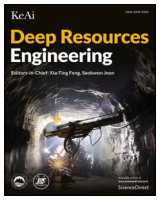
Deep Resources Engineering 3 (2026) 100233

Contents lists available at ScienceDirect

Deep Resources Engineering

journal homepage: www.keaipublishing.com/en/journals/deepre/



Full length article

An innovative approach for oil well bottomhole pressure forecasting using Kolmogorov-Arnold Neural Networks (KANs): A case study in an offshore oilfield

Adeboye Adeyinka [a,*], Opedamola Oriola [b], Olusegun Stanley Tomomewo [a,*]

[a] Energy and Petroleum Engineering Department, College of Engineering and Mines, University of North Dakota, ND 58202, USA
[b] Computer Vision Department, Inveniam, France

ARTICLE INFO

Keywords:
Time-series forecasting
Kolmogorov-Arnold networks
Interpretable machine learning
Bottomhole pressure prediction
Reservoir engineering
Numerical reservoir simulation

ABSTRACT

This study introduces a novel method for bottomhole pressure (BHP) prediction utilizing the Kolmogorov-Arnold Network (KAN) machine learning algorithm. The KAN model is an interpretable neural network that employs learnable non-linear activation functions to model data peculiarities. To enhance reservoir management and optimize production results, the reservoir engineer must accurately forecast dynamic reservoir properties such as BHP and its temporal changes. BHP trends are conventionally acquired by utilizing pressure gauges in designated wells and incorporating the data into reservoir simulation or nodal analysis models. Conducting these pressure surveys can be challenging due to the associated expenses and missed production opportunities when performed on a flowing well lacking a permanent downhole gauge. The KAN model was utilized to forecast BHP for a single well in the Volve oilfield using exclusively surface-measured data and engineered features. We employed several machine learning algorithms, including linear regression, KAN, neural networks and tree-based models (XGBoost and GBM), enabling reservoir engineers to prioritize the acquisition of specific surveillance data based on the highly correlated features within the models. The models were evaluated using the coefficient of determination (R^2), mean absolute error (MAE), and root mean squared error (RMSE) metrics, with the KAN model demonstrating superior performance compared to the other regression models. The KAN model also outperformed the full-field history-matched reservoir simulation model provided by the oilfield operator and gave the best results when compared to actual downhole gauge data, demonstrating its superior predictive capabilities.

1. Introduction

This paper focuses on the prediction of bottomhole pressure (BHP) and its evolution over time. BHP is a dynamic property that plays a critical role in the reservoir engineer's understanding of the subsurface as it directly influences reservoir management practices, production optimization actions, and well performance analysis. BHP improves reservoir performance predictions by acting as a boundary condition in reservoir simulation models. Accurate BHP estimates can assist in the calculation of well productivity index and calibration of simulation models required for production optimization purposes [1-3].

Understanding BHP trends is important because BHP is an indicator of the energy available within the reservoir to lift fluids from the bottomhole to the surface and it can be managed by balancing withdrawal and injection volumes within the reservoir [4-6]. In practice, the engineer regulates production and injection volumes by adjusting choke sizes [7]. This measure of control could be important for efficient reservoir management and achieving specific reservoir states, for instance, where it has been pre-determined that the reservoir is to be produced above bubble point to prevent the formation of a secondary gas-cap [8-10].

In addition, field development decisions rely on accurate economic evaluations, which are closely linked to BHP predictions as pressure trends affect production decline forecasts, recovery factors, reserves and investment actions [11,12]. The reservoir engineer should therefore understand the near-wellbore pressure conditions needed to optimize flow rates and reduce coning effects or water breakthrough [13-15]. BHP prediction also plays an important role in enhanced oil recovery

Peer review under responsibility of Editorial Board of of Deep Resources Engineering.
* Corresponding authors.
E-mail addresses: adeboye.adeyinka@und.edu (A. Adeyinka), damola.oriola@inveniam.fr (O. Oriola), olusegun.tomomewo@und.edu (O.S. Tomomewo).

https://doi.org/10.1016/j.deepre.2025.100233
Received 27 July 2025; Received in revised form 19 October 2025; Accepted 5 December 2025
Available online 7 December 2025

(EOR) processes where gas, chemical or steam injection are used. Achieving optimum pressure control in these EOR systems helps maintain a stable flood front, improve sweep efficiency and ultimately leads to higher recovery [16-18].

For subsurface safety and integrity considerations, maintaining an appropriate BHP ensures the wellbore pressure remains within safe operational limits thereby preventing formation damage. Excessive drawdowns could lead to sand production or completion failure which could negatively affect well productivity and reservoir integrity [19-21]. In assets where there is active drilling, reliable BHP predictions help achieve wellbore stability during drilling operations thereby preventing safety risks such as wellbore collapse or blowouts [22-24].

To properly characterize the dynamic behavior of a reservoir and understand the BHP trend over time, the reservoir engineer must incorporate data from multiple sources to aid the understanding of the subsurface [25-27]. Seismic, core, petrophysical, production, injection and pressure data are integrated into a reservoir simulation model which mimics the dynamic behavior of the subsurface through a time-consuming, iterative history matching process with the end goal being to forecast fluid withdrawal rates, fluid saturation and pressures (static and flowing) over time [28-30]. Reservoir simulation models predict BHP based on subsurface properties and fluid flow guided by differential equations. Multiphase flow correlations are then used to generate the pressure drop along the tubing using correlations that may sometimes fall short due to the complex and dynamic nature of downhole and wellbore conditions [31].

History matching and forecasting with a reservoir simulation model is a labor and resource intensive process that must be repeated at a predetermined frequency (quarterly, bi-annually or annually) by the asset team as new surveillance data becomes available [32-34]. Recent advances in machine learning now offer an alternative, quick and reliable way to achieve the same results for identified dynamic properties [35, 36].

The impact of using interpretable machine learning models for dynamic reservoir property prediction has not yet been fully explored in the literature. Using interpretable models instead of black box models can significantly contribute to the reservoir engineer's understanding of the subsurface and focus surveillance activities on gathering the data that matters. This work aims to fill this research gap by exploring the application of interpretable machine learning techniques (Kolmogorov-Arnold Neural Network, KAN and linear regression) to predict downhole bottom hole pressure (BHP) in the Volve field. The results are compared to neural network model results and tree-based algorithms. The ability to predict subsurface dynamic properties such as downhole pressure and fluid saturation as they vary with time and distance is an important step in dynamic reservoir characterization [37-39].

In this paper, we propose a machine learning approach for BHP prediction using only surveillance data acquired for surface-measured variables. This study contributes to the currently limited literature available on the application of KAN for BHP predictions utilizing real-world data. The results show that KAN achieves a reliable degree of prediction for BHP values. The study also shows that linear regression models, when augmented with some reservoir engineering knowledge and features, can give improved results. These machine learning approaches are easy to set-up and update and offer significant time-savings when compared to the full-scale reservoir simulation model with the machine learning models used in this work all taking less than 5 min to run.

The key contribution of this research is that it provides a method of applying an interpretable machine learning technique that facilitates quick decision-making and estimation of BHP trends in the subsurface. Some of the merits of the method used in this paper are that usable analytical equations are generated by the KAN and linear regression machine learning models, the machine learning models are an order of magnitude faster to run than full-field simulation models and the machine learning models are built with much less data and work hours than is needed for a reservoir simulation model. To the best of the authors' knowledge, this is the first work that uses the Kolmogorov-Arnold interpretable neural network to predict BHP trends using real world data from an oil field.

Section 2 provides a review of relevant literature relating to BHP prediction using machine learning techniques. Literature specific to the work of previous authors on the Volve field were also reviewed. Section 3 looks at the data provided, the features that can be created from the data and the machine learning algorithms and metrics used. Section 4 discusses the results and compares the machine learning model predictions making comparisons with the gauge data and reservoir simulation data while Section 5 discusses the conclusions of the study.

2. Literature review

Some of the advancements in BHP prediction accuracy have been brought about by the increased deployment of machine learning algorithms. Accepted industry methods such as numerical reservoir simulation capture the complexities of real-world reservoirs but require significant time and computing power [40-42]. Another industry-accepted method is empirical or mechanistic models which tend to approximate reality and are applicable within the limits of the simplifying assumptions made [43,44]. Machine learning methods have created new opportunities, enabling data-driven solutions to predict BHP with accuracy [45]. Thus, this literature review explores the evolution of machine learning approaches for BHP prediction, examining key methodologies, applications, and insights.

Reservoir simulation models have been the conventional approach when predicting subsurface properties such as fluid saturation and pressure which varies over space and time. These simulation models have found great acceptance and wide applicability being used to model petroleum reservoirs systems. The simulation model employs a combination of partial differential equations, mechanistic models and correlations which are usually time consuming and computationally intensive to solve. In addition, the subsurface is usually complex and dynamic with numerous factors such as temperature, pressure, fluid properties and geology interacting in different ways. This complexity makes it difficult to accurately predict subsurface properties such as BHP.

Machine learning algorithms have shown a lot of promise in addressing the limitations of industry-accepted BHP prediction techniques. Navratil et al. [46,47] demonstrated the effectiveness of artificial neural networks (ANNs) in modeling nonlinear relationships in reservoir data in shorter time and with better computational performance and comparable results to numerical simulations in dynamic conditions.

In some instances, physics-informed models have been used to complement the capabilities of neural networks. Jin et al. [48] integrated physics-based and data-based ANN models for continuous BHP estimation. Though both models demonstrated good results, the physics-based approach displayed a higher degree of accuracy during the later phase of production under steady-state conditions. Models based on data produced more favorable outcomes in the initial stages of production when flow was still transient. Zalavadia et al. [49] employed both physics-based and data-based models in a large U.S. unconventional shale basin with more than 3000 producing wells providing accurate BHP estimates. Zheng et al. [50] integrated physics-based models (based on wellbore multi-phase flow theory) with machine learning models (XGBoost and Feedforward neural network) to form a knowledge-guided machine learning model that incorporated domain knowledge from the physics-based model and applied it as an extra penalty on the loss function. This method improved the prediction of BHP from permanent downhole gauge data and gave robust results across different wells in the field that it was tested on.

Several authors have used hybrid neural network models in their work to extend the functionality of ANN. Zhu et al. [51] went further by developing a hybrid neural network model combining backpropagation,

long short-term memory (LSTM), and one-dimensional convolutional neural networks to predict BHP during managed pressure drilling with the hybrid model achieving improved accuracy by combining the advantages of the different single models. Ashena and Moghadasi [52] developed a hybrid model combining ANNs and genetic algorithms. This hybrid model outperformed the ANN standalone model in the estimation of BHP. Ahmadi and Chen [53] similarly built an ANN model and optimized the hyperparameters using a hybrid genetic algorithm and particle swarm optimization approach with a maximum of 10 % error when compared to the measured pressure data. Panja et al. [54] used convolutional LSTM to predict spatiotemporal parameters in the SACROC oil field. The models were trained for saturation and pressure prediction with promising results seen. The authors observed that though the model exhibited superior performance during the initial phase, the deviations increased with time due to accumulated errors from past predictions.

The scope of machine learning applications in BHP prediction has been expanded by some authors to include ensemble techniques such as random forest and gradient boosting algorithms. Ignatov et al. [55] used gradient boosting algorithms and highlighted their ability to handle complex non-linear input relationships and noisy datasets while [56] showed the importance of feature selection algorithms in improving model performance by using a gradient boosting regressor to evaluate feature importance for a shallow marine reservoir. Eltahan et al. [57] implemented a novel approach by building machine learning algorithms to predict the difference between the empirical estimates and actual values of BHP using data from 11 producing wells. They concluded that the random forest model gave the best results and helped quantify the uncertainty in BHP prediction.

With the increase in deployment of permanent downhole gauges and Internet of things (IoT) devices, real-time data can be processed through machine learning models for insights. Tian and Horne [58] processed real-time data and connected nonlinear features to the pressure targets. They successfully modelled the pressure data from the permanent downhole gauges by using the features as an approximation of model characteristics. Mishra et al. [59] explored data integration techniques, showing that incorporating geological and operational data into machine learning models enhanced the robustness of the prediction. The authors demonstrated the value of utilizing machine learning techniques to develop data-driven models for predicting bottomhole pressure and temperature from wellhead measurements and injection rates.

Several authors have applied machine learning techniques for BHP prediction specifically in the Volve field. Olamigoke and Onyeali [60] used three machine learning models, including Random Forests and Long Short-Term Memory (LSTM) to predict BHP pressure as a time-series problem for two wells in the Volve field. Their results showed reasonable accuracy in BHP prediction. They also investigated the feature combinations that gave superior prediction performance and concluded that the most indicative features were bottomhole temperature (BHT), oil flow rate, gas flow rate, choke size and onstream hours. Akinsete and Adesiji [61] built an Artificial Neural Network (ANN) to predict the BHP in Volve field. They found that ANN outperformed other models such as Support Vector Machines (SVM), Decision Trees (DT) and Random Forest (RF) in the prediction of BHP. Blessing and Agbons [62] developed machine learning models using RF and Gradient Boosted Models (GBM) to predict the BHPs in the Volve field. The developed models predicted BHPs for the wells with an acceptable degree of accuracy. Li et al. [63] used Recurrent Neural Network (RNN) models to analyze well behavior and predict BHP in the Volve field as a time series. The production data from 1 well (15/9-F-1 C) which produced from April 2014 to April 2016 was used for both training and testing the models. Single parameters were used to train a single layer gated recurrent unit (GRU) model to predict the flowing BHP, the best parameter (based on the least relative square error) was found to be bottomhole temperature, followed by the choke opening percentage and then water production. A summary is provided in Table 1 below.

Table 1
Review of some selected studies specific to Volve field BHP Prediction.

Author	Model employed	Comments
Olamigoke, Onyeali [60]	LSTM, RF, SVM	Wells F15D and F1C analyzed. Downhole variables were used in the study. LSTM gave the best results regardless of feature combinations with test R^2 of 92.2 %.
Akinsete, Adesiji [61]	ANN, RF, DT, SVM	All 7 wells analyzed with only 16 % of the dataset selected for use. Authors also included pressure drop in tubing which is data leakage. Downhole variables were used in the study. ANN achieved the best results with test R^2 of 99.9 %.
Blessing, Agbons [62]	RF, GBM	All 7 wells analyzed with only 23 % of the dataset selected for use. Downhole variables were used in the study. RF achieved the best results with test R^2 of 97.8 %.
Li et al. [63]	RNN, LSTM+GRU, LSTNet	Well F1C analyzed. Downhole variables were used in the study. LSTNet yielded the best test score. R^2 was not used as a metric.
This work	LR, KAN, XGBOOST, GBM, Bi-LSTM	Well F1C analyzed with over 97 % of the dataset selected for use. No downhole variable was used. KAN yielded the best results with test R^2 of 96.4 %.

The machine learning techniques used in this paper exclusively used only on data measured at the surface as an input in the prediction of BHP. Surface measurements such as on-stream hours, tubing head pressure, choke size, pressure drop across choke and fluid rates (for oil, gas and water) are used as features that feed into the model. Based on the authors' reservoir engineering knowledge, some additional features were also created to mimic subsurface properties. The best features were selected from this dictionary of features and used to predict BHP which was then compared to the actual BHP taken in gauges and the simulated BHP from a history-matched reservoir simulation model. The results indicate that machine learning models can significantly improve the accuracy of BHP predictions, making them a complementary tool that can provide quicker and accurate results compared to reservoir simulation models.

3. Methodology

This research aims to add to the existing literature by applying machine learning techniques to develop proxy models using real-life data from the Volve oil field. The proposed workflow processed the available dataset using the following steps: data audit, data cleaning, feature generation, machine learning modeling. The aim was to predict BHP in a single well (F1C) using surface-measured data and engineered features.

3.1. Data description

The data used for this work was obtained from the Volve oil field. Equinor officially released the complete set of data from the Volve field in the North Sea for research, study, and development purposes in 2018 [64]. The Volve Field is located in Block 15/9 (see Fig. 1) approximately 200 km west of Sleipner A Platform in water depths of about 90 m [65]. The field was discovered in 1993 and production commenced in 2008 till 2016. A total of 24 wells were drilled including oil and associated gas producers, water source wells, water injection wells and some abandoned wells.

The data provided by the operator covers 5 oil producers and 2 water injectors. One of the water injectors (F5AH) was converted into a producer later in the life of the field and produced for less than 6 months (see Fig. 2 below). A history-matched reservoir simulation model was

Fig. 1. Geographic location of the volve field [66].

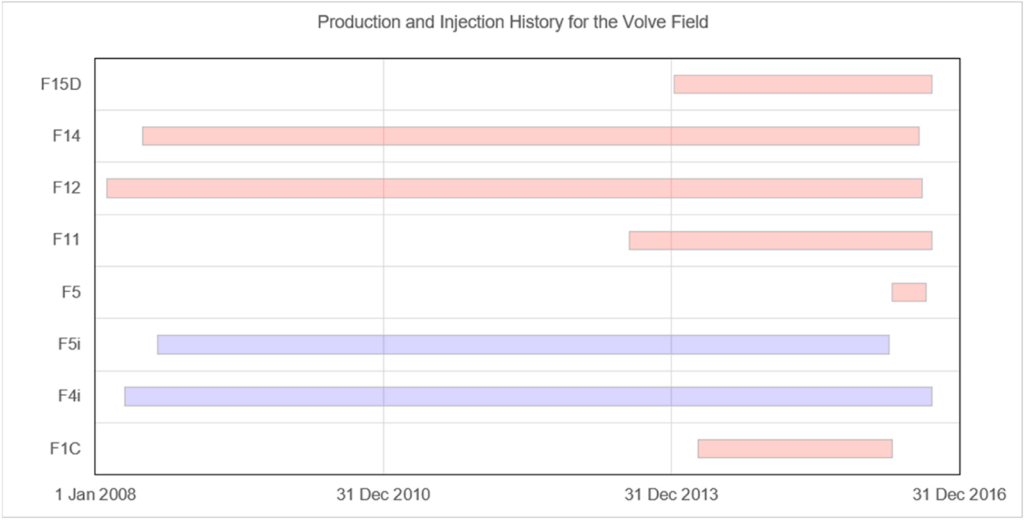


Fig. 2. Production (red) and water injection (blue) periods for the Volve wells.

also provided (see Fig. 3 below).

3.1.1. Data audit

The data provided by Equinor for all the producer and injector wells had 15,634 rows and 24 columns. Some of the columns were dropped since they did not have any significance to the study. Data such as the Field Code, Wellbore Code, Facility Name and Field Name whose rows essentially contained the same data were dropped to reduce the number of feature inputs into the model. Of all the data supplied, we found only 8 features that were numerical and relevant to the study.

During the data audit, it was also discovered that the tubing pressure drop (dp_tubing) was derived directly from the subtraction of the tubing head pressure recorded at the surface (THP) from the BHP read by the downhole gauge, i.e. dp_tubing = BHP – THP. Since BHP is supposed to be the predicted quantity, we dropped the dp_tubing from the input features to avoid a case where the machine learning model already had the answer right from the features. This action was important as some of the researchers in the literature included dp_tubing in their feature inputs thus inadvertently creating data leakage.

3.1.2. Data cleaning

The data was cleaned based on reservoir engineering domain knowledge. Some features which had no direct bearing on the BHP predictions were dropped. Bottomhole temperature (BHT) feature was also dropped to ensure that all feature inputs into the machine learning model were surface-based readings, and no downhole feature was included. Rows with BHP = 0 were also dropped. This constraint was applied based on the reservoir engineering perspective that drawdown in producer wells is managed based on the field's operating philosophy and it is very unlikely that the wells would be allowed to produce at BHP

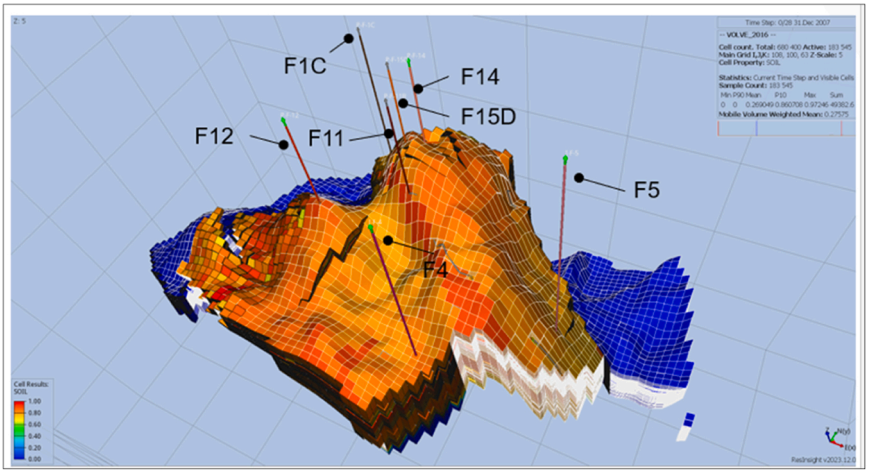


Fig. 3. Image from the Volve field reservoir simulation model showing production and injection wells [64].

= 0 which could damage the formation or cause borehole collapse.

3.1.3. Feature generation

Additional features were generated from the initial dataset of 8 variables based on reservoir engineering domain knowledge to help the machine learning models properly calibrate the drawdown and build up pressure responses. Since F1C was the well of interest, we created an expanded dataset that included the production and injection rates, pressure data (casing, tubing head), choke size and well uptime for all the other wells during the producing life of F1C as new features. This would help capture the interaction between the wells since they were all completed in the same reservoir. Features such as cumulative field oil production (FOPT), cumulative field gas production (FGPT), cumulative field water production (FWPT), cumulative field water injection (FWIT) and well cumulative production and injection volumes were added to help the model understand the total volumes that had been taken out of the subsurface tank and the voidage replacement ratio. Table A.1 in Appendix A has a full list of the expanded dataset that contains the generated features.

3.2. Machine learning methods

Five machine learning models were used: linear regression (LR), Kolmogorov-Arnold Neural network (KAN), eXtreme Gradient Boosting (XGBoost), Gradient Boosting Machine (GBM) and Bidirectional LSTM (BiLSTM). LR and KAN models were selected because they output equations that if needed, a reservoir engineer could easily input into a spreadsheet for quick computations of the target variable. A brief overview of each model is provided in the following sections.

3.2.1. KAN

Kolmogorov–Arnold networks (KANs) are based on the Kolmogorov–Arnold representation theorem which states that a multivariate continuous function on a bounded domain can be represented as a sum of continuous univariate functions.

KAN differs from conventional neural networks which utilize multilayer perceptron (MLP) in the architecture of the neural network. MLPs rely on the universal approximation theorem which states that a feedforward neural network with at least one hidden layer, nonlinear activation functions and a finite number of neurons in the hidden layer can approximate any continuous function arbitrarily well [67]. The universal approximation theorem is represented as:

$$f(x) \approx \sum_{i=1}^{N(\epsilon)} a_i \sigma(w_i \cdot x + b_i) \tag{1}$$

Where $f(x)$ is the target function to be approximated, a_i, w_i and b_i are the weights and biases of the hidden neurons, x is the input feature and σ is the activation function of the hidden neurons. $N(\epsilon)$ is the number of neurons (or hidden units) in the network and ϵ is the desired accuracy or margin of error between the true value of a target variable and the neural network's prediction. In practice, ϵ is minimized using a loss function.

While MLPs have fixed activation functions at the nodes of the neural network, KANs have learnable functions on the edges of the neural network. The KAN model is tuned by adjusting the learnable functions, which are represented by splines, and then subsequently added up at the nodes. MLPs learn by adjusting the weights and bias used to multiply the input features before passing them through a fixed nonlinear activation function such as ReLU (Rectified Linear Unit) or sigmoid at the nodes [68]. In contrast, for KAN, each input feature is passed through a learnable nonlinear activation function instead of a single scalar weight and bias. The Kolmogorov-Arnold representation theorem is presented as:

$$f(x) = f(x_1, \quad \ldots, \quad x_n) = \sum_{q=1}^{2n+1} \Phi_q \left(\sum_{p=1}^{n} \phi_{q,p}(x_p) \right) \tag{2}$$

where $\phi_{q,p}(x_p)$ are the univariate functions and Φ_q takes the univariate functions and combines them. The equation above represents a KAN that has only two layers and a limited number of terms ($2n + 1$). This was improved upon and generalized into arbitrary widths and depths in the work of [69].

The learnable nonlinear activation functions that KAN uses to approximate continuous univariate functions are called b-splines. Splines are functions that are piecewise polynomials smoothly connected at certain points called knots. The splines use different polynomial expressions over different intervals of the domain to be approximated. Thus, KANs can model data peculiarities and patterns but might require significant computational resources when used on big datasets with a lot of input features. The KAN package used in this work can be installed using the python manager PIP, with the following command: pip install pykan.

3.2.2. Linear regression

A linear regression (LR) machine learning model is a supervised machine learning model that predicts a continuous target variable by assuming a linear relationship between the input features and the target variable.

$$y = b + \quad w_1 x_1 + w_2 x_2 + \ldots + w_n x_n \tag{3}$$

where y is the target variable, x_1, x_2, …, x_n are the input features and b, w_1, w_2, …, w_n are the bias and weights for each feature.

Linear regression models are simple, easy to understand and relatively fast to train. Linear regression minimizes the square error between the actual and predicted data points to find a best-fit line. Linear regression models may, however, not perform well with complex nonlinear relationships unless the input features are modified by creating pseudo-features, or using polynomial interactions or log transformations that can appropriately capture the nonlinearity.

3.2.3. Extreme gradient boosting

XGBoost (eXtreme Gradient Boosting) is an ensemble learning method that sequentially builds a series of weak learners using decision trees, to create a strong predictive model where each subsequent tree is trained to correct the residual errors of the previous ones by minimizing a differentiable loss function.

XGBoost uses a regularization term that incorporates Lasso and Ridge penalties to control model complexity and prevent overfitting, which is a key advantage over traditional gradient boosting.

3.2.4. Gradient Boosting Machine

Gradient Boosting Machine (GBM) is an ensemble learning technique that builds the model in a stepwise fashion by iteratively adding weak learners to the ensemble of decision trees. GBMs leverage on the principle of gradient descent optimization where at each iteration, a new decision tree is trained to minimize the residual errors of the existing ensemble. This is achieved by fitting the new tree to the negative gradient of a predefined loss function, computed with respect to the current model's predictions. The final prediction is the sum of the predictions from all the individual trees in the sequence.

3.2.5. Bidirectional Long Short-Term Memory

Bidirectional Long Short-Term Memory (BiLSTM) networks is an extension of the recurrent neural networks (RNN) that can capture past and future context in sequential data. It enhances the standard Long Short-Term Memory (LSTM) network, which uses a gated cell mechanism to learn long-term dependencies and handle the vanishing gradient problem.

BiLSTM consists of two parallel LSTM layers: one processes the input sequence in the forward direction, while the other processes it in reverse. The outputs from both directions are then combined, allowing the model to leverage information from the entire input sequence at each time step.

3.3. Evaluation metrics

Three metrics: coefficient of determination (R^2), mean absolute error (MAE) and root mean squared error (RMSE) were used to evaluate the performance of the machine learning models. A brief explanation for these metrics is provided below.

The coefficient of determination, R^2, is a metric that represents the proportion of the variance in the dependent variable that is predictable from the independent variables. R^2 measures goodness of fit and typically ranges from 0 to 1, with a higher R^2 suggesting a better fit. The formula is given as:

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2} \tag{4}$$

where y_i is the actual value of the i-th data point of the target variable, $\hat{y}_i$ is the predicted value of the i-th data point of the target variable, $\bar{y}$ is the mean value and n is the number of data points.

The mean absolute error, MAE, measures the average magnitude of errors in a set of predictions without prejudice to whether the prediction was an overestimation or underestimation. A lower MAE implies a more accurate model with the formula given as:

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^{n} |y_i - \hat{y}_i| \tag{5}$$

where y_i is the actual value of the i-th data point of the target variable, $\hat{y}_i$ is the predicted value of the i-th data point of the target variable and n is the number of data points.

The root mean square error, RMSE, measures the average magnitude of the square of errors in a set of actual versus predicted data values. Due to the squaring operation, it is more sensitive to outliers since the errors are further magnified. A lower RMSE implies a more accurate model. The formula is given as:

$$\text{RMSE} = \sqrt{\left[\frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2\right]} \tag{6}$$

where y_i is the actual value of the i-th data point of the target variable, $\hat{y}_i$ is the predicted value of the i-th data point of the target variable and n is the number of data points.

4. Results

F1C was selected as the candidate well for this analysis. The selection of this producer was based on the premise that machine learning models typically require a sizeable dataset to be effectively trained and as such, using the producer with the shortest production time would be a good test for the applicability of the models.

4.1. Data preparation

F1C produced from Apr 2014 till Apr 2016 with the original dataset provided having 746 rows and 24 features. A total of 13 features were dropped since they had no immediate relevance to the analysis, contained zero or empty values, or contained the same data all through. Table A.2 in Appendix A contains a summary of the dropped features. In addition, downhole variable, BHT, and the tubing pressure drop (dp_tubing) were also dropped due to reasons stated earlier in the data audit section. Rows with zero, empty or spurious BHP values were also dropped leaving the dataset with 720 rows and 11 features (including the target variable (BHP), date and well name features). The target feature (BHP) had a dimension of 720 rows and 1 column.

Additional features were then generated based on reservoir engineering knowledge with a total of 720 rows and 93 features available for analysis (Table A.1 in Appendix A). The training dataset was from start of production in April 2014 till September 2015. The validation set was from October 2015 to Dec 2015 while the test dataset was January to March 2016. This corresponded approximately to a dataset split of 75 % train, 12.5 % validation and 12.5 % test. There was no shuffling of the data before splitting it into train, validation and test data to prevent data leakage where the training model learns using data from the validation or test set. The scaler was applied on the training set only (to prevent information leakage during model training) and thereafter used on the validation and test. The standard scaler formula is as shown below:

$$X_{\text{new}} = \frac{X - \quad X_{\text{mean}}}{X_{\text{std}}} \tag{7}$$

Where X_{new} is the scaled value of X, X_{mean} and X_{std} are the mean value and standard deviation of the feature respectively.

4.2. Model results and discussion

In the process of this study, five different scenarios were tested. The first was a baseline linear regression model to predict BHP using only

Table 2
Summary of models used.

Model	Comments	Features used
baseLR	Baseline linear regression model	8 features from the original Volve dataset
optLR	Optimized linear regression model	14 features: mix of original and newly generated
altLR	Alternative linear regression model Predicted dp_tubing instead of BHP	12 features: mix of original and newly generated
baseKAN	Base KAN model	Same features as baseLR
optKAN	Optimized KAN model	Same features as optLR

input features from the original dataset. The second was an optimized linear regression model to predict BHP, with input features from both the original and expanded dataset. The third was an alternate linear regression model to predict dp_tubing instead of BHP, with input features from both the original and expanded dataset. The fourth, a base KAN model to predict BHP using only input features from the original dataset and lastly an optimized KAN model to predict BHP using same input features as the optimized linear regression model. The linear regression models were run in multiple scenarios in a bid to find an improvement in the linear regression model that would make it match the KAN model in performance. A summary of these scenarios is provided in the Table 2 below.

Table 3
Features used to build baseline linear regression model.

Feature in volve dataset	Renamed as	Comments
ON_STREAM_HRS	on_stream_hours	daily onstream production hours for F1C
AVG_CHOKE_SIZE_P	choke_size	F1C choke size
AVG_WHP_P	thp	F1C tubing head pressure
AVG_WHT_P	tht	F1C tubing head temperature
DP_CHOKE_SIZE	dp_choke_size	F1C pressure drop across choke
BORE_OIL_VOL	wopr	F1C oil production rate
BORE_GAS_VOL	wgpr	F1C gas production rate
BORE_WAT_VOL	wwpr	F1C water production rate

4.2.1. *Baseline linear regression model, baseLR*

The baseLR scenario was built using just 8 numerical features in the data supplied with the original Volve dataset. We did not consider any newly created features in the expanded dataset. The idea was to build a baseline linear regression model that would be used as a benchmark to judge the performance of the other models. Table 3 below shows a list of the 8 numeric features used.

4.2.2. *Optimized linear regression model, optLR*

The optimized linear regression model was built from a combination of the original 8 numeric features and an additional 85 features generated in the expanded dataset. Since it would be impractical to use 93 features in a linear regression model, we selected the best set of features using sequential forward feature selection (SFS) and recursive feature elimination (RFE). The SFS gave the better results and was carried forward in the optLR model.

SFS is a greedy algorithm that iteratively adds features based on how they improve the performance metrics of the model [70]. This SFS method significantly saves time as it is not as computationally expensive (needing only 93 runs) compared to an exhaustive feature selection for all possible combinations of 93 features which would require 9.9×10^{27} runs! Using the R^2 of the validation dataset as a scoring criterion, 14 features were selected using the SFS algorithm. The selected features and coefficients are captured in Fig. 4 below. Of the 14 best features selected, 7 are directly related to F1C (dp_choke_size, wwpr, on_stream_hours, wopr, thp, wopt, wwpt) while the remaining 7 features (wtime_cum_F4, wgpt_F14, wwpr_F12, thp_F15D, casing_pressure_F15D, casing_pressure_F14, wgpt_F12) capture interactions with other wells and field withdrawals.

4.2.3. *Alternative linear regression model, altLR*

The alternative linear regression model follows the same approach as the optLR model but instead of using BHP as the target variable, dp_tubing was used as the target variable and then added to the corresponding THP to get the BHP prediction. This gave more accurate results and is discussed further in the section on the discussion of results.

Based on the SFS algorithm used and its results on the validation dataset, 12 features were selected for use in the final altLR model. The selected features and coefficients are listed in Fig. 5 below. Of the 12 best

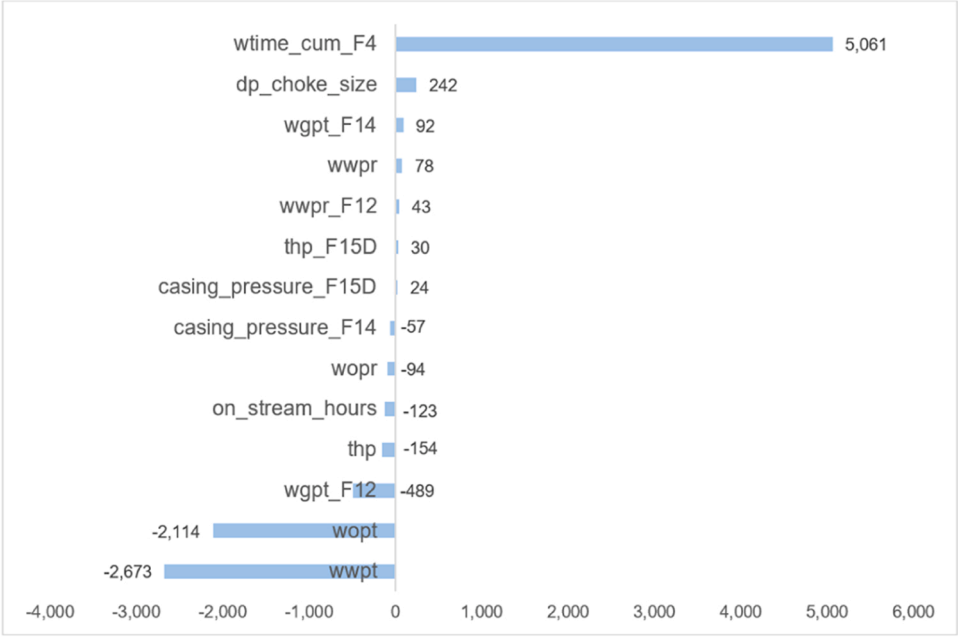


Fig. 4. Best selected features for predicting BHP in F1C and the correlation coefficients for optLR.

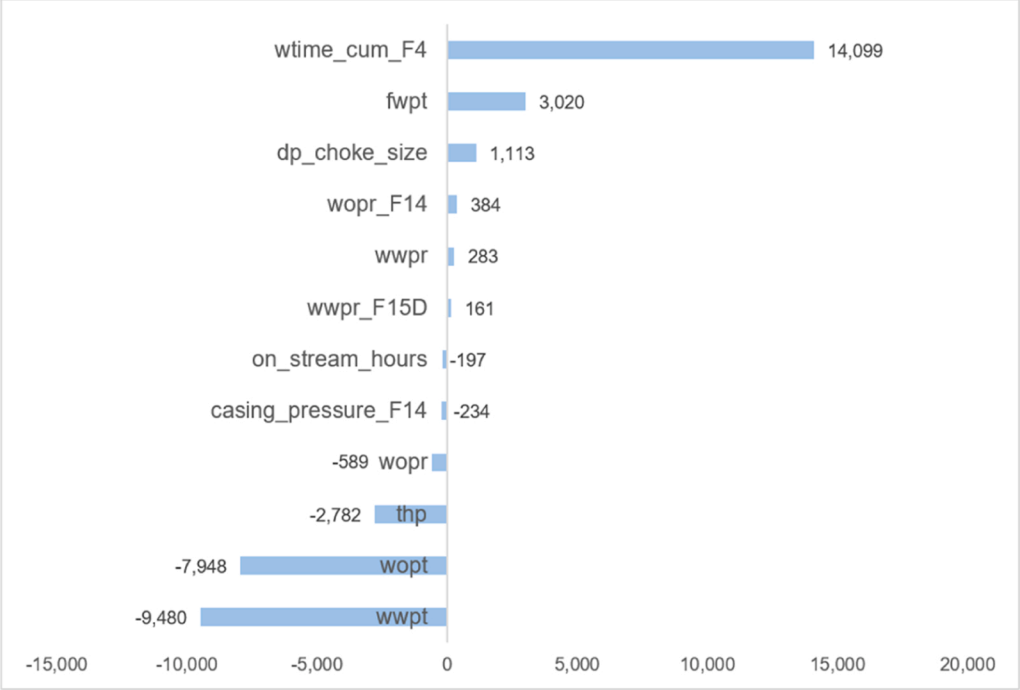


Fig. 5. Best selected features for predicting dp_tubing in F1C and correlation coefficients for altLR.

Table 4
Hyperparameters tested for the base KAN model.

Hyperparameter	Values tested
learning rate	[1, 1.5, 2]
hidden neurons	[6, 8, 10, 12]
grid	[1, 2, 3, 4, 5]
k	[1, 2, 3, 4]

Table 5
Hyperparameters tested for the optKAN model.

Hyperparameter	Range
learning rate	0.00001 – 5
hidden neurons	1 – 128
grid	1 – 20
k	1 – 40
optimizer	Adam, LBFGS

features selected, 6 are directly related to F1C (dp_choke_size, wwpr, on_stream_hours, wopr, wopt, wwpt) while the remaining 6 features capture interactions with other wells and field withdrawals.

4.2.4. *Base KAN model, baseKAN*

The base KAN model was built using the same 8 numeric features in the original volve dataset that was used to build the baseLR model with hyperparameter tuning carried out subsequently. The selected KAN model had 10 hidden neurons, grid of 2, k of 2 and learning rate of 1.5 with Adam (Adaptive Moment Estimation) as the optimizer. A table of the hyperparameters tested is shown in Table 4 below.

4.2.5. *Optimized KAN model, optKAN*

The optimized KAN model was built using the same 14 numeric features in the optimized linear regression model with hyperparameter tuning carried out subsequently. The selected KAN model had 16 hidden neurons, grid of 1, k of 22 and learning rate of 0.278 with LBFGS (Limited-memory Broyden Fletcher Goldfarb Shanno) as the optimizer. The hyperparameter search space for the optKAN model is reported in Table 5 below. The random seed was not part of the search space as it was fixed at a value of 15 for all the runs.

4.2.6. *Discussion of results*

The performance of the 5 models built is shown in Table 6 below. The results of the reservoir simulation model are also shown as ResSim. The baseLR had a test R^2 of 0.70 which was used as a baseline to judge the usefulness of any subsequent model built. It is therefore noteworthy that the optLR, altLR, baseKAN and optKAN models all outperformed the baseLR model.

We first compare the performance of the baseLR model vs the baseKAN model since they both used the same 8 input variables. The baseKAN with a R^2 of 0.85 outperforms the baseLR model with a R^2 of 0.7 on the test dataset. This shows that the baseKAN model explains 85 % of the variance in the target variable. The baseKAN also

Table 6
Summary of results for all models run.

Model	Runtime (seconds)	Train			Test		
		R^2	RMSE	MAE	R^2	RMSE	MAE
ResSim	-	0.754	1047.191	1021.619	0.878	1046.849	1025.172
baseLR	0.12	0.697	224.803	164.487	0.699	180.531	150.862
optLR	0.07	0.933	105.569	78.917	0.801	146.992	125.487
altLR	0.10	0.931	107.455	79.299	0.861	122.960	104.385
baseKAN	7.03	0.865	150.135	99.147	0.849	127.997	101.871
optKAN	218.73	0.991	38.642	26.462	0.906	100.941	85.603

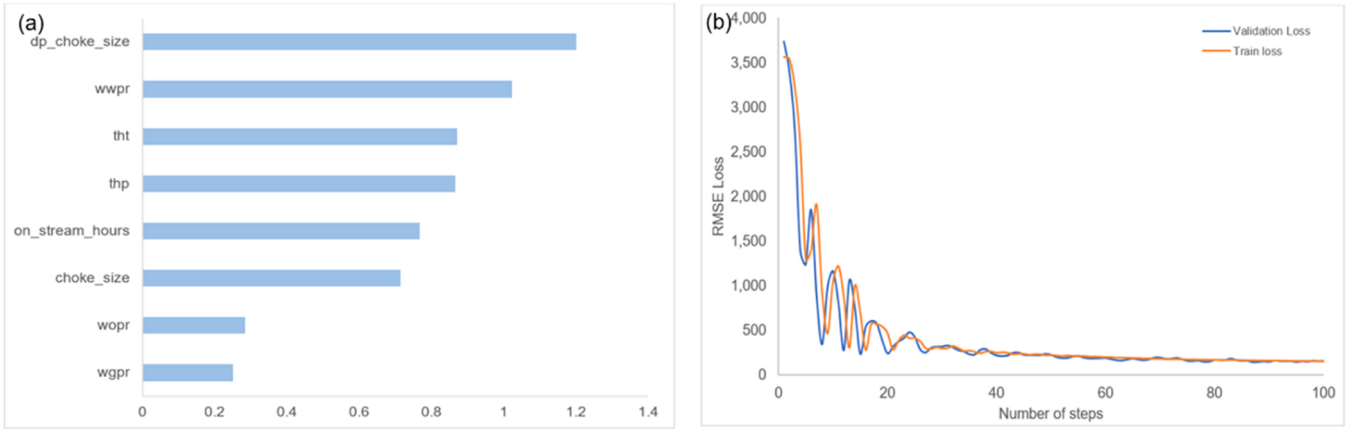


Fig. 6. (a) Feature importance of the baseKAN model; (b) Loss decay and stabilization during the training of the baseKAN model.

outperforms the baseLR in the RMSE and MAE metrics.

The feature importance of the baseKAN model is shown in Fig. 6 below with the top 5 most important features used to predict BHP given as dp_choke_size, wwpr, tht, thp and on_stream_hours. The loss function during the building of the model for the train and validation data is also shown where the loss decays and then stabilizes for both the train and validation data as the model progressively learns the patterns in the data. Observing a similar trend in train and validation data gives a measure of confidence that the model is not overfitting the data. The analytical expression generated for the baseKAN model is shown in Appendix B.

The optLR model with the expanded dataset consisting of features from the original Volve dataset and the newly created features achieved a R^2 of 0.80 on the test data. For the altLR model, redefining the target variable as dp_tubing instead of BHP and then using the predicted dp_tubing to subsequently calculate the corresponding BHP by adding the THP, led to further improvement with a R^2 of 0.86 on the test data. The altLR model also showed that it is possible to re-imagine the parameter used as the target variable in a bid to achieve better outcomes. It is important to note that the test data for dp_tubing was closer to normal distribution compared to the BHP and this may have influenced the improved prediction accuracy. The power of the baseKAN model is clearly demonstrated in that by using just the 8 original features from the Volve dataset, it achieved a significant R^2 value of 0.85 on the test dataset. When further optimized using the same 14 variables used in optLR, the optKAN model achieves a R^2 of 0.91 on the test dataset. This shows the potential of KANs to effectively learn complex patterns in the data.

The ResSim model with its full-field reservoir simulation, incorporating all available surface and subsurface data and leveraging the power of flow and mechanistic equations was still outperformed by the optKAN model. The ResSim model achieved a test R^2 of 0.88 compared to the 0.91 of the optKAN model. This highlights how results from machine learning models can complement those results from reservoir simulation models given that the machine learning models can be built in shorter time and with less computational power with similarly accurate results.

Though the R^2 results from the ResSim model are comparable to the other models, Table 6 above shows that the RMSE and MAE from the ResSim model are unusually high. This is because though the ResSim model follows the trend of test data time series, it is consistently higher than the actual downhole gauge data and doesn't appropriately capture the magnitude of the dips and rises. On the other hand, the optLR, altLR, baseKAN and optKAN models give good visual results and appropriately

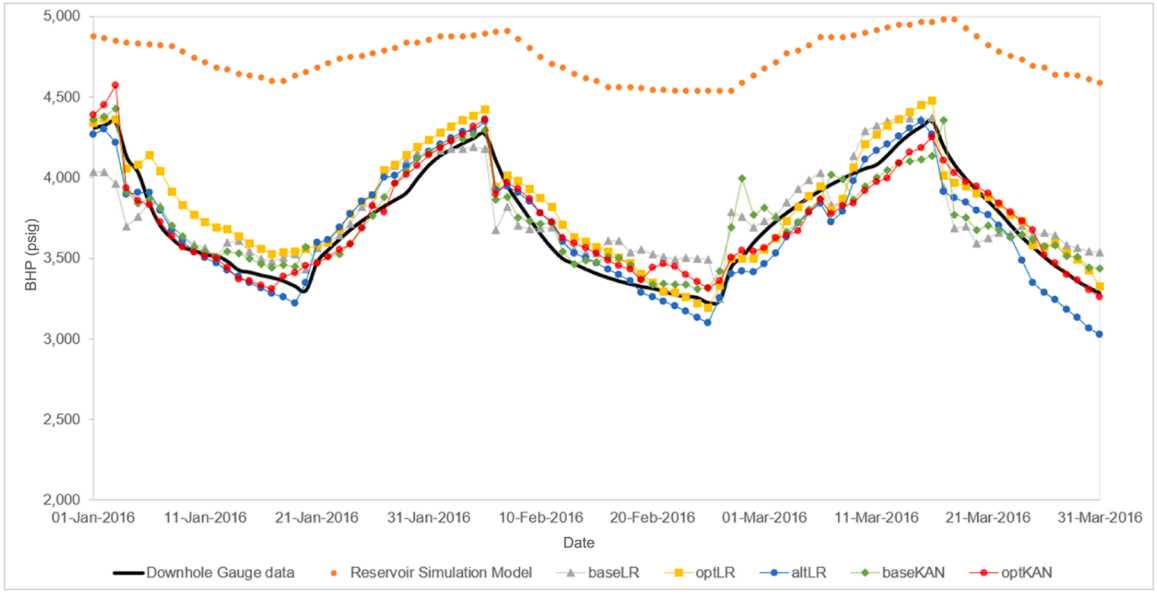


Fig. 7. Plot of actual downhole gauge data, reservoir simulation, optLR, altLR and KAN predictions.

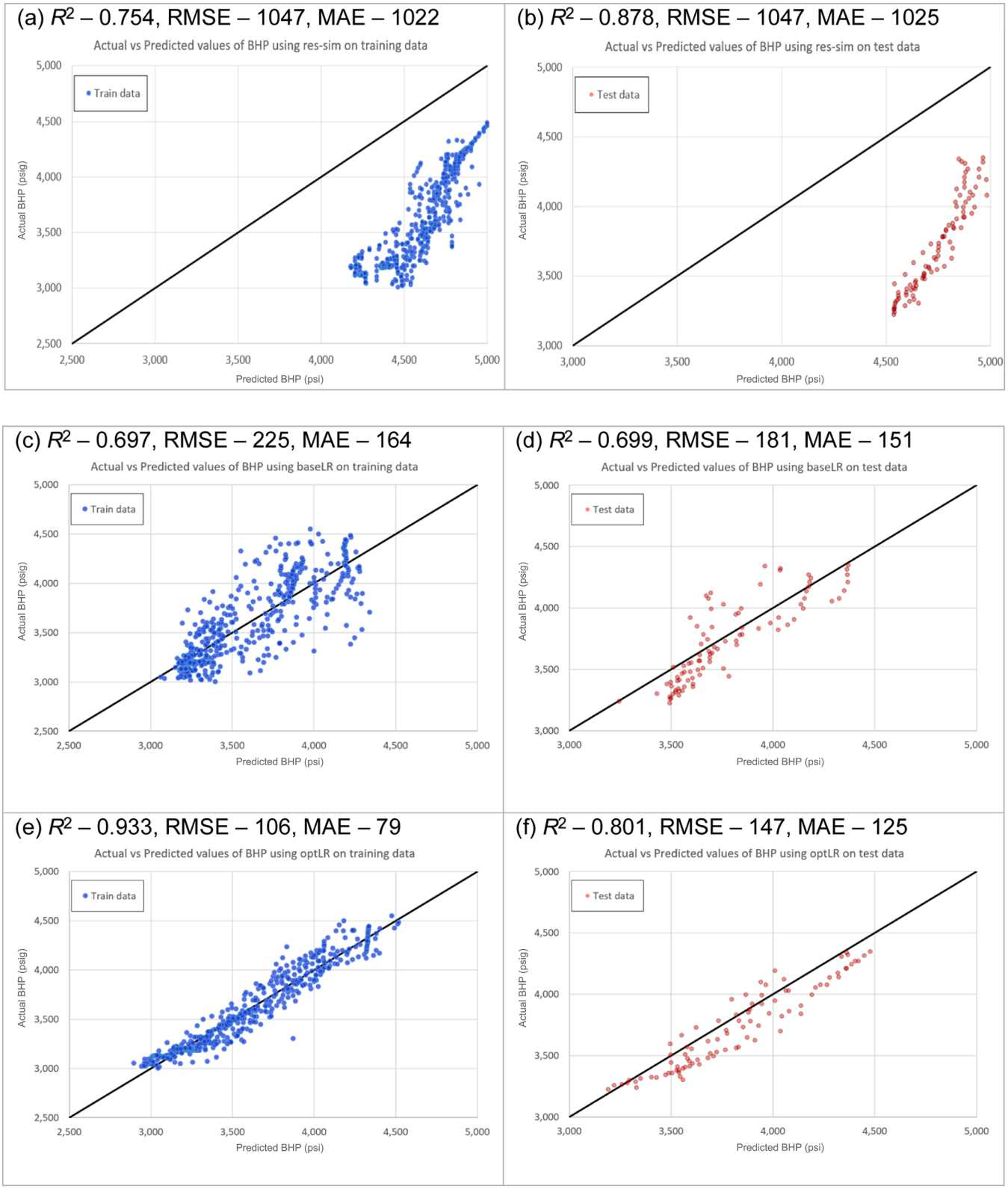


Fig. 8. (a),(b) Scatter plot of train and test data for ResSim; (c),(d) Scatter plot of train and test data for baseLR; (e),(f) Scatter plot of train and test data for optLR; (g),(h) Scatter plot of train and test data for altLR; (i),(j) Scatter plot of train and test data for baseKAN; (k),(l) Scatter plot of train and test data for optKAN.

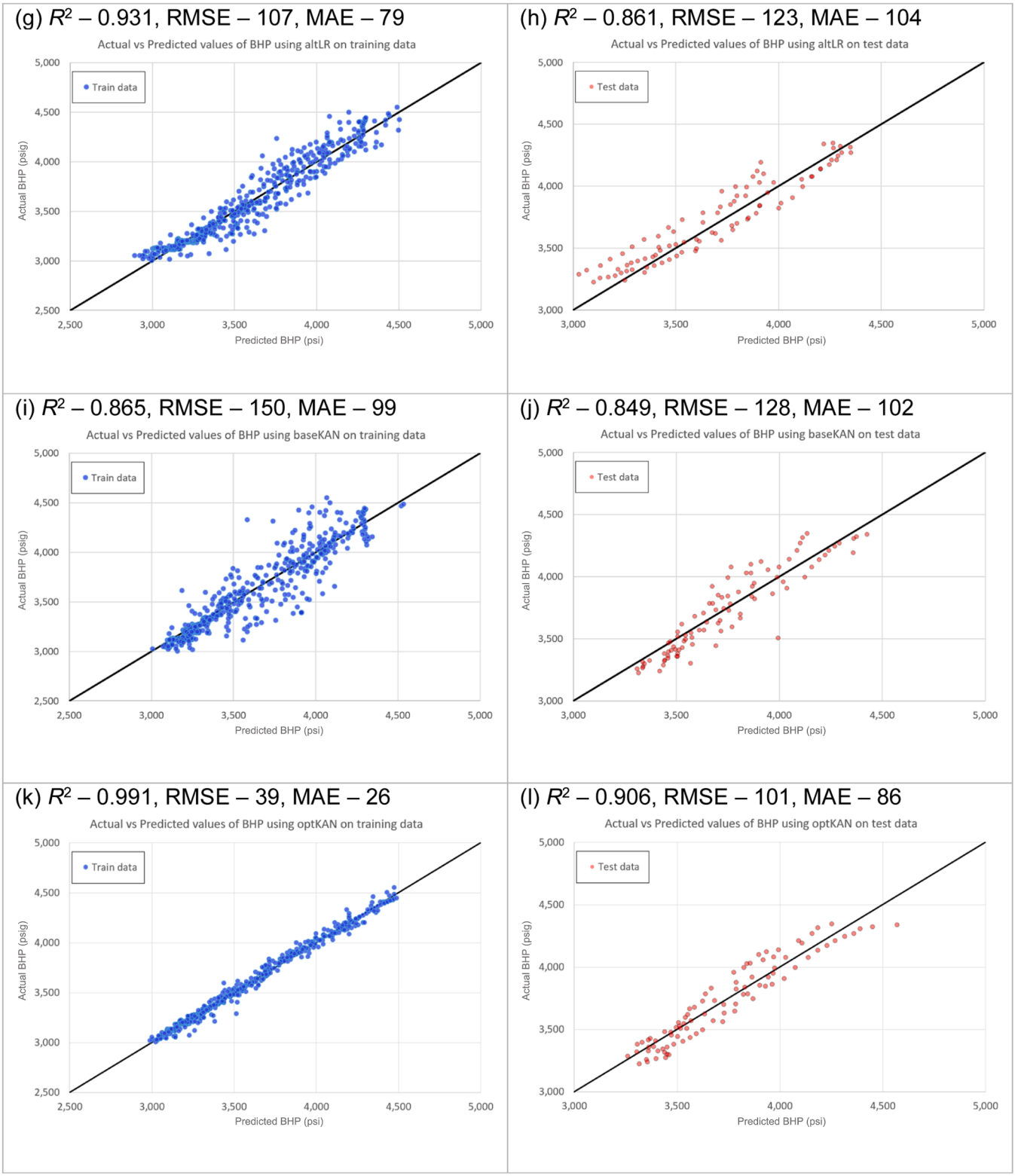


Fig. 8. (*continued*).

capture the trends for the drawdown and buildup sections of the data as seen in Fig. 7 below.

Finally, we show in Fig. 8, the train and test scatter plots for all the models tested as a visual test of the quality of the match. The optKAN model shows more cluster of the points around the 45-degree line for both the train and test data indicating a better representation of the relationships between the input variables and the target variable.

Table 7
Summary of lag features for well oil production rate.

WOPR lag feature	Interpretation
wopr_diff	one timestep lag
wopr_ema_10	10-period exponentially weighted moving average
wopr_rolling_mean_10	10-period moving average
wopr_rolling_pct_change_10	10-period moving average for the percentage change

Table 8
Best model parameters obtained after using 18 features.

Model	Runtime (seconds)	Train			Test		
		R^2	RMSE	MAE	R^2	RMSE	MAE
optKAN	189.73	0.999	12.777	7.799	0.964	62.695	48.030
XGBoost	0.54	0.999	0.984	0.701	0.932	86.019	69.067
GBM	0.84	0.999	12.675	7.556	0.927	89.120	72.926
BiLSTM	53.95	0.938	97.702	71.528	0.908	98.500	77.876

4.3. Further model results and discussion

Models were also built for the XGBoost, GBM and BiLSTM using the same 14 features used for the optLR and optKAN models. The test R^2 results for the tree-based models: XGBoost, GBM were 36 % and 48 % respectively. BiLSTM gave good results with a R^2 of 84 % due to its inherent ability to handle time series data very well. To improve the tree-based model results, some lag features were introduced for the well oil production rate variable. This was based on reservoir engineering knowledge that the historical oil withdrawal rate would be the variable that would likely have the most significant impact on the BHP. Four new lag features were added: wopr_diff, wopr_ema_10, wopr_rolling_mean_10 and wopr_rolling_pct_change_10. The definitions of these lag features are provided in the Table 7: Summary of lag features for well oil production rate below.

These lag features were added to the initial 14 features to make a total of 18 features. The optKAN, XGBoost, GBM and BiLSTM models were then run on these 18 features. The results are shown in Table 8 with KAN once again outperforming all the other models in every metric except model runtime. A major drawback of the KAN model is that it takes much longer computational time.

5. Conclusion

In this study, we have shown the applicability of interpretable machine learning models, particularly KAN to predict BHP using real-world data from the Volve field. By using interpretable forms of machine learning models, it is possible to see the main features affecting the BHP prediction accuracy. This information can be used in a capital-constrained business environment to tailor surveillance budgets to focus on gathering data for the strongly correlated features and use the value of information to drive improved reservoir management practices.

The analytical expressions generated by the linear regression and KAN models can also be easily coded into a spreadsheet for future predictions making them a very handy tool for the reservoir engineer to have. Furthermore, this study has also shown that leveraging machine learning can help reduce the frequency of reservoir simulation history-matches and can be used for quick operational decision-making since machine learning forecasts of BHP up to 3 months in the future showed reasonable accuracy.

In conclusion, reservoir engineers can employ machine learning models, which simplify the complex physics underlying reservoir simulators, to reduce time to decision and test multiple scenarios. The results obtained in this study have shown that the simplest machine learning method: linear regression, can give reasonably accurate predictions when combined with reservoir engineering knowledge. This means that machine learning methods should be considered complementary to reservoir simulation and seen as another tool that the engineer uses to make accurate and quick operational decisions.

This study focused on using only surface measured variables to predict BHP. There is an opportunity to expand the input feature dataset and improve the machine learning models. This can be achieved by incorporating downhole measurements such as bottom-hole temperature and previous values of downhole-measured BHP as lagged features. This will enable the model to better represent the distribution of BHP in the specific well and account for temporal dependencies. This is however outside the scope of this study.

As part of a future study, we plan to extend this work to cover machine learning assisted, multivariable, spatio-temporal predictions of saturation and pressure trends thus fully mimicking the expected output from reservoir simulators.

CRediT authorship contribution statement

Adeboye Adeyinka: Writing – original draft, Software, Formal analysis, Conceptualization. **Opedamola Oriola:** Writing – original draft, Software, Formal analysis, Data curation. **Olusegun Stanley Tomomewo:** Writing – review & editing, Supervision, Methodology.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work the authors used Quillbot AI tool in order to paraphrase some of the text. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors would like to thank Equinor and the Volve license partners for making the Volve field dataset available for scientific research.

Appendix A

Table A.1
Expanded dataset features used to build optLR and altLR linear regression models.

Feature name in expanded dataset	Comments	Candidate wells	Variable Count
on_stream_hours	daily onstream production hours for the well	F1C, F4, F5, F11, F12, F14, F15D	7
choke_size	well choke size	F1C, F5, F11, F12, F14, F15D	6
thp	tubing head pressure	F1C, F5, F11, F12, F14, F15D	6
tht	tubing head temperature	F1C, F5, F11, F12, F14, F15D	6
dp_choke_size	pressure drop across choke	F1C, F5, F11, F12, F14, F15D	6
wopr	well oil production rate	F1C, F5, F11, F12, F14, F15D	6
wgpr	well gas production rate	F1C, F5, F11, F12, F14, F15D	6
wwpr	well water production rate	F1C, F5, F11, F12, F14, F15D	6
wopt	well oil production total	F1C, F5, F11, F12, F14, F15D	6
wgpt	well gas production total	F1C, F5, F11, F12, F14, F15D	6
wwpt	well water production total	F1C, F5, F11, F12, F14, F15D	6
wtime_cum	cumulative onstream time for the well	F1C, F4, F5, F11, F12, F14, F15D	7
casing_pressure	casing pressure	F5, F11, F12, F14, F15D	5
wwir	well water injection rate	F4, F5	2
fopr	field oil production rate	field	1
fgpr	field gas production rate	field	1
fwpr	field water production rate	field	1
fwir	field water injection rate	field	1
ftime	field onstream time	field	1
ftime_cum	cumulative onstream time for the field	field	1
fopt	field oil production total	field	1
fgpt	field gas production total	field	1
fwpt	field water production total	field	1
fwit	field water injection total	field	1
fvrr	field voidage replacement ratio	field	1
fvrr_cum	cumulative field voidage replacement ratio	field	1
TOTAL			93

Table A.2
Dropped features during the data preparation for F1C.

Feature name in original volve dataset	Comments
WELL_BORE_CODE	unique well identifier
NPD_WELL_BORE_CODE	unique well identifier
NPD_FIELD_CODE	field code
NPD_FIELD_NAME	field name
NPD_FACILITY_CODE	facility code
NPD_FACILITY_NAME	facility name
AVG_CHOKE_UOM	choke unit of measurement
FLOW_KIND	producer or injector well identifier
AVG_ANNULUS_PRESS	casing pressure
BORE_WI_VOL	well water injection rate, wwir
WELL_TYPE	producer or injector well identifier
AVG_DOWNHOLE_TEMPERATURE	bht
AVG_DP_TUBING	dp_tubing

Appendix B

KAN model analytical expression where x_1 is on_stream_hours, x_2 is choke_size, x_3 is thp, x_4 is tht, x_5 is dp_choke_size, x_6 is wopr, x_7 is wgpr and x_8 is wwpr. Eq. (B.1):

$$
\begin{aligned}
&BHP \\
&= 21.8862x_1 - 37.6092x_2 + 12.4841x_3 + 0.6638x_4 + 46.3019x_5 + \boxed{99.6409x_8} + 23.1759(0.8652 - x_6)^2 \\
&- 15.1919(1 - 0.7048x_1)^2 - 55.4764(1 - 0.6683x_5)^2 + 20.45(1 - 0.9893x_7)^2 \\
&+ \frac{873.9368e^{0.0485(0.3345-x_3)^2-0.071(1-0.6877x_2)^2-0.1161(-0.6852x_1-1)^2+0.1608\sin(0.7392x_4+3.6421)+0.3809\sin(0.2971x_5-0.867)+0.2199\sin(0.486x_8-7.6168)-0.1448e^{-0.391x_6}}}{(1 - 0.1958x_7)^{0.3263}} \\
&- 2.5421(-0.8809x_1 - 1)^2 + 0.6576(-0.7687x_1 - 1)^2 - 0.1203(-0.6702x_1 - 1)^2 + 43.2403(-0.6193x_1 - 1)^2 \\
&- 3.5581(-x_2 - 0.2459)^2 + 3.582(-0.9566x_6 - 1)^2 + 0.396(-x_7 - 0.6305)^2 + 3.5703(-0.9508x_7 - 1)^2 \\
&- 7.2698(-0.754x_7 - 1)^2 + 1.6433(-0.7203x_8 - 1)^2 - 0.8789e^{1.5623x_4} + \boxed{199.7769log(4.5077x_5 + 9.4294)} \\
&+ 242.6166\sin(0.2x_2 - 9.9994) + 935.8876\sin(0.3152x_3 + 4.2962) - 0.2048\sin(0.4134x_3 - 7.1767) \\
&- 106.0297\sin(0.8094x_3 + 1.5995) - 49.7981\sin(0.8311x_3 + 6.3628) + 85.6297\sin(0.2697x_4 + 0.629) \\
&+ 96.3963\sin(0.6719x_4 - 2.8602) + 111.6923\sin(0.7878x_4 - 2.5905) + 0.119\sin(1.021x_4 - 3.0149) \\
&+ 0.3076\sin(0.366x_5 - 3.1594) - 0.0644\sin(0.5628x_6 - 0.801) + 112.71\sin(0.5716x_6 - 1.0445) \\
&- 0.1184\sin(0.7174x_6 - 3.8795) + 0.0558\sin(0.6722x_7 - 3.9266) - 0.0941\sin(0.7394x_7 + 2.5437) \\
&- 0.2716\sin(0.2841x_8 - 3.6678) - 44.3579\sin(0.4693x_8 - 3.8905) + 0.0393\sin(0.537x_8 - 1.5838) \\
&+ 26.59\sin(0.6614x_8 - 1.8086) \\
&+ 15.3511\sin(-0.2151x_3 - 0.0239x_4 - 0.0708(-0.9456x_1 - 1)^2 + 0.0198(-x_6 - 0.554)^2 \\
&+ 0.0398(-x_7 - 0.4297)^2 + 0.42\sin(0.3812x_2 + 8.1814) + 0.4642\sin(0.495x_5 - 3.98) \\
&+ 0.4025\sin(0.4941x_8 - 0.9015) + 2.585) \\
&+ 4.2486\sin(0.0013x_3 - 0.1317(0.86 - x_5)^2 + 1.3058(-0.3215x_1 - 1)^2 + 0.2798\sin(1.6266x_4 - 4.4047) \\
&+ 0.4794\sin(1.384x_6 - 4.1222) - 0.3862\sin(1.6022x_7 + 5.3272) - 0.4707\sin(0.256x_8 + 0.6001) + 3.4256 \\
&+ 1.0984e^{-0.348x_2}) \\
&+ 576.0547\sin(0.1086(0.7369 - x_2)^2 - 0.1224(1 - 0.4602x_7)^2 + 0.17(-0.6424x_1 - 1)^2 \\
&- 0.3372\sin(0.6002x_3 - 1.756) + 0.2415\sin(0.6854x_4 + 0.3291) - 0.4852\sin(0.4299x_8 + 6.117) - 3.5418 \\
&+ 0.0728e^{-0.5821x_6} + 0.167e^{-1.1685x_3}) + 3543.4593 + 333.5735e^{-0.1426x_6} - 0.259e^{-0.9969x_3}
\end{aligned}
$$

The learnable univariate functions in KAN tie a particular feature to a node. This node is usually a sum of univariate functions from each of the individual features feeding into it. We demonstrate this using Fig. 9 below where the relationship between choke differential pressure (dp_choke_size) and water production rate (wwpr) as they sum into a node is shown. The KAN model approximates a log function for the choke differential pressure (dp_choke_size) and a linear relationship for the water production rate (wwpr). It is important to note that there are several other branches within the KAN network that capture more relationships for these 2 features, but these have just been selected as an example. The log function and linear function are also highlighted using red boxes in Eq. (B.1) above where choke differential pressure (dp_choke_size) is x_5 and water production rate (wwpr) is x_8.

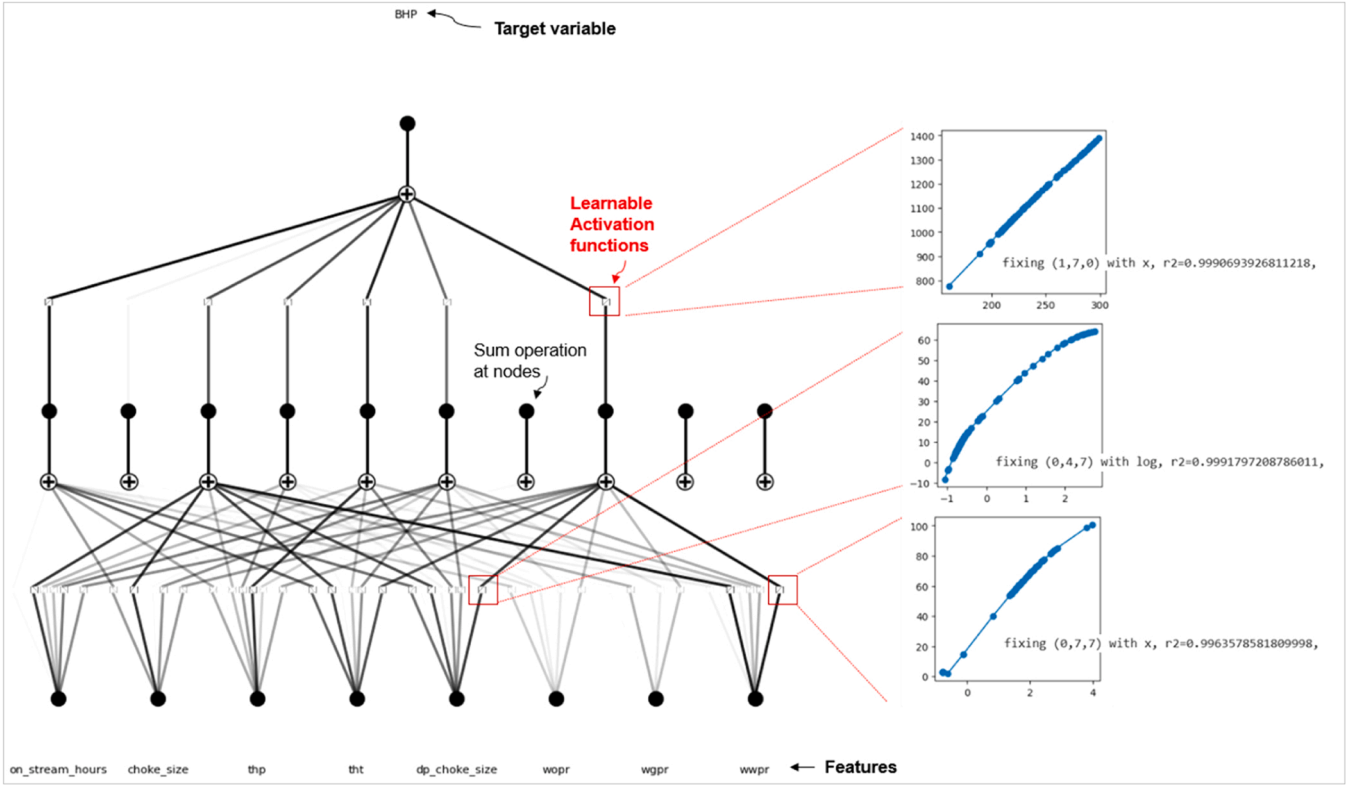


Fig. 9. KAN network plot showing activation functions and nodes. Note that some of the branches appear as faint lines due to their diminished effect on the target variable.

Appendix C

Fig. 10 below shows the SFS scoring curve versus the subset size. The subset size on the chart has been limited to the first 20 features for ease of readability. The chart shows an increase in the R^2 of the train as the feature subset size increased. It is important to point out that the SFS algorithm had access to only the train data. The validation folds were run after the SFS process and the best R^2 values for validation were between 13 and 15 features; beyond this, overfitting occurs and while the train continues to increase slowly, all the R^2 values for the 3 validate folds begin a downward trend. It was based on this chart that the 14 features were selected.

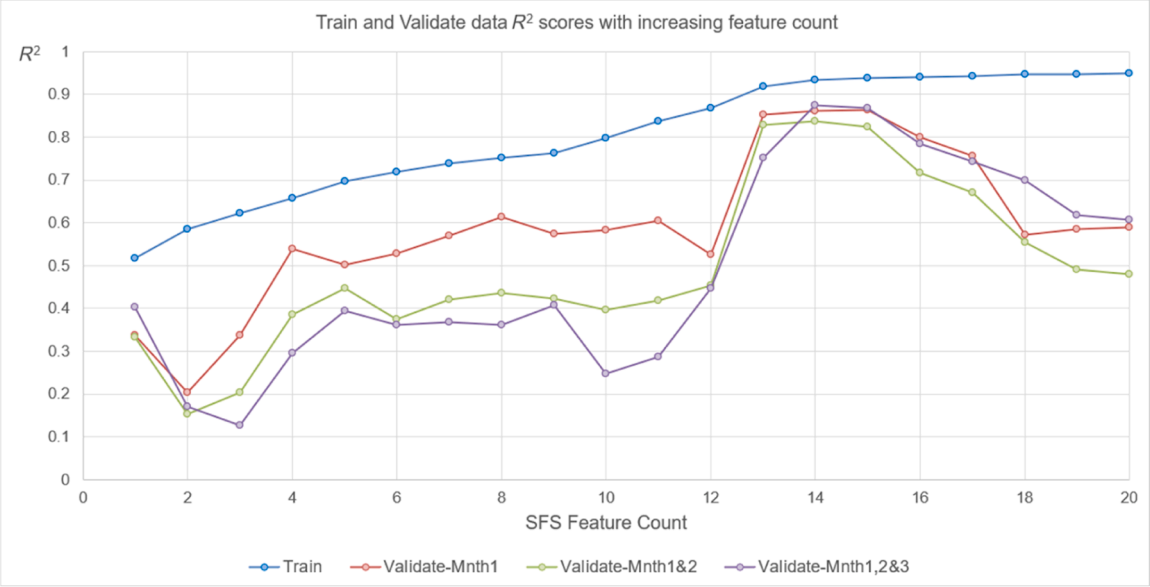


Fig. 10. Plot of train and validate R^2 scores versus feature size for SFS algorithm.

Appendix D

Table D.1 shows rolling origin cross validation results for some of the models used. The results show stability in the validation metrics across the splits used. The average metrics outperform those reported in Table 6 and Table 8. This is because the train sets used here have been expanded to assimilate more data over the three splits leading to a better trained model.

Table D.1
Rolling origin validation results for the models.

Apr-2014 till Sep-2015	Oct-2015	Nov-2015	Dec-2015		OptLR	OptKAN	XGBoost	GBM
Train	Validate			R^2	0.862	0.994	0.923	0.947
				RMSE	149.8	31.0	112.1	92.7
				MAE	130.7	20.7	85.2	72.7
Train		Validate		R^2	0.868	0.963	0.978	0.971
				RMSE	124.4	65.8	51.0	58.1
				MAE	88.9	57.7	35.8	47.5
Train			Validate	R^2	0.922	0.970	0.936	0.937
				RMSE	129.8	81.0	117.5	116.4
				MAE	107.2	69.4	98.5	97.0
	Average validation metrics across splits			Avg. R^2	0.884	0.976	0.946	0.952
				Avg. RMSE	134.7	59.3	93.5	89.1
				Avg. MAE	108.9	49.3	73.1	72.4

Table D.2 and Table D.3 below show results for a monthly look-ahead forecast model. The train data is in a fixed window and not expanded as in the rolling origin approach. The model is built using the 3-month validation data and then used to forecast the monthly validate and test values. The results across the different monthly look-ahead forecasts remain stable.

Table D.2
Month-ahead forecasting for optLR and optKAN.

Apr-2014 till Sep-2015	Oct-2015	Nov-2015	Dec-2015	Jan to Mar-2016		OptLR	OptKAN	XGBoost	GBM
Train	Validate				R^2	0.862	0.990	0.923	0.949
					RMSE	149.8	40.8	112.1	91.2
					MAE	130.7	27.9	85.2	71.5
Train		Validate			R^2	0.800	0.967	0.909	0.912

(continued on next page)

Table D.2 (*continued*)

Apr-2014 till Sep-2015	Oct-2015	Nov-2015	Dec-2015	Jan to Mar-2016		OptLR	OptKAN	XGBoost	GBM
					RMSE	153.6	62.5	103.3	101.8
					MAE	115.7	49.4	77.3	79.4
Train			Validate		R^2	0.920	0.897	0.865	0.877
					RMSE	131.6	149.3	170.8	163.5
					MAE	114.3	114.8	140.4	132.8

Table D.3
Month-ahead forecasting for optLR and optKAN.

Apr-2014 till Sep-2015	Oct till Dec-2015	Jan-2016	Feb-2016	Mar-2016		OptLR	OptKAN	XGBoost	GBM
Train		Test			R^2	0.620	0.914	0.877	0.938
					RMSE	184.9	88.0	105.2	74.6
					MAE	164.9	76.0	93.0	61.5
Train			Test		R^2	0.846	0.969	0.973	0.958
					RMSE	135.5	60.5	56.3	71.2
					MAE	116.2	44.1	45.6	59.3
Train				Test	R^2	0.863	0.957	0.936	0.851
					RMSE	109.9	61.7	75.0	114.5
					MAE	94.7	53.2	62.6	97.7

Data availability

Data used is from a publicly accessible repository cited in the article. The codes used in this work are available upon request.

References

[1] I. Jahanandish, B. Salimifard, H. Jalalifar, Predicting bottomhole pressure in vertical multiphase flowing wells using artificial neural networks, J. Pet. Sci. Eng. 75 (3-4) (2011) 336–342.
[2] Y.D. Kim, L.J. Durlofsky, Convolutional–recurrent neural network proxy for robust optimization and closed-loop reservoir management, Comput. Geosci. 27 (2) (2023) 179–202.
[3] Y. Li, F.O. Alpak, V. Jain, et al., History-matching and forecasting production rate and bottomhole pressure data using an enhanced physics-based data-driven simulator, SPE Reserv. Eval. Eng. 26 (3) (2023) 957–974.
[4] T. Ahmed, P. McKinney, Advanced Reservoir Engineering, Elsevier, 2011.
[5] L.P. Dake, Fundamentals of Reservoir Engineering, Elsevier, 1983.
[6] R.N. Horne, Modern Well Test Analysis, Petroway Inc, 1995.
[7] M. Stanko. Underground Reservoirs: Fluid Production and Injection, Norwegian University of Science and Technology, 2024.
[8] L.P. Dake, The practice of Reservoir Engineering (revised edition)., Elsevier, 2001.
[9] A. Muggeridge, A. Cockin, K. Webb, et al., Recovery rates, enhanced oil recovery and technological limits, Philos. Trans. R. Soc. A Math. Phys. Eng. Sci. 372 (2014) 20120320.
[10] K.D. Scott, W.-C. Chu, R.W. Flumerfelt, Application of real-time bottom-hole pressure to improve field development strategies in the Midland Basin Wolfcamp Shale. Unconventional Resources Technology Conference, Society of Exploration Geophysicists, San Antonio, Texas, 2015.
[11] B. Agbodike, U. Osokogwu, G. Achumba, Solution to limited pressure bhp data in brown fields; material balance equation approach, Proceedings of the SPE Nigeria Annual International Conference and Exhibition, Lagos, Nigeria (2019) 10.2118/198785-MS.
[12] S. Sun, D.A. Pollitt, Optimising development and production of naturally fractured reservoirs using a large empirical dataset, Pet. Geosci. 27 (2) (2021) petgeo2020-079.
[13] S.O. Inikori, Numerical study of water coning control with Downhole Water Sink (DWS) completions in vertical and horizontal wells. PhD Thesis, Louisiana State University and Agricultural and Mechanical College, 2002.
[14] A.N. Okon, D. Appah, J.U. Akpabio, Water coning prediction review and control: Developing an integrated approach, J. Sci. Res. Rep. 14 (4) (2017) 1–24.
[15] A.A. Roozshenas, H. Hematpur, R. Abdollahi, et al., Water production problem in gas reservoirs: concepts, challenges, and practical solutions, Math. Probl. Eng. 2021 (1) (2021) 9075560.
[16] B. Dogah, V. Atashbari, M. Ahmadi, et al., Enhanced oil recovery using CO_2 in Alaska, Geosciences 11 (2) (2021) 98.
[17] B.V. Malozyomov, N.V. Martyushev, V.V. Kukartsev, et al., Overview of methods for enhanced oil recovery from conventional and unconventional reservoirs, Energies 16 (13) (2023) 4907.
[18] W. Ozowe, G.O. Daramola, I.O. Ekemezie, Innovative approaches in enhanced oil recovery: A focus on gas injection synergies with other EOR methods, Magna Sci. Adv. Res. Rev. 11 (1) (2024) 311–324.
[19] A. Acock, T. ORourke, D. Shirmboh, et al., Practical approaches to sand management, Oilfield Rev. 16 (1) (2004) 10–27.
[20] A. Qatari, A workflow to estimate critical drawdown pressure utilizing 1-D geomechanics modelling, SPE Middle East Oil, Gas and Geosciences Show, Society of Petroleum Engineers, Manama, Bahrain (2023) SPE-213325-MS.
[21] X. Yang, K. Qiu, Y. Zhang, et al., Analyzing unexpected sanding issues in the high-pressure/high-temperature, tight-sandstone Keshen gas reservoir, Western China, SPE Drill. Complet. 33 (3) (2018) 192–208.
[22] B.K. Edan, H.A. Hussein, Geomechanics analysis of well drilling instability: A review, J. Eng. 29 (8) (2023) 94–105.
[23] E. Tercan, Managed pressure drilling techniques, equipment & applications. M.S. Thesis, Middle East Technical University, 2010.
[24] J.E. Udegbunam, Improved well design with risk and uncertainty analysis. PhD Thesis, University of Stavanger, 2014.
[25] A.E. Esiri, D.D. Jambol, C. Ozowe, Enhancing reservoir characterization with integrated petrophysical analysis and geostatistical methods, Open Access Res. J. Multidiscip. Stud. 7 (2) (2024) 168–179.
[26] O.I. Fagbemi, A.I. Olayinka, M.A. Oladunjoye, et al., Focused reservoir characterization: Analysis of selected sand units using well log and 3-D seismic data in 'Kukih' field, Onshore Niger Delta, Nigeria, Sci. Rep. 14 (1) (2024) 13763.
[27] C. Yang, C. Zhao, X. Meng, et al., Reservoir characterization and an integrated approach of reservoir modeling for Miano gas field, Middle Indus Basin, Energies 16 (2) (2023) 856.
[28] B. Solanke, F.B. Onita, O.J. Ochulor, et al., Techniques for improved reservoir characterization using advanced geological modeling in the oil and gas industry, Int. J. Appl. Res. Soc. Sci. 6 (9) (2024) 2060–2088.
[29] M. Tang, Y. Liu, L.J. Durlofsky, A deep-learning-based surrogate model for data assimilation in dynamic subsurface flow problems, J. Comput. Phys. 413 (2020) 109456.
[30] J. Wang, Integrated reservoir characterization and simulation studies in stripper oil and gas fields. PhD Thesis, Texas A&M University, 2008.
[31] M.A. Ahmadi, M. Galedarzadeha, S.R. Shadizadeh, Low parameter model to monitor bottom hole pressure invertical multiphaseflow in oil production wells, Petroleum 2 (3) (2016) 258–266.
[32] F.N. Abdulrazzaq, O.F. Hasan, A review of automatic history matching, Mater. Today Proc. 80 (2023) 3817–3822.
[33] E. Barrela, V. Demyanov, L. Azevedo, Coupling geologically consistent geostatistical history matching with parameter uncertainty quantification, arXiv Prepr. (2018) 1810.06642.
[34] C.D. Martin, The strength of massive Lac du Bonnet granite around underground openings. PhD Thesis, University of Manitoba, Winnipeg, Canada, 1993.
[35] A. Samnioti, V. Gaganis, Applications of machine learning in subsurface reservoir simulation—a review—part I, Energies 16 (16) (2023) 6079.
[36] B.T.H. Marbun, D.E. Prasetyo, H. Prabowo, et al., Well integrity evaluation prior to converting a conventional gas well to CO_2 injector well – gundih CCS pilot project in Indonesia (phase 1), Int. J. Greenh. Gas. Control 88 (2019) 447–459, https://doi.org/10.1016/j.ijggc.2019.06.006.
[37] D. Badawi, E. Gildin, Neural operator-based proxy for reservoir simulations considering varying well settings, locations, and permeability fields, Comput. Geosci. 196 (2025) 105826.
[38] E.J. Coutinho, M. Dall'Aqua, E. Gildin, Physics-aware deep-learning-based proxy reservoir simulation model equipped with state and well output prediction, Front. Appl. Math. Stat. 7 (2021) 651178.

[39] A.A. Emerick, Estimation of pressure and saturation fields from time-lapse impedance data using the ensemble smoother, J. Geophys. Eng. 11 (3) (2014) 035007.
[40] K. Elser, K. Mukundakrishnan, V. Natoli, et al., Realizing the potential of GPUs for reservoir simulation. Proceedings of the ECMOR XIV - 4th European Conference on the Mathematics of Oil Recovery, European Association of Geoscientists & Engineers, 2014, pp. 1–16.
[41] T.C. Halsey, Computational sciences in the upstream oil and gas industry, Philos. Trans. R. Soc. A Math. Phys. Eng. Sci. 374 (2078) (2016) 20150429.
[42] F.A. Portella, D. Buchaca, J.R. Rodrigues, et al., A tuning algorithm strategy for reservoir simulation workloads, J. Comput. Sci. 63 (2022) 101811.
[43] M.M. Ahmed, M.A. Ayoub, A comprehensive study on the current pressure drop calculation in multiphase vertical wells; Current trends and future prospective, J. Appl. Sci. 14 (23) (2014) 3162–3171.
[44] G. Takacs, C.G. Guffey. Prediction of flowing bottomhole pressures in gas wells., SPE Gas Technology Symposium, Dallas, Texas, 1989, https://doi.org/10.2118/19107-MS.
[45] O.E. Agwu, S. Alatefi, A. Alkouh, et al., Artificial intelligence models for flowing bottomhole pressure estimation: State-of-the-art and proposed future research directions, Int. J. Adv. Sci. Eng. Inf. Technol. 14 (6) (2024) 1868–1879.
[46] J. Navratil, A. King, J. Rios, et al., Accelerating physics-based simulations using neural network proxies: An application in oil reservoir modeling, Front. big Data 2 (2019) 471450.
[47] O. Sudakov, D. Koroteev, B. Belozerov, et al., Artificial neural network surrogate modeling of oil reservoir: A case study. Advances in Neural Networks–ISNN 2019: 16th International Symposium on Neural Networks, Springer International Publishing, Moscow, Russia, 2019.
[48] M. Jin, H. Emami-Meybodi, M. Ahmadi, Flowing bottomhole pressure during gas lift in unconventional oil wells, SPE J 29 (5) (2024) 2432–2444, https://doi.org/10.2118/214832-PA.
[49] H. Zalavadia, V. Sabharwal, S. Sankaran, Continuous bottomhole pressure estimation using machine learning and physics-based methods – a field case study. International Petroleum Technology Conference, SPE, Dhahran, Saudi Arabia, 2024, https://doi.org/10.2523/IPTC-24528-MS.
[50] H. Zheng, B. Lin, J. Jiang, et al., Knowledge-guided machine learning method for downhole gauge record prediction in deep water gas field, Offshore Technology Conference Asia, SPE, Kuala Lumpur, Malaysia, 2024, https://doi.org/10.4043/34844-MS.
[51] Z. Zhu, X. Song, R. Zhang, et al., A hybrid neural network model for predicting bottomhole pressure in managed pressure drilling, Appl. Sci. 12 (13) (2022) 6728.
[52] R. Ashena, J. Moghadasi, Bottom hole pressure estimation using evolved neural networks by real coded ant colony optimization and genetic algorithm, J. Pet. Sci. Eng. 77 (3-4) (2011) 375–385.
[53] M.A. Ahmadi, Z. Chen, Machine learning models to predict bottom hole pressure in multi-phase flow in vertical oil production wells, Can. J. Chem. Eng. 97 (11) (2019) 2928–2940.
[54] P. Panja, W. Jia, A. Nelson, et al., Convolutional long short-term memory (convLSTM) for spatio-temporal forecastings of saturations and pressure in the SACROC field, Pet. Sci. Technol. (2024) 1–25, https://doi.org/10.48550/arXiv.2212.00796.
[55] D.I. Ignatov, K. Sinkov, P. Spesivtsev, et al., Tree-based ensembles for predicting the bottomhole pressure of oil and gas well flows. Analysis of Images, Social Networks and Texts: 7th International Conference, AIST 2018, Springer International Publishing, Moscow, Russia, 2018.
[56] D.A. Otchere, T.O. Ganat, J.O. Ojero, et al., Application of gradient boosting regression model for the evaluation of feature selection techniques in improving reservoir characterisation predictions, J. Pet. Sci. Eng. 208 (2022) 109244.
[57] E. Eltahan, R. Ganjdanesh, W. Yu, et al., Machine learning approach to improve calculated bottom-hole pressure. Unconventional Resources Technology Conference, Unconventional Resources Technology Conference (URTeC), Houston, Texas, 2021.
[58] C. Tian, R.N. Horne, Applying machine-learning techniques to interpret flow-rate, pressure, and temperature data from permanent downhole gauges, SPE Reserv. Eval. Eng. 22 (2) (2019) 386–401.
[59] S. Mishra, B. Hill, A. Haagsma, et al., Machine learning applications for geologic data integration and operational data analysis in geologic carbon dioxide storage systems. Proceedings of the 16th International Conference on Greenhouse Gas Control Technologies GHGT-16, SSRN, Lyon, France, 2022.
[60] O. Olamigoke, D.C. Onyeali, Machine learning prediction of bottomhole flowing pressure as a time series in the volve field, Int. J. Front. Eng. Technol. Res. 2 (2) (2022) 20–29, https://doi.org/10.53294/ijfetr.2022.2.2.003.
[61] O. Akinsete, B.A. Adesiji, Bottom-hole pressure estimation from wellhead data using artificial neural network. Proceedings of the Nigeria Annual International Conference and Exhibition, SPE, Lagos, Nigeria, 2019.
[62] O. Blessing, I.S. Agbons, Bottom-hole pressure prediction from wellhead data using developed machine learning models, NIPES J. Sci. Technol. Res. 3 (3) (2021).
[63] Y. Li, R. Sun, R. Horne, Deep learning for well data history analysis. SPE Annual Technical Conference and Exhibition, SPE, Calgary, Alberta, Canada, 2019.
[64] Equinor. (2018). Volve data village. Retrieved from Equinor Data Portal: ⟨https://data.equinor.com/dataset/Volve⟩ (Accessed 1st January 2024).
[65] N.M. Ahmad, M. Yaser, M. Abdel-Basset, et al., Sustainable redevelopment of the volve field amidst recent energy security challenges faced by northwestern Europe. A case study from central North Sea, Proceedings of the International Petroleum Technology Conference, IPTC, Dhahran, Saudi Arabia, 2024.
[66] B. Wang, J. Sharma, J. Chen, et al., Ensemble machine learning assisted reservoir characterization using field production data - an offshore field case study, Energies 14 (4) (2021) 1052.
[67] K. Hornik, M. Stinchcombe, H. White, Multilayer feedforward networks are universal approximators, Neural Netw. 2 (5) (1989) 359–366.
[68] G. Cybenko, Approximation by superpositions of a sigmoidal function, Math. Control Signals Syst. 2 (4) (1989) 303–314.
[69] Z. Liu, Y. Wang, S. Vaidya, et al., KAN: Kolmogorov-Arnold, Netw. arXiv Prepr. arXiv 2404 (2024) 19756.
[70] J. Bemister-Buffington, A.J. Wolf, S. Raschka, et al., Machine learning to identify flexibility signatures of class A GPCR inhibition, Biomolecules 10 (3) (2020) 454.

Adeboye Adeyinka is a reservoir engineer with 15 years' experience working in the energy industry. He has over a decade experience with an oil and gas major and his research interests include machine learning applications in reservoir engineering and carbon sequestration. Adeboye is currently a graduate student on the PhD Energy Engineering program at Energy and Petroleum Engineering Department, College of Engineering and Mines, University of North Dakota, ND, USA.

Opedamola Oriola has over 5 years' experience working with an international oil servicing company specializing in wireline data acquisition. Now as a Computer Vision Engineer, he develops targeted AI solutions for energy and telecom companies, helping optimize operations and drive innovation. Beyond his professional work, he has created award-winning AI solutions in global competitions hosted by ThinkOnward, Google, NASA and Fossil. His projects span seismic analysis, kerogen analysis, sea conservation and demand forecasting, demonstrating his ability to apply data-driven solutions to real-world industry challenges

Olusegun Stanley Tomomewo is an Assistant Professor and faculty member in the Department of Energy and Petroleum Engineering at the College of Engineering and Mines, University of North Dakota. He is the graduate director of the Energy Engineering program. His expertise encompasses the design, implementation, and optimization of diverse energy systems, including conventional power plants, oil and gas technologies, and renewable energy sources such as biofuels, solar, and wind power.