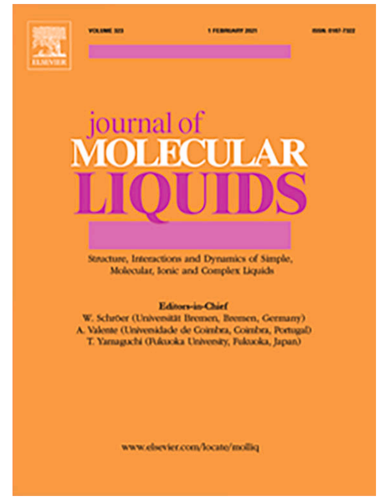
Journal Pre-proofs



Advanced machine learning-based modeling of interfacial tension in the crude oil-brine-diethyl ether system: Insights into the effects of temperature and salinity

Amir Mohammadi, Mahsa Parhizgar Keradeh, Alireza Keshavarz, Mohsen Farrokhrouz

PII: S0167-7322(24)00917-6
DOI: https://doi.org/10.1016/j.molliq.2024.124861
Reference: MOLLIQ 124861

To appear in: *Journal of Molecular Liquids*

Received Date: 21 January 2024
Revised Date: 19 March 2024
Accepted Date: 28 April 2024

Please cite this article as: A. Mohammadi, M. Parhizgar Keradeh, A. Keshavarz, M. Farrokhrouz, Advanced machine learning-based modeling of interfacial tension in the crude oil-brine-diethyl ether system: Insights into the effects of temperature and salinity, *Journal of Molecular Liquids* (2024), doi: https://doi.org/10.1016/j.molliq.2024.124861

Advanced Machine Learning-Based Modeling of Interfacial Tension in the Crude Oil-Brine-Diethyl Ether System: Insights into the Effects of Temperature and Salinity

Amir Mohammadi [a,*], Mahsa Parhizgar Keradeh [a], Alireza Keshavarz [b], Mohsen Farrokhrouz [b]

[a] Faculty of Petroleum and Natural Gas Engineering, Sahand University of Technology, Tabriz, PO. Box: 51335-1996, Iran

[b] School of Engineering, Edith Cowan University, Joondalup, WA 6027, Australia

Email address: mohammadiamir1994@yahoo.com

Abstract

Solvent injection, a well-established method for enhanced oil recovery (EOR), has demonstrated significant improvements in oil recovery when compared with conventional water flooding techniques. The interfacial tension (IFT) is pivotal in determining the displacement efficiency and overall performance of innovative techniques like dimethyl ether-enhanced waterflooding (DEW), which has gained substantial attention in recent years. In this study, following laboratory measurements of IFT, six advanced machine learning (ML) techniques were employed: Generalized Linear Model (GLM), Gradient Additive Model (GAM), Random Forest (RF), Support Vector Machine (SVM), Extreme Gradient Boosting (XGBoost), and Boosted Regression Tree (BRT) to model the IFT in both oil-brine and oil-brine-diethyl ether (DEE) systems. The analysis is based on an extensive dataset comprising 7,017 data points for oil-brine and 6,949 data points for oil-brine-DEE systems obtained from experimental studies. The findings indicate that the developed RF model excels in predicting IFT, boasting a remarkable coefficient of determination ($R^2 = 0.99$) along with the lowest root mean squared error (RMSE = 0.2), mean squared error (MSE = 0.04), and mean absolute error (MAE = 0.13). The study underscores the significance of optimizing salinity levels to achieve the most substantial reduction in IFT. This reduction is attributed to the enhanced migration of polar components, such as asphaltene molecules, to the interface of the oil-brine system. Moreover, the research highlights a synergistic decrease in IFT when both DEE and soluble ions are present, resulting in the lowest IFT at around 2 mN/m in 40,000 ppm salinity (S_2) at 70°C (T_3). This indicates that the adsorption of DEE at the water–oil interface forms a layer capable of adsorbing ions, thereby enhancing the layer's thickness. As a result, the oil–solvent-ion layer becomes thicker compared to the oil-ion layer, leading to the maximum decrease in IFT. Additionally, with increasing temperature up to 70°C, the IFT of both systems demonstrated a downward trend, as evidenced by all experiments. The outcomes of this study have the potential to enhance our comprehension of the underlying mechanisms involved in water-soluble solvent EOR techniques.

Key Words: Diethyl ether, Interfacial tension, Machine learning algorithms, Mutual solvent, EOR

1. Introduction

The increasing global energy demand has intensified the focus on Enhanced Oil Recovery (EOR) [1–4]. According to the Organization of Petroleum Exporting Countries (OPEC), projections indicate that by 2040, the demand for oil is expected to reach approximately 11.1 million barrels per day. This represents a 23.1 % increase compared to the current value [5]. Therefore, low-cost EOR methods are highly favorable as a result [6–9]. A wide range of EOR techniques proposed and developed, including thermal, chemical, and gas injection methods; different techniques had certain set of advantages and disadvantages [10,11]. Multiple studies propose utilizing ether as a cost-effective, non-hydrocarbon solvent for recovering crude oil. Shell initially introduced this method based on the favorable combination of ether properties, water, and oil, establishing it as a highly effective solvent for oil recovery [12–15]. The capacity to dissolve in both brine and oil offers an economic benefit [16,17]. During ether-enhanced water flooding, the transfer of ether mass into the oleic phase leads to the mobilization of residual oils through mechanisms such as swelling, viscosity reduction, and a decrease in interfacial tension (IFT), ultimately resulting in substantial oil recovery [18–20]. According to the previous research, ethers exhibit higher solubility in the aqueous phase when compared to CO_2. Particularly, Dimethyl ether (DME) has been observed to increase oil swelling up to four times more than CO_2 [21]. The extent of oil swelling and viscosity reduction due to DME mass transfer from the water to the oil phase, known as the partition coefficient, plays a crucial role in ether-enhanced water flooding. This process enhances oil mobility and decreases residual oil saturation [22–24]. Additionally, DME and CO_2 possess global warming potentials of 0.1 and 1.0, respectively, when considered over a 500-year timeframe. This suggests that DME could be deemed environmentally friendly [25]. Consequently, due to their relatively minor negative environmental effects compared to many other EOR methods, ethers emerge as a promising solution to address sustainability challenges in the oil industry [26–28]. In detail, DME represents the simplest form of ether. Under standard atmospheric conditions, DME typically exists in the gaseous state. In contrast, Diethyl ether (DEE), with the chemical formula $(C_2H_5)_2O$, remains in a liquid state at room temperature, characterized by a relatively low melting point of -116.3°C [29]. The low melting and boiling points make DEE easily manipulable in laboratory settings and contribute to its widespread use as a solvent [30]. Due to the reasons mentioned above, working with DEE under laboratory conditions is feasible. It is anticipated that the methods used to recover oil with DME and DEE share similar production mechanisms. As a result, it is practical to conduct a single study and utilize experimental data from DEE tests and apply it for DME. Given the substantial potential to improve oil recovery through mechanisms like IFT reduction, it is beneficial to gain insights into the interfacial behavior of ether in the oil-brine system.

Generally, IFT plays a vital role in all EOR processes. The IFT is significantly influenced by the composition of the two phases, as well as the pressure and temperature conditions. Experimental techniques, including pendent and spinning drop methods, are commonly employed to measure this property, but they can be both costly and time-intensive, and in certain situations, they may pose significant challenges. Consequently, the consideration of an alternative approach, such as calculating IFT through modeling, becomes crucial [31]. Artificial intelligence (AI) techniques have found application in the petroleum industry for diverse purposes[32–36], offering an alternative solution for tasks such as estimating IFT. AI, functioning as an intuitive mechanism, encompasses various capabilities, including observation, learning, and reasoning [37]. As an interdisciplinary science, AI employs multiple approaches, with notable successful applications in areas such as "classification," "forecasting," "control systems," and "optimization and decision-making"[38]. Machine learning (ML), a subset of AI and computer science, centers on utilizing data and algorithms to emulate human learning processes, gradually enhancing its accuracy over time. ML is specifically concerned with the development of computer programs capable of adapting when exposed to new data [39,40]. As a result, the utilization of a ML approach emerges as an intelligent strategy for modeling IFT using straightforward inputs from experimental data, including temperature, salinity, and other crucial parameters. With these inputs, a robust ML model has the potential to provide accurate and timely predictions.

It's important to highlight that, given the prevalence of oil and water as the dominant fluids in reservoirs, the majority of ML models have been designed to predict IFT in various systems, including oil-brine [41,42], water-hydrocarbon [43], brine-hydrocarbon [44,45], and CO_2-brine [46–48]. For example, In Barati Harooni et al.'s (2016) study, they utilized the combination of Least Square Support Vector Machine (LS-SVM) and Coupled Simulated Annealing (CSA) to model oil-brine IFT. The results revealed that their developed model accurately predicts experimental IFT data [49]. In 2019, Menad Nait Amar et al. presented the Gradient Boosting Decision Tree (GBDT) model as superior in predicting IFT for oil-brine systems, achieving an R-squared value of 0.9977 across all data, surpassing the AdaBoost SVR method. Their research contributed to the field by developing and statistically validating two ML models, with the GBDT model showing high accuracy and utility for estimating IFT [45]. In 2020, Alexsandro Kirch et al. showcased the effectiveness of ML, particularly the gradient boosted algorithm, in forecasting oil-brine IFT with an R-squared score of 0.97, surpassing the less precise linear regression approach [50]. In the 2024 study by Yousefmarzi et al., six ML algorithms were employed to predict IFT between gas-water and oil-water systems. These algorithms included Support Vector Regression (SVR), Random Forests (RF), Decision Tree (DT), Gradient Boosting (GB), Catboosting (CB), and XGBoosting (XGB). The study revealed that SVR and CB outperformed other algorithms in terms of accuracy and robustness [51]. However, to the best of our knowledge, no investigation has been conducted on modeling the IFT of oil-brine-DEE systems using ML methods. Despite the past decade witnessing limited research dedicated to investigating the use of ethers as an EOR agent, their potential merits further exploration.

In this regard, some experimental studies have examined that molecular diffusion and partition coefficient of DME play a significant role in oil recovery during DME solvent injection in comparison with relative permeability and capillary pressure [52]. However, the oil recovery rate attributed to DME molecular diffusion is notably slower than that driven by capillary forces. Khalifi et al. (2019) conducted a study on the DME diffusion coefficient in Athabasca bitumen, examining pressures ranging from 0.689 to 2.757 MPa and temperatures from 50°C to 110°C. The research findings revealed variations in DME molecular diffusion within bitumen, ranging between 0.2–2 and 10–9 m²/s [53]. The investigation conducted by Fayazi and Kantzas (2019) employed magnetic resonance imaging (MRI) to determine the diffusion coefficients of various solvents, including DME, propane, ethane, and CO_2, within heavy oil. The findings revealed a notable swelling effect of DME on heavy oil in comparison to the other solvents, with a density of 0.9887 g/cm^3 at 15.56°C. Specifically, DME exhibited dynamic and equilibrium swelling factors of 18.4 % and 37.7 %, respectively, under a pressure of 0.55 MPa [54].

Additional research has been devoted to assessing the EOR capabilities of DME for enhanced waterflooding. This innovative chemical EOR technology has gained prominence across various reservoirs, representing a distinct approach capable of enhancing oil recovery in reservoirs characterized by both sandstone and carbonate formations with low permeability. Notably, successful implementation of DME injection has been demonstrated in the Hatter's Pond field in Alabama [55]. Parsons et al. explored DEW core flood experiments in the presence of live oil (25 cp) within Berea sandstone. Their study involved the injection of 3.2 pore volumes (PV) of freshwater, succeeded by 1 PV of DME slug (10 % DME in water). This injection sequence led to a notable surge in incremental oil recovery, reaching 25 % beyond the 45 % achieved through the conventional waterflooding [20].

The results of a study assessing the effectiveness of injecting DME into a fractured chalk reservoir were noteworthy [56]. Core-flooding experiments with DME-brine showed a significant 44.2% increase in oil recovery. This increase can be attributed to the migration of DME from the DME-brine solution into the oil phase, resulting in an increase in the density of the DME-brine solution and a decrease in the density of the oil phase. These changes are anticipated to improve the vertical sweep efficiency of the oil by the DME-brine solution. Further investigations by Javanmard et al. explored DME-brine injection (at 3 and 6 PV) into tight chalk core samples, revealing an additional oil recovery of 31.4 % compared to conventional waterflooding. Notably, in chalk reservoirs, DME injection exhibited distinct advantages over CO_2 flooding, as it did not lead to precipitation or mineral dissolution. The observed increase in differential pressure was attributed to the significant swelling of oil in the presence of DME-brine solution [57].

Recently, there have been numerical simulation investigations into the injection of DME. The solubility and partition coefficient of DME in the DME-brine-oil system are influenced by the interactions among its various components. To comprehensively capture these interactions and model the equation of state (EOS) with flow behavior, it is imperative to integrate numerical simulations with experimental studies [58]. In the context of the DEW model, a compositional simulation-based numerical analysis was developed to examine the phase behavior performance and clarify the transport mechanism, utilizing the GEM compositional simulator from the Computer Modeling Group (CMG). Notably, the IFT between DME and reservoir oil was found to be very small, while that between CO_2 and reservoir oil was approximately 5 dyne/cm. Consequently, DME exhibited a more pronounced reduction in oil viscosity compared to CO_2 flooding. The study revealed that DEW application led to a 34 % and 12 % improvement in oil recovery compared to conventional waterflooding and CO_2 flooding, respectively [59]. Chernetsky et al. developed a dynamic model to conduct history matching, interpretation, and sensitivity analyses at various stages of core-flooding experiments. The DME solubility and partition coefficient data were fine-tuned using the cubic-plus-association (CPA) EOS model, which combines the Soave-Redlich-Kwong (SRK) equation of state with the association term from the Wertheim theory. Through both experiments and simulations focused on oil recovery during core floods, they observed a high level of harmony between the experimental findings and the simulated models. Following the injection of DME/brine, noteworthy alterations in the wettability of carbonate cores were observed, shifting towards a more water-wet condition [60].

Their research findings highlighted a significant knowledge gap pertaining to the impact of DEE on the IFT within the oil-brine system.

While interest in utilizing ethers for EOR is increasing, the existing literature predominantly examines DME solubility in aqueous and oleic phases across various salinity, pressure, and temperature conditions, crucial for oil recovery. Some studies have also investigated DME diffusion coefficients in crude oil, given their significant impact on oil recovery during solvent injection. Several researchers have evaluated DME's EOR effectiveness in enhanced waterflooding, with some conducting pilot tests of this method. However, there remains a notable gap in our understanding of how ethers, particularly DEE, affect IFT in oil-water systems, as the effect of DEE on IFT has not been examined. Additionally, our research and comprehensive literature review have revealed the absence of a dedicated AI model for predicting the IFT relationship between oil and brine in the presence of DEE. To bridge this gap, our study aims to introduce six advanced ML models: the Generalized Linear Model (GLM), Gradient Additive Model (GAM), Support Vector Machine (SVM), Random Forest (RF), Extreme Gradient Boosting (XGBoost), and Boosted Regression Tree (BRT). The selected models strike a balance between linear and nonlinear interpretations. Additionally, their successful application in related fields provides a strong foundation for predicting IFT values in the oil-brine-diethyl ether system.

In this study, Firstly, the data were generated by measuring the IFT of oil-brine-DEE across various salinities and temperatures, while maintaining constant pressure conditions (2000 psi) using the pendant drop method. Secondly, the experimental data underwent modeling utilizing the six aforementioned ML models. Thirdly, the proposed models' validity was evaluated, followed by a rigorous statistical analysis to determine the optimal model. This research endeavor seeks to contribute valuable insights to the field of water-based EOR projects by addressing a crucial knowledge gap regarding the influence of ethers on IFT dynamics in oil-water systems.

2. Methodology

2.1 Workflow

Fig. 1 illustrates the full set of steps utilized in this study, consisting of six primary steps. In Steps 1 and 2, initially, the 40,000 ppm solution (S_2) was prepared. To assess the influence of salinity on the IFT, two additional solutions with salinities of 4000 ppm (S_1) and 80,000 ppm (S_3) were also prepared. Subsequently, the study examined the effects of varying temperatures of 30ºC (T_1), 50ºC (T_2), and 70ºC (T_3) on the IFT of the oil-brine system, serving as the base case. Then, DEE with maximum solubility was introduced into the aqueous phase across all solutions. The IFT was then measured under consistent experimental conditions, and a comparative analysis was conducted between the two sets of experiments. This was done to elucidate the impact of DEE on the IFT of the oil-brine system. Moving on to Step 3, A dataset comprising 7017 entries for the oil-brine system and 6949 entries for the oil-brine-DEE system was assembled to develop the ML models. The entire dataset was divided into two distinct sets: the training set and the testing set. For this study, 70 % of the data points were randomly selected and employed in constructing the prediction models. The remaining 30 % of the experiment samples represented the testing phase of the ML paradigms. Step 4 includes selecting advanced ML algorithms specifically tailored to model the IFT. The research then proceeds to Step 5, where the selected ML models are trained on the prepared data, and their performance is rigorously evaluated using statistical indices. Finally, Step 6 selects the best model based on their performance in modeling the IFT of oil-brine in the presence of DEE.

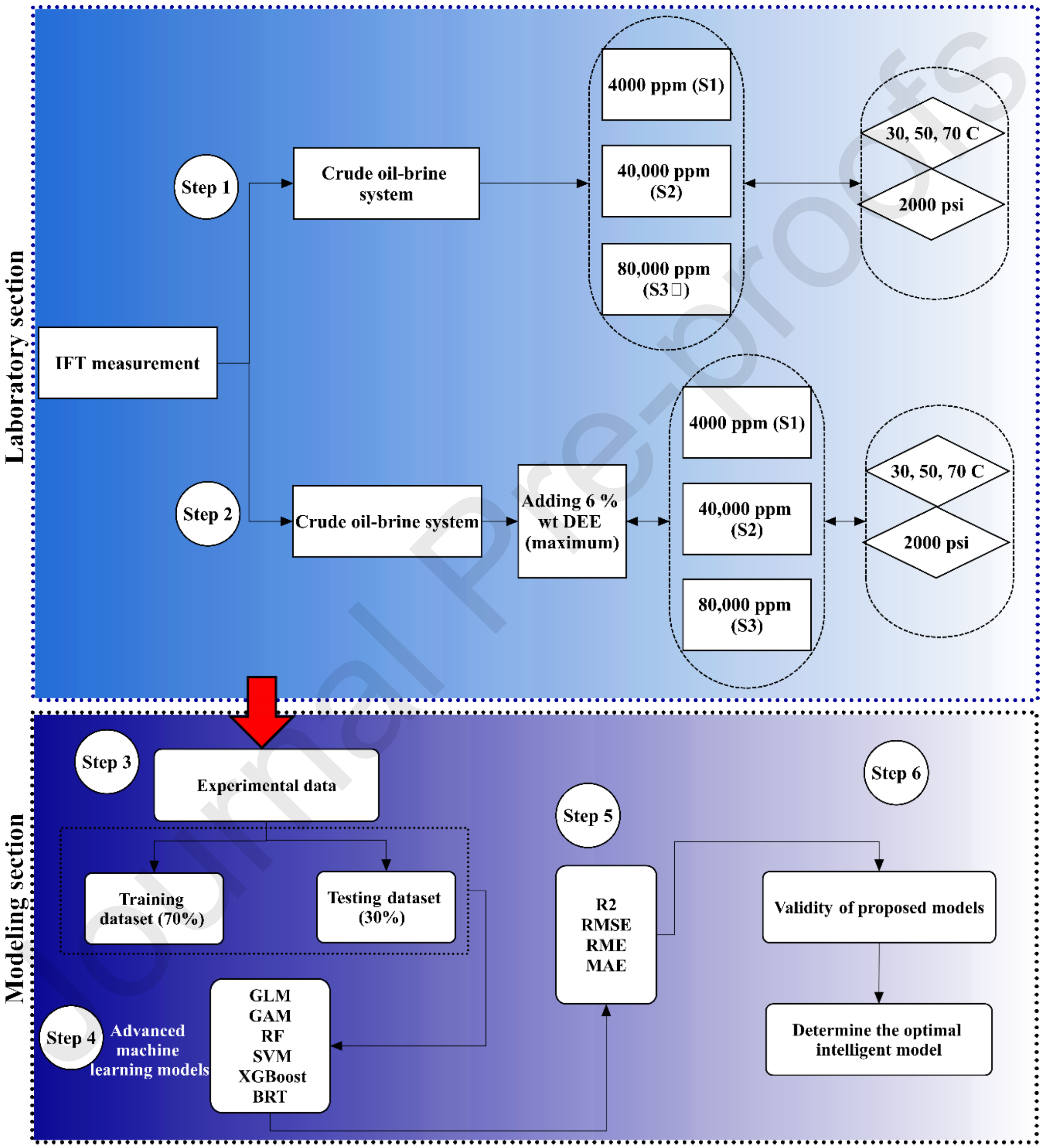


Fig. 1. The series of steps for this research

2.2 Brine and crude oil

In the present study, an S_2 solution was prepared, and its analysis is detailed in Table 1. Additionally, S_1 and S_3 solutions were prepared separately with distilled water to investigate the impact of salinity on the IFT of oil-brine-DEE. The oil utilized in this research originates from the Bangestan reservoir in Iran, characterized by carbonate rock formations, specifically limestone varieties. The characteristics of the oil utilized in this study are detailed in Table 2.

Table 1. Composition of various brines (in ppm) utilized in this research

Ion	4000 ppm (S_1)	40,000 ppm (S_2)	80,000 ppm (S_3)
Na^+	1724	17243	34486
Cl^-	1981	19807	39614
K^+	40	400	800
HCO^{3-}	5	50	100
Mg^{2+}	214	2143	4286
Ca^{2+}	46	460	920
SO_4^{2-}	150	1497	2994
TDS	4162	41659	83318
Ionic strength, mol/L	0.083	0.832	1.664

Table 2. Physical characteristics of the dead oil employed in this research

Properties	value
°API	29

Viscosity (at 30ºC), centipoise	11
Asphaltene (wt. %)	2.5
Density (at 30ºC), gr/cm^3	0.885

2.3 DEE-brine solutions

The study utilized high-purity DEE (≥99 %) obtained from Merck Company, and we determined the maximum solubility of DEE in brine (6 % wt). Literature indicates that DEE's solubility in deionized water at 25°C and ambient pressure is 60.5 g/mL [61]. To guarantee maximal dissolution of DEE in brine, 20cc of seawater was introduced into a test tube, and DEE was gradually introduced with continuous stirring on a magnetic stirrer for approximately 1 hour. Upon complete mixing, the solution underwent inspection for single-phase consistency. It is important to mention that throughout the experimentation, the test tube remained tightly sealed to prevent DEE evaporation.

2.4 IFT measurement

The IFT measurement between oil and various solutions was conducted using the pendant drop method, which entails deducing the IFT by analyzing the shape profile of a droplet of one liquid suspended within another at mechanical equilibrium. The high-pressure/high-temperature IFT apparatus is equipped with two pumps, a 300cc chamber with glass windows, a sealed metal needle connected to the hydrocarbon pump at the chamber's bottom, a heating jacket with ±1°C accuracy, a pressure gauge with ±10 psi accuracy, a light projection system, a camera, and a computer featuring image processing software. This software is employed for measuring IFT values under both static and dynamic conditions, as depicted in Fig. 2. In this procedure designed to evaluate the IFT between two liquids, the process involves introducing the solution into the chamber using a pump. Subsequently, an oil droplet is suspended from the needle within the solution. The pump is utilized to control the oil droplet, while a digital camera captures images of the oil droplets at various time intervals. The software, based on Eq. (1), calculates the IFT between the oil droplets and the aqueous solution. The IFT measurements are conducted at three distinct temperatures, specifically T_1, T_2, and T_3, which are maintained and regulated by a temperature controller.

$$\gamma = \frac{\Delta \rho g D^2}{H} \tag{1}$$

where, "γ" represents the IFT (mN/m), "g" corresponds to the force of gravity, "D" denotes the equatorial diameter of the drop (cm), "$\Delta\rho$" stands for the density disparity between two immiscible liquids (g/cm³), and "H" signifies the coefficient determining the drop's shape [62].

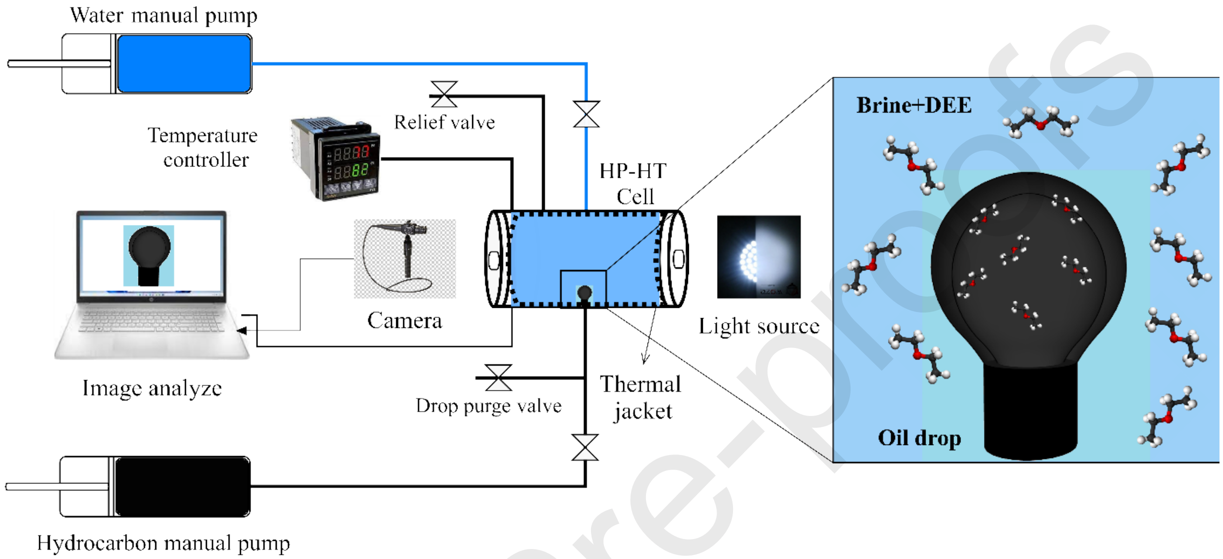


Fig. 2. Apparatus and schematic of the high-pressure/high-temperature IFT utilized in this study

2.5 Machine learning model developments

In this section, we expound on the theoretical principles underlying six advanced ML algorithms and provide statistical metrics for assessing their relative performance. The subsequent segment specifies the theoretical foundations and essential functionalities of the GLM, GAM, RF, SVM, XGBoost, and BRT models utilized in the study. All data analysis, running the models and visualization were conducted in R software.

2.5.1 Generalized Linear Model (GLM)

In this study, the GLM was utilized to model the IFT of two systems: oil-brine and oil-brine-DEE under laboratory conditions. The GLM methodology was implemented using the MASS package within the R 4.2.2 software environment. Serving as an extension of conventional linear regression, GLM proves versatile in handling both linear and non-linear datasets for regression analysis. Despite its inherent simplicity, GLM has found widespread application in predicting and has exhibited commendable performance relative to other modeling methods. GLM leverages multivariate regression to express conditional factors as functions of the presence or absence of predictand-related factors. Notably, one of the key advantages of GLM, compared to traditional linear regression models, is its independence from the assumption of normal distribution for observed data, contributing to its applicability in diverse contexts. More information about this model can be found at [63].

2.5.2 Gradient Additive Model (GAM)

GAM, often referred to as "wiggly models" [64], was pioneered by previous research [65] as an amalgamation of GLM and additive models. In GAM, the linear association between the dependent and independent variables is substituted with non-linear smooths. GAM adopts an additive approach where the appropriate functional form is chosen based on the data without prior knowledge of the model's functional aspects [66]. Initially designed to combine the advantages of GLMs and additive models within a single framework [67], GAMs serve as nonparametric extensions of GLMs. Their primary advantage over the latter lies in the ability to model complex and non-linear relationships. The response variable (Y) in GAMs is not confined to following a normal distribution; instead, it can be fitted by various distributions such as Poisson or binomial distributions.

2.5.3 Support Vector Machine (SVM)

The SVM, a member of the Generalized Linear Classifier family and founded on the Vapnik-Chervonenkis Dimension theory, was initially devised by Vladimir N. Vapnik in 1963 for linear models [68]. It was later expanded to handle non-linear training data in 1995 by Cortes and Vapnik in 1995 [69]. SVM constitutes a collection of supervised ML models employing a kernel function for regression (SVR) and implementing nonlinear classification (SVC). It constructs an optimal separating hyperplane that transforms a low-dimensional input vector into a higher-dimensional feature space, utilizing the Vapnik-Chervonenkis Dimension theory to ensure robust generalization capabilities [69].

2.5.4 Random Forest (RF)

The RF concept is a fusion of classifier and repressor, enabling decision-making through a collection of tree-based decisions. RF functions by creating multiple Decision Trees (DTs) using randomly selected subsets of input features and training data. Each DT is trained on a separate subset of the data, and the ultimate prediction is obtained by combining the predictions from each individual DT [70]. By employing multiple DTs, RF mitigates overfitting and improves prediction accuracy by diminishing the impact of biased or inconsistent individual trees [71]. Additionally, RF can manage high-dimensional data and incomplete observations, thereby enhancing its adaptability for various applications [32,72].

2.5.5 Extreme Gradient Boosting (XGBoost)

The core principle behind XGBoost lies in utilizing a boosting algorithm, which combines multiple weak prediction models to generate more accurate overall predictions [73]. In the gradient boosting (GB) approach, successive iterations incorporate additional models into the ensemble, with each model focusing on instances inaccurately predicted by previous ones. Noteworthy for its versatility, XGBoost can handle various input data structures, such as sparse, missing, and categorical data [73]. XGBoost offers a range of hyperparameters that can be adjusted to improve its effectiveness for a given task, including the learning rate, maximum depth of each DT, number of trees in the ensemble, and regularization parameters [74]. Consequently, XGBoost has shown superior performance compared to other commonly used ML algorithms, such as RF and neural networks [33,74].

2.5.6 Boosted Regression Tree (BRT)

The BRT method, as introduced by Elith (2008)[75] is a powerful ensemble algorithm that combines two key techniques: boosting and regression trees. This approach harnesses the strengths of tree algorithms, allowing for the handling of diverse predictor variable types and accommodating missing values through surrogate splitting. The BRT leverages the advantages of boosting, a sequential process that enhances predictive performance by combining multiple trees. The core concept involves giving increased attention to observations poorly modeled by existing trees, specifically those with high deviations from the mean. This emphasis continues until the algorithm minimizes predictive deviance [76]. Notably, BRT addresses the limitations of standalone regression trees, effectively overcoming their suboptimal predictive performance through the boosting algorithm [77].

BRT employs a forward, stage-wise process, maintaining existing trees unchanged while adding new trees in each iteration. The residuals of each observation are updated to account for the contribution of the newly added tree. Final predictions are determined by a weighted sum of each tree's predictions. This approach, detailed by previous studies [78–80], builds on Friedman's gradient boosting introduced in 2001 [81]. Friedman later [82] modified the procedure in 2002 by incorporating random subsampling of training data to improve prediction performance, reduce the risk of overfitting, and enhance computation efficiency [83].

2.5.7 Model evaluation indices

To assess the performance of the ML models employed in this study for modeling the IFT of the oil-brine and oil-brine-DEE systems in the laboratory condition, four commonly used statistical indices such as coefficient of determination (R^2), root mean squared error (RMSE), mean squared error (MSE) and mean absolute error (MAE) were employed. The following indices were utilized to compare the predicted IFT values with the observed values [84–86]. R^2 is a dimensionless value ranging from 0 to 1. When equals 1, it signifies that the data points are precisely aligned along a line with no dispersion. This indicates a model that fits the data perfectly. A lower RMSE, which measures the average magnitude of the differences between predicted and observed values, suggests a better fit, indicating smaller errors, but it is sensitive to outliers. MSE, representing the average squared differences, is also minimized for better model fit, though it is influenced by outliers due to the squaring operation. Both RMSE and MSE are on the same scale as the original data. On the other hand, MAE, which measures average absolute differences, is less sensitive to outliers and provides a direct interpretation in the original units.

$$R^2 = 1 - \frac{\sum (y_e - y_p)^2}{\sum (y_e - \overline{y_e})^2} \tag{2}$$

$$RSME = \sqrt{\frac{1}{n} + \sum_{i=1}^{n} (y_i - \hat{y}_i)^2} \tag{3}$$

$$MSE = \frac{\sum (y_e - y_p)^2}{N} \tag{4}$$

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

where, y_e and y_p are the experimental and observed values. y_i is also the mean of the experimental values and N is the number of values.

3. Results and discussion

3.1 Effect of the temperature and salinity on IFT of the oil-brine-DEE system

In this study, we investigated the potential differences in the responses of two different systems, one with the presence of DEE and the other without, under varying temperature and salinity levels. Fig. 3 illustrates the variations in IFT (mN/m) through box plot representation across different salinity levels S_1, S_2, and S_3 and temperatures T_1, T_2, and T_3 for two systems: oil-brine and oil-brine-DEE. The box plot is the favored technique for visually examining a single parameter. It displays key statistics such as the minimum, first quartile (Q_1), median, third quartile (Q_3), maximum, and mean values of a parameter. The upper limit (maximum) and lower limit (minimum) are determined as $[Q3 + 1.5 * (Q3 - Q1)]$ and $[Q1 - 1.5 * (Q3 - Q1)]$, respectively. Values exceeding or falling below these limits are classified as outliers. The robust Mann-Whitney U test was employed to assess the significance of differences between two systems, specifically tailored for non-normally distributed data. The null hypothesis (H_0) assumed no difference in the groups' distributions, while the alternative hypothesis (H_1) suggested a significant difference. We assessed significance using standard thresholds: $p < 0.001$ (***) for extreme significance, $0.001 \leq p < 0.01$ (**) for very significance, $0.01 \leq p < 0.05$ (*) for significance, $0.05 \leq p < 0.1$ ('.') for marginal significance, and $p \geq 0.1$ for no significance. The resulting p-values, all below 0.0001 (****), strongly support the rejection of the null hypothesis, indicating significant differences between the two systems at specific salinity levels and temperatures. Our findings consistently demonstrate that the use of DEE leads to a significant reduction in IFT when compared to the oil-brine system under various experimental conditions. The most notable contrast between the two systems occurred in S_1 solution, where the average IFT difference exceeded 10 mN/m. Moreover, through box plot analysis, it is evident that the distribution of calculated IFT values around the mean and the frequency of outliers notably decreased with higher salinity levels. Specifically, the lowest distribution was observed in S_3 solution for both systems.

Figs. 4 and 5 offer a detailed examination of IFT variations in the oil-brine-DEE system under diverse temperature and salinity levels, providing enhanced insights into the impact of these parameters on IFT. In Fig. 4, it is evident that increasing temperature from T_1 to T_3 leads to a consistent decrease in IFT across all salinity concentrations for the groups. The IFT exhibits significant differences for all salinity levels. Notably, at an S_1 solution, IFT varies between 5 mN/m to 10 mN/m, while for S_2 and S_3 solutions, the average IFT falls below 5 mN/m for all temperatures. Remarkably, the temperature of T_3 also reduces the IFT deviation from the mean compared to T_1 and T_2. This is evident in the narrower boxplots and reduced occurrence of outliers, indicating a more concentrated distribution of IFT values around the median. Contrarily, the results at T_1 and T_3 exhibit more outliers and generally wider boxplots, signifying higher differences between quartiles.

In general, the impact of temperature on IFT is contingent upon variables like the oil type, aqueous solutions, and specific temperature and pressure conditions [87]. The observed reduction in IFT is likely a consequence of diminished intermolecular forces that bind molecules within the oil and brine. With elevated temperatures, these forces weaken, promoting increased molecular mobility and ultimately leading to a more pronounced reduction in IFT [88,89]. Another noteworthy consideration is that an increase in temperature reduces the number of hydrogen bonds formed among water molecules, thereby decreasing the energy required to create a unit area of free water. Consequently, this leads to a reduction in IFT. Furthermore, heightened temperatures enhance the miscibility of water and DEE, contributing to a further decrease in IFT [90]. It's important to highlight that at lower temperatures, the reduction in IFT is not as pronounced because the thermal motion of solvent molecules is considerably weaker [91].

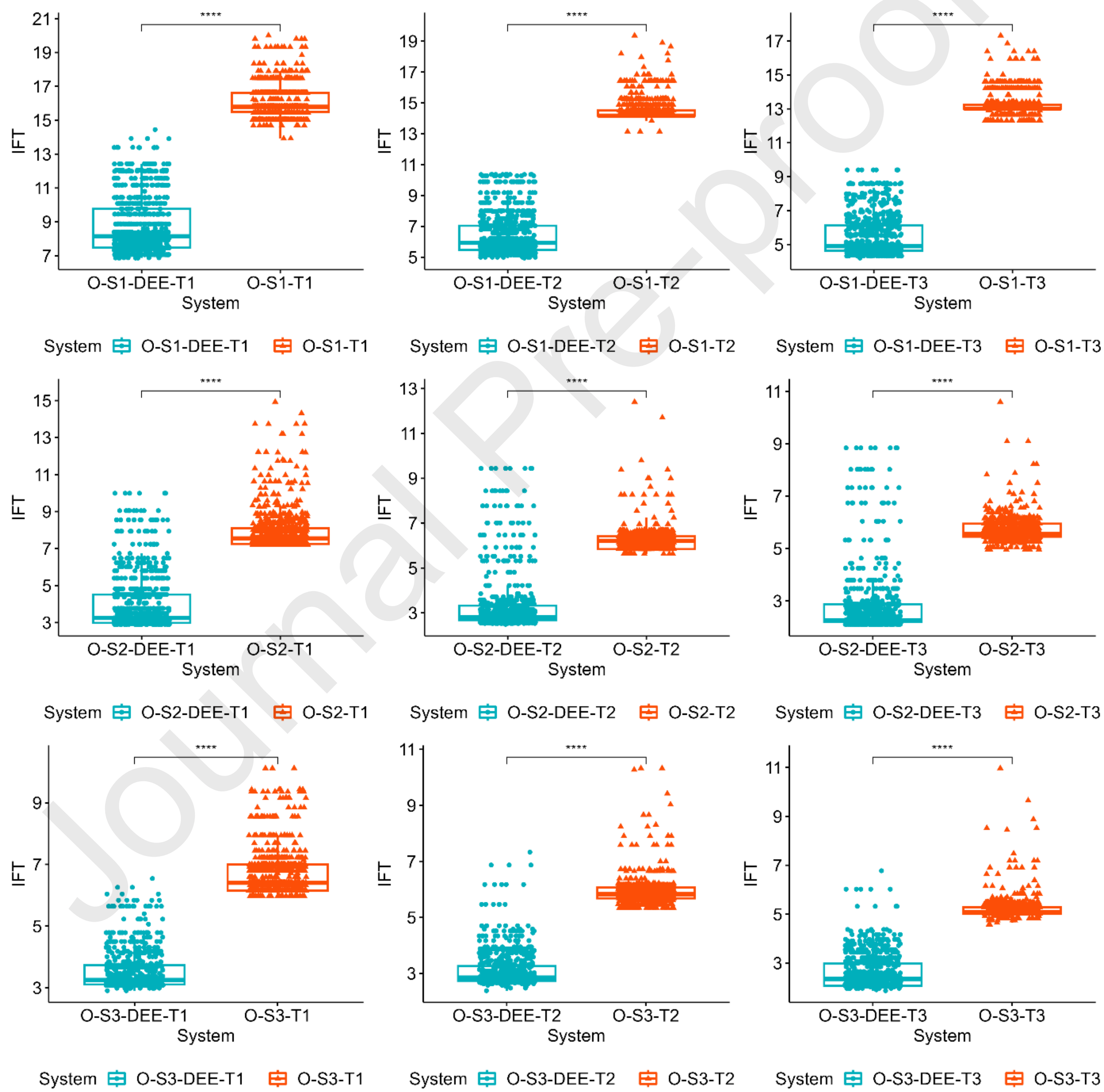


Fig. 3. Box plot representation of the IFT (mN/m) variations at salinity levels S_1, S_2, and S_3 and temperatures T_1, T_2, and T_3 for two systems: oil-brine and oil-brine-DEE

Fig. 5 delves into the influence of varying salinities on the IFT within the oil-brine-DEE system. According to the salinity results, a significant difference between groups is observed at all salinities. Notably, IFT experiences a pronounced and statistically significant decrease (P-value < 0.0001) with the rise in brine from S_1 to S_3, reaching its lowest value, approximately 2 mN/m, at T_3. A noteworthy observation is at S_3 brine, where the boxplots for the three temperatures are narrower than other experimental conditions, indicating a more concentrated distribution of IFT values around the median with fewer outliers. In contrast, the widest boxplot with the highest deviation of IFT is observed at the lowest salinity.

Overall, regarding the impact of salinity on the IFT in the oil-brine-DEE system, our findings indicate that as the salinity level rises from S_1 to S_3, there is a corresponding decrease in IFT. The minimum IFT value was observed in S_2 solution at T_3, reaching approximately 2 mN/m. Subsequent increases in salinity from S_2 to S_3 did not result in a further reduction in IFT; instead, the IFT remained within the range of approximately 2 mN/m. This stabilization is likely attributed to salting-out effect. This pattern can be elucidated by the notion that an ideal salinity range promotes the migration of polar components such as asphaltene molecules to the interface of the oil-brine system, resulting in a subsequent decrease in IFT [92–96]. The salting-out mechanism in high-salinity water involves the increased likelihood of organic components migrating to the oil phase. This is due to the reduced solubility of polar organic components in the aqueous phase. On the other hand, the salting-in mechanism is characterized by the affinity of organic particles to dissolve in water [8,97,98]. In a pure water environment, strong structural arrangements among water molecules occur on one side of the aqueous phase interface, resulting in minimal disturbance to the interface [99–101]. A collaborative impact on reducing IFT occurs when both a solvent (DEE in the current study) and soluble ions are present in the solution. This implies that the adsorption of solvents at the water–oil interface forms a layer capable of adsorbing ions, thereby augmenting the layer. Consequently, the oil–solvent-ion layer becomes thicker compared to the oil-ion layer. With an elevation in salinity, the solubility of solvents in aqueous solutions diminishes, resulting in an increased transfer of solvent mass from the aqueous phase to the oil phase [102,103]. At elevated salinity levels, the activity coefficient of DEE rises, accompanied by a decrease in solubility in the aqueous phase, thereby intensifying the influence of DEE [104].

Fig. 6 supplements these findings by presenting the histogram of IFT in two systems including oil-brine and oil-brine-DEE system, providing a comprehensive visualization of the distribution classes of the IFT under different salinity levels S_1, S_2, and S_3 and temperatures T_1, T_2, and T_3 for two systems. When the variance is too small or too large, a histogram is preferred to show the distribution of a numerical parameter. A histogram shows the numbers of values within an interval on vertical and variable on horizontal [105]. In the presented histograms, the IFT variable was explored across nine distinct graphs for two mentioned systems. These graphs were systematically organized based on three varying brine compositions, denoted as S_1, S_2, and S_3, as well as three different temperatures labeled T_1, T_2, and T_3. The vertical axis of each histogram signified the "count," representing the frequency or number of occurrences corresponding to specific IFT values on the horizontal axis.

It is important to highlight that in this study, we examined the impact of DEE on IFT under varying conditions of salinity and temperature. However, the influence of pressure was not evaluated during the IFT measurements in the presence of DEE. Furthermore, it is recommended that future research investigate the effectiveness of DEE in altering rock wettability, as this could potentially lead to tangible improvements in oil recovery.

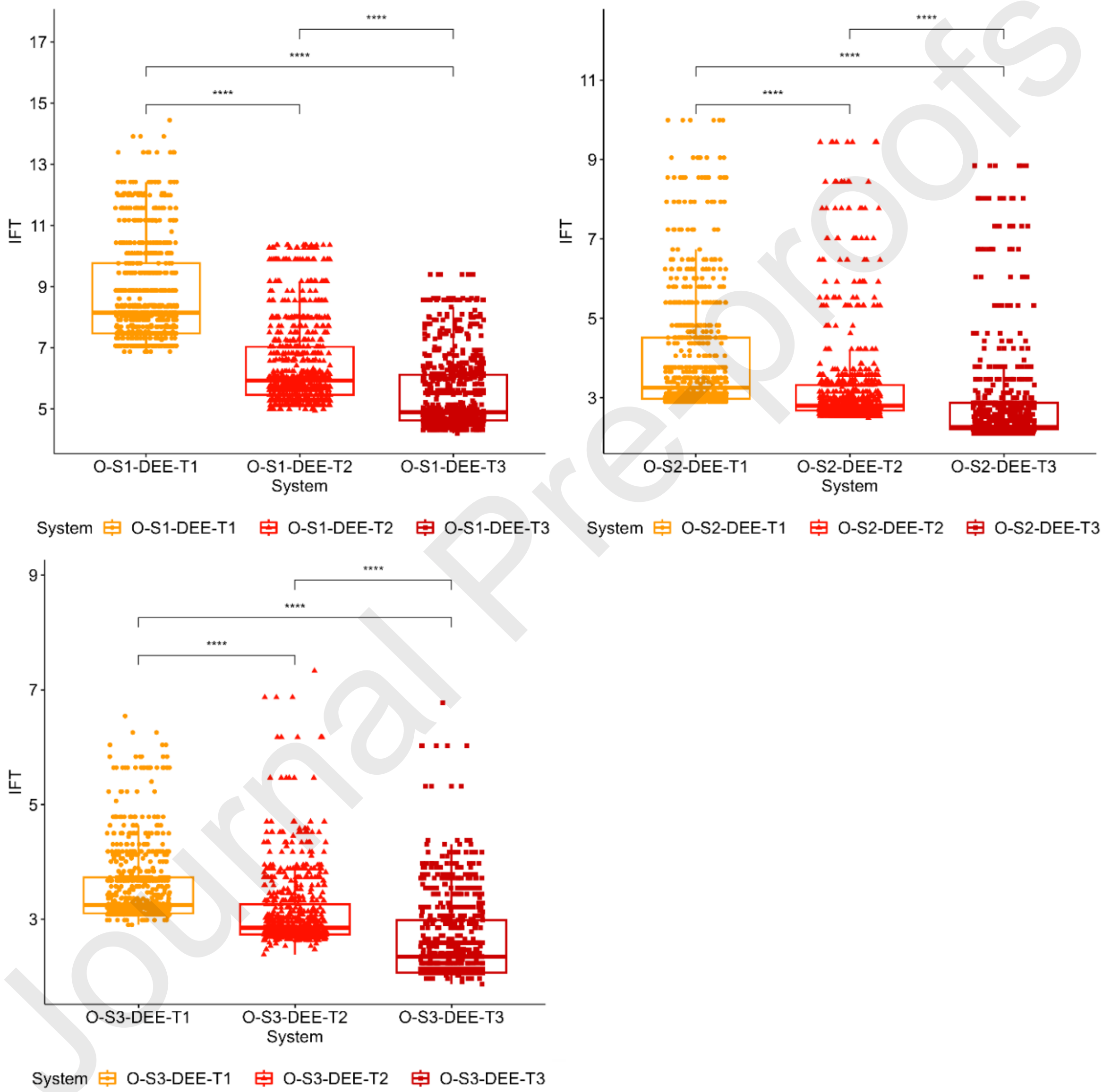


Fig. 4. Box plot representation of the IFT (mN/m) variations of the oil-brine-DEE system at different temperatures T_1, T_2, and T_3

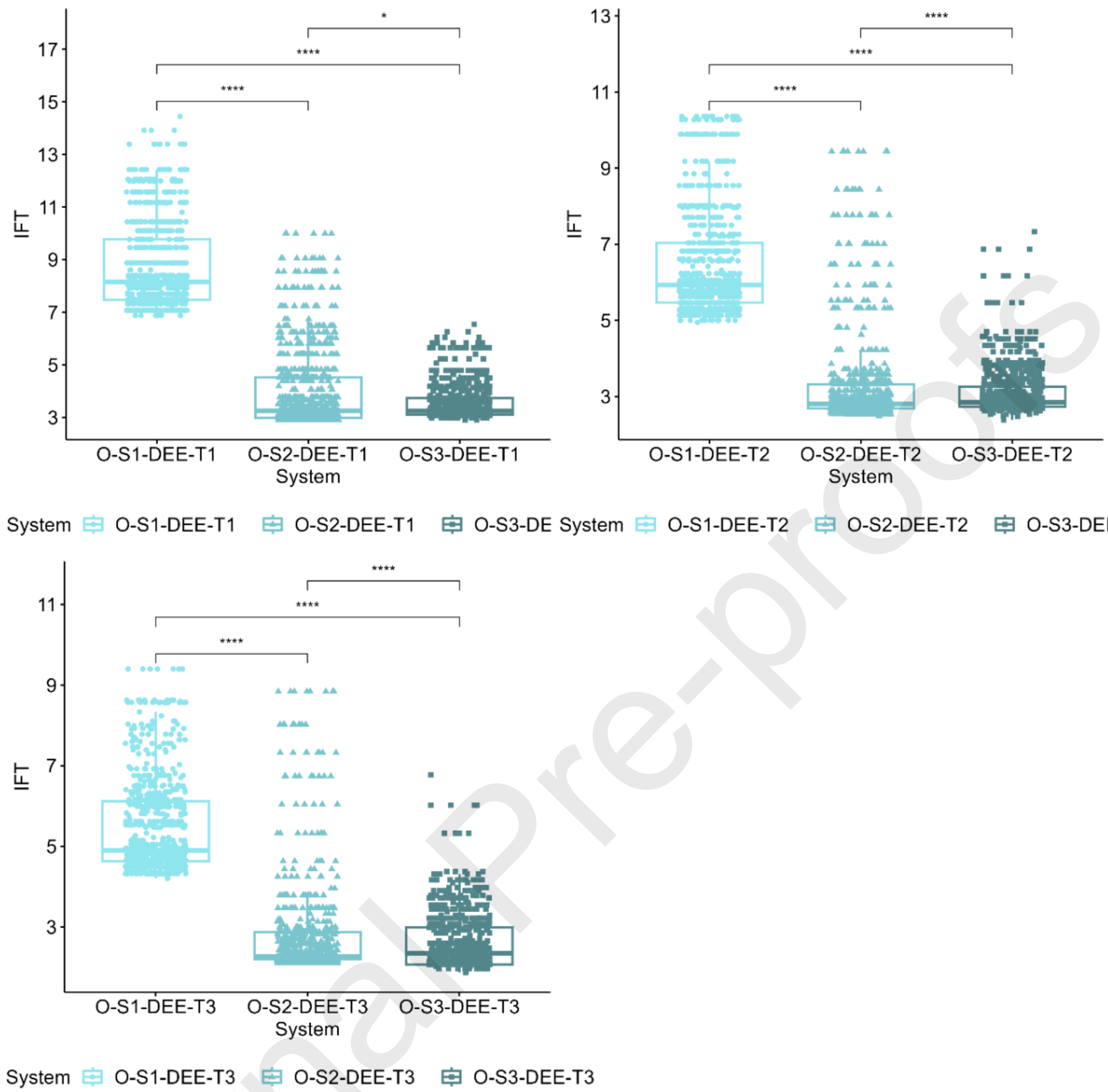


Fig. 5. Box plot representation of the IFT (mN/m) variations of the oil-brine-DEE system at different salinity levels S_1, S_2, and S_3

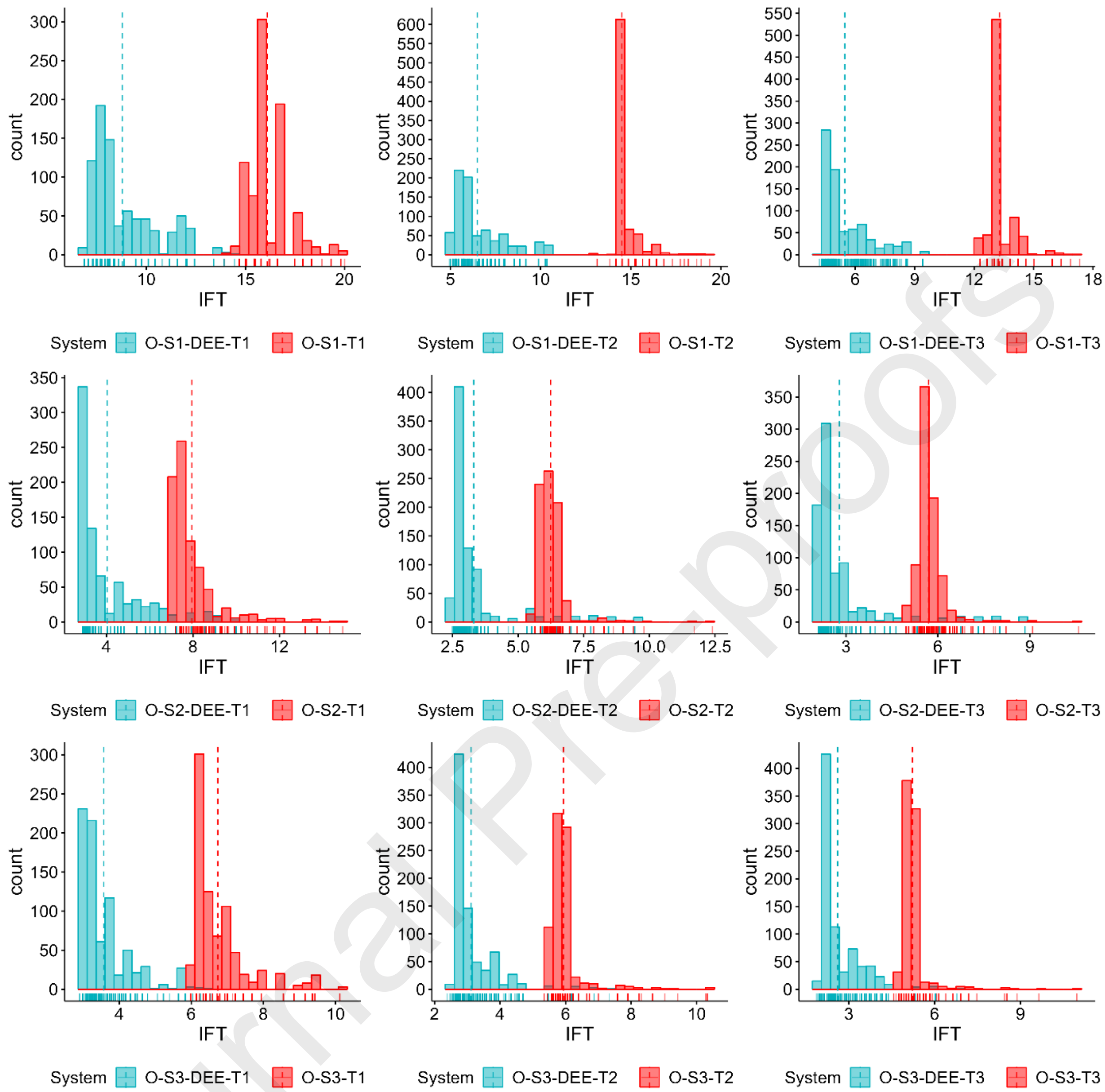


Fig. 6. Histogram depicting the IFT of oil-brine-DEE at salinity levels S_1, S_2, and S_3 and temperatures T_1, T_2, and T_3

3.2 IFT modeling using advanced ML algorithms

In Table 3, the presentation of statistical parameters obtained from each model is showcased. Before executing the models, we partitioned the complete dataset into a 70 % training set and a 30 % testing set. Interestingly, R^2, RMSE, MSE, and MAE indicated that the models generalize well from the training to the testing set, with minimal performance degradation. This suggests that the models not only fit the training data effectively but also maintain their accuracy when applied to test data. Such consistency across training and testing sets is crucial for ensuring the reliability of the models in real-world scenarios. All methods demonstrate satisfactory performance, evident from the high R^2 (>0.71) presented in Table 3. The outcomes, showcased in Figs. 7 and 8, illustrate the estimation results generated by the unified model for both the training and testing sets of the two systems. A visual assessment of the data unveils a noticeable disparity in the model fitting between the two systems. Specifically, in the presence of DEE (Fig. 7), the data points demonstrate a more closely aligned pattern with the model, indicating a superior fit to the training set compared to the base system. This is particularly evident in the reduced dispersion of points, suggesting a more precise depiction of the underlying patterns in the data.

Moreover, a specific observation can be made concerning the RF, BRT, and XGBoost algorithms. In the training dataset, the points associated with these algorithms closely approximate the 45-degree line, which is the ideal line, R^2 is increasing, and the relative error is decreasing, suggesting a high level of accuracy and consistency in their predictions (Figs. 7 and 8). Furthermore, the results consistently exhibit less scattered points in the testing set across all algorithms, indicating the robust generalization performance of the models beyond the training data. To put it more simply, it is clear that there is variability in the performance among the developed models. In accordance with Table 3, the results of the oil-brine-DEE system revealed that RF, BRT, and XGBoost consistently outperformed other models, attaining remarkable R^2 values of 0.99 and highlighting minimal RMSE, MSE, and MAE, emphasizing their robust generalization capabilities. SVM also demonstrated robust predictive capabilities with a high R^2 of 0.94. While GLM and GAM exhibited commendable performance, the ensemble methods stood out with superior accuracy (Table 3). The models exhibiting the least favorable performance were GLM and GAM, producing R^2 values of 0.72 and 0.9, along with RMSE values of 1.23 and 0.73, respectively. This observation is reinforced by the considerable spread of data points from the unit slope line in Figs. 7 and 8.

Regarding oil-brine system, RF, XGBoost, and BRT stood out prominently again for modeling of IFT with perfect R^2 values of 1.00 in both the training and testing sets, highlighting their exceptional ability to explain the variance in the data. These top-performing models consistently demonstrated lower values across all metrics, including RMSE, MSE, and MAE, indicative of minimized errors in their predictions. Specifically, RF displayed the smallest RMSE (0.24) and MSE (0.06) in the testing set, closely followed by XGBoost and BRT. While GAM achieved impressive results, the ensemble methods displayed superior overall performance. The comparison between training and testing sets indicates the robust generalization of RF, XGBoost, and BRT, maintaining consistently low error metrics in both scenarios.

Table 3. Comparative statistical results across developed models

oil-brine-DEE system								
Models	**R^2 Test**	**RMSE Test**	**MSE Test**	**MAE Test**	**R^2 Training**	**RMSE Training**	**MSE Training**	**MAE Training**

GLM	0.72	1.23	1.51	0.98	0.71	1.25	1.56	0.99
GAM	0.90	0.73	0.53	0.58	0.89	0.76	0.57	0.60
RF	0.99	0.20	0.04	0.13	0.99	0.19	0.04	0.12
SVM	0.94	0.61	0.37	0.42	0.93	0.64	0.40	0.42
XGBoost	0.99	0.31	0.10	0.19	0.99	0.32	0.10	0.19
BRT	0.99	0.20	0.04	0.15	0.99	0.19	0.04	0.14

oil-brine system

Models	**R^2 Test**	**RMSE Test**	**MSE Test**	**MAE Test**	**R^2 Training**	**RMSE Training**	**MSE Training**	**MAE Training**
GLM	0.77	1.97	3.88	1.75	0.78	1.94	3.78	1.73
GAM	0.98	0.58	0.34	0.42	0.98	0.56	0.31	0.40
RF	1.00	0.24	0.06	0.17	1.00	0.23	0.05	0.16
SVM	0.97	0.68	0.46	0.53	0.98	0.67	0.44	0.53
XGBoost	1.00	0.32	0.10	0.25	1.00	0.31	0.10	0.25
BRT	1.00	0.28	0.08	0.21	1.00	0.27	0.07	0.20

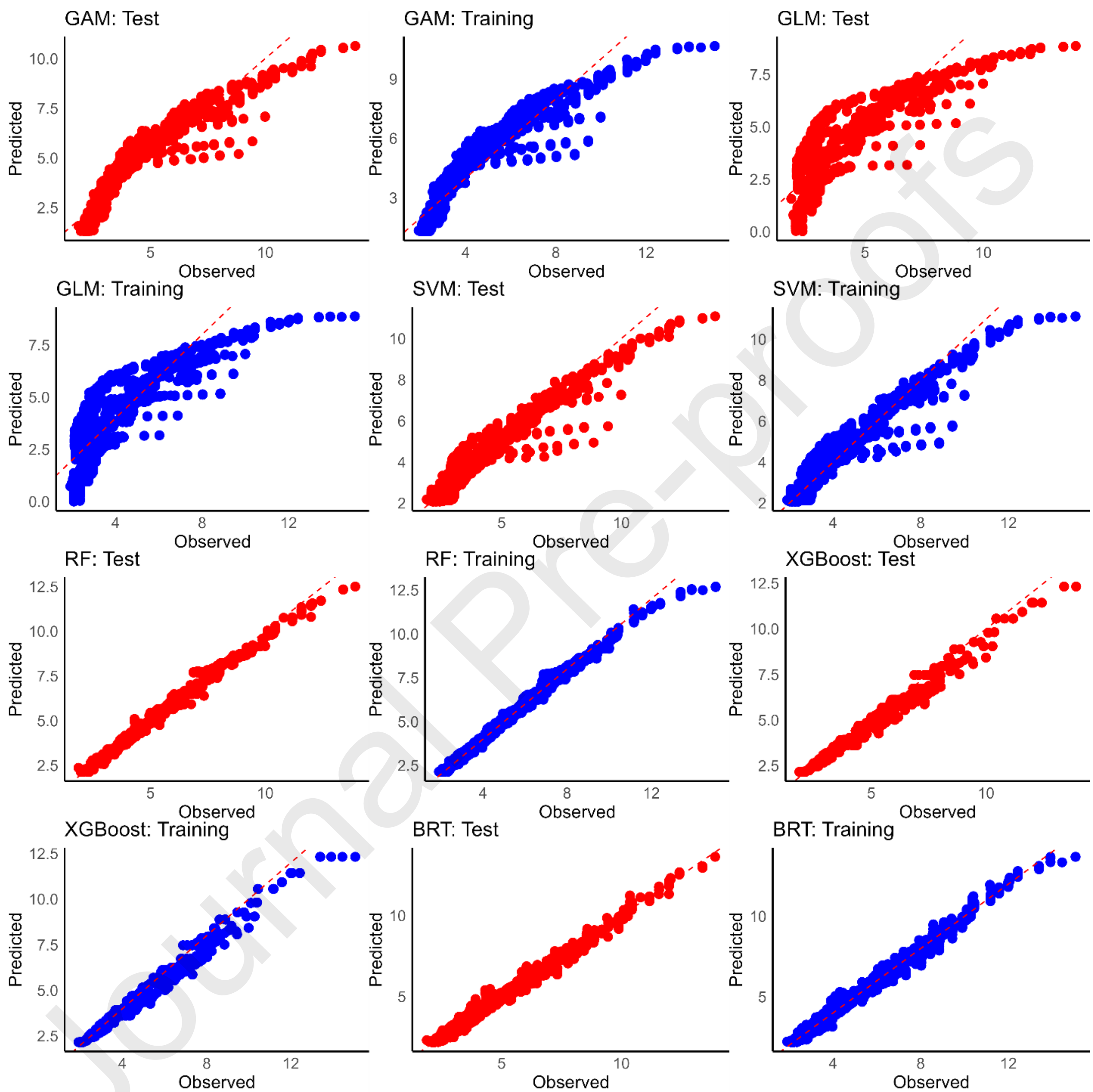


Fig. 7. Comparison of the predicted and observed plots for the IFT of the oil-brine-DEE system using advanced ML models (GAM, GLM, SVM, RF, XGBoost, and BRT)

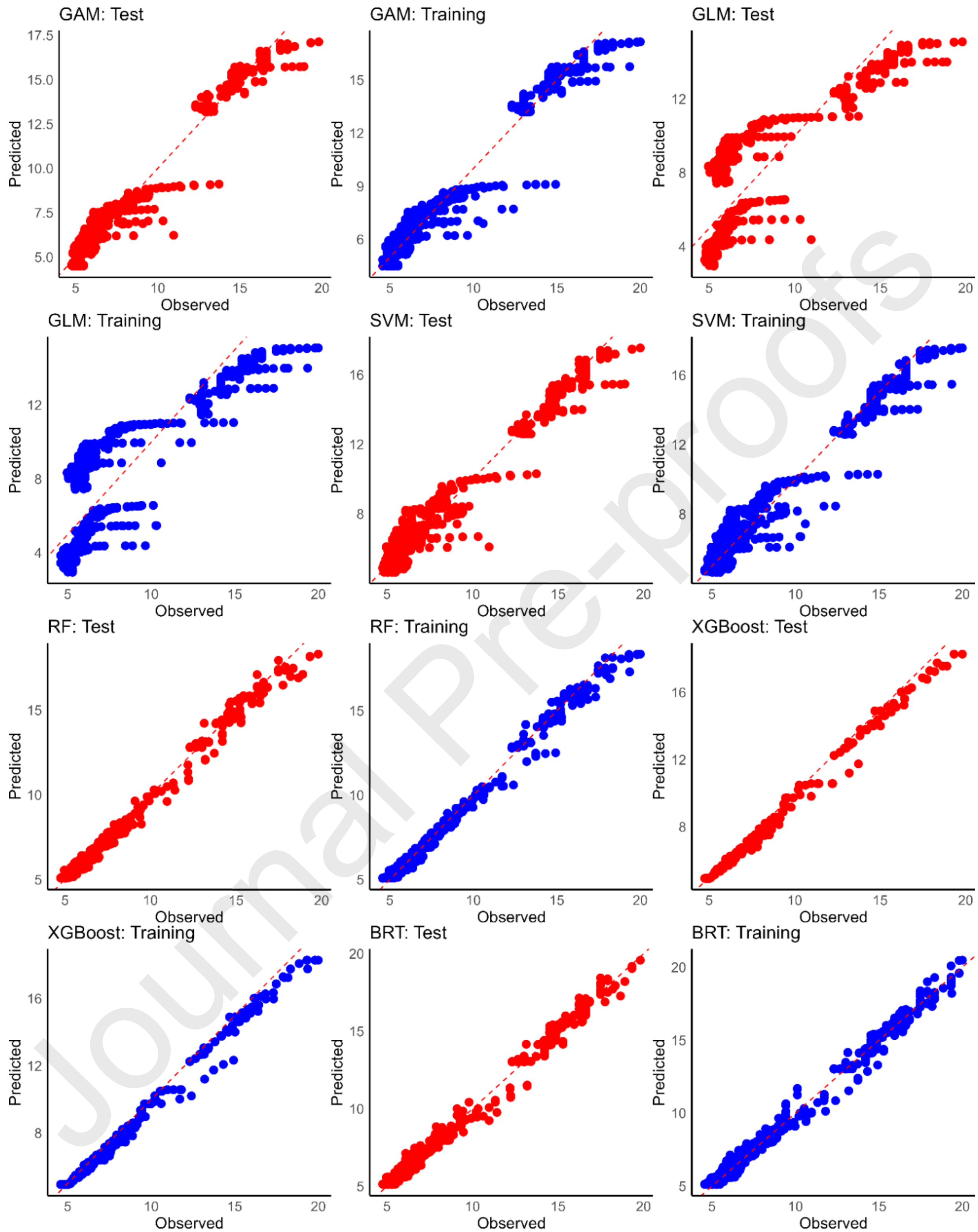


Fig. 8. Comparison of the predicted and observed plots for the IFT of the oil-brine system using advanced ML models (GAM, GLM, SVM, RF, XGBoost, and BRT)

3.3 Validity of proposed models

Emphasizing the significance of the fact that oil and water are commonly found as the primary fluids in reservoirs, most ML models have been created to forecast IFT across different systems. However, this marks the initial application of advanced ML in the oil-brine-DEE system, and all examined models yielded impressive outcomes. Notably, the RF method outperforms alternative approaches, boasting the highest overall R^2 (0.99), as evidenced in Table 3 and illustrated in Fig. 7. In terms of accuracy metrics, the RF model excels with the lowest RMSE (0.2), MSE (0.04), and MAE (0.13). Likewise, the BRT and XGBoost models demonstrate performance on par with the RF system. The BRT model achieves an overall R^2 of 0.99, accompanied by corresponding RMSE, MSE, and MAE values of 0.2, 0.04, and 0.15, respectively. In general, RF stands out as a widely utilized ensemble learning algorithm renowned for its effectiveness in addressing a multitude of classification and regression challenges [106,107]. This algorithm operates by amalgamating predictions from multiple decision trees, resulting in accurate and stable forecasts [108,109]. Noteworthy advantages distinguish RF from other ML algorithms, including its high accuracy and resilience to data noise and outliers [110]. Moreover, RF boasts ease of implementation, and its performance can be enhanced through hyperparameter tuning, such as adjusting the number of decision trees and the number of features selected at each split [111,112]. Furthermore, the XGBoost model secures the third position, delivering a similarly noteworthy overall R^2 of 0.99. The outstanding performance of the XGBoost can be attributed to its inherent technical advantages, which effectively address model variances and mitigate the risk of overfitting [113,114].

The study provides valuable insights into the optimization of salinity levels and the synergistic effects of DEE and soluble ions on IFT reduction. These insights could contribute to the enhanced design and performance of EOR projects. However, there may be a risk of overreliance on ML techniques without sufficient consideration of the underlying physical mechanisms governing IFT behavior in complex oil-brine-solvent systems. This oversight could limit the generalizability and interpretability of the models.

4. Summary and Conclusions

This research explored the application of ML algorithms to predict oil-brine-DEE IFT reliably when complex experimental data is unavailable. A comprehensive dataset comprising 7017 entries for the oil-brine system and 6949 entries for the oil-brine-DEE system was assembled to develop the ML models. Six intelligent models, namely GLM, GAM, SVM, RF, XGBoost, and BRT, were developed to forecast IFT considering temperature and salinity variables at constant pressure. Various statistical metrics such as R^2, RMSE, MSE, and MAE were employed to identify the most robust and reliable model. Additionally, a graphical analysis was conducted to assess the accuracy of the models. The major findings of the study are as follows:

- With increasing salinity and temperature, there was a decrease in IFT. However, this study emphasizes the optimal salinity range for achieving the most substantial reduction in IFT. This is attributed to the facilitation of the migration of polar components, such as asphaltene molecules, to the interface of the oil-brine system. A synergistic reduction in IFT is observed when both DEE and soluble ions are present in the solution, leading to the lowest IFT at approximately 2 mN/m in the S_2 solution at T_3. This suggests that the adsorption of DEE at the water–oil interface forms a layer capable of adsorbing ions, thereby enhancing the layer. Consequently, the oil–solvent-ion layer becomes thicker compared to the oil-ion layer, resulting in the maximum decrease in IFT.

- The RF model proved to be the most accurate in estimating the IFT for both the oil-brine and oil-brine-DEE systems, displaying the highest R^2 (1, 0.99) and the lowest RMSE (0.24, 0.2), MSE (0.06, 0.02), and MAE (0.17, 0.13) compared to the other models suggested. Conversely, two robust intelligent models, namely BRT and XGBoost, also produced reliable predictions for IFT outcomes.
- The GLM technique displayed the least satisfactory performance, as evidenced by the lowest observed R^2 values of (0.72, 0.77) in the two systems, oil-brine and oil-brine-DEE.
- A key benefit of this method is the swift application of ML in the field. Once a ML model is trained and demonstrates effectiveness on unseen data, it serves as a rapid tool for predicting the optimal EOR method for a given reservoir. Additionally, this approach ensures that screening results remain unbiased, devoid of any influence from expert knowledge or opinions.

CRediT authorship contribution statement

Amir Mohammadi: Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Mahsa Parhizgar Keradeh:** Writing – original draft, Software, Data curation, Validation, Conceptualization. **Alireza Keshavarz:** Writing – review & editing, Supervision, Project administration, Conceptualization. **Mohsen Farrokhrouz:** Writing – review & editing, Supervision, Project administration, Conceptualization.

Declaration of Competing Interest

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

Data availability

Data will be made available on request

Acknowledgements

The authors wish to extend their heartfelt gratitude to Dr. Maziar Mohammadi for his invaluable support throughout the development of this paper. His expertise in modeling and insightful advice have made significant contributions to its refinement and overall quality.

Nomenclature

AI	Artificial Intelligence	BRT	Boosted Regression Tree
DEE	Diethyl Ether	DEW	DME-Enhanced Waterflooding
DME	Dimethyl Ether	EOR	Enhance Oil Recovery
GAM	Gradient Additive Model	GLM	Generalized Linear Model
IFT	Interfacial Tension (mN/m)	MAE	Mean Absolute Error (mN/m)
ML	Machine Learning	MSE	Mean Squared Error $(mN/m)^2$
R^2	Coefficient of Determination	RF	Random Forest
RMSE	Root Mean Squared Error (mN/m)	S_1	4000 ppm Solution
S_2	40,000 ppm Solution	S_3	80,000 ppm Solution
SVM	Support Vector Machine	T_1	Temperature of 30°C
T_2	Temperature of 50°C	T_3	Temperature of 70°C
XGBoost	Extreme Gradient Boosting		

References

[1] K. Guo, H. Li, Z. Yu, In-situ heavy and extra-heavy oil recovery: A review, Fuel 185 (2016) 886–902. https://doi.org/10.1016/j.fuel.2016.08.047.

[2] J. Griffin, A.-M. Fantini, World Oil Outlook, OPEC, 2014. https://doi.org/ISBN 978-3-9503936-0-6.

[3] T. Ahmad, D. Zhang, A critical review of comparative global historical energy consumption and future demand: The story told so far, Energy Reports 6 (2020) 1973–1991. https://doi.org/10.1016/j.egyr.2020.07.020.

[4] M. Parhizgar Keradeh, S.A. Tabatabaei-Nezhad, A comprehensive evaluation of the effect of key parameters on the performance of DTPA chelating agent in modifying sandstone surface charge, Heliyon 9 (2023) e21990. https://doi.org/10.1016/J.HELIYON.2023.E21990.

[5] M.R. Mojid, B.M. Negash, H. Abdulelah, S.R. Jufar, B.K. Adewumi, A state – of – art review on waterless gas shale fracturing technologies, J Pet Sci Eng 196 (2021) 108048. https://doi.org/10.1016/j.petrol.2020.108048.

[6] T. Ahmad, D. Zhang, A critical review of comparative global historical energy consumption and future demand: The story told so far, Energy Reports 6 (2020) 1973–1991. https://doi.org/10.1016/j.egyr.2020.07.020.

[7] M.P. KERADEH, S.A. TABATABAEI-NEZHAD, Investigation of the effect of diethylene triamine pentaacetic acid chelating agent as an enhanced oil recovery fluid on wettability alteration of sandstone rocks, Petroleum Exploration and Development 50 (2023) 675–687.

[8] M.P. Keradeh, S.A. Tabatabei-Nezhad, Comprehensive analysis of the effect of reservoir key parameters on the efficacy of DTPA chelating agent in minimizing interfacial tension and enhanced oil recovery, Results in Engineering (2023) 101316.

[9] M. Parhizgar Keradeh, S.A. Tabatabaei-Nezhad, Enhanced Oil Recovery from Heavy Oil Sandstone Reservoirs Using DTPA Chelating Agent/SW Solution, Arab J Sci Eng 48 (2023) 17049–17066. https://doi.org/10.1007/s13369-023-08361-z.

[10] F. Zhang, D.S. Schechter, Gas and foam injection with CO2 and enriched NGL's for enhanced oil recovery in unconventional liquid reservoirs, J Pet Sci Eng 202 (2021) 108472. https://doi.org/10.1016/j.petrol.2021.108472.

[11] M. Almobarak, Z. Wu, D. Zhou, K. Fan, Y. Liu, Q. Xie, A review of chemical-assisted minimum miscibility pressure reduction in CO2 injection for enhanced oil recovery, Petroleum 7 (2021) 245–253. https://doi.org/10.1016/j.petlm.2021.01.001.

[12] C. Parsons, A. Chernetsky, D. Eikmans, P. te Riele, Introducing a Novel Enhanced Oil Recovery Technology, in: SPE Improved Oil Recovery Conference, 2016: pp. 1–10. https://doi.org/10.2118/179560-MS.

[13] A. Alkindi, N. Al-Azri, D. Said, K. AlShuaili, P. Te Riele, Persistence in EOR - Design of a Field Trial in a Carbonate Reservoir using Solvent-based Water-Flood Process, SPE EOR Conference at Oil and Gas West Asia (2016). https://doi.org/10.2118/179838-MS.

[14] R.R. Ratnakar, B. Dindoruk, L. Wilson, Use of DME as an EOR agent: Experimental and modeling study to capture interactions of DME, brine and crudes at reservoir conditions, in: Proceedings - SPE Annual Technical Conference and Exhibition, 2016: pp. 1–15. https://doi.org/10.2118/181515-ms.

[15] J.A.W.M. Groot, A. Chernetsky, P.M. Te Riele, J. Cui, L.C. Wilson, R. Ratnakar, Representation of Phase Behavior and PVT Workflow for DME Enhanced Water-Flooding, SPE EOR Conference at Oil and Gas West Asia, 21-23 March, Muscat, Oman (2016).

[16] M. Mahdizadeh, A.A. Eftekhari, H. M. Nick, Theory and Application of DME Enhanced Waterflooding in Low Permeable Heterogeneous Reservoirs, in: 2018. https://doi.org/10.3997/2214-4609.201801112.

[17] R.R. Ratnakar, B. Dindoruk, L. Wilson, Experimental investigation of DME-water-crude oil phase behavior and PVT modeling for the application of DME-enhanced waterflooding, Fuel 182 (2016) 188–197. https://doi.org/10.1016/j.fuel.2016.05.096.

[18] A. Chernetsky, S. Masalmeh, D. Eikmans, C.A. Parsons, A. Parker, D.M. Boersma, A Novel Enhanced Oil Recovery Technique: Experimental Results and Modelling Workflow of the DME Enhanced Waterflood Technology, in: Abu Dhabi International Petroleum Exhibition and Conference, Society of Petroleum Engineers, 2015. https://doi.org/10.2118/177919-MS.

[19] A. Haddadnia, B. Azinfar, M. Zirrahi, H. Hassanzadeh, J. Abedi, Thermophysical properties of dimethyl ether/Athabasca bitumen system, Canadian Journal of Chemical Engineering 96 (2018) 597–604. https://doi.org/10.1002/cjce.23009.

[20] C. Parsons, A. Chernetsky, D. Eikmans, P. te Riele, Introducing a Novel Enhanced Oil Recovery Technology, in: SPE Improved Oil Recovery Conference, 2016: pp. 1–10. https://doi.org/10.2118/179560-MS.

[21] P. te Riele, C. Parsons, P. Boerrigter, J. Plantenberg, Implementing a Water Soluble Solvent Based Enhanced Oil Recovery Technology - Aspects of Field Development Planning, SPE EOR Conference at Oil and Gas West Asia (2016).

[22] A. Fayazi, A. Kantzas, Determining Diffusivity, Solubility, and Swelling in Gaseous Solvent–Heavy Oil Systems, Ind Eng Chem Res 58 (2019) 10031–10043. https://doi.org/10.1021/acs.iecr.9b01510.

[23] P. te Riele, C. Parsons, P. Boerrigter, J. Plantenberg, Implementing a Water Soluble Solvent Based Enhanced Oil Recovery Technology - Aspects of Field Development Planning, SPE EOR Conference at Oil and Gas West Asia (2016).

[24] A. Alkindi, N. Al-Azri, D. Said, K. AlShuaili, P. Te Riele, Persistence in EOR - Design of a Field Trial in a Carbonate Reservoir using Solvent-based Water-Flood Process, SPE EOR Conference at Oil and Gas West Asia (2016). https://doi.org/10.2118/179838-MS.

[25] D. STEPANENKO, Z. KNEBA, DME as alternative fuel for compression ignition engines – a review, Combustion Engines 177 (2019) 172–179. https://doi.org/10.19206/ce-2019-230.

[26] A. Bakhsh, L. Zhang, H. Wei, A. Shaikh, N. Khan, S. Khan, R. Shaoran, The approach of dimethyl ether-enhanced waterflooding (DEW) for oil recovery: a review, Arabian Journal of Geosciences 15 (2022) 520. https://doi.org/10.1007/s12517-022-09747-3.

[27] N.N. Petrukhina, A.L. Maximov, Use of Dimethyl Ether in Technologies for Enhancing the Oil Recovery from Reservoirs (A Review), Petroleum Chemistry 63 (2023) 67–73. https://doi.org/10.1134/S0965544123020019.

[28] Y.J. Choi, K. Seo, K.S. Lee, Techno-economical optimization of water-alternating-CO 2 /dimethyl ether process for enhanced oil recovery, Pet Sci Technol (2023) 1–19. https://doi.org/10.1080/10916466.2023.2241562.

[29] T.H. Fleisch, A. Basu, M.J. Gradassi, J.G. Masin, Dimethyl ether: a fuel for the 21st century, in: Stud Surf Sci Catal, Elsevier, 1997: pp. 117–125.

[30] B. Bailey, J. Eberhardt, S. Goguen, J. Erwin, Diethyl ether (DEE) as a renewable diesel fuel, SAE Transactions (1997) 1578–1584.

[31] E. Salehi, M.-R. Mohammadi, A. Hemmati-Sarapardeh, V.R. Mahdavi, T. Gentzis, B. Liu, M. Ostadhassan, Modeling Interfacial Tension of N2/CO2 Mixture + n-Alkanes with Machine Learning Methods: Application to EOR in Conventional and Unconventional Reservoirs by Flue Gas Injection, Minerals 12 (2022) 252. https://doi.org/10.3390/min12020252.

[32] H. Vo Thanh, H. Zhang, Z. Dai, T. Zhang, S. Tangparitkul, B. Min, Data-driven machine learning models for the prediction of hydrogen solubility in aqueous systems of varying salinity: Implications for underground hydrogen storage, Int J Hydrogen Energy 55 (2024) 1422–1433. https://doi.org/10.1016/j.ijhydene.2023.12.131.

[33] H. Vo Thanh, Z. Dai, Z. Du, H. Yin, B. Yan, M.R. Soltanian, T. Xiao, B. McPherson, L. Abualigah, Artificial intelligence-based prediction of hydrogen adsorption in various kerogen types: Implications for underground hydrogen storage and cleaner production, Int J Hydrogen Energy 57 (2024) 1000–1009. https://doi.org/10.1016/j.ijhydene.2024.01.115.

[34] A.F. Ibrahim, Prediction of shale wettability using different machine learning techniques for the application of CO2 sequestration, Int J Coal Geol 276 (2023) 104318. https://doi.org/10.1016/j.coal.2023.104318.

[35] M. Rajabi, O. Hazbeh, S. Davoodi, D.A. Wood, P.S. Tehrani, H. Ghorbani, M. Mehrad, N. Mohamadian, V.S. Rukavishnikov, A.E. Radwan, Predicting shear wave velocity from conventional well logs with deep and hybrid machine learning algorithms, J Pet Explor Prod Technol 13 (2023) 19–42.

[36] E. Salehi, M.-R. Mohammadi, A. Hemmati-Sarapardeh, V.R. Mahdavi, T. Gentzis, B. Liu, M. Ostadhassan, Modeling Interfacial Tension of N2/CO2 Mixture + n-Alkanes with Machine Learning Methods: Application to EOR in Conventional and Unconventional Reservoirs by Flue Gas Injection, Minerals 12 (2022) 252. https://doi.org/10.3390/min12020252.

[37] A. Ghosh, D. Chakraborty, A. Law, Artificial intelligence in Internet of things, CAAI Trans Intell Technol 3 (2018) 208–218. https://doi.org/10.1049/trit.2018.1008.

[38] S.A. Kalogirou, Applications of artificial neural-networks for energy systems, Appl Energy 67 (2000) 17–35. https://doi.org/10.1016/S0306-2619(00)00005-2.

[39] P. Ongsulee, Artificial intelligence, machine learning and deep learning, in: 2017 15th International Conference on ICT and Knowledge Engineering (ICT&KE), IEEE, 2017: pp. 1–6. https://doi.org/10.1109/ICTKE.2017.8259629.

[40] C. Janiesch, P. Zschech, K. Heinrich, Machine learning and deep learning, Electronic Markets 31 (2021) 685–695. https://doi.org/10.1007/s12525-021-00475-2.

[41] A. Najafi-Marghmaleki, A. Tatar, A. Barati-Harooni, A. Mohebbi, M. Kalantari-Meybodi, A.H. Mohammadi, On the prediction of interfacial tension (IFT) for water-hydrocarbon gas system, J Mol Liq 224 (2016) 976–990. https://doi.org/10.1016/J.MOLLIQ.2016.10.083.

[42] M.N. Amar, M. Shateri, A. Hemmati-Sarapardeh, A. Alamatsaz, Modeling oil-brine interfacial tension at high pressure and high salinity conditions, J Pet Sci Eng 183 (2019) 106413. https://doi.org/10.1016/j.petrol.2019.106413.

[43] M. Kalantari Meybodi, A. Shokrollahi, H. Safari, M. Lee, A. Bahadori, A computational intelligence scheme for prediction of interfacial tension between pure hydrocarbons and water, Chemical Engineering Research and Design 95 (2015) 79–92. https://doi.org/10.1016/j.cherd.2015.01.004.

[44] M.H.E. Baghdadi, H. Darvish, H. Rezaei, M. Savadinezhad, Applying LSSVM algorithm as a novel and accurate method for estimation of interfacial tension of brine and hydrocarbons, Pet Sci Technol 36 (2018) 1170–1174. https://doi.org/10.1080/10916466.2018.1465963.

[45] H. Mehrjoo, M. Riazi, M.N. Amar, A. Hemmati-Sarapardeh, Modeling interfacial tension of methane-brine systems at high pressure and high salinity conditions, J Taiwan Inst Chem Eng 114 (2020) 125–141. https://doi.org/10.1016/J.JTICE.2020.09.014.

[46] E. Niroomand-Toomaj, A. Etemadi, A. Shokrollahi, Radial basis function modeling approach to prognosticate the interfacial tension CO2/Aquifer Brine, J Mol Liq 238 (2017) 540–544. https://doi.org/10.1016/J.MOLLIQ.2017.04.135.

[47] A. Kamari, M. Pournik, A. Rostami, A. Amirlatifi, A.H. Mohammadi, Characterizing the CO2-brine interfacial tension (IFT) using robust modeling approaches: A comparative study, J Mol Liq 246 (2017) 32–38. https://doi.org/10.1016/J.MOLLIQ.2017.09.010.

[48] M. Nait Amar, Towards improved genetic programming based-correlations for predicting the interfacial tension of the systems pure/impure CO2-brine, J Taiwan Inst Chem Eng 127 (2021) 186–196. https://doi.org/10.1016/j.jtice.2021.08.010.

[49] A. Barati-Harooni, A. Soleymanzadeh, A. Tatar, A. Najafi-Marghmaleki, S.J. Samadi, A. Yari, B. Roushani, A.H. Mohammadi, Experimental and modeling studies on the effects of temperature, pressure and brine salinity on interfacial tension in live oil-brine systems, J Mol Liq 219 (2016) 985–993. https://doi.org/10.1016/J.MOLLIQ.2016.04.013.

[50] A. Kirch, Y.M. Celaschi, J.M. de Almeida, C.R. Miranda, Brine–oil interfacial tension modeling: assessment of machine learning techniques combined with molecular dynamics, ACS Appl Mater Interfaces 12 (2020) 15837–15843.

[51] F. Yousefmarzi, A. Haratian, J. Mahdavi Kalatehno, M. Keihani Kamal, Machine learning approaches for estimating interfacial tension between oil/gas and oil/water systems: a performance analysis, Sci Rep 14 (2024). https://doi.org/10.1038/s41598-024-51597-4.

[52] J.A.W.M. Groot, A. Chernetsky, P.M. Te Riele, J. Cui, L.C. Wilson, R. Ratnakar, Representation of Phase Behavior and PVT Workflow for DME Enhanced Water-Flooding, SPE EOR Conference at Oil and Gas West Asia, 21-23 March, Muscat, Oman (2016).

[53] M. Khalifi, M. Zirrahi, H. Hassanzadeh, J. Abedi, Measurements of Molecular Diffusion Coefficient and Solubility of Dimethyl Ether in Bitumen at T = (323.15–383.15 K) and P = (0.69–2.76 MPa), J Chem Eng Data 64 (2019) 5935–5945. https://doi.org/10.1021/acs.jced.9b00763.

[54] A. Fayazi, A. Kantzas, Determining Diffusivity, Solubility, and Swelling in Gaseous Solvent–Heavy Oil Systems, Ind Eng Chem Res 58 (2019) 10031–10043. https://doi.org/10.1021/acs.iecr.9b01510.

[55] S. Kong, G. Feng, Y. Liu, K. Li, Potential of dimethyl ether as an additive in CO2 for shale oil recovery, Fuel 296 (2021) 120643. https://doi.org/10.1016/J.FUEL.2021.120643.

[56] H. Javanmard, M. Seyyedi, S.A. Jones, S.M. Nielsen, Dimethyl Ether Enhanced Oil Recovery in Fractured Reservoirs and Aspects of Phase Behavior, Energy and Fuels 33 (2019) 10718–10727. https://doi.org/10.1021/acs.energyfuels.9b02600.

[57] H. Javanmard, M. Seyyedi, S.M. Nielsen, On Oil Recovery Mechanisms and Potential of DME–Brine Injection in the North Sea Chalk Oil Reservoirs, Ind Eng Chem Res 57 (2018) 15898–15908. https://doi.org/10.1021/acs.iecr.8b04278.

[58] R.R. Ratnakar, B. Dindoruk, L.C. Wilson, Phase behavior experiments and PVT modeling of DME-brine-crude oil mixtures based on Huron-Vidal mixing rules for EOR

applications, Fluid Phase Equilib 434 (2017) 49–62. https://doi.org/https://doi.org/10.1016/j.fluid.2016.11.021.

[59] J. Cho, T.H. Kim, K.S. Lee, Compositional modeling and simulation of dimethyl ether (DME)-enhanced waterflood to investigate oil mobility improvement, Pet Sci 15 (2018) 297–304. https://doi.org/10.1007/s12182-017-0212-z.

[60] A. Chernetsky, S. Masalmeh, D. Eikmans, C.A. Parsons, A. Parker, D.M. Boersma, A Novel Enhanced Oil Recovery Technique: Experimental Results and Modelling Workflow of the DME Enhanced Waterflood Technology, in: Abu Dhabi International Petroleum Exhibition and Conference, Society of Petroleum Engineers, 2015. https://doi.org/10.2118/177919-MS.

[61] M. Windholz, S. Budavari, R.F. Blumetti, E.S. Otterbein, The Merck Index, Merck & Co, Inc., Rahway, NJ 1051 (1983).

[62] A.W. Adamson, A.P. Gast, others, Physical chemistry of surfaces, Interscience publishers, New York, 1967.

[63] U. Olsson, Generalized linear models, An Applied Approach. Studentlitteratur, Lund 18 (2002).

[64] M. Lyons, Generalised additive models (gams): An introduction, in: Environ. Comput, 2018.

[65] S. Arcuti, C. Calculli, A. Pollice, G. D'Onghia, P. Maiorano, A. Tursi, Spatio-temporal modelling of zero-inflated deep-sea shrimp data by Tweedie generalized additive, Statistica 73 (2013) 87–101.

[66] K. Jones, S. Almond, Moving out of the linear rut: the possibilities of generalized additive models, Transactions of the Institute of British Geographers (1992) 434–447.

[67] S. Arcuti, C. Calculli, A. Pollice, G. D'Onghia, P. Maiorano, A. Tursi, Spatio-temporal modelling of zero-inflated deep-sea shrimp data by Tweedie generalized additive, Statistica 73 (2013) 87–101.

[68] V.N. Vapnik, Pattern recognition using generalized portrait method, Automation and Remote Control 24 (1963) 774–780.

[69] C. Cortes, V. Vapnik, Support-vector networks, Mach Learn 20 (1995) 273–297.

[70] S. Misra, Y. Wu, Machine learning assisted segmentation of scanning electron microscopy images of organic-rich shales with feature extraction and feature ranking, Machine Learning for Subsurface Characterization 289 (2019) 4.

[71] T. Shaikhina, D. Lowe, S. Daga, D. Briggs, R. Higgins, N. Khovanova, Decision tree and random forest models for outcome prediction in antibody incompatible kidney transplantation, Biomed Signal Process Control 52 (2019) 456–462.

[72] L. Breiman, Random forests, Mach Learn 45 (2001) 5–32.

[73] T. Chen, C. Guestrin, XGBoost: A scalable tree boosting system, in: Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM, 2016: pp. 785–794. https://doi.org/10.1145/2939672.2939785.

[74] M. Ye, L. Zhu, X. Li, Y. Ke, Y. Huang, B. Chen, H. Yu, H. Li, H. Feng, Estimation of the soil arsenic concentration using a geographically weighted XGBoost model based on hyperspectral data, Science of The Total Environment 858 (2023) 159798. https://doi.org/10.1016/J.SCITOTENV.2022.159798.

[75] J. Elith, J.R. Leathwick, T. Hastie, A working guide to boosted regression trees, Journal of Animal Ecology 77 (2008) 802–813. https://doi.org/10.1111/j.1365-2656.2008.01390.x.

[76] D. Saha, P. Alluri, A. Gan, Prioritizing Highway Safety Manual's crash prediction variables using boosted regression trees, Accid Anal Prev 79 (2015) 133–144.

[77] J. Elith, J.R. Leathwick, T. Hastie, A working guide to boosted regression trees, Journal of Animal Ecology 77 (2008) 802–813. https://doi.org/10.1111/j.1365-2656.2008.01390.x.

[78] G. De'Ath, Boosted trees for ecological modeling and prediction, Ecology 88 (2007) 243–251.

[79] J. Döpke, U. Fritsche, C. Pierdzioch, Predicting recessions with boosted regression trees, Int J Forecast 33 (2017) 745–759.

[80] D. Saha, P. Alluri, A. Gan, Prioritizing Highway Safety Manual's crash prediction variables using boosted regression trees, Accid Anal Prev 79 (2015) 133–144.

[81] J.H. Friedman, Greedy function approximation: a gradient boosting machine, Ann Stat (2001) 1189–1232.

[82] J.H. Friedman, Stochastic gradient boosting, Comput Stat Data Anal 38 (2002) 367–378.

[83] G. De'Ath, Boosted trees for ecological modeling and prediction, Ecology 88 (2007) 243–251.

[84] R. Taherdangkoo, A. Tatomir, M. Taherdangkoo, P. Qiu, M. Sauter, Nonlinear autoregressive neural networks to predict hydraulic fracturing fluid leakage into shallow groundwater, Water (Basel) 12 (2020) 841.

[85] J. Zhang, Y. Sun, L. Shang, Q. Feng, L. Gong, K. Wu, A unified intelligent model for estimating the (gas + n-alkane) interfacial tension based on the eXtreme gradient boosting (XGBoost) trees, Fuel 282 (2020) 118783. https://doi.org/10.1016/j.fuel.2020.118783.

[86] J. Zhang, Q. Feng, S. Wang, X. Zhang, S. Wang, Estimation of CO2–brine interfacial tension using an artificial neural network, J Supercrit Fluids 107 (2016) 31–37. https://doi.org/10.1016/j.supflu.2015.08.010.

[87] C. Hu, N.C. Garcia, R. Xu, Interfacial Properties of Asphaltenes at the Heptol-Brine Interface, Energy and Fuels 30 (2016) 80–87. https://doi.org/10.1021/acs.energyfuels.5b01855.

[88] M.E. Hassan, R.F. Nielsen, J.C. Calhoun, others, Effect of pressure and temperature on oil-water interfacial tensions for a series of hydrocarbons, Journal of Petroleum Technology 5 (1953) 299–306.

[89] B. Kumar, H. Yarranton, E. Baydak, Effect of Salinity on the Interfacial Tension of Crude Oil, in: University of Calgary, 2012. https://doi.org/10.3997/2214-4609.20143757.

[90] C. Jian, M.R. Poopari, Q. Liu, N. Zerpa, H. Zeng, T. Tang, Mechanistic Understanding of the Effect of Temperature and Salinity on the Water/Toluene Interfacial Tension, Energy & Fuels 30 (2016) 10228–10235. https://doi.org/10.1021/acs.energyfuels.6b01995.

[91] Z. Chen, X. Zhao, Enhancing Heavy-Oil Recovery by Using Middle Carbon Alcohol-Enhanced Waterflooding, Surfactant Flooding, and Foam Flooding, Energy & Fuels 29 (2015) 2153–2161. https://doi.org/10.1021/ef502652a.

[92] P. Armenante, H. Karlsson, Salting-out parameters for organic acids, American Chemical Society, 1982. https://doi.org/10.1021/je00028a016.

[93] J.-M. Bai, W.-Y. Fan, G.-Z. Nan, S.-P. Li, B.-S. Yu, Influence of interaction between heavy oil components and petroleum sulfonate on the oil--water interfacial tension, J Dispers Sci Technol 31 (2010) 551–556.

[94] F. Moeini, A. Hemmati-sarapardeh, M. Ghazanfari, M. Masihi, S. Ayatollahi, Towards Mechanistic Understanding of Heavy Crude Oil/Brine Interfacial Tension: the Roles of Salinity, Temperature and Pressure, Fluid Phase Equilib 375 (2014) 191–200. https://doi.org/10.1016/j.fluid.2014.04.017.

[95] T. Al-Sahhaf, A. Elkamel, A. Suttar Ahmed, A.R. Khan, The influence of temperature, pressure, salinity, and surfactant concentration on the interfacial tension of the n-octane-water system, Chem Eng Commun 192 (2005) 667–684.

[96] B. Kumar, H. Yarranton, E. Baydak, Effect of Salinity on the Interfacial Tension of Crude Oil, in: University of Calgary, 2012. https://doi.org/10.3997/2214-4609.20143757.

[97] I. Nowrouzi, A.K. Manshad, A.H. Mohammadi, Effects of TiO2, MgO, and γ-Al2O3 nano-particles in carbonated water on water-oil interfacial tension (IFT) reduction in chemical enhanced oil recovery (CEOR) process, J Mol Liq 292 (2019) 111348. https://doi.org/10.1016/j.molliq.2019.111348.

[98] A. Esfandiarian, A. Maghsoudian, M. Shirazi, Y. Tamsilian, S. Kord, J.J. Sheng, Mechanistic Investigation of the Synergy of a Wide Range of Salinities and Ionic Liquids for Enhanced Oil Recovery: Fluid–Fluid Interactions, Energy & Fuels 35 (2021) 3011–3031. https://doi.org/10.1021/acs.energyfuels.0c03371.

[99] A. Nourinia, A.K. Manshad, S.R. Shadizadeh, J.A. Ali, S. Iglauer, A. Keshavarz, A.H. Mohammadi, M. Ali, Synergistic Efficiency of Zinc Oxide/Montmorillonite Nanocomposites and a New Derived Saponin in Liquid/Liquid/Solid Interface-Included Systems: Application in Nanotechnology-Assisted Enhanced Oil Recovery, ACS Omega 7 (2022) 24951–24972. https://doi.org/10.1021/acsomega.1c07182.

[100] J. Kaliyugarasan, Surface Chemistry Study of Low Salinity Waterflood, Centre for Integrated Petroleum Research (Uni CIPR), Department of Chemistry, University of Bergen, 2013.

[101] J.-M. Bai, W.-Y. Fan, G.-Z. Nan, S.-P. Li, B.-S. Yu, Influence of interaction between heavy oil components and petroleum sulfonate on the oil--water interfacial tension, J Dispers Sci Technol 31 (2010) 551–556.

[102] I. Nowrouzi, A.H. Mohammadi, A.K. Manshad, Utilization of methanol and acetone as mutual solvents to reduce interfacial tension (IFT) in enhanced oil recovery process by carbonated smart water injection, J Mol Liq 304 (2020) 112733. https://doi.org/10.1016/j.molliq.2020.112733.

[103] I. Nowrouzi, A.H. Mohammadi, A.K. Manshad, Effects of a ketone mutual solvent on the dynamic and equilibrium behaviors of crude oil swelling in enhanced oil recovery process by carbonated seawater flooding, J Pet Sci Eng 196 (2021) 108005. https://doi.org/10.1016/j.petrol.2020.108005.

[104] N.S. Al Maskari, A. Saeedi, Q. Xie, Alcohol-Assisted Waterflooding in Carbonate Reservoirs, Energy & Fuels 33 (2019) 10651–10658. https://doi.org/10.1021/acs.energyfuels.9b02497.

[105] H. Ding, N. Zhang, Y. Zhang, M. Wei, B. Bai, Experimental Data Analysis of Nanoparticles for Enhanced Oil Recovery, Ind Eng Chem Res 58 (2019) 12438–12450. https://doi.org/10.1021/acs.iecr.9b02132.

[106] A.A. Silva, M.W. Tavares, A. Carrasquilla, R. Misságia, M. Ceia, Petrofacies classification using machine learning algorithms, Geophysics 85 (2020) WA101--WA113.

[107] F. Anifowose, J. Labadin, A. Abdulraheem, Improving the prediction of petroleum reservoir characterization with a stacked generalization ensemble model of support vector machines, Appl Soft Comput 26 (2015) 483–496.

[108] D.S. Palmer, N.M. O'Boyle, R.C. Glen, J.B.O. Mitchell, Random forest models to predict aqueous solubility, J Chem Inf Model 47 (2007) 150–158.

[109] V. Svetnik, A. Liaw, C. Tong, J.C. Culberson, R.P. Sheridan, B.P. Feuston, Random forest: a classification and regression tool for compound classification and QSAR modeling, J Chem Inf Comput Sci 43 (2003) 1947–1958.

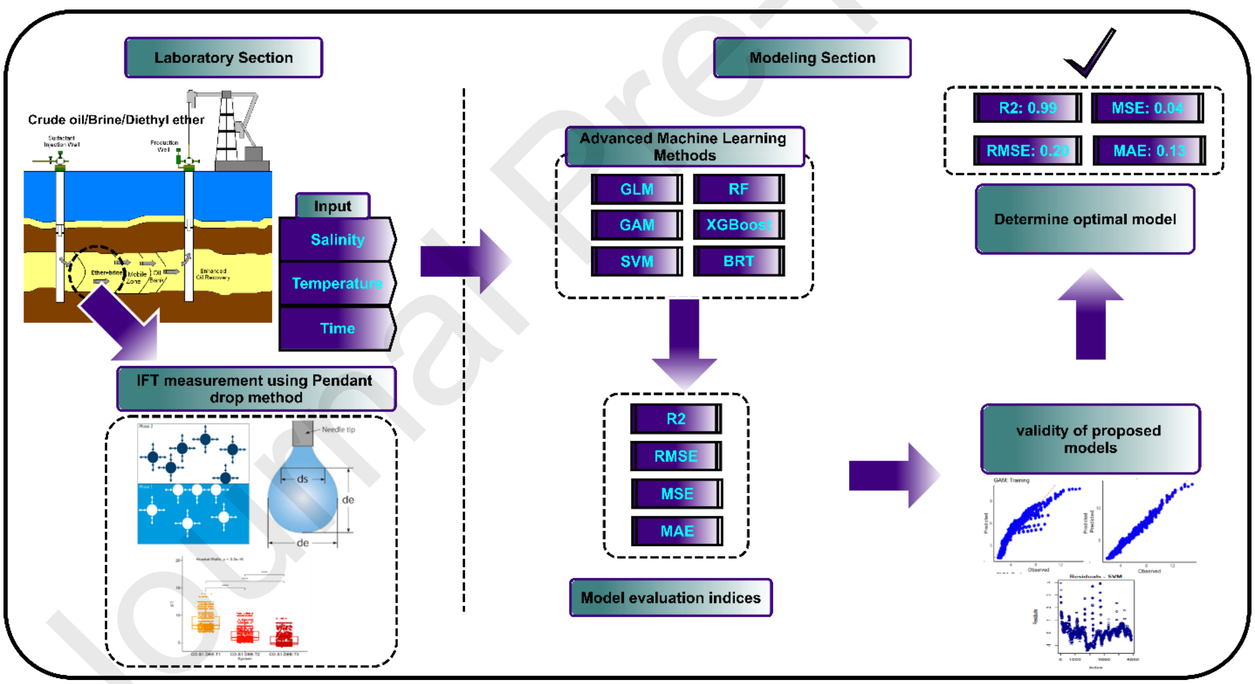
[110] P.M. Granitto, C. Furlanello, F. Biasioli, F. Gasperi, Recursive feature elimination with random forest for PTR-MS analysis of agroindustrial products, Chemometrics and Intelligent Laboratory Systems 83 (2006) 83–90.

[111] L. Breiman, Random forests, Mach Learn 45 (2001) 5–32.

[112] A. Kumar, S. Ridha, T. Ganet, P. Vasant, S.U. Ilyas, Machine learning methods for herschel--bulkley fluids in annulus: Pressure drop predictions and algorithm performance evaluation, Applied Sciences 10 (2020) 2588.

[113] T. Chen, T. He, M. Benesty, V. Khotilovich, Y. Tang, H. Cho, K. Chen, R. Mitchell, I. Cano, T. Zhou, others, Xgboost: extreme gradient boosting, R Package Version 0.4-2 1 (2015) 1–4.

[114] J. Zhang, Y. Sun, L. Shang, Q. Feng, L. Gong, K. Wu, A unified intelligent model for estimating the (gas + n-alkane) interfacial tension based on the eXtreme gradient boosting (XGBoost) trees, Fuel 282 (2020) 118783. https://doi.org/10.1016/j.fuel.2020.118783.



Research Highlights

- The impact of diethyl ether on the interfacial tension (IFT) of crude oil-brine was examined under varying salinity and temperature conditions.
- An experimental database comprising 7,017 sets of crude oil/brine/DEE IFT data was acquired.
- Six advanced machine-learning models were developed to accurately estimate the IFT of Crude oil–Brine-DEE.

- Statistical analysis was conducted, demonstrating outstanding predictions for a wide range of input variables.

CRediT authorship contribution statement

Amir Mohammadi: Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Mahsa Parhizgar Keradeh:** Writing – original draft, Software, Data curation, Validation, Conceptualization. **Alireza Keshavarz:** Writing – review & editing, Supervision, Project administration, Conceptualization. **Mohsen Farrokhrouz:** Writing – review & editing, Supervision, Project administration, Conceptualization.

Declaration of interests

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

☐ The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: