policy*, 37(2), 441-464.
- Kogbe, C. A., & Burolet, P. F. (1990). A review of continental sediments in Africa. *Journal of African Earth Sciences (and the Middle East)*, 10(1-2), 1-25.
- Kolawole, F., & Evenick, J. C. (2023). Global distribution of geothermal gradients in sedimentary basins. *Geoscience Frontiers*, 14(6), 101685.
- Kundu, S. N. (2023). *Geoscience for Petroleum Engineers*. Springer.
- Kundu, S. N. (2023). Unconventional Hydrocarbons Resources. In *Geoscience for Petroleum Engineers* (173-192) Springer.
- Labenski, F., Reid, P., & Santos, H. (2003, October). Drilling fluids approaches for control of wellbore instability in fractured formations. In SPE/IADC Middle East Drilling Technology Conference and Exhibition (pp. SPE-85304). SPE.
- Lai, J., Wang, G., Ran, Y., Zhou, Z., & Cui, Y. (2016). Impact of diagenesis on the reservoir quality of tight oil sandstones: The case of Upper Triassic Yanchang Formation Chang 7 oil layers in Ordos Basin, China. *Journal of Petroleum Science and Engineering*, 145, 54-65.
- Lai, J., Wang, G., Wang, S., Cao, J., Li, M., Pang, X., ... & Qin, Z. (2018). Review of diagenetic facies in tight sandstones: Diagenesis, diagenetic minerals, and prediction via well logs. *Earth-Science Reviews*, 185, 234-258.
- Lake, L. W., Johns, R., Rossen, B., & Pope, G. A. (2014). *Fundamentals of enhanced oil recovery* (Vol. 1, p. 1). Richardson, TX: Society of Petroleum Engineers.
- Lang, P. S., Paluszny, A., Nejati, M., & Zimmerman, R. W. (2018). Relationship between the orientation of maximum permeability and intermediate principal stress in fractured rocks. *Water Resources Research*, 54(11), 8734-8755.
- Lash, G. G., & Engelder, T. (2011). Thickness trends and sequence stratigraphy of the Middle Devonian Marcellus Formation, Appalachian Basin: Implications for Acadian foreland basin evolution. *AAPG bulletin*, 95(1), 61-103.
- Lash, G. G., & Engelder, T. (2011). Thickness trends and sequence stratigraphy of the Middle Devonian Marcellus Formation, Appalachian Basin: Implications for Acadian foreland basin evolution. *AAPG bulletin*, 95(1), 61-103.
- Laubach, S. E., Olson, J. E., & Gross, M. R. (2009). Mechanical and fracture stratigraphy. *AAPG bulletin*, 93(11), 1413-1426.
- Lawal, A., Yang, Y., He, H., & Baisa, N. L. (2024). Machine learning in oil and gas exploration: A review. *Ieee Access*, 12, 19035-19058.
- Layachi, A. (2013). The changing geopolitics of natural gas: the case of Algeria.
- Lazaruk, Y., & Karabyn, V. (2020). Shale gas in Western Ukraine: Perspectives, resources, environmental and technogenic risk of production.
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *nature*, 521(7553), 436-444.
- Li, D., & Du, Y. (2017). *Artificial intelligence with uncertainty*. CRC press.
- Li, F., Zhang, Y., Liu, J., Zhang, L., & Fang, S. (2016). The dynamical response characteristics of elastic-plastic coal under dynamic load. *Journal of Natural Gas Science and Engineering*, 29, 497-505.
- Li, J. (2017). Assessing the accuracy of predictive models for numerical data: Not r nor r2, why not? Then what?. *PloS one*, 12(8), e0183250.
- Li, S., Li, S., Guo, R., Zhou, X., Wang, Y., Chen, J., ... & Qiu, J. (2021). Occurrence state of soluble organic matter in shale oil reservoirs from the Upper Triassic Yanchang formation in the Ordos

- Basin, China: insights from multipolarity sequential extraction. *Natural Resources Research*, 30(6), 4379-4402.
- Li, X. F., Li, X., Li, H. B., Zhang, Q. B., & Zhao, J. (2018). Dynamic tensile behaviours of heterogeneous rocks: the grain scale fracturing characteristics on strength and fragmentation. *International Journal of Impact Engineering*, 118, 98-118.
- Li, Z., Li, L., Li, M., Zhang, L., Zhang, Z., Huang, B., & Tang, C. A. (2018). A numerical investigation on the effects of rock brittleness on the hydraulic fractures in the shale reservoir. *Journal of Natural Gas Science and Engineering*, 50, 22-32.
- Lian, Z., Zhang, Q., Lin, T., & Wang, F. (2015). Experimental and numerical study of drill string dynamics in gas drilling of horizontal wells. *Journal of Natural Gas Science and Engineering*, 27, 1412-1420.
- Liang, W., Luo, S., Zhao, G., & Wu, H. (2020). Predicting hard rock pillar stability using GBDT, XGBoost, and LightGBM algorithms. *Mathematics*, 8(5), 765.
- Liaw, A., & Wiener, M. (2002). Classification and regression by randomForest. *R news*, 2(3), 18-22.
- Liégeois, J. P., Latouche, L., Boughrara, M., Navez, J., & Guiraud, M. (2003). The LATEA metacraton (Central Hoggar, Tuareg shield, Algeria): behaviour of an old passive margin during the Pan-African orogeny. *Journal of African Earth Sciences*, 37(3-4), 161-190.
- Linardatos, P., Papastefanopoulos, V., & Kotsiantis, S. (2020). Explainable ai: A review of machine learning interpretability methods. *Entropy*, 23(1), 18.
- Linol, B., de Wit, M. J., Milani, E. J., Guillocheau, F., & Scherer, C. (2014). New regional correlations between the Congo, Paraná and Cape-Karoo basins of southwest Gondwana. In *Geology and resource potential of the Congo Basin* (pp. 245-268). Berlin, Heidelberg: Springer Berlin Heidelberg.
- Liu, H. (2017). *Principles and applications of well logging* (pp. 237-269). Berlin, Heidelberg: Springer Berlin Heidelberg.
- Liu, S., Yang, Y., Deng, B., Zhong, Y., Wen, L., Sun, W., ... & Peng, H. (2021). Tectonic evolution of the Sichuan Basin, southwest China. *Earth-Science Reviews*, 213, 103470.
- Liu, Y., Zhu, R., Zhai, S., Li, N., & Li, C. (2023). Lithofacies identification of shale formation based on mineral content regression using LightGBM algorithm: A case study in the Luzhou block, South Sichuan Basin, China. *Energy Science & Engineering*, 11(11), 4256-4272.
- Lu, Y. (2019). Artificial intelligence: a survey on evolution, models, applications and future trends. *Journal of management analytics*, 6(1), 1-29.
- Lučić, D., & Bosworth, W. (2018). Regional geology and petroleum systems of the main reservoirs and source rocks of North Africa and the Middle East. In *The Geology of the Arab World---An Overview* (pp. 197-289). Cham: Springer International Publishing.
- Macgregor, D. S., & Reeves, C. V. (2025). A Paleotectonic Atlas of the African Plate: Permian to Recent. *Journal of Petroleum Geology*, 48(4), 231-279.
- Maddaloni, F. (2021). LITHOSPHERIC STRUCTURE AND TECTONIC EVOLUTION OF INTRACRATONIC BASINS: THE CONGO BASIN A NATURAL LABORATORY.
- Mahmoud, A. A., Gamal, H., Elkatatny, S., & Chen, W. (2022). Real-time evaluation of the dynamic Young's modulus for composite formations based on the drilling parameters using different machine learning algorithms. *Frontiers in Earth Science*, 10, 1034704.
- Mahmoud, A. A., Gamal, H., Mutrif, O., & Elkatatny, S. (2021, June). Artificial neural networks-based equation for real-time estimation of the dynamic Young's modulus. In *ARMA US Rock Mechanics/Geomechanics Symposium* (pp. ARMA-2021). ARMA.
- Malozyomov, B. V., Martyushev, N. V., Kukartsev, V. V., Tynchenko, V. S., Bukhtoyarov, V. V., Wu, X., ... & Kukartsev, V. A. (2023). Overview of methods for enhanced oil recovery from conventional and unconventional reservoirs. *Energies*, 16(13), 4907.

- Maltman, A. (Ed.). (2012). The geological deformation of sediments. Springer Science & Business Media.
- Mangalathu, S., Hwang, S. H., & Jeon, J. S. (2020). Failure mode and effects analysis of RC members based on machine-learning-based SHapley Additive exPlanations (SHAP) approach. *Engineering Structures*, 219, 110927.
- Mann, P., Gahagan, L., & Gordon, M. B. (2003). Tectonic setting of the world's giant oil and gas fields.
- Manzocchi, T. (2002). The connectivity of two-dimensional networks of spatially correlated fractures. *Water Resources Research*, 38(9), 1-1.
- Marr, B. (2019). Artificial intelligence in practice: how 50 successful companies used AI and machine learning to solve problems. John Wiley & Sons.
- Martins, C. M. S., Celino, J. J., Cerqueira, J. R., Garcia, K. S., & Queiroz, A. F. D. S. (2023). Thermal effect of igneous intrusions on organic-rich Irati Formation and the implications for petroleum systems: a case study in the Paraná Basin, Brazil. *Brazilian Journal of Geology*, 53(1), e20220008.
- Mastalerz, M., & Drobniak, A. (2020). Coalbed methane: Reserves, production, and future outlook. In *Future energy* (pp. 97-109). Elsevier.
- Mavko, G., Mukerji, T., & Dvorkin, J. (2020). The rock physics handbook. Cambridge university press.
- Mayer, B., Klassen, P., Cheung, K., Taylor, S., & Shevalier, M. (2007). Isotopic approaches for monitoring potential contamination of shallow aquifers with produced fluids or gases from coalbed methane operations in Alberta, Canada. *Advances in Isotope Hydrology and its Role in Sustainable Water Resources Management (IHS—2007)*, 21, 519.
- McCarthy, J., Minsky, M. L., Rochester, N., & Shannon, C. L. (1956). The Dartmouth summer research project on artificial intelligence. *Artificial intelligence: past, present, and future*.
- McCarthy, K., Rojas, K., Niemann, M., Palmowski, D., Peters, K., & Stankiewicz, A. (2011). Basic petroleum geochemistry for source rock evaluation. *Oilfield Review*, 23(2), 32-43.
- Milad, M., Junin, R., Sidek, A., Imqam, A., & Tarhuni, M. (2021). Huff-n-puff technology for enhanced oil recovery in shale/tight oil reservoirs: Progress, gaps, and perspectives. *Energy & Fuels*, 35(21), 17279-17333.
- Milani, E. J., & De Wit, M. J. (2008). Correlations between the classic Paraná and Cape-Karoo sequences of South America and southern Africa and their basin infills flanking the Gondwanides: du Toit revisited.
- Milani, E. J., & Zalán, P. V. (1999). An outline of the geology and petroleum systems of the Paleozoic interior basins of South America. *Episodes Journal of International Geoscience*, 22(3), 199-205.
- Minisini, D., Fryklund, B., Gerali, F., & Fantín, M. (2020). The first economical unconventional play outside North America: Context, history, and “coopetition”.
- Minkoff, S. E., Stone, C. M., Bryant, S., Peszynska, M., & Wheeler, M. F. (2003). Coupled fluid flow and geomechanical deformation modeling. *Journal of Petroleum Science and Engineering*, 38(1-2), 37-56.
- Mohammadinia, F., Ranjbar, A., Kafi, M., & Keshavarz, R. (2023). Application of machine learning algorithms in classification the flow units of the Kazhdumi reservoir in one of the oil fields in southwest of Iran. *Journal of Petroleum Exploration and Production Technology*, 13(6), 1419-1434.
- Molnar, P., Anderson, R. S., & Anderson, S. P. (2007). Tectonics, fracturing of rock, and erosion. *Journal of Geophysical Research: Earth Surface*, 112(F3).
- Moore, T. A. (2012). Coalbed methane: a review. *International Journal of Coal Geology*, 101, 36-81.
- Muggeridge, A., Cockin, A., Webb, K., Frampton, H., Collins, I., Moulds, T., & Salino, P. (2014). Recovery rates, enhanced oil recovery and technological limits. *Philosophical Transactions of the Royal Society A : Mathematical, Physical and Engineering Sciences*, 372(2006), 20120320.
- Müller, M. E., Nazarian, S., Koch, P., & Schatzker, J. (2012). The comprehensive classification of fractures of long bones. Springer Science & Business Media.

- Musolf, A. M., Holzinger, E. R., Malley, J. D., & Bailey-Wilson, J. E. (2022). What makes a good prediction? Feature importance and beginning to open the black box of machine learning in genetics. *Human Genetics*, 141(9), 1515-1528.
- Muthur, T., Qureshi, H. A., Syed, F. I., Aziz, H., Siyal, A., Dahaghi, A. K., & Negahban, S. (2022). Unconventional hydrocarbon resources: geological statistics, petrophysical characterization, and field development strategies. *Journal of Petroleum Exploration and Production Technology*, 12(6), 1463-1488.
- Mutrif Siddig, O., & Elkatatny, S. (2023, March). Utilizing Drilling Data and Machine Learning in Real-Time Prediction of Poisson's Ratio. In *SPE Middle East Oil and Gas Show and Conference* (p. D021S073R006). SPE.
- Nagel, N. B., & Sanchez-Nagel, M. A. (2015, June). On the importance and impact of key geomechanical parameters in unconventional play developments. In *ARMA US Rock Mechanics/Geomechanics Symposium* (pp. ARMA-2015). ARMA.
- Nagel, N. B., Sanchez-Nagel, M., & Lee, B. (2012). Gas shale hydraulic fracturing: a numerical evaluation of the effect of geomechanical parameters. In *SPE Hydraulic Fracturing Technology Conference and Exhibition* (pp. SPE-152192). SPE.
- Naghibi, S. A., Hashemi, H., Berndtsson, R., & Lee, S. (2020). Application of extreme gradient boosting and parallel random forest algorithms for assessing groundwater spring potential using DEM-derived factors. *Journal of Hydrology*, 589, 125197.
- Narimani, S., Davarpanah, S. M., & Vásárhelyi, B. (2024). Estimation of the Poisson's Ratio of the Rock Mass. *Periodica Polytechnica Civil Engineering*, 68(1), 274-288.
- Natekin, A., and A. Knoll. 2013. Gradient boosting machines, a tutorial. *Frontiers in Neuroinformatics* 7:21.
- Nelson, R. A. (1985). *Geologic analysis of naturally fractured reservoirs* (Vol. 1). Gulf Professional Publishing.
- Neuendorf, K. K. (2005). *Glossary of geology*. Springer Science & Business Media.
- Nohara, Y., Matsumoto, K., Soejima, H., & Nakashima, N. (2019, September). Explanation of machine learning models using improved shapley additive explanation. In *Proceedings of the 10th ACM international conference on bioinformatics, computational biology and health informatics* (pp. 546-546).
- Norvig, P., & Russell, S. J. (1995). *Artificial intelligence a modern approach*.
- Odling, N. E., Gillespie, P., Bourguine, B., Castaing, C., Chiles, J. P., Christensen, N. P., ... & Watterson, J. (1999). Variations in fracture system geometry and their implications for fluid flow in fractures hydrocarbon reservoirs. *Petroleum Geoscience*, 5(4), 373-384.
- Ogg, J. G. (2015). The mysterious mid-Carnian “Wet Intermezzo” global event. *Journal of Earth Science*, 26(2), 181-191.
- Olsen, P. E. (1997). Stratigraphic record of the early Mesozoic breakup of Pangea in the Laurasia-Gondwana rift system. *Annual Review of Earth and Planetary Sciences*, 25(1), 337-401.
- Onalo, D., Oloruntobi, O., Adedigba, S., Khan, F., James, L., & Butt, S. (2018). Static Young's modulus model prediction for formation evaluation. *Journal of Petroleum Science and Engineering*, 171, 394-402.
- Opportunities and challenges to coal bed methane production in Australia. *Journal of Petroleum Science and Engineering*, 88, 1-4.
- Orujov, A., Coddington, K., & Aryana, S. A. (2023). A review of CCUS in the context of foams, regulatory frameworks and monitoring. *Energies*, 16(7), 3284.
- Pan, S. J., & Yang, Q. (2009). A survey on transfer learning. *IEEE Transactions on knowledge and data engineering*, 22(10), 1345-1359.

- Papon, P. (2025). AIE (Agence internationale de l'énergie). World Energy Outlook 2024, Paris: AIE, octobre 2024, 398 p. URL: <https://iea.blob.core.windows.net/assets/6a25abba-1973-4580-b6e3-ba014a81b458/WorldEnergyOutlook2024.pdf>. Consulté le 28 novembre 2024. *Futuribles*, 464(1), 140-143.
- Paterson, M. S., & Wong, T. F. (2005). *Experimental rock deformation—the brittle field*. Berlin, Heidelberg: Springer Berlin Heidelberg.
- Paul, A., Mukherjee, S., & Biswas, M. sequestration, *Results in Earth Sciences*, (2025).
- Payne, M. L., & Abbassian, F. (1997). Advanced Torque-and-Drag Considerations in Extended-Reach Wells. *SPE Drilling & Completion*, 12(01), 55-62.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. *the Journal of machine Learning research*, 12, 2825-2830.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O. & Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *the Journal of machine Learning research*, 12, 2825-2830.
- Popa, A. S., & Cassidy, S. (2012, October). Artificial intelligence for heavy oil assets: the evolution of solutions and organization capability. In *SPE Annual Technical Conference and Exhibition?* (pp. SPE-159504). SPE.
- Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A. V., & Gulin, A. (2018). CatBoost: unbiased boosting with categorical features. *Advances in neural information processing systems*, 31.
- Qiu, Y., Zhou, J., Khandelwal, M., Yang, H., Yang, P., & Li, C. (2022). Performance evaluation of hybrid WOA-XGBoost, GWO-XGBoost and BO-XGBoost models to predict blast-induced ground vibration. *Engineering with Computers*, 38(Suppl 5), 4145-4162.
- Questiaux, J. M., Couples, G. D., & Ruby, N. (2010). Fractured reservoirs with fracture corridors. *Geophysical Prospecting*, 58(2), 279-295.
- Qun, L. E. I., Yun, X. U., Bo, C. A. I., Baoshan, G., Xin, W. A. N. G., Guoqiang, B. I., ... & Tao, L. (2022). Progress and prospects of horizontal well fracturing technology for shale oil and gas reservoirs. *Petroleum Exploration and Development*, 49(1), 191-199.
- Radwan, A. E. (2024). The north African neoproterozoic, phanerozoic sedimentary basins: tectonostratigraphic events, Lithostratigraphy, Petroleum systems. In *The Geology of North Africa* (pp. 221-249). Cham: Springer International Publishing.
- Raimi, D., Zhu, Y., Newell, R. G., Prest, B. C., & Bergman, A. (2023). Global energy outlook 2023: sowing the seeds of an energy transition. *Resources for the future*, 1(1), 1-44.
- Rashid, F., Hussein, D., Lorinczi, P., & Glover, P. W. J. (2023). The effect of fracturing on permeability in carbonate reservoir rocks. *Marine and Petroleum Geology*, 152, 106240.
- Ravi, K., Bosma, M., & Gastebled, O. (2002, February). Improve the economics of oil and gas wells by reducing the risk of cement failure. In *SPE/IADC Drilling Conference and Exhibition* (pp. SPE-74497). SPE.
- Reading, H. G. (Ed.). (2009). *Sedimentary environments: processes, facies and stratigraphy*. John Wiley & Sons.
- Recent advances in decision trees: an updated survey. *Artificial Intelligence Review*, 56(5), 4765-4800.
- Rengasamy, D., Mase, J. M., Kumar, A., Rothwell, B., Torres, M. T., Alexander, M. R., ... & Figueredo, G. P. (2022). Feature importance in machine learning models: A fuzzy information fusion approach. *Neurocomputing*, 511, 163-174.
- Rocha, H. D. V., Mendes, M., Pereira, Z., Rodrigues, C., Fernandes, P., Lopes, G., ... & de Sousa, M. L. (2020). New palynostratigraphic data of the Irati (Assistência Member) and the Corumbataí formations, Paraná Basin, Brazil, and correlation with other south American basins. *Journal of South American Earth Sciences*, 102, 102631.

- Rogner, H. H. (1997). An assessment of world hydrocarbon resources. *Annual review of energy and the environment*, 22(1), 217-262.
- Rosa, L., & D'Odorico, P. (2019). The water-energy-food nexus of unconventional oil and gas extraction in the Vaca Muerta Play, Argentina. *Journal of cleaner production*, 207, 743-750.
- Rosenblatt, F. (1958). The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological review*, 65(6), 386.
- Ross, D. J., & Bustin, R. M. (2009). The importance of shale composition and pore structure upon gas storage potential of shale gas reservoirs. *Marine and petroleum Geology*, 26(6), 916-927.
- Rossi, C., Kálin, O., Arribas, J., & Tortosa, A. (2002). Diagenesis, provenance and reservoir quality of Triassic TAGI sandstones from Ourhoud field, Berkine (Ghadames) Basin, Algeria. *Marine and petroleum geology*, 19(2), 117-142.
- Roure, F., Swennen, R., Schneider, F., Faure, J. L., Ferket, H., Guilhaumou, N., ... & Vandeginste, V. (2005). Incidence and importance of tectonics and natural fluid migration on reservoir evolution in foreland fold-and-thrust belts. *Oil & Gas Science and Technology*, 60(1), 67-106.
- Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *nature*, 323(6088), 533-536.
- Ruse, C. M. (2024). *Natural Fracture and Pore Space Characterization in the Tuscaloosa Marine Shale*. University of Louisiana at Lafayette.
- Sahimi, M. (2011). *Flow and transport in porous media and fractured rock: from classical methods to modern approaches*. John Wiley & Sons.
- Sahin, E. K. (2022). Comparative analysis of gradient boosting algorithms for landslide susceptibility mapping. *Geocarto International*, 37(9), 2441-2465.
- Sahour, H., Gholami, V., Torkaman, J., Vazifedan, M., & Saeedi, S. (2021). Random forest and extreme gradient boosting algorithms for streamflow modeling using vessel features and tree-rings. *Environmental Earth Sciences*, 80(22), 747.
- Salih, A. A., & Abdulazeez, A. M. (2021). Evaluation of classification algorithms for intrusion detection system: A review. *Journal of Soft Computing and Data Mining*, 2(1), 31-40.
- Salman, H. A., Kalakech, A., & Steiti, A. (2024). Random forest algorithm overview. *Babylonian Journal of Machine Learning*, 2024, 69-79.
- Samuel, A. L. (1959). Some studies in machine learning using the game of checkers. *IBM Journal of research and development*, 3(3), 210-229.
- Samuel, A. L. (2000). Some studies in machine learning using the game of checkers. *IBM Journal of research and development*, 44(1.2), 206-226.
- Sander, R., Pan, Z., & Connell, L. D. (2017). Laboratory measurement of low permeability unconventional gas reservoir rocks: A review of experimental methods. *Journal of Natural Gas Science and Engineering*, 37, 248-279.
- Sanei, M., Ramezanzadeh, A., & Delavar, M. R. (2023). Applied machine learning-based models for predicting the geomechanical parameters using logging data. *Journal of Petroleum Exploration and Production Technology*, 13(12), 2363-2385.
- Sanei, M., Ramezanzadeh, A., & Delavar, M. R. (2023). Applied machine learning-based models for predicting the geomechanical parameters using logging data. *Journal of Petroleum Exploration and Production Technology*, 13(12), 2363-2385.
- Sarica, A., Cerasa, A., & Quattrone, A. (2017). Random forest algorithm for the classification of neuroimaging data in Alzheimer's disease: a systematic review. *Frontiers in aging neuroscience*, 9, 329.
- Schapiro, R. E. (2013). Explaining adaboost. In *Empirical inference: festschrift in honor of vladimir N. Vapnik* (pp. 37-52). Berlin, Heidelberg: Springer Berlin Heidelberg.

- Schonlau, M., & Zou, R. Y. (2020). The random forest algorithm for statistical learning. *The Stata Journal*, 20(1), 3-29.
- Schultz, R., Skoumal, R. J., Brudzinski, M. R., Eaton, D., Baptie, B., & Ellsworth, W. (2020). Hydraulic fracturing-induced seismicity. *Reviews of Geophysics*, 58(3), e2019RG000695.
- Scotese, C. R. (2021). An atlas of Phanerozoic paleogeographic maps: the seas come in and the seas go out. *Annual Review of Earth and Planetary Sciences*, 49(1), 679-728.
- Segall, P., & Fitzgerald, S. D. (1998). A note on induced stress changes in hydrocarbon and geothermal reservoirs. *Tectonophysics*, 289(1-3), 117-128.
- Settari, A. (2002). Reservoir compaction. *Journal of petroleum technology*, 54(08), 62-69.
- Settari, A., & Walters, D. A. (2001). Advances in coupled geomechanical and reservoir modeling with applications to reservoir compaction. *Spe Journal*, 6(03), 334-342.
- Shang, J., Dong, Z., Tan, W., Zhang, Y., Liang, T., Xing, L., & Wang, Z. (2024). A review of fracturing and enhanced recovery integration working fluids in tight reservoirs. *Processes*, 12(6), 1241.
- Sharma, S., Saxena, A., & Saxena, N. (2019). *Unconventional resources in India: the way ahead*. Switzerland: Springer.
- Shuichang, Z., Baomin, Z., Benliang, L., Guangyou, Z., Jin, S., & Xiaomei, W. (2011). History of hydrocarbon accumulations spanning important tectonic phases in marine sedimentary basins of China: Taking the Tarim Basin as an example. *Petroleum Exploration and Development*, 38(1), 1-15.
- Siddig, O., & Elkatatny, S. (2021). Workflow to build a continuous static elastic moduli profile from the drilling data using artificial intelligence techniques. *Journal of Petroleum Exploration and Production Technology*, 11(10), 3713-3722.
- Siddig, O., Gamal, H., Elkatatny, S., & Abdulraheem, A. (2021). Real-time prediction of Poisson's ratio from drilling parameters using machine learning tools. *Scientific Reports*, 11(1), 12611.
- Sieminski, A., & Administrator, U. (2016). *Energy information administration*. International Energy Outlook.
- Singh, A. K., & Hajra, P. N. (2018). *Coalbed methane in India: Opportunities, issues and challenges for recovery and utilization*. Springer.
- Smith, D. C., Richards, J. M., & Colwell, R. J. (2017). Where 'shale' we go from here: opportunities and challenges in shale plays located outside the USA. *The Journal of World Energy Law & Business*, 10(3), 159-219.
- Snow, D. T. (1968). Rock fracture spacings, openings, and porosities. *Journal of the Soil Mechanics and Foundations Division*, 94(1), 73-91.
- Soeder, D. J. (2018). The successful development of gas and oil resources from shales in North America. *Journal of Petroleum Science and Engineering*, 163, 399-420.
- Solomatine, D. P., & Shrestha, D. L. (2004, July). AdaBoost. RT: a boosting algorithm for regression problems. In 2004 IEEE international joint conference on neural networks (IEEE Cat. No. 04CH37541) (Vol. 2, pp. 1163-1168). IEEE.
- Sone, H., & Zoback, M. D. 2013. Mechanical properties of shale-gas reservoir rocks—Part 1: Static and dynamic elastic properties and anisotropy. *Geophysics* 78(5): D381–D392.
- Sontrach/Schlumberger, 2007 : Algérie WEC 2007, French version.
- Soua, M. (2014). Paleozoic oil/gas shale reservoirs in southern Tunisia: an overview. *Journal of African Earth Sciences*, 100, 450-492.
- Spiess, A. N., & Neumeyer, N. (2010). An evaluation of R2 as an inadequate measure for nonlinear models in pharmacological and biochemical research: a Monte Carlo approach. *BMC pharmacology*, 10(1), 6.
- Staff, P. E. S. A. (2024). *Global Gas Outlook 2050*.

- Stambouli, A. B. (2011). Promotion of renewable energies in Algeria: Strategies and perspectives. *Renewable and sustainable energy reviews*, 15(2), 1169-1181.
- Stampfli, G. M., Hochard, C., V  rard, C., & Wilhem, C. (2013). The formation of Pangea. *Tectonophysics*, 593, 1-19.
- Stearns, D. W., & Friedman, M. (1972). Reservoirs in fractured rock.
- Stulberg, A. N. (2015). Out of gas?: Russia, Ukraine, Europe, and the changing geopolitics of natural gas. *Problems of Post-Communism*, 62(2), 112-130.
- Sutton, R. S., & Barto, A. G. (1998). Reinforcement learning: An introduction (Vol. 1, No. 1, pp. 9-11). Cambridge: MIT press.
- Swarbrick, R. E., & Osborne, M. J. (1998). Mechanisms that generate abnormal pressures: an overview.
- Syed, F. I., AlShamsi, A., Dahaghi, A. K., & Neghabhan, S. (2022). Application of ML & AI to model petrophysical and geomechanical properties of shale reservoirs—A systematic literature review. *Petroleum*, 8(2), 158-166.
- Syed, F. I., AlShamsi, A., Dahaghi, A. K., & Neghabhan, S. (2022). Application of ML & AI to model petrophysical and geomechanical properties of shale reservoirs—A systematic literature review. *Petroleum*, 8(2), 158-166.
- Syed, F. I., Dahaghi, A. K., & Muther, T. (2022). Laboratory to field scale assessment for EOR applicability in tight oil reservoirs. *Petroleum Science*, 19(5), 2131-2149.
- Szabo N P 2011 Shale volume estimation based on the factor analysis of well-logging data; *Acta Geophys*. 59(5) 935–953.
- Tabatabaie, S. H., & Pooladi-Darvish, M. (2017). Multiphase linear flow in tight oil reservoirs. *SPE Reservoir Evaluation & Engineering*, 20(01), 184-196.
- Taghavinejad, A., Ostadhassan, M., & Daneshfar, R. (2021). Unconventional reservoirs: rate and pressure transient analysis techniques: a reservoir engineering approach. Springer Nature.
- Takbiri-Borujeni, A., Fathi, E., Kazemi, M., & Belyadi, F. (2019). An integrated multiscale model for gas storage and transport in shale reservoirs. *Fuel*, 237, 1228-1243.
- Taleghani, A. D., Gonzalez, M., & Shojaei, A. (2016). Overview of numerical models for interactions between hydraulic fractures and natural fractures: Challenges and limitations. *Computers and Geotechnics*, 71, 361-368.
- Tang, H., Liang, H., Zhang, L., & Li, H. (2022). Fully 3D simulation of hydraulic fracture propagation in naturally fractured reservoirs using displacement discontinuity method. *SPE Journal*, 27(03), 1648-1670.
- Tang, M., Zhao, Q., Ding, S. X., Wu, H., Li, L., Long, W., & Huang, B. (2020). An improved lightGBM algorithm for online fault detection of wind turbine gearboxes. *Energies*, 13(4), 807.
- Tariq, Z., Abdulraheem, A., Mahmoud, M., & Ahmed, A. (2018). A rigorous data-driven approach to predict Poisson's ratio of carbonate rocks using a functional network. *Petrophysics*, 59(06), 761-777.
- Tariq, Z., Elkatatny, S. M., Mahmoud, M. A., Abdulraheem, A., Abdelwahab, A. Z., & Woldeamanuel, M. (2017). Estimation of rock mechanical parameters using artificial intelligence tools. In *ARMA US Rock Mechanics/Geomechanics Symposium* (pp. ARMA-2017). ARMA.
- Tenaglia, M., Eberli, G. P., Weger, R. J., Blanco, L. R., S  nchez, L. E. R., & Swart, P. K. (2020). Total organic carbon quantification from wireline logging techniques: A case study in the Vaca Muerta Formation, Argentina. *Journal of Petroleum Science and Engineering*, 194, 107489.
- Thongsuwan, S., Jaiyen, S., Padcharoen, A., & Agarwal, P. (2021). ConvXGB: A new deep learning model for classification problems based on CNN and XGBoost. *Nuclear Engineering and Technology*, 53(2), 522-531.
- Tiab, D., & Donaldson, E. C. (2024). *Petrophysics: theory and practice of measuring reservoir rock and fluid transport properties*. Elsevier.

- Tolmachev, O., Urunov, A., Muminova, S., Dvoichenkova, G., & Davydov, I. (2020). Review of unconventional hydrocarbon resources: production technologies and opportunities for development. *Mining of mineral deposits*.
- Troudi, H., Chevalier, F., Alouani, W., Mzoughi, W., & Abdelkader, O. (2018, November). Insight on the exploration potential of the ordovician gas play in Tunisia Ghadames Basin, North Africa. In *Abu Dhabi International Petroleum Exhibition and Conference* (p. D011S013R003). SPE.
- Ukar, E., Gale, J. F., Fall, A., López, R. G., Hryb, D., Manceda, R., ... & Cobbold, P. R. (2020). Natural fractures: From core and outcrop observations to subsurface models.
- van Golf-Racht, T. D. (1982). *Fundamentals of fractured reservoir engineering* (Vol. 12). Elsevier.
- Van Wyk, A. (2023). *Machine Learning with LightGBM and Python: A practitioner's guide to developing production-ready machine learning systems*. Packt Publishing Ltd.
- Vapnik, V. (2013). *The nature of statistical learning theory*. Springer science & business media.
- Varoquaux, G., Buitinck, L., Louppe, G., Grisel, O., Pedregosa, F., & Mueller, A. (2015). Scikit-learn: Machine learning without learning the machinery. *GetMobile: Mobile Computing and Communications*, 19(1), 29-33.
- Vedachalam, N., Srinivasalu, S., Rajendran, G., Ramadass, G. A., & Atmanand, M. A. (2015). Review of unconventional hydrocarbon resources in major energy consuming countries and efforts in realizing natural gas hydrates as a future source of energy. *Journal of Natural Gas Science and Engineering*, 26, 163-175.
- Vergés, J., & Fernández, M. (2012). Tethys–Atlantic interaction along the Iberia–Africa plate boundary: The Betic–Rif orogenic system. *Tectonophysics*, 579, 144-172.
- Viswanathan, H. S., Ajo-Franklin, J., Birkholzer, J. T., Carey, J. W., Guglielmi, Y., Hyman, J. D., ... & Tartakovsky, D. M. (2022). From fluid flow to coupled processes in fractured rock: Recent advances and new frontiers. *Reviews of Geophysics*, 60(1), e2021RG000744.
- Walsh, J. B. (1965). The effect of cracks in rocks on Poisson's ratio. *Journal of Geophysical Research*, 70(20), 5249-5257.
- Wang, H. (2019). Hydraulic fracture propagation in naturally fractured reservoirs: Complex fracture or fracture networks. *Journal of Natural Gas Science and Engineering*, 68, 102911.
- Wang, H., Li, Y., Song, D., Lin, M., Guo, X., & Shi, X. (2024). Mechanism and model analysis of ultralow-temperature fluid fracturing in low-permeability reservoir: Insights from liquid nitrogen fracturing. *Processes*, 12(6), 1117.
- Wang, Q., Qi, J., Hosseini, S., Rasekh, H., & Huang, J. (2023). ICA-LightGBM algorithm for predicting compressive strength of geo-polymer concrete. *Buildings*, 13(9), 2278.
- Wang, R. (2012). AdaBoost for feature selection, classification and its relation with SVM, a review. *Physics Procedia*, 25, 800-807.
- Wang, X., & Sterling, R. L. (2007). Stability analysis of a borehole wall during horizontal directional drilling. *Tunnelling and Underground Space Technology*, 22(5-6), 620-632.
- Wang, X., Yang, S., Zhao, Y., & Wang, Y. (2018). Improved pore structure prediction based on MICP with a data mining and machine learning system approach in Mesozoic strata of Gaoqing field, Jiyang depression. *Journal of Petroleum Science and Engineering*, 171, 362-393.
- Wang, Z., Luo, X., Lei, Y., Zhang, L., Shi, H., Lu, J., ... & Jiang, T. (2020). Impact of detrital composition and diagenesis on the heterogeneity and quality of low-permeability to tight sandstone reservoirs: An example of the Upper Triassic Yanchang Formation in Southeastern Ordos Basin. *Journal of Petroleum Science and Engineering*, 195, 107596.
- Warpinski, N. R., & Teufel, L. W. (1987). Influence of Geologic Discontinuities on Hydraulic Fracture Propagation: *Journal of Petroleum Technology*.

- Warpinski, N. R., Mayerhofer, M. J., Vincent, M. C., Cipolla, C. L., & Lolon, E. P. (2009). Stimulating unconventional reservoirs: maximizing network growth while optimizing fracture conductivity. *Journal of Canadian Petroleum Technology*, 48(10), 39-51.
- Warren, J. E., & Root, P. J. (1963). The behavior of naturally fractured reservoirs. *Society of Petroleum Engineers Journal*, 3(03), 245-255.
- Wei, S., Jin, Y., Xia, Y., & Lin, B. (2020). The flowback and production analysis in sub-saturated fractured shale reservoirs. *Journal of Petroleum Science and Engineering*, 186, 106694.
- Wei, Z., Yang, W., Chen, Q., Liang, D., & Wu, G. (2024). Uniaxial compressive strength prediction based on measurement while drilling data: A laboratory study. *Journal of Applied Geophysics*, 229, 105499.
- Wennberg, O. P., Casini, G., Jonoud, S., & Peacock, D. C. (2016). The characteristics of open fractures in carbonate reservoirs and their impact on fluid flow: a discussion. *Petroleum Geoscience*, 22(1), 91-104.
- Wenzhi, Z., Suyun, H., Lianhua, H., Tao, Y., Xin, L., Bincheng, G. U. O., & Zhi, Y. (2020). Types and resource potential of continental shale oil in China and its boundary with tight oil. *Petroleum Exploration and Development*, 47(1), 1-11.
- Wibberley, C. A., Petit, J. P., & Rives, T. (2000). Micromechanics of shear rupture and the control of normal stress. *Journal of Structural Geology*, 22(4), 411-427.
- Willmott, C. J., & Matsuura, K. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate research*, 30(1), 79-82.
- Worden, R. H., Armitage, P. J., Butcher, A. R., Churchill, J. M., Csoma, A. E., Hollis, C., ... & Omma, J. E. (2018). Petroleum reservoir quality prediction: overview and contrasting approaches from sandstone and carbonate communities.
- Wright, E., Landry, C. J., & Eichhubl, P. (2023). Occurrence and origin of nanoscale grain boundary channels under diagenetic conditions. *Journal of Geophysical Research: Solid Earth*, 128(7), e2023JB026961.
- Wu, Z., Cui, C., Jia, P., Wang, Z., & Sui, Y. (2022). Advances and challenges in hydraulic fracturing of tight reservoirs: A critical review. *Energy Geoscience*, 3(4), 427-435.
- Xi, Y., Lian, W., Fan, L., Tao, Q., & Guo, X. (2021). Research and engineering application of pre-stressed cementing technology for preventing micro-annulus caused by cyclic loading-unloading in deep shale gas horizontal wells. *Journal of Petroleum Science and Engineering*, 200, 108359.
- Xie, J., Cheng, W., Wang, R., Jiang, G., Sun, D., & Sun, J. (2018). Experiments and analysis on the influence of perforation mode on hydraulic fracture geometry in shale formation. *Journal of Petroleum Science and Engineering*, 168, 133-147.
- Xu, C., Kang, Y., You, Z., & Chen, M. (2016). Review on formation damage mechanisms and processes in shale gas reservoir: Known and to be known. *Journal of Natural Gas Science and Engineering*, 36, 1208-1219.
- Xu, K., Xia, Z., Cheng, M., & Tan, X. (2023). Carbon price prediction based on multiple decomposition and XGBoost algorithm. *Environmental Science and Pollution Research*, 30(38), 89165-89179.
- Xu, Y., Liu, X., Cao, X., Huang, C., Liu, E., Qian, S., ... & Zhang, J. (2021). Artificial intelligence: A powerful paradigm for scientific research. *The Innovation*, 2(4).
- Yahi, N., Schaefer, R. G., & Littke, R. (2001). Petroleum generation and accumulation in the Berkine basin, eastern Algeria. *AAPG bulletin*, 85(8), 1439-1467.
- Yang, J., Sun, J., Bai, Y., Lv, K., Zhang, G., & Li, Y. (2022). Status and prospect of drilling fluid loss and lost circulation control technology in fractured formation. *Gels*, 8(5), 260.

- Yang, R. Z., Zhao, Z. G., Peng, W. J., Gu, Y. B., Wang, Z. G., & Zhuang, X. Q. (2013). Integrated application of 3D seismic and microseismic data in the development of tight gas reservoirs. *Applied Geophysics*, 10, 157-169.
- Yao, X., Song, X., Han, L., Zhang, H., Zhu, Z., Li, B., & Zhang, R. (2023, June). Real-Time Prediction of Formation Pore Pressure Using Hybrid LSTM-BP Model. In *ARMA US Rock Mechanics/Geomechanics Symposium* (pp. ARMA-2023). ARMA.
- Yin, S., Radwan, A. E., Xiao, F., Xie, G., & Lai, P. (2024). Developmental characteristics of vertical natural fracture in low-permeability oil sandstones and its influence on hydraulic fracture propagation. *Geomechanics and Geophysics for Geo-Energy and Geo-Resources*, 10(1), 165.
- Ying, C., Qi-Guang, M., Jia-Chen, L., & Lin, G. (2013). Advance and prospects of AdaBoost algorithm. *Acta Automatica Sinica*, 39(6), 745-758.
- Yu, H., Xu, W., Li, B., Huang, H., Micheal, M., Wang, Q., ... & Wu, H. (2023). Hydraulic fracturing and enhanced recovery in shale reservoirs: theoretical analysis to engineering applications. *Energy & Fuels*, 37(14), 9956-9997.
- Yue, X., Yue, Z., Yan, Y., Gao, D., & Li, Y. (2023). Sensitivity analysis and rock strength prediction study of rotary drilling with drilling parameters. *Geoenergy Science and Engineering*, 230, 212169.
- Zerrouki, A. A., Aïfa, T., & Baddari, K. (2014). Prediction of natural fracture porosity from well log data by means of fuzzy ranking and an artificial neural network in Hassi Messaoud oil field, Algeria. *Journal of Petroleum Science and Engineering*, 115, 78-89.
- Zhang, D., & Tingyun, Y. A. N. G. (2015). Environmental impacts of hydraulic fracturing in shale gas development in the United States. *Petroleum Exploration and Development*, 42(6), 876-883.
- Zhang, H., Zeng, J., Wang, M., Qiao, J., & Cui, H. (2024). Comprehensive Multi-Method Characterization of Pore Structure and Its Impact on Fluid Mobility of Tight Sandstones—A Case Study of the Shaximiao Formation in the Central Sichuan Basin. *Processes*, 12(11), 2470.
- Zhang, J., Yin, S.-X., 2017. Fracture gradient prediction: an overview and an improved method. *Petrol. Sci.* 14, 720–730. <https://doi.org/10.1007/s12182-017-0182-1>.
- Zhang, L., Wang, B., Hu, M., Shi, X., Yang, L., & Zhou, F. (2025). Research Progress on Optimization Methods of Platform Well Fracturing in Unconventional Reservoirs. *Processes*, 13(6), 1887.
- Zhang, P. B., & Yang, Z. X. (2016). A novel AdaBoost framework with robust threshold and structural optimization. *IEEE transactions on cybernetics*, 48(1), 64-76.
- Zhang, T., Ellis, G. S., Ruppel, S. C., Milliken, K., & Yang, R. (2012). Effect of organic-matter type and thermal maturity on methane adsorption in shale-gas systems. *Organic geochemistry*, 47, 120-131.
- Zhang, W., He, Y., Wang, L., Liu, S., & Meng, X. (2023). Landslide Susceptibility mapping using random forest and extreme gradient boosting: A case study of Fengjie, Chongqing. *Geological Journal*, 58(6), 2372-2387.
- Zhang, W., Wu, C., Zhong, H., Li, Y., & Wang, L. (2021). Prediction of undrained shear strength using extreme gradient boosting and random forest based on Bayesian optimization. *Geoscience Frontiers*, 12(1), 469-477.
- Zhang, X., Jeffrey, R. G., & Thiercelin, M. (2009). Mechanics of fluid-driven fracture growth in naturally fractured reservoirs with simple network geometries. *Journal of Geophysical Research: Solid Earth*, 114(B12).
- Zhang, Y., Yu, W., Li, Z., Raza, S., & Cao, H. (2021). Detecting ethereum ponzi schemes based on improved lightgbm algorithm. *IEEE Transactions on Computational Social Systems*, 9(2), 624-637.
- Zhang, Z., Li, X., He, J., Wu, Y., & Li, G. (2017). Numerical study on the propagation of tensile and shear fracture network in naturally fractured shale reservoirs. *Journal of Natural Gas Science and Engineering*, 37, 1-14.

- Zhao, A., Yu, Y., Wang, B., Liu, Y., Liu, J., Fu, X., ... & Tian, F. (2025). Real-Time Intelligent Recognition and Precise Drilling in Strongly Heterogeneous Formations Based on Multi-Parameter Logging While Drilling and Drilling Engineering. *Applied Sciences*, 15(10), 5536.
- Zhao, B., Ratnakar, R., Dindoruk, B., & Mohanty, K. (2020). A hybrid approach for the prediction of relative permeability using machine learning of experimental and numerical proxy SCAL data. *SPE Journal*, 25(05), 2749-2764.
- Zhao, C., Jiang, Y., & Wang, L. (2022). Data-driven diagenetic facies classification and well-logging identification based on machine learning methods: a case study on Xujiache tight sandstone in Sichuan Basin. *Journal of Petroleum Science and Engineering*, 217, 110798.
- Zheng, H., Gao, X., Yang, M., Yang, X., Li, Y., & Ding, Y. (2025). Complex product quality prediction method based on an improved light gradient boosting machine. *Applied Intelligence*, 55(4), 248.
- Zhong, C., Geng, F., Zhang, X., Zhang, Z., Wu, Z., & Jiang, Y. (2021, May). Shear wave velocity prediction of carbonate reservoirs based on CatBoost. In 2021 4th International Conference on Artificial Intelligence and Big Data (ICAIBD) (pp. 622-626). IEEE.
- Zoback, M. D. (2010). *Reservoir geomechanics*. Cambridge university press.
- Zoback, M. D. 2007. *Reservoir Geomechanics*. 1st ed. Cambridge, United Kingdom: Cambridge University Press.
- Zoback, M. D., & Kohli, A. H. (2019). *Unconventional reservoir geomechanics*. Cambridge University Press.
- Zou, C. (2017). *Unconventional petroleum geology*. Elsevier.
- Zou, C., Zhao, Q., Dong, D., Yang, Z., Qiu, Z., Liang, F., ... & Hu, Z. (2017). Geological characteristics, main challenges and future prospect of shale gas. *Journal of Natural Gas Geoscience*, 2(5-6), 273-288.
- Zou, C., Zhu, R., Liu, K., Su, L., Bai, B., Zhang, X., ... & Wang, J. (2012). Tight gas sandstone reservoirs in China: characteristics and recognition criteria. *Journal of Petroleum Science and Engineering*, 88, 82-91.

Appendix

Table 13. Hyperparameter setting applied to the three configured ML models

Algorithm	Hyperparameter	Optimum value selected
GB	Learning rate	0.01
	Max depth	3
	Min samples leaf	4
	Min samples split	2
	n estimators	50
	subsample	1.0
XGB	Colsample_bytree	0.3
	n_estimators	100
	max_depth	5
	Learning rate	0.1
RF	n_estimators	100
	max_depth	10

Table 14. Overview of Model Optimization Methods and Hyperparameter Tuning

Algorithms	Hyperparameter	Optimum value selected
RF	n_estimators	100
	max_features	1.0
	min_samples_leaf	1
	min_samples_split	2
	random_state	42
	max_depth	10
XGB	n_estimators	100
	max_depth	5
	learning_rate	0.1
	random_state	42
	verbosity	0
	Subsample	1.0
GB	learning_rate	0.1
	max_depth	3
	min_samples_leaf	1
	min_samples_split	2
	n_estimators	100
	random_state	42
	subsample	1.0
AB	learning_rate	1.0
	n_estimators	100
	random_state	42
CB	n_estimators	100

LGBM	random_state	42
	colsample_bytree	1.0
	learning_rate	0.1
	max_depth	-1
	min_child_samples	20
	min_child_weight	0.001
	n_estimators	100
	num_leaves	31
	random_state	42
	subsample	1.0

Webography

Figure 5 <https://water.fanack.com/publications/algerias-groundwater-resources/>

Figure 12 <https://www.red-gate.com/simple-talk/business-intelligence/data-science/introduction-to-artificial-intelligence/>

Figure 13 <https://medium.com/data-science/machine-learning-types-2-c1291d4f04b1>

Figure 14 <https://medium.com/@vipinjanghu/supervised-machine-learning-a-beginners-guide-f54057572fba>

Figure 15 <https://aditi22prerna.medium.com/unsupervised-learning-a24caf362e79>

Figure 16 <https://www.geeksforgeeks.org/machine-learning/what-is-reinforcement-learning/>

Figure 17 <https://lamarr-institute.org/blog/deep-neural-networks/>

Figure 18 <https://williamkoehrsen.medium.com/random-forest-simple-explanation-377895a60d2d>