ABUAD Journal of Engineering Research and Development (AJERD)
ISSN (online): 2645-2685; ISSN (print): 2756-6811

Volume 8, Issue 2, 233-241

Improved Correlation of Oil Recovery Factor for Water Driven Reservoirs in the Niger Delta

Ayodele DANIEL[1], Faisal Olasubomi OLANIPEKUN[2], Sunday OloruntobaISEHUNWA[2]

[1]*Afe Babalola University, Km 8.5, Ado-Ikare road, Ado-Ekiti, Ekiti state, Nigeria*
daniela@abuad.edu.ng

[2]*University of Ibadan, Ibadan, Oyo state, Nigeria*
olanipekufo@ui.edu.ng/isehunwaso@ui.edu.ng

Corresponding Author: daniela@abuad.edu.ng, +2348122993053
Date Submitted: 04/04/2025
Date Accepted: 27/07/2025
Date Published: 29/07/2025

Abstract*: Recovery factor is one of the most important variables for a reservoir engineer as it plays a major role in determining the economic viability of oil and gas projects and by implication what projects to mature. Over the years, many different approaches have been taken to estimating recovery factor of oil and gas reservoirs generally and these include simulations, volumetric method and correlations. All these methods have their high inherent cost except for correlations which are not only easy and quick to use but also low cost. Even though correlations have been developed in the past for the recovery factor of Niger Delta crude, none has employed the data analytics and machine learning techniques. Data from strong water driven crude oil reservoir in the Niger Delta was used in this study. After data cleaning and quality checking, cleaned data was used to train the machine learning model using multiple linear regression algorithms optimized with batch gradient descent method. This was implemented using Python code developed for this work. The model developed had an excellent performance on the training set as the coefficient was about 0.84. The mean absolute error is about 0.018. The results obtained showed better model performance and generalization than any previously existing model.*

Keywords: *Recovery Factor, Python, Machine Learning, Correlations, Model*

1. INTRODUCTION

Recovery factor is the ratio of recoverable amount of hydrocarbon to the hydrocarbon initially in place. It can be mathematically stated as:

$$RecoveryFactor = \frac{UR}{HCIIP} \tag{1}$$

where UR is ultimate recovery and HCIIP is hydrocarbon initially in place.

Recovery factor is in two main categories. These are the primary recovery factor which depends on primary recovery energy and supplementary recovery factor, a function of improved recovery methods adopted for the reservoir. Recovery Factor (RF) may also be typified as Technical and Non- technical. Technical Recovery factor is purely a function of the rock and fluid properties, reservoir geology, and drive mechanism and most importantly the available technologies employed in production. Non-technical recovery, on the other hand, involves other considerations like economics, environment and ecology.

There is a direct correlation between the RF and commercial viability of a particular hydrocarbon accumulation.Recovery factor estimation methods include Analogy Method, Volumetric Method,Performance Methods and Empirical Correlations. While comparing bottom and edge water drive,Nnemgbu et al [15] concluded that bottom water drive results in a 0.4% increase over edge water drive. The method of Khusnutdinova [13] demonstrates that statistical tools are useful for determination of recovery factor. Knowing the primary RF helps to establish a target oil or gas reservoir for implementation of improved recovery processes or otherwise. It is therefore, necessary to develop ways to know the recovery factor of a reservoir prior to commencement of production. Such methods need to be simple, yet effectivefor predicting oil and gas recovery factor. Authors such as Craze & Buckley (1945), Muskat& Taylor [14], Guthrie & Greenberger [11], Arps & Roberts [4], Isehunwa & Nwankwo [12], Onolemhemhen & Isehunwa [16], etc have done extensive work on recovery factor correlations.

This paper presents an empirical correlation in reservoirs under water drive obtained by using Python to determine a linear regression model. The model relates various reservoir and fluids properties with recovery factor and further closes the gap between actual and predicted results for water driven reservoirs in the Niger Delta.

2. MATERIALS AND METHODS

Data cleaning, pre-processing, feature selection, model training, model testing and validation were performed by programming using the Python programming language. The following libraries: NumPy, Pandas, Scikit-learn, Matplotlib, Seaborn were used for data manipulation and computation. Jupyter was used as the compiler of choice for this paper.
Microsoft Excel was also used for data processing, visualization and presentation of final results in tables and charts.
Multiple Linear Regression using gradient descent for model optimization was used mainly because of the opportunity to pull an actual equation from the training process and minimize model error.
We define sum of squares error; SSE or Q as follows:

$$Q = \sum_{i=1}^{n}(y_i - (a_0 + a_1x_{i1} + a_2x_{i2} + \ \ldots . + a_kx_{ik}))^2 \quad (2)$$

The way to find the model estimators (or coefficients) is by differentiating, partially, equation 2 with respect to each of the (k+1) independent variables, one at a time. Each of these equations is then equated to zero. Solving this system of simultaneous equations yields the solutions for the model estimators: $\hat{a}_0, \hat{a}_1, \hat{a}_2, \widehat{a}_3, \ldots \ldots . \hat{a}_k$

2.1 Materials

The data employed for the development of this model were obtained from an oil and gas company operating in the Niger Delta area of Nigeria. The data consists of about 99 reservoirs with strong water drive mechanism comprising the following fluid and reservoir properties: initial oil in place, porosity, water saturation, minimum oil column, average net pay, initial reservoir pressure, reservoir temperature, oil viscosity, water viscosity, oil formation volume factor, initial solution gas ratio, specific gravity of oil and the recovery factor.

2.2 Methods

2.2.1 Data cleaning and quality checking

The data cleaning was done using MS Excel. First, the null value point and then a range of 30-70% recovery factor was used because this is common in the area. The correlations of Isehunwa& Nwankwo[12], Onolemhemhen & Isehunwa[16] in the Niger Delta for strong water drive recovery factors. They employed a tolerance range of about +/- 2.5% as an acceptable tolerance which were employed to confirm water drive reservoirs by using the OR operator. Any reservoir that agreed with any of the two correlations was picked as a definite water drive reservoir. OR was used as against AND to prevent data bias towards any of the available correlations. For this study, the 99 strong water drive reservoirs were split randomly using sci-kit learn into:

i. Training Set: a subset used to train a model (60%)
ii. Validation Set: a subset also used to train the model (20%)
iii. Test Set: a subset used to test model (20%)

2.3 Model Optimization Technique

Gradient descent was used for optimizing the values of the coefficients by iteratively minimizing the error of the model on the training data until no further improvement is possible. The batch gradient descent was used in this study.

2.4 Feature Selection

For this work, a combination of expert domain knowledge and iterative feature selection was used. These features as extracted from literature are:

i. Initial oil in place (IIP, MMstb)
ii. Porosity (Ø)
iii. Water saturation (Sw)
iv. Minimum oil column ($H_{min,}$ ft)
v. Average net pay (H, ft)
vi. Initial reservoir pressure (Pi, psia)
vii. Reservoir temperature (T, ^{o}F)
viii. Oil viscosity ($\mu_{o,}$ cp)
ix. Water viscosity (μ_w, cp)
x. Oil formation volume factor (B_{oi}, bbls/stb)
xi. Residual oil saturation (S_{or})
xii. Initial solution gas ratio (Rsi, scf/bbl) and
xiii. Specific gravity of oil

In addition well logand polynomial transforms of some of these above listed variables were used.After this, iterative feature selection is then used as shown in Figure 1. The method of forward selection which is an iterative procedure in which we start with no features in the model. In each iteration, we keep adding features which best improves the model until additional feature additions do not lead to any marked improvement in the performance of the model.

During this work, the first feature added was residual oil saturation and then initial water saturation because of their prominence in the work done on Niger Delta crude oil recovery factor. Other features were then added as per forward selection. The effect and significance of each one were checked using all the model evaluation parameters of the model as they are included in the training process. The most important features were then chosen and were the basis on which further model tuning was done.

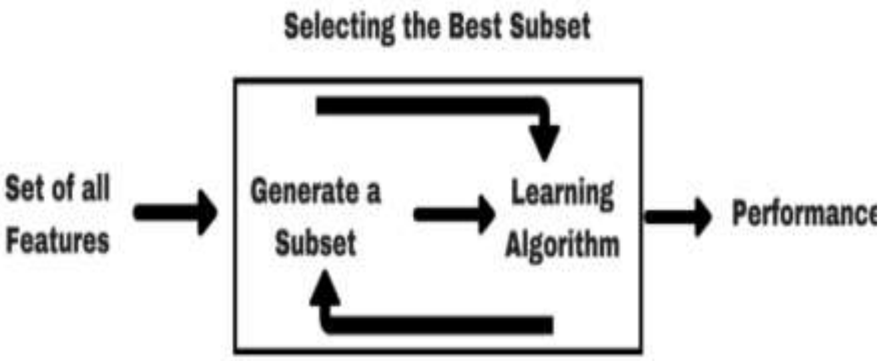

Figure1: Iterative feature selection algorithm

2.5 Model Evaluation Techniques

In developing the model, a combination evaluation techniques were chosen as against just one measure of error or accuracy so as to eliminate any form of model bias, prevent instances of overfitting and hence ensuring model robustness. The evaluation methods used for the model tuning and training are:

2.5.1 Coefficient of determination (R^2): It is simply defined as the measure of the variation of the dependent variable that is accounted for by a regressor. The best R^2- score is 1 and this signifies that the model completely accounts for all the variation in the dependent variable.

$$R^2 = \frac{SSE_{REG}}{SSE_{TOTAL}} \tag{3}$$

In using R^2, care must be taken to ensure that the model is not over-fitted by adding too many model variables or even by using a polynomial transform of too high a degree. To combat this, two measures were put in place.

i. The use of mean absolute error as a measure to give a feel when the model starts to over-fit.
ii. The R^2- score on the test set also gives a measure of how well the model is generalizing which is something that does not coincide with overfitting.

2.5.2 Mean absolute error: It is a measure of the magnitude of errors in a set of predictions. It is the average of the absolute difference between predicted and experimental values of the dependent variable. It is very key to actually know how close to the actual values the model is predicting. Mathematically,

$$MAE = \frac{1}{N}\sum_{j=1}^{n}\left|y_j - \hat{y}_j\right| \tag{4}$$

2.6 Cross-validation

Cross-validation is a statistical method of evaluating generalization performance that is more stable and thorough than traditional technique of using a split into training and a test set. In cross-validation, the data is instead split repeatedly and multiple models are trained. The version used is called k-fold cross-validation, where k is a user-specified number depending on amount of data available. For the purpose of this study, k was chosen as 5. When performing five-fold cross-validation, for example, the data is first partitioned into five parts of (approximately) equal size, called folds. Next, a sequence of models is trained. The first model is trained using the first fold as the test set, and the remaining folds (2-5) are used as the training set. This process is repeated using folds 2,3,4 and 5 as test sets one at a time with the other four used as training sets. The tools used in this reporting and visualization process are;

i. Matplotlib library
ii. Excel (Charts and Tables)

3. RESULTS AND DISCUSSION

3.1 Results

Data cleaning produced Table 1 which shows the various reservoir and fluid properties and the parameter ranges.The method chosen is to use available correlations (Isehunwa & Nwankwo [12], Onolemhemhen & Isehunwa [16] in the Niger Delta for strong water drive recovery factors with a tolerance range of about +/- 2.5% as an acceptable tolerance. Using the OR operator, any reservoir that agreed with any of the two correlations was picked as a definite water drive reservoir. OR was used as against AND to ensure data bias is prevented towards any of these correlations. A total of 99 reservoirs were classified as strong water drive using this method.

By using the same evaluation techniques as stages 1 and 2, the performance of this data set was clearly best and seem satisfactory enough for further tuning and feature selection. The result is as shown below.

Accuracy on training set: $R^2 = 0.8831364$
Accuracy on test set: $R^2 = 0.7019287$
The model metrics are the best we have seen all through and it must be noted that this is without any form of parameter tuning. What these R^2-scores tell us is the model is performing very well on both the training and test data set.

Table 1: Cross section of reservoir and fluid properties

	MIN	**MAX**
IIP MMstb	0.5	598.3
UR MMstb	0.3	373.4
Porosity (%)	15	37
Sw (%)	5	40
Initial P psi	1766	6840
Viscosity cp	0.09	4.26
Bo bbl/stb	1.05	2.9
Rsi scf/bbl	150	3069
RF	0.40	0.75

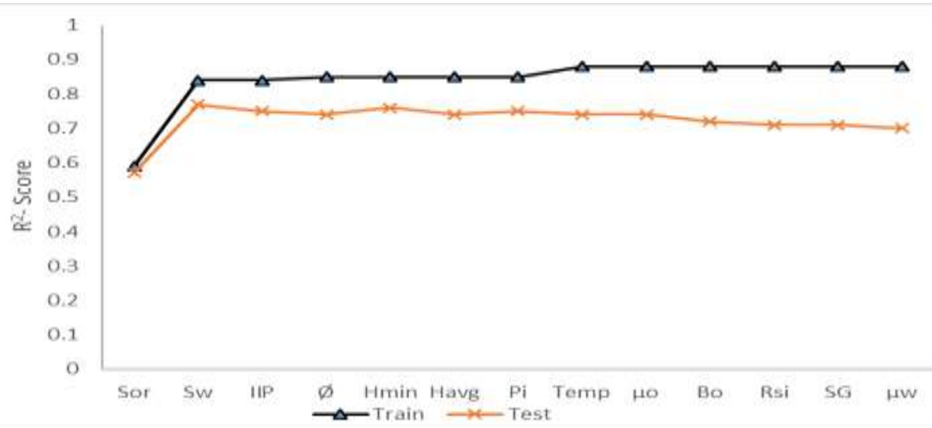

Figure2: R^2 -score obtained by adding each parameter

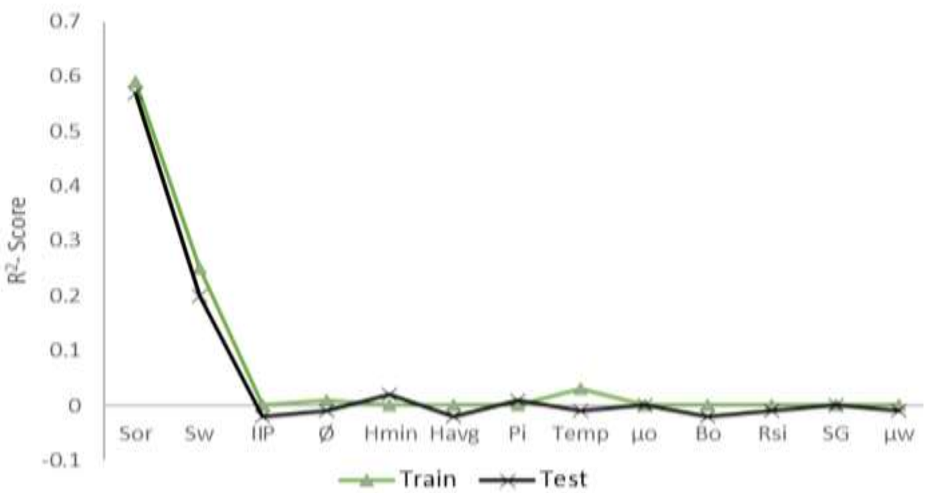

Figure3: Comparison of differential accuracy (R^2 -score) vs parameter

Figure 3 is a summary of the results obtained from the forward selection iterative process. Residual oil saturation S_{or}, was determined the most correlated variable (see Isehunwa & Nwankwo, 1999) using univariate statistical analysis. With S_{or} as the first variable, variables were added one after the other and the model re-trained and new R^2-scores obtained. The result is displayed in Figure 2 for both the training and the test sets. This illustrates the variables and the dependence (or lack) of recovery factor on them noted as follows:

i. That oil residual saturation is the most strongly correlated to recovery factor of strong water drive reservoirs and as such is the most important variable accounting for about 59% of the variations in recovery factor (RF) in the training set and about 57% in the test set.
ii. That the initial water saturation has the next best contribution to the predictor. It accounts for, together with the S_{or}, about 84% of variations in the recovery factor in the training set as much as 77% in the test set as shown in Table 2.
iii. That other parameters like porosity, oil initially in place, minimum oil column, average oil column, initial reservoir pressure, oil viscosity, oil formation volume factor, initial solution gas-oil-ratio, oil gravity and water viscosity and their various transforms all account for only about 4% additional variation of RF in the training set with the test accuracy reducing by about 7%.

Figure 3 is a chart built on Figure2. It is a representation of the contribution of each parameter to the predictor's accuracy for both the training and test sets. Just like in Figure 3, the parameters follow the same trend in terms of their importance to the model accuracy. From Figures 2 and 3, it is clear that only S_{or} and S_{wi} had a significant impact on model accuracy. Other parameters mostly either had no effect at all or having too little effect on training accuracy to justify their effect on test accuracy and hence model generalisation. Summary of the results are in the Table 2 below.

Table 2: The effect of each variable on model accuracy

	R^2- score	
Variables	**Train**	**Test**
S_{or}	0.59	0.57
S_{or}, S_w	**0.84**	**0.77**
S_{or}, S_w, IIP	0.84	0.75
S_{or}, S_w, IIP, Ø	0.85	0.74
S_{or}, S_w, IIP, Ø, H_{min}	0.85	0.76
S_{or}, S_w, IIP, Ø, H_{min}, H_{avg}	0.85	0.74
S_{or}, S_w, IIP, Ø, H_{min}, H_{avg}, P_i	0.85	0.75
S_{or}, S_w, IIP, Ø, H_{min}, H_{avg}, P_i, Temp	0.88	0.74
S_{or}, S_w, IIP, Ø, H_{min}, H_{avg}, P_i, Temp, μ_o	0.88	0.74
S_{or}, S_w, IIP, Ø, H_{min}, H_{avg}, P_i, Temp, μ_o, B_o	0.88	0.72
S_{or}, S_w, IIP, Ø, H_{min}, H_{avg}, P_i, Temp, μ_o, B_o, Rsi	0.88	0.71
S_{or}, S_w, IIP, Ø, H_{min}, H_{avg}, P_i, Temp, μ_o, B_o, Rsi, SG	0.88	0.71
S_{or}, S_w, IIP, Ø, H_{min}, H_{avg}, P_i, Temp, μ_o, B_o, Rsi, μ_w	0.88	0.7

The best combination of variables for this model work is therefore the one in **bold**: S_{or}, S_w because it shows the highest R^2 score for test data. The reason for this choice is supported by Ockham's razor. The principle of Ockham's razor quite simply says that the best model is the simplest model that best describes the problem. Here, the choice of variable combination is S_{or} and S_w as against S_{or}, S_w, IIP, Ø, H_{min} , for example, even though their accuracies are comparable is because the former is simpler and as a result requires less features than the latter would which might have a role to play in better model generalization. Using the obtained parameter combination as a basis, further model tuning was carried out to optimize the model even further before carrying out prediction test and validation sets.

3.2 Test Results

To truly test how good the model is, the results from the training sets are simply insufficient. For an independent, true and fair assessment of the model accuracy, it must not only be tested on the training set but also on data that are not part of the training process. For the purposes of this study, the model was tested on a validation set (which was set aside from the model data set), a test set which is the same one as was used in validating the two available correlations for strong water driven Niger Delta reservoirs. Figure 4shows the relationship between the actual and predicted RF for test data.

3.3 Validation results

The validation set consists of 18 reservoirs set aside from the initial model data set. The excellent R^2- score of the model here (0.80) gives an indication that the model generalizes well on new data. The chart summarising this result is as shown in Figure 5. Mean Absolute Error is only 0.017.

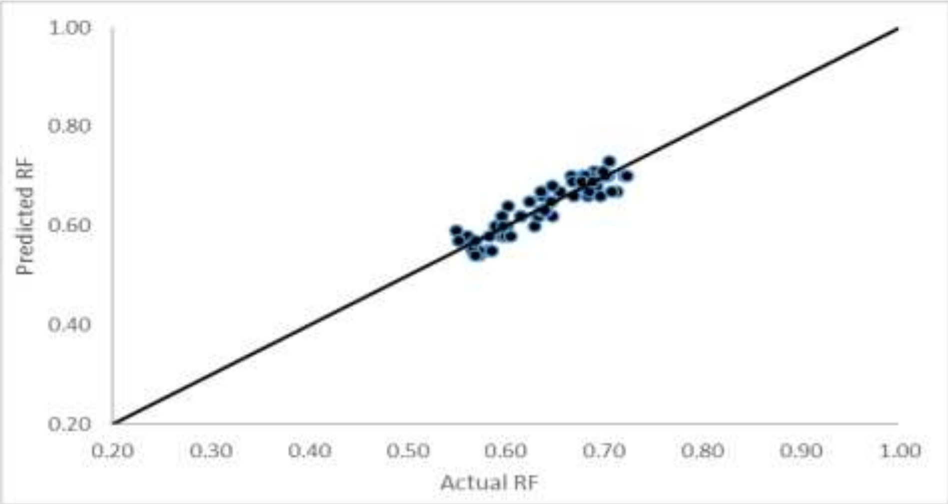

Figure 4: Model training result

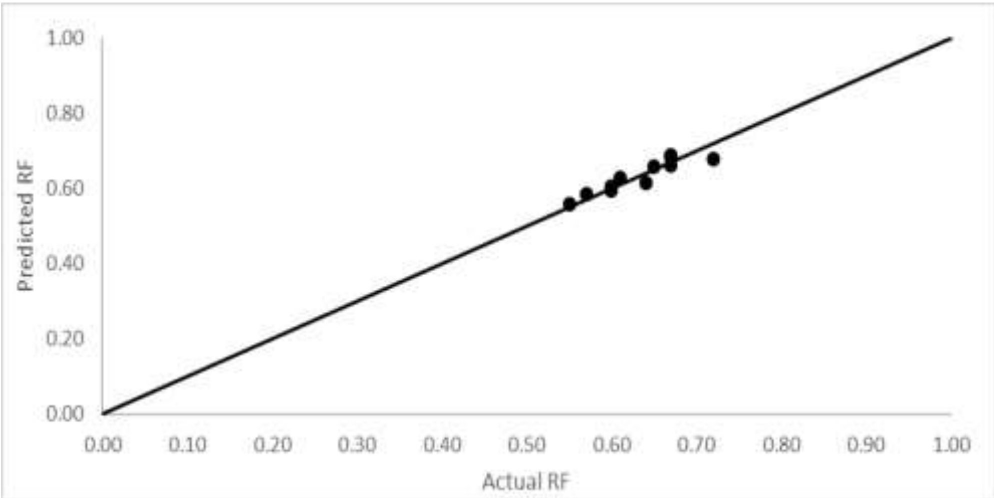

Figure 5: Model results in 18 reservoirs (Validation set)

3.4 Prediction Results

The test set is used to test model robustness and generalization. The test set are 16 reservoir data points that were held back from the training set. The results of the model performance are expressed in a table and charts that will be seen subsequently. Also, the results of the models developed by Isehunwa & Nwankwo [12] and Onolemhemhen & Isehunwa [16] will be used as a basis for comparison. The procedures in the previous paragraph are repeated for the test sets used by both studies that already exist to establish model robustness, generalisation and well-rounded application.

3.5 Final model

The model developed, presented as equation 5, had an excellent performance on the training set as the coefficient was about 0.84. The mean absolute error is about 0.018. Putting the accuracy of the model into perspective, the model trains on the training set RF to within 0.01 absolute error 27% of the time, within 0.02 absolute error 52% of the time and to an error of about 0.03 82% of the time. The maximum training error is 0.042. The visualization of the model training results is as shown in Figure 6.

It can be seen from Table 3 that the best model here is the newly developed one (this study) because MAE is the least.The results of the newly developed correlation are compared with all the other two available for the Niger Delta. This study's model clearly outperforms the best available correlation (Onolemhemhen & Isehunwa [16]) as it has higher RF (up to 84%) in this set. Also, the developed model has a better MAE, albeit marginally, compared to the best available.

Table 4 shows the results of each of the three models and their performance on the literature data used for the validation of the two existing models for RF in strong water crude oil reservoirs in the Niger Delta.

Table 3: Test set (1) results model result on the test set

S/N	S_w	Viscosity (cp)	S_{or}	RF O&I	RF I&N	This Study	RF
1	0.39	0.27	0.14	0.70	0.77	0.72	0.70
2	0.25	0.86	0.27	0.64	0.64	0.64	0.63
3	0.17	0.30	0.28	0.68	0.67	0.68	0.70
4	0.35	1.98	0.25	0.60	0.62	0.63	0.62
5	0.30	1.00	0.25	0.64	0.65	0.64	0.62
6	0.28	1.10	0.26	0.63	0.63	0.63	0.62
7	0.25	0.90	0.27	0.64	0.63	0.64	0.64
8	0.15	1.47	0.36	0.61	0.58	0.60	0.62
9	0.26	0.34	0.23	0.68	0.69	0.68	0.69
10	0.14	0.34	0.30	0.67	0.65	0.67	0.64
11	0.20	0.30	0.26	0.68	0.68	0.69	0.70
12	0.11	0.24	0.30	0.67	0.66	0.67	0.65
13	0.25	0.20	0.21	0.69	0.72	0.71	0.67
14	0.34	0.22	0.16	0.70	0.76	0.72	0.70
15	0.15	0.12	0.25	0.69	0.71	0.71	0.70
16	0.25	0.21	0.21	0.69	0.72	0.71	0.73
R^2-	0.770	0.540	0.836				
MAE	0.017	0.055	0.016				

Table 4: Further test results model result on the I & N test set

S/N	S_w	Viscosity (cp)	S_{or}	RF O&I	RF I&N	This Study	RF
1	0.12	0.20	0.29	0.68	0.57	0.69	0.67
2	0.22	0.50	0.27	0.66	0.56	0.66	0.65
3	0.19	0.25	0.26	0.68	0.58	0.69	0.67
4	0.15	1.00	0.34	0.64	0.51	0.61	0.64
5	0.17	0.91	0.32	0.64	0.52	0.63	0.61
6	0.25	2.40	0.32	0.58	0.49	0.58	0.57
7	0.30	2.20	0.28	0.59	0.50	0.60	0.60
8	0.42	0.70	0.16	0.66	0.61	0.68	0.72
9	0.47	4.20	0.21	0.58	0.52	0.59	0.60
10	0.23	4.10	0.35	0.58	0.46	0.56	0.55
11	0.30	0.61	0.23	0.66	0.57	0.66	0.67
12	0.15	0.24	0.28	0.68	0.57	0.68	0.67
R^2-score	0.760	0.927	0.857				
MAE	0.018	0.097	0.015				

3.6 Results Summary

From all the results obtained, it is statistically obvious that the model obtained during the course of this study is better than all the other correlations that were previously in use as seen in table 4. Not only does it account for a larger percentage of the variation in RFs, it also generalises better than other as shown by its performance on data sets of different sources. Not only that, it uses the exact same number of features (three) as the existing model hence will not be requiring any additional data.

The model type of choice for this study being Linear Regression has one very key merit over most machine learning algorithms which is that we can extract the key parameters of the model obtained and write an equation from that and thereby making it possible to use the outcomes of the work without necessarily having the code.

The equation obtained empirically for the estimation of the recovery factor for strong water driven crude oil reservoirs in the Niger Delta is as expressed below in terms of residual oil saturation and initial water saturation:

$$RF = 1.09176735 - 1.168614085S_{or} - 0.5366645S_{wi} \tag{5}$$

where all terms have their usual meanings with RF, S_{or} and S_{wi} in fraction.

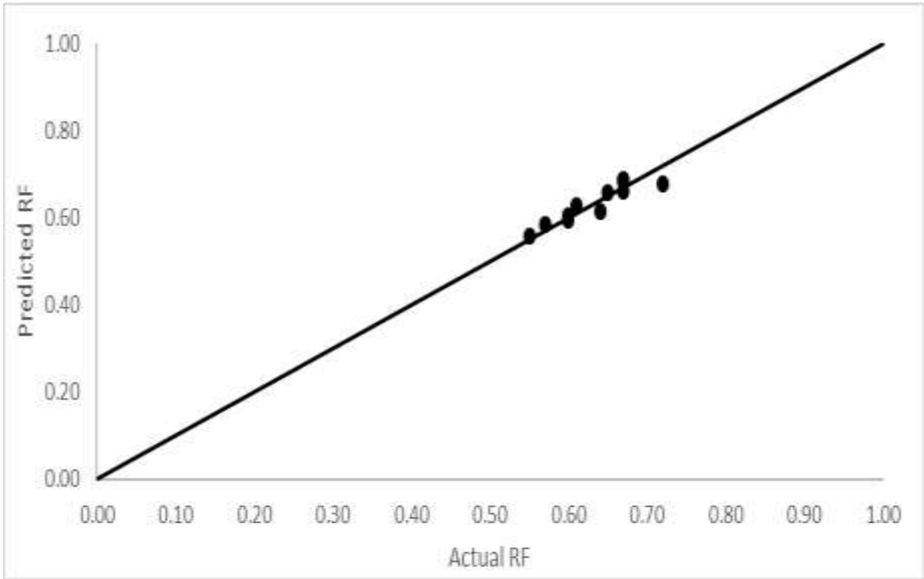

Figure6: Model results of 16 tested reservoirs

4. CONCLUSION

The approach developed in this study is the best available model for evaluating in recovery factor in strong water drive reservoirs in the Niger Delta. It has the best R^2-score and MAE in the reservoirs studied. The model developed had an excellent performance on the training set as the coefficient was about 0.84. The mean absolute error is about 0.018. The results obtained showed better model performance and generalization than any previously existing model.

The Python code used in model development can be deployed in various eco-systems including but not limited to the academia and the industry once the right adjustments to the code are correctly made.

The inputs going into this developed model are just two: the initial water saturation and the residual oil saturation. Both of these are easily obtainable by core analysis and in turn making the deployment of the model very easy.

REFERENCES

[1] API Bulletin D14 (1967): A Statistical Study of Recovery Efficiency, American Petroleum Institute, ProductionDepartment, Dallas, Texas.

[2] API Bulletin D14 (1984): A Statistical Study of Recovery Efficiency and Recovery Efficiency, 2nd Edition, American Petroleum Institute, Production Department, Dallas, Texas.

[3] Andreas C.M. and Sarah G. (2017): *Introduction to Machine Learning with Python*, O'Reilly Media Inc., California.

[4] Arps J.J. & Roberts T.G (1955).: The Effect of the Relative Permeability, Oil Gravity and the Solution Gas-OilRatio On the Primary Recovery from A Depletion Type Reservoir, *Trans.,AIME*, 204 (01): 120–127. https://doi.org/10.2118/469-G.

[5] Barlow, W.H. & Berwald, W.B. (1945): Optimum Oil-Well Spacing, Drilling and Production Practices, API, 129.

[6] Craze, R.C. & Buckley, S.E. (1945): "A Factual Analysis of the Effect of Well Spacing on Oil Recovery," Drilling and Production Practices, API, 144.

[7] Dake, L.P. (1978): Fundamentals of Reservoir Engineering, Elsevier Scientific Publishing Co. Amsterdam.

[8] Ferruh, D. (2007): Reserves Estimation: The Challenge for the Industry, Journal of Petroleum Technology, *Distinguished Author Series*, SPE, 103434.

[9] Gareth J., Daniela W., Trevor H. & Robert T. (2013): An Introduction to Statistical Learning, Science Business Media, Springer New York.

[10] Gulstad, R.L. (1995): The Determination of Hydrocarbon Reservoir Recovery Factors by using Modern Multiple Linear Regression Techniques, M.Sc. Project, Texas Tech. University.
[11] Guthrie, R.K., & Greenberger, M.H. (1955): The Use of Multiple Correlation Analysis for Interpreting Petroleum Engineering Data, Drilling and Production Practices, API, 130.
[12] Isehunwa, S.O. & Nwankwo, S.U. (1999): A Correlation of Oil Recovery Factor forWater Drive Reservoirs in The Niger Delta," NAPE Bulletin, Annual International Conference and Exhibition, Paper 20.
[13] Khusnutdinova, R. Yalovega, D.,,Yusupova, L., Gimaztdinova, E., & Gutorov, A.(2025), Methods for determining the oil recovery factor in an oil field E3S Web of Conferences **623**, 02018: https://doi.org/10.1051/e3sconf/202562302018
[14] Muskat, M. & Taylor, M.O. (1946): Effect of Reservoir Fluid and Rock Characteristics or Production Histories of Gas Drive Reservoirs, Trans., AIME, 165 (01): 78–93. https://doi.org/10.2118/946078-G
[15] Nmegbu C. G. J, Ebube, O, F. &Edet, E. A. (2021), Comparative Study of Oil Recovery Factor Determination for Edge and Bottom Water Drive Mechanism Using Water Influx Models, European Journal of Technology and Engineering Research, 6, 5
[16] Onolemhemhen, R.U., Isehunwa, S.O. and Salufu, S.O. (2016): Development of Recovery Factor Model for Water Drive and Depletion Drive Reservoirs in the Niger Delta, 184283 at the SPE Nigeria Annual International Conference and Exhibition.
[17] Petroleum Resources Management System. (2007). Richardson, Texas, Society of Petroleum Engineers. https://www.spe.org/industry/docs/Petroleum_Resources_Management_System_2007 .pdf (Accessed on July 29, 2018).
[18] Talluru, G. and Xingru, W. (2017): Using Data Analytics on Dimensionless Number to Predict the Ultimate Recovery Factors for Different Drive Mechanisms of Gulf of Mexico Oil Fields, 187269 at the SPE AnnualTechnical Conference and Exhibition, San Antonio, Texas.