Improved Model for Predicting Water Production in a Typical Dry Gas Well

Ebifagha Bebenimibo[1]; Victor Joseph Aimikhe[2]

[1]Energy Technology Institute (ETI), University of Port-Harcourt, Port-Harcourt, Nigeria
[2]Department of Petroleum and Gas Engineering, University of Port-Harcourt, Port-Harcourt, Nigeria

Publication Date: 2025/12/04

Abstract: Water production is a major challenge in many gas wells leading to a substantial reduction in the net gas deliverability, high operational costs associated with large-scale water separation facilities, increased energy consumption for lifting fluids, and significant logistical and regulatory burden of produced water disposal. This study focused on the development of a more accurate model for predicting and diagnosing water production in dry gas wells to enhance reservoir management and optimize production strategies. The model was developed by investigating the effectiveness of various machine learning models in predicting water production rates from dry gas wells using data sourced from Niger Delta. Five algorithms were trained and validated on a dataset comprising 249 daily records across eight dry gas wells in three reservoirs. The results showed CatBoost as an effective tool for petroleum reservoir forecasting demonstrating its superiority over widely used algorithms such as Random Forest, XGBoost, KNN and Support Vector Regression in predicting water production. Furthermore, the model predictions revealed that permeability had the strongest positive correlation with water production, whereas reservoir temperature and dew point pressure exhibited significant negative correlations. Overall, the findings underscore the superiority of ensemble-based models in capturing the complex, nonlinear relationships inherent in the Niger Delta field data. The developed predictive model is crucial for engineers to size equipment properly and avoid both under-capacity (leading to facility bottlenecks) and over-design (leading to wasted capital). Further research will focus on expanding the dataset to include a wider range of operational and reservoir parameters to enhance the model robustness and applicability.

Keywords: *Liquid Loading, Water Production, Machine Learning, Dry Gas Wells, Modelling.*

How to Cite: Ebifagha Bebenimibo; Victor Joseph Aimikhe (2025) Improved Model for Predicting Water Production in a Typical Dry Gas Well. *International Journal of Innovative Science and Research Technology*, 10(11), 2446-2460.
https://doi.org/10.38124/ijisrt/25nov1436

I. INTRODUCTION

A prominent part of the world's energy mix, natural gas is becoming increasingly important in the continuous shift to cleaner energy sources. Globally, and especially in areas like Nigeria, gas deposits are a significant resource for industrialization and economic growth. However, the unavoidable problem of undesired water production frequently hinders the effective and sustainable production of natural gas, particularly in established dry gas fields.

Gas wells that have high water-gas ratio have several detrimental effects. In addition to speeding up corrosion and scaling in wellbore and surface facilities, it lowers the volume of saleable gas thereby raising operating expenditure (OPEX) because of the costs of water lifting, separation, treatment, and disposal. This may result in environmental issues with produced water management, safety risks, and equipment damage. Uncontrolled water inflow ultimately leads to premature well abandonment, wasteful reservoir depletion, and the avoidance of valuable gas reserves [1].

In the oil and gas sector, the "water burden" on gas wells is a serious and expanding issue. Water invasion becomes more noticeable when gas fields age and reservoir pressures drop, which raises water-gas ratios (WGRs) [2]. This phenomenon directly results in a significant decrease in a well's net gas deliverability. The presence of water results in disproportionately high operating expenses in addition to the immediate loss of gas. These include the substantial logistical and legal burden of disposing of produced water, the higher energy consumption for lifting fluids, and the initial and ongoing costs of large-scale water separation and treatment facilities [2].

Furthermore, produced water often contains gases and dissolved particles (such as CO_2 and H_2S) that hasten the corrosion of surface flowlines, downhole tubulars, and processing machinery. This results in more safety hazards, expensive maintenance, and equipment replacement. Additionally, water can lead to hydrate development and scaling, which can drastically limit or even entirely block flow channels, causing costly workovers and unanticipated production delays [3]. From the standpoint of the reservoir, high water production frequently indicates poor sweep, in which encroaching water avoids gas and lowers the valuable hydrocarbon's eventual recovery factor ([4]; [5]).

Although there are several water management technologies available, including cement squeezes, polymer gels, and sophisticated completions, their use is frequently ad hoc, reactive, and devoid of thorough, model-driven optimization [6]. Current water breakthrough or WGR progression prediction models sometimes rely on oversimplified hypotheses or empirical correlations that could not fully account for the fluid complexity, geological heterogeneity, or dynamic operating conditions typical of numerous gas fields including those in the Niger Delta. As a result, the economic life of gas wells is not maximized, ineffective treatments are chosen, and intervention timing is suboptimal [7]. An enhanced, integrated modelling method that can offer precise forecasts, trustworthy diagnostics, and optimized intervention recommendations catered to the unique difficulties of gas production is therefore desperately needed.

➢ *To Accomplish this Overall Goal, the Following Goals were Developed:*

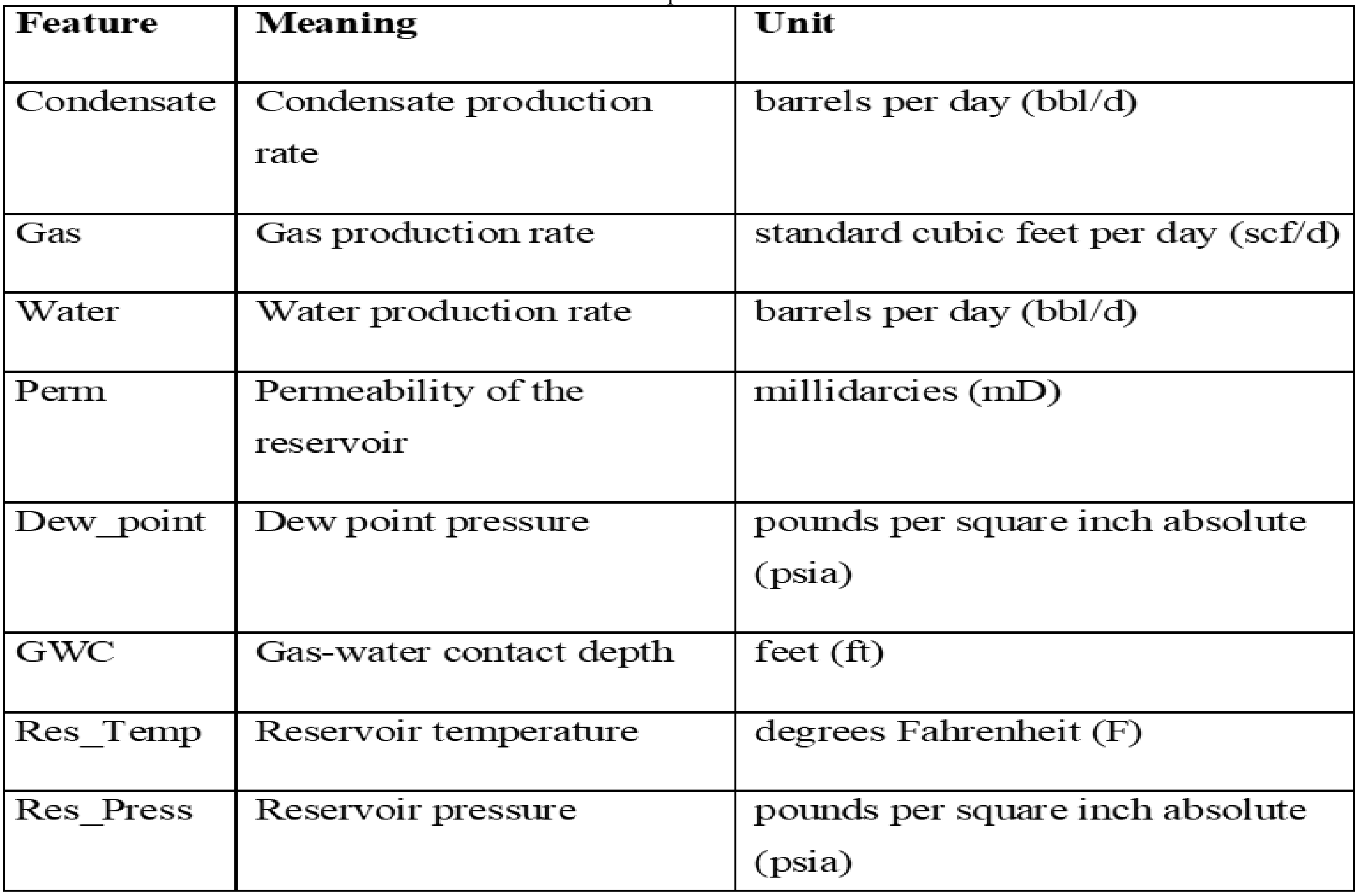
- Obtain primary data from eight dry gas wells and three gas reservoirs in Niger Delta.
- Create a data-driven model to forecast a typical gas well's water production.
- Conduct a comparison with other conventional methods of forecasting water production
- Determine the significant parameters that mostly influence water production.

II. MATERIALS AND METHODS

This study utilized data derived from historical production and reservoir records, comprising 249 daily entries from eight dry gas wells situated in three gas reservoirs inside the Niger Delta. Gas rate, water rate, condensate rate, permeability, dew point pressure, reservoir temperature, reservoir pressure, and gas-water contact depth are some of the parameters in the data (see Table 1). Field engineers kept operations logs and took field measurements to get these data. Standard field sensors and recording tools were used to gather the data.

Data cleaning and preprocessing guaranteed that the dataset was reliable. Box plots were used to find outliers, and Tukey's Fences were used to deal with them based on the interquartile range. Normalization was also used to make feature scales more consistent, which made the model work better. The analysis is reliable because there are no missing values and strong preprocessing methods are used. The dataset had features that were on very different scales, which can have a big effect on how well machine learning models work.

Table 1 Description of Features



Feature	Meaning	Unit
Condensate	Condensate production rate	barrels per day (bbl/d)
Gas	Gas production rate	standard cubic feet per day (scf/d)
Water	Water production rate	barrels per day (bbl/d)
Perm	Permeability of the reservoir	millidarcies (mD)
Dew_point	Dew point pressure	pounds per square inch absolute (psia)
GWC	Gas-water contact depth	feet (ft)
Res_Temp	Reservoir temperature	degrees Fahrenheit (F)
Res_Press	Reservoir pressure	pounds per square inch absolute (psia)

To fix this problem, data normalisation was used to put the characteristics on the same scale so that each one had the same effect on model training. [25]. The Standard Scaler method was used to normalize the dataset. This method changes the data so that the mean of each characteristic is 0 and the standard deviation is 1. This transformation centers the data and makes the distribution normal so that all the characteristics work together in the model.

The formula for Standard Scaler is:

$$z = \frac{x-\mu}{\sigma} \qquad (1)$$

Where:

z is the scaled value,

x is the original value,

μ is the mean of the feature,

σ is the standard deviation of the feature.

It is important to note that the target variable (water production) was not normalized because it represents the actual values to be predicted and scaling it can make the interpretation of the model outputs less intuitive. The data was partitioned into two datasets for training and testing using 85% by 15% random split. The ML models were trained using the training dataset and then evaluated using the testing dataset. Figure 1 shows the research methodology.

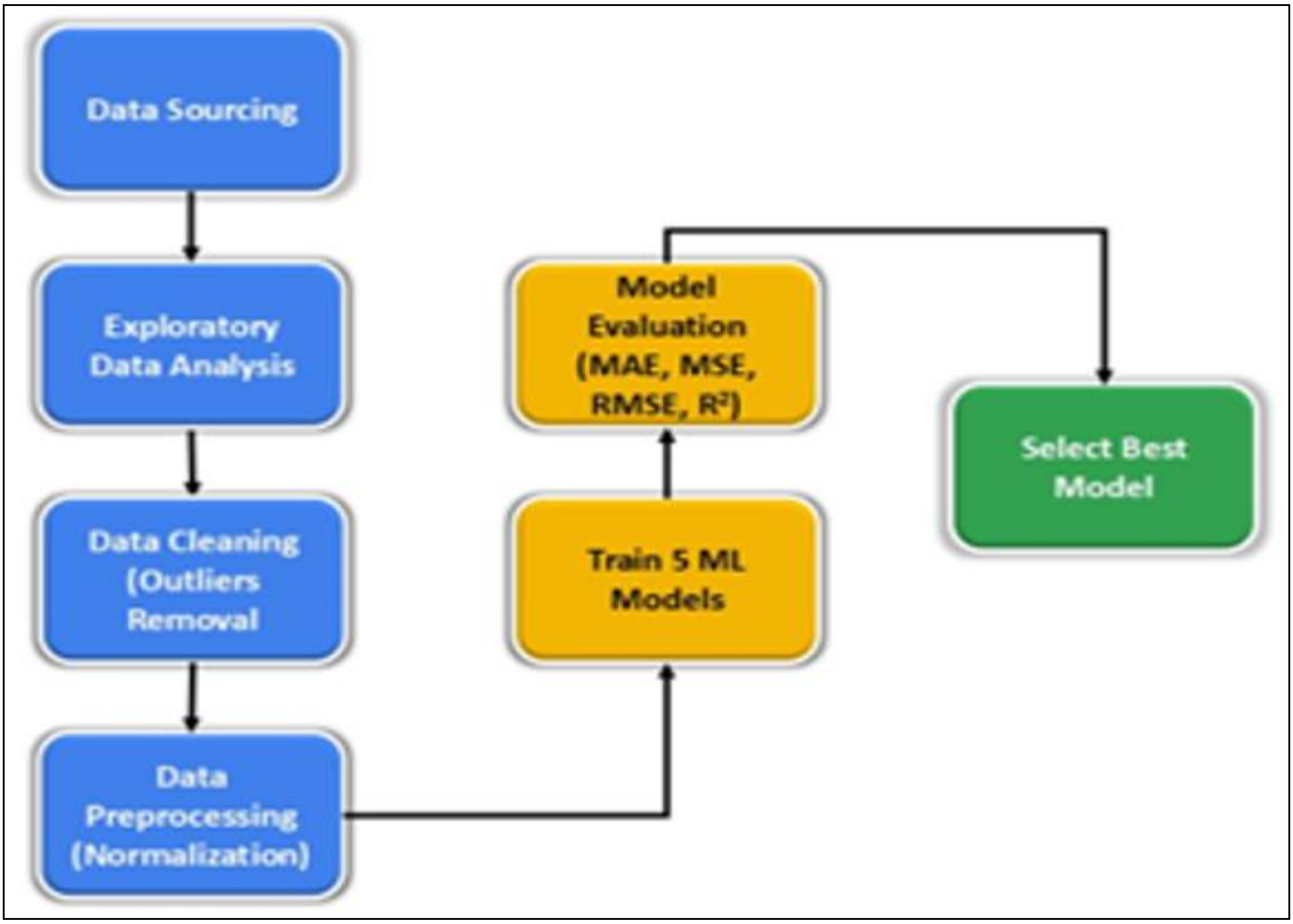

Data Sourcing
Exploratory Data Analysis
Data Cleaning (Outliers Removal
Data Preprocessing (Normalization)
Train 5 ML Models
Model Evaluation (MAE, MSE, RMSE, R²)
Select Best Model

Fig 1 Research Methodology Workflow

The study employed machine learning-based regression analysis. Five models, namely CatBoost, Random Forest, XGBoost, K-Nearest Neighbors (KNN), and Support Vector Regression (SVR), were trained and evaluated using the Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and R^2 Score (Coefficient of Determination), performance metrics.

The evaluation metrics were used to evaluate each of the five models' performances on the test data. These metrics help evaluate the accuracy and quality of regression models by quantifying the difference between predicted and actual values on different scales and perspectives. The predictive models were developed and evaluated using Python in a Jupyter Notebook environment. Libraries such as Pandas, Scikit-learn, XGBoost, CatBoost, Matplotlib, and Seaborn were employed to build, train, and assess the performance of the regression algorithm.

III. RESULTS AND DISCUSSION

The summary statistics in Table 2 highlight a clear distinction between the dynamic and highly variable nature of production rates versus the more stable and consistent characteristics of the underlying reservoir properties. The wide range and high variability in production rates, particularly the extreme values in condensate and water suggest fluctuating operational conditions, transient reservoir responses, or potential anomalies such as water coning, breakthrough events, or liquid loading. Meanwhile, the low variability in reservoir properties suggests a relatively homogeneous physical environment or consistent measurement conditions.

Table 2 Summary Statistics of Dataset Used in the Study

Features	Mean	Min	Q1	Median	Q3	Max
Condensate	2175.82	0.00	43.00	248.00	1752.65	67964.69
Gas	774467.90	6385.21	369656.90	618503.80	955790.20	4182610.00
Water	19005.08	0.00	72.68	1219.47	11575.23	449842.53
Perm	2055.32	1535.00	1535.00	2095.00	2095.00	3500.00
Dew_point	1895.88	1215.00	1896.00	1896.00	2195.00	2195.00
GWC	5393.90	5200.00	5200.00	5462.00	5462.00	5811.00
Res_Temp	136.56	132.00	137.00	137.00	138.00	138.00
Res_Press	2289.34	2268.00	2268.00	2278.00	2322.00	2322.00

The relationships between the features and the water production rate were analyzed and presented in Figures 2 and 3.

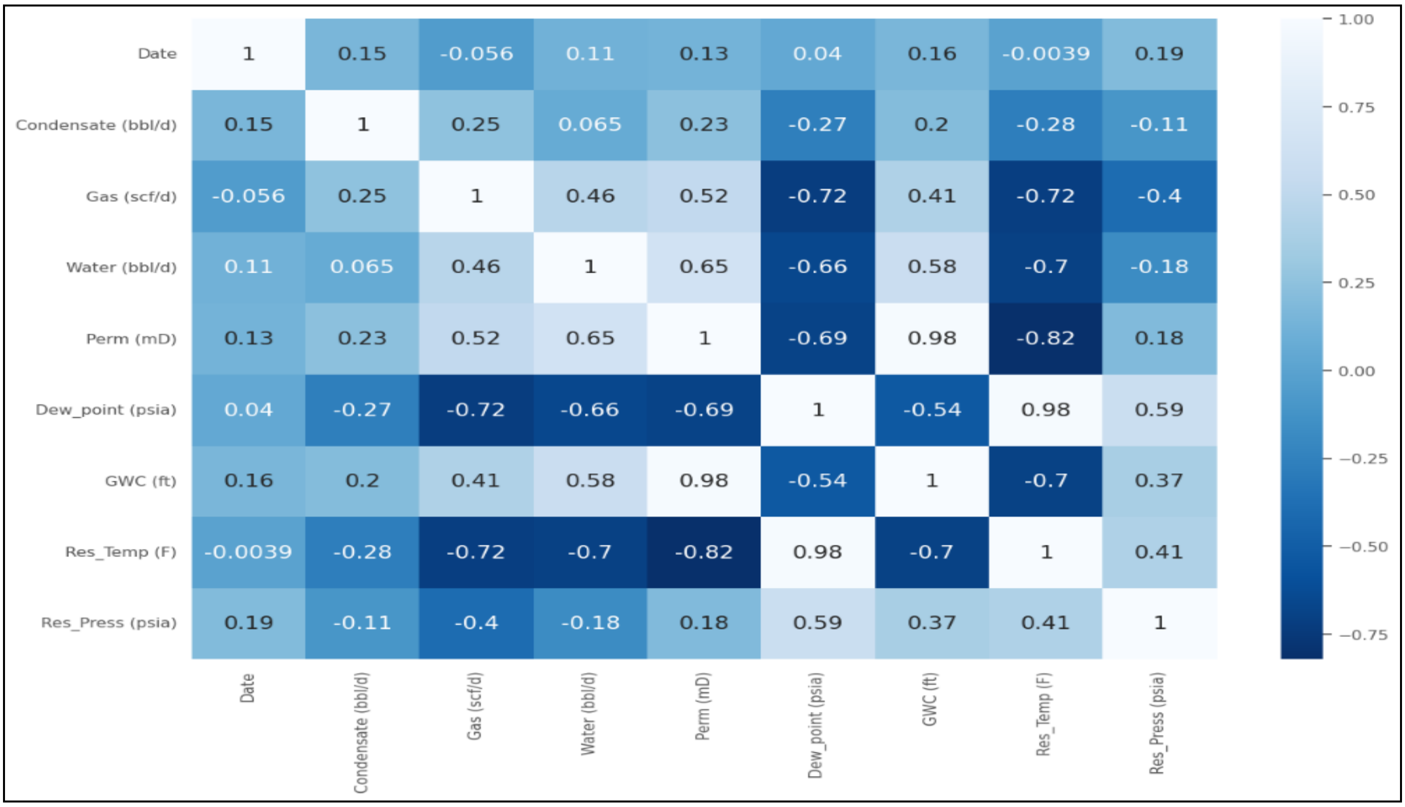
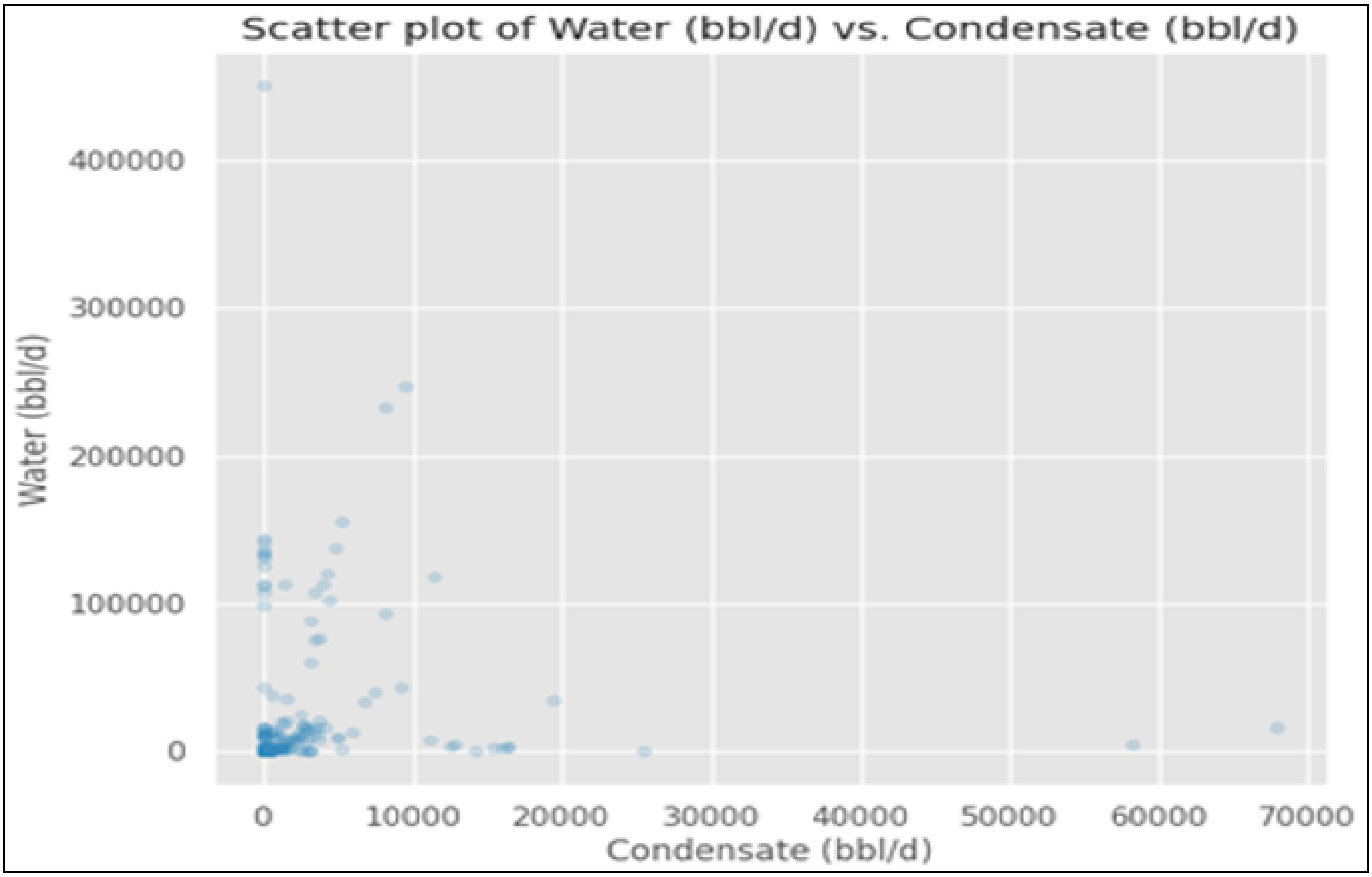
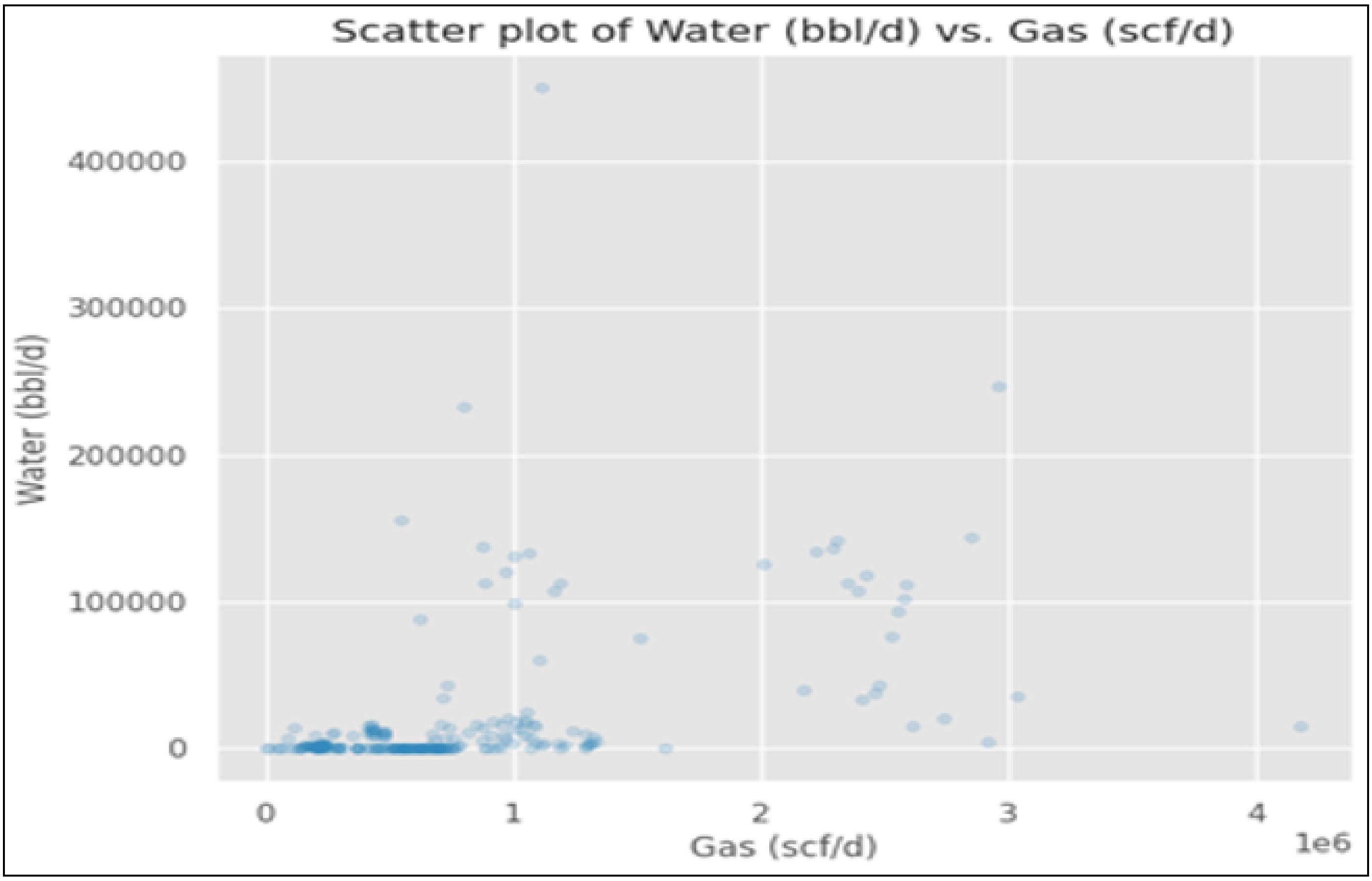


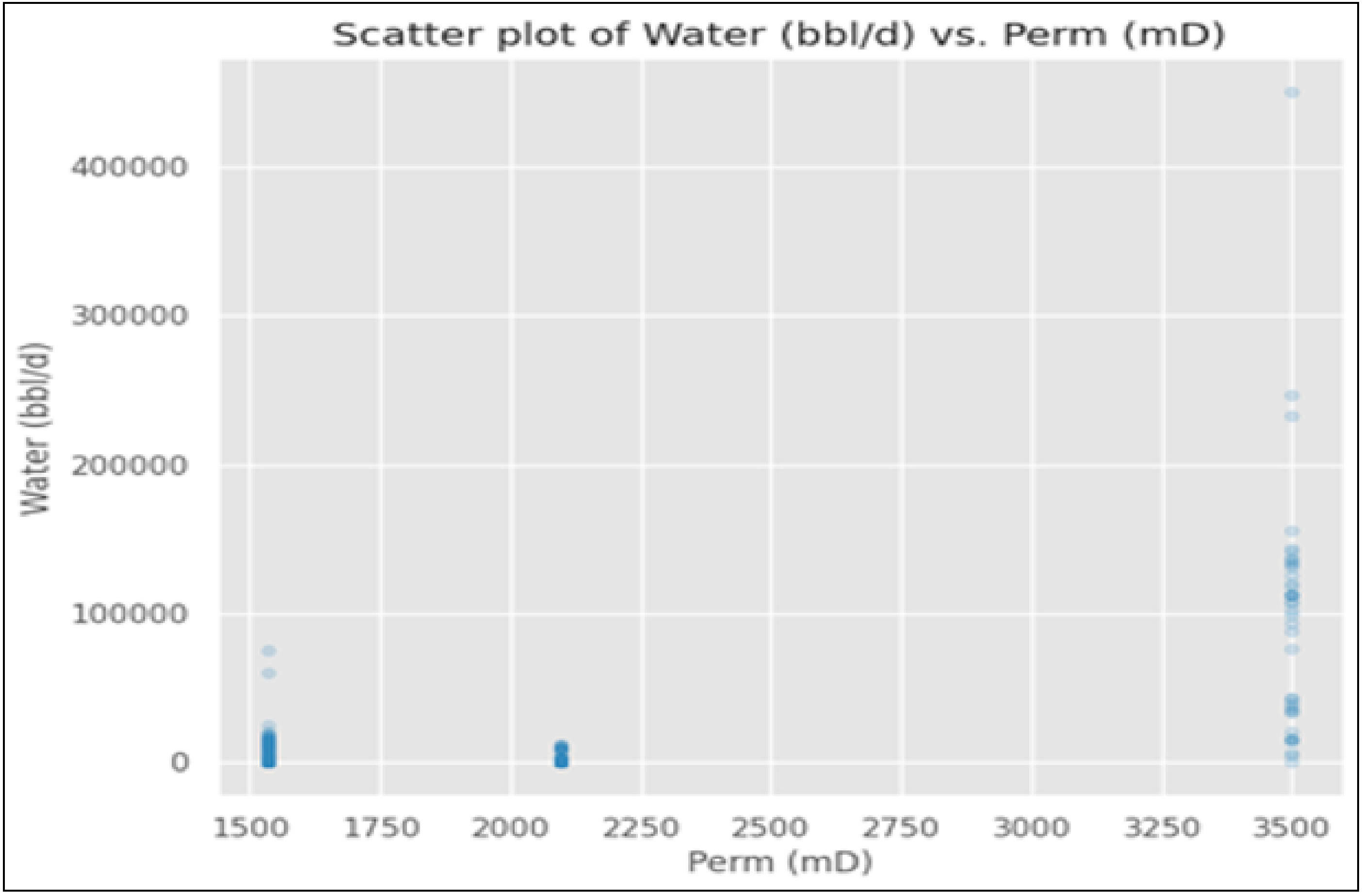
Fig 2 Correlation Matrix of the Features


Scatter plot of Water (bbl/d) vs. Condensate (bbl/d)
Water (bbl/d)
400000
300000
200000
100000
0
0
10000
20000
30000
40000
50000
60000
70000
Condensate (bbl/d)

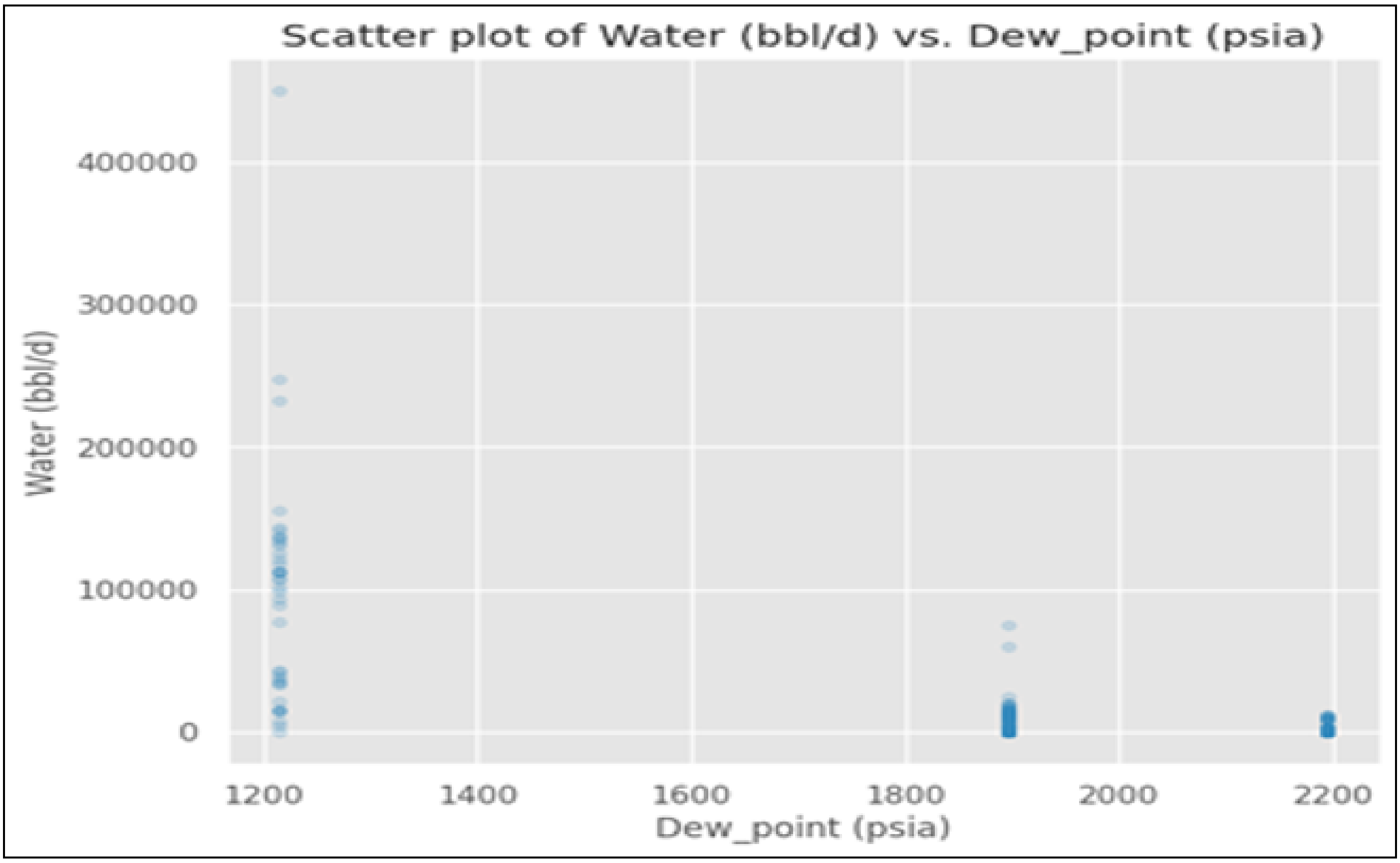
(a)


Scatter plot of Water (bbl/d) vs. Gas (scf/d)
Water (bbl/d)
400000
300000
200000
100000
0
0
1
2
3
4
1e6
Gas (scf/d)

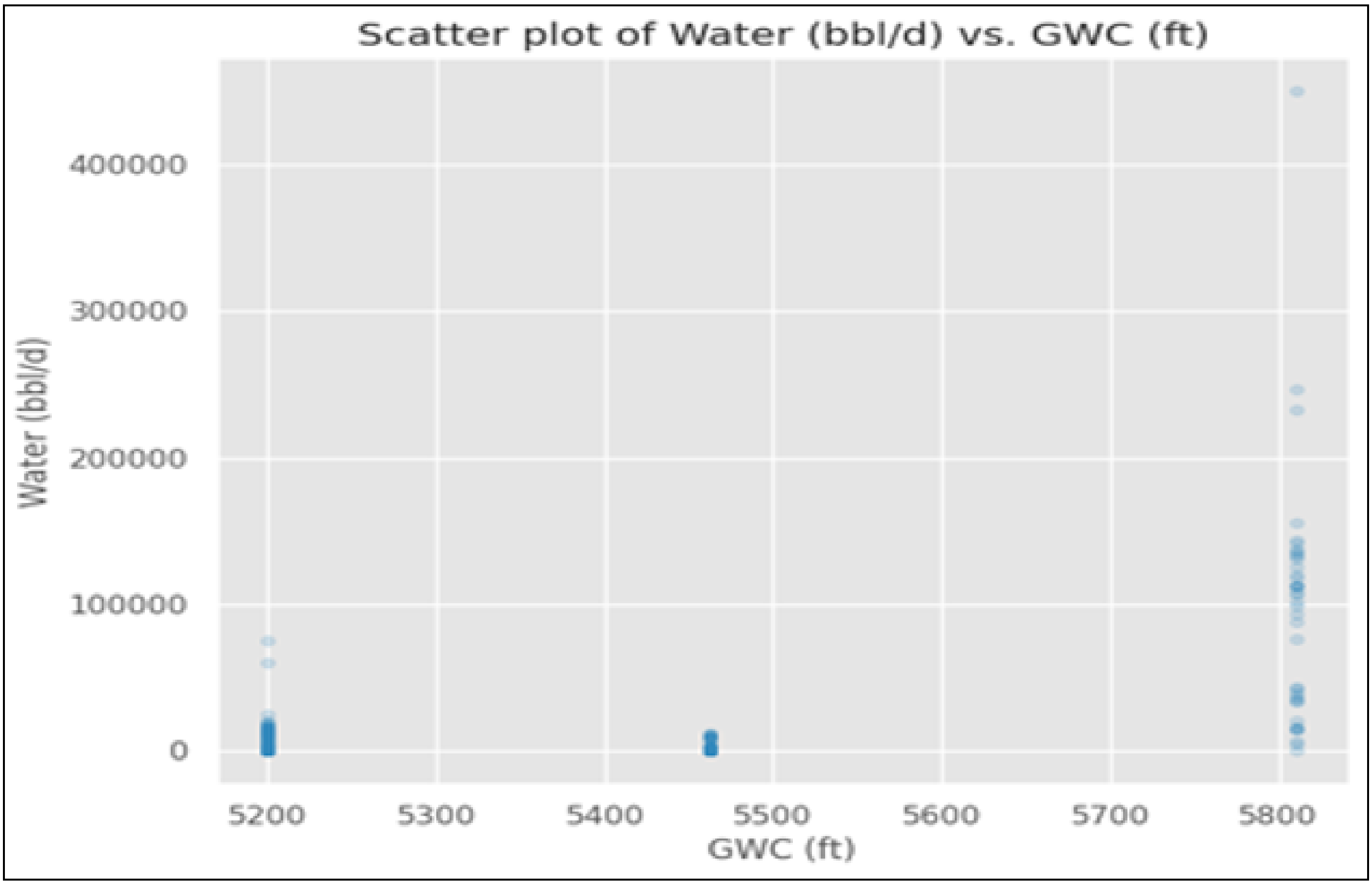
(b)


Scatter plot of Water (bbl/d) vs. Perm (mD)
Water (bbl/d)
0
100000
200000
300000
400000
1500
1750
2000
2250
2500
2750
3000
3250
3500
Perm (mD)

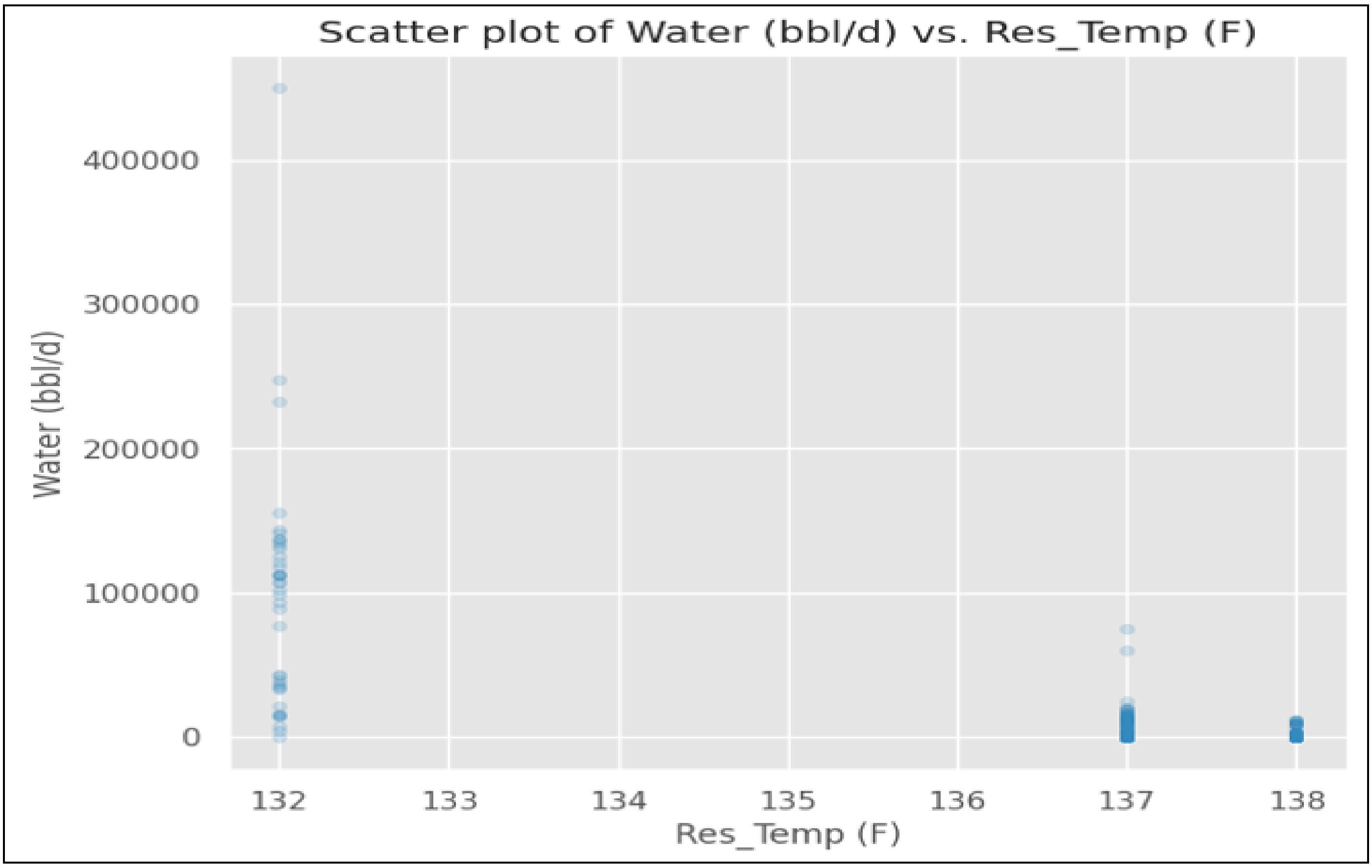
(c)


Scatter plot of Water (bbl/d) vs. Dew_point (psia)
Water (bbl/d)
0
100000
200000
300000
400000
1200
1400
1600
1800
2000
2200
Dew_point (psia)

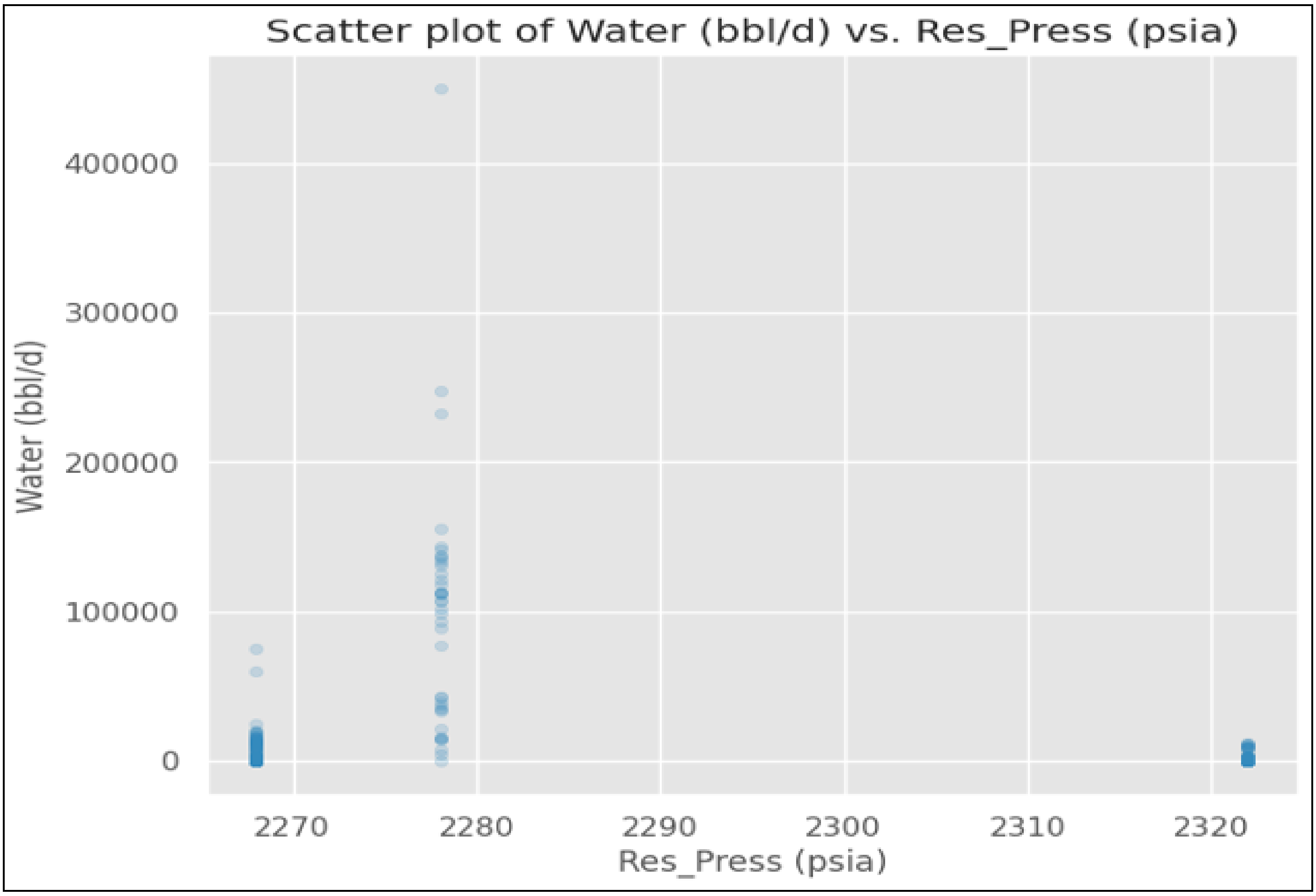
(d)


Scatter plot of Water (bbl/d) vs. GWC (ft)
Water (bbl/d)
0
100000
200000
300000
400000
5200
5300
5400
5500
5600
5700
5800
GWC (ft)

(e)


Scatter plot of Water (bbl/d) vs. Res_Temp (F)
Water (bbl/d)
0
100000
200000
300000
400000
132
133
134
135
136
137
138
Res_Temp (F)

(f)



(g)

Fig 3 Scatter Plots Showing the Features' Relationships with Water Production and (a) Condensate Rate (bbl/d); (b) Gas Rate (scf/d); (c) Permeability (mD); (d) Dew Point Pressure (Psia); (e) GWC (ft); (f) Reservoir Temperature (°F); (g) Reservoir Pressure (Psia).

In Figure 2, the features with the most significant relationships to water production were permeability, reservoir temperature, and dew point. The strongest positive correlation was with permeability, indicating that as the reservoir's ability to allow fluid flow increases, the water production rate also tends to increase. This is a physically logical relationship, as higher permeability facilitates the movement of all fluids, including water. The strongest negative correlations were with reservoir temperature and dew point. This means that as reservoir temperature and dew point pressure increase, the water production rate tends to decrease. These are particularly notable inverse relationships that could point to specific reservoir conditions or fluid properties influencing water production. Figure 3 shows the relationship of the selected features with water production. Some features showed very weak or no significant linear relationship with water production. The correlation with condensate rate is very weak and close to zero, indicating that changes in condensate production are largely unrelated to changes in water production. Similarly, the correlation with reservoir pressure is weak, suggesting that the factor has a very limited influence on the water production rate in this dataset. However, it is crucial to recognize that these are just linear correlations. A high correlation coefficient does not imply causation; it merely shows a strong linear association. For example, while higher permeability is strongly correlated with greater water production, this might be due to other underlying geological factors that influence both variables.

The MAE, RMSE, and R-squared score of each of the five models are displayed in Figures 4 and 5. The results reveal significant differences in the predictive ability of the machine learning models applied to water production forecasting in the Niger Delta field.

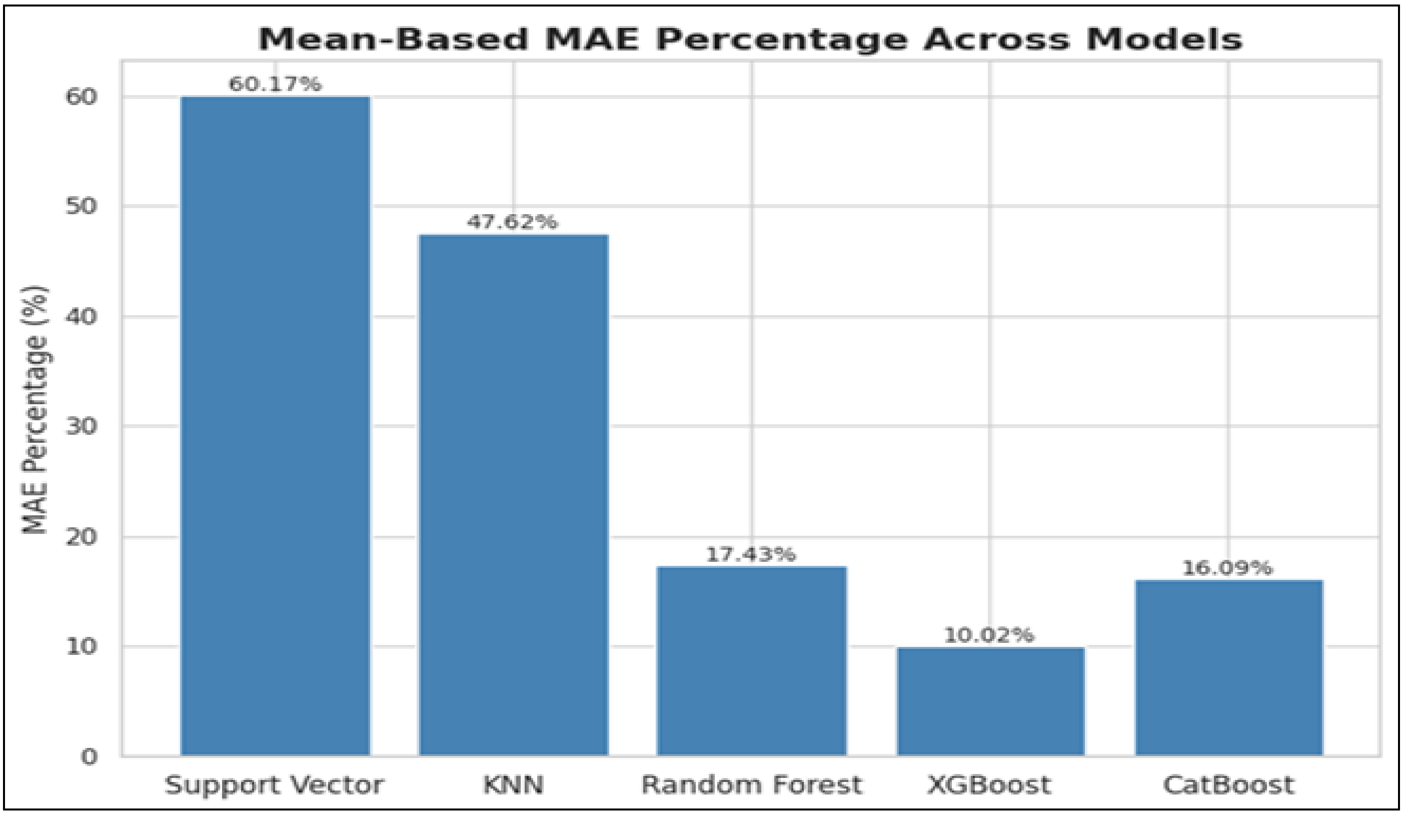


Fig 4 MAE Scores of the Five ML Models

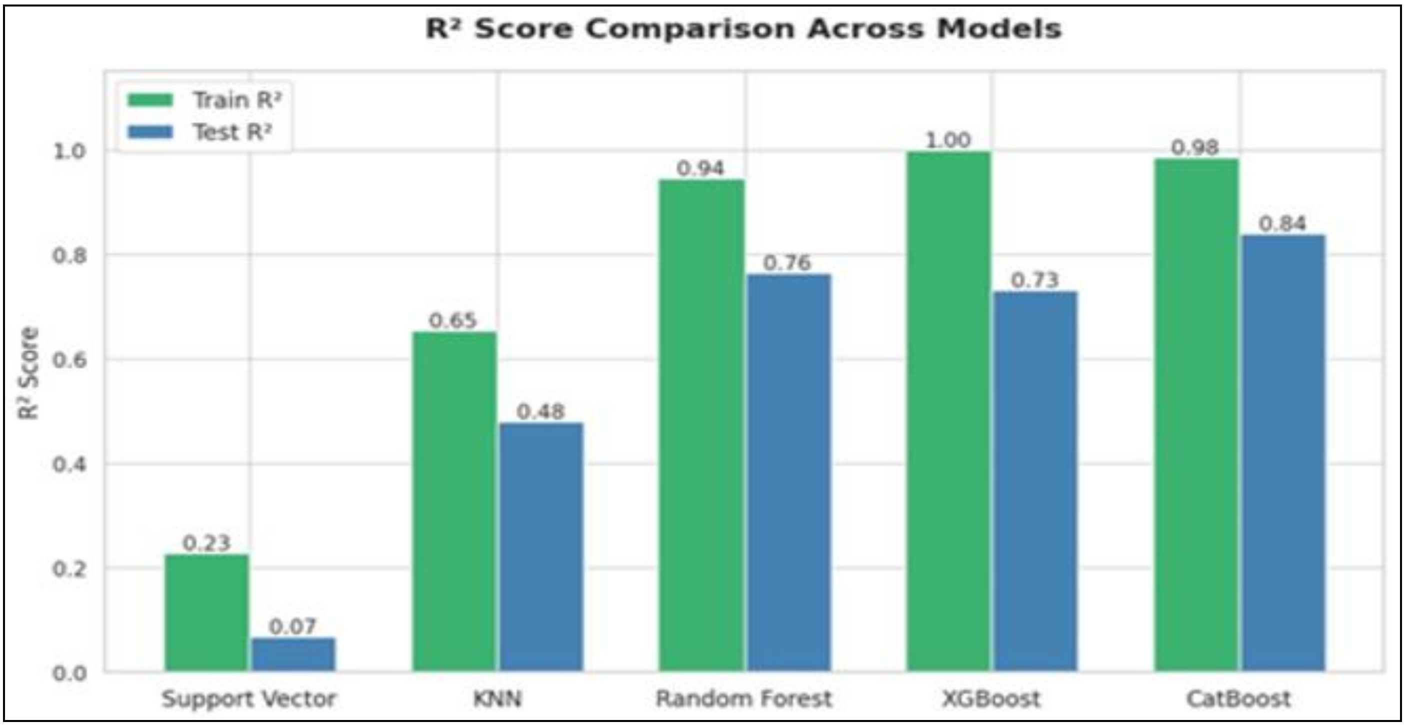


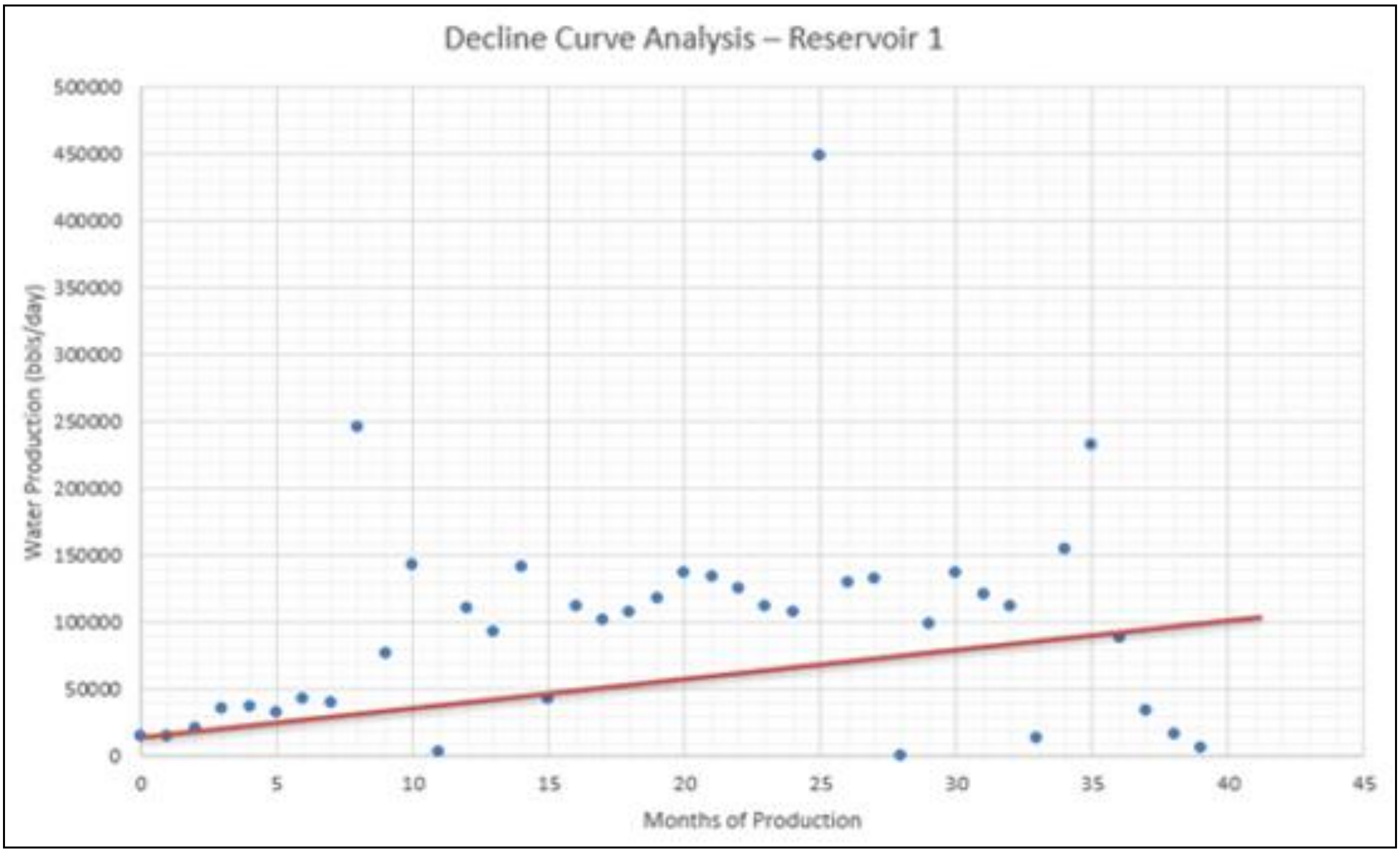
Fig 5 R^2 Scores of the Five ML Models

Using the conventional method (decline curve analysis), Figure 6 shows the mean absolute percentage error in forecasting water production in the natural dry gas wells.

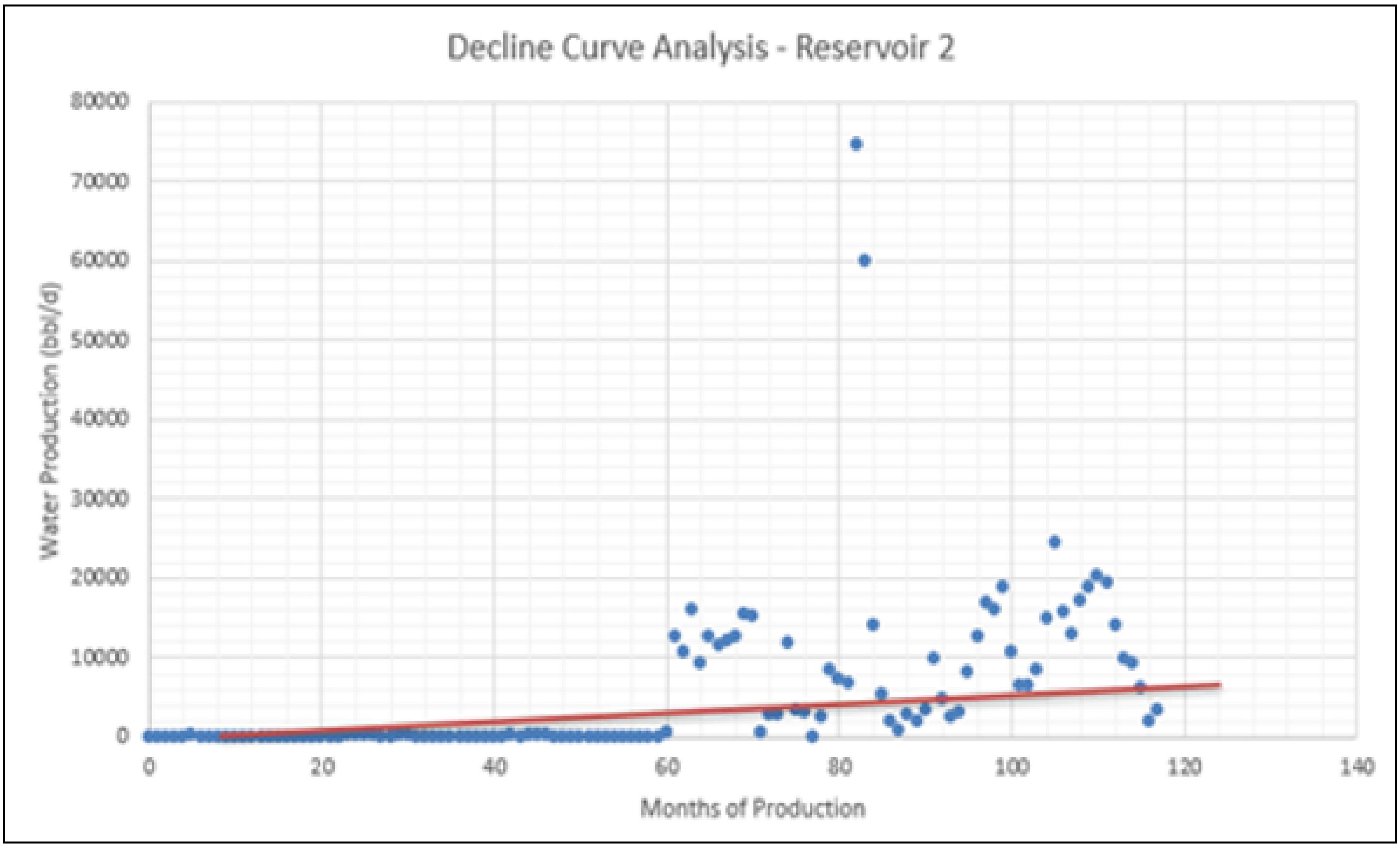
➢ *The Steps Taken in Figure 6 Involves the Following:*

- Compile the reservoir's water production monthly.
- The production rate was plotted against months of production.

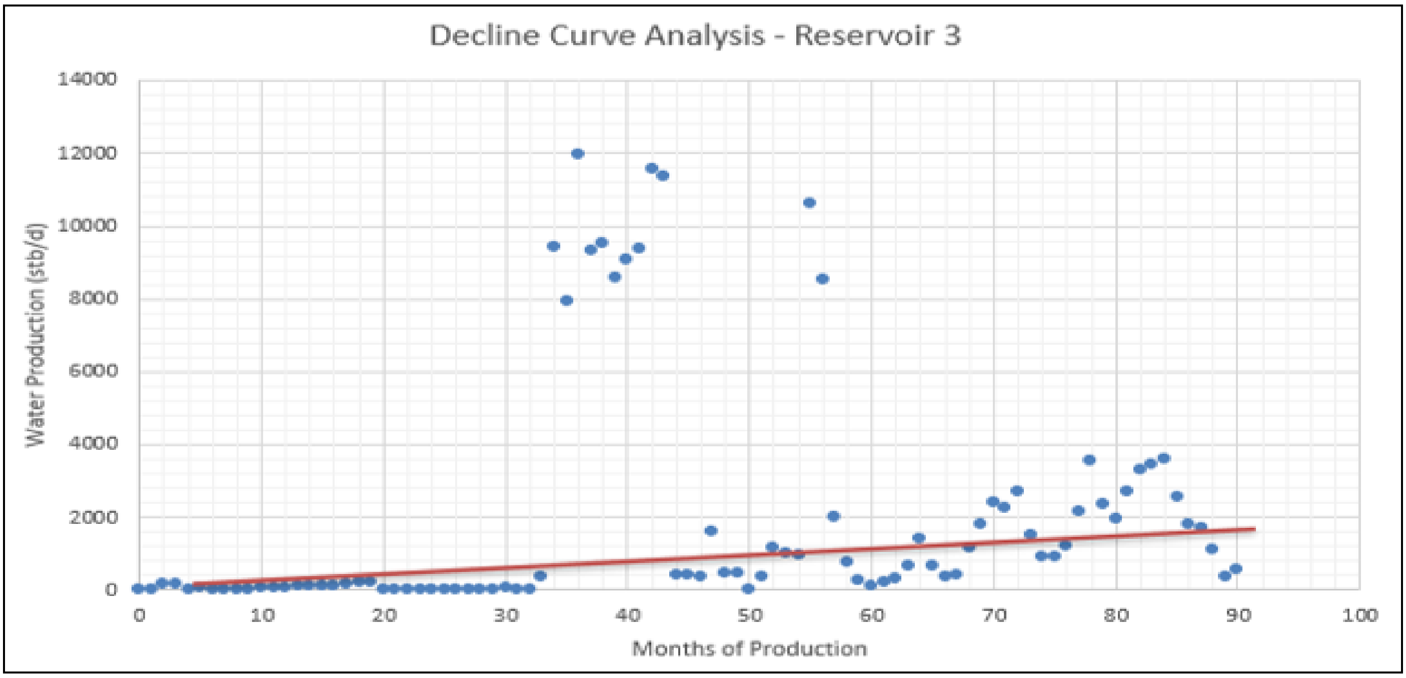
Using exponential decline, the historical data is fitted using regression analysis and then projected into the future to predict water production


Decline Curve Analysis – Reservoir 1
Water Production (bbls/day)
500000
450000
400000
350000
300000
250000
200000
150000
100000
50000
0
0
5
10
15
20
25
30
35
40
45
Months of Production

(a)


Decline Curve Analysis - Reservoir 2
Water Production (bbl/d)
80000
70000
60000
50000
40000
30000
20000
10000
0
0
20
40
60
80
100
120
140
Months of Production

(b)



(c)

Fig 6 DCA Model for Water Production Forecast

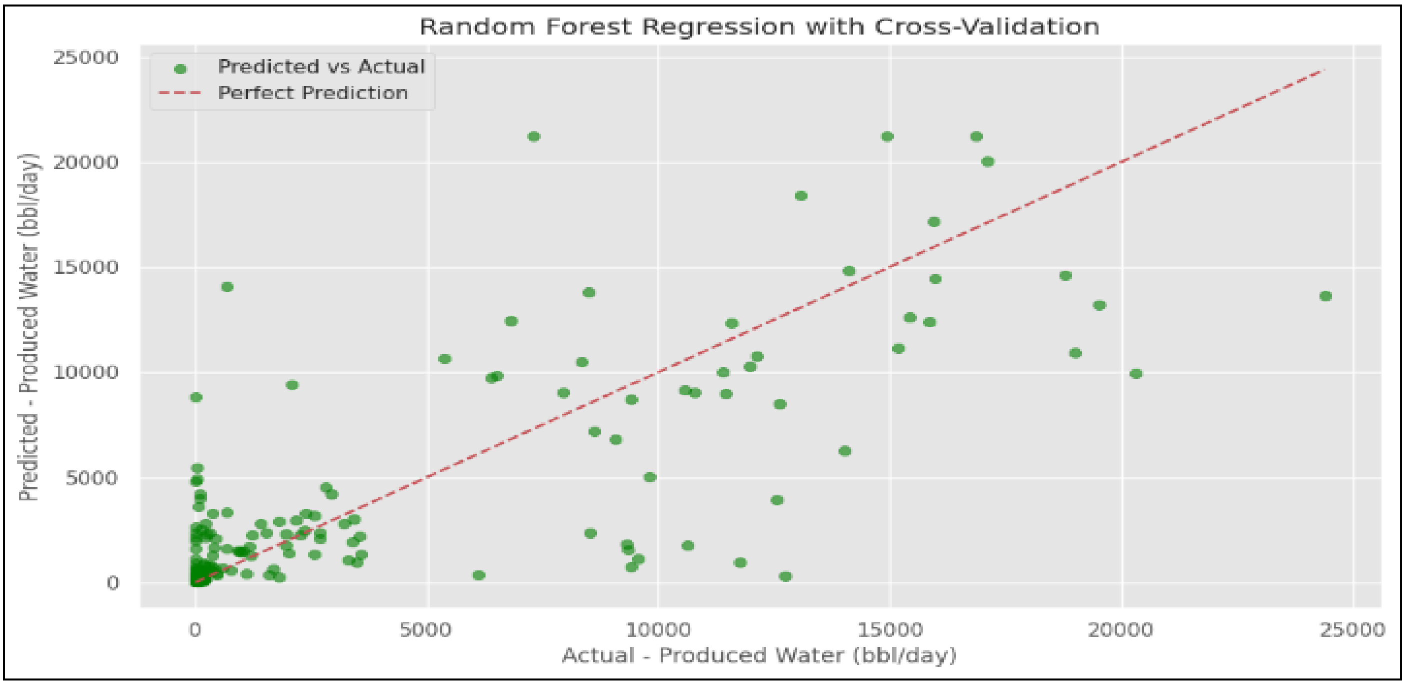
In Figure 6(a), the predicted water production forecast differed from the actual water production with an initial water production of 14,856 bbls/month by a mean absolute error of 63% for reservoir 1.

In Figure 6(b), the predicted water production forecast differed from the actual water production with an initial water production of 0.68 bbls/month by a mean absolute error of 25% for reservoir 2.

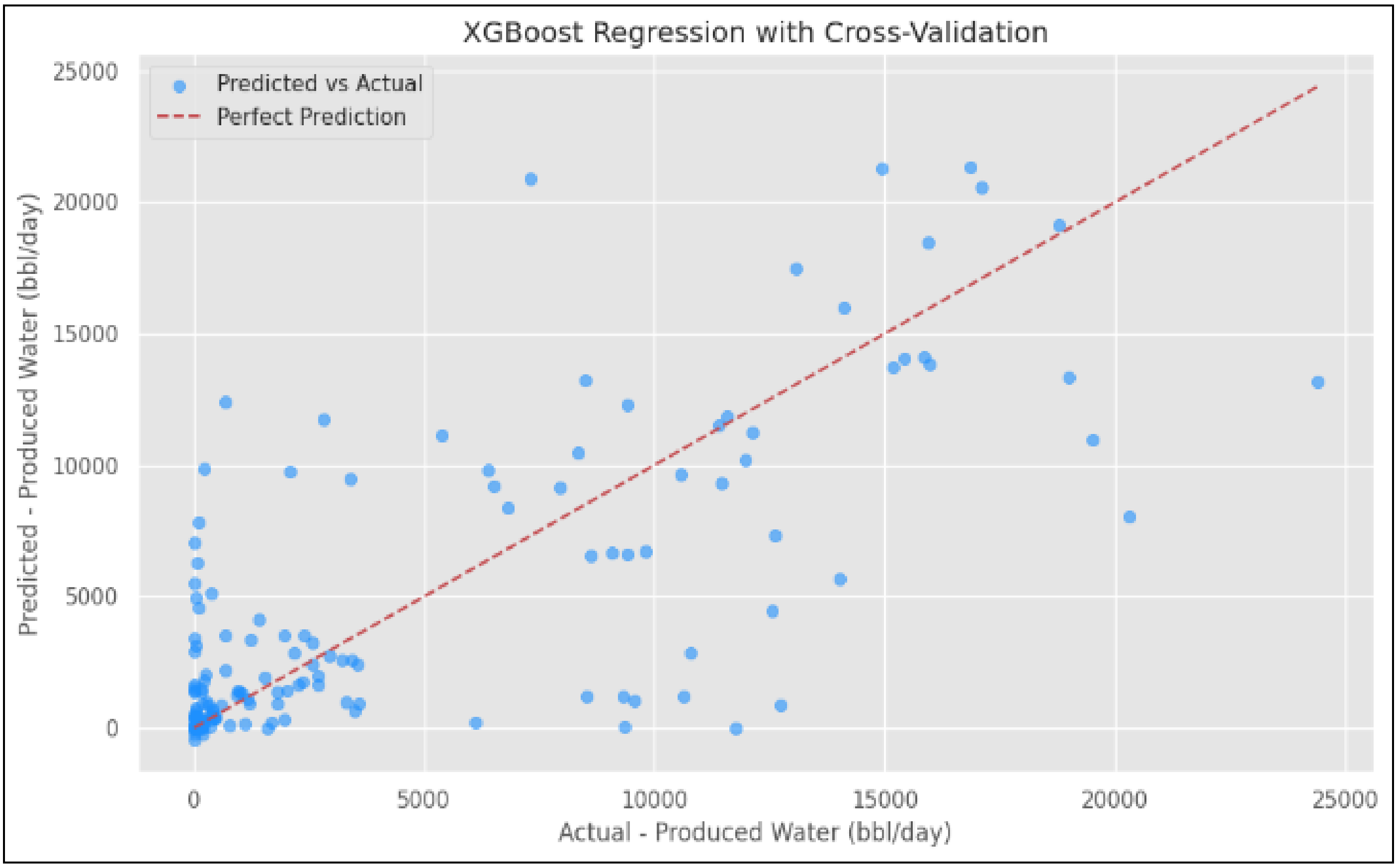
In Figure 6(c), the predicted water production forecast differed from the actual water production with an initial water production of 2.89 bbls/month by a mean absolute percentage error of 72% for reservoir 2.

Hence, the DCA model for forecasting water production for reservoir 1, 2 and 3 showed mean absolute percentage error of 63%, 25% and 72% for (a), (b) and (c) respectively where none were less than the CatBoost or XGBoost model.

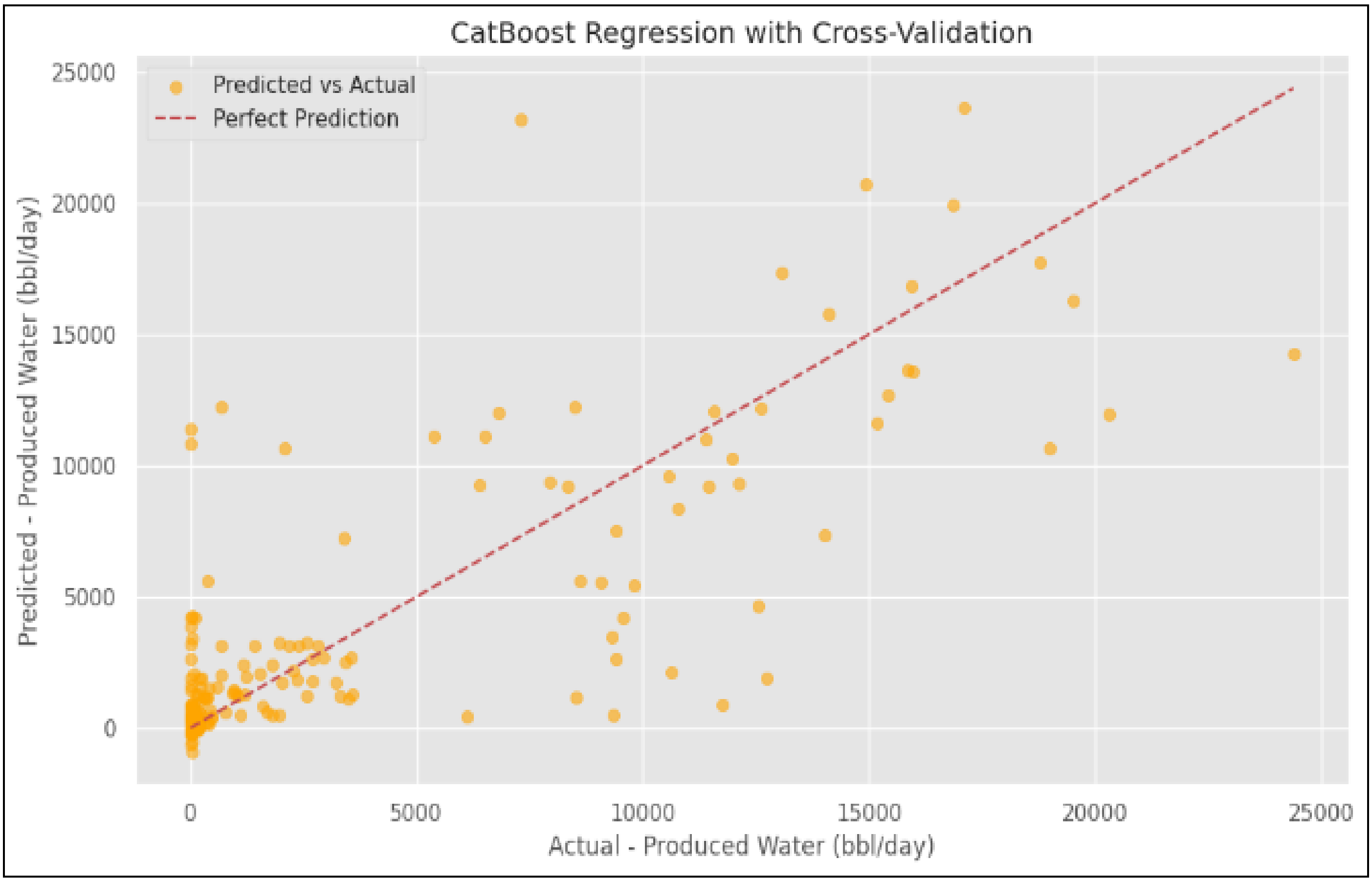
Figure 7 shows the actual production compared with the predicted production for the five machine learning models used.



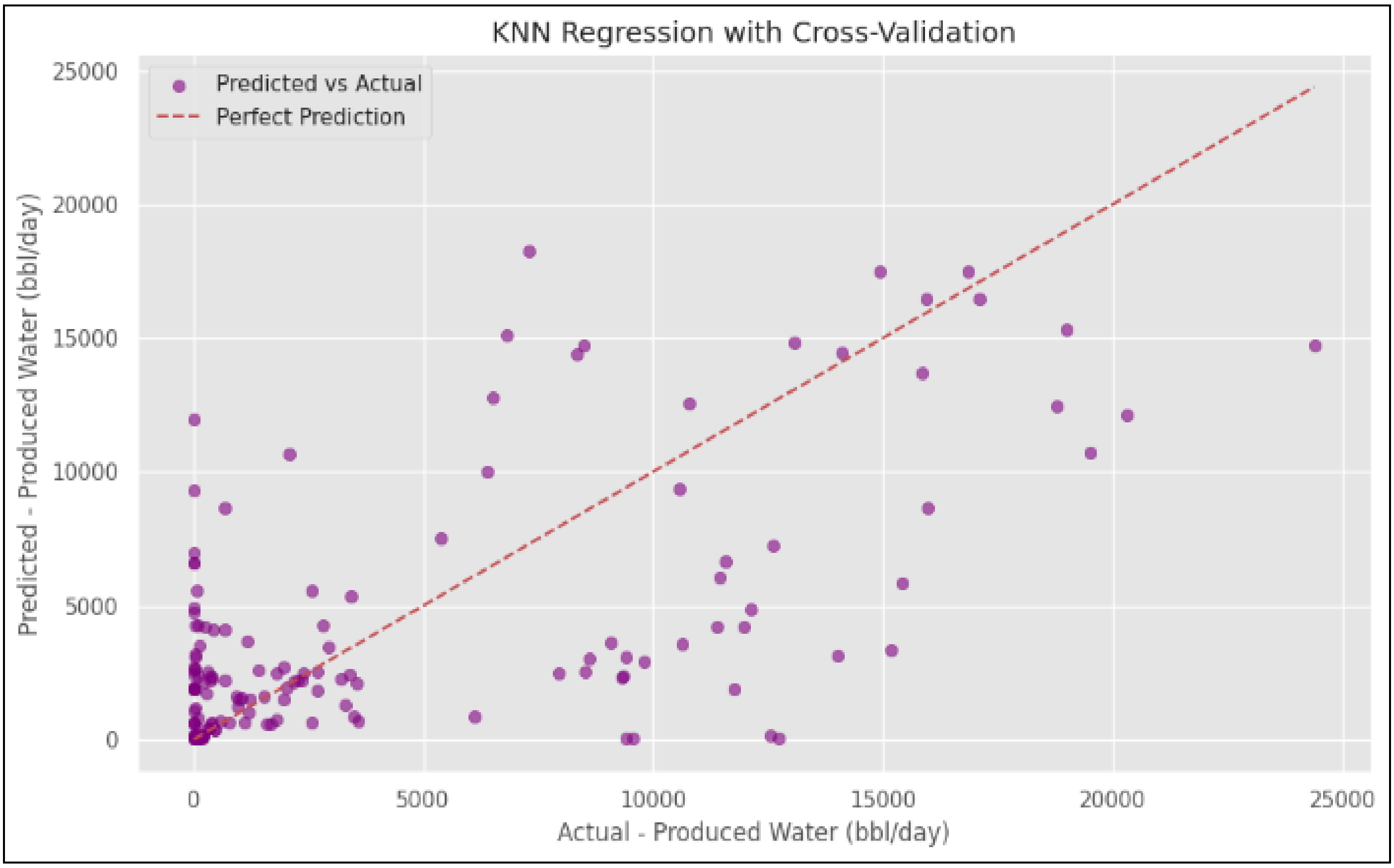
(a)


XGBoost Regression with Cross-Validation
Predicted vs Actual
Perfect Prediction
Predicted - Produced Water (bbl/day)
Actual - Produced Water (bbl/day)
0
5000
10000
15000
20000
25000

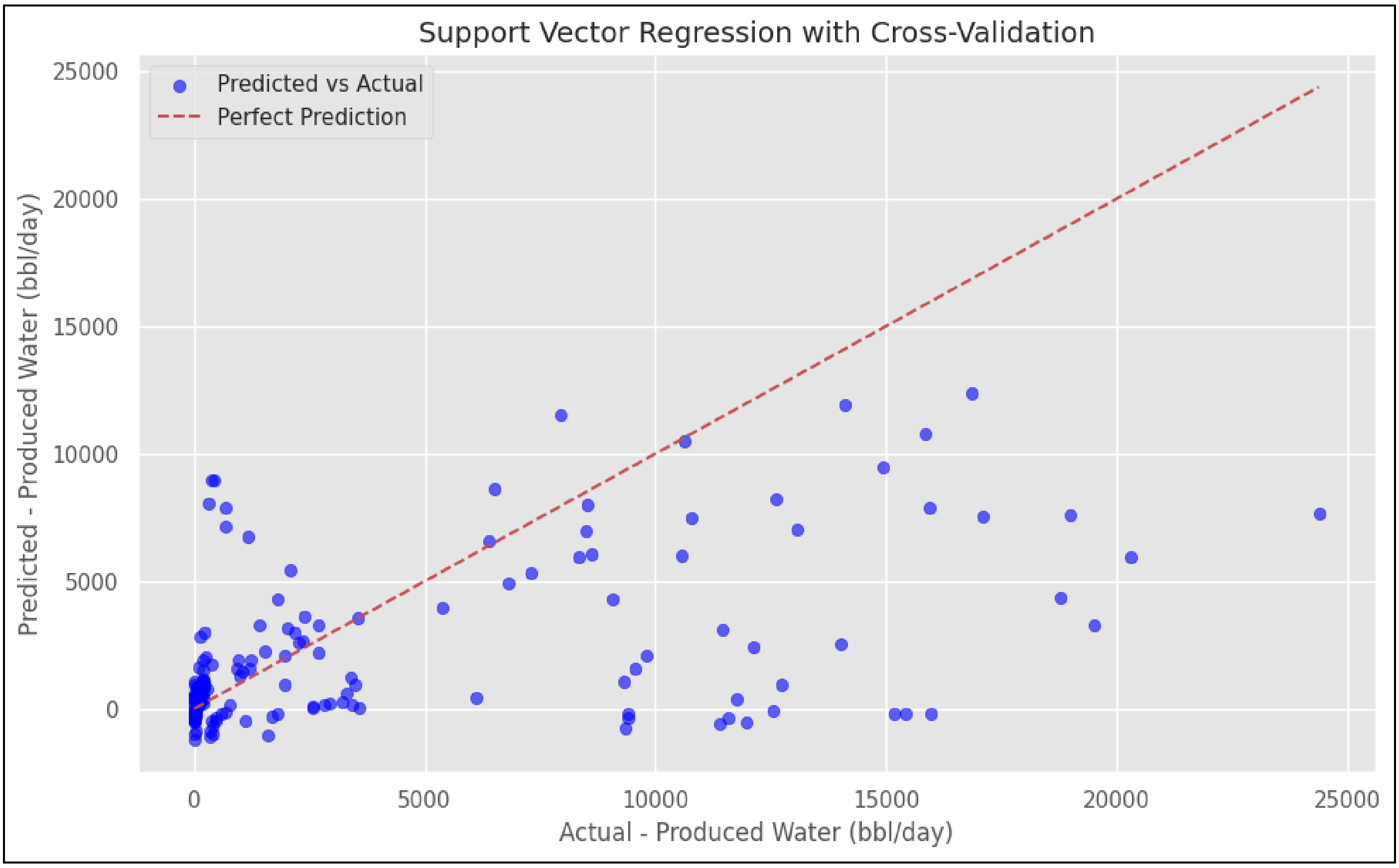
(b)


CatBoost Regression with Cross-Validation
Predicted vs Actual
Perfect Prediction
Predicted - Produced Water (bbl/day)
Actual - Produced Water (bbl/day)
0
5000
10000
15000
20000
25000

(c)



(d)



(e)

Fig 7 Plot of Actual vs. Predicted. (a) Random Forest; (b) XGBoost. (c) CatBoost. (d) KNN (e) SVM

In Figure 7, the CatBoost model, with an R² of 0.84 and the lowest error values (MAE = 1508.73 bbl/day, RMSE = 2468.45 bbl/day), provides the most dependable forecasts of daily water production. For example, if the actual water production from a well is around 20,000 bbl/day, CatBoost's predictions would typically be within ±1,500 to 2,500 bbl/day of that value. In operational terms, this level of accuracy is crucial for designing water handling facilities such as separators, pumps, and disposal systems. Engineers could rely on these predictions to size equipment properly and avoid both under-capacity (leading to facility bottlenecks) and over-design (leading to wasted capital).

The Random Forest model, with an R² of 0.78, also produced strong results, with average errors around 1,800–3,000 bbl/day. If a well was expected to produce 15,000 bbl/day of water, Random Forest might forecast between 12,000 and 18,000 bbl/day. While this is still useful for field development planning, such an error margin could cause challenges in wells where production swings are highly dynamic. For instance, underestimating water by 3,000 bbl/day could overwhelm existing disposal capacity, forcing emergency flaring or production curtailment. Thus, Random Forest may work well for field-scale estimates but may be slightly less reliable for well-level operational decisions.

The XGBoost model, with an R² of 0.73, is somewhat less precise. For a producing well delivering 10,000 bbl/day of water, XGBoost's forecasts might deviate by ±2,500 to 3,200 bbl/day. In practice, this could be acceptable for long-term projections of water handling requirements across an asset, where errors average out across multiple wells. However, for short-term facility scheduling, such as allocating produced water between injection and disposal wells, XGBoost's higher error could lead to overloaded injection pumps or downtime in treatment plants.

The KNN model, with an R² of 0.48, performed poorly. Its error range of about 2,600–4,400 bbl/day is too wide for practical use. For instance, if a surface facility is designed to handle 25,000 bbl/day of produced water, and KNN underestimates actual flow by 4,000 bbl/day, the excess water could exceed separator or reinjection capacity. This would force operators to either choke back production (reducing gas output) or incur additional unplanned disposal costs. On the other hand, if it overpredicts water production, capital might be wasted on oversized water treatment equipment that is never fully utilized.

Finally, the Support Vector Regression model, with an R² of only 0.07, showed extremely poor predictive ability. Its errors of over 3,300–5,900 bbl/day are unacceptable in a field context. For example, if actual water production is 8,000 bbl/day, SVR might predict anywhere from 2,000 to 14,000 bbl/day. This kind of inaccuracy would make it impossible to design or operate water handling infrastructure effectively. If operators relied on these predictions, they could end up with severely undersized disposal wells or grossly oversized treatment plants, both of which would lead to major economic and operational inefficiencies.

Among these models, though XGBoost demonstrated the highest predictive accuracy, achieving a Percentage Mean Absolute Error (MAE) of 10%, CatBoost model with a predictive accuracy of 16% is preferred for prediction as it has better training and test data model compared to XGBoost that has a problem with overfitting.

Hence, the findings emphasize the superiority of ensemble-based models (CatBoost, Random Forest, and XGBoost) for forecasting water production in complex petroleum systems such as the Niger Delta. These models not only minimize predictive error but also provide better generalization, making them valuable tools for reservoir management and water control strategies.

IV. CONCLUSION

This research investigated the effectiveness of various machine learning models in predicting water production rates from gas wells, using data sourced from the Niger Delta region. The study employed five algorithms, namely CatBoost, Random Forest, XGBoost, K-Nearest Neighbors (KNN), and Support Vector Regression (SVR), on a dataset comprising 249 daily records across eight reservoir and production parameters.

Among these models, though XGBoost demonstrated the highest predictive accuracy, achieving a Percentage Mean Absolute Error (MAE) of 10%, CatBoost model with a predictive accuracy of 16% is preferred for prediction as it has better training and test data model as compared to XGBoost that has a problem in overfitting. This result shows a better predictive accuracy than traditional method of using decline curve analysis.

However, the analysis was based on a dataset restricted to a very limited number of reservoirs and production parameters, which may not fully represent the dynamic complexity of the reservoir. Moreover, the relatively limited sample size, coupled with potential noise in the data, could have reduced the robustness of the predictive models.

DECLARATIONS

➢ *Funding*
No funding was received to assist with the preparation of this manuscript.

➢ *Competing Interest*
The authors have no competing interests to declare that are relevant to the content of this article.

➢ *Author Contribution*
All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by Ebifagha Bebenimibo and Victor Joseph Aimikhe. The first draft of the manuscript was written by Ebifagha Bebenimibo and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

REFERENCES

[1]. Khatib, Z., & Verbeek, P. (2003). Water to value - Produced water management for sustainable field development of mature and green fields. JPT, Journal of Petroleum Technology, 55(1). https://doi.org/10.2118/0103-0026-JPT

[2]. Gao, S., Zhang, J., Liu, H., Ye, L., Zhu, W., & Xiong, W. (2023). Water-gas ratio characteristics and development concepts for water-producing gas reservoirs. In Heliyon (Vol. 9, Issue 8). Elsevier Ltd. https://doi.org/10.1016/j.heliyon.2023.e19201

[3]. Aimikhe V.J and Adeyemi M.A (2019). A Critical Evaluation of Natural Gas-water Formula Correlations. Journal of Scientific Research & Reports. 25(6): 1-20

[4]. Radwan, M. F. (2017). Feasibility evaluation of using downhole gas-water separation technology in gas reservoirs with bottom water. SPE Middle East Oil and Gas Show and Conference, MEOS, Proceedings, 2017-March. https://doi.org/10.2118/183739-ms

[5]. Wei, M., Ren, K., Duan, Y., Chen, Q., & Dejam, M. (2019). Production decline behavior analysis of a vertical well with a natural water influx/waterflood. Mathematical Problems in Engineering, 2019. https://doi.org/10.1155/2019/1683989

[6]. Bin Marta, Hammouda, Tantawy, Khamis, & Wahba. (2023a). Diagnosing and Controlling Excessive Water Production: State-of-the-Art Review. Journal of Petroleum and Mining Engineering,0(0),0–0. https://doi.org/10.21608/jpme.2023.233102.

[7]. Falcon, & Barbosa (2013). State-of-the-art Review of Liquid Loading in Gas Wells. DGMK.

[8]. Joseph, A., & Ajienka, J. A. (2010). A review of water shutoff treatment strategies in oil fields. Society of Petroleum Engineers - Nigeria Annual International Conference and Exhibition 2010, NAICE, 1. https://doi.org/10.2118/136969-ms

[9]. Roozshenas, A. A., Hematpur, H., Abdollahi, R., & Esfandyari, H. (2021). Water Production Problem in Gas Reservoirs: Concepts, Challenges, and Practical Solutions. Mathematical Problems in Engineering, 2021. https://doi.org/10.1155/2021/9075560

[10]. Francis, M., & Ogbeide, P. O. (2021). Application of Chan Plot in Water Control Diagnostics for Field Optimization: Water/Gas Coning and Cusping. NIPES - Journal of Science and Technology Research, 3(4), 227–232. https://doi.org/10.37933/nipes/3.4.2021.23x

[11]. Guo, B., & Lee, R. L. H. (1993). Simple approach to optimization of completion interval in oil/water coning systems. SPE Reservoir Engineering (Society of Petroleum Engineers), 8(4). https://doi.org/10.2118/23994-pa

[12]. Okwananke, A., & Isehunwa, S. O. (2008). Analysis of Water Cresting in Horizontal Wells. SPE 119733.

[13]. Pirson, S. J., & Mehta, M. M. (1967). A Study of Remedial Measures for Water-Coning By Means of a Two-Dimensional Simulator. https://doi.org/10.2118/1808-ms

[14]. Inikori, S. (2002). Numerical study of water coning control with Downhole Water Sink (DWS) completions in vertical and horizontal wells. https://repository.lsu.edu/gradschool_dissertations

[15]. Reynolds, R. R. (2003). Produced Water and Associated Issues. Oklahoma Geological Survey.

[16]. Muskat, M., & Wycokoff, R. D. (1935). An Approximate Theory of Water-coning in Oil Production. Transactions of the AIME, 114(01). https://doi.org/10.2118/935144-g

[17]. Chaperon, I. (1986). Theoretical study of coning toward horizontal and vertical wells in anisotropic formations: Subcritical and critical rates. Proceedings - SPE Annual Technical Conference and Exhibition. https://doi.org/10.2523/15377-ms

[18]. Fetkovich, M. J., Reese, D. E., & Whitson, C. H. (1999). Application of a general material balance for high-pressure gas reservoirs. SPE Reprint Series, 52.

[19]. Nmegbu, G., Festus Awara, L., Bariakpoa Kinate, B., Cgj, N., Festus, L., & Bariakpoa KINATE, B. (2020). Diagnosis and Control of Excessive Water Production in Niger Delta Oil Wells. International Journal of Advancements in Research & Technology, 9(10). https://www.researchgate.net/publication/345682999

[20]. Guo, B., Lyons, W. C., & Ghalambor, A. (2007). Petroleum Production Engineering, A Computer-Assisted Approach. In Petroleum Production Engineering, A Computer-Assisted Approach. https://doi.org/10.1016/B978-0-7506-8270-1.X5000-2

[21]. Poe, B. D. (2003). Production Diagnostic Analyses with Incomplete or No Pressure Records. Proceedings - SPE Annual Technical Conference and Exhibition. https://doi.org/10.2118/84224-ms

[22]. Yortsos, Y. C., Choi, Y., Yang, Z., & Shah, P. C. (1999). Analysis and interpretation of water/oil ratio in waterfloods. SPE Journal, 4(4). https://doi.org/10.2118/59477-PA

[23]. Da Silva, D. V. A., & Jansen, J. D. (2015). A Review of Coupled Dynamic Well-Reservoir Simulation. 2nd IFAC Workshop on Automatic Control in Offshore Oil and Gas Production, May 27-29, Florianópolis, Brazil

[24]. Rabiei, M., Gupta, R., Cheong, Y. P., & Soto, G. A. S. (2009). Excess water production diagnosis in oil fields using ensemble classifiers. Proceedings - 2009 International Conference on Computational Intelligence and Software Engineering, CiSE 2009. https://doi.org/10.1109/CISE.2009.5362732

[25]. Ha, J., Kambe, M., & Pe, J. (2011). Data Mining: Concepts and Techniques. In Data Mining: Concepts and Techniques. https://doi.org/10.1016/C2009-0-61819-5