SOCAR

SOCAR Proceedings

Reservoir Engineering

journal home page: http://proceedings.socar.az

WEB OF SCIENCE EMERGING SOURCES CITATION INDEX CLARIVATE ANALYTICS

Scopus

APPLICATION OF MACHINE LEARNING TECHNIQUE IN WATER, CHEMICAL, AND THERMAL FLOODING: REVIEW PAPER

Rihab A. Deabl*, Mohammed S. Al-Jawad

College of Engineering, University of Baghdad, Iraq

A B S T R A C T

This article aims to provide literature for the development of new technologies in enhanced oil recovery (EOR), with a focus on reservoir modelling using machine learning (ML). The inaccuracy and inefficiency of traditional physics-based numerical simulations necessitate the development of faster and more intelligent tools. This review serves as a comprehensive resource on applied ML approaches for reservoir modelling. This research classifies the previous artificial intelligence (AI) research in EOR into three methods: water flooding (WF), chemical enhanced oil recovery (CEOR), and thermal flooding. The comprehensive classification is based on the algorithm used, dataset, purpose, inputs, results, and evaluation for each method in each paper. A novel method for simulating dynamic fluid distributions in WF, the Conditional Deep Convolutional Generative Adversarial Network (CDC-GAN) significantly lowers computational costs while handling complex nonlinear relationships. Particle Swarm Optimization (PSO) combined with Bayesian Random Forest (BRF) provides reliable proxy modelling and optimization, also the Echo State Network (ESN) improves prediction accuracy, however it requires high-quality historical data. While Adaptive Neuro-Fuzzy Inference Systems (ANFIS) efficiently handle uncertainties, Least Square Support Vector Machines (LSSVM) and Artificial Neural Networks (ANNs) demonstrate predictive capabilities for nonlinear relationships in the CEOR. This review emphasizes the role of Reinforcement Learning (RL) in thermal, along with the incorporation of Principal Component Analysis (PCA) and clustering techniques for improved data interpretation. Consequently, this study presents a comprehensive analysis of AI techniques in Enhanced Oil Recovery (EOR) from 2009 to 2024, offering researchers and technical experts insights for future investigations.

Keywords: Bayesian random forest; particle swarm optimisation; least square support vector machine; ANNs; reinforcement learning.

Date submitted: 16.07.2025 **Date accepted:** 05.02.2026

Introduction & background

Artificial intelligence has been applied in the petroleum industry for a wide range of tasks, including seismic and log interpretation, precise modelling, optimized drilling operations, well dynamics forecasting, and safety alerts. Machine learning has gained prominence among petroleum engineering experts by transforming large datasets into meaningful insights that facilitate rapid and reliable decision-making. Data-driven modelling methods, which use a lot of data analysis and machine learning to predict system behavior, are a very interesting option. EOR is an important stage in field development planning. It is used when primary and secondary recovery methods are insufficient to extract hydrocarbons from remaining reserves. EOR techniques encompass gas flooding, chemical flooding, and thermal flooding. This method enables oil companies to evaluate the field's economic viability by modifying the system inputs (e.g., number of new production and injection wells, well locations, injection flow rates and, compositions, production flow rates, etc.).

The main goal of EOR is to find and exploit oil and gas resources to ensure energy availability, meet global energy needs, and maximize profits. Reservoir engineers need subsurface reservoir simulation to reach these goals. This better understanding lowers the uncertainty about the reservoir's properties, boundaries, and heterogeneity, which in turn improves the model's quality and reliability. Moreover, the predictive capabilities of machine learning can be employed to enhance production plans. Reservoir engineers typically apply screening criteria to determine suitable EOR technologies for a certain reservoir. The EOR screening criteria delineate the validity intervals for each significant reservoir and fluid properties, based on successful field experiments, technical assessments, and expert evaluations [1]. In the era of big data, artificial intelligence has gained attention as a powerful tool for increasing production and operational efficiency in industrial sectors, and machine learning, espe-

**E-mail:* rehab.abbas2208@coeng.uobaghdad.edu.iq

http://dx.doi.org/10.5510/OGP20260201196

cially deep learning, has advanced significantly [2].

It is estimated that two-thirds of the original oil remains after the primary and secondary recovery phases. The residual oil quantity has garnered the interest of industries and researchers towards the development of novel procedures known as tertiary oil recovery methods [3]. The advancement of artificial intelligence has facilitated the transition of traditional oil and gas fields into the era of intelligent, digital operations, significantly enhancing their intelligence. This change reduces development costs and enhances processes and efficiency. These three EOR methods are now important tools for improving production and research in the field of petroleum engineering. However, the determination of optimal parameters for these methods may require costly tests, production processes, or numerical simulations. Researchers can use ML to predict likely parameters or come up with optimization design strategies by training on sample data.

Artificial intelligence (AI) has become essential in many fields, including the oil and gas industry, where it is applied to drilling, reservoirs, and production. In oil and gas extraction, traditional techniques, such as reservoir simulation, are employed to forecast the oil production rates. This simulation necessitates extensive data, resulting in prolonged and costly process steps. AI is desperately required and may serve as a solution in this instance. AI has emerged as a prominent topic in contemporary applications [4]. The primary benefits of employing AI over traditional approaches include faster processing times and accurate predictions [5-8]. Numerous research studies have illustrated the application of AI in the petroleum sector to enhance operational efficiency [9-11]. An application of AI in the drilling sector was executed by [12]. Various employed AI methodologies include particle swarm optimization-adaptive neuro-fuzzy inference system (PSO-ANFIS), (ANFIS), radial basis function algorithm (RBF), and (LSSVM-GA). These strategies are employed to predict the performance and select the most appropriate model. [13] Employed AI for predicting oil flow rates through the implementation of five machine learning techniques: RBF, ANFIS, Multilayer Perceptron (MLP), LSSVM-GA, and Gene Expression Programming. MLP produced a more precise model for forecasting oil flow rates with 830 training data points.

This study produces a reliable forecasting model with high accuracy [14]. Using input parameters such as wellhead pressure, fluid properties, variable speed drive sensor data, and downhole electric submersible pump sensor readings, four ANN models were developed. The ANN models demonstrated cost-effectiveness, simplicity, and efficiency [15]. The Python programming language was used to build the predictive models for all three approaches. Traditional methods are limited by complex geological conditions and oil field layouts, which leads to lower water injection efficiency and higher costs. ML is described as the study of a computer's capacity to learn from data [16]. However, due to complex geological conditions and oil field setups, traditional methods often fail to achieve the best results, leading to lower injection efficiency, higher costs, and longer production times [17-19].

A pertinent literature evaluation indicates that an optimization methodology for ANN design can be employed to effectively identify the most suitable neural network topology (including the number of hidden layers, hidden neurons, and transfer functions) [20]. This trend is due to most oilfields worldwide entering the decline phase. Consequently, the postponement of reservoir abandonment has become a subject of close attention for scholars globally. Their study typically emphasizes the critical need for the development of novel techniques, commonly classified as tertiary oil recovery systems, which possess the capability to sustain economic production rates. The CEOR process considers both macroscopic and microscopic sweep efficiency. Chemical agents are used in the first case, while polymers are used to increase the mobility ratio by increasing the shear viscosity of water in the second.

CEOR is implemented in reservoirs as a tertiary approach when waterflooding has attained its recovery efficiency threshold, estimated to be around 30–50 % of the original oil in place (OOIP). This approach effectively recovers residual oil entrapped by capillary forces during waterflooding. Chemical substances are introduced into the reservoirs to modify the wettability of the fluids, enhance sweep efficiency, and reduce the interfacial tension between water and oil. Polymer flooding is a promising approach to CEOR that functions as a thickening agent to augment the viscosity of injected water. The rheological properties of polymers can enhance oil recovery compared to traditional water injection methods. Polymer injection is a well-established EOR technology, supported by laboratory experiments and practical field operations for over 50 years [21]. Surfactants are special chemicals that help reduce the tension between oil and water, which leads to a higher capillary number, directly linked to the removal of leftover oil after water flooding. Al-Dousari reported a strong correlation between chemistry and an expert system that employed ANNs to forecast the oil recovery factor (RF). Ahmadi investigated surfactant flooding by concentrating on the optimization of the ANN architecture to improve the prediction of the RF and project cost [3]. Thermal techniques are considered the most common methods for enhanced oil recovery, which include cyclic steam injection (CSI), steam flooding, in-situ combustion, and steam-assisted gravity drainage (SAGD). The primary aim of applying heat to a heavy-oil reservoirs is to reduce the fluid viscosity. The inaugural application of CSI occurred in the late 1950s. The rapid payout and the feasibility of injection and production from a single well-rendered field operation were more favourable than research developments in earlier periods. Thermal recovery processes seek to reduce oil viscosity by injecting heat into the reservoir [22].

Steam flooding is the oldest and most effective commercial EOR method, having been in use since the 1960s. It is acknowledged as one of the most efficient techniques for extracting oil from various reservoir types, as it yields a higher ultimate oil recovery compared to other EOR methods [23-26]. Recently, AI has emerged as a prominent subject, with an increasing number of AI methodologies being employed in the oil business for sophisticated data analysis. Steam flooding has been quite effective in shallow heavy oil fields. The steam-assisted gravity drainage (SAGD) is the most appropriate thermal recovery technology, demonstrating superior performance relative to other methods in heavy oil and bituminous sandstone reservoirs [27]. Approximately 68–75 % of the oil can be recovered in situ using the SAGD process, which uses two parallel horizontal wells, usually one drilled 5 meters below the other, with the upper well used for steam circulation in the reservoir and the lower well used for oil extraction [28]. SAGD is mainly recommended for thick, shallow reservoirs containing high-viscosity oil and more than 15 meters deep [29].

The success of SAGD depends on various factors, including reservoir fluidflow characteristics, oil saturation, differences in the reservoir's structure, and the conditions of steam injection. Directional fluid mobility, the amount of oil present, reservoir variations, the presence of water and gas zones, and the rate, pressure, and quality of steam injected have a significant impact on the effectiveness of SAGD. Shin [30] identified reservoir thickness, permeability, oil saturation, and porosity as four critical parameters for predicting SAGD performance and assessing the economics of a SAGD project. Also, steam huff and puff injection, or cyclic steam stimulation, is a thermal process that involves the cyclical injection of steam, alternating with oil extraction. The issue of cost and time inefficiency in reservoir simulation continues to affect the design of a steam huff and puff injection method. This approach accounted for almost 40% of overall EOR generation in 2015 [31]. The thermal EOR technology operates by decreasing oil viscosity through elevated reservoir temperatures. Following the soaking period, the well is reopened and oil production is enhanced. Since a large portion of the steam condenses during the soaking period, the well produces both heated water and heated oil. Previous applications of this strategy have been documented in Cold Lake Field, in Canada [32, 33], Midway Sunset Field,in California, Duri Field,in Indonesia [34-36]; and Tia Juana Field, in Venezuela [37, 38]. Due to the requirement for continuous surface steam generation, SAGD is significantly limited by the need for large quantities of steam, particularly in thin, low-quality reservoirs [39]. AI's use in the oil industry can be divided into two categories: pattern recognition and efficiency/parameter forecasting [40- 43].

Steam injection is expensive, with steam comprising over fifty percent of operational costs; thus, process optimization is essential to enhance the net present value (NPV). Researchers have employed several decision variables and objective functions throughout the years to optimize steam allocation for bitumen and heavy oil reservoirs. In oil fields, accurately forecasting the thermal characteristics along wellbores is a primary responsibility for practicing engineers. Viscous oil is a matter of global significance, as conventional oil reserves are being exhausted [44]. Steam flooding is a thermal recovery technique that involves the continual injection of hot steam into the reservoir. The distribution of steam in the reservoir affects the outcome of a steam flooding project. The development scenario, reservoir thermal characteristics, residual oil saturation, reservoir heterogeneity, oil viscosity, and well spacing are some of the many factors that affect a thermal recovery project's effectiveness.

This paper explores the application of ML to screen for EOR in reservoir development projects. It highlights the challenges with ML, such as making misinterpreting about EOR screening as a classification problem, lack of data for learning that can be used in other situations, and lack of input features.

Application of machine learning to water flooding

Waterflooding is extensively implemented globally as the secondary recovery method to enhance reservoir efficiency. Investigating the application of AI in water injection development within oil fields has considerable practical ramifications. Waterflooding, is a process involving the injection of water into a reservoir to augment pressure and enhance sweep efficiency, is widely employed as a secondary oil recovery method to increase oil production. Historically, the favourite method for estimating the future performance of a waterflood project is numerical reservoir simulation. Nonetheless, despite the considerable strength and insight offered by simulation models, this potential is occasionally unattainable due to substantial processing demands. Recently, certain studies employed machine learning to deduce the injector/producer relationship in waterflooding reservoirs, using the injection rate as input and the production rate as output [45, 46].

This study involves the design and implementation of a proxy model using a conditional deep convolutional generative adversarial network (CDC-GAN) to efficiently compute the dynamic fluid distribution in a heterogeneous reservoir during waterflooding. Zero-sum game theory underpins the CDC-GAN, which comprises a duo of generative discriminative models. This study involves the development of a CDC-GAN proxy model designed to reliably forecast fluid saturation. One portion constituted the training dataset for the proxy model, while the other portion served as the testing dataset to assess the proxy model's performance. This method demonstrates that CDC-GAN can acquire the nonlinear mapping function between input and output spaces [47].

In waterflood projects, important decisions are made to maximise the net present value (NPV) for a given period, maintain reservoir pressure, minimise water production. In this study, a new proxy model presented based on machine learning for predicting waterflood performance and use it for optimising production to get the best future well control. This data-driven model is implemented using an ESN within the «Reservoir Computing» (RC) paradigm. In contrast to conventional ESNs, this approach incorporates a feedback loop from the output into the high-dimensional «reservoir.» The Teacher Forcing (TF) technique facilitates the incorporation of the feedback information without adding extra recurrent loops to the training process. All computations herein possess analytical answers that ensure much enhanced speed, thereby reducing the processing requirements during the optimisation process. This method is more appropriate for mature fields, as dependable production data following breakthrough from each producer can significantly enhance the training process. It can also be applied in situations where no reservoir model has been developed. The linear output link significantly reduces computing burden and enhances simplicity compared to most other proxies [48].

This study intends to apply AI methodologies to predict oil production rates based on water injection rates. Three wells are interconnected using a direct line drive pattern. Three distinct AI methodologies were employed, namely multiple linear polynomial regression (PR), multiple linear regression (MLR), and ANN, to develop models for predicting oil production rates. The dataset has been randomly divided into 80% for training and 20% for testing subgroups. The training data is used to construct prediction models, whilst the testing data serves to assess model performance. The comparative analysis identifies the model exhibiting the minimal root mean square error (RMSE) and the maximal test value. The training data is utilised to construct prediction models, whilst the testing data serves to assess model performance. The MLR and PR techniques serve as efficient and expedient statistical prediction methodologies. MLR can create a linear correlation between response variables and predictors. Concurrently, PR is used as a predictive technique for data exhibiting nonlinear relationships [49].

The Bayesian Random Forest (BRF) approach is used as the model's training algorithm. The model is assessed by computing the mean square error and coefficient of determination, drawing an error distribution histogram, and creating a cross-plot of simulation data versus verification data. The production forecast model for waterflooding reservoirs is constructed by training an architecture of neural networks with characteristics specifically targeted to enhance performance. A multitude of features are developed, and via a brute force methodology, those that do not substantially enhance the model's performance are omitted in the final iteration of the model. This study used 14 years of production data to train, validate, and test an artificial neural network. The data was split into four parts: 80% for training, 5% for testing, 5% for validation, and 10% for a blind evaluation of the model. The model presented in this research exhibits the lowest average absolute percent error of 14.732%, compared to the empirical models of Standing and Vazquez-Beggs. It can yield effective generalization for noisy datasets, such as the oil, gas, and water production rates used in this work [50].

The data assimilation technology is initially employed to align geological model variables with observed well dynamics. The flow relationships between injectors and producers within the block are determined using an automated identification method for layered injection-production flow relationships in mature fields. Finally, the production projections and optimization of the precise water injection plan are accomplished using the PSO algorithm. Case study data shows an 8.2% increase in cumulative oil production 12 months after optimization compared to non-optimized production [51]. By merging the BRF and PSO, an ensemble proxy-model-assisted optimization framework for the injection production system is constructed. The BRF model outperforms competing machine learning models when used as a stand-in for the injection manufacturing system. On-site adjustment of the injection mode is possible using the Pareto frontier. Selecting the particle injection mode that corresponds to the maximum oil rate is an option when developing a reservoir with high oil output as a goal. When producers encounter a high gas oil ratio (GOR), they have the option to choose the injection mode that minimises GOR [52].

Predicting future oil output from an established onshore field that is being injected with both water and steam at the same time is the issue at hand. However, reliance solely on conventional wisdom may be insufficient. Recurrent neural networks (RNNs) and linear regression (LR) are the two ML algorithms used. Both methods can produce accurate production projections. Characteristics derived from material-balance analysis are (i) total injection rate and (ii) number of producing wells. Recurrent ANNs can explain previous data due to loops. The outputs of earlier steps are used as inputs in the current phase. Currently, the data-driven methodologies are utilized to analyze oil reservoirs subjected to gas-flood and water-flood scenarios [53].

This paper provides a brief overview of intelligent field deployment for mature reservoirs, involving the recent developments in data foundation creation, research on applied AI algorithms, potential use cases, and the complete deployment process. International and domestic oil and gas companies are in the process of starting to build and distribute lake data on a massive scale [54].

The previous papers collectively demonstrate the role of AI in optimizing waterflooding processes. Numerical reservoir simulations, though powerful, face scalability issues, prompting the adoption of AI-driven proxies like CDC-GANs and ESN for efficient fluid distribution and production forecasting. Statistical methods (MLR, PR) provide simplicity but lack robustness for nonlinear relationships. Techniques like, BRF and PSO enhance model accuracy and decision-making. The production rate, which acts as a critical performance indicator, and the injection rate, and has a direct impact on reservoir pressure, are the main variables influencing AI approaches for forecasting oil production rates. Geological features and well configuration are also important factors influencing fluid dynamics and recovery efficiency. More accurate predictions are ensured through improved model training and validation using historical production data. Model performance is improved by incorporating different features. Important chemicals that have been reported for EOR applications in different oil fields are listed in table 1. Different studies reported in table 1 represent the ML transitions into the EOR process by WF and the development of EOR methods to improve the oil recovery in the last decades. The table summarizes the method of ML and optimization algorithms for water flooding for green and mature carbonate fields, which outlines the input variables and their impacts on output, the objective of the study, and the result, evaluation & limitation.

Categorization of machine learning algorithms for optimizing the water flooding method

The literature reviewed above for WF presents a classification of machine learning methods. The CDC-GAN proxy model is the first method for modelling dynamic fluid distribution in heterogeneous reservoirs. It uses a cutting-edge AI method to predict water and oil saturation distributions, enabling optimization and uncertainty analysis. This method significantly reduces computational costs compared to traditional numerical reservoir simulations and handles complex nonlinear relationships effectively. BRF with PSO is the second method, combining two powerful techniques: BRF for accurate proxy modelling and PSO for optimization. It achieves high prediction accuracy with reduced training time and has been proven to optimize injection strategies and improve cumulative oil production. However, its complexity and reliance on specific reservoir conditions make it slightly less versatile than CDC-GAN. ESN with TF is the third method, implementing a reservoir computing paradigm with feedback loops for enhanced prediction accuracy. It requires less computational effort compared to traditional numerical models or deep learning and is suitable for mature fields with extensive production data. However, its performance heavily depends on the quality and history of training data and has limited applicability to reservoirs lacking sufficient

Table 1

ML on water flooding (compiled from literature)

Author(s)	Method	Dataset	Goal	Inputs	Result	Evaluation	limitation
Diyah Rosiani [49]	MLR, PR, and ANN	1180 field data included water injection rates from two wells and oil production history from one well (South Sumatra, Indonesia)	This data enhances predictive models, and oil production accuracy could be improved by incorporating additional important variables alongside the water injection rate	(Water injection rate from well 1 &well 2) and the rate of oil production is the only output variable	In comparison to the PR and MLR approaches, the ANN obtains the best test value of 16.2% and the smallest (RMSE) of 0.142	One quick and easy way to estimate oil production rates is the ANN model	–
JIA Deli [51]	PSO	Complex faulted reservoir in eastern China	Using ML to judge how to adjust water injection	The porosity and permeability multipliers for each layer, the Kv/Kh ratio, and the initial capillary pressure at the oil-water contact are considered	Integration of petroleum reservoir & production engineering, water injection intelligent optimization, mature fields adjustment	Determine the efficacy of water injection in various well layers and to carry out the modification of the water injection direction	Considerable investment in time, reliance on the expertise of reservoir engineers, and constrained optimization strategies
Shu-Yi Du [52]	Proxy-model-combining BRF with PSO	Gather the dynamic data of injectors and producers from the reservoir, informed by actual geological condition, production measures, and theoretical frameworks	Demonstrating enhanced accuracy in predicting dynamic parameters	Thickness of the reservoir, Initial reservoir pressure, initial bubble point pressure, average porosity, permeability range. Oil volume factor, oil compressibility	The proxy model showed superior prediction accuracy relative to deep learning, taking less training time, while simultaneously reducing the GOR and enhancing oil output by over 10% in carbonate formations	The contribution rates of the production measures & injection parameters are 60% & 40%, respectively. The ideal injection pattern is determined, and the operational strategy is adjusted in conjunction with Pareto front.	Choke size and injection pressure are determined based on the site development conditions
NEGASH Berihun [50]	ANN & BRF	Reservoir in Malay basin	Forecast oil, gas, and water production rates for a water-flooded reservoir	Pressure & Temperature in the tubing head, gas lift rate, casing pressure, Production manifold pressure, Water injection rate, and Water injection pressure	Forecasting methods for production, including numerical reservoir modelling, do not consider the past injection rate as a factor affecting the present production rate	The prediction model achieves a factor of determination exceeding 0.9, which demonstrates consistency between the simulation results and practical data in predicting production rates	The local minimum is determined through the interpolation of the Gauss-Newton algorithm (GNA) and the gradient descent method
Zhi Zhong [47]	Proxy model using CDC-GAN	Use a numerical reservoir simulator (CMG IMEX) to produce the training and testing dataset from multiple reservoirs	Fast compute the dynamic distribution of fluid within a heterogeneous reservoir during waterflooding and (Sw) prediction	The reservoir properties, specifically permeability distribution, and forecast time are considered as inputs, while water saturation is the intended output	The generative model learns the input-output relationship and generates output similar to training data, while the discriminative model efficiently distinguishes fake output from real data	The estimated model parameters from history matching enhance reservoir characterisation and enable water flooding optimization and analysis with minimal computational cost	Production rate & the number of productions wells
Lichi Deng [48]	ESN under the paradigm of «Reservoir Computing» with TF technique	Mature fields	Introduce a machine-learning proxy model for waterflood prediction and use it for production optimization to optimize future well location control and learn from past production history	Tuned some parameters: dynamic «reservoir» size, leaky integrator α, spectral radius of the recurrent weight matrix $\rho(W)$ & output weight matrix	Optimize (NPV), minimize water production, and ESN proxy model predicts future waterflood performance	Optimisation lowered the ultimate field water cut by 1.54%, indicating that the optimised scenario balanced water displacement & sweep efficiency.	Data quality, sites, completion design, and capacity. Training won't be on green fields; the trained model is confined to a similar well.
Leonardo Kubota [53]	LR and RNNs	50-year production history and 2000+ wells.	The production forecasts may be established without a geological model, fluid and rock parameters, or porous media fluid flow knowledge	Injection history, production history and number of producers	They proved that both techniques (LR and RNN) forecast well	Performing reservoir engineers have two choices (LR, RNN).	
Deli Jia [54]	Surrogate model	China reservoir	extensive AI research, particularly smart deployment for mature waterflood reservoirs	Geological & reservoir properties from previous research	Create a «whole data» framework for reservoir research, including exploration and development	Assess the impact of multi-well stratified water injection, analyze its direction, and establish the final plan using a big data intelligent optimization system	Re and adjust inability to sustain high water injection rates over time

production history. ANN is the fourth method, proven to outperform simpler regression methods with lower RMSE and higher test values. It is well-suited for noisy datasets and complex relationships between injection and production rates. However, it is computationally intensive and requires significant effort in feature selection and dataset preparation. RNNs are the fifth method but are computationally expensive and require fine-tuning and a significant amount of training data. PR and MLR are near the bottom due to their simplicity and computational efficiency but poor performance in handling complex, nonlinear relationships in real-world waterflooding scenarios.

Application of machine learning to chemical flooding

CEOR reduces the interfacial tension between the aqueous and oil phases or increases the aqueous phase's viscosity, enhancing the mobility ratio and sweep efficiency. Polymer flooding, with its low injection cost and oil production increases, is one of the most widely used chemical EOR procedures. Polymers are introduced into the supply to decrease water mobility and increase clear efficiency. Polymer flooding the adjoint approach can calculate the NPV gradient to determine optimal control variables, including injection and production rates, polymer concentrations, and polymer grading. Conventional reservoir simulators are computationally intensive, especially for complex fields, but ML methods can predict the RF and NPV, two crucial parameters of chemical flooding operations. This thinking about employments manufactured fabricating histories from a high-fidelity numerical reenactment show to form ANN proxies.

The essential factors distinguished as affecting chemical adsorption on rocks are concentration, salinity, temperature, and PH. The purpose of this study is to summarize research on surfactant adsorption on various reservoir rocks and under various situations, as well as the effects of surfactant concentration, salinity, temperature, and pH. It is possible that surfactant concentrations will decrease due to adsorption on rock surfaces, which could impact their efficiency and efficacy in practical EOR applications. The primary role of surfactants is to significantly lower the oil-water-rock system's interfacial tension (IFT) to an extremely low value. To keep the whole process from failing, it is essential to reduce the quantity of surfactant adsorption. The adsorption capacity grows as the concentration of the surfactant rises. Beyond the critical molecular concentration (CMC), the adsorption behaviour remains unaffected by concentration increases. This describes the interaction between salt ions and surfactants. Molecules show that saltiness changes how surfactants stick to surfaces. Because the repulsive interactions between adsorbed molecules reduce as the salinity of reservoir brine increases, surfactants are more effectively adsorbed on rock surfaces. As the temperature rises, the kinetic energy of the species increases, which in turn causes a significant drop in surfactant adsorption. The amount of surfactant adsorption varies at different pH levels as the surfactant's charge interacts with other things. Rephrase Surface charges show how much pH affects the way surfactants stick to surfaces [55].

The proxies consider the fluid and reservoir rock properties as well as the project design criteria. In this study, the outcomes of comprehensive blind tests that were conducted to validate the proxy models to get the NPV of the polymer. For injection projects, a hybrid approach combining expert systems with PSO method was proposed. The expert system enables engineers to model various optimization scenarios by adjusting economic parameters, including capital cost, oil price, and injectant cost. ANN model reduced the complexity of the simulation model 200 times faster than commercial simulators and obtained a 90% correlation for test samples with a mean absolute deviation < 10% [56]. The primary objective of this research is to identify effective chemical flooding strategies for petroleum reservoirs by using a novel AI approach known as «LSSVM» to make predictions. Applying LSSVM to the existing literature databases allowed to more accurately forecast the RF and chemical flooding costs, enabling targets to be achieved. Chemical flooding operations can benefit greatly from the suggested model's predictive capabilities [3].

To forecast the RF and NPV for an SP chemical flooding process, [57] built a multilayer perceptron (MLP) model that considers variables such as the size and concentration of surfactant slugs, the concentration and drive size of polymers, the salinity of the polymer drives, and more. For more accurate predictions, the network's design was improved. The model has been shown to be very accurate. The results show that using MLP-based chemical flooding is dependable and can be calculated easily and affordable. Finally, [58] used API gravity, viscosity, porosity, permeability, chemical flooding performance, and other parameters to train ANN models with Genetic Algorithms (GAs). There is no mention of this novel methodology in the current literature; instead, the ANN model architecture is fine-tuned using a GA-fold cross-validation hybrid. The RF may be accurately predicted by the models that were developed. Traditional methods, including policy iteration and backward induction, become computationally difficult and intractable when the state space gets massive and complicated. The alternative approach known as approximate dynamic programming (ADP) was employed to address these concerns while reducing the number of computational resources needed. ADP is a strong method that can tackle difficult, big problems and find a very beneficial solution. The approximate dynamic approach that is used for improving the decision's timing and value by considering both the information available prior and following the decision, as well as any future information that could be useful for making decisions, results in significant improvements in economy. The Tree-Based Pipeline Optimisation Tool (TPOT) is an Auto ML method that's employed in much research. Olson and Moore (2019) were the first to suggest TPOT. A stochastic search approach, like genetic programming, is used by TPOT to optimize ML pipelines. The decision will influence the learning process over time, and this study provides valuable insights into the reservoir development plan [59].

In this study to assess the output EOR performance, fuzzy neural networks like ANFIS were used, together with MLPs. The proposed models predict data with high accuracy (R^2=0.9990 and RMSE=0.0002 for MLP, R^2=0.9973 and RMSE=0.0008 for RBF, and R^2=0.9729 and RMSE=0.0150 for ANFIS neural network). These statistics demonstrate the MLP neural network's exceptional predictive performance. The parameters most affecting polymer flooding

EOR performance are API gravity, salinity, permeability, porosity, and salt concentration. The data records that made up the collected data sets were split evenly between training (70%), validation (15%), and testing (15%) data. In this paper, tangent sigmoid (tansig), log-sigmoid (log-sig), and the linear transfer function (purelin) are all activation functions employed in MLP. Tansig is utilized as the connection between the input and hidden layers [60]. In this study, a tool was developed to predict how well a polymer will work in flooding using a data-driven model based on ANN. Data-driven models are created to assist in decision making on a specific oil reservoir. To develop a predictive tool for polymer injection, two main steps were used: building a numerical model of the reservoir and developing an ANN. The numerical reservoir model is made using a black-oil reservoir simulator, specifically Computer Modelling Group (CMG). The main goal is to demonstrate the effectiveness of data-driven models compared to traditional numerical simulators, specifically for polymer injection in the Daqing Oil Field, the largest polymer-flooding field in the world. Using prolonged polymer flooding helps keep the water-to-oil ratio lower. Nevertheless, the RF by itself cannot determine if polymer flooding is superior to alternative injection techniques. Polymer injection can greatly lower the volume of water injected and produced. Data-driven technologies can offer a rapid and low-cost method of information evaluation and analysis. The suggested method (forward and inverse models) can help determine the best way to inject chemicals to meet expected production goals [61].

They trained a reverse ANN using information about the area's features, starting conditions, permeability levels, and oil production data to predict the design details and how water production and pressure change over time. The forward model showed much better results. In the second case, the author performed simulations to create a training dataset. This dataset included similar details about space, starting conditions, fluid qualities, and permeability rates. It also covered properties of polymer, like how much it sticks to surfaces, salt levels, and thickness. The results focused on water production, oil production, and pressure at the injection well. The writer created the same models as before. The forward model showed much better results [62].

In this study, 1100 training instances were generated for ANN to cover reservoir properties and chemical kinds in a five-spot pattern. The design details were connected to the reservoir features through various methods to improve the ANN models and boost their performance. In this study, five ANN models for a sandstone reservoir were created and how the reservoir responded was measured using Computer Modelling Group-Specialized Treatment for Arthroplasty Recovery and Stability (CMG-STARS). One forward model (Model #1) and four inverse models using back-propagation techniques. Model #1 predicts the reservoir's behaviour, including oil production, water content, pressure at the injection point, total oil produced, and project details (size of the pattern and the amount and strength of chemicals used). Model #2 finds out the reservoir features by comparing them with previous results from the reservoir. Model #3 predicts project details using known reservoir behaviour and features. Models #4 and #5 help predict important details for oil projects based on how much oil is needed and how long the project will take. Improve the performance of the ANN models by changing the structure and using more training data. These adjustment made the errors smaller, and the new models also worked much faster, taking about 10000 times less time to compute. The results indicated that the suggested models can accurately recreate the results from the simulator and do so much faster [63].

Jiang created a complex neural network to predict how much oil will be produced and the amount of water that comes out during a special flooding process using a polymer. They found that chaotic neural networks can make predictions about nonlinear time series more accurate, and this has promising potential for future use [64, 65]. In this study, an ML method was used to create a better way to choose polymer gels for use in injection wells. Logistic regression models were trained using previous data from four real-time gel systems: bulk, high-temperature, colloidal dispersion, and weak gel. Three methods were used: univariate entropy R2, stepwise regression, and the area under the ROC curve (AUC). The results indicate that logistic classification models and similar methods can accurately predict the right gel technology in over 85% of the projects in both the training and validation samples. The three models used in this study are referred to as G2, G3, and G4. The findings for the three models highlight the following points: Five parameters (rock space, oil layer thickness, oil thickness, water content, and oil recovery) have a significant role. Logistic regression models are effective in the G4 and G3 models, but they don't predict well in the G2 model. Five important factors were identified: porosity, net thickness, oil viscosity, water cut, and RF. In the end, a logistic regression method was found to be suitable for EOR screening, despite complicated data patterns and several variables [66].

While prior laboratory studies primarily investigated the influence of ultra-low oil–water interfacial tension on oil recovery efficiency, this paper examines the impact of active surfactant polymer (ASP) concentration for different rock types. Over fifty flood tests using fake models were carried out. These models were created with different levels of how easily fluids can pass through them. Sixteen ASP solutions were made and used in tests. The ASP flood tests looked at how the amount of NaOH and the balance of reducing surface tension and adding oil affected the recovery. This crucial viscosity of displacement should be a key factor in enhancing the chemical composition of an ASP flood for a particular reservoir [67]. In most of those predictions, only one type of plastic was used to achieve the study's goal. This paper aims to suggest an effective tool for predicting recovery rates at various stages of injecting two polymer slugs during polymer flooding, using an ANN back-propagation algorithm with six input factors used to predict three output factors, using a hidden layer that had 10 neurons. The performance of the ANN model was compared using multiple linear regression. The ANN tool accurately predicted RF with less than 1% error, with the second polymer concentration having a smaller effect due to the first polymer injection altering permeability.

A study was conducted using simulations and computer models to understand how effective two types of thick liquid (polymers) are in helping recover oil. ANN tool accurately predicted RF with < 1% errors. The RMSE values were 0.11, 0.37, and 0.36 % for the RF at three different times. The results showed that the speed of injection and the amount of polymer in sizes 1 and 2, along with the size of the water drive, were the most important factors. This study demonstrated that the suggested ANN model (hidden layer composed of 10 neurons) can accurately predict how much oil can be recovered during polymer flooding using two types of polymers [68].

The reviewed studies demonstrate the significant impact of ML in optimizing CEOR processes, particularly polymer and surfactant flooding. ML models such as ANNs, MLPs, LSSVMs, and hybrid approaches like ANN-GA have shown remarkable accuracy in predicting key parameters like RF, NPV, and adsorption effects while drastically reducing computational costs compared to traditional simulators. Advanced methods like ANFIS, TPOT, and ADP further enhance decision-making, model optimization, and reservoir development. Some of the most significant variables that affect adsorption, recovery efficiency, and project economics are surfactant concentration, salinity, temperature, pH, polymer properties, and reservoir characteristics. Logistic regression and ANN-based methods have shown that they can forecast how well EOR will work in a variety of situations. But there are difficulties, such as the necessity for high-quality data, the capacity to apply the models to different reservoirs, and the difficulty of adding complicated reservoir physics to ML models. In table 2, many methods from previous works are reported, representing how the ML developed the EOR process by CEOR to increase the EOR in the last decades. The table summarizes the method of ML and optimization algorithms, which showed the type of ML, the type of chemical that was used, the inputs that were taken and their effects on output, the objective of the study, and the result, evaluation, and limitation.

Categorization of machine learning algorithms for optimizing the chemical flooding method

The literature review above for CEOR presents a classification of machine learning methods. LSSVM is a good method; it has high reliability and accuracy, as well as the ability to handle non-linear relationships among input features, making it suitable for complex EOR scenarios and effective integration with existing literature databases for improved predictions of RF and costs. ANNs are strong due to their exceptional predictive performance with very high R^2 values and low RMSE and flexibility in modelling complex non-linear relationships, which are common in EOR scenarios, where multiple interacting factors (such as chemical concentration, salinity, and temperature) are present. The use of ANNs in combination with GAs for model optimization enhances their performance; studies indicate that ANNs not only match but often exceed the performance of these conventional methods in terms of speed and accuracy and their ability to perform forward and inverse modelling.

MLPs are an effective, specific type of ANN that has shown strong performance in predicting chemical flooding outcomes and can be fine-tuned for better accuracy. The model utilizes various activation functions and can be fine-tuned for better accuracy. ANFIS is useful for combining neural networks with fuzzy logic to handle uncertainty and provide accurate results. PSO is beneficial for optimizing parameters in combination with other methods, although it is not a standalone predictive model, but it increases the performance of other models by optimizing the search space. ADP is considered useful for solving complex decision problems but may require computational resources. It is more suited for specific scenarios rather than broad prediction projects. TPOT is an Auto ML method that automates the optimization of machine learning pipelines.

Application of machine learning to thermal

The steam-assisted gravity drainage (SAGD) method is a complicated way, takes a long time, can have many different effects, and may vary a lot in different areas because of the complex dynamic of this process. In this paper a simulation model of a reservoir made with information from a reservoir in Alberta has one set of wells (one for injecting and one for producing) and looks at a production period of 250 days. The model is measured by the total amount of oil and water produced, as well as the amount of water injected. For each step in time, there are three possible actions: increase, decrease, or keep the same amount of steam injection. Stochastic gradient descent is employed to estimate the action-value function, so this process goes on for several runs of the changing situation until convergence is achieved. This paper introduces how RL can be used instead of traditional methods to improve the management of reservoirs. The goal is to find the best steam injection rates at each moment that will increase the NPV. RL surpasses both constant rate steam injection and traditional optimization techniques by increasing NPV and lowering the steam-oil ratio [70].

Earlier studies found that using steam flooding (SF) and CSI together led to the best increase in oil recovery, but more research is needed to find the best approach. The RF and NPV are the main goals, and they examined how different controlled factors affect these goals. A quadratic mathematical proxy model is developed to assist in evaluating and reaching the intended objective. The CMG-STARS simulator is used to model how heat recovery works in a reservoir. The proxy model created has been tested using statistics and shows a strong relationship in its results. Using the proxy model can greatly reduce the time needed for the evaluation, making it possible to solve the complex optimization problem much faster [71].

This paper talks about three models using ANNs. These three models were the Forward model (to estimate viscosity contours & performance indicators), Inverse Model (to guess the design details for CSI), and Inverse Model 2 (to forecast important features of a reservoir). Some of the results show higher error rates, like matrix permeability, but the error rates for the other results are acceptable. Viscosity contours reflect the operational performance indicators. When the contours are near the well, the size of the area that can be stimulated is small, while when the viscosity contours are far from the well, the area that gets treated is bigger, because the injected steam spreads out over a wide area; the distance between fractures determines the stimulated area. So, flow rate at the start becomes high because the steam adds energy that is kept in a much smaller space, although the oil was coming out quickly at first [72].

This paper uses ANN to forecast production from SAGD in non-homogeneous reservoirs. The Arps decline parameters from real data have been successfully tested for creating a total production profile. PCA is used to make the data simpler, improve the accuracy of predictions, and prevent

Table 2

Summary of the recent publication regarding CEOR methods

Author(s)	Chemical	Method	Dataset	Goal	Inputs	Result	Evaluation	Limitation
Mohammad Ali Ahmadi [3]	Surfactant	LSSVM, & GA	The initial dataset contained 202 data	When experimen-tal data is unavailable, the LSSVM approach can predict chemical flooding efficiency in oil reservoirs using statistical criteria such as MSE and R^2	Size & concentration of surfactant slug, polymer concentration in surfactant slug, Kv/Kh ratio, Size, concentration, & salinity in polymer	Stronger & accurate when considering both economic pers-pectives (NPV) and (RF). This tool enhances the accuracy of commercial reservoir simulators like Eclipse and CMG	LSSVM model rovides acceptable reliability, integrity, and robustness based on RMSE, correlation coefficient, and average absolute relative deviation	–
Jestril Ebaga [68]	Two types of polymers	ANN & back-propagation algorithm, MLR	A five-spot well pattern in a sandstone reservoir includes four production wells, one injector well, and 400 data points	RF was predicted at three unique intervals (RF1, RF2, and RF3) using a simulation study and ANN modeling	Size & concentration of polymer slug size 1 & 2, water drive size, and injection rate (Q)	Due to permeability changes following the first polymer slug, the second polymer concentration had less effect. When comparing ANN and MLR results, ANN won	Overall correlations between simulated results and ANN predictions were 0.999 for training, validating, and testing processes	–
Mohammed Alghazal [69]	Polymer Gel	ANN	80% training, 10% validation, and 10% testing sets from fractured reservoir	Provides a soft-computing, data-driven machine learning model for deep polymer gel. Surrogate modelling using ML to describe complex polymer gel kinetics and dynamics of flow	Temperature, injection rate, gel concentration, BHP, drainage radius, porosity, fracture spacing, permeability	ANN model reduced the complexity of the simulation model 200 times faster than commercial simulators and obtained a 90% correlation for test samples with a mean absolute deviation <10%	A second-order reaction system for gel formulation was used to explain the chemical component. The input parameters are sufficient to train the model and predict the conformance treatment's performance, including oil and water rate improvement after polymer gel treatment	–
Qian Sun [56]	polymer	ANN based proxies, & PSO		Quickly analyze polymer injection projects' (techno-economic feasibility) and provide a forward-looking and inverse design expert ANN system for screening and optimizing polymer	Initial state, reserve, oil & polymer characteristics, relative permeability, design parameter	A project screening strategy combining an expert system (forward-looking proxy) & PSO approach is proposed to maximize polymer injection	The forward-looking expert system is used for forecasting and screening, predicting project reactions. The inverse-looking ANN predicts project design concepts	Polymer supply, surface water treatment limits, and the capacity of the injection pump
Amine Tadjer [59]	Polymer	ADP	Reservoir development plan	Value of Information (VOI) analyses can be performed more consistently to handle sequential decision problems & provide knowledge about how decisions are made	Initial reservoir pressure, porosity, polymer concentration, water and oil density, viscosity, and compressibility	Choose the optimal moment to start a polymer flooding procedure. ADP evaluates uncertainties in dynamic and state variables, including economic indicators	VOI can be used for further domains, such as well-placement and energy storage optimization	–

Table 2 (continued)

Summary of the recent publication regarding CEOR methods

Author(s)	Chemical	Method	Dataset	Goal	Inputs	Result	Evaluation	Limitation
Hossein Saberi [60]	Hydrolyzed Polyacrylamide (HPAM)	MLP, & ANFIS	847 data records	Select the proper EOR methods	This includes polymer concentration, salt concentration, rock type, Soi, porosity, permeability, pore volume flooding, temperature, API gravity, polymer molecular weight, and salinity	with MLP being the most accurate	The MLP neural network accurately predicts data within and outside its built-in range	–
Negar Zarepakzad [61]	polymer	ANN-based data-driven model	Daqing Oil	Forecast polymer flooding performance in heteroge-neous reservoirs	Reservoir Properties & Operational Parameters	The created technology predicts performance indicators with an average absolute error difference of 0.017 and an R^2 of 0.94	Simulator outputs were used to train models with three performance indicators: efficiency, water cut, and recovery. Each scenario included no injection, water-only flooding, polymer-only flooding	All parameters were evenly distributed within boundaries
Mohammad Abdullah [63]	ASP	ANN	1100 training cases	The models cut computing time by four orders of mag-nitude	Response of reservoir (oil rate, water cut, injector BHP, cumulative oil), reservoir characteristics (permeability, thickness, Sor, chemical adsorption), and parameters design (pattern size, size & concen-tration of chemical slug)	Changing reservoir features and design parameters and recording the CMGSTARS reservoir reaction	Included a back-check loop that feeds the forward ANN model with expected variables derived from the inverse ANN models	–
Munqith Aldhaheri [66]	polymer gels	Logistic regression ML method	Datasets for 19 properties or parameters	Determine the best gel technology for a specific reservoir. & undetected injection wells	porosity, net thickness, oil viscosity, water cut, and RF	Predictively of classifiers was evaluated using three global indicators and visual monitoring with the prediction profile	Utilize machine learning to create an effective selection strategy for polymer gels in injection wells	–

mistakes from having too many details. It has been shown that using cluster analysis to identify patterns and groups in the data before applying ANN greatly improves the reliability and accuracy of predictions. SAGD works worse when there are a lot of long, continuous layers of shale. A new measure called the shale continuity indicator (SI) has been created; the indicator shows how the nearest shaly barrier affects the amount of oil during the SAGD process. Since important information like bottom-hole pressures, work schedules, fluid types, flowability, and thermal properties is usually not available to the public. These models are created without any individual preferences, which makes them advantageous for comparing how different fields perform solely based on their reservoir characteristics. The result helps identify if one field is doing better or worse than others with similar reservoir qualities. This research indicated that the steam-oil ratio is an important measure for both costs and operations [73].

The aim of this study is to build a computer system that can predict ultimate recovery. The ANN model included more than 250 different data points from various sources. The collected data was used in a special type of computer model called a feed-forward back-propagation to estimate RF with less than a 10% mistake. Feedforward networks take the data from input points through the hidden. The best ANN model has five layers. The first four layers are hidden layers, and the last layer is the output layer. The neural network interface was found to be correct with a 2–10 % margin of error. When this method was tested with a new set of 64 data, it made predictions that were 83% accurate according to linear regression [74].

The Fast-SAGD method is a new way; in this method, extra wells are drilled to inject and produce at regular times. This reduces the expenses associated with SAGD paired injection-production wells and results in increased output in a shorter time frame. The main goal of this research was to identify a better way to make the Fast-SAGD process work faster and more accurately. A GA was used to improve this study and choose a function that combines RF and cyclic steam oil ratio (CSOR). To make the optimization faster, advantageous variables were changed to fixed numbers. The position of the SAGD and nearby wells stays the same in the X direction, but the depth of the wells from the bottom of the reservoir can change. Other parameters can improve to optimize, including when to start the injection for the offset well, soak time, & time it takes to inject into the offset well. In thermal recovery, the CSOR-to-RF ratio is kept as small as possible (objective function = CSOR/RF). Adding more steam or increasing the number of cycles increases the saturation of steam in the surrounding area, which eventually causes steam to flow around the fracture. The results showed that the amount of oil produced using two-cycle and one-cycle methods was nearly the same, resulting in similar recovery rates. However, the two-cycle system used more injected steam than the one-cycle system, which led to a higher CSOR value in the two-cycle method. The RMSE values were found using two methods: ANN-K means and ANN-fuzzy. When the soak time is increased to 60 days, RF goes up and CSOR goes down, so it is the best amount of time for this model. The results showed that the one-cycle method is more important for CSOR and RF values [75, 76].

In this research, predictive models were developed for the steam huff and puff method by two types of ML techniques: regular polynomial regression and an ANN. Using a model of a single well-shaped synthetic oil reservoir, different experiments were created and tested regarding reservoir characteristic and six output factors that show how the reservoir produced and its characteristics after the injection phase and training data 80% & testing data 20%. The results, like total oil produced, the highest amount of oil produced, and the oil production rate at the end of the process from each simulation, are used to create the prediction model. The models were checked for fitting the data using three measures: the R-square value, MAE, and RMSE. The Sobol analysis shows that Swi, the viscosity of the oil, is the key factor for predicting oil performance. The PR method is used to analyze the relationship between inputs & results. The PR module was used to run the simulations by recording the results and changing the hyperparameters to create the accurate model. This method can be used to create similar models for different types of reservoirs or injection techniques to estimate performance for one cycle in a reservoir [77].

In this study, hierarchical clustering analysis was used (HCA) along with another technique called PCA to look at steam flooding projects around the world and discover patterns in these projects, which helps us identify the most similar cases. In this study, a technique known as principal component analysis was employed to transform the original data into a new structure. HCA uses a better optimization with five groups, a specific way to measure distance called Euclidean distance (Ward's linkage). Comparing the results with and without PCA before using HCA shows that using PCA with HCA facilitates to see clear clusters. Using the calculation of 30 measures and the grouping pattern found that the optimum is 5 groups, which means there are 5 stable patterns [78]. Data from an operating steam flood are constantly collected and used in a model called Data Physics (Ensemble Kalman Filter (EnKF)) to optimize NPV and injection costs. This case study focuses on a thick, heavy oil area. It demonstrates the application of data physics modelling to analyze future strategies. This paper gives a comparison between the field execution and model estimation, which permits demonstrating approval and highlights openings for assisted advancement. The model's predictions were reliable enough to trust for making future decisions [79].

In this study, machine learning suggested a way to plan steam use that can understand how different steam injection patterns affect heavy oil recovery. A study is conducted using a 3D model of an area involved in the SAGD process. For each well, a polynomial model is created using time data to predict key performance indicators (KPIs). The suggested smart steam allocation system will help make the operations of digital heavy oil facilities more efficient. These improvement will lead to higher profits and a smaller carbon footprint. This study shows the best use of steam in an SAGD facility when the condition is constantly changing. It provides instances of varying degrees of recovery: high, medium, and low. Improving processes based on different amounts of steam available: a) When there is 100% steam available, meaning all the steam needed is present, b) 100% or less steam available with ongoing injection c) Changing steam availability from 80 to 100 % of the operating capacity with ongoing injection rates, and d) The best steam availability needed with continuous injection rates. This research employs a model of an artificial reservoir consisting of three sets of horizontal wells. The wells were warmed up for four months before they started producing oil. The benefit of this suggested process is that it updates the reservoir prediction model in real-time, helping to make quick decisions. Unlike first-principal models that need a lot of computational resources to make decisions, the suggested workflow can be easily added to everyday decision-making in field operations [80].

In this paper, a new model is suggested to predict the heat-related properties of superheated steam (SHS) in injection wells and measure heat efficiency. This model is created to predict pressure, temperature, and superheat levels in SHS injection wells, using direct and indirect techniques to measure how effectively heat is employed within the well. Perform an analysis to assess sensitivity. This paper provides simple guidance for engineers on improving injection settings and discusses the prediction of thermal and physical characteristics of SHS. This paper presents four new ideas: (1). A new math model is created to predict the thermal and physical properties of SHS in wellbores, considering phase change. (2) A new method to measure heat is introduced and used in SHS injection wells. (3) The flow patterns of the SHS in the wells have been created. (4) The study also explores the impact of injection settings on the temperature at the bottom of the well. This study explains some key flows of SHS in oil wells. To understand how heat and flow move in wellbores during SHS, a few basic assumptions are: (1) The injection settings at the top of the well stay the same during the injection. (2) The heat transfer rate inside the wells must be constant, while the heat transfer rate in the surroundings changes. (3) Radiation and natural convection are the two main ways heat moves between the outer wall of the tube and the inner wall of the casing. (4) The physical and thermal properties are unaffected by changes in temperature or the depth of the well. Engineers need to choose a level of superheating that gets lower as the injection pressure increases. To achieve a higher superheat level at the bottom of the well, engineers should avoid setting the injection pressure too high. It's important to carefully consider the boiler's capacity and how well the heat moves in the well to find the optimal temperature for injection at the wellhead [81].

This study aims to create a proxy model that uses data to predict the total amount of oil produced during the SAGD process. When creating a model, ANN is used as an efficient helper for the physics-based model & the von Bertalanffy performance indicator to connect the physics-based model with the ANN. The average error of the performance indicator in the test data is 77%. The data-based proxy model could be used for quick studies by using the Monte Carlo method. The results of this study could improve our understanding of combining physics-based models with data-based models. In this study, ten features are chosen to use in the number-based simulation model. Due to the limited number of features, all ten attributes are utilized as input parameters for the proxy model based on data. This helps to gather as many different configurations as possible. The proxy model considers starting conditions, reservoir features, and operational settings. This research employs a combination of the «Reduce LR On Plateau» optimizer, ReLU activation function, and L1 loss function. Enhancing the predictive capabilities of the neural network, various techniques were implemented, including L2 regularization and modifications to the output [82].

The reviewed studies explore advanced techniques for optimizing thermal recovery operations, emphasizing ML, ANNs, and proxy models. RL worked well to dynamically optimize steam injection, raise NPV, and lower CSOR. ANNs and PCA improved production forecasting, especially in non-homogeneous reservoirs, while clustering methods enhanced data organization and interpretation. Fast-SAGD methods, supported by GAs, optimized CSOR and RF but required careful parameter tuning. Many challenges face thermal flooding, like data heterogeneity, shale barriers, and operational constraints. To maximize recovery efficiency, steam injection rates are one of the main factors influencing the different approaches for forecasting oil production in thermal processes. The reservoir's permeability and viscosity have a huge impact on how much oil and gas can be made and how quickly it can flow. The historical production data, which includes the total amount of water and oil produced, is used to train prediction models. These inputs help machine learning make better choices. Also, issues like when to inject and how long to soak are critical for attaining the optimum performance. In table 3 we can discuss all the methods of ML and optimization algorithms, which showed the type of ML, the type of thermal, the inputs, the objective of the study, and the result, evaluation & limitation.

Categorization of machine learning algorithms for optimizing thermal methods

The literature review above for thermal flooding presents a classification of ML methods. RL is a very effective way to improve SAGD operations. It can increase profits, cut computing costs, and change automatically based on the conditions of the reservoir by figuring out the best ways to inject and regularly interact with its surroundings. ANN-based models for predicting SAGD performance are very accurate and save time while still being precise. Another method is the combination of PCA and cluster analysis, which helps make predictions more accurate and make complicated data easier to understand. Fast-SAGD using GA optimization results in improved RF and reduced CSOR because it uses advanced methods (GAs) to modify its operation and aims to use less steam while getting the most out of recovery. Proxy models for optimizing SAGD use methods like ANN and PR to save time by using quick data-based estimates instead of running detailed computer simulations. The Smart Steam Allocation System uses live data to make quick decisions, which helps improve the value of an investment by 25-50 % by distributing steam more efficiently; this lowers carbon emissions and makes operations run better. HCA and PCA for steam flooding analysis find patterns and group similar projects, helping with decision-making, and PCA makes data simpler and keeps most of the important information, which helps with grouping similar items better. Data Physics Modelling, specifically the EnKF, is good for predicting the future and reducing injection costs. Finally, the quadratic mathematical proxy model for optimization depends on easier math models and is less flexible in changing or different reservoir situations.

Table 3

Machine learning models in the literature for thermal EOR

Author(s)	Type	Method	Dataset	Goal	Inputs	Result	Evaluation	Limitation
Ehsan Amirian [73]	SAGD	ANN	Alberta, Canada, these 150 reservoir models	Predict heterogeneous reservoir SAGD production	Characteristics & production/injection settings	Indicates increased prediction quality and limited overfitting by analyzing lot of difficult field data	Predicted using many characteristics, including shale barrier distance, thickness, and proportions	K-means clustering
J. L. Guevara [70]	SAGD	a model-predictive control (MPC) strategy with an RL algorithm	data from the northern Alberta heavy oil reservoir	Determine the optimal steam injection rates for each time step to maximize NPV	injection rate	RL optimizes steam injection policy, which improves NPV by at least 30%, reducing computation costs by 60%	Interactions with the environment help determine the appropriate policy	–
Xiaohu Dong [25]	Processes include CSS, SF, SAGD, in-situ combustion, thermal solvent, thermal-Non-condensate gas (NCG), and thermal-chemical	Review paper	China, North America, the North Sea, Bohai Bay, Canada, Venezuela, and Xinjiang	This critical review will help to detect future exploitation focusses on heavy oil & bitumen	experimental and field data	The hybrid thermal-solvent and thermal-NCG techniques are used to reduce oil viscoity and increase oil quality. Observe the physical qualities of formations have altered relative to their previous state	Point out that the hybrid thermal-chemical technique is used to manage a steam channeling path and increase sweep efficiency	–
Ahmet Ersahin, [72]	CSI	ANN	The data set consists of 1428 cases	Smart proxy models can imitate the usual numerical models	Features of the reservoir, properties of the fluid, rock and fluid interaction, injection design	It was noticed that ANN models can provide results in just a few seconds, while commercial software takes over 30 minutes and has a low error rate	It is shown that ANN models can quickly produce accurate results in just a short time	Viscosity contours, which are operational performance indicators
Yu Li [83]	CSI	RF	50 wells with a 30-year production history (1990 –2020)	To solve the problem of steam breakthrough, it is important to understand how steam moves through channels	The permeability ratio, steam injection speed, steam quality, steam-channelling reconstruction, and prediction set	Forecast the steam channelling during the CSI, which may affect breakthrough	Explain the state of steam channelling and estimate steam propaga-tion in the future cycle	The properties of formation, distance, and location of well
N. Sibaweihi [80]	SAGD	non-linear optimization	A designed reservoir model includes three pairs of horizontal wells. (90% training, 10% validation data)	Iimproves profits and reduces the environmental impact of steam injection on heavy oil by using ML	Reservoir properties	This strategy is successful; NPV increases by about 25 to 50 % compared to the basic situation, with steam injection remaining the same	Objective function suggested KPIs as high as possible by adjusting the amount of steam and lowering the steam-oil ratio	Steam availability, pressure levels. Shut in well work and surface equipment

Table 3 (continued)

Machine learning models in the literature for thermal EOR

Author(s)	Type	Method	Dataset	Goal	Inputs	Result	Evaluation	Limitation
Areeba Ansari [74]	SAGD	ANN	Western Canada, a total of 268 data points	Build an ANN that can correctly estimate the ultimate RF of oil reservoirs	Oil viscosity, steam injection rate, porosity, horizontal permeability, permeability ratio, reservoir thickness, and the pressure of the injected steam	A sensitivity analysis showed that horizontal permeability had the biggest effect on recovery, but porosity was found to be the least important. ANN could predict recovery about 90% accurately	Comparisons with other studies that were done on the same topic. The field shows that's need more real data points to create correct cases	Quality of data
Yang Yu [82]	SAGD	ANN, von Bertalanffy, Monte Carlo method	MacKay River oil sands	Develop an alternative model to forecast oil production	initial conditions, reservoir characteristic, operating parameters	The results indicate the possibility of a solution time series issues, & quick parametric studies, fast uncertainty analysis, and predicting daily oil production	The basic layout of the ANN has three hidden layers with 21, 16, and 15 neurons	–
Na Zhang [78]	Steam flooding	HCA & PCA	Testing Cases and Analogue Results from (China, Trinidad, Tobago, Canada, and Germany)	The aim of this research is to put similar SF projects together in groups and also get the optimal design and production experiences from these similar cases to make decision	Porosity, permeability, net thickness, gravity, viscosity of oil, depth, temp, & soi	Each group finds a specific range of property, and the analogue cases show that fields under similar reservoir/fluid conditions might have a similar operational design	It reduces the data from 8 to 2 dimensions while keeping about 90% of the important information	The threshold depth is 3000 ft
Fengrui Sun [81]	SHS	Mathematical model	The Kmk oil field is located in Aktyubinsk, which is in the northwest region of Kazakhstan	A novel model is showen for anticipating thermophysical properties (pressure, temperature, superheat degree, and wellbore heat efficiency) in SHS wells	Wellbore parameters (interior & exterior radius of internal tubing, external tubing, casing, and well depth). Injection parameters (temp., pressure, rate of injection)	Results indicate a superheat degreeat the well bottom increases with the increase of temperature, rate of injection, and decreases with the pressure of injection	The amount of heat flux is received as the index to approve the viability of a numerical model. And the average error of the model is 6.1%, so it's reliable	Heat transfer rate
Pallav Sarma [79]	steam flood	EnKF & evolutionary optimization algorithms	San Joaquin Basin of California	detect the optimal temporal and spatial distribution of steam injection that will improve future economics &recovery	Production & injection parameter, temperature, completion, steam quality	Optimize steam injection rates to maximize profits & minimize the cost of injection. provide quick, useful predictions	Showed a 60% increase in injection from the base case	–

Table 3 (continued)

Machine learning models in the literature for thermal EOR

Author(s)	Type	Method	Dataset	Goal	Inputs	Result	Evaluation	Limitation
Kaveh Mohammadi [76]	Fast-SAGD	GA	Naturally fractured heavy oil reservoirs in Iran (Mond field)	Exploring strategies of the Fast-SAGD process by improving time and accuracy	pressure & rate & period for injection and production wells, heights of offset well	It showed a 3% increase in RF and a 4.7% decrease in CSOR when using one injection cycle instead of two cycles	The optimization results using analysis of variance (ANOVA) showed that the injected steam and RF were 10% and 6% lower than the results from GA. The speed of this optimization algo-rithm is increasing from 50 to 70%	Operational constraints
Boni Swadesi [71]	SF & CSS	Proxy model	Sumatera field, Indonesia	Optimize the development scenario for a combined CSS-SF approach	Average permeability, oil viscosity, rock & fluid heat capacity, injection rate, time of injection &production& soaking, cycle number	The proxy model was used to assess the objective function and find the opti-mized variable. At higher permeability and lower viscosity to achieve the optimum result	This work shows that a combination of Proxy model develop-ment and optimization lead to the best scenario	–
Mohammad Galang Merdeka [77]	Steam huff and puff	ANN & PR	6043 different experiments with 28 inputs	Used an ML system to create a proxy. Models for one of the cycles will help to create simulation scenarios and save a lot of time	Proxy input (Sw, viscosity, press, thickness, area, permeability, temp, sor, swc, inj. rate), operating parameters (inj. rate, quality, temp of steam & soaking time)	ANN outperforms the polynomial regression model in accurate predictions and lower MAE and RMSE values	Sobol analysis was conducted to determine the sig-nificance of each variable in the model	This analysis notes that the proxy model has a limitation: it is only applicable when the input para-meter values fall within the model's specified parameter intervals

Conclusion

This study investigates the applications of artificial intelligence in WF & EOR. The application of AI algorithms improves the intelligence level of integrated analysis software. This paper provides a classification of these uses in three main strategies of EOR: water flooding, chemical flooding, and thermal techniques. AI and ML methods are increasingly gaining application in various sections of the petroleum industry as effective substitute for conventional methods. The proposed intelligence-based model works as an alternative for monitoring the efficiency of EOR in the absence of requisite experimental data. This paper provides a comprehensive review of practical considerations, including screening tools for selecting EOR projects, their goals, inputs, results, evaluation, and classification, as well as the optimization algorithm used in water injection, surfactant injection, and thermal injection. The novelty of this research is that it provides a complete study for all AI methods in EOR from 2009 to 2024, and to achieve the optimum design, implementation, and evaluation by references. Which facilitates the choice for engineers in future projects if the conditions of the required project are like the conditions and case of the reservoir used in this article.

This paper explains how different ML algorithms are perfect for EOR simulations where multiple parameters must be considered at once because they can handle high-dimensional and multivariate data to improve the efficiency, accuracy, and speed of reservoir management and decision-making procedures. As technology continues to advance and more data becomes available, ML models will likely become even more integrated into reservoir simulations, leading to even greater optimization and efficiency in EOR.

This paper introduces a strong process that can discover different ways to develop projects to reach the desired oil recovery goals. The objective of our work is to pioneer a novel invention of new techniques, typically classified as secondary and tertiary oil recovery methods, that maintain the economic production rate. This review focuses on ML techniques and how variables like production rate, salinity, temperature, pH, well configuration, geological features, historical production data, reservoir characteristics (viscosity and permeability), and operational parameters (soak time and injection timing) affect performance. Several novel procedures remain restricted to laboratory scales and necessitate further research. To fully utilize AI in the field of EOR, future studies should concentrate on hybrid models, real-time applications, and sustainable reservoir management techniques. This paper aims to offer commentary on the current state of ML applications in oil recovery, identify challenges, and suggest avenues for future research to optimize production strategies.

References

1. Alvarado, V., Manrique, E. (2010). Enhanced oil recovery: field planning and development strategies. *Gulf Professional Publishing*.

2. Hashim, A. H., Al-Jawad, M. S., Klati, K. A. (2023). Utilizing reservoir model to optimize future oil production for hydraulic fracture wells in tight reservoir. *The Iraqi Geological Journal*, 56(2B), 21-36.

3. Ahmadi, M. A., Pournik, M. (2016). A predictive model of chemical flooding for enhanced oil recovery purposes: Application of least square support vector machine. *Petroleum*, 2(2), 177-182.

4. Kuang, L., He, L. I. U., Yili, R. E. N., et al. (2021). Application and development trend of artificial intelligence in petroleum exploration and development. *Petroleum Exploration and Development*, 48(1), 1-14.

5. Salleh, N., Yuhaniz, S. S., Azmi, N. F. M. (2022). Modelling orbital propagation using regression technique and artificial neural network. *International Journal on Advanced Science, Engineering and Information Technology*, 12(3), 1279–1286.

6. Azim, R. A. (2022). A new correlation for calculating wellhead oil flow rate using artificial neural network. *Artificial Intelligence in Geosciences*, 3, 1-7.

7. Sircar, A., Yadav, K., Rayavarapu, K., et al. (2021). Application of machine learning and artificial intelligence in oil and gas industry. *Petroleum Research*, 6, 379–391.

8. Alkinani, H. H., Al-Hameedi, A. T. T., Dunn-Norman, S., Flori, R. E. (2019). Applications of artificial neural networks in the petroleum industry: A review. In: *The SPE Middle East Oil and Gas Show and Conference, Manama, Bahrain.*

9. Bahaloo, S., Mehrizadeh, M., Najafi-Marghmaleki, A. (2023). Review of application of artificial intelligence techniques in petroleum operations. *Petroleum Research*, 8, 167–182.

10. Solanki, P., Baldaniya, D., Jogani, D., et al. (2022). Artificial intelligence: New age of transformation in petroleum upstream. *Petroleum Research*, 7(1), 106-114.

11. Tontiwachwuthikul, P., Chan, C. W., Zeng, F., et al. (2020). Recent progress and new developments of applications of artificial intelligence (AI), knowledge-based systems (KBS), and machine learning (ML) in the petroleum industry. *Petroleum*, 6, 319–320.

12. Syah, R., Ahmadian, N., Elveny, M., et al. (2021). Implementation of artificial intelligence and support vector machine learning to estimate the drilling fluid density in high-pressure high-temperature wells. *Energy Reports*, 7, 4106-4113.

13. Ghorbania, H., Wood, D. A., Choubineh, A., et al. (2018). Prediction of oil flow rate through an orifice flow meter: Artificial intelligence alternatives compared. *Petroleum*, 6(4), 404-414.

14. Chakra, N. C., Song, K. Y., Gupta, M. M., Saraf, D. N. (2013). An innovative neural forecast of cumulative oil production from a petroleum reservoir employing higher-order neural networks (HONNs). *Journal of Petroleum Science and Engineering*, 106, 18–33.

15. AL-Assad, E. A., Hamd-Allah, S. M. (2024). Applying artificial neural networks with Bourgoyne and Young model to predict rate of penetration in Al-Garraf oil field. *Journal of Engineering*, 30(12), 146-166.

16. Anifowose, F. A., Labadin, J., Abdulraheem, A. (2017). Ensemble machine learning: An untapped modelling paradigm for petroleum reservoir characterization. *Journal of Petroleum Science and Engineering*, 151, 480–487.

17. Zhang, J., Yan, W. (1992). Summary of small well spacing waterflooding development and enhanced oil recovery test in Daqing oilfield. *Petroleum Geology & Oilfield Development in Daqing*, 2, 70–74.

18. Wang, D., Wang, L., Zhang, F., et al. (2022). Research status and development direction of separated layer water injection technology in Bohai oilfield. *China Offshore Oil and Gas*, 34(2), 125–137

19. Yang, Y., Cao, X. (2024). New comprehensive management method and engineering practice path for mature oilfields with high water cut. *Petroleum Geology and Recovery Efficiency*, 31(1), 54–62.

20. Ketineni, S. P., Ertekin, T., Anbarci, K., Sneed, T. (2015, September). Structuring an integrative approach for field development planning using artificial intelligence and its application to an offshore oilfield. SPE-174871-MS. In: *SPE Annual Technical Conference and Exhibition, Houston, Texas, USA*.

21. Al-Husseini, A. K., Hamd-Allah, S. M., Hamzah, S. H. (2025). Enhancing sweep efficiency and delaying water breakthrough in carbonate reservoir: The impact of polymer flooding. *Research Square*

22. Al-Dousari, M. M., Garrouch, A. A. (2013). An artificial neural network model for predicting the recovery performance of surfactant polymer floods. *Journal of Petroleum Science and Engineering*, 109, 51–62.

23. Al-Rikaby, A. S., Al-Jawad, M. S. (2025). A novel method to recognise faults and reefs from history matching of gas production-case study: Mishrif reservoir, Southern Iraq. *Iraqi Journal of Science*, 66(1), 209-218.

24. Shafiei, A., Ahmadi, M., Dusseault, M., et al. (2018). Data analytics techniques for performance prediction of steamflooding in naturally fractured carbonate reservoirs. *Energies*, 11(2), 292.

25. Dong, X., Liu, H., Chen, Z., et al. (2019). Enhanced oil recovery techniques for heavy oil and oilsands reservoirs after steam injection. *Applied energy*, 239, 1190-1211.

26. Pang, Z., Wang, L., Yin, F., Lyu, X. (2021). Steam chamber expanding processes and bottom water invading characteristics during steam flooding in heavy oil reservoirs. *Energy*, 234, 121214.

27. Shin, H., Choe, J. (2009). Shale barrier effects on the SAGD performance. SPE-125211-MS. In: *The SPE/EAGE Reservoir Characterization and Simulation Conference, Abu Dhabi, UAE, October.*

28. Al-Obaidi, D. A., Al-Jawad, M. S. (2020). Immiscible CO2-Assisted gravity drainage process for enhancing oil recovery in bottom water drive reservoir. *Association of Arab Universities Journal of Engineering Sciences*, 27(2), 60-66.

29. Al-Obaidi, D. A., Al-Jawad, M. S. (2020). Numerical simulation of immiscible CO_2-assisted gravity drainage process to enhance oil recovery. *Iraqi Journal of Science*, 61(8), 2004-2016.

30. Shin, H., Polikar, M. (2005). New economic indicator to evaluate SAGD performance. SPE-94024-MS. In: *The SPE Western Regional Meeting, Irvine, California, March. Society of Petroleum Engineers.*

31. McGlade, C., Sondak, G., Han, M. (2021). Whatever happened to enhanced oil recovery? *International Energy Agency*.

32. Buckles, R. S. (1979). Steam stimulation heavy oil recovery at Cold Lake, Alberta. SPE-7994-MS. In: *The SPE California Regional Meeting, Ventura, California, April.*

33. Beattie, C. I., Boberg, T. C., McNab, G. S. (1991). Reservoir stimulation in the Cold Lake oil sands. *SPE Reservoir Engineering*, 6, 200–206.

34. Gael, B. T., Gross, S. J., McNaboe, G. J. (1995). Development planning and reservoir management in the Duri steam flood. SPE-29668-MS. In: *The SPE Western Regional Meeting, Bakersfield, California, March.*

35. Bae, W. S., Masduki, A., Permadi, A. K., Abdurrahman, M. (2017). EOR in Indonesia: Past, present, and future. *International Journal of Oil, Gas and Coal Technology*, 16, 250-270.

36. Pearce, J. C., Megginson, E. A. (1991). Current status of the Duri steamflood project Sumatra, Indonesia. SPE-21527-MS. In: *The SPE International Thermal Operations Symposium, Bakersfield, California, February.*

37. De Haan, H. J., van Lookeren, J. (1969). Early results of the first large-scale steam soak project in the Tia Juana Field, Western Venezuela. *Journal of Petroleum Technology*, 21(01), 101–110.

38. Puig, F., Schenk, L. (1984). Analysis of the performance of the M-6 area of the Tia Juana Field, Venezuela, under primary, steam-soak, and steam drive conditions. SPE-12656-MS. In: *The SPE Enhanced Oil Recovery Symposium, Tulsa, Oklahoma, April.*

39. Canbolat, S. (2020). Assessment of SAGD heat transfer mechanism via no condensable gas co-injection. *Journal of Petroleum Science and Engineering*, 195, 107936.

40. Berry, M. W., Mohamed, A., Yap, B. W. (2019). Supervised and unsupervised learning for data science. *Springer*.

41. Lam, D., Wunsch, D. C. (2014). Academic press library in signal processing: Volume 1. Signal processing theory and machine learning. *Elsevier*.

42. Li, B., Billiter, T. C., Tokar, T. (2021). Rescaling method for improved machine-learning decline curve analysis for unconventional reservoirs. *SPE Journal*, 26, 1759–1772.

43. Kasha, A. A., Sakhaee-Pour, A., Hussein, I. A. (2022). Machine learning for capillary pressure estimation. *SPE Reservoir Evaluation & Engineering*, 25, 1–20.

44. Guo, K., Li, H., Yu, Z. (2016). In-situ heavy and extra-heavy oil recovery: A review. *Fuel*, 185, 886–902.

45. Taher, W. I., Hamd-Allah, S. M. (2019). Feasibility study of applying a pattern waterflood without and with peripheral injectors for maximizing oil recovery under limited water sources availability a case study of heterogeneous giant carbonate reservoir. *Journal of Engineering and Applied Sciences*,14(Special Issue 7), 9923-9935.

46. Ahmed, B. I., Al-Jawad, M. S. (2020). Geomechanical modelling and two-way coupling simulation for carbonate gas reservoir. *Journal of Petroleum Exploration and Production Technology*, 10(8), 3619-3648.

47. Zhong, Z., Sun, A. Y., Wang, Y., Ren, B. (2020). Predicting field production rates for waterflooding using a machine learning-based proxy model. *Journal of Petroleum Science and Engineering*, 194, 107574.

48. Deng, L., Pan, Y. (2021). Data-driven proxy model for waterflood performance prediction and optimization using Echo State Network with Teacher Forcing in mature fields. *Journal of Petroleum Science and Engineering*, 197, 107981.

49. Siggel, L., Santa, M., Hansch, M., et al. (2012, April). A new class of viscoelastic surfactants for enhanced oil recovery. SPE-153969-MS. In: *The SPE Improved Oil Recovery Symposium, Tulsa, Oklahoma, USA.*

50. Negash, B. M., Yaw, A. D. (2020). Artificial neural network based production forecasting for a hydrocarbon reservoir under water injection. *Petroleum Exploration and Development*, 47, 383–392.

51. Jia, D., Liu, H., Zhang, J., et al. (2020). Data-driven optimization for fine water injection in a mature oilfield. *Petroleum Exploration and Development*, 47, 674–682.

52. Du, S. Y., Zhao, X. G., Xie, C. Y., et al. (2023). Data-driven production optimization using particle swarm algorithm based on the ensemble-learning proxy model. *Petroleum Science*, 20, 2951–2966.

53. Kubota, L., Reinert, D. (2019). Machine learning forecasts oil rate in mature onshore field jointly driven by water and steam injection. SPE-196152-MS. In: *Proceedings of the SPE Annual Technical Conference and Exhibition. Society of Petroleum Engineers.*

54. Jia, D., Zhang, J., Li, Y., et al. (2023). Recent development of smart field deployment for mature waterflood reservoirs. *Petroleum Exploration and Development*, 15, 784.

55. Belhaj, A. F., Elraies, K. A., Mahmood, S. M., et al. (2020). The effect of surfactant concentration, salinity, temperature, and pH on surfactant adsorption for chemical enhanced oil recovery: a review. *Journal of Petroleum Exploration and Production Technology*, 10, 125-137.

56. Sun, Q., Ertekin, T. (2020). Screening and optimization of polymer flooding projects using artificial-neural-network (ANN) based proxies. *Journal of Petroleum Science and Engineering*, 185, 106617.

57. Karambeigi, M. S., Zabihi, R., Hekmat, Z. (2011). Neuro-simulation modelling of chemical flooding. *Journal of Petroleum Science and Engineering*, 78, 208–219.

58. Amirian, E., Dejam, M., Chen, Z. (2018). Performance forecasting for polymer flooding in heavy oil reservoirs. *Fuel*, 216, 83–100.

59. Tadjer, A., Bratvold, R. B., Hong, A., Hanea, R. (2021). Application of machine learning to assess the value of information in polymer flooding. *Petroleum Research*, 6(4), 309-320.

60. Saberi, H., Esmaeilnezhad, E., Choi, H. J. (2021). Artificial neural network to forecast enhanced oil recovery using hydrolyzed polyacrylamide in sandstone and carbonate reservoirs. *Polymers*, 13(16), 2606.

61. Zarepakzad, N., Artun, E., Durgut, I. (2021). A comparative analysis and rapid performance prediction of polymer flooding process by coupling reservoir simulation with neural networks. *International Journal of Oil, Gas and Coal Technology*, 27(3), 227-246.

62. Sun, Q. (2017). The development of an artificial-neural-network-based toolbox for screening and optimization of enhanced oil recovery projects. Doctoral Thesis. *The Pennsylvania State University, Energy and Mineral Engineering, University Park, PA, USA.*

63. AlAbdullah, M., Emami-Meybodi, H., Ertekin, T. (2019, June). Development and application of an artificial neural network tool for chemical EOR field implementations. SPE-195492-MS. In : *The SPE Europec featured at 81st EAGE Conference and Exhibition, London, England, UK, June*.

64. Aihara, T., Takabe, M., Toyoda, K. (1990). Chaotic neural networks. *Physics Letters A*, 144(6), 333–340.

65. Jiang, J., Shao, K., Wei, Y., Tian, T. (2007). Chaotic neural network model for output prediction of polymer flooding. In: *Proceedings of the International Conference on Mechatronics and Automation, Harbin, China.*

66. Aldhaheri, M., Wei, M., Bai, B., Alsaba, M. (2017). Development of machine learning methodology for polymer gels screening for injection wells. *Journal of Petroleum Science and Engineering*, 151, 77-93.

67. Hou, J. R., Liu, Z. C., Dong, M. Z., et al. (2006). Effect of viscosity of alkaline/surfactant/polymer (ASP) solution on enhanced oil recovery in heterogeneous reservoirs. PETSOC-06-11-03. *Journal of Canadian Petroleum Technology*, 45(11).

68. Ebaga-Ololo, J., Chon, B. H. (2017). Prediction of polymer flooding performance with an artificial neural network: a two-polymer-slug case. *Energies*, 10(7), 844.

69. Alghazal, M., Ertekin, T. (2020, November). Modelling of deep polymer gel conformance treatments using machine learning. SPE-203450-MS. In: *The Abu Dhabi International Petroleum Exhibition & Conference, Abu Dhabi, UAE.*

70. Guevara, J. L., Patel, R. G., Trivedi, J. J. (2018, December). Optimization of steam injection for heavy oil reservoirs using reinforcement learning. SPE-193769-MS. In: *The SPE International Heavy Oil Conference and Exhibition, Kuwait City, Kuwait*.

71. Morvan, M., Degré, G., Leng, J., et al. (2009, April). New viscoelastic fluid for chemical EOR. SPE-121675-MS. In: The SPE International Symposium on Oilfield Chemistry, *The Woodlands, Texas.*

72. Ersahin, A., Ertekin, T. (2020). Artificial neural network modelling of cyclic steam injection process in naturally fractured reservoirs. *SPE Reservoir Evaluation & Engineering*, 23(03), 0979-0991.

73. Amirian, E., Leung, J. Y., Zanon, S., Dzurman, P. (2015). Integrated cluster analysis and artificial neural network modelling for steam-assisted gravity drainage performance prediction in heterogeneous reservoirs. *Expert Systems with Applications*, 42(2), 723-740.

74. Ansari, A., Heras, M., Nones, J., et al. (2020). Predicting the performance of steam assisted gravity drainage (SAGD) method utilizing artificial neural network (ANN). *Petroleum*, 6(4), 368-374.

75. Sarapardeh, A., Kiasari, H. H., Alizadeh, N., et al. (2013). Application of fast-SAGD in naturally fractured heavy oil reservoirs: A case study. SPE-164418-MS. In: *Proceedings of the SPE Middle East Oil and Gas Show and Conference. Society of Petroleum Engineers.*

76. Mohammadi, K., Ameli, F. (2019). Toward mechanistic understanding of Fast SAGD process in naturally fractured heavy oil reservoirs: Application of response surface methodology and genetic algorithm. *Fuel*, 253, 840-856.

77. Merdeka, M. G., Ridha, S., Negash, B. M., Ilyas, S. U. (2022). Reservoir performance prediction in steam huff and puff injection using proxy modelling. *Applied Sciences*, 12(6), 3169.

78. Zhang, N., Wei, M., Bai, B., et al. (2022). Pattern recognition for steam flooding field applications based on hierarchical clustering and principal component analysis. *ACS Omega*, 7(22), 18804-18815.

79. Sarma, P., Lawrence, K., Zhao, Y., et al. (2018, December). Implementation and assessment of production optimization in a steamflood using machine-learning assisted modelling. SPE-193680-MS. In: *The SPE International Heavy Oil Conference and Exhibition, Kuwait City, Kuwait*.

80. Sibaweihi, N., Patel, R. G., Guevara, J. L., et al. (2019, April). Real-time steam allocation workflow using machine learning for digital heavy oil reservoirs. SPE-195312-MS. In: *The SPE Western Regional Meeting, San Jose, California, USA.*

81. Sun, F., Yao, Y., Chen, M., et al. (2017). Performance analysis of superheated steam injection for heavy oil recovery and modelling of wellbore heat efficiency. *Energy*, 125, 795-804.

82. Yu, Y., Liu, S., Liu, Y., et al. (2021). Data-driven proxy model for forecasting of cumulative oil production during the steam-assisted gravity drainage process. *ACS Omega*, 6(17), 11497-11509.

83. Li, Y., Liu, H., Jiao, P., et al. (2023). Machine-learning-assisted identification of steam channeling after cyclic steam stimulation in heavy-oil reservoirs. *Geofluids*, 2023, 1-13.