FUDMA Journal of Sciences (FJS)
ISSN online: 2616-1370
ISSN print: 2645 - 2944
Vol. 8 No. 3, June, 2024, pp 235 - 242
DOI: https://doi.org/10.33003/fjs-2024-0803-2400

MACHINE LEARNING PREDICTION OF VOLUME FRACTION OF GAS-HYDRATES IN NATURAL GAS PIPELINES IN OFFSHORE NIGER DELTA

***Abubakar Tanko, Mukhtar Abdulkadir, Afolabi Eyitayo Amos and Lamido Inuwa**

Department of Chemical Engineering, Federal University of Technology Minna, Nigeria

*Corresponding authors' email: abubakartanko87@gmail.com Phone: +2348034580686

ABSTRACT

This study employs multiphase simulations with OLGA software to investigate volume fractions of hydrate in an offshore gas system and develops machine-learning models to predict these fractions. Annually, substantial operating expenditures are allocated to hydrate prevention, with significant costs associated with inhibition (Wang *et al.,* 2022). Hydrate formation along natural gas pipelines is recognized as a critical threat to the success of gas field operations. Despite the importance, no machine learning model has been validated for predicting volume fractions of hydrate in the Niger Delta study area, making this development crucial. Key findings indicate significant hydrate jamming risks in Niger Delta offshore flowlines and risers, with a peak volume fraction of 0.54, highlighting the need for proactive management strategies. Hydrate formation begins at 750 m where fluid temperatures fall below formation thresholds, with a sudden increase in volume at 2971 m, peaking at 3022 m before declining. Machine Learning model comparisons show Random Forest's superior accuracy (correlation coefficient of 0.9391, mean absolute error of 0.0271), while Linear Regression provides interpretable insights for future predictions. All models perform well, with Random Forest leading in accuracy. Regression analysis reveals relationships between volume fractions of hydrate and various parameters, guiding management strategies. The Random Forest and Linear Regression models are valuable for estimating hydrate volumes and enhancing management approaches in natural gas pipelines due to their accuracy and interpretability. These findings underscore the importance of proactive hydrate management in offshore gas systems and the potential of Machine Learning models to optimize these strategies.

Keywords: Hydrates, Machine Learning, Models, Prediction, Niger Delta, Volume fraction

INTRODUCTION

Natural gas hydrates are solid, non-stoichiometric crystalline compounds that occur when water molecules attach themselves by hydrogen bonding and create cavities filled by either a gas or volatile liquid molecule (Wang *et al.,* 2022). Formation of hydrate requires like light ends of petroleum such as methane, ethane, propane, butane and adequate water in low temperature and high-pressure conditions (Seo *et al.,* 2021). The formation and accumulation of gas hydrate particles foster flow resistance in the pipeline, thus reducing the energy obtainable for conveying hydrocarbon fluids, requiring a higher operating pressure, key equipment destruction, and shutdown of production (Qin, 2020).

In the oil and gas sector, gas hydrate is a major problem; it can form in the reservoir during enhanced oil recovery with carbon dioxide (Gbaruko, 2004). During wellbore drilling, it is also possible to create hydrates due to the kicking of gas (Barker, 1989). Hydrates may develop in pipelines during the extraction of oil (Dorstewitz, 1995) and due to Joule Thomson's cooling, hydrates can form when gas is expanded. Nonetheless, they can be handled well by means of hydrate inhibitors (Odutola *et al.,* 2015). Obstruction of flow by blockage of the pipeline caused by formation of hydrate may lead to monetary losses and certain environmental and safety problems (Saeed, 2021). Currently Nigeria has about 28.5 billion barrels of crude oil and a gas reserve of more than 166 trillion natural gas, according to the Nigerian National Petroleum Investment Management Services (NAPIMS, 2023). Due to the frequent vandalism of onshore installations and large reservoirs which are normally located within the maritime environment, operators carry out a lot of oil exploration and production activities in Nigeria. While there is available technology to encourage offshore production in Nigeria, hydrate formation is however a significant issue in offshore production in the Niger delta.

Hydrate formation caused the shutdown of production, which resulted in a massive revenue loss due to production interruption and dissociation of hydrate interventions (Odutola, 2022). Consequently, A precise prediction of the equilibrium conditions for gas hydrates needs to be made accordingly.

To date, temperatures and pressures of gas hydrate equilibrium have been acquired from experiments, thermodynamics models (Moradi, 2013), empirical correlations (Kummamuru *et al.,* 2021), and artificial intelligence techniques (Zhong *et al.,* 2019). In certain cases, the costs and time needed for experimental measurements of gas hydrate equilibrium conditions are high while the thermodynamic models do not do not provide valid forecasts at low temperatures or higher pressures (Garapati, 2014). A new model with a high degree of efficiency and accuracy should therefore be proposed. Artificial intelligence is the collection of algorithms and techniques that are used to create computer systems which can learn from data in order to make forecasts or predictions (Swamynathan, 2017).The field of Artificial Intelligence comprises subareas symbolizing different methods for development of systems that can understand and learn from data. This information can thereafter be utilized by the system to solve jobs and reach certain goals (Kaplan & Haenlein, 2019). Machine Learning (ML), which is an essential part and subfield of AI, involves methods that help the software to learn with no programming so as to have a clear understanding of patterns and more complicated information coming out of data. Nazifi *et al.,* (2024) stated that ML models are trained on datasets to identify patterns, allowing them to execute tasks for instance classification, clustering or regression,.

In recent years, several researchers have researched the use of machine learning in the development of prediction models due to the fact that these techniques are very promising in

prediction accuracy (Zarei *et al.*, 2021). One of the fastest growing areas of computer science is machine learning as a branch of artificial intelligence, which aims to imitate human intelligence by learning from the environment (Yu & Tian, 2022).

In summary, the study makes significant contributions to knowledge in the field of hydrate-blocking risk management, spanning from the identification of risks to the development of predictive models. It offers practical insights and tools for industry practitioners to improve operational safety and efficiency.

MATERIALS AND METHODS

Description and Simulation of System

This study involved the modelling of an offshore natural gas production system consisting of a wellbore, wellhead, flowline, riser and topside, for an existing gas platform located in the offshore Niger Delta field by utilizing Schlumberger's Multiphase dynamic flow software, OLGA (Schlumberger, 2022). The system conveys multiphase fluids in offshore natural gas production. The offshore system comprises a well tubing pipeline with a 2800 m true vertical depth, a 160 m long wellhead pipe, a 3200 m flowline extending to a 400 m vertical riser and a 120 m long horizontal topside pipe. The entire production is managed using the choke. The well has an inlet and outlet ambient heat of 52 °C and 4.5 °C respectively. A temperature of 4.5 °C exists in both the riser and flow line. A pressure of 210 bara exists in the reservoir while the temperature is 52 °C. The temperature of seawater is 4.5 °C. The heat transmission coefficient on external and internal walls is adjusted to 500 W/m^2K and 10 W/m^2K respectively. The pipeline internal diameter is 0.3048 m, while the roughness 0.000045 m. The pipeline is subdivided into 22 sections. The schematic of the model is shown in Figure 1.

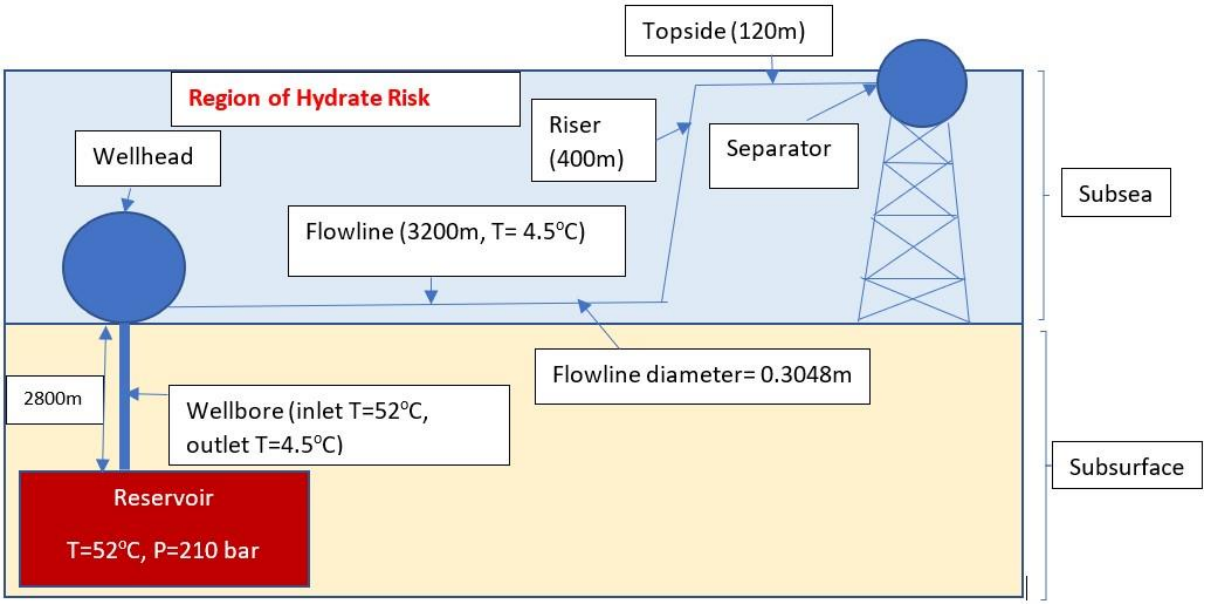

Figure 1:Schematic of the modelled offshore gas system

Simulation

The curve for hydrate formation (Figure 3) was obtained using MultiFlash according to the composition of the fluid (Infochem, 2015). The calculated equilibrium hydrate data was exported into a table with Pressure-Temperature data. OLGA brought in the table and computed the fluid properties at specific pressures and temperatures as needed for the simulations through interpolation in the PT tables. Ordinarily, continuous inhibitor injection is employed to prevent hydrate formation. Consequently, this research presumed no inhibitors were injected, and formation of hydrate could occur. Therefore, the simulation was conducted and the results are shown in Figure 4. The fluid constitution is predominantly gas, as shown in Table 1, as such it is a gas well. The temperature of the fluid can be lower than the flow line hydrate temperature. To forestall hydrate blocking risks, sections susceptible to high hydrate blockage risks should be identified. The methodology utilized is shown by Figure 2.

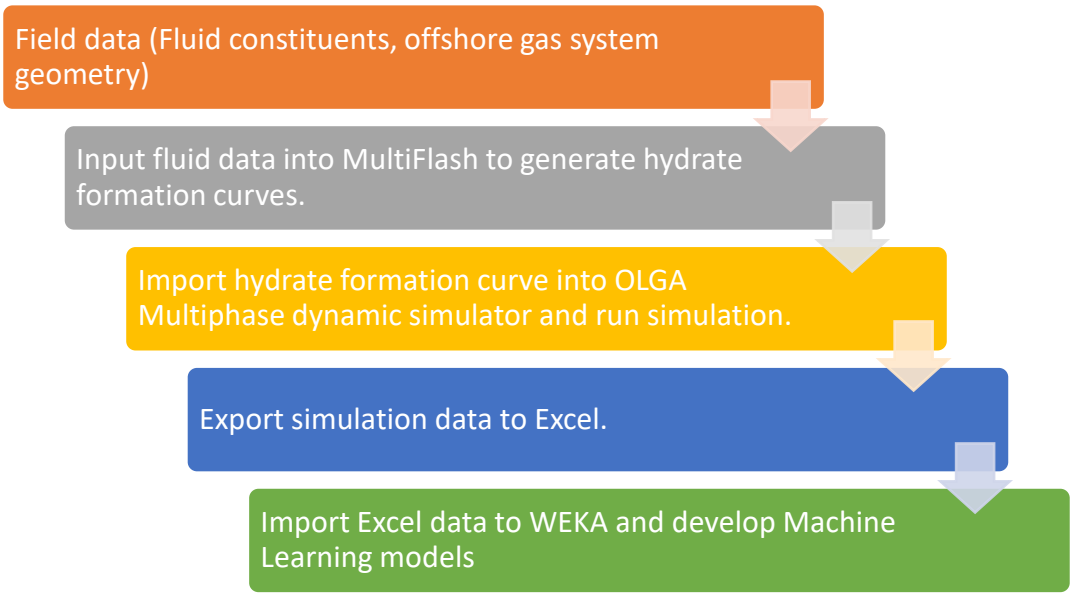

Figure 2: Workflow employed in the study

Table 1: Constituents of fluid

Constituents	Amount (moles)
Nitrogen	0.72
Carbon Dioxide	1.31
Methane	85.91
Ethane	6.74
Propane	3.12
N-Butane	0.90
I-Butane	0.71
N-Pentane	0.59
Water	10

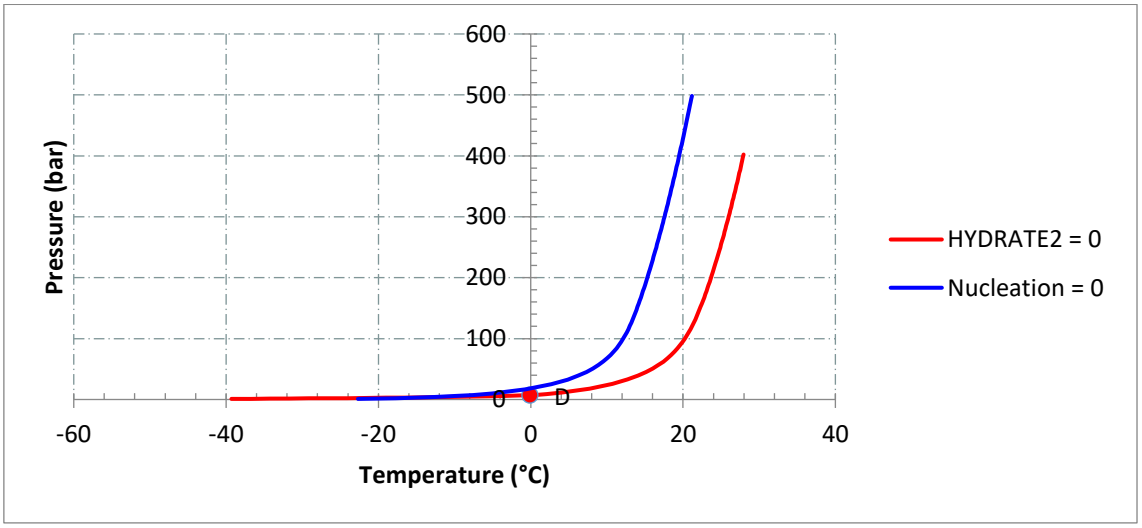

Figure 3: Equilibrium hydrate curves

Following the flow of fluid in the flowline, it dipped down to a subcooling temperature, below the equilibrium temperature of hydrates. Following subcooling, the system temperature increased swiftly to the temperature of hydrate development, and additional hydrate development was restrained by the removal of heat from the flowline. Given that the formation of hydrate is exothermic, as the heat is released, it raises the fluid temperature. Hydrate started appearing at about 750 m, where the temperature of fluid fell below the temperature of hydrate formation. At a length of 2971 m from the wellhead, the volume of hydrate in the pipeline increased suddenly, and peaked at 3022 m, with a maximum fraction of 0.54. This means that 54% of the flowline is covered by gas hydrates. After 3022 m, the rate of hydrate development began to decline. The temperature, meanwhile, stayed unchanged, however the pressure kept on dwindling. With the force driving hydrate development diminishing, the volume of hydrate eventually fell to 0 at about 3600 m. The full simulation cycle was 5 hours as shown in Figure 4.

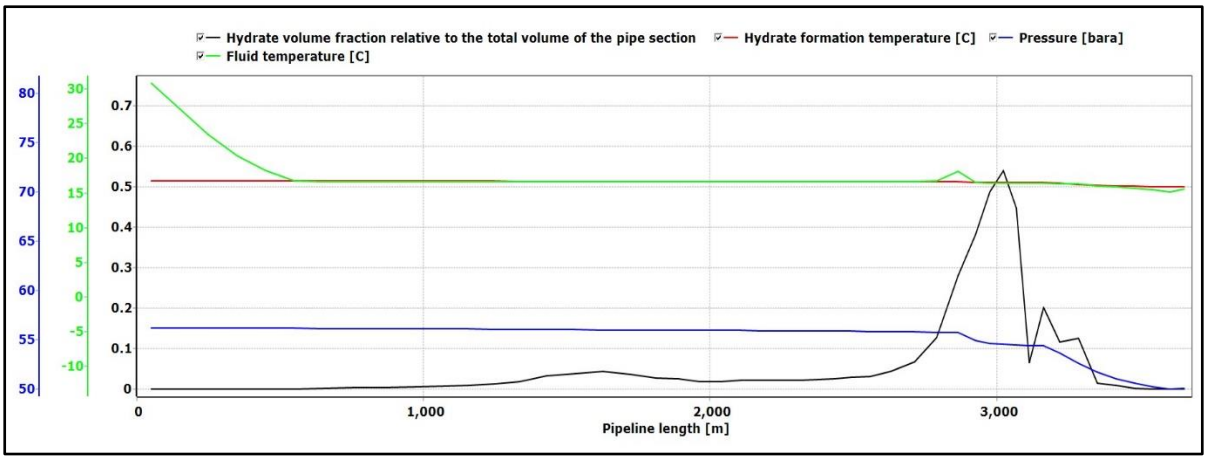

Figure 4: Simulation of the offshore gas system after 5 hours

Predicting of Volume fraction of gas hydrate using Machine Learning Models

Four models were used to predict the Volume fraction of hydrate, observed in the simulation data. Variables like water cut, mixture velocity, hydrate formation temperature, hydrate formation rate, fluid temperature, hydrate formation pressure, pipeline length, pressure, and pressure drop were employed as input variables to the models. The models' outputs were the Volume fraction of hydrate in the pipe. Figures 5 and 6 depict the parameters employed in developing the linear regression model and the other three models respectively. The cross-validation results are shown in Figure 7, where the Volume Fraction of hydrate forecasted by the models were charted against the actual volume fraction of hydrate obtained in the simulation.

RESULTS AND DISCUSSION

Linear Regression Model

$$Hydrate\ Volume\ Fraction = 0.0001 * Pipeline\ length + 0.0442 * Hydrate\ formation\ rate - 0.5472 * Mixture\ velocity + 0.011 * Pressure + 0.0049 * Fluid\ temperature + 0.1375 \quad (1)$$

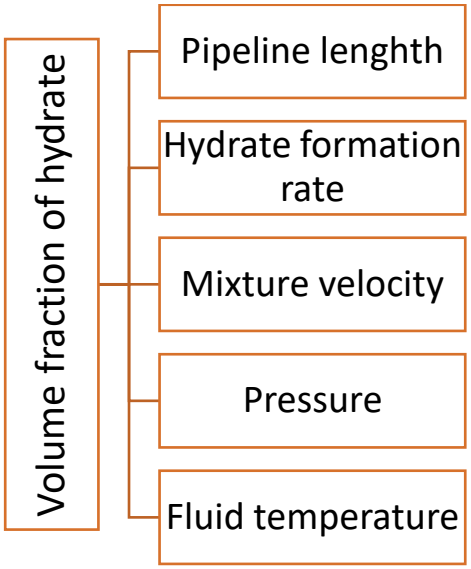

Figure 5: Parameters employed by Linear regression model to predict volume fraction of hydrate

Equation 1, is a linear regression model (models that simulates a linear relationship in one response variable and an explanatory variable (Hackeling, 2014) that provides an interpretation of the nature of the dependence of the Volume fraction of hydrate on the independent variables. A positive correlation between the Volume fraction of hydrate and pipeline length, hydrate formation rate, pressure, and fluid temperature is obtained, whereas a negative correlation is obtained between the Volume fraction of hydrate and Mixture velocity. The signs of the regression equation help in the interpretation of the type of relationship between the dependent and independent variables. Equation 1 aids in understanding of how increasing pipeline length, hydrate formation rate, pressure, and fluid temperature can increase the Volume fraction of hydrate, but decreasing mixture velocity can decrease the Volume fraction of hydrate. Furthermore, the prediction accuracy of the Linear regression model was very good, yielding a Correlation coefficient of 0.9307. The regression results are presented in **Error! Reference source not found.**.

Table 2: Results from the linear regression model to forecast volume fraction of hydrate

Predictor	Coefficients	Standard Errors	t	p-value (significance)
Intercept	0.13745	0.32581	0.422	0.675
Pipeline length [m]	7.15E-05	1.01E-05	7.08	< .001
Hydrate formation rate per unit volume [kg/m^3-s]	0.04415	0.01782	2.477	0.017
Mixture velocity (m/s)	-0.54724	0.03187	-17.171	< .001
Pressure (bara)	0.01097	0.00501	2.191	0.034
Fluid temperature [^{o}C]	0.0049	0.00275	1.781	0.082

The outputs of regression in Table 2 indicate that mixture velocity, pipeline length, hydrate formation rate, pressure and fluid temperature are significant statistically owing to the elevated t-statistic and lower P-values as employed by Srivastava (2018). Figure 7 presents a chart for the predicted and actual Volume fraction of hydrate values. Figure 7 implies that a linear model is capable of accurately forecasting the actual volume fraction of hydrate obtained in the simulation.

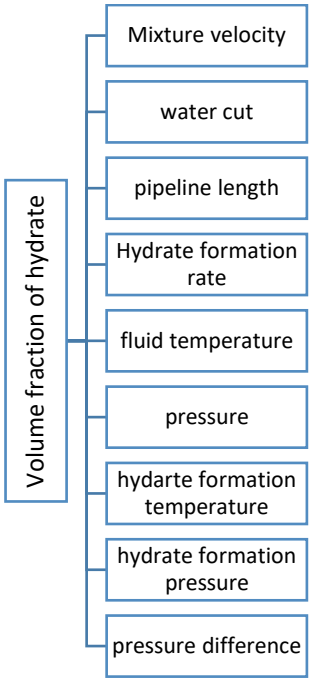

Figure 6: Parameters employed by the other three models to predict the volume fraction of hydrate

Table 3 presents the results of predictions of the four (4) models. A bar chart is plotted to show the accuracy of the models concerning the Correlation Coefficient and Mean Absolute Error (Figure 8). The correlation coefficient values are plotted in blue while the mean absolute error values are in orange.

Based on the results obtained from our analysis comparing different regression models, it is evident that each model exhibits varying degrees of predictive performance.

The Decision Tree model demonstrates a moderate correlation coefficient of 0.6244 and a mean absolute error (MAE) of 0.0583. While it shows some level of predictive capability, its performance falls short compared to other models evaluated.

In contrast, Linear Regression, Multi-Layer Perceptron (MLP), and Random Forest models showcase significantly higher correlation coefficients, indicating stronger linear relationships between the features and target variable. Among these models, Random Forest exhibits the highest correlation coefficient of 0.9391, closely followed by the MLP and Linear Regression with correlation coefficients of 0.931 and 0.9307, respectively. This suggests that these models are better able to capture the underlying patterns in the data. Additionally, when considering the mean absolute error (MAE), which measures the average magnitude of errors between predicted and actual values, it is observed that both Random Forest and Linear Regression models outperform the others, with lower MAE values of 0.0271 and 0.0237, respectively. The MLP follows closely with an MAE of 0.0314. The models were ranked in the following order of prediction accuracies: Random Forest > Linear Regression > Multilayer Perceptron > Decision Tree. These results agree with the findings of the preceding investigation by Yu & Tian, (2022), which emphasized the superiority of Random Forest in hydrate forecasting.

Table 3: Volume fraction of hydrate Predictions using four models

Actual	Linear regression	Multi Layer Perceptron	Random Forest	Decision Tree
0	-0.016	-0.04	0.004	0.078
0.025	0.035	0.019	0.027	0.078
0	-0.013	-0.035	0.004	0.078
0.023	0.035	0.036	0.025	0.078
0.019	0.036	0.01	0.045	0.078
0.039	0.044	0.047	0.036	0.036
0.032	0.036	0.051	0.036	0.036
0.067	0.036	0.063	0.073	0.036
0.005	-0.002	0.002	0.003	0.036
0	-0.016	-0.008	0.003	0.036
0.019	0.037	0.01	0.042	0.066
0.014	0.021	0.005	0.015	0.066
0	-0.006	0.029	0.031	0.066
0.009	0.015	0.004	0.008	0.066
0.54	0.436	0.474	0.363	0.066
0.203	0.086	0.06	0.049	0.025
0.021	0.036	0.026	0.024	0.025
0.448	0.508	0.517	0.296	0.436
0.066	0.285	0.233	0.147	0.436
0.382	0.458	0.463	0.433	0.436
0.125	0.07	0.092	0.123	0.024
0.025	0.024	0.065	0.044	0.024
0.127	0.063	0.082	0.118	0.024
0	-0.013	0.016	0.002	0.024
0.029	0.025	0.067	0.041	0.024
0	-0.049	-0.002	0.003	0.03
0.036	0.046	0.008	0.026	0.03
0.043	0.049	0.006	0.028	0.03
0.022	0.034	0.026	0.025	0.03
0.002	0.001	-0.004	0.005	0.03
0.018	0.027	0.015	0.017	0.003
0.044	0.035	0.076	0.044	0.028
0.487	0.432	0.449	0.397	0.343
0.024	0.035	0.054	0.024	0.028
0.002	0.006	0.006	0.006	0.016
0	0.004	0.006	0.002	0.027
0.022	0.039	0.03	0.023	0.027
0.003	-0.007	0.001	0.002	0.027
0.007	0.01	0.002	0.007	0.027
0.028	0.037	0.028	0.03	0.027
0.002	0.007	-0.052	0.014	0.016
0.033	0.035	-0.013	0.029	0.039
0.116	0.141	0.276	0.114	0.152
0.014	0.053	0.04	0.059	0.152
0.28	0.175	0.203	0.129	0.19
0.001	0	-0.01	0.006	0.001
0.008	0.022	0.018	0.013	0.19
0.006	0.004	0.025	0.004	0.003

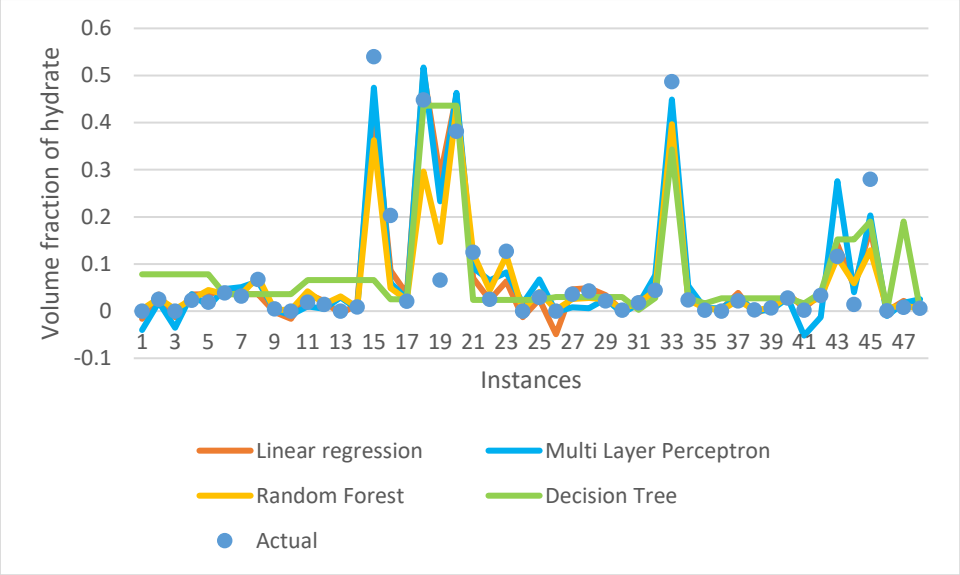

Figure 7: Predicted volume fraction of hydrate using the four models

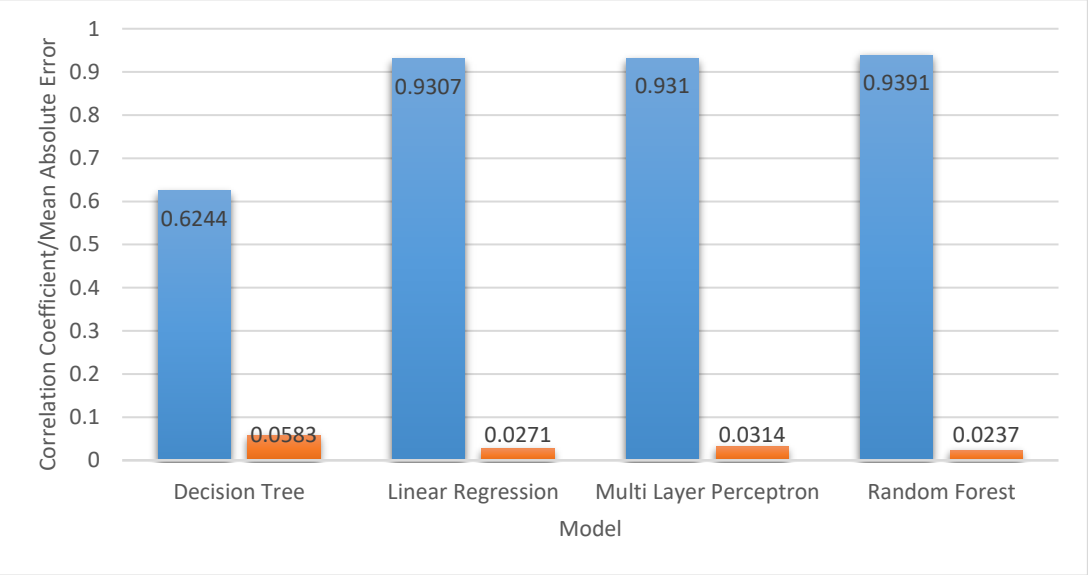

Figure 8: Accuracy of models' predictions

In a nutshell, while the Decision Tree model shows some predictive capability, its performance is overshadowed by the other models evaluated. Linear Regression, Multi-Layer Perceptron, and Random Forest models exhibit stronger predictive performance, as evidenced by their higher correlation coefficients and lower mean absolute errors. Among these models, the Random Forest model stands out as it achieves the highest correlation coefficient and the lowest mean absolute error, suggesting its superiority in capturing the underlying relationships within the data and making accurate predictions. Therefore, for this particular regression task of forecasting the volume fraction of hydrates, the Random Forest model may be the most suitable choice.

CONCLUSION

This study examined the volume fraction of hydrates in an offshore gas system using multiphase simulations and machine learning models. Simulations indicated significant hydrate formation risk in the Niger Delta offshore gas flowlines, with a peak volume fraction of 0.54, necessitating management measures. Hydrate formation began at 750 m, peaked at 0.54 around 3022 m, and dropped to zero by 3600 m, with pressures decreasing and temperatures stable. The Random Forest model achieved the highest accuracy (correlation coefficient of 0.9391, mean absolute error of 0.0271), while Linear Regression offered better interpretability. All ML models performed well, with Random Forest ranked highest. Regression analysis showed longer pipeline lengths, higher hydrate formation rates, pressures, and fluid temperatures increase hydrate volume fraction, while higher mixture velocities decrease it. Random Forest and Linear Regression are recommended for their accuracy and interpretability, aiding in advanced hydrate management techniques and improving operational safety and efficiency.

REFERENCES

Barker, J. W. (1989). Formation of Hydrates during Deep water Drilling Operations, . *SPE/ADC 16130, presented at the 1987 SPE/LADC Drilling Conference.* New Orleans, LA.: SPE/ADC.

Dorstewitz F. Mewes, D. (1995). Hydrate Formation in Pipelines. *Presented at the Fifth International Offshore and Polar Engineering Conference*

Garapati, B. J., N. A. (2014). Statistical thermodynamics model and empirical correlations for predicting mixed hydrate phase equilibria. *Fluid Phase Equilibria*, *373*(NA), 20–28. https://doi.org/10.1016/j.fluid.2014.03.010

Gbaruko, B. C. (2004). Asphaltenes, oil recovery and down-hole upgrading in the Nigerian petroleum industry. In *ACS National Meeting Book of Abstracts* (Vol. 228, Issue 2). 228 ACS National meeting.

Hackeling, G. (2014). *Mastering Machine Learning with scikit-learn*. Birmingham: Packt Publishing.

Infochem, Technologies. (2015). *Multiflash for Windows*. Surrey: KBC Advanced Technologies Ltd.

Kaplan, A., & Haenlein, M. (2019). Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. *Business Horizons*, *62*(1), 15–25. https://doi.org/10.1016/j.bushor.2018.08.004

Kummamuru P.; Lenaerts, Silvia, N. B. . P. (2021). A New Generalized Empirical Correlation for Predicting Methane Hydrate Equilibrium Conditions in Pure Water. *Industrial & Engineering Chemistry Research*, *60*(8), 3474–3483. https://doi.org/10.1021/acs.iecr.0c05833

Moradi, E., G. K. (2013). Modeling of hydrate formation conditions for CH4, C2H6, C3H8, N2, CO2 and their mixtures using the PRSV2 equation of state and obtaining the Kihara potential parameters for these components. *Fluid Phase Equilibria*, *38*(NA), 179–187. https://doi.org/10.1016/j.fluid.2012.11.010

NAPIMS. (2023). *Crude Oil Reserves/ Production*. http://www.napims.com/crudeoil.html .

Nazifi, S., O. G. (2024). Crop Yield Prediction Using Selected Machine Learning Algorithms. *FUDMA Journal of Sciences (FJS)*, 61 - 68.

Odutola, T. (2022). *Engineering and Scientific Research (UJESR) Comparing the Effectiveness of Methanol , Ethanol and Monoethylene Glycol at Preventing Hydrate Formation in a Hydrate Flow Loop*. *6*(2), 6–13.

Odutola, T. O., Ikiensikimama, S. S., & Ajienka, J. A. (2015). Effective Hydrate Management during Gas Expansion. *Effective Hydrate Management during Gas Expansion Conference Paper at Nigeria Annual International Conference and Exhibition (2015) Society of Petroleum Engineers/2015/ SPE 178342*.

Patrice, K., & Lenaerts, S. N. (2021). A New Generalized Empirical Correlation for Predicting Methane Hydrate Equilibrium Conditions in Pure Water. *Industrial & Engineering Chemistry Research*, *60 (8)* , 3474–3483.

Qin, H. (2020). *Hydrate Film Growth and Risk Management in Oil/Gas Pipelines using Experiments, Simulations, and Machine Learning*.

Saeed, Z. and E. A. (2021). Modelling the Formation of Gas Hydrate in the Pipelines. *Petroleum & Petrochemical Engineering Journal*, *5*(1), 1–14. https://doi.org/10.23880/ppej-16000259

Schlumberger. (2022). *Release notes*. Schlumberger.

Seo, Y., Kim, B., Lee, J., & Lee, Y. (2021). Development of ai-based diagnostic model for the prediction of hydrate in gas pipeline. *Energies*, *14*(8), 1–22. https://doi.org/10.3390/en14082313

Srivastava, V. (2018). *Quantitative Risk Modeling of Hydrate Bedding Using Mechanistic, Statistical, and Artifical Neural Network Frameworks*.

Swamynathan, M. (2017). *Mastering Machine Learning with Python in Six Steps*. Bangalore, Karnataka, India: Freepik.

Wang, J., Wang, Q., Meng, Y., Yao, H., Zhang, L., Jiang, B., Liu, Z., Zhao, J., & Song, Y. (2022). Flow characteristic and blockage mechanism with hydrate formation in multiphase transmission pipelines: In-situ observation and machine learning predictions. *Fuel*, *330*(August), 125669. https://doi.org/10.1016/j.fuel.2022.125669

Yu, Z., & Tian, H. (2022). Application of Machine Learning in Predicting Formation Condition of Multi-Gas Hydrate. *Energies*, *15*(13). https://doi.org/10.3390/en15134719

Zarei Mostafa; Mohammadi, Amir H.; Moosavi, Ali, M. M. H. (2021). Model development for estimating calcium sulfate dihydrate, hemihydrate, and anhydrite solubilities in multicomponent acid and salt containing aqueous solutions over wide temperature ranges. *Journal of Molecular Liquids*, *328*(NA), 115473-NA. https://doi.org/10.1016/j.molliq.2021.115473

Zerpa, L.E., E. S. (2012). Hydrate Risk Assessment and RestartProcedure Optimization of an Offshore Well. *Offshore Technology Conference Brasil* (pp. 49-56). Rio de Janeiro: Society of Petroleum Engineers.

Zhong Qi-Long; Wang, Yu; He, Yufa; Li, Zi-Han, H.-Q. Y. (2019). A Graphical Alternating Conditional Expectation to Predict Hydrate Phase Equilibrium Conditions for Sweet and Sour Natural Gases. *Mathematical Problems in Engineering*, *2019*(NA), 1–15. https://doi.org/10.1155/2019/2383961