23rd International Conference on Modeling & Applied Simulation, 019
21st International Multidisciplinary Modeling & Simulation Multiconference

2724-0037 © 2024 The Authors.
doi: 10.46354/i3m.2024.mas.019

Leveraging ML for On-Site Bubble Point Pressure Prediction in Sudanese Crude Oil: A Resilient Approach During Conflict

Mohaiad Elbasheer[1], Hassan Suliman[2], Caterina Fusto[1], Giovanni Mirabelli[1], Antonio Padovano[1] and Vittorio Solina[1,*]

[1]Department of Mechanical, Energy and Management Engineering (DIMEG), University of Calabria, Italy
[2]Petroleum Labs Researches & Studies Ministry of Oil and Gas, Khartoum, Sudan

[*]Corresponding author. Email address: vittorio.solina@unical.it

Abstract

This study proposes a novel machine learning (ML)-driven framework for on-site Bubble Point Pressure (BBP) prediction tailored to Sudanese crude oil. Given the current disruption of traditional laboratory-based Pressure-Volume-Temperature (PVT) analysis due to the conflict in Sudan, there is an urgent need for reliable, decentralized solutions within the energy sector. Empirical correlations, though widely used in this context, often provide inaccurate predictions for BBP due to the unique properties of Sudanese crude.

In this work, we introduce and evaluate multiple ML models, with a focus on optimizing real-time BBP predictions. Among the evaluated models, the Multi-Layer Perceptron (MLP) and XGBoost demonstrated superior predictive accuracy, with MLP achieving the most significant improvements over traditional empirical methods. This advancement directly addresses the limitations of existing models by offering enhanced precision and adaptability. The contribution of this study lies in the development of a field-deployable, ML-powered solution that enables real-time BBP predictions without the need for centralized laboratory infrastructure. This approach not only ensures continuity of oil analysis & production during the conflict but also provides a robust foundation for post-conflict recovery, supporting long-term operational and economic resilience in the Sudanese oil sector.

Keywords: Bubble Point Pressure; Conflict; War Zone; Machine Learning; AI; Resilience

1. Introduction

1.1. General Context & Motivation

The oil sector has long been a cornerstone of Sudan's economy, contributing significantly to national revenues. According to recent reports, the oil sector accounted for approximately 50% of the government's income and 95% of export revenues prior to the conflict [3, 24]. Sudan's annual oil production was approximately 60,000 barrels per day (bpd) in recent years, down from a peak of 500,000 bpd in 2011 due to the secession of South Sudan and subsequent loss of major oil fields [23]. Despite this reduction, the oil industry remains vital to the country's economic health and future prospects. However, the ongoing conflict in Sudan has severely disrupted the industry. The war has impacted critical infrastructure, including processing plants, refineries, and key laboratory infrastructure, making it nearly impossible to conduct traditional laboratory analyses necessary for efficient oil production. As a result, production activities have slowed down, and the sector's overall contribution to the economy has been significantly reduced [11]. In addition to the immediate impacts, the long-term aftermath of the conflict presents further challenges. Reconstructing damaged infrastructure, particularly specialized laboratories, will take considerable time and resources. In this context, there is a growing demand for decentralized, field-deployable technologies that can bypass the need for central laboratories. Developing such solutions is crucial to sustaining oil production, minimizing the economic impact of the conflict, and supporting the eventual recovery of Sudan's oil sector. This work contributes by addressing these urgent needs through the development of a practical solution that can be deployed in the field, ensuring continuity in crude oil analysis during and after the conflict.

1.2. Significance of Bubble Point Pressure Estimation

The Bubble Point Pressure (BBP) is a critical parameter in crude oil analysis, representing the pressure at which the first bubble of gas forms in a liquid mixture during pressure reduction [10]. This pressure signifies the onset of gas liberation from the liquid phase, marking the boundary between single-phase liquid and two-phase (liquid-gas) conditions. Theoretical understanding of the bubble point is essential in reservoir engineering and production operations because it influences key decisions regarding the development, management, and operations of oil reservoirs [2]. For instance, maintaining reservoir pressure above the bubble point can prevent gas breakout, thereby optimizing oil recovery and maintaining reservoir drive. Accurate knowledge of the BBP also assists in designing surface separation facilities and managing phase behavior throughout the production process, ensuring efficient and safe handling of the extracted hydrocarbons [14, 22].

Traditionally, the BBP can be determined in two ways: through laboratory experiments using Pressure-Volume-Temperature (PVT) tests, or by employing empirical correlations that leverage field data [12]. PVT tests are considered the most reliable method but are often costly, time-consuming, and require specific conditions for accurate sampling and analysis [12]. Due to these challenges, empirical correlations, like the ones developed by Standing [20], Glaso [9], Al-Marhoun [1], Vasquez-Beggs [25], and Petrosky-Farshad [16], have long been used to estimate BBP, but these methods often rely on generalized datasets that fail to account for the unique properties of local crude oils. This can lead to inaccurate predictions, which in turn may result in inefficient production practices and costly operational errors. Therefore, In war-affected regions like Sudan, where access to traditional laboratory-based PVT analysis is limited, accurate field-based BBP prediction is particularly valuable.

2. Literature Review

2.1. Empirical Correlations for BBP Estimation

Over the decades, various empirical correlations have been developed to simplify BBP estimation without the need for extensive laboratory testing. These correlations, built upon large datasets from specific regions and crude oil types, offer practical solutions but also come with limitations, particularly when applied to diverse crude oil compositions.

One of the earliest and most influential correlations was introduced by Standing in 1947 [20]. Standing's model utilized key variables such as the solution Gas-Oil Ratio (GOR), gas-specific gravity (**SG gas**), oil-specific gravity (**SG oil**), and reservoir temperature to estimate BBP. His work, based on data from Californian crude oils, quickly gained widespread adoption due to its straightforward approach and its utility for light and medium crude oils. Building upon this foundation, Glaso extended the applicability of Standing's method by incorporating additional complexities to better account for higher API gravities and pressures in the North Sea region [9]. This extension helped to capture more nuanced behavior in Pressure-Volume-Temperature (PVT) data, particularly for crude oils with a broader range of physical characteristics. Recognizing the need for more region-specific correlations, Al-Marhoun developed a method in 1988 that focused on Middle Eastern crude oils [1]. His correlation employed similar variables as Standing and Glaso but tailored the parameters to accommodate the higher viscosities and distinct properties of regional crude oils. This allowed operators in the Middle East to make more accurate BBP predictions without relying on empirical data from other regions. In contrast to these region-specific models, the Vasquez-Beggs correlation aimed for a more universally applicable method by segregating crude oils based on their API gravity. This approach provided greater accuracy for low-gravity oils, which are often underrepresented in datasets used for previous correlations [25]. Petrosky and Farshad later improved upon this idea by focusing specifically on heavy oils, incorporating variables such as reservoir temperature, solution GOR, and oil gravity to optimize predictions for heavier crude oil types [16].

2.2. ML Models for BBP Prediction

With the limitations of empirical correlations, machine learning (ML) techniques began gaining traction for their ability to capture non-linear relationships in PVT data. Early models such as Support Vector Regression (SVR) and Artificial Neural Networks (ANNs) showed potential but struggled with sparse data and limited extrapolation capabilities [27]. Building on these initial efforts, hybrid models that integrated multiple algorithms, like Particle Swarm Optimization (PSO) and Multiple Extreme Learning Machines (MELM), demonstrated superior prediction accuracy [17, 13], particularly in handling field data. More advanced neurocomputing models, including Multilayer Perceptron (MLP) and Radial Basis Function (RBF), further improved the handling of high-dimensional, complex datasets [8]. Recent advancements in ensemble learning and deep learning models, such as XGBoost, LightGBM, and random forests, have further pushed the boundaries of BBP prediction. These models, leveraging compositional data and robust cross-validation, have achieved superior predictive accuracy [7, 21].

While significant progress has been made, there remains a gap in applying these advanced techniques to specific crude oils, such as Sudanese crude, which possess unique properties not fully captured by global datasets. This study aims to bridge that gap by applying and optimizing these advanced machine learning techniques for Sudanese crude oil, contributing to both the local industry and the broader field of BBP prediction.

3. Methodology

3.1. BBP Estimation in the Lab setting

To accurately determine the BBP of crude oil in the lab setting, a series of carefully controlled experimental procedures are conducted, typically involving the use of a PVT cell as demonstrated in Figure 1.

As demonstrated in Figure 2, the experimental setup begins with the collection of a representative sample of reservoir fluid, which is introduced into the PVT cell. This cell is capable of withstanding the high pressures and temperatures required to simulate reservoir conditions. The first step in the experimentation process is to heat the sample to the reservoir temperature, ensuring that the fluid mimics the in-situ state as closely as possible. Following this, the pressure in the PVT cell is gradually reduced in a controlled manner. As the pressure decreases, the volume of the fluid is carefully monitored. This phase of the experiment involves observing the point at which the first gas bubble forms, which is the BBP. The formation of the gas bubble indicates the pressure at which the liquid phase begins to separate into gas and liquid phases. To ensure accuracy, the experiment is repeated multiple times, and additional parameters such as temperature, oil and gas volumes, and pressure are recorded throughout the process. The collected data, including pressure-volume relationships and density values, is then analyzed to determine the BBP accurately. The process involves plotting the pressure-volume data to identify the characteristic point where the slope changes, marking the onset of gas bubble formation. This rigorous experimental approach ensures that the determined BBP is reliable and representative of the actual reservoir conditions, thereby aiding in accurate reservoir management and fluid characterization.

3.2. Data Modeling

The dataset used for this experimentation represents an aggregation from various Sudanese oil wells. For each well, the physical characteristics of the crude oil are estimated in the field, providing crucial insights into the thermodynamic behavior of the oil. To estimate the BBP accurately, the dataset includes measurements from 217 different wells. These measurements capture a range of conditions and compositions, forming a robust basis for ML modeling and prediction. The following input parame-

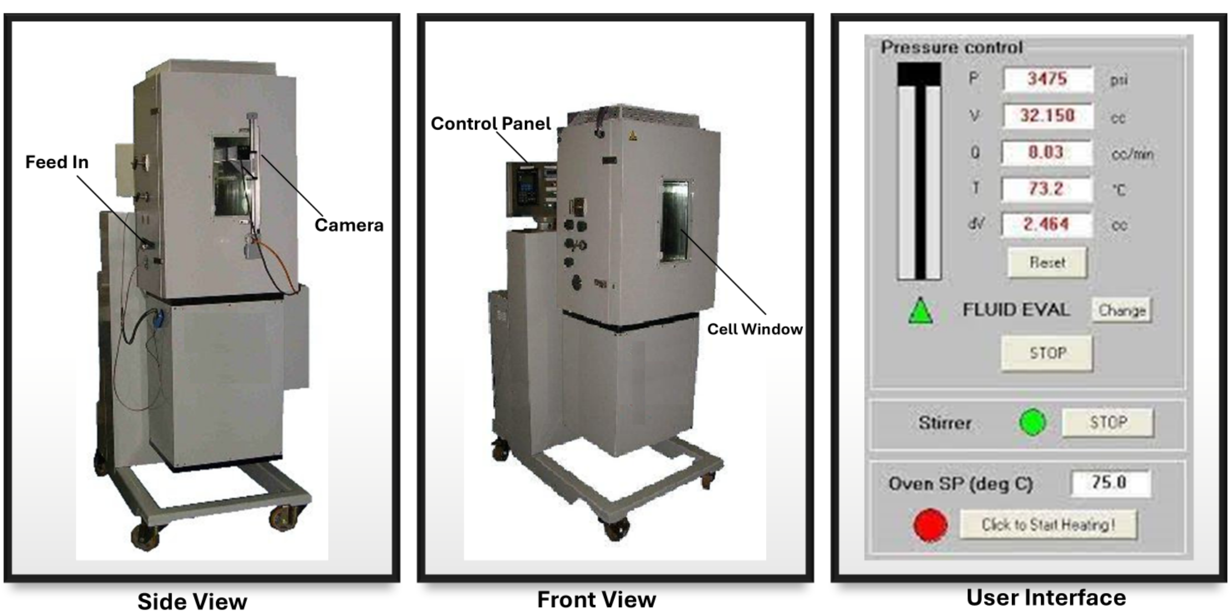

Figure 1. Pressure Volum Temprature (PVT) Cell

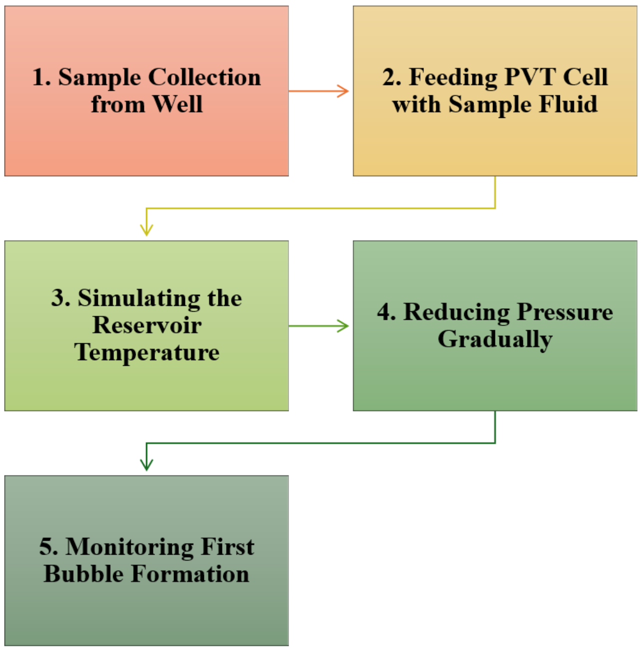

Figure 2. Experimental Steps for Determining the BBP in Lab Settings

ters are crucial for understanding the phase behavior and BBP of crude oil:

- **Temperature (T°R)**: Measured in Rankine, this parameter is vital as it directly influences the fluid's thermodynamic state. The temperature of the crude oil at reservoir conditions is a key determinant of phase changes.
- **Gas-Oil Ratio (GOR)**: This is a measure of the volume of gas that can be dissolved in a unit volume of oil at specific temperature and pressure conditions. GOR provides insight into the fluid's capacity to release gas as the pressure decreases, which is critical for BBP determination.
- **Specific Gravity of Gas (SG gas)**: This is the ratio of the density of the gas phase to the density of air at standard conditions. It reflects the molecular composition of the gas, influencing how the gas interacts with the oil phase.
- **Specific Gravity of Oil (SG oil)**: The density of the oil phase compared to water density, which helps characterize the heaviness or lightness of the crude oil. Variations in SG oil can affect the BBP as different compositions might react differently under pressure changes.

The dataset captures a wide array of samples, each identified by a unique well number, which ensures that the modeling accounts for variability across different geographical and geological settings. This diversity is essential for developing models that can generalize well to new data. To gain deeper insights into the dataset, the Box plots in Figure 3 are used to visually represent the data, highlighting the central tendencies, spread, and presence of any outliers that might influence the modeling process.

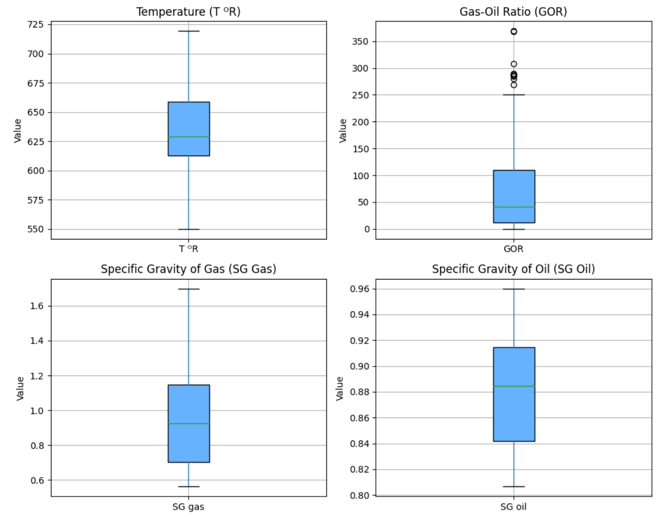

Figure 3. Input Variables for BBP Prediction

3.3. ML Experimentation

In this study, a ML framework was employed to predict the BBP of Sudanese crude oil, incorporating advanced regression techniques and model optimization strategies. The workflow guiding this experimentation, as depicted in Figure 4, outlines the major stages from data collection and preprocessing through to model evaluation and comparison with traditional empirical correlations.

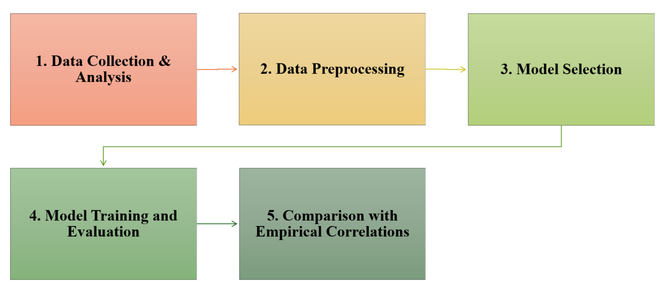

Figure 4. ML Workflow for Predicting Sudanese Oil BBP

3.3.1. Data Collection and Preprocessing

The raw dataset used in this study consisted of 217 wells, with features extracted to characterize the physical and thermodynamic properties of the crude oil. The target variable for prediction was the Measured BBP, representing laboratory-determined values for each reservoir sample.

Data preprocessing involved two key steps:

1. **Outlier Removal**: The Interquartile Range (IQR) method was employed to detect and eliminate extreme outliers from the dataset [26], ensuring that anomalous data points would not skew the ML models.
2. **Standardization**: Features were standardized using the StandardScaler to ensure that all input variables had a

mean of 0 and a standard deviation of 1 [15].

3.3.2. *Model Selection and Hyperparameter Optimization*

To ensure the most robust and accurate predictions, four different, widely used, regression models were selected based on their ability to handle nonlinear relationships and their popularity in tabular datasets. which makes them particularly relevant for the building a predictive model on BBP of Sudanese crude oil:

1. **Linear Regression (Baseline)**: A basic linear model was employed as a baseline for comparison [19].
2. **Random Forest Regressor**: An ensemble method capable of capturing complex nonlinear relationships by constructing multiple decision trees [4].
3. **Support Vector Regressor (SVR)**: A model capable of handling high-dimensional feature spaces using kernel tricks [5].
4. **XGBoost Regressor**: An advanced boosting algorithm known for its high predictive accuracy in structured datasets [6].
5. **Multi-Layer Perceptron (MLP)**: A neural network model, optimized to capture complex patterns in the data [18].

Each model was trained and optimized using GridSearchCV to identify the best hyperparameters for each algorithm. For Random Forest, SVR, XGBoost, and MLP, a 5-fold cross-validation was used to avoid overfitting and improve the generalization of the models. Table 1 depicts the hyperparameters tuned for each model.

Table 1. Hyperparameters for each machine learning model.

Model	Hyperparameter	Values Considered
Random Forest Regressor	n_estimators	[100, 200, 300]
	max_depth	[10, 20, 30]
SVR	C	[0.1, 1, 10]
	kernel	['linear', 'rbf']
XGBoost Regressor	n_estimators	[100, 200, 300]
	learning_rate	[0.01, 0.1, 0.2]
	max_depth	[3, 5, 7]
MLP	hidden_layer_sizes	[(128, 64, 32), (100, 50), (256, 128, 64)]
	activation	['relu']
	alpha	[0.0001, 0.001, 0.01]
	learning_rate	['constant', 'adaptive']
	learning_rate_init	[0.001, 0.01]

Hyperparameters were optimized using 5-fold cross-validation via GridSearchCV.

3.3.3. *Model Training & Evaluation*

The models were trained on 80% of the data, with the remaining 20% reserved for testing. Training was conducted using the optimal hyperparameters determined through the cross-validation process. The models were then evaluated on the unseen test set to gauge their generalization capability. The evaluation was carried out using several regression performance metrics. Root Mean Squared Error (RMSE) was used to assess the average magnitude of prediction errors, where larger errors are penalized more heavily. Mean Absolute Error (MAE) was calculated to provide an overall measure of error magnitude without exaggerating large errors. The Mean Absolute Percentage Error (MAPE) was computed to show the percentage-based prediction error, useful for comparing model performance across different scales. Finally, the R^2 (coefficient of determination) was used to measure the proportion of variance explained by the model, indicating how well the predictions fit the actual data.

4. Results

Figure 5 shows the distributions of the input features after preprocessing, including Temperature, Gas-Oil Ratio (GOR), Specific Gravity of Gas (SG gas), and Specific Gravity of Oil (SG oil). These variables are key to understanding the physical and thermodynamic properties of the crude oil, and their distribution characteristics provide insight into the dataset's complexity. The Temperature distribution is approximately normal, with most values clustered around the mean, suggesting that the majority of wells have similar temperature conditions. In contrast, the GOR exhibits a right-skewed distribution, indicating that most wells have low gas content, with fewer wells showing higher GOR values. The SG gas presents a wider spread and multimodal tendencies, reflecting variability in gas compositions across the sampled wells. Finally, SG oil shows a bimodal distribution, pointing to the presence of both lighter and heavier crude oil samples.

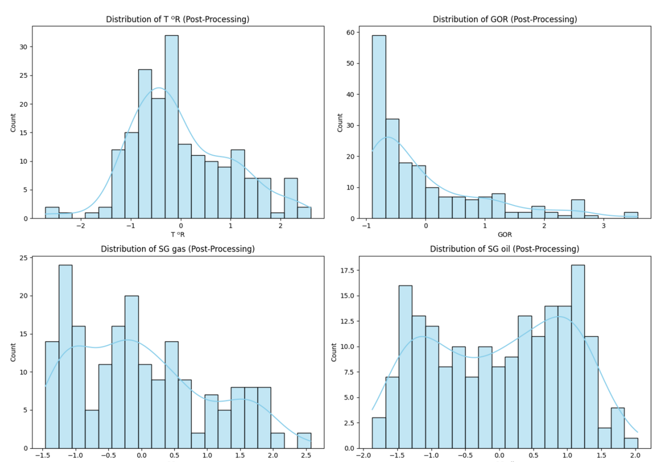

Figure 5. Post-processing Input Variable's Distribution

4.1. Performance of ML Models

The performance of each model was assessed using key regression metrics, including RMSE, MAE, MAPE, and the R^2. The results for each model are shown in Table 2, highlighting the significant differences in performance across the various techniques employed. Figure 6 demonstrates that all the different ML models has a reasonable predictability of the BBP.

The Linear Regression model, serving as a baseline, produced an R^2 of 0.8610 and an RMSE of 170.92. While it captured some of the underlying relationships between the input variables and BBP, its relatively high RMSE and moderate R^2 indicate that the linear model struggled with the nonlinearity in the data.

In contrast, Random Forest showed marked improvement, with an R^2 of 0.8937 and a reduced RMSE of 149.49. This model, by leveraging the power of ensemble learning, was able to better capture complex patterns within the dataset. The optimized hyperparameters for the Random Forest model, including a depth of 20 and 200 estimators, allowed for greater model flexibility without overfitting. The SVR, while effective in some cases, did not perform as well on this dataset. With an R^2 of 0.8146 and the highest RMSE among the machine learning models (197.39), SVR's linear kernel struggled to adequately model the nonlinearity present. Despite optimization, the SVR results suggest that a nonlinear kernel or alternative regression technique might have been better suited for this task. XGBoost, on the other hand, was one of the top-performing models, achieving an R^2 of 0.9181 and an RMSE of 131.15. XGBoost's ability to iteratively refine its predictions through boosting allowed it to capture more intricate relationships between the input features and BBP. The tuned hyperparameters, including a learning rate of 0.2 and a maximum depth of 5, provided an optimal balance between learning efficiency and model complexity. The MLP achieved the best overall performance, with an R^2 of 0.9229 and the lowest RMSE at 127.26. The deep learning architecture of the MLP enabled it to model highly nonlinear relationships between the features and the target variable, particularly benefiting from the ReLU activation function and its multilayer structure. However, its slightly higher MAE (95.57) compared to XGBoost suggests that while MLP was very accurate on average, it may have produced larger errors in certain instances.

Overall, the results highlight the superior performance of advanced models such as XGBoost and MLP, which were able to effectively capture the nonlinear dynamics of the dataset. These models outperformed the more traditional methods like Linear Regression and SVR, particularly in terms of RMSE and R^2, making them highly suitable for predicting BBP in this context.

4.2. Residual Analysis for ML & Empirical Correlation

Figure 7 shows the residuals for the two best-performing machine learning models (XGBoost and MLP) compared with the empirical correlations (Standing, Glasos, Marhouns, and Petrosky-Farshads). Residuals represent the difference between actual and predicted BBP, with smaller residuals indicating better model performance.

XGBoost exhibits the tightest residual distribution around zero, demonstrating the highest predictive accuracy, followed by MLP. Both models show a concentration of errors near zero, suggesting fewer large deviations and more reliable predictions. In contrast, the empirical correlations show much larger and more dispersed residu-

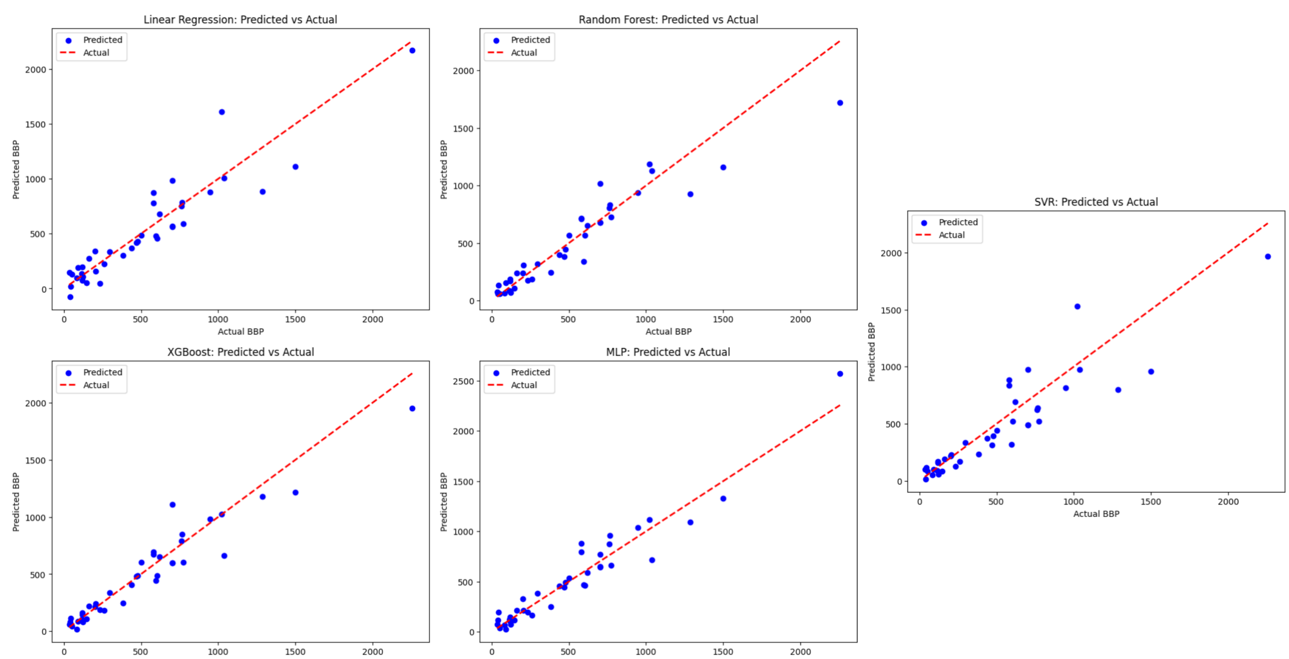

Figure 6. Scattered Plot (Predicted BBP VS Actual BBP)

Table 2. Performance comparison between ML models and Empirical Correlations for BBP Prediction.

Model	RMSE	MAE	MAPE	R^2	Best Parameters / Empirical Model
Linear Regression	170.92	119.68	0.4703	0.8610	N/A
Random Forest	149.49	98.41	0.3112	0.8937	max_depth: 20, n_estimators: 200
SVR	197.39	142.08	0.3716	0.8146	C: 10, kernel: linear
XGBoost	**131.15**	**88.16**	**0.2674**	**0.9181**	learning_rate: 0.2, max_depth: 5, n_estimators: 100
MLP	**127.26**	**95.57**	**0.3629**	**0.9229**	activation: relu, alpha: 0.0001, hidden_layer_sizes: (128, 64, 32), learning_rate: constant, learning_rate_init: 0.001
Standing	199.26	119.59	0.5521	0.8432	*Empirical Model*
Glasos	183.44	126.65	0.4944	0.8671	*Empirical Model*
Marhouns	232.98	149.32	0.8558	0.7857	*Empirical Model*
Petrosky-Farshads	913.07	845.45	11.7003	−2.2918	*Empirical Model*

Machine learning models were trained and evaluated using GridSearchCV for hyperparameter tuning. Empirical models were evaluated directly based on their performance on the dataset.

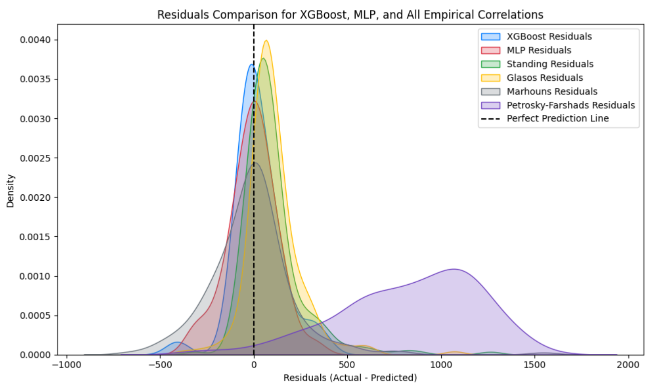

Figure 7. Residuals Comparison for XGBoost, MLP, and Empirical Correlations

als, particularly for Petrosky-Farshads, which has a notable bias and high variance, indicating poor generalization to the data. Marhouns also exhibits wide residuals, while Standing and Glasos correlations, although better, still show greater error spread compared to XGBoost and MLP. This analysis highlights the superior performance of machine learning models, particularly XGBoost, over traditional empirical correlations, which tend to have less reliable predictions for BBP in Sudanese crude oil.

5. Discussion

The results of this study demonstrate the clear superiority of machine learning models, particularly XGBoost and MLP, over traditional empirical correlations in predicting BBP for Sudanese crude oil. XGBoost achieved an R^2 of 0.9181 with an RMSE of 131.15, while MLP performed even better, with an R^2 of 0.9229 and an RMSE of 127.26, highlighting the ability of these models to capture the complex, nonlinear relationships in the data. In contrast, empirical models such as Standing, Glasos, and Marhouns exhibited significantly larger residuals, with Petrosky-Farshads performing the worst due to its limited applicability to the unique characteristics of Sudanese crude oil. These findings underscore the potential of advanced ML models in providing more accurate, real-time BBP predictions in the field, making them well-suited for deployment in resource-constrained environments.

In this study, the MLP model was selected to power the developed on-site BBP prediction application due to its superior performance in capturing complex, nonlinear relationships in the Sudanese crude oil data. As shown in Figure 8, the app interface enables field engineers to input key reservoir parameters—Temperature, GOR, SG gas, and SG oil—with the predicted BBP displayed instantly. This real-time capability facilitates critical decisions without requiring laboratory facilities.

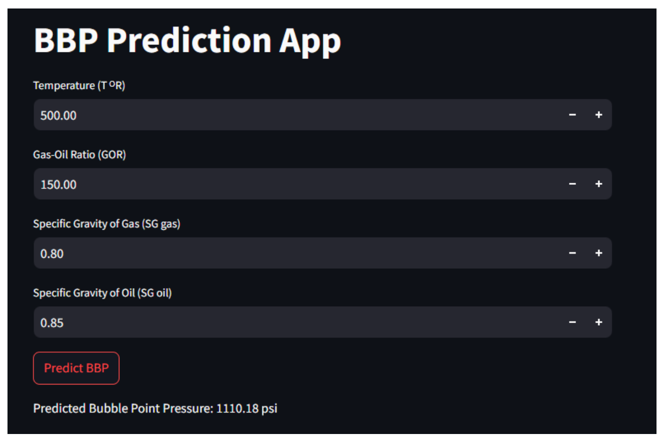

Figure 8. Interface for On-Site BBP Prediction APP

The ongoing conflict in Sudan has made conventional laboratory testing difficult, with key infrastructures often being inaccessible or destroyed. In this context, the on-site application addresses an urgent need for decentralized tools that allow oil producers to continue operations. The app enables users to gather essential data in the field and generate accurate BBP predictions, bypassing the logistical challenges of transporting samples to central labs, which are either damaged or too far away. Looking ahead, this tool will also play a crucial role in post-war recovery. Rebuilding laboratory facilities can take years, but with this on-site solution, the industry can maintain momentum and avoid unnecessary downtime. The application's portability and accessibility ensure that it can be deployed in remote areas or under challenging conditions, providing vital support for the continuity of oil production and reservoir management. The on-site BBP prediction application contributes to resilience in the face of adversity. It provides an adaptive solution that not only addresses current limitations caused by the war but also positions the industry for a smoother post-conflict recovery.

6. Conclusion

This study demonstrates how ML models, particularly the MLP model, significantly improve the accuracy of BBP predictions of Sudanese crude compared to traditional empirical correlations. The superior performance of ML models, validated through various metrics, highlights their potential in addressing the challenges of BBP estimation, especially in regions affected by war, where conventional laboratory facilities are either unavailable or inaccessible. By providing a resilient on-site solution, this research underscores the importance of innovation in sustaining oil production and reservoir management during and after conflict.

Looking to the future, there are considerable opportunities for advancing this model. Incorporating deeper neural networks, such as Long Short-Term Memory (LSTM) networks or more recent architectures like TabNet, could further enhance prediction accuracy by capturing more intricate patterns within the data. These advancements may provide even more robust solutions, particularly when dealing with large-scale, real-time datasets. Moreover, integrating the on-site BBP application with sensor technology opens new possibilities for real-time, automatic BBP estimation within an Internet of Things (IoT) framework. This integration would enable continuous monitoring of reservoir conditions, facilitating more responsive and data-driven decision-making in the field, thereby optimizing production and reducing downtime in both conflict and post-war recovery phases. This approach paves the way for a more resilient, adaptive oil sector that can thrive even in the most challenging environments

References

1. Al-Marhoun, M. A. (1988). Pvt correlations for middle east crude oils. *Journal of Petroleum Technology*, 40(05):650–666.
2. Al-Shammasi, A. (1999). Bubble point pressure and oil formation volume factor correlations. In *SPE Middle East Oil and Gas Show and Conference*, pages SPE–53185. SPE.
3. Ali, S., Murshed, S. M., and Papyrakis, E. (2023). Oil, export diversification and economic growth in sudan: evidence from a var model. *Mineral Economics*, 36(1):77–96.
4. Breiman, L. (2001). Random forests. *Machine learning*, 45:5–32.
5. Burges, C. J. and Schölkopf, B. (1996). Improving the accuracy and speed of support vector machines. *Advances in neural information processing systems*, 9.
6. Chen, T. and Guestrin, C. (2016). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, pages 785–794.
7. Gholami, A., Asoodeh, M., and Bagheripour, P. (2014). How committee machine with svr and ace estimates bubble point pressure of crudes. *Fluid Phase Equilibria*, 382:139–149.
8. Ghorbani, H., Wood, D., Choubineh, A., Mohamadian, N., Tatar, A., Farhangian, H., and Nikooey, A. (2020). Performance comparison of bubble point pressure from oil pvt data: several neurocomputing techniques compared. exp comput multiph flow 2: 225–246.
9. Glaso, O. (1980). Generalized pressure-volume-temperature correlations. *Journal of Petroleum Technology*, 32(05):785–795.
10. Lasater, J. (1958). Bubble point pressure correlation. *Journal of Petroleum Technology*, 10(05):65–67.
11. Liptrot, T. (2024). Keeping the oil flowing: Sudan's oil sector during the 2023–24 war.
12. Moradi, B., Malekzadeh, E., Amani, M., Boukadi, F. H., and Kharrat, R. (2010). Bubble point pressure empirical correlation. In *SPE Trinidad and Tobago Section Energy Resources Conference?*, pages SPE–132756. SPE.
13. Naveshki, M., Naghiei, A., Soltani Tehrani, P., Ahmadi Alvar, M., Ghorbani, H., Mohamadian, N., and Moghadasi, J. (2021). Prediction of bubble point pressure using new hybrid computational intelligence models. *Journal of Chemical and Petroleum Engineering*, 55(2):203–222.
14. Oyeneyin, B. (2015). Introduction to the hydrocarbon composite production system. In *Developments in Petroleum Science*, volume 63, pages 11–128. Elsevier.
15. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., et al. (2011). Scikit-learn: Machine learning in python. *the Journal of machine Learning research*, 12:2825–2830.
16. Petrosky Jr, G. and Farshad, F. (1993). Pressure-volume-temperature correlations for gulf of mexico crude oils. In *SPE Annual Technical Conference and Exhibition?*, pages SPE–26644. SPE.
17. Rashidi, S., Mehrad, M., Ghorbani, H., Wood, D. A., Mohamadian, N., Moghadasi, J., and Davoodi, S. (2021). Determination of bubble point pressure & oil formation volume factor of crude oils applying multiple hidden layers extreme learning machine algorithms. *Journal of Petroleum Science and Engineering*, 202:108425.
18. Rumelhart, D. E., Hinton, G. E., and Williams, R. J. (1986). Learning representations by back-propagating errors. *nature*, 323(6088):533–536.
19. Seber, G. A. and Lee, A. J. (2012). *Linear regression analysis*. John Wiley & Sons.

20. Standing, M. B. et al. (1977). Volumetric and phase behavior of oil field hydrocarbon systems.
21. Tazikeh, S., Davoudi, A., Shafiei, A., Parsaei, H., Atabaev, T. S., and Ivakhnenko, O. P. (2022). A comparison between the perturbed-chain statistical associating fluid theory equation of state and machine learning modeling approaches in asphaltene onset pressure and bubble point pressure prediction during gas injection. *ACS omega*, 7(34):30113–30124.
22. Tewari, R. D., Dandekar, A. Y., and Ortiz, J. M. (2018). *Petroleum fluid phase behavior: characterization, processes, and applications*. CRC Press.
23. The International Trade Administration (2022). Sudan - country commercial guide. Accessed: 2024-09-06.
24. The World Bank (2024). Sudan overview. Accessed: 2024-09-06.
25. Vazquez, M. and Beggs, H. D. (1977). Correlations for fluid physical property prediction. In *SPE Annual Technical Conference and Exhibition?*, pages SPE–6719. SPE.
26. Vinutha, H., Poornima, B., and Sagar, B. (2018). Detection of outliers using interquartile range technique from intrusion dataset. In *Information and decision sciences: Proceedings of the 6th international conference on ficta*, pages 511–518. Springer.
27. Wood, D. A. and Choubineh, A. (2020). Transparent open-box learning network and artificial neural network predictions of bubble-point pressure compared. *Petroleum*, 6(4):375–384.