machine learning & knowledge extraction

Review

A Comprehensive Summary of the Application of Machine Learning Techniques for CO_2-Enhanced Oil Recovery Projects

Xuejia Du, Sameer Salasakar and Ganesh Thakur *

Department of Petroleum Engineering, Cullen College of Engineering, University of Houston, Houston, TX 77204, USA; xdu8@central.uh.edu (X.D.); ssalasak@cougarnet.uh.edu (S.S.)
* Correspondence: gcthakur@central.uh.edu

Abstract: This paper focuses on the current application of machine learning (ML) in enhanced oil recovery (EOR) through CO_2 injection, which exhibits promising economic and environmental benefits for climate-change mitigation strategies. Our comprehensive review explores the diverse use cases of ML techniques in CO_2-EOR, including aspects such as minimum miscible pressure (MMP) prediction, well location optimization, oil production and recovery factor prediction, multi-objective optimization, Pressure–Volume–Temperature (PVT) property estimation, Water Alternating Gas (WAG) analysis, and CO_2-foam EOR, from 101 reviewed papers. We catalog relative information, including the input parameters, objectives, data sources, train/test/validate information, results, evaluation, and rating score for each area based on criteria such as data quality, ML-building process, and the analysis of results. We also briefly summarized the benefits and limitations of ML methods in petroleum industry applications. Our detailed and extensive study could serve as an invaluable reference for employing ML techniques in the petroleum industry. Based on the review, we found that ML techniques offer great potential in solving problems in the majority of CO_2-EOR areas involving prediction and regression. With the generation of massive amounts of data in the everyday oil and gas industry, machine learning techniques can provide efficient and reliable preliminary results for the industry.

Keywords: machine learning; CO_2-EOR; minimum miscible pressure (MMP); water-alternating-gas (WAG); system review

check for updates

Citation: Du, X.; Salasakar, S.; Thakur, G. A Comprehensive Summary of the Application of Machine Learning Techniques for CO_2-Enhanced Oil Recovery Projects. *Mach. Learn. Knowl. Extr.* **2024**, *6*, 917–943. https://doi.org/10.3390/make6020043

Academic Editors: Jaroslaw Krzywanski, Marcin Sosnowski, Karolina Grabowska, Dorian Skrobek, Ghulam Moeen Uddin and Andreas Holzinger

Received: 15 February 2024
Revised: 22 April 2024
Accepted: 26 April 2024
Published: 29 April 2024

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

There is a strong correlation between energy consumption and economic growth. Liquid fossil fuels are a key component of the energy mix, contributing up to about 35% of worldwide energy usage. While energy sources are diversifying, liquid fossil fuels are still a key energy source in developing countries such as India and China. The rapid development of these economies will most likely intensify energy generation from fossil fuels. This will unsurprisingly lead to CO_2 emissions. CO_2 emissions have been rising worldwide. The IPCC report on "Global Warming of 1.5 °C" declared a major concern that unless CO_2 emissions are reduced by 50% by the year 2030, major changes will occur in the ocean and on the land, and unfortunately, they may be permanent in nature.

The time is of essence to globally transition to new energy systems. Bloomberg news mentions "Climate change is not a problem with a single solution. And it is not a challenge that any one group—governments, companies, scientists or individual citizens—can solve alone". Working together, one can build a healthier and more sustainable future for the generations to come. Utilizing a variety of technologies, e.g., solar, wind, geo-thermal, nuclear, extended batteries, and hydrogen, and strong government support, dedicated companies, universities and research centers, regulatory agencies and others, we have a great opportunity to solve the problem.

We can distinguish two main strategies for reducing atmospheric concentrations of CO_2. The first strategy includes reducing the emissions of CO_2 to the atmosphere by increasing energy efficiency and switching to low-carbon fuel sources, utilizing proven and existing technologies, e.g., solar, wind and nuclear at a large scale and fast pace. The second strategy includes the deployment of negative emission technologies to remove carbon from the atmosphere and sequester it reliably. Some examples of this strategy may include DAC (direct air capture), CCS and CCUS (e.g., CO_2 EOR). The potential impact of these technologies on reducing CO_2 emissions is immense and should not be underestimated.

Our knowledge of the reservoir management of an oil and gas field from primary to tertiary recovery phases yields an understanding of its key properties. Hence, the use of mature or declining oil and gas reservoirs to store CO_2 significantly reduces subsurface uncertainties. CO_2 injection is a well-documented method for improving hydrocarbon production rates and increasing recovery. Thus, in light of climate concerns, using CO_2 injection for the dual objectives of enhancing oil recovery and carbon storage is a powerful choice.

Petroleum resources have been deemed as the principal source of fossil-fuel-based energy to meet the world's energy demands since the early 20th century. The importance of enhancing oil reservoir extraction efficiency has grown due to the restricted supply of reserves. Over two-thirds of the original oil in place (OOIP) remains trapped after primary and secondary recovery processes. Furthermore, extracting the remaining oil from mature reservoirs in complicated geological formations is more challenging. EOR methods are initiated to recover the remaining oil from reservoirs after both primary and secondary recovery methods have been exhausted. Surfactant flooding, chemical flooding, polymer flooding, steam stimulation, microbial flooding, gas injection, and so forth [1,2] are the common EOR approaches. Carbon dioxide (CO_2) is very successful since it increases oil production by increasing mobility and reducing oil viscosity and saturation, which works well with both conventional and some unconventional formations. CO_2-EOR is one of the popular techniques, occupying around 20% of 1120 worldwide EOR projects (Figure 1). It may recover 15% to 25% of the OOIP of the light or medium oil fields that are close to depletion due to flooding [3].

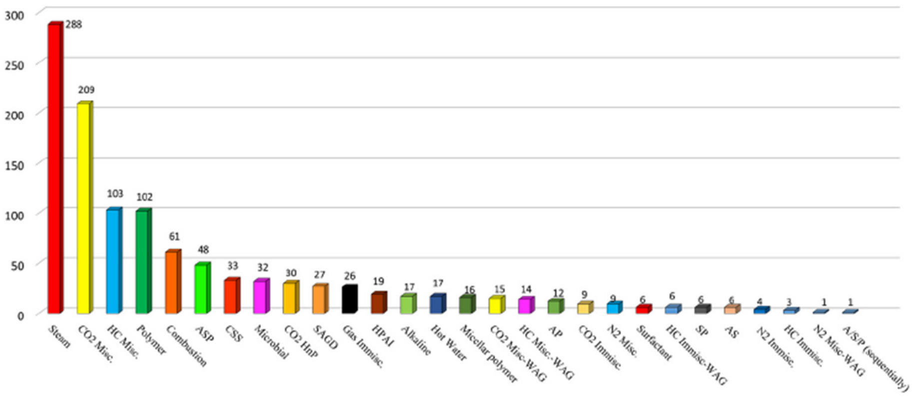


Figure 1. Distribution of different EOR projects worldwide [4].

The utilization of CO_2 in EOR can significantly improve oil recovery; at the same time, it plays an essential role in environmental preservation. The importance of CO_2-EOR as part of carbon capture, use, and storage (CCUS) schemes becomes more vital as the petroleum industry works toward decarbonization to mitigate greenhouse gas emissions. If reinjection is not considered, approximately 60% of injected CO_2 can be trapped in the reservoir at the CO_2 breakthrough [5]. This approach, efficiently utilizing CO_2 in oil recovery, aligns with an environmentally friendly protocol while simultaneously enhancing resource efficiency and contributing substantially to sustainability goals [6].

Machine learning (ML) approaches have drawn considerable interest as emerging technologies in the oil and gas industry over the past 20 years. Applying the ML approaches to examine issues in the oilfield development process has acquired new life with the advent of intelligent oilfields and big data technology. Indeed, ML shows the feasibility of offering a more straightforward and quicker method than rigorous and numerous simulations or experiments. Many ML correlations have emerged with the development of computer tools, particularly in reservoir characterization, CO_2 storage, production, and drilling operations [7–10].

Many literature reviews have been conducted in the past to summarize the application of ML in the oil and gas industry [11]. However, no study on global research trends analyzed the dominant input parameters and evaluated the research work on CO_2-EOR projects. The evaluations could help researchers get a preliminary idea about the current research trend on CO_2-EOR and whether their recent research impacts a particular field. Furthermore, few studies have systematically summarized and examined all the literature on ML for CO_2-EOR. Few reviews find the most critical topics, objectives, input parameters, evaluations, and research gaps in ML for CO_2-EOR. This study aims to offer insight into current trends and technological development indicators, which will help identify the viewpoint for the following research areas and prospects. Thus, data extraction analysis was carried out to ascertain the research advancement and trends in ML for CO_2-EOR, whereby a systematic review accomplishes the closure of research gaps on this subject.

This paper aims to summarize and evaluate the various ML models in CO_2-EOR and provide insightful analysis with 101 papers reviewed. The rest of the paper is organized as follows: Section 2 describes the mechanisms and processes of CO_2-EOR; Section 3 provides the most popular ML and optimization methods employed in the literature; Section 4 summarizes the work that has applied ML in the CO_2-EOR process, including MMP prediction, WAG, well placement optimization, oil production or recovery factor prediction, multiple objectives optimization, PVT properties estimation, and CO_2-foam; and Section 5 outlines the benefits and limitations of the application of ML in the CO_2-EOR process, before ending this survey paper with concluding remarks.

2. Mechanisms and Processes of CO_2-EOR

CO_2 is generally injected into the reservoir under the following conditions: (a) miscible injection; (b) immiscible front displacement after water flooding; (c) water alternating gas (WAG) displacement; and (d) CO_2 dissolved in brine flooding, also referred to as carbonated water injection (CWI) [12]. Miscible displacement has been successful over the years. It occurs at pressures above a minimum miscible pressure (MMP) of the oil, where the injected gas and the hydrocarbons are entirely miscible and form a single-phase fluid. The main advantages of miscible displacement are that it can promote oil swelling, reduce fluid viscosity, increase mobility, reduce remaining oil saturation, and improve oil production.

CO_2 has been historically favored over other gases due to its low MMP. Furthermore, CO_2 gas injection can potentially mitigate greenhouse gas emissions while improving oil recovery. CO_2-miscible flooding, whether initiated upon first contact or multiple contacts, results in the remaining oil and CO_2 becoming miscible, which leads to near-zero interfacial tension (IFT), no capillary pressure, and improved volumetric sweep (Ev) and displacement efficiency (Ed) [13]. Conversely, in the case of CO_2-immiscible flooding, the IFT is not near zero, maintaining the capillary pressure and causing some residual oil saturation. The oil recovery efficacy is contingent upon the efficiency of fluid displacement, volumetric sweep, and CO_2 solubility in the oleic phase, consequently increasing oil mobility. These characteristics are influenced by various factors, including gravity, rock wettability, reservoir heterogeneity, crude oil phase behavior, and phenomena such as viscous fingering, etc. [12,14].

3. Summary of Machine Learning Approaches

Machine learning (ML) involves the development of computational models and algorithms capable of learning patterns and making data-driven predictions or decisions without being explicitly programmed. ML algorithms employ data to automatically identify and generalize patterns, which may be applied for classification, regression, clustering, and more tasks. ML can be categorized into four main types: supervised learning, unsupervised learning, semi-supervised learning, and reinforcement learning. Figure 2 provides some examples of different ML algorithms. Among these various algorithms, supervised learning is most applied in the oil and gas industry [11].

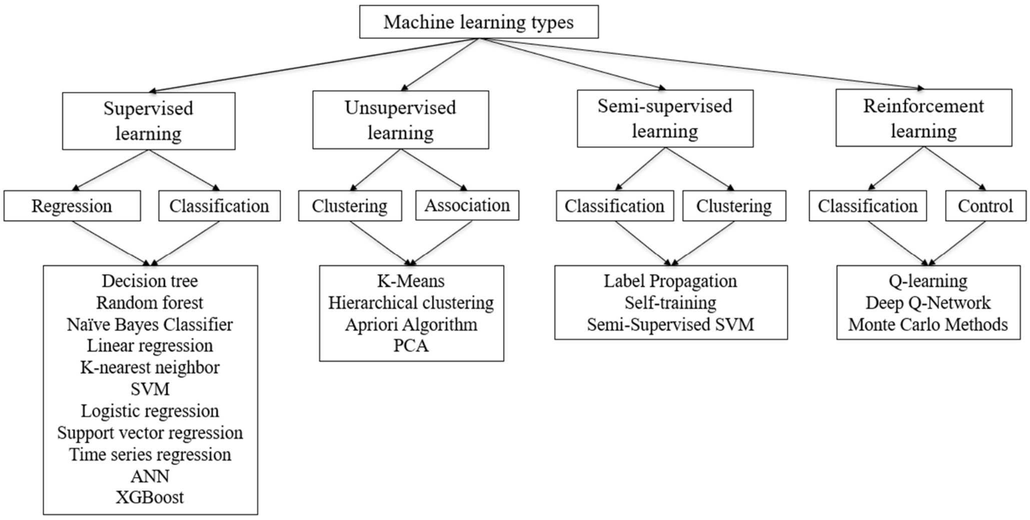


Figure 2. Examples of different machine learning algorithms.

For instance, ANNs have demonstrated remarkable efficacy in providing user-friendly, cost-effective, reliable, and expedited solutions to a variety of complex challenges encountered in the oil and gas industry. This is primarily attributed to the inherent complexity and non-linear nature of oil and gas datasets, which often have intricate relationships between input variables and output parameters. ANNs excel in capturing these complex relationships by effectively modeling non-linear functions. Moreover, oil and gas data are frequently characterized by noise, incompleteness, and heterogeneity. ANNs exhibit superior capability in handling such diverse data types and can adeptly adapt to varying data distributions, thereby making them highly versatile for addressing various tasks across different domains within the industry.

Furthermore, the enhancement of the ML process involves optimization techniques to determine optimal values for control parameters, including the spreading coefficient, number of neurons, biases, and weights. Several optimization methods, such as the Levenberg–Marquardt (LM) algorithm, genetic algorithm (GA), and smart nature-inspired swarm algorithms like particle swarm optimization (PSO), grey wolf optimization (GWO), and ant colony optimization (ACO), have demonstrated their efficacy in achieving significant improvements in these tasks. There are two categories in intelligent optimization algorithms: single-objective optimization and multi-objective optimization (Figure 3).

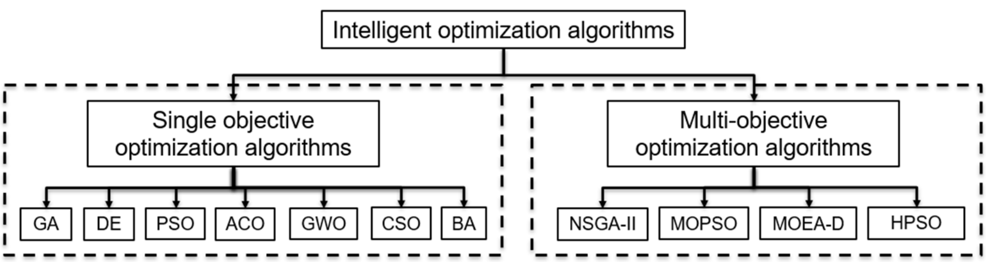


Figure 3. Representative intelligent optimization algorithms [15].

4. Application of ML in CO_2-EOR

4.1. Minimum Miscibility Pressure (MMP)

In miscible gas injection, MMP is one of the most important parameters to determine the accuracy of miscible CO_2 flooding into the reservoir. Traditionally, MMP is defined as the pressure at which 80% of the OOIP is extracted from the reservoir upon the breakthrough of CO_2 [16]. Because CO_2 flooding is more expensive than waterflooding, an accurate estimation of MMP can help better design miscible CO_2 flooding, ultimately leading to cost savings. In the literature, researchers have proposed various MMP estimation approaches, including the following:

(a) experimental methods such as slim-tube tests [17]; rising-bubble apparatus [18]; vanishing interfacial tension [19];

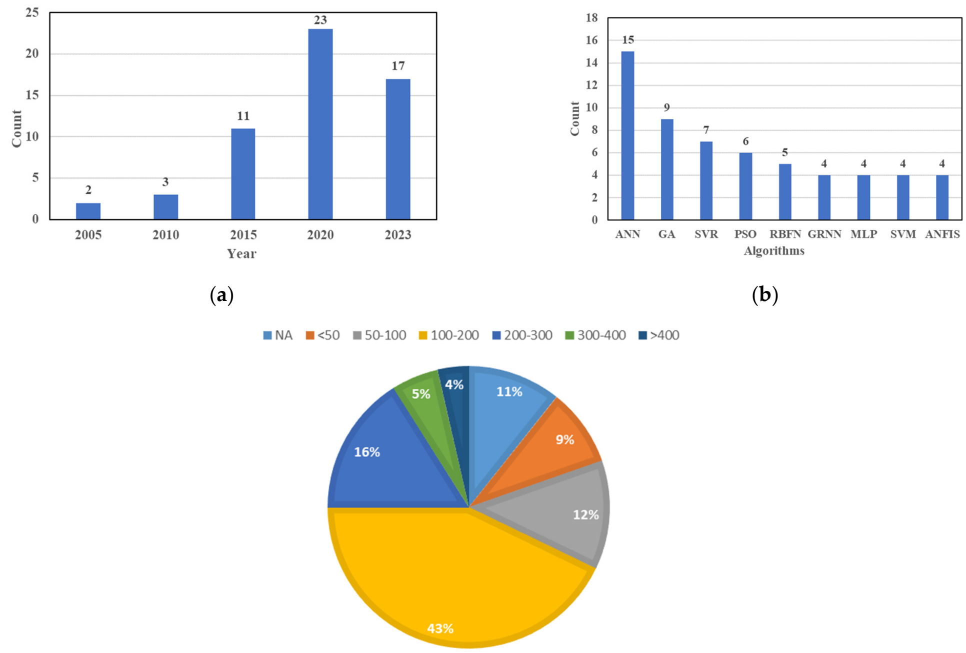
(b) empirical correlations [17,20–22] and computational techniques such as single mixing-cell and multiple mixing-cell approaches [23].

However, though accurate and reliable, experimental methods are time-consuming and expensive. Most empirical correlations and computation techniques do not consider different thermodynamic and reservoir properties. Moreover, they exhibit limitations in accurately estimating the trend of MMP concerning their input parameters [24]. In contrast, the advent of ML has provided various robust algorithms in problems involving regression/classification. Consequently, considerable research studies dedicated to the precise modeling of MMP and the successful application of ML in this domain have been well documented.

The earliest application of ML on CO_2-EOR MMP can be traced back to 2003, when Huang et al. [25] first introduced ANN into this field. Subsequently, Emera and Sarma [26] employed the GA to optimize the MMP prediction processes. Following the year 2010, there has been a gradual increase in the adoption of ML algorithms and optimization techniques, accompanied by a significant expansion of the available dataset. Nowadays, the application of ML in predicting MMP has evolved into a more mature state. A comprehensive survey of the literature review in the field of CO_2-oil MMP estimation applying ML, spanning the period from 2003 to the present, is summarized in Table 1. Each reviewed paper is scrutinized and synthesized with respect to the employed algorithms, dataset size, data splitting methods, input variables, outcomes, our assessment, and a rating score. A paper deserving a high rating ought to exhibit certain characteristics, such as the following: a substantial dataset, typically comprising no fewer than 100 data points; a demonstration of effective model generalization without signs of overfitting, where the training dataset constitutes a maximum of 80% of the total data; and validation through empirical evidence derived from experimental and/or field data. Furthermore, a high-rated paper should demonstrate depth in result analysis, including a thorough examination of the outcomes in comparison to other existing models.

Figure 4 presents a statistical analysis from 56 research papers. It reveals a remarkable surge in the adoption of ML methodologies within this domain. ANN and GA have emerged as the most favored choices among many ML and optimization algorithms. ANNs, particularly RBFNN and MLP, are prominently employed. We have provided a separate categorization for RBFNN and MLP to afford a more detailed perspective on their individual utilization patterns.

Furthermore, an essential factor impacting the efficacy of ML models in MMP predictions is the size of the dataset. It is widely recognized that an inadequately sized dataset can lead to overfitting, potentially compromising the model's generalizability. A substantial proportion of the examined papers (64%) have datasets with fewer than 200 data points, with a noteworthy subset (21%) relying on datasets with fewer than 100 data points. This stark discrepancy in dataset size necessitates critically examining the quality and robustness of models trained on such limited data. Therefore, it becomes paramount to consider the trade-offs between the advantages of ML applications and the constraints posed by data scarcity in the context of MMP prediction.



(**a**) (**b**) (**c**)

Figure 4. (**a**). Rise of ML application papers in MMP prediction; (**b**) occurrence of different ML algorithms; (**c**) distribution of dataset size.

As summarized in Table 1, the most dominant parameters affecting pure CO_2 MMP are reservoir temperature, the molecular weight of C_{5+} or C_{7+}, the mole fraction of volatile oil elements, and the mole fraction of intermediate oil elements. Meanwhile, for impure CO_2 MMP, additional parameters such as the mole fraction of gas, including C_1 to C_4, CO_2, N_2, and H_2S are also considered. Some studies included volatile oil components (C_1 and N_2) as well.

A more rigorous way to investigate the impact of each input variable involves conducting sensitivity analysis, a widely employed way to analyze the effect of each input parameter on model predictions. It plays a crucial role in model interpretation, validation, and feature selection, ultimately improving the trustworthiness and transparency of machine learning models. Methods like SHAP (Shapley Additive exPlanations) and relevancy factors are commonly used for sensitivity analysis. Nevertheless, few existing studies [27,28] have performed a sensitivity analysis, while the majority of research only compares their models with experimental and/or empirical results. Future research endeavors should allocate attention toward sensitivity analysis, thereby enhancing the completeness and credibility of machine learning studies.

Table 1. Summary of ML application on CO_2-EOR MMP.

Authors	Methods	Dataset	Splitting	Inputs	Results	Evaluation	Limitations	Rating *
Huang et al. [25]	ANN	N/A	N/A	Pure CO_2 (T_R, x_{vol}, MW_{C5+}, x_{int}), impure CO_2 (y_{H2S}, y_{N2}, y_{CH4}, y_{SO2}, F_{imp})	ANN can predict MMP.	First applied ANN. ANN is better than other statistical models.	Need to separate pure CO_2 and impure CO_2.	7
Emera and Sarma [26]	GA	N/A	N/A	T_R, MW_{C5+}, $x_{vol}/(y_{C1} + y_{H2S} + y_{CO2} + y_{N2} + y_{C2\text{-}C4})$.	GA is best for predicting MMP and impurity factors.	First used GA. Limited input parameters (only 3 variables).	Pure CO_2. MW_{C7+} only up to 268.	7
Dehghani et al. [29]	GA	55	80% train + 20% test	T_R, T_C, MW_{C5+}, x_{vol}/x_{int}.	GA is better than conventional methods.	Can predict pure and impure CO_2. But limited input parameters and data points.	Limited input parameters and data points.	6

Table 1. *Cont.*

Authors	Methods	Dataset	Splitting	Inputs	Results	Evaluation	Limitations	Rating *
Shokir [22]	ACE	45	50% train + 50% test	T_R, MW_{C5+}, y_{CO2}, y_{H2S}, y_{N2}, y_{C1}, $y_{C2\text{-}C4}$, x_{C1+N2}, x_{int}	Can predict relatively accurate MMP for pure and impure CO_2.	Can predict pure and impure CO_2. But very limited data points. It may have overfitting.	valid only for C1, N_2, H_2S, and C2–C4 contents in the injected CO_2 stream.	6
Dehghani et al. [30]	ANN-GA	46	N/A	T_R, MW_{C5+}, y_{CO2}, y_{H2S}, y_{N2}, y_{C1}, $y_{C2\text{-}C4}$, x_{C1+N2}, x_{int}	GA-ANN is better than Shokir [22], Emera and Sarma [26].	It can be applied to both CO_2 and natural gas streams.	Limited data points and only ANN architecture is tested.	6
Nezhad et al. [31]	ANN	179	N/A	T_R, x_{vol}, MW_{C5+}, y_{CO2}, $y_{volatile}$, $y_{intermediate}$	ANN is acceptable.	Acceptable data points but not detailed explanations.	Local minima or overfitting	8
Shokrollahi et al. [32]	LSSVM	147	80% train + 10% test + 10% validate	T_R, x_{vol}, MW_{C5+}, y_{CO2}, y_{C1}, y_{H2S}, y_{N2}, $y_{C2\text{-}C5}$	First applied LSSVM.	It can be used for both pure and impure CO_2. Also applied outlier analysis.	Valid only for the impurity contents of C1, N_2, H_2S, and C_2–C_5.	8
Tatar et al. [33]	RBFN	147	80% train + 20% test	T_R, MW_{C5+}, y_{CO2}, y_{H2S}, y_{N2}, y_{C1}, $y_{C2\text{-}C5}$, $(x_{C1} + x_{N2})/(x_{C2\text{-}C4} + x_{H2S} + x_{CO2})$	Better than Emera and Sarma's model.	Compared with almost all available empirical correlations.	Limited data points	8
Zendehboudi et al. [34]	ANN-PSO	350	71% train + 29% test	T_R, x_{vol}, MW_{C5+}, y_{CO2}, y_{C1}, y_{H2S}, y_{N2}, $y_{C2\text{-}C4}$	ANN-PSO is best.	Though it has large datasets, but only suitable for fixed input parameters.	Only valid for specific conditions	8
Chen et al. [35]	ANN	83	70% train + 30% test	T_R, MW_{C5+}, x_{vol}, x_{int}, y_{CO2}, y_{H2S}, y_{C1}, and y_{N2}	ANN provides the least errors.	May have overfitting.	Small datasets	7
Asoodeh et al. [36]	CM (NN-SVR)	55	N/A	T_R, MW_{C5+}, x_{vol}/x_{int}, $y_{C2\text{-}C4}$, y_{CO2}, y_{H2S}, y_{C1}, and y_{N2}	CM is better than NN and SVR.	Limited data points and may have overfitting.	Small datasets	6
Rezaei et al. [37]	GP	43	N/A	T_R, MW_{C5+}, x_{vol}/x_{int}	GP provides the best estimation.	Limited data points and may have overfitting.	Small datasets and only consider pure CO_2.	6
Chen at al. [38]	GA-BPNN	85	75% train + 25% test	T_R, MW_{C7+}, x_{vol}, $x_{C5\text{-}C6}$, y_{CO2}, y_{H2S}, y_{N2}, y_{C1}, $y_{C2\text{-}C4}$, x_{int}	Both pure and impure CO_2, better than other correlations.	It can be applied to both pure and impure CO_2 but may have overfitting.	Limited data points. GA is time-consuming.	7
Ahmadi and Ebadi [39]	FL	59	93% train + 7% test	T_R, MW_{C5+}, x_{vol}/x_{int}, T_C	The curve shape membership function has the lowest error.	Limited data points and a high possibility of overfitting.	Only four experimental results for testing.	6
Sayyad et al. [40]	ANN-PSO	38	N/A	T_R, x_{vol}, MW_{C5+}, y_{CO2}, y_{H2S}, y_{C1}, y_{N2}, $y_{C2\text{-}C5}$	Better than Emera and Sarma, Shokir.	Only valid for fixed inputs.	Limited data points	6
Zargar et al. [41]	GRNN	N/A	N/A	T_R, MW_{C5+}, x_{vol}/x_{int}, $y_{C2\text{-}C4}$, y_{CO2}, y_{H2S}, y_{C1}, and y_{N2}.	GRNN is an efficient computational structure. GA reduces the runs of GRNNs.	Though compared with most known correlations, but unknown about the data source.	Need more information about the treatment of data.	6
Bian et al. [42]	SVR-GA	150	67% train + 23% test and 83% train + 17% test	T_R, MW_{C5+}, x_{vol}, y_{CO2}, y_{H2S}, y_{C1}, y_{N2}.	Better than other empirical correlations.	Can be used for pure and impure CO_2 and low AARD.	Separate pure and impure CO_2.	9
Hemmati-Sarapardeh et al. [43]	MLP	147	70% train + 15% test + 15% validate	T_R, T_C, MW_{C5+}, x_{vol}/x_{int}	Can predict both pure and impure CO_2.	Simple and reliable.	Treatment of inputs may be too simple.	8
Zhong and Carr [44]	MKF-SVM	147	90% train + 10% test	T_R, T_C, MW_{C5+}, x_{vol}/x_{int}	The mixed kernel provides better performance.	Treatment of inputs may be too simple.	Did not consider the effect of N_2, H_2S.	8
Fathinasab and Ayatollahi [45]	GP	270	80% train + 20% test	T_R, T_{cm}, MW_{C5+}, x_{vol}/x_{int}	GP provides the best prediction.	Relatively large datasets but may simplify the inputs.	AARE is a little high (11.76%).	7

Table 1. *Cont.*

Authors	Methods	Dataset	Splitting	Inputs	Results	Evaluation	Limitations	Rating *
Alomair and Garrouch [46]	GRNN	113	80% train + 20% test	T_R, MW_{C5+}, MW_{C7+}, x_{C1}, x_{C2}, x_{C3}, x_{C4}, x_{C5}, x_{C6}, x_{C7+}, x_{CO2}, x_{H2S}, x_{N2}.	GRNN is better than five empirical correlations.	Too many inputs and no further comparison between GRNN and other ML methods.	Does not consider the purity of CO_2.	7
Karkevandi-Talkhooncheh et al. [47]	ANFIS	270	80% train + 20% test	T_R, T_C, MW_{C5+}, x_{vol}, x_{int}	ANFIS-PSO is the best among the five optimization methods.	Very comprehensive comparison with available models and different optimizations.	Further applicability may be needed.	9
Ahamdi et al. [48]	GEP	N/A	N/A	T_R, T_{cm}, MW_{C5+}, x_{vol}/x_{int}	GEP is better than traditional correlations.	Unknown about datasets.	Further validation may be needed.	6
Karkevandi-Talkhooncheh et al. [49]	RBF-GA/ PSO/ICA/ ACO/DE	270	80% train + 20% test	T_R, MW_{C5+}, x_{vol}, $x_{C2\text{-}C4}$, y_{CO2}, y_{H2S}, y_{C1}, y_{N2}.	ICA-RBF is best.	Comparable large datasets. Five algorithms were applied.	Further applicability may be needed.	9
Tarybakhsh et al. [50]	SVR-GA, MLP, RBF, GRNN	135	92.5% train + 7.5% test	T_R, $MW_{C2\text{-}C6\,(OIL)}$, MW_{C7+}, SG_{C7+}, $MW_{C2\text{-}C6\,(GAS)}$, y_{CO2}, y_{H2S}, y_{C1}, y_{N2}.	SVT-GA is best.	Too many input parameters may cause a high possibility of overfitting.	The R^2 is as high as 0.999. Too perfect to be reliable.	6
Dong et al. [51]	ANN	122	82% train + 18% test	H_2S, CO_2, N_2, C_1, $C_2 \ldots C_{36+}$	ANN can be used to predict MMP.	Too many inputs. No dominant input selection.	Input variables were assumed based on the availability of data.	7
Hamdi and Chenxi [52]	ANFIS	48	73% train + 27% test	T_R, MW_{C5+}, x_{vol}, x_{int}	Gaussian MF is the best among the five MFs. ANFIS is better than ANN.	Though applied five MF but limited data points.	Limited data points and does not consider the existence of CO_2.	6
Khan et al. [53]	ANN, FN, SVM	51	70% train + 30% test	T_R, MW_{C7+}, x_{C1}, $x_{C2\text{-}C6}$, MW_{C2+}, x_{C2}	ANN is best.	Compared three methods but input parameters are overlapping.	Limited data points and does not consider the existence of CO_2.	6
Choubineh et al. [54]	ANN	251	75% train + 10% test + 15% validate	T_R, MW_{C5+}, x_{vol}/x_{int}, SG	ANN is best compared with empirical correlations.	Relatively large dataset. Use gas SG instead.	Further applicability may be needed.	8
Li et al. [55]	NNA, GFA, MLR, PLS	136	N/A	T_R, T_C, MW_{C5+}, x_{vol}/x_{int}, $y_{C2\text{-}C5}$, y_{CO2}, y_{H2S}, y_{C1}, y_{N2}.	ANN is best among both empirical and other algorithms.	Unclear about how to split the data.	Further applicability may be needed.	8
Hassan et al. [56]	ANN, RBF, GRNN, FL	100	70% train + 30% test	T_R, MW_{C7+}, $x_{C2\text{-}C6}$	RBF provides the highest accuracy.	Only three input parameters may simplify the model.	Does not consider the purity of CO_2 and the limited dataset.	7
Sinha et al. [57]	Linear SVM/KNN/ RF/ANN	N/A	67% train + 33% test	T_R, MW_{C7+}, MW_{Oil}, x_{C1}, x_{C2}, x_{C3}, x_{C4}, x_{C5}, x_{C6}, x_{C7+}, x_{CO2}, x_{H2S}, and x_{N2}.	Modified correlation with linear SVR and hybrid method with RF is best.	Only need oil composition and TR. Does not consider the purity of CO_2.	MMP range 1000–4900 psi.	7
Nait Amar and Zerabi [9]	SVR-ABC	201	87% train + 13% test	T_R, T_C, MW_{C5+}, x_{vol}/x_{int}, $x_{C2\text{-}C4}$	SVR-ABC is better SVR-TE.	The choice of inputs is limited	Limited comparison.	8
Dargahi-Zarandi et al. [58]	AdaBoost SVR, GDMH, MLP	270	67% train + 33% test	T_R, T_C, MW_{C5+}, x_{vol}, $x_{C2\text{-}C4}$, y_{CO2}, y_{H2S}, y_{C1}, y_{N2}.	AdaBoost SVR is best.	Create a 3-D plot for better visualization.	Further applicability was limited	9
Tian et al. [59]	BP-NN (GA, MEA, PSO, ABC, DA)	152	80% train + 20% test	T_R, MW_{C5+}, x_{C1}, x_{C2}, x_{C3}, x_{C4}, x_{C5}, x_{C6}, x_{C7+}, y_{CO2}, y_{H2S}, y_{N2}.	DA-BP has the highest accuracy.	Compared with empirical correlations and GA-SVR.	Too many input parameters may have overfitting.	8
Ekechukwu et al. [60]	GPR	137	90% train + 10% test	T_R, T_C, MW_{C5+}, x_{vol}/x_{int}	GPR has higher accuracy than other models.	Very comprehensive comparison. A larger dataset may be better.	Further validation with experiments may be needed.	8

Table 1. *Cont.*

Authors	Methods	Dataset	Splitting	Inputs	Results	Evaluation	Limitations	Rating *
Saeedi Dehaghani and Soleimani [61]	SGB, ANN, ANN-PSO, ANN-TLBO	144	75% train + 25% test	T_R, MW_{C5+}, x_{vol}, x_{int}, y_{CO2}, y_{C1}, y_{int}, y_{N2}.	PSO and TLBO can help improve the accuracy of the ANN model. SGB is better than ANN.	First applied SGB. Maybe compared with other optimization methods will be better.	Further validation with experiments may be needed.	8
Dong et al. [62]	FCNN	122	82% train + 18% test	x_{CO2}, x_{H2S}, x_{N2}, x_{C1}, x_{C2}, x_{C3}, x_{C4}, x_{C5}, x_{C6}, ..., x_{C36+}.	L2 regularization and Dropout can help reduce overfitting.	Alleviate overfitting but small datasets.	Small datasets.	7
Chen et al. [63]	SVM	147	80% train + 20% test	T_R, MW_{C7+}, x_{vol}, $x_{C2\text{-}C4}$, $x_{C5\text{-}C6}$, y_{CO2}, y_{HC}, y_{C1}, and y_{N2}.	POLY kernel is more accurate. MW_{C7+} and $x_{C5\text{-}C6}$ should not be considered.	Very complete and comprehensive. Includes optimization and evaluation.	More persuasive with a large dataset.	9
Ghiasi et al. [64]	ANFIS, AdaBoost-CART	N/A	90% train + 10% test	T_R, T_C, MW_{C5+}, x_{vol}/x_{int}, y_{CO2}, y_{H2S}, $y_{C1\text{-}C5}$, and y_{N2}	The novel AdaBoost-The CART model is the most reliable.	The size of the dataset is unknown. First one to use AdaBoost.	May have overfitting and validation is not strong.	7
Chemmakh et al. [65]	ANN, SVR-GA	147 (pure CO_2), 200 (impure CO_2)	NA	T_R, T_C, MW_{C5+}, x_{vol}/x_{int}	ANN and SVR-GA are reliable to use.	The novelty of work is not clear.	Only compared with empirical correlations.	7
Pham et al. [66]	FCNN	250	80% train + 20% test	T_R, x_{vol}/x_{int}, MW, y_{C1}, y_{C2+}, y_{CO2}, y_{H2S}, y_{N2}	Multiple FCN together with Early Stopping and K-fold cross validation has high prediction of MMP.	Applied deep learning—multiple FCN to predict MMP. Limited comparisons and validations.	Only compared with decision tree and random forest.	7
Haider et al. [67]	ANN	201	70% train + 30% test	T_R, MW_{C7+}, x_{CO2}, x_{C1}, x_{C2}, x_{C3}, x_{C4}, x_{C5}, x_{C6}, x_{C7}, y_{CO2}, y_{H2S}, y_{C1}, y_{N2}.	An empirical correlation is developed based on ANN.	Too many inputs and a high possibility of overfitting.	Need further validation with other reservoir fluid and injected gas.	7
Huang et al. [68]	CGAN-BOA	180	60% train + 20% test + 20% validate	T_R, MW_{C7+}, x_{CO2}, x_{C1}, x_{C2}, x_{C3}, x_{C4}, x_{C5}, x_{C6}, x_{C7+}, x_{N2}, y_{CO2}, y_{H2S}, y_{N2}, y_{C1}, y_{C2}, y_{C3}, y_{C4}, y_{C5}, y_{C6}, y_{C7+}.	CGAN-BOA and ANN are better than SVR-RBF and SVR-POLY.	Proved deep learning has the potential for predicting MMP.	May have overfitting problems given 21 input parameters.	8
He et al. [69]	GBDT-PSO	195	85% train + 15% test	T_R, x_{CO2}, x_{C1}, x_{C2}, x_{C3}, x_{C4}, x_{C5}, x_{C6}, x_{C7+}, x_{N2},	GBDT is better than LR, RR, RF, MLP.	Improved GBDT by using PSO. But not a comprehensive comparison.	Only GBDT was optimized. Other algorithms could also be tuned and compared.	7
Hou et al. [70]	GPR-PSO	365	80% train + 20% test	T_R, T_C, MW_{C5+}, x_{vol}/x_{int}, y_{CO2}, y_{H2S}, y_{C1}, $y_{C2\text{-}C5}$, y_{N2}.	GPR-PSO provides the highest accuracy.	Comprehensive comparison and large datasets.	The model was only validated with literature data.	9
Rayhani et al. [71]	SFS, SBS, SFFS, SBFS, LR, RFFI	812	80% train + 20% test	T_R, T_C, MW_{C7+}, MW_{gas}, x_{C5}, x_{C6}, $x_{C2\text{-}C6}$	SBFS provides the highest accuracy.	Large datasets. Comprehensive data selection and model comparison.	Further applicability with field data or commercial simulation was limited.	9
Shakeel et al. [72]	ANN, ANFIS	105	70% train + 30% test	T_R, MW_{C7+}, x_{vol}, $x_{C2\text{-}C4}$, $x_{C5\text{-}C6}$, y_{CO2}, y_{H2S}, y_{C1}, y_{HC}, y_{N2}.	ANN is better than ANFIS; the trainlm performs best.	Demonstrated good accuracy but lack of uncertainty analysis.	Limited dataset and only two ML algorithms were tested.	7
Shen et al. [73]	XGBoost, TabNet, KXGB, KTabNet	421	80% train + 20% test	T_R, MW_{C5+}, x_{vol}/x_{int}, y_{CO2}, y_{H2S}, y_{C1}, $y_{C2\text{-}C5}$, y_{HC}, and y_{N2}	KXGB is best. KTabNet can be used as an alternative.	Large datasets. Comprehensive model comparison. New insights into deep learning.	Need improvement of feature comprehensiveness.	9
Lv et al. [24]	XGBoost, CatBoost, LGBM, RF, deep MLN, DBN, CNN	310	80% train + 20% test	T_R, T_C, MW_{C5+}, x_{vol}/x_{int}	CatBoost outperforms than other AI techniques.	Comprehensive model comparison and evaluation. New insights into deep learning.	The accuracy depends on the databank. A larger dataset will be more robust.	9

Table 1. *Cont.*

Authors	Methods	Dataset	Splitting	Inputs	Results	Evaluation	Limitations	Rating *
Hamadi et al. [74]	MLP-Adam, SVR-RBF, XGBoost	193	84% train + 16% test	T_R, T_C, MW_{C5+}, x_{vol}/x_{int}	XGBoost provides the best prediction for both pure and impure CO_2.	Not comprehensive comparison and a limited dataset.	Limited dataset and only two ML algorithms were tested	7
Huang et al. [27]	1D-CNN, SHAP	193	NA	T_R, MW_{C7+}, x_{CO2}, x_{C1}, x_{C2}, x_{C3}, x_{C4}, x_{C5}, x_{C6}, x_{C7+}, x_{N2}, y_{CO2}, y_{H2S}, y_{N2}, y_{C1}, y_{C2}, y_{C3}, y_{C4}, y_{C5}, y_{C6}, y_{C7+}.	MMPs from the slim tube and rising bubble are different. 1D-CNN performs best.	It is novel in the SHAP application, but the comparison with other ML models is limited.	Further applicability with field data or commercial simulation was limited.	8
Al-Khafaji et al. [28]	MLR, SVR, DT, RF, KNN	147 (type 1), 197 (type 2), 28 (type 3)	80% train + 20% test	Type 1: T_R, MW_{C5+}, x_{vol}/x_{int}; Type 2: MW_{C7+}, x_{vol}, x_{int}, $x_{C5\text{-}C6}$, x_{C7+}, y_{CO2}, y_{H2S}, y_{N2}, y_{C1}, $y_{C2\text{-}C6}$, y_{C7+}.; Type 3: T_R, MW_{C6+}, x_{vol}, x_{int}, x_{C6+}, API, sp.gr, Pb.	KNN has the highest efficient accuracy and lowest complexity.	Have a broad range of data including both experimental and field data. Performed thorough comparisons.	Only pure CO_2.	9
Sinha et al. [75]	Light GBM	205	80% train + 20% test	T_R, MW_{C7+}, MW_{Oil}, x_{C1}, x_{C2}, x_{C3}, x_{C4}, x_{C5}, x_{C6}, x_{C7+}, x_{CO2}, x_{N2}, x_{H2S}	An expanded range is developed with Light GBM.	Compared with empirical and EOS correlations. First used Light GBM in MMP prediction.	Further applicability with field data or commercial simulation was limited.	8

*: On a scale of 1 to 10, a higher score indicates higher quality of the article.

4.2. Water-Alternating-Gas (WAG)

WAG injection, a widely adopted method in EOR techniques, cyclically injects water and gas, typically CO_2 or CO_2-hydrocarbon blends, to increase sweep efficiency and maximize oil recovery. Optimizing parameters such as the WAG ratio, duration of each cycle, and reservoir properties is pivotal for achieving favorable economic outcomes. The application of ML methods on WAG has been developed more recently. The earliest application of ML in WAG started in 2016; Hosseinzadeh Helaleh and Alizadeh [76] employed SVM together with three optimization methods, ACO, PSO, and GA, to predict fractional oil recovery. In 2018, Nait Amar et al. [77] used time-dependent multi-ANN to predict the total field oil production. Later on, Nait Amar and Zeraibi [78] successfully applied SVR to construct a dynamic proxy of a field in Algeria, complemented by genetic algorithms (GAs) for optimizing water-alternating CO_2 gas parameters. A more detailed summary is listed in Table 2. Figure 5 provides statistical analysis based on 26 papers. Similar to MMP, the most popular ML algorithm is ANN, and the most preferred optimization is GA.

Table 2. Summary for ML applications on WAG.

Authors	Methods	Dataset	Splitting	Objectives	Inputs	Results	Evaluations	Limitations	Rating *
Hosseinzadeh Helaleh and Alizadeh [76]	SVM (ACO, GA, PSO)	200	80% train + 20% test	Fractional oil recovery	RL_C, RL_D, N_{gAO}, N_{gGO}, M_{SWAG}, N_C, SGR, N_{Pe}, N_{SCon}, N_B, N_α, N_σ, λ^*_{Dx}, N_n, He	ACO has high accuracy and low computational time compared to ANN, GA, and PSO.	Evaluate with both experiments and simulations. Limited to a similar geological model.	Only has SVM model.	8
Le Van and Chon [79]	ANN	223 (simulation)	45% train + 20% test + 35% validation	Oil recovery factor, oil rate, GOR, accumulative CO_2 production, net CO_2 storage	Swi, kv/kh, WAG ratio, duration of each cycle	ANN models can support numerical simulation of CO_2-EOR projects. WAG ratio less than 1.5 is best.	Evaluated multiple objectives but only limited to ANN.	Only have simulation results as trained data.	8
Van and Chon [80]	ANN	263 (simulation)	50% train + 20% test + 30% validation	Oil recovery + net CO_2 storage + cumulative gaseous CO_2 production	Kv/Kh, WAG ratio, Sw, well distance between each injector, T	ANN can help estimate oil recovery and CO_2 storage. Injection cycle 25 is best.	Evaluate different WAG ratios but limited to ANN models only.	Only have simulation results as trained data.	7
Mohagheghia [81]	GA, PSO	2000 (simulation)	NA	NPV + incremental recovery factor	Water and gas injection rates, BHP of producers, cycle ratio, cycle time, injected gas composition, total WAG period	PSO is capable of optimizing WAG variables and projects at field scale.	First used GA in WAG at field scale. Evaluated with three case studies. Limited to specific geological models.	Only GA and PSO are evaluated. Specific to E-segment.	9
Nwachukwu, Jeong, Sun et al. [82]	XGBoost, MADS	1000 (simulation)	50% train + 50% test	Oil/water/gas production rates, well locations, NPV	Well x-coordinates, well y-coordinates, water/gas injection rates, well block ϕ/k, well block Swi	The new model combined XGBoost and MADS provided high accuracy.	Demonstrated with a case study in which underlying geology is uncertain. Limited to one model.	Only XGBoost is employed.	8
Nait Amar et al. [77]	ANN/GA, ACO	85	88% train + 12% test	Field oil production total	Gas/water injection rates, gas/water injection half-cycle, WAG ratio, and slug size	Both GA and ACO are highly effective in the optimization of the WAG process.	Demonstrated the application of a time-dependent proxy model for the WAG process. Without further application of the case study.	Restricted to specific geological models. Limited simulation runs	8
Belzareg et al. [83]	Regression, GDMH	4290	70% train + 30% test	Incremental recovery factor	kh, kv, API, gas gravity, water viscosity, solution GOR, WAG ratio, WAG cycle, land coefficient, reservoir pressure, PV of injected water, PV of injected gas	GMDH performed better in selecting effective input parameters and optimizing the model structure.	Novel approach but did not apply real field WAG pilot data to validate.	Limited to two ML methods.	8
Jaber et al. [84]	CCD	81	NA	Oil recovery	k, ϕ, kv/kh, cyclic length, BHP, WAG ratio, CO_2 slug size	The new proxy model can predict oil recovery. The optimum WAG ratio is 1.5.	Developed a new proxy model based on CCD, but limited to one model.	Limited data points and only from simulation runs.	7

Table 2. *Cont.*

Authors	Methods	Dataset	Splitting	Objectives	Inputs	Results	Evaluations	Limitations	Rating *
Menad and Noureddine [85]	MLP (LMA, BR, SCG) + NSGA-II	From 2010 to 2018	NA	FOPR, FWPR	Time, FWIR, FGIR, the value of the needed parameter at the previous time step	MLP-LMA has the highest accuracy and lowest computation time.	Developed a dynamic proxy model for multiple objectives. But limited to one geological model.	The database was generated based on multiple runs of the simulation.	8
Nait Amar and Zeraibi [9]	SVR, GA	75	NA	Field oil production total	Injection rates of water and gas, half-cycle injection time, WAG ratio, slug size, initialization time of the process	SVR-GA provides high accuracy and reasonable CPU time.	Established a dynamic proxy model based on SVR-GA, but no comparison with other algorithms.	Limited data points and only one model evaluated.	7
Yousef et al. [86]	ANN	8 years × 37 wells	85% train + 15% test	Oil/gas/water production rate, GOR, infill well location	Well trajectory data, well logs, seismic data, production and injection history, reservoir pressure, choke opening, and WHP history	Implementing ANN for top-down modeling can predict reservoir performance under WAG.	Can predict the reservoir performance 3 months ahead. But simplify the data gathering, modeling, and validation process.	Unknown about specific input data. No comparison with other models or field case studies.	6
Belazreg and Mahmood [87]	GDMH	177	70% train + 30% test	Incremental oil recovery factor	Rock type, WAG process type, reservoir horizontal permeability, API, oil viscosity, reservoir pressure and temperature, and hydrocarbon pore volume of injected gas	GDMH models can predict three WAG incremental recovery factors: sandstone immiscible gas injection, sandstone miscible gas injection, and carbonate miscible gas injection	Proved GDMH can model the WAG process and has good potential. More data and validation are needed to improve model robustness and applicability.	Limited published WAG pilot data.	8
You et al. [88]	ANN	820	80% train + 10% test + 10% validation	Oil recovery, CO_2 storage, and project NPV	Water injection time, CO_2 injection time, producer BHP, water injection rate	The ANN proxy model can help improve the prediction performance.	Could handle two or three objectives very well when a limited number of control parameters	Only suitable for limited input parameters.	8
You et al. [89]	Gaussian SVR-PSO	217	NA	Hydrocarbon recovery + CO_2 sequestration volume + NPV	FOPR × 2, gas cycle × 5, water cycle × 5	The proposed method can optimize the WAG process with high accuracy.	Nice sensitivity studies of CO_2 price and oil price on NPV. Limited comparison with other ML models.	Restricted to specific geological models.	8
Enab and Ertekin [90]	ANN	2000	80% train + 10% test + 10% validation	Production prediction, production schemes design, history matching	25 inputs including reservoir rock characteristics, initial conditions, oil composition, well design parameters, and injection strategy parameters	ANN provides a faster prediction for fish-bone structure in low permeability reservoirs.	Nice project design and economic analysis, but limited to ANN model only.	Limitations were	

Table 2. *Cont.*

Authors	Methods	Dataset	Splitting	Objectives	Inputs	Results	Evaluations	Limitations	Rating *
Afzali et al. [91]	GEP	96	67% train + 33% test	Recovery factor	Oil viscosity, gas/water injection rates, k, PVI, number of cycles	The developed model is successful when compared with experimental results.	Novelty in using GEP. The dataset is from mathematical correlation.	imposed by defining the range of each variable.	8
Lv et al. [92]	ANN-PSO	2100	70% train + 15% test + 15% validation	Oil production	So, Pi, k, ϕ, h, Pwf, water injection rate, water cut before gas flooding, gas injection rate, water injection volume, cycle time, water injection time, production rate, grid size	ANN-PSO provides a good model for parameter optimization of CO_2 WAG-EOR.	Routine procedures, not too much novelty in applying ANN-PSO.	Limited and less supportive dataset.	8
Nait Amar et al. [93]	MLP-LM, RBFNN-ACO/GWO	82	88% train + 12% test	Field oil production total	Water/gas injection rates, injection half-cycle, downtime, WAG ratio, gas slug size	MLP-LMA is best. The proxy model can significantly reduce simulation time and conserve high accuracy.	The application of GWO is novel. Limited runs and may have overfitting problems.	No comparison with other ML models.	7
Junyu et al. [94]	Gaussian-SVR	1400	NA	Cumulative oil production and cumulative CO_2 storage.	Water/gas cycle, producer BHP, water injection rate, etc. (91 variables in total)	Gaussian-SVR performs best.	Showed the possibility to design a CO_2-WAG project using as many inputs as possible.	Water cut is limited to 50%. Reservoir pressure must be higher than MMP.	8
Sun et al. [95]	SVR, MLNN, RSM	600	83% train + 17% test	Oil production, CO_2 storage, NPV.	Duration of CO_2 and water injection cycles, water injection rate, production well specifications, oil price, CO_2 price, etc. (62 parameters)	The MLNN model can handle problems with large input and output dimensions.	Compared three different methods. But only suitable for specific geological models.	Given the large number of input parameters, the dataset may not be large enough.	7
Huang et al. [96]	LSTM	5404	90% train + 10% test	Oil production, GOR, water cut	Daily liquid rate, daily oil/gas/water rate, GIR, WIR, reservoir pressure, WHFP, choke size of producers	The calculation time of LSTM is 864% less than the simulation, while the prediction error of the LSTM method is 261% less than the simulation.	The model is based on real reservoir data over 15 years. But limited to one ML model.	The average reservoir pressure must be between 3700–5400 psi.	8
Li et al. [97]	RF	216	70% train + 30% test	Cumulative oil production, CO_2 storage amount, CO_2 storage efficiency	CO_2-WAG period, CO_2 injection rate, water-gas ratio, reservoir properties, oil properties, depth, layer thickness, Soi, well operation	CO_2-WAG cycle time has a slight influence on oil production. Random forest can predict oil production and CO_2 storage.	Proved RF has high computation efficiency and accuracy in CO_2-WAG projects. But no comparison of different ML models.	Only one ML model is considered. No comparison with other models.	7

Table 2. *Cont.*

Authors	Methods	Dataset	Splitting	Objectives	Inputs	Results	Evaluations	Limitations	Rating *
Andersen et al. [98]	LSSVM—PSO/GA/GWO/GSA	2500	70% train + 15% test + 15% validation	Oil recovery factor	Water-oil and gas-oil mobility ratios, water-oil and gas-oil gravity numbers, reservoir heterogeneity factor, two hysteresis parameters, and water fraction	LSSVM with GWO or PSO performed better than GA or GSA.	Very detailed and thorough study. The dataset is relatively large. Some limitations of input parameters.	Small dataset and only one ML model is studied.	7
Singh et al. [99]	DNN-GA	2200	70/80% train + 30/20% test	Maximize oil recovery	Water injection rates, gas-to-water ratio, slug size	DNN-GA workflow can identify improved WAG parameters over the baseline recovery, with incremental increases of 0.5–2%.	Presents a novel workflow for WAG optimization using ML. Requires a large number of simulation runs (2200 here) to initially train DNN.	Several important parameters were not varied much.	9
Asante et al. [100]	LSTM	2345 × 3	80% train + 20% test	Oil production rate, oil recovery factor	Bottom-hole pressure at injector and producer, water and gas injection volumes, WAG cycle	LSTM can model complex time-series data without the use of the geological model.	Shows the ability of LSTM to perform time series analysis. But the input parameters are restricted.	Limited to optimizing WAG parameters.	7
Matthew et al. [101]	ANN-NSGA-II	68 + 97	NA	Maximize oil produced and CO_2 storage	Water and gas injection rate, half-cycle length, time step	The developed proxy model can predict both simple and complex models.	Developed a dynamic proxy model for multiple objectives. But the dataset size is limited.	Requires large amounts of quality field data.	7

*: On a scale of 1 to 10, a higher score indicates higher quality of the article.

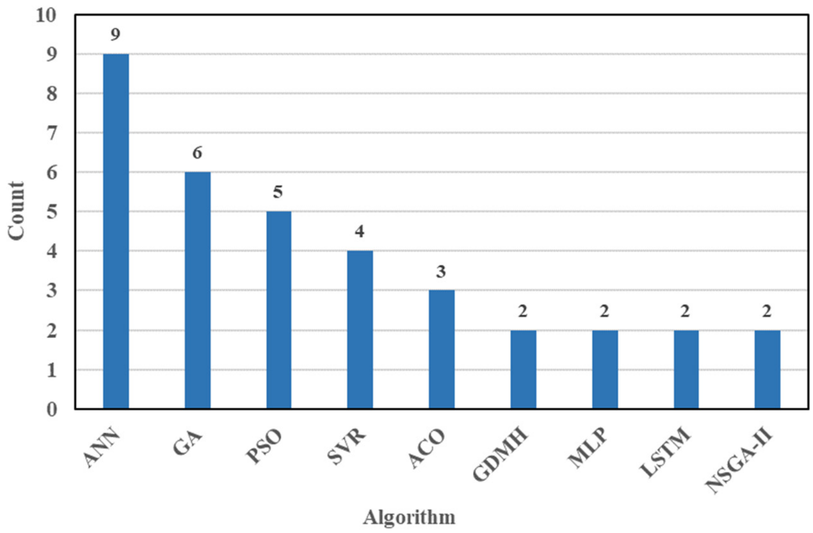


Figure 5. Occurrence of ML algorithms in WAG.

4.3. Well Placement Optimization (WPO)

WPO plays an essential role in reservoir management and development for many reasons. It can help maximize oil recovery and economic considerations (because drilling and maintaining wells is expensive). However, it has been considered one of the most challenging tasks due to the necessity of evaluating numerous computation scenarios to identify the optimal location for wells and achieve maximum production. The complexity of geological heterogeneities, such as variations in permeability and porosity, the existence of multiple facies, and stratigraphic and structural boundary conditions, requires extensive computational efforts. Furthermore, small changes in well locations can lead to significant changes in oil recovery prediction, making the optimization more challenging. Numerous simulations for hundreds or thousands of scenarios need to be run to make the best decision.

In recent years, studies suggesting the integration of ML approaches have been proposed in the literature as a potential solution. They hold the potential to accelerate computation processes, enabling the quicker attainment of accurate scenarios within numerical simulations. Despite the recognized importance of optimizing well placement, the investigations of CO_2 injector locations for optimal oil recovery and storage are relatively infrequent (Table 3). Most research is focused on waterflood injector selection [102].

Table 3. Summary of ML applications in well location optimization.

Authors	Methods	Dataset	Splitting	Objectives	Inputs	Results	Evaluations	Limitations	Rating *
Nwachukwu et al. [103]	XGBoost	200, 500, 1000	NA	Total profit, cumulative oil/gas produced, net CO_2 stored	Well-to-well pairwise connectivity, injector block k and ϕ, initial injector block saturations	Quick evaluation of well placement using well-to-well connectivity was successful with 1000 simulation runs and R^2 = 0.92.	No co-optimization of oil recovery and CO_2 storage, only ML proxy usage.	The dataset is from simulation runs. Only suitable for one geological model.	8
Selveindran et al. [104]	AdaBoost, RF, ANN	3000, 2000, 1000	70% train + 30% test	Incremental oil production	K, ϕ, PV, initial fluid saturation, pressure, time of flight, well-to-well distances, distance to the injector, injection rate, and injection depth	Stacked learner is better than an individual learner. ML helps rapidly identify the areas that are optimal for injection.	Detailed and comprehensive analysis, including posterior sampling.	Heavily rely on the geological model.	8

*: On a scale of 1 to 10, a higher score indicates higher quality of the article.

4.4. Oil Production/Recovery Factor

The recovery factor, defined as the ratio of produced oil to OOIP, is one of the most crucial success metrics for evaluating all EOR projects, as it determines how much incremental oil or ultimate oil is produced. Accurately predicting the recovery factor is challenging because it depends on diverse factors, including reservoir characteristics and heterogeneity, fluid properties, well design, injection condition, and the composition of the injected fluid. Reservoir simulations, together with laboratory experiments at reservoir conditions, can help predict the recovery factor. After that, a small-scale pilot test is conducted before undertaking larger-scale operations [105]. Although this approach may provide solutions to numerous problems, it is costly and time consuming. Therefore, ML methods emerge as more practical, affordable, rapid, and accurate alternatives.

Alternatively, ML methods have obtained popularity in predicting oil recovery. For example, Ahmadi et al. [106] applied LSSVM to predict the ultimate oil recovery factor of the miscible CO_2-EOR injection operations at different rock, fluid, and process conditions. Karacan [107] employed fuzzy logic to predict the recovery factors of the major past and existing U.S. field applications of miscible CO_2-EOR. Table 4 provides further information on ML applications on the CO_2-EOR recovery factor.

Table 4. Summary of ML applications on oil production/recovery factor.

Authors	Methods	Dataset	Splitting	Objectives	Inputs	Results	Evaluations	Limitations	Rating *
Ahmadi et al. [106]	LSSVM	46	80% train + 20% test	Oil recovery factor	BHP of injection well, CO_2 injection rate, CO_2 injection concentration, BHP of production well, oil production rate	The hybridization of LSSVM and BBD is statistically correct for predicting RF.	Provided the possibility of using ML and comparing it with commercial software. But limited dataset.	Small dataset and only suitable for similar oil reservoirs. Only valid for the same input parameters range.	7
Chen and Pawar [108]	MARS, SVR, RF	500, 250, 100	NA	Recovery factor	Thickness, depth, k, Sor, CO_2 injection rate, BHP of production well	MARS has the best performance.	Applied to 5 fields in Permian Basin and had good matches. Heavily relies on a base model and may not fully represent diverse ROZs.	Significant assumptions are made regarding uncertain parameters like residual oil saturation.	8
Karacan [107]	FL	24	83% train + 17% test	Recovery factor	Lithology, API, ϕ, k, HCPV, depth, net pay, Pi, well spacing, Sorw	FL provided a reasonably accurate prediction.	Though a small dataset, it provides the possibility of using ML in recovery factor prediction.	Too difficult to draw statistical conclusions from such a small dataset.	7
Iskandar and Kurihara [109]	AR, MLP, LSVM	3653 × 8 wells	40% train + 20% test + 40% validation	Oil, gas, and water production	ϕ, k, formation thickness, BHP, flow capacity, storage capacity	The AR model is best, with long and consistent forecast horizons across wells. LSTM performs well but has shorter forecast horizons. MLP has high variability and short forecast horizons.	First time series forecasting study. No model updating/retraining over time. Overall, it is a solid study.	Limited hyper-parameter tuning is performed. Only three models were tested.	9

*: On a scale of 1 to 10, a higher score indicates higher quality of the article.

4.5. Multi-Objective Optimization

As the name indicates, multi-objective optimization optimizes multiple objections simultaneously, such as the oil recovery factor or cumulative oil production, CO_2 storage, and net present value (NPV). For each objective, running high-fidelity numerical models provides possible solutions to figure out the optimum. However, finding optimal solutions to all the objectives simultaneously is not always guaranteed since objectives can compete with each other. For example, to maximize oil recovery, more CO_2 may be needed, leading to higher oil production. However, this might also mean that more CO_2 is used, potentially increasing the project's cost, which will also adversely affect the project NPV [110]. This requires sophisticated optimization techniques to identify solutions that balance these objectives, considering all the constraints involved in the problem. Therefore, ML techniques outperform other methods as an effective, reliable, and stable workflow to co-optimize crude oil recovery, CO_2 sequestration, NPV, and related factors.

Given the complexity of multi-objective optimization, the application of ML on CO_2-EOR is very limited (Table 5) and is strongly restricted by the geological model. Once the reservoir characteristics have changed, the model must be rebuilt and retrained. The development of the ML and optimization workflow is challenging and requires more effort in different oil and gas fields.

Table 5. Summary of ML applications on multi-objective optimizations.

Authors	Methods	Dataset	Splitting	Objectives	Inputs	Results	Evaluations	Limitations	Rating *
Ampomah et al. [111]	GA	NA	NA	Oil recover + CO_2 storage	NA	The proxy models to determine the optimal operational parameters, including injection/production rates, pressure, and WAG cycles.	First used proxy models and GA to optimize oil recovery and CO_2 storage simultaneously. But relies heavily on having an accurate reservoir mode.	Optimal parameters are specific to this reservoir—and not necessarily generalizable.	7
You et al. [112]	RBFNN	160	N/A	Cumulative oil production + CO_2 storage + NPV	water cycle, gas cycle, BHP of producer, water injection rate	The proxy model is built based on RBFNN for optimization.	The overall prediction is acceptable, but the CO_2 storage prediction is much higher.	The CO_2 storage optimization is 18% higher than the baseline.	7
You et al. [113]	ANN-PSO	820 (numerical model)	80% train + 10% test + 10% validation	Cumulative oil production + CO_2 storage + NPV	water cycle, gas cycle, BHP of producer, water injection rate	The optimization study showed promising results for multiple objectives.	Developed a novel hybrid optimization for multiple objective functions. But only validated with field case.	Only four input parameters are considered.	7
Vo Thanh et al. [114]	ANN-PSO	351 (numerical model)	80% train + 10% test + 10% validation	Cumulative oil production + cumulative CO_2 storage +cumulative CO_2 retained	ϕ, k, Sorg, Sorw, BHP of producer, CO_2 injection rate	ANN can forecast the performance of CO_2 EOR and storage in a residual oil zone.	The ANN provides R^2 of 0.99 and MSE of less than 2%, but the application in other types of reservoirs is questionable.	Case specific.	7

*: On a scale of 1 to 10, a higher score indicates the higher quality of the article.

4.6. PVT Properties

For any CO_2-flooding project, it is imperative to comprehend the intricate physical and chemical interactions between CO_2 and the reservoir oil, even when primarily exploring recovery potential. Laboratory investigations and the utilization of available modeling or correlation packages serve as viable methods for analyzing the influence of CO_2 on the physical properties of oil. Nonetheless, conducting a comprehensive laboratory study to obtain an extensive dataset is costly and time consuming. Furthermore, the available correlation packages are limited in their applicability, rendering them unsuitable for many scenarios.

ML is being increasingly harnessed for tasks such as predicting CO_2 solubility and interfacial tension (IFT), as briefly presented in Table 6. Intriguingly, a majority of the studies incorporated the same dataset sourced from Emera and Sarma [115]. Given the relatively small dataset size comprising only 106 data points, the risk of overfitting looms large, casting doubt on the accuracy and generalizability of their ML models. It is evident that a larger and more diverse dataset is required to facilitate a deeper comprehension of the performance of ML techniques in this context.

Table 6. Summary of ML application on PVT properties.

Authors	Methods	Dataset	Splitting	Objectives	Inputs	Results	Evaluations	Limitations	Rating *
Emera and Sarma [115]	GA	106 (dead oil), 74 (live oil)	NA	CO_2 solubility, oil swelling factor, CO_2-oil density, and viscosity	API, Ps, T, MW	The GA-base correlations provided the highest accuracy.	First applied GA in CO_2-oil properties prediction. Will be more helpful if a full dataset is provided.	Validated over a certain data range. May not be reliable if it is out of data range.	8
Rostami et al. [116]	ANN, GEP	106 (dead oil), 74 (live oil)	80% train + 20% test	CO_2 solubility	Ps, T, MW, γ, Pb	GEP is more accurate than ANN for dead oil.	Compared with several empirical methods. More comparisons between ML models will be more persuasive.	Limited dataset on live oil.	8
Rostami et al. [117]	LSSVM	106 (dead oil), 74 (live oil)	70% train + 15% test + 15% validation	CO_2 solubility	Ps, T, MW, γ	LSSVM showed higher accuracy compared to previous empirical correlations.	More rigorous validation against experimental data equations of state models would be useful.	Only a few literature models were compared.	7
Mahdaviara et al. [118]	MLP, RBF (GA, DE, FA), GMDH	NA	NA	CO_2 solubility	Ps, T, MW, γ, Pb	MLP-LM and MLP-SCG are better at predicting solubility. GMDH is better than LSSVM.	Compared with various models and optimization methods. But unknown for the dataset.	Not known for the dataset.	8
Hamadi et al. [74]	MLP-Adam, SVR-RBF, XGBoost	105 (dead oil), 74 (live oil)	80% train + 20% test	CO_2 solubility, IFT	Ps, T, MW, γ, Pb	SVR-RBF provided the best accuracy.	Limited comparisons between different models.	Given the year that this paper was published, the dataset is small.	7

*: On a scale of 1 to 10, a higher score indicates higher quality of the article.

4.7. CO_2-Foam Flooding

The implementation of CO_2 injection in Enhanced Oil Recovery (EOR) demonstrates significant potential, but it is accompanied by inherent limitations, including suboptimal sweep efficiency, asphaltene precipitation, and the corrosion of well infrastructure. In response to these challenges, the utilization of CO_2 foam has emerged as a promising strategy to enhance the effectiveness of CO_2-EOR flooding. Foams offer distinct advantages, primarily due to their elevated viscosities compared to pure gases, a property that equips foams with the capability to displace oil from reservoir formations more efficiently [119]. Furthermore, by obstructing highly permeable pore pathways, foams redirect displaced fluids toward unswept reservoir regions, thereby improving both the sweep efficiency and the storage capacity of CO_2 within the reservoir matrix. While ML models have found extensive applications in EOR research, their application in the context of CO_2 foam is still in its nascent stages, and the existing body of literature on this subject remains limited, as evidenced in Table 7.

Table 7. Summary of ML application on CO_2-foam EOR.

Authors	Methods	Dataset	Splitting	Objectives	Inputs	Results	Evaluations	Limitations	Rating *
Moosavi et al. [120]	MLP, RBF (GA, COA)	214	80% train + 20% test; 75% train + 25% test; 90% train + 10% test	Oil flow rate and recovery factor	Surfactant kind, ϕ, K, PV of core, Soi, injected foam PV	Both MLP and RBF provide high accuracy with R^2 up to 0.99.	The earliest research on CO_2-foam EOR. Only focus on laboratory data.	Only studied two methods, and there was no comparison among other ML algorithms.	8
Raha Moosavi et al. [121]	RBF (TLBO, PSO, GA, ICA)	214	80% train + 20% test	Oil flow rate and recovery factor	Surfactant kind, ϕ, K, PV of core, Soi, injected foam PV	RBF-TLBO provides the highest accuracy.	Proved ML can provide high accuracy (R^2 can reach 0.999), but is only limited to coreflood.	Limited to laboratory experiments.	8
Iskandarov et al. [119]	DT, RF, ERT, GB, XGBoost, ANN	145	70% train + 30% test	Surfactant stabilized CO_2 apparent foam viscosity	Shear rate, Darcy velocity, surfactant concentration, salinity, foam quality, T, and pressure	ML can provide reliable prediction, and ANN provides the highest accuracy.	Proved ML can predict for both bulk and sandstone formation under various conditions.	The dataset size is relatively small and may have overfitting.	8
Khan et al. [122]	XGBoost	200	70% train + 30% test	Oil recovery factor	Foam type, Soi, total PV tested, ϕ, K, injected foam PV	XGBoost can provide high accuracy.	Proved XGBoost can be used for CO_2-foam. Limited to laboratory data.	Only one ML is applied. No other comparisons.	7
Vo Thanh et al. [123]	GRNN, CFNN-LM, CFNN-BR, XGBoost	260	70% train + 30% test	Oil recovery factor	IOIP, TPVT, ϕ, K, injected foam PV	Porosity is the most significant parameter. GRNN has the highest accuracy.	Comprehensive and detailed description.	Limited to laboratory experiments.	9

*: On a scale of 1 to 10, a higher score indicates the higher quality of the article.

5. Benefits and Limitations of ML

ML exhibits high efficiency when compared with conventional reservoir simulators. Typically, these simulators are performed on 3D grids comprising one million to several billion cells. Computations tend to be time-consuming, imposing constraints on the feasibility of conducting multiple iterations. Consequently, this limitation reduces the optimization potential for meticulous field development planning. A pivotal role of ML techniques is their capacity to speed up reservoir modeling computations. These models can predict time-dependent variables at 100 to 1000 times faster speeds than traditional simulators. This acceleration in computation velocity via ML methods maintains an equivalent level of functionality [11].

Furthermore, extensive research findings have proved the impressive performance of ML methods, consistently yielding accuracy levels exceeding 90% based on statistical quality assessments. This high degree of accuracy demonstrates the confidence in ML's reliability and portends a promising future within the oil and gas industry.

While the advantages of employing ML are widely acknowledged, it is imperative to recognize the associated limitations inherent in ML-based methodologies. A central challenge confronting researchers is obtaining authentic data from experimental and/or field sources. The limited availability of large datasets is also a concern, impacting both the training accuracy and the overall efficacy of the ML models. When faced with restricted data, researchers often use single-shot learning strategies, wherein models are pre-trained on similar datasets and subsequently refined through experience.

Overfitting is a prevalent issue in ML applications, primarily driven by insufficient training data and the absence of well-defined stopping criteria during training. In total, 12% of the reviewed research papers contain datasets with fewer than 100 data points, heightening the risk of overfitting. Addressing this problem may involve adjusting the model's structure, including weight modifications. However, it is important to recognize that such alterations can increase model complexity, potentially limiting its generalization beyond the specific dataset.

More efforts are needed to advance ML applications within the oil and gas industry. For instance, integrating knowledge from multiple disciplines, such as geology, reservoir engineering, and petrophysics, into ML models could enhance model accuracy and interpretability. Future endeavors may involve the development of hybrid models implementing ML techniques with physics-based methodologies. Another improvement is reducing data scarcity and heterogeneity, which requires concerted efforts to address bias and model generalization. Researchers could focus on deploying data augmentation techniques, employing transfer learning methodologies, and refining ML algorithms to handle sparse and noisy data effectively.

6. Conclusions

In this work, we have investigated and summarized the employment of ML methods in the application of CO_2-EOR from several areas: MMP, WAG, well location placement, oil production/recovery factor, multi-objective optimization, PVT properties, and CO_2 foam. We have listed the input parameters, objectives, data sources, results, evaluation, and rating for each area based on the data quality, ML process, and results analysis. The important highlights of this work are summarized by seven key points as follows:

- Most reports on model performance indicators are limited to the size of the data bank, with 12% of the investigated papers having a database of less than 100 data points, making it difficult to accurately assess the quality of the model over time or track its drift with new data;
- Regarding validation and verification, the CO_2-EOR has many reliable, dependable, and well-established techniques for verification and validation procedures for ML models; the research highlights several issues with current ML models, including scalability, validation and verification deficiencies, and an absence of published data regarding the establishment costs of ML models;
- Most CO_2-EOR research focused on MMP predictions and WAG design, with 56 out of 101 papers devoted to MMP prediction and 26 of 101 papers to WAG design; the applications in the recovery factor, well placement optimization, and PVT properties are limited;
- ANN is the most employed ML algorithm, and GA is the most popular optimization algorithm based on 101 reviewed papers. ANN has proven to be flexible enough to be implemented to build intelligent proxies; while oil and gas data are frequently characterized by noise, incompleteness, heterogeneity, and nonlinearity, ANNs exhibit superior capability in handling such diverse data types and can adeptly adapt to varying data distributions;
- ML algorithms have the potential to greatly reduce the computational cost and time to perform compositional simulation runs; however, ML applications for well placement and multi-objective optimizations in CO_2-EOR are very limited given the complexity of the problem. Furthermore, the reliability of coupled ML-metaheuristic paradigms based on reservoir simulation results needs further investigation;
- The application of ML in the oil and gas industry still requires further exploration and development. Future work can focus on integrating knowledge from multiple disciplines, such as geology, reservoir engineering, and petrophysics with ML models to enhance accuracy and interpretability; another focus area could be the development of hybrid models that implement ML techniques alongside physics-based methodologies, providing robust and reliable support;
- In summary, this study provides a comprehensive overview of the application of ML and optimization techniques in CO_2-EOR projects; our work significantly contributes to the advancement of knowledge in the field by providing a synthesis of the latest research; these methods have demonstrated their ability to improve the efficiency, production forecast, and economic viability of CO_2-EOR operations; the insights gained from this study provide valuable guidance for the future direction of ML applications in CO_2-EOR R&D (research and development) and deployment.

Author Contributions: Conceptualization, G.T. and X.D.; methodology, X.D.; validation, X.D. and G.T.; investigation, S.S.; resources, X.D. and S.S.; data curation, X.D.; writing—original draft preparation, X.D.; writing—review and editing, X.D. and G.T.; visualization, X.D.; supervision, G.T.; All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflicts of interest.

Nomenclature

AARD	Average absolute relative deviation
AARE	Average absolute relative error
ABC	Artificial bee colony
ACO	Ant colony optimization
ACE	Alternating conditional expectation
AR	Auto-regressive
ANN	Artificial Neural Network
ANFIS	Adaptive neuro-fuzzy inference system
BA	Bee algorithm
BOA	Bayesian optimization algorithm
BPNN	Backpropagation algorithm neural network
BR	Bayesian regularization
CatBoost	Categorical boosting
CCD	Central composite design
CFNN	Cascade forward neural network
CGAN	Conditional generative adversarial network
CM	Committee machine
CNN	Convolutional neural network
COA	Cuckoo optimization algorithm
CSO	Cuckoo search optimization
DA	Dragonfly algorithm
DBN	Deep belief network
DE	Differential evolution
DNN	Dense neural network
ERT	Extremely randomized trees
FCNN	Fully connected neural network
FGIR	Field gas injection rate
FL	Fuzzy logic
FN	Functional network
GA	Genetic algorithm
GB	Gradient boosting
GBDT	Gradient boosting decision tree
GBM	Gradient boost method
GEP	Gene expression programming
GFA	Genetic function approximation
GIR	Gas injection rate
GMDH	Group method of data handling
GP	Genetic programming
GPR	Gaussian process regression
GRNN	Generalized regression neural network
GSA	Gravitational search algorithm
GWO	Grey wolf optimization
He	Hurst exponent
HPSO	Hybrid particle swarm optimization
ICA	Imperialist competitive algorithm

KXGB	Knowledge-based XGB
LGBM	Light gradient boosting machine
LM	Levenberg—Marquardt
LR	Lasso regression
LSSVM	Least-squares support vector machine
LSTM	Long short-term memory
MADS	Mesh adaptive direct search
MARS	Multivariate Adaptive Regression Splines
MASRD	Mean absolute symmetric relative deviation
MEA	Mind evolutionary algorithm
MF	Membership function
MKF	Mixed kernels function
MLP	Multi-layer perceptron
MLR	Multiple linear regression
MLNN	Multi-layer neural networks
MOPSO	Multi-objective particle swarm optimization
MSE	Mean squared error
NNA	Neural network analysis
NPV	Net present value
NSGA-II	Non-dominated sorting genetic algorithm version II
PLS	Partial least squares
POLY	Polynomial function
PSO	Particle swarm optimization
RBFN	Radial-based function networks
RFFI	Random forest feature importance
RR	Ridge regression
RSM	Response surface models
SBFS	Sequential backward floating selection
SBS	Sequential backward selection
SCG	Scaled conjugate gradient
SFS	Sequential forward selection
SFFS	Sequential forward floating selection
SGB	Stochastic gradient boosting
SGR	Solution gas ratio
SHAP	Shapley Additive explanations
SVR	Support vector regression
SVM	Support vector machine
TLBO	Teaching learning-based optimization
TPVT	Total pore volume tested
WIR	Water injection rate
WHFP	Well head flow pressure
XGBoost	Extreme gradient boosting
$\lambda^*{}_{Dx}$	Effective correlation length

References

1. Green, D.W.; Willhite, P.G. *Enhanced Oil Recovery*; Henry, L., Ed.; Doherty Memorial Fund of AIME, Society of Petroleum Engineers: Richardson, TX, USA, 1998; Volume 6.
2. Yang, G.; Jiang, R.; Li, X.; Jiang, Y. Evaluation of Polymer Flooding Performance Using Water-Polymer Interference Factor for an Offshore Oil Field in Bohai Gulf: A Case Study. In Proceedings of the SPE Improved Oil Recovery Conference, Tulsa, OK, USA, 14–18 April 2018; SPE: Tulsa, OK, USA, 2018.
3. Yongmao, H.; Zenggui, W.; Binshan, J.; Yueming, C.; Xiangjie, L.; Petro, X. Laboratory Investigation of CO_2 Flooding. In Proceedings of the Nigeria Annual International Conference and Exhibition, Abuja, Nigeria, 2 August 2004; SPE: Abuja, Nigeria, 2004.
4. Cheraghi, Y.; Kord, S.; Mashayekhizadeh, V. Application of Machine Learning Techniques for Selecting the Most Suitable Enhanced Oil Recovery Method; Challenges and Opportunities. *J. Pet. Sci. Eng.* **2021**, *205*, 108761. [CrossRef]
5. Gozalpour, F.; Ren, S.R.; Tohidi, B. CO_2 EOR and Storage in Oil Reservoir. *Oil Gas Sci. Technol.* **2005**, *60*, 537–546. [CrossRef]
6. Hasan, M.M.F.; First, E.L.; Boukouvala, F.; Floudas, C.A. A Multi-Scale Framework for CO_2 Capture, Utilization, and Sequestration: CCUS and CCU. *Comput. Chem. Eng.* **2015**, *81*, 2–21. [CrossRef]

7. Ghoraishy, S.M.; Liang, J.T.; Green, D.W.; Liang, H.C. Application of Bayesian Networks for Predicting the Performance of Gel-Treated Wells in the Arbuckle Formation, Kansas. In Proceedings of the SPE Symposium on Improved Oil Recovery, Tulsa, OK, USA, 19–23 April 2008; SPE: Tulsa, OK, USA, 2008.
8. Liu, M.; Fu, X.; Meng, L.; Du, X.; Zhang, X.; Zhang, Y. Prediction of CO_2 Storage Performance in Reservoirs Based on Optimized Neural Networks. *Geoenergy Sci. Eng.* **2023**, *222*, 211428. [CrossRef]
9. Nait Amar, M.; Zeraibi, N. Application of Hybrid Support Vector Regression Artificial Bee Colony for Prediction of MMP in CO_2-EOR Process. *Petroleum* **2020**, *6*, 415–422. [CrossRef]
10. You, J.; Lee, K.J. Pore-Scale Numerical Investigations of the Impact of Mineral Dissolution and Transport on the Heterogeneity of Fracture Systems During CO_2-Enriched Brine Injection. *SPE J.* **2022**, *27*, 1379–1395. [CrossRef]
11. Ng, C.S.W.; Nait Amar, M.; Jahanbani Ghahfarokhi, A.; Imsland, L.S. A Survey on the Application of Machine Learning and Metaheuristic Algorithms for Intelligent Proxy Modeling in Reservoir Simulation. *Comput. Chem. Eng.* **2023**, *170*, 108107. [CrossRef]
12. Kumar, N.; Augusto Sampaio, M.; Ojha, K.; Hoteit, H.; Mandal, A. Fundamental Aspects, Mechanisms and Emerging Possibilities of CO_2 Miscible Flooding in Enhanced Oil Recovery: A Review. *Fuel* **2022**, *330*, 125633. [CrossRef]
13. Satter, A.; Thakur, G.C. *Integrated Petroleum Reservoir Management: A Team Approach*; PennWell Books: Tulsa, OK, USA, 1994.
14. Yang, G.; Li, X. Modified Peng-Robinson Equation of State for CO_2/Hydrocarbon Systems within Nanopores. *J. Nat. Gas Sci. Eng.* **2020**, *84*, 103700. [CrossRef]
15. Wang, L.; Yao, Y.; Luo, X.; Daniel Adenutsi, C.; Zhao, G.; Lai, F. A Critical Review on Intelligent Optimization Algorithms and Surrogate Models for Conventional and Unconventional Reservoir Production Optimization. *Fuel* **2023**, *350*, 128826. [CrossRef]
16. Holm, L.W.; Josendal, V.A. Mechanisms of Oil Displacement By Carbon Dioxide. *J. Pet. Technol.* **1974**, *26*, 1427–1438. [CrossRef]
17. Yellig, W.F.; Metcalfe, R.S. Determination and Prediction of CO_2 Minimum Miscibility Pressures (Includes Associated Paper 8876). *J. Pet. Technol.* **1980**, *32*, 160–168. [CrossRef]
18. Christiansen, R.L.; Haines, H.K. Rapid Measurement of Minimum Miscibility Pressure with the Rising-Bubble Apparatus. *SPE Reserv. Eng.* **1987**, *2*, 523–527. [CrossRef]
19. Rao, D.N.; Lee, J.I. Application of the New Vanishing Interfacial Tension Technique to Evaluate Miscibility Conditions for the Terra Nova Offshore Project. *J. Pet. Sci. Eng.* **2002**, *35*, 247–262. [CrossRef]
20. Alston, R.B.; Kokolis, G.P.; James, C.F. CO_2 Minimum Miscibility Pressure: A Correlation for Impure CO_2 Streams and Live Oil Systems. *Soc. Pet. Eng. J.* **1985**, *25*, 268–274. [CrossRef]
21. Orr, F.M.; Jensen, C.M. Interpretation of Pressure-Composition Phase Diagrams for CO_2/Crude-Oil Systems. *Soc. Pet. Eng. J.* **1984**, *24*, 485–497. [CrossRef]
22. Shokir, E.M.E.-M. CO_2–Oil Minimum Miscibility Pressure Model for Impure and Pure CO_2 Streams. *J. Pet. Sci. Eng.* **2007**, *58*, 173–185. [CrossRef]
23. Ahmadi, K.; Johns, R.T. Multiple-Mixing-Cell Method for MMP Calculations. *SPE J.* **2011**, *16*, 733–742. [CrossRef]
24. Lv, Q.; Zheng, R.; Guo, X.; Larestani, A.; Hadavimoghaddam, F.; Riazi, M.; Hemmati-Sarapardeh, A.; Wang, K.; Li, J. Modelling Minimum Miscibility Pressure of CO_2-Crude Oil Systems Using Deep Learning, Tree-Based, and Thermodynamic Models: Application to CO_2 Sequestration and Enhanced Oil Recovery. *Sep. Purif. Technol.* **2023**, *310*, 123086. [CrossRef]
25. Huang, Y.F.; Huang, G.H.; Dong, M.Z.; Feng, G.M. Development of an Artificial Neural Network Model for Predicting Minimum Miscibility Pressure in CO_2 Flooding. *J. Pet. Sci. Eng.* **2003**, *37*, 83–95. [CrossRef]
26. Emera, M.K.; Sarma, H.K. Use of Genetic Algorithm to Estimate CO_2–Oil Minimum Miscibility Pressure—A Key Parameter in Design of CO_2 Miscible Flood. *J. Pet. Sci. Eng.* **2005**, *46*, 37–52. [CrossRef]
27. Huang, C.; Tian, L.; Wu, J.; Li, M.; Li, Z.; Li, J.; Wang, J.; Jiang, L.; Yang, D. Prediction of Minimum Miscibility Pressure (MMP) of the Crude Oil-CO_2 Systems within a Unified and Consistent Machine Learning Framework. *Fuel* **2023**, *337*, 127194. [CrossRef]
28. Al-Khafaji, H.F.; Meng, Q.; Hussain, W.; Khudhair Mohammed, R.; Harash, F.; Alshareef AlFakey, S. Predicting Minimum Miscible Pressure in Pure CO_2 Flooding Using Machine Learning: Method Comparison and Sensitivity Analysis. *Fuel* **2023**, *354*, 129263. [CrossRef]
29. Dehghani, S.A.M.; Vafaie Sefti, M.; Ameri, A.; Kaveh, N.S. A Hybrid Neural-Genetic Algorithm for Predicting Pure and Impure CO_2 Minimum Miscibility Pressure. *Iran. J. Chem. Eng.* **2006**, *3*, 44–59.
30. Dehghani, S.A.M.; Sefti, M.V.; Ameri, A.; Kaveh, N.S. Minimum Miscibility Pressure Prediction Based on a Hybrid Neural Genetic Algorithm. *Chem. Eng. Res. Des.* **2008**, *86*, 173–185. [CrossRef]
31. Nezhad, A.B.; Mousavi, S.M.; Aghahoseini, S. Development of an Artificial Neural Network Model to Predict CO_2 Minimum Miscibility Pressure. *Nafta* **2011**, *62*, 105–108.
32. Shokrollahi, A.; Arabloo, M.; Gharagheizi, F.; Mohammadi, A.H. Intelligent Model for Prediction of CO_2—Reservoir Oil Minimum Miscibility Pressure. *Fuel* **2013**, *112*, 375–384. [CrossRef]
33. Tatar, A.; Shokrollahi, A.; Mesbah, M.; Rashid, S.; Arabloo, M.; Bahadori, A. Implementing Radial Basis Function Networks for Modeling CO_2-Reservoir Oil Minimum Miscibility Pressure. *J. Nat. Gas Sci. Eng.* **2013**, *15*, 82–92. [CrossRef]
34. Zendehboudi, S.; Ahmadi, M.A.; Bahadori, A.; Shafiei, A.; Babadagli, T. A Developed Smart Technique to Predict Minimum Miscible Pressure-Eor Implications. *Can. J. Chem. Eng.* **2013**, *91*, 1325–1337. [CrossRef]

35. Chen, G.; Wang, X.; Liang, Z.; Gao, R.; Sema, T.; Luo, P.; Zeng, F.; Tontiwachwuthikul, P. Simulation of CO_2-Oil Minimum Miscibility Pressure (MMP) for CO_2 Enhanced Oil Recovery (EOR) Using Neural Networks. *Energy Procedia* **2013**, *37*, 6877–6884. [CrossRef]
36. Asoodeh, M.; Gholami, A.; Bagheripour, P. Oil-CO_2 MMP Determination in Competition of Neural Network, Support Vector Regression, and Committee Machine. *J. Dispers. Sci. Technol.* **2014**, *35*, 564–571. [CrossRef]
37. Rezaei, M.; Eftekhari, M.; Schaffie, M.; Ranjbar, M. A CO_2-Oil Minimum Miscibility Pressure Model Based on Multi-Gene Genetic Programming. *Energy Explor. Exploit.* **2013**, *31*, 607–622. [CrossRef]
38. Chen, G.; Fu, K.; Liang, Z.; Sema, T.; Li, C.; Tontiwachwuthikul, P.; Idem, R. The Genetic Algorithm Based Back Propagation Neural Network for MMP Prediction in CO_2-EOR Process. *Fuel* **2014**, *126*, 202–212. [CrossRef]
39. Ahmadi, M.-A.; Ebadi, M. Fuzzy Modeling and Experimental Investigation of Minimum Miscible Pressure in Gas Injection Process. *Fluid Phase Equilib.* **2014**, *378*, 1–12. [CrossRef]
40. Sayyad, H.; Manshad, A.K.; Rostami, H. Application of Hybrid Neural Particle Swarm Optimization Algorithm for Prediction of MMP. *Fuel* **2014**, *116*, 625–633. [CrossRef]
41. Zargar, G.; Bagheripour, P.; Asoodeh, M.; Gholami, A. Oil-CO_2 Minimum Miscible Pressure (MMP) Determination Using a Stimulated Smart Approach. *Can. J. Chem. Eng.* **2015**, *93*, 1730–1735. [CrossRef]
42. Bian, X.-Q.; Han, B.; Du, Z.-M.; Jaubert, J.-N.; Li, M.-J. Integrating Support Vector Regression with Genetic Algorithm for CO_2-Oil Minimum Miscibility Pressure (MMP) in Pure and Impure CO_2 Streams. *Fuel* **2016**, *182*, 550–557. [CrossRef]
43. Hemmati-Sarapardeh, A.; Ghazanfari, M.-H.; Ayatollahi, S.; Masihi, M. Accurate Determination of the CO_2-Crude Oil Minimum Miscibility Pressure of Pure and Impure CO_2 Streams: A Robust Modelling Approach. *Can. J. Chem. Eng.* **2016**, *94*, 253–261. [CrossRef]
44. Zhong, Z.; Carr, T.R. Application of Mixed Kernels Function (MKF) Based Support Vector Regression Model (SVR) for CO_2—Reservoir Oil Minimum Miscibility Pressure Prediction. *Fuel* **2016**, *184*, 590–603. [CrossRef]
45. Fathinasab, M.; Ayatollahi, S. On the Determination of CO_2–Crude Oil Minimum Miscibility Pressure Using Genetic Programming Combined with Constrained Multivariable Search Methods. *Fuel* **2016**, *173*, 180–188. [CrossRef]
46. Alomair, O.A.; Garrouch, A.A. A General Regression Neural Network Model Offers Reliable Prediction of CO_2 Minimum Miscibility Pressure. *J. Pet. Explor. Prod. Technol.* **2016**, *6*, 351–365. [CrossRef]
47. Karkevandi-Talkhooncheh, A.; Hajirezaie, S.; Hemmati-Sarapardeh, A.; Husein, M.M.; Karan, K.; Sharifi, M. Application of Adaptive Neuro Fuzzy Interface System Optimized with Evolutionary Algorithms for Modeling CO_2-Crude Oil Minimum Miscibility Pressure. *Fuel* **2017**, *205*, 34–45. [CrossRef]
48. Ahmadi, M.A.; Zendehboudi, S.; James, L.A. A Reliable Strategy to Calculate Minimum Miscibility Pressure of CO_2-Oil System in Miscible Gas Flooding Processes. *Fuel* **2017**, *208*, 117–126. [CrossRef]
49. Karkevandi-Talkhooncheh, A.; Rostami, A.; Hemmati-Sarapardeh, A.; Ahmadi, M.; Husein, M.M.; Dabir, B. Modeling Minimum Miscibility Pressure during Pure and Impure CO_2 Flooding Using Hybrid of Radial Basis Function Neural Network and Evolutionary Techniques. *Fuel* **2018**, *220*, 270–282. [CrossRef]
50. Tarybakhsh, M.R.; Assareh, M.; Sadeghi, M.T.; Ahmadi, A. Improved Minimum Miscibility Pressure Prediction for Gas Injection Process in Petroleum Reservoir. *Nat. Resour. Res.* **2018**, *27*, 517–529. [CrossRef]
51. Dong, P.; Liao, X.; Chen, Z.; Chu, H. An Improved Method for Predicting CO_2 Minimum Miscibility Pressure Based on Artificial Neural Network. *Adv. Geo-Energy Res.* **2019**, *3*, 355–364. [CrossRef]
52. Hamdi, Z.; Chenxi, D. Accurate Prediction of CO_2 Minimum Miscibility Pressure Using Adaptive Neuro-Fuzzy Inference Systems. In Proceedings of the SPE Gas & Oil Technology Showcase and Conference, Dubai, United Arab Emirates, 21–23 October 2019.
53. Khan, M.R.; Kalam, S.; Khan, R.A.; Tariq, Z.; Abdulraheem, A. Comparative Analysis of Intelligent Algorithms to Predict the Minimum Miscibility Pressure for Hydrocarbon Gas Flooding. In Proceedings of the Abu Dhabi International Petroleum Exhibition & Conference, Abu Dhabi, United Arab Emirates, 11 November 2019; SPE: Abu Dhabi, United Arab Emirates, 2019.
54. Choubineh, A.; Helalizadeh, A.; Wood, D.A. Estimation of Minimum Miscibility Pressure of Varied Gas Compositions and Reservoir Crude Oil over a Wide Range of Conditions Using an Artificial Neural Network Model. *Adv. Geo-Energy Res.* **2019**, *3*, 52–66. [CrossRef]
55. Li, D.; Li, X.; Zhang, Y.; Sun, L.; Yuan, S. Four Methods to Estimate Minimum Miscibility Pressure of CO_2-Oil Based on Machine Learning. *Chin. J. Chem.* **2019**, *37*, 1271–1278. [CrossRef]
56. Hassan, A.; Elkatatny, S.; Abdulraheem, A. Intelligent Prediction of Minimum Miscibility Pressure (MMP) During CO_2 Flooding Using Artificial Intelligence Techniques. *Sustainability* **2019**, *11*, 7020. [CrossRef]
57. Sinha, U.; Dindoruk, B.; Soliman, M. Prediction of CO_2 Minimum Miscibility Pressure MMP Using Machine Learning Techniques. In Proceedings of the SPE Improved Oil Recovery Conference, Tulsa, OK, USA, 31 August–4 September 2020; SPE: Tulsa, OK, USA, 2020.
58. Dargahi-Zarandi, A.; Hemmati-Sarapardeh, A.; Shateri, M.; Menad, N.A.; Ahmadi, M. Modeling Minimum Miscibility Pressure of Pure/Impure CO_2-Crude Oil Systems Using Adaptive Boosting Support Vector Regression: Application to Gas Injection Processes. *J. Pet. Sci. Eng.* **2020**, *184*, 106499. [CrossRef]
59. Tian, Y.; Ju, B.; Yang, Y.; Wang, H.; Dong, Y.; Liu, N.; Ma, S.; Yu, J. Estimation of Minimum Miscibility Pressure during CO_2 Flooding in Hydrocarbon Reservoirs Using an Optimized Neural Network. *Energy Explor. Exploit.* **2020**, *38*, 2485–2506. [CrossRef]
60. Ekechukwu, G.K.; Falode, O.; Orodu, O.D. Improved Method for the Estimation of Minimum Miscibility Pressure for Pure and Impure CO_2–Crude Oil Systems Using Gaussian Process Machine Learning Approach. *J. Energy Resour. Technol.* **2020**, *142*, 123003. [CrossRef]

61. Saeedi Dehaghani, A.H.; Soleimani, R. Prediction of CO_2-Oil Minimum Miscibility Pressure Using Soft Computing Methods. *Chem. Eng. Technol.* **2020**, *43*, 1361–1371. [CrossRef]
62. Dong, P.; Liao, X.; Wu, J.; Zou, J.; Li, R.; Chu, H. A New Method for Predicting CO_2 Minimum Miscibility Pressure MMP Based on Deep Learning. In Proceedings of the SPE/IATMI Asia Pacific Oil & Gas Conference and Exhibition, Bali, Indonesia, 25 October 2020; SPE: Bali, Indonesia, 2020.
63. Chen, H.; Zhang, C.; Jia, N.; Duncan, I.; Yang, S.; Yang, Y. A Machine Learning Model for Predicting the Minimum Miscibility Pressure of CO_2 and Crude Oil System Based on a Support Vector Machine Algorithm Approach. *Fuel* **2021**, *290*, 120048. [CrossRef]
64. Ghiasi, M.M.; Mohammadi, A.H.; Zendehboudi, S. Use of Hybrid-ANFIS and Ensemble Methods to Calculate Minimum Miscibility Pressure of CO_2—Reservoir Oil System in Miscible Flooding Process. *J. Mol. Liq* **2021**, *331*, 115369. [CrossRef]
65. Chemmakh, A.; Merzoug, A.; Ouadi, H.; Ladmia, A.; Rasouli, V. Machine Learning Predictive Models to Estimate the Minimum Miscibility Pressure of CO_2-Oil System. In Proceedings of the Abu Dhabi International Petroleum Exhibition & Conference, Abu Dhabi, United Arab Emirates, 15–18 December 2021; SPE: Abu Dhabi, United Arab Emirates, 2021.
66. Pham, Q.C.; Trinh, T.Q.; James, L.A. Data Driven Prediction of the Minimum Miscibility Pressure (MMP) Between Mixtures of Oil and Gas Using Deep Learning. In *Proceedings of the Volume 1: Offshore Technology*; American Society of Mechanical Engineers: Cincinnati, OH, USA, 2021.
67. Haider, G.; Khan, M.A.; Ali, F.; Nadeem, A.; Abbasi, F.A. An Intelligent Approach to Predict Minimum Miscibility Pressure of Injected CO_2-Oil System in Miscible Gas Flooding. In Proceedings of the ADIPEC, Abu Dhabi, United Arab Emirates, 31 October–3 November 2022; SPE: Abu Dhabi, United Arab Emirates, 2022.
68. Huang, C.; Tian, L.; Zhang, T.; Chen, J.; Wu, J.; Wang, H.; Wang, J.; Jiang, L.; Zhang, K. Globally Optimized Machine-Learning Framework for CO_2-Hydrocarbon Minimum Miscibility Pressure Calculations. *Fuel* **2022**, *329*, 125312. [CrossRef]
69. He, Y.; Li, W.; Qian, S. Minimum Miscibility Pressure Prediction Method Based On PSO-GBDT Model. *Improv. Oil Gas Recovery* **2023**, *6*. [CrossRef]
70. Hou, Z.; Su, H.; Wang, G. Study on Minimum Miscibility Pressure of CO_2–Oil System Based on Gaussian Process Regression and Particle Swarm Optimization Model. *J. Energy Resour. Technol.* **2022**, *144*, 103002. [CrossRef]
71. Rayhani, M.; Tatar, A.; Shokrollahi, A.; Zeinijahromi, A. Exploring the Power of Machine Learning in Analyzing the Gas Minimum Miscibility Pressure in Hydrocarbons. *Geoenergy Sci. Eng.* **2023**, *226*, 211778. [CrossRef]
72. Shakeel, M.; Khan, M.R.; Kalam, S.; Khan, R.A.; Patil, S.; Dar, U.A. Machine Learning for Prediction of CO_2 Minimum Miscibility Pressure. In Proceedings of the SPE Middle East Oil and Gas Show and Conference, MEOS, Proceedings, Abu Dhabi, United Arab Emirates, 23–25 May 2023; Society of Petroleum Engineers (SPE): Abu Dhabi, United Arab Emirates, 2023.
73. Shen, B.; Yang, S.; Gao, X.; Li, S.; Yang, K.; Hu, J.; Chen, H. Interpretable Knowledge-Guided Framework for Modeling Minimum Miscible Pressure of CO_2-Oil System in CO_2-EOR Projects. *Eng. Appl. Artif. Intell.* **2023**, *118*, 105687. [CrossRef]
74. Hamadi, M.; El Mehadji, T.; Laalam, A.; Zeraibi, N.; Tomomewo, O.S.; Ouadi, H.; Dehdouh, A. Prediction of Key Parameters in the Design of CO_2 Miscible Injection via the Application of Machine Learning Algorithms. *Eng* **2023**, *4*, 1905–1932. [CrossRef]
75. Sinha, U.; Dindoruk, B.; Soliman, M. Physics Guided Data-Driven Model to Estimate Minimum Miscibility Pressure (MMP) for Hydrocarbon Gases. *Geoenergy Sci. Eng.* **2023**, *224*, 211389. [CrossRef]
76. Hosseinzadeh Helaleh, A.; Alizadeh, M. Performance Prediction Model of Miscible Surfactant-CO_2 Displacement in Porous Media Using Support Vector Machine Regression with Parameters Selected by Ant Colony Optimization. *J. Nat. Gas Sci. Eng.* **2016**, *30*, 388–404. [CrossRef]
77. Nait Amar, M.; Zeraibi, N.; Redouane, K. Optimization of WAG Process Using Dynamic Proxy, Genetic Algorithm and Ant Colony Optimization. *Arab. J. Sci. Eng.* **2018**, *43*, 6399–6412. [CrossRef]
78. Amar, M.N.; Zeraibi, N.; Jahanbani Ghahfarokhi, A. Applying Hybrid Support Vector Regression and Genetic Algorithm to Water Alternating CO_2 Gas EOR. *Greenh. Gases Sci. Technol.* **2020**, *10*, 613–630. [CrossRef]
79. Le Van, S.; Chon, B.H. Evaluating the Critical Performances of a CO_2–Enhanced Oil Recovery Process Using Artificial Neural Network Models. *J. Pet. Sci. Eng.* **2017**, *157*, 207–222. [CrossRef]
80. Van, S.L.; Chon, B.H. Effective Prediction and Management of a CO_2 Flooding Process for Enhancing Oil Recovery Using Artificial Neural Networks. *J. Energy Resour. Technol.* **2018**, *140*, 032906. [CrossRef]
81. Mohagheghian, E.; James, L.A.; Haynes, R.D. Optimization of Hydrocarbon Water Alternating Gas in the Norne Field: Application of Evolutionary Algorithms. *Fuel* **2018**, *223*, 86–98. [CrossRef]
82. Nwachukwu, A.; Jeong, H.; Sun, A.; Pyrcz, M.; Lake, L.W. Machine Learning-Based Optimization of Well Locations and WAG Parameters under Geologic Uncertainty. In Proceedings of the SPE Improved Oil Recovery Conference, Tulsa, OK, USA, 14–18 April 2018; SPE: Tulsa, OK, USA, 2018.
83. Belazreg, L.; Mahmood, S.M.; Aulia, A. Novel Approach for Predicting Water Alternating Gas Injection Recovery Factor. *J. Pet. Explor. Prod. Technol.* **2019**, *9*, 2893–2910. [CrossRef]
84. Jaber, A.K.; Alhuraishawy, A.K.; AL-Bazzaz, W.H. A Data-Driven Model for Rapid Evaluation of Miscible CO_2-WAG Flooding in Heterogeneous Clastic Reservoirs. In Proceedings of the SPE Kuwait Oil & Gas Show and Conference, Mishref, Kuwait, 13–16 October 2019; SPE: Mishref, Kuwait, 2019.
85. Menad, N.A.; Noureddine, Z. An Efficient Methodology for Multi-Objective Optimization of Water Alternating CO_2 EOR Process. *J. Taiwan Inst. Chem. Eng.* **2019**, *99*, 154–165. [CrossRef]

86. Yousef, A.M.; Kavousi, G.P.; Alnuaimi, M.; Alatrach, Y. Predictive Data Analytics Application for Enhanced Oil Recovery in a Mature Field in the Middle East. *Pet. Explor. Dev.* **2020**, *47*, 393–399. [CrossRef]
87. Belazreg, L.; Mahmood, S.M. Water Alternating Gas Incremental Recovery Factor Prediction and WAG Pilot Lessons Learned. *J. Pet. Explor. Prod. Technol.* **2020**, *10*, 249–269. [CrossRef]
88. You, J.; Ampomah, W.; Sun, Q. Development and Application of a Machine Learning Based Multi-Objective Optimization Workflow for CO_2-EOR Projects. *Fuel* **2020**, *264*, 116758. [CrossRef]
89. You, J.; Ampomah, W.; Morgan, A.; Sun, Q.; Huang, X. A Comprehensive Techno-Eco-Assessment of CO_2 Enhanced Oil Recovery Projects Using a Machine-Learning Assisted Workflow. *Int. J. Greenh. Gas Control.* **2021**, *111*, 103480. [CrossRef]
90. Enab, K.; Ertekin, T. Screening and Optimization of CO_2-WAG Injection and Fish-Bone Well Structures in Low Permeability Reservoirs Using Artificial Neural Network. *J. Pet. Sci. Eng.* **2021**, *200*, 108268. [CrossRef]
91. Afzali, S.; Mohamadi-Baghmolaei, M.; Zendehboudi, S. Application of Gene Expression Programming (GEP) in Modeling Hydrocarbon Recovery in WAG Injection Process. *Energies* **2021**, *14*, 7131. [CrossRef]
92. Lv, W.; Tian, W.; Yang, Y.; Yang, J.; Dong, Z.; Zhou, Y.; Li, W. Method for Potential Evaluation and Parameter Optimization for CO_2-WAG in Low Permeability Reservoirs Based on Machine Learning. *IOP Conf. Ser. Earth Environ. Sci.* **2021**, *651*, 032038. [CrossRef]
93. Nait Amar, M.; Jahanbani Ghahfarokhi, A.; Ng, C.S.W.; Zeraibi, N. Optimization of WAG in Real Geological Field Using Rigorous Soft Computing Techniques and Nature-Inspired Algorithms. *J. Pet. Sci. Eng.* **2021**, *206*, 109038. [CrossRef]
94. Junyu, Y.; William, A.; Qian, S. Optimization of Water-Alternating-CO_2 Injection Field Operations Using a Machine-Learning-Assisted Workflow. In Proceedings of the SPE Reservoir Simulation Conference, Online, 26 October–1 November 2021; SPE: Abu Dhabi, United Arab Emirates, 2021.
95. Sun, Q.; Ampomah, W.; You, J.; Cather, M.; Balch, R. Practical CO_2—WAG Field Operational Designs Using Hybrid Numerical-Machine-Learning Approaches. *Energies* **2021**, *14*, 1055. [CrossRef]
96. Huang, R.; Wei, C.; Li, B.; Yang, J.; Wu, S.; Xu, X.; Ou, Y.; Xiong, L.; Lou, Y.; Li, Z.; et al. Prediction and Optimization of WAG Flooding by Using LSTM Neural Network Model in Middle East Carbonate Reservoir. In Proceedings of the Abu Dhabi International Petroleum Exhibition & Conference, Abu Dhabi, United Arab Emirates, 15–18 November 2021; SPE: Abu Dhabi, United Arab Emirates, 2021.
97. Li, H.; Gong, C.; Liu, S.; Xu, J.; Imani, G. Machine Learning-Assisted Prediction of Oil Production and CO_2 Storage Effect in CO_2-Water-Alternating-Gas Injection (CO_2-WAG). *Appl. Sci.* **2022**, *12*, 10958. [CrossRef]
98. Andersen, P.Ø.; Nygård, J.I.; Kengessova, A. Prediction of Oil Recovery Factor in Stratified Reservoirs after Immiscible Water-Alternating Gas Injection Based on PSO-, GSA-, GWO-, and GA-LSSVM. *Energies* **2022**, *15*, 656. [CrossRef]
99. Singh, G.; Davudov, D.; Al-Shalabi, E.W.; Malkov, A.; Venkatraman, A.; Mansour, A.; Abdul-Rahman, R.; Das, B. A Hybrid Neural Workflow for Optimal Water-Alternating-Gas Flooding. In Proceedings of the SPE Reservoir Characterization and Simulation Conference and Exhibition, Abu Dhabi, United Arab Emirates, 24–26 January 2023; SPE: Abu Dhabi, United Arab Emirates, 2023.
100. Asante, J.; Ampomah, W.; Carther, M. Forecasting Oil Recovery Using Long Short Term Memory Neural Machine Learning Technique. In Proceedings of the SPE Western Regional Meeting, Anchorage, AK, USA, 22–25 May 2023; SPE: Anchorage, AK, USA, 2023.
101. Matthew, D.A.M.; Jahanbani Ghahfarokhi, A.; Ng, C.S.W.; Nait Amar, M. Proxy Model Development for the Optimization of Water Alternating CO_2 Gas for Enhanced Oil Recovery. *Energies* **2023**, *16*, 3337. [CrossRef]
102. Xiong, X.; Lee, K.J. Data-Driven Modeling to Optimize the Injection Well Placement for Waterflooding in Heterogeneous Reservoirs Applying Artificial Neural Networks and Reducing Observation Cost. *Energy Explor. Exploit.* **2020**, *38*, 2413–2435. [CrossRef]
103. Nwachukwu, A.; Jeong, H.; Pyrcz, M.; Lake, L.W. Fast Evaluation of Well Placements in Heterogeneous Reservoir Models Using Machine Learning. *J. Pet. Sci. Eng.* **2018**, *163*, 463–475. [CrossRef]
104. Selveindran, A.; Zargar, Z.; Razavi, S.M.; Thakur, G. Fast Optimization of Injector Selection for Waterflood, CO_2-EOR and Storage Using an Innovative Machine Learning Framework. *Energies* **2021**, *14*, 7628. [CrossRef]
105. Ding, M.; Yuan, F.; Wang, Y.; Xia, X.; Chen, W.; Liu, D. Oil Recovery from a CO_2 Injection in Heterogeneous Reservoirs: The Influence of Permeability Heterogeneity, CO_2-Oil Miscibility and Injection Pattern. *J. Nat. Gas Sci. Eng.* **2017**, *44*, 140–149. [CrossRef]
106. Ahmadi, M.A.; Zendehboudi, S.; James, L.A. Developing a Robust Proxy Model of CO_2 Injection: Coupling Box–Behnken Design and a Connectionist Method. *Fuel* **2018**, *215*, 904–914. [CrossRef]
107. Karacan, C.Ö. A Fuzzy Logic Approach for Estimating Recovery Factors of Miscible CO_2-EOR Projects in the United States. *J. Pet. Sci. Eng.* **2020**, *184*, 106533. [CrossRef]
108. Chen, B.; Pawar, R.J. Characterization of CO_2 Storage and Enhanced Oil Recovery in Residual Oil Zones. *Energy* **2019**, *183*, 291–304. [CrossRef]
109. Iskandar, U.P.; Kurihara, M. Time-Series Forecasting of a CO_2-EOR and CO_2 Storage Project Using a Data-Driven Approach. *Energies* **2022**, *15*, 4768. [CrossRef]
110. You, J.; Ampomah, W.; Sun, Q.; Kutsienyo, E.J.; Balch, R.S.; Dai, Z.; Cather, M.; Zhang, X. Machine Learning Based Co-Optimization of Carbon Dioxide Sequestration and Oil Recovery in CO_2-EOR Project. *J. Clean. Prod.* **2020**, *260*, 120866. [CrossRef]
111. Ampomah, W.; Balch, R.S.; Grigg, R.B.; McPherson, B.; Will, R.A.; Lee, S.; Dai, Z.; Pan, F. Co-optimization of CO_2-EOR and Storage Processes in Mature Oil Reservoirs. *Greenh. Gases Sci. Technol.* **2017**, *7*, 128–142. [CrossRef]

112. You, J.; Ampomah, W.; Sun, Q.; Kutsienyo, E.J.; Balch, R.S.; Cather, M. Multi-Objective Optimization of CO_2 Enhanced Oil Recovery Projects Using a Hybrid Artificial Intelligence Approach. In Proceedings of the SPE Annual Technical Conference and Exhibition, Calgary, AB, Canada, 30 September–2 October 2019; SPE: Calgary, AB, Canada, 2019.
113. You, J.; Ampomah, W.; Kutsienyo, E.J.; Sun, Q.; Balch, R.S.; Aggrey, W.N.; Cather, M. Assessment of Enhanced Oil Recovery and CO_2 Storage Capacity Using Machine Learning and Optimization Framework. In Proceedings of the SPE Europec featured at 81st EAGE Conference and Exhibition, London, UK, 3–6 June 2019; SPE: London, UK, 2019.
114. Vo Thanh, H.; Sugai, Y.; Sasaki, K. Application of Artificial Neural Network for Predicting the Performance of CO_2 Enhanced Oil Recovery and Storage in Residual Oil Zones. *Sci. Rep.* **2020**, *10*, 18204. [CrossRef] [PubMed]
115. Emera, M.K.; Sarma, H.K. A Genetic Algorithm-Based Model to Predict CO-Oil Physical Properties for Dead and Live Oil. *J. Can. Pet. Technol.* **2008**, *47*, 52–61. [CrossRef]
116. Rostami, A.; Arabloo, M.; Kamari, A.; Mohammadi, A.H. Modeling of CO_2 Solubility in Crude Oil during Carbon Dioxide Enhanced Oil Recovery Using Gene Expression Programming. *Fuel* **2017**, *210*, 768–782. [CrossRef]
117. Rostami, A.; Arabloo, M.; Lee, M.; Bahadori, A. Applying SVM Framework for Modeling of CO_2 Solubility in Oil during CO_2 Flooding. *Fuel* **2018**, *214*, 73–87. [CrossRef]
118. Mahdaviara, M.; Nait Amar, M.; Hemmati-Sarapardeh, A.; Dai, Z.; Zhang, C.; Xiao, T.; Zhang, X. Toward Smart Schemes for Modeling CO_2 Solubility in Crude Oil: Application to Carbon Dioxide Enhanced Oil Recovery. *Fuel* **2021**, *285*, 119147. [CrossRef]
119. Iskandarov, J.; Fanourgakis, G.S.; Ahmed, S.; Alameri, W.; Froudakis, G.E.; Karanikolos, G.N. Data-Driven Prediction of in Situ CO_2 Foam Strength for Enhanced Oil Recovery and Carbon Sequestration. *RSC Adv.* **2022**, *12*, 35703–35711. [CrossRef]
120. Moosavi, S.R.; Wood, D.A.; Ahmadi, M.A.; Choubineh, A. ANN-Based Prediction of Laboratory-Scale Performance of CO_2-Foam Flooding for Improving Oil Recovery. *Nat. Resour. Res.* **2019**, *28*, 1619–1637. [CrossRef]
121. Raha Moosavi, S.; Wood, D.A.; Samadani, A. Modeling Performance of Foam-CO_2 Reservoir Flooding with Hybrid Machine-Learning Models Combining a Radial Basis Function and Evolutionary Algorithms. *Comput. Res. Prog. Appl. Sci. Eng. CRPASE* **2020**, *6*, 1–8.
122. Khan, M.R.; Kalam, S.; Abu-khamsin, S.A.; Asad, A. Machine Learning for Prediction of CO_2 Foam Flooding Performance. In Proceedings of the ADIPEC, Abu Dhabi, United Arab Emirates, 31 October–3 November 2022; SPE: Abu Dhabi, United Arab Emirates, 2022.
123. Vo Thanh, H.; Sheini Dashtgoli, D.; Zhang, H.; Min, B. Machine-Learning-Based Prediction of Oil Recovery Factor for Experimental CO_2-Foam Chemical EOR: Implications for Carbon Utilization Projects. *Energy* **2023**, *278*, 127860. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.