Machine Learning Modeling of Wettability and Contact Angle Behavior in CO_2-Water-Rock System

by

Yernar Tiyntayev

THESIS SUPERVISOR

Associate Professor Dr. Masoud Riazi

THESIS CORE SUPERVISOR

Assistant Professor Dr. Shams Kalam

Thesis submitted to the School of Mining and Geosciences of Nazarbayev University in Partial Fulfillment of the Requirements for the Degree of
Master of Science in Petroleum Engineering

Nazarbayev University
10.04.2025

ORIGINALITY STATEMENT

I, Yernar Tiyntayev, hereby declare that this submission is my work and to the best of my knowledge it contains no materials previously published or written by another person, or substantial proportions of material which have been accepted for the award of any other degree or diploma at Nazarbayev University or any other educational institution, except where due acknowledgement is made in the thesis.

Any contribution made to the research by others, with whom I have worked at NU or elsewhere, is explicitly acknowledged in the thesis.

I also declare that the intellectual content of this thesis is the product of my work, except to the extent that assistance from others in the project's design and conception, or style, presentation, and linguistic expression is acknowledged.

Signed on 10.04.2025

ABSTRACT

The focus of this research project is to understand how machine learning algorithms work with regarding wettability and contact angles in CO_2-water-rock combinations. A broad dataset containing static contact angles along with advancing and receding contact angle measurements originated from published experimental studies that studied different types of rocks across a range of pressures, temperatures, and salinities. The analysis implemented six machine learning algorithms, including Decision Trees, Random Forests, XGBoost, Gradient Boosting Regressor, Support Vector Machines, and Artificial Neural Networks. Both training and testing datasets received evaluation through different error metrics, which assessed the performance of the models. The GBR model delivered optimum performance in static contact angle prediction for CO_2-water-rock systems by reaching $R^2 = 0.99$ value during training and $R^2 = 0.92$ value during testing. The high accuracy value shows that GBR effectively identifies intricate non-linear patterns in wettability patterns. The GBR model produced the most accurate results for testing dataset advancing and receding contact angles with a calculation error rate of $R^2 = 0.93$ training and $R^2 = 0.88$ testing. GBR model demonstrate excellent proficiency for understanding the changing behavior of liquid wetting in such systems. Different input conditions such as pressures and temperatures and rock types and salinities were subject to sensitivity analysis to check model prediction accuracy and to investigate the impact on the advancing, receding and static contact angles. The research shows that pressure produces the largest impact on contact angle measurement whereas temperature and salinity impacts differ based on the rock type. The predictive capabilities of developed models worked effectively in predicting the contact angles of silicate, clay, carbonate and basalt and effectively captured their complex non-linear relationships.

ACKNOWLEDGMENT

I am thankful to my thesis supervisor Associate Professor Dr. Masoud Riazi for his support during my master’s research along with his expert supervision. I am entirely thankful for the guidance provided by my core supervisor Assistant Professor Dr. Shams Kalam whose deep insights led to the successful conclusion of my thesis work.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

1.1 Background

The crucial aspect of CO_2 sequestration project performance depends on wettability, which scientists measure through contact angle analysis. Recent research has proven that factors such as pressure and salinity, alongside temperature and rock mineralogy define the contact angles within CO_2-brine-rock systems (Arif et al., 2016). Study results indicate that pressure tends to enhance CO_2 wetness by increasing contact angles between rocks and water (Broseta et al., 2012). The solid CO_2 exhibits a greater affinity for CO_2 at high pressures because of increased CO_2 density. Research findings about temperature influence on wettability present conflicting results since some reports indicate lower contact angles occur at elevated temperatures (Arif et al., 2017), but other studies show unchanged or elevated angles (Al-Yaseri et al., 2016). The inconsistent findings demonstrate the multiple reciprocal relationships that exist between thermal conditions and wettability modifications. The research by Arif et al. (2016) shows that elevated salt levels promote higher contact angles, which create more CO_2-wet conditions. The modification of interfacial tensions and rock surface electrical double layers affects this phenomenon according to Chiquet et al. (2009). Rock mineral composition determines how CO_2-brine solutions affect wettability behaviour, as each mineral displays unique reactions to these solutions. Wettability behaviour stands distinct between quartz and calcite minerals even when subject to similar conditions (Iglauer, 2017). The wettability characteristics change dramatically when organic matter and surface roughness exist together (Al-Arif et al., 2017). The exact relationships between wettability behaviour factors remain unclear because advanced modelling methods are required to analyze their complex interactions. Advanced modeling methods allow for predicting wettability more accurately because they handle experimental challenges that occur under different conditions. ANNs have proven effective at identifying complex patterns and self-learning operations through which they analyze wettability complexity (Goodfellow et al., 2016). XGBoost demonstrates superior application performance by providing high-accuracy contact angle predictions through parameter optimization, thus enabling users to achieve effective predictions (Chen & Guestrin, 2016). Implementation of machine learning techniques helps to create better predictive models that model the complicated interactions between CO_2-brine-rock type systems. The developed models will advance the analysis of wettability characteristics during CO_2 storage.

1.2 Objectives of the Thesis

The main research objective is to develop and evaluate machine learning models for predicting wettability patterns and contact angle behavior in CO_2-brine-rock systems. Specific objectives include:

1. Collect a comprehensive set of experimental contact angle data points for CO_2-brine-rock systems from published scientific literature.

2. Implement multiple machine learning algorithms using the collected experimental datasets to predict static, advancing, and receding contact angles. These models will use different rock types, pressure, temperature, and brine salinity as input parameters.

3. Conduct sensitivity analyses to assess how input parameters (pressure, temperature, rock type, and salinity) affect the prediction of advancing, receding, and static contact angles.

4. Evaluate the accuracy of six machine learning models using both training and testing datasets. Performance will be assessed using error metrics including RMSE, MSE, R^2, AIC, and BIC.

1.3 Scope of Work

The chapters of the thesis are structured as follows:

Chapter 1. Introduction: The chapter introduces the research by explaining how wettability and contact angle behavior influence the behavior of CO_2-water-rock systems. The discussion focuses on established measurement problems and demonstrates how machine learning improves prediction results. The chapter also presents the research objectives alongside the problem statement and the significance of the study.

Chapter 2. Literature Review: This chapter reviews previous studies on wettability and contact angle behavior in CO_2-brine-rock systems. The importance of different types of contact angles is explained in detail. In addition, different experimental methods for contact angle measurement and the entire measurement process for each technique are explained. Furthermore, it studies the factors affecting contact angles, including pressure, temperature, salinity, and rock type. The causes of contact angle hysteresis are also explained in detail. The working process of 6 different ML models and their hyperparameters was described. At the end of the chapter, the application of different ML models for predicting the wettability of mineral/CO_2/brine systems was described in detail.

Chapter 3. Methodology: The data collection process was described in detail, including the sources of experimental contact angle data sets. Each source provided detailed information regarding the lead author's name and publication year, together with contact angle measurement methods and overall wettability patterns, and studied rock types as well as experimental pressures and temperatures and brine salinity, and the number of data points obtained from contact angle measurements. Then, there were pre-processing methods using pair plots, statistical analysis, box plots, and heat maps to remove outliers from the dataset. This contributes to a more accurate prediction. Also at this stage, the parameters of pressure, temperature, rock type and salinity were converted into a single unit of measurement for the convenience of training ML models. At the data processing stage, 6 machine learning models were briefly described, their hyperparameters, and the error metrics that will be used to evaluate each model.

Chapter 4. Results and Discussion: Sensitivity analyses were performed to evaluate how the input parameters of pressure, temperature, rock type, and salinity affected the prediction results of advancing, receding, and static contact angle. A detailed discussion was provided for each sensitivity analysis. In addition, the accuracy of the six machine learning models was assessed on the training and test datasets using the RMSE, MSE, R^2, AIC, and BIC error metrics.

Chapter 5. Conclusion: The main findings are summarized, highlighting the effectiveness of each ML in predicting the behavior of the contact angle. It discusses the future plan to improve the different models for more accurate prediction. Recommendations for the application area of the developed ML models were given.

1.4 Significance of the study

Experimental methods for contact angle measurement are time-consuming and expensive. In addition, experimental methods require high-pressure and high-temperature chambers, precise imaging systems, and long equilibration times, making them both labor-intensive and expensive. Using an alternative method such as machine learning, it is possible to quickly and more accurately predict the contact angle for a system. In addition, machine learning can significantly reduce costs by not using experimental equipment and materials, thereby reducing resources. There will be no need for repeated laboratory experiments. The implementation of ML models will improve decision-making for CO_2 storage and wettability analysis.

2. LITERATURE REVIEW

2.1 Introduction to contact angle

According to Erbil (2014) wettability represents the essential characteristic of solid surfaces which determines their ability to attract liquids while defining their ability to spread (adhesion). Contact angle measurement provides the main technique to quantify wettability by determining the angular relationship between solid surface and liquid-vapor interface tangent at the three-phase contact point (Kwok & Neumann, 1999). Young's Equation 1 relates the contact angle to the tensions at the solid-liquid-vapor phase boundaries.

$$cos\theta = \frac{\gamma_{sv}-\gamma_{sl}}{\gamma_{lv}} \tag{1}$$

Young established Equation 1 that combines solid-vapor (γ_{sv}), solid-liquid (γ_{sl}) and liquid-vapor (γ_{lv}) interfacial tensions. Equation 1 serves as a theoretical basis which explains the connection between surface energies and how they affect wettability. Figure 1 illustrates the different types of contact angles.

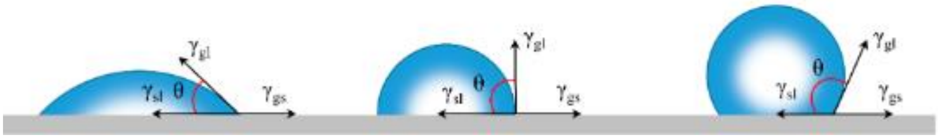


Figure 1. Illustration of contact angles (Guan et al., 2019)

2.2 Importance of static contact angle

The static contact angle (θ) describes the interface angle between a solid surface and the liquid-gas tangent at the three-phase contact line according to Yuan & Lee (2013). Static contact angle measurement in CO_2-water-rock systems determines the surface preference of rock materials between water and CO_2 by evaluating angles below or above 90° wherein below 90° indicates water-wet conditions (Broseta et al., 2012). This type of measurements remains popular because they offer simple and reproducible operation even though they have clear performance restrictions. The obtained contact angle measurements establish an equilibrium state which fails to accurately depict the ongoing fluid interactions within real geologic formations like those found in porous media (Al-Menhali & Krevor, 2016). Laboratory measurements of contact angles exhibit significant variation when compared to in-situ behavior because surface roughness and chemical heterogeneity along with pore geometry influence the obtained values (AlRatrout et al., 2018).

2.3 Importance of advancing and receding contact angle

Advancing contact angle measurement is essential for fluid invasion alongside capillary pressure relationships (Iglauer et al., 2015). The advancing contact angle (θa) describes the angle developed by a solid surface and the tangent that touches the liquid-gas interface during liquid movement across solid material (Yuan & Lee, 2013). The measurement of advancing contact angle in CO_2-water-rock systems shows a value larger than static contact angle because it indicates fluid flow resistance in porous media. Measurements of the advancing contact angle hold significant importance due to multiple key factors. The measurement of advancing contact angles gives crucial information about how well CO_2 performs during EOR operations. The measurement of advancing contact angles supports the estimation of trapping mechanisms involved in geological CO_2 storage (Gershenzon et al., 2017). Analysis of advancing contact angles enables the determination of existing CO_2 leakage potential through storage formation cap rocks (Chiquet et al., 2007). Receding contact angle shows how easily it can separate from the rock surface during drainage. It affects capillary trapping. In addition, receding contact angle is necessary to retain liquids in the porous medium. A small angle shows that the efficiency of CO_2 trapping in a geological repository will be high. A large angle can lead to a weakening of capillary forces and the risk of CO_2 leakage. As shown in Figure 2, the illustration demonstrates the advancing and receding contact angles.

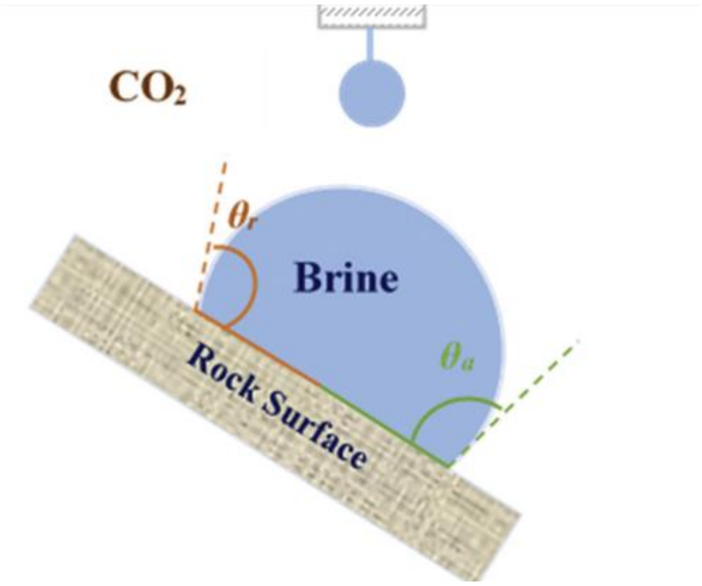


Figure 2. Illustration demonstrating the advancing and receding contact angles (Arif et al., 2019)

2.4 Contact angle hysteresis

Contact angle hysteresis occurs when there is a difference between the advancing and receding contact angles. This phenomenon is influenced by very different factors. Rocks of sandstone and carbonate nature exhibit inherent surface roughness, which produces local variations of surface wettability properties, thus pinning the contact point and increasing fluid flow resistance (Espinoza & Santamarina, 2010). Fluid displacement efficiency is modified by the surface energy distribution changes resulting from the chemical heterogeneity in carbonate reservoirs that contains quartz, calcite, and clays (Wang et al., 2013). As CO_2 dissolves in water, it causes a pH decrease, which initiates chemical reactions between minerals and produces new surface compositions that adjust advancing and receding contact angles (Iglauer et al., 2015). When CO_2 adsorbs onto mineral surfaces of carbonate reservoirs, it leads to wettability alteration into CO_2-wet conditions that affect both capillary trapping and residual saturation behavior (Arif et al., 2017). The fluctuations of capillary pressure together with varying pressure and temperature conditions disrupt the stability of the CO_2-water interface and modify three-phase contact line forces, leading to increased contact angle hysteresis (Sedaghatinasab et al., 2021). When CO_2 is absorbed by the surface of carbonate rocks, it makes them more CO_2-wet. This changes the way CO_2 is held in the pores. This affects how much CO_2 stays in the rock and how easily it can escape (Arif et al., 2017). Changes in pressure and temperature also affect fluid behavior. They cause the interface between CO_2 and water to become unstable. This upsets the balance of forces that hold the fluid in place. As a result, the difference between the contact angle as the fluid moves back and forth becomes larger. This is called contact angle hysteresis (Sedaghatinasab et al., 2021).

2.5 Experimental methods for measuring contact angle

2.5.1 Sessile drop method

Researchers use the sessile drop method to apply brine drops onto rock surfaces while subjecting them to CO_2-exposed high-pressure extended exposure under elevated temperatures. A methodology exists for testing combinations of high-pressure and high-temperature reservoir conditions where the system measures the contact angle precisely. The first step of measurement requires rock sample preparation through cleaning, followed by optional polishing before mounting into the holder. Surface preparation stands as a critical step because improper methods will lead to inaccurate contact angle measurements (Arif et al., 2016). The rock sample enters the HPHT chamber at pressure and temperature conditions where the chamber receives CO_2 to evacuate air. At this point, the chamber reaches equilibrium while under CO_2 pressure at 9 conditions and an optimal temperature. The chamber needs time to reach experimental stability before measurements through an equilibration period. The aqueous solution gets injected onto rock surfaces through needle or capillary tubing but occupies only a tiny volume. Special attention is paid to the injection method to create sessile drops with suitable dimensions. The system requires time for equilibrium establishing during a period between minutes to hours based on research requirements and experimental conditions. The droplet needs time to establish equilibrium while clearing out any remaining dynamic processes. The sessile drop generates high resolution images through an imaging system implementation. A 3D reconstruction requires multiple angle-based image captures using the specified imaging system (Kaveh et al., 2014). The specialized software analyzes captured images to determine the contact angle. The analysis starts by matching the drop shape to mathematical models to measure the angle at the three-phase contact line. Figure 3 shows the steps for measuring advancing and receding contact angles using the sessile drop technique.

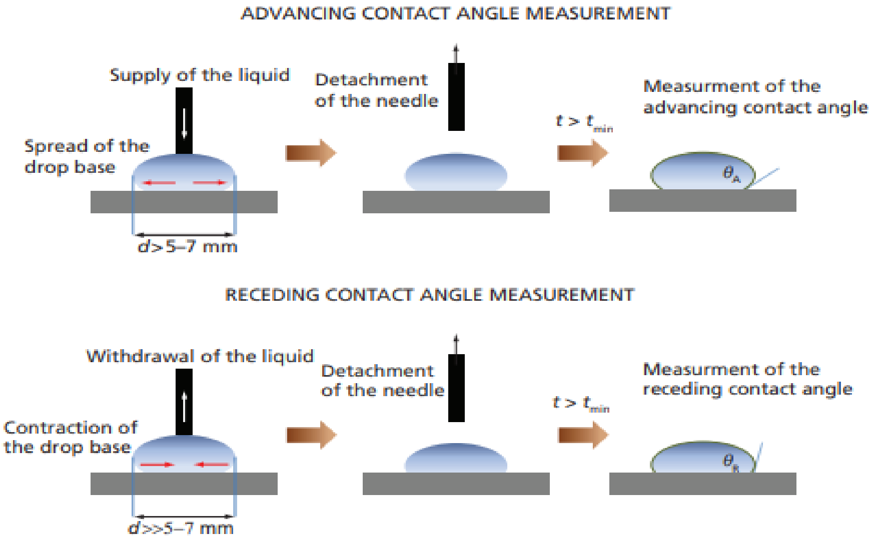


Figure 3. Schematic view of sessile drop technique (Marmur et al., 2017).

2.5.2 *Needle method*

The needle method establishes CO_2 bubbles through rock sample immersion in brine under elevated pressure and temperature conditions according to Arif et al. (2019). The procedure starts with proper rock sample preparation that includes cleaning, followed by polishing and mounting within the sample holder. After brine filling the high-pressure cell, researchers submerge the rock sample. The pressure cell receives CO_2 pressure after cell sealing to reach the target condition before allowing the system to stabilize until brine reaches complete CO_2 saturation (Arif et al., 2017). The needle injection system produces a CO_2 bubble below the rock surface after the system reaches its target pressure and temperature (Sarmadivaleh et al., 10 2015). Another equilibration period occurs to establish a stable status for the CO_2 bubble. Imaging system records high-definition pictures of the CO_2 bubble during its interaction with the rock surface. Specialized software analyzes captured images to measure angles between the CO_2 bubble and rock surface contact and brine (Arif et al., 2019). Figure 4 illustrates a schematic view of the needle technique, where the needle dispenses the solution by touching the droplet without contacting the surface.

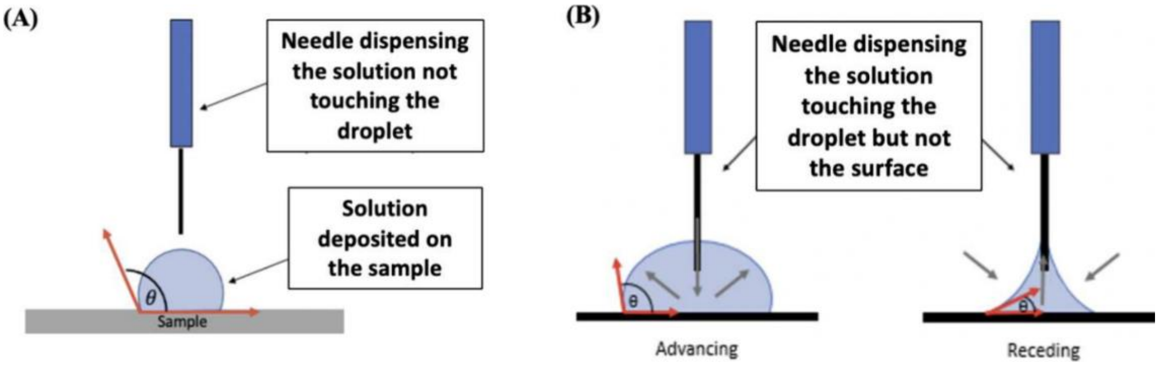


Figure 4. Schematic view of needle technique (Zou et al., 2024).

2.5.3 Tilting method

The dynamic wetting behavior of a flat rock sample can be evaluated through a tilting procedure that involves stepwise motion of the rock until the drop starts to move. The measurement process commences through proper rock sample care involving cleaning, followed by possible polishing to place the sample on the tilting stage inside the holder. The procedure for surface preparation proves essential because inadequate preparation affects contact angle evaluation results. Although the HPHT chamber receives the rock sample, it undergoes sealing afterward. A flush of CO_2 removes all air inside the chamber while additional CO_2 creates the required pressure. Temperature adjustment for the target value occurs before system equilibrium is reached (Arif, Barifcani, & Iglauer, 2016). Brine drops form sessile droplets on rock surfaces through the use of a micro-injection delivered by needles or capillary tubes. The system requires an additional equilibrium period before the drop reaches the equilibrium state (Kaveh, Rudolph, van Hemert, Rossen, & Wolf, 2014). After the drop achieves stability, the process of rock sample tilting can begin. The experiment controls the gradual increase of rock sample tilt angle at rates under 1 degree per second. The drop remains on the rock as the imaging system records numerous pictures or videos during this time. The drop will change its shape because of gravity during increases in the tilt angle measurement. During contact angle measurement, the advancing value comes from observing the slope-facing drop boundary, and the receding value emerges from the opposing upslope edge. The experimentation process continues its motion controls until movement occurs between the drop and the surface. The critical sliding angle represents the specific angle at which movement of the drop starts to occur. The detection of droplet-surface stickiness depends on this measurement. Image capture of the drop shape and position occurs during the entire tilting operation. Specialized software assesses captured images to calculate advancing and receding contact angles with various tilt angles and the critical sliding angle. Figure 5 shows a schematic view of the tilting technique for measuring dynamic contact angles.

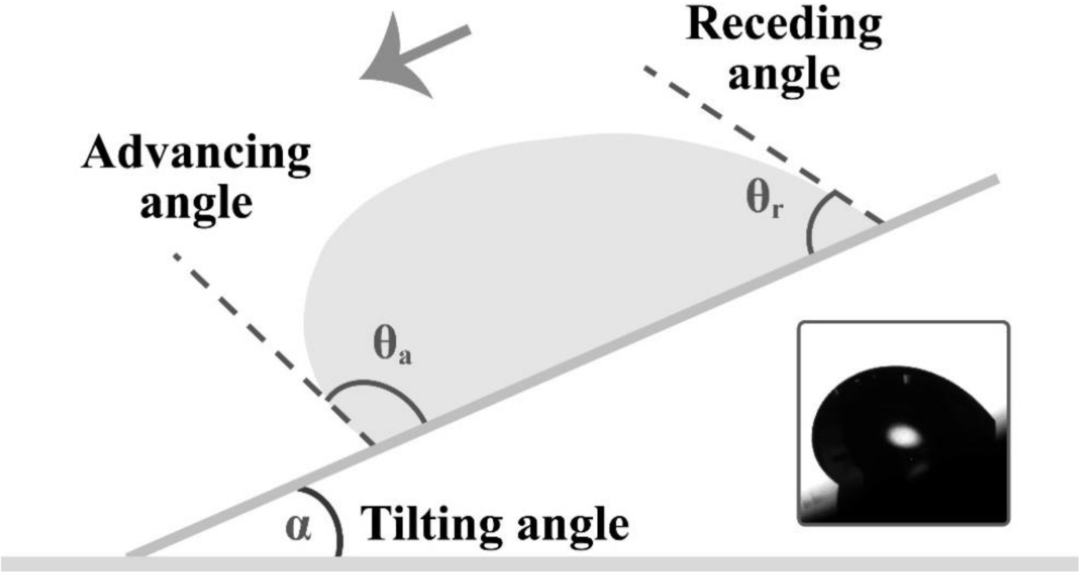


Figure 5. Schematic view of tilting technique (Kus et al., 2019).

2.6 Factors influencing on contact angle

2.6.1 Pressure effect

The wettability behavior of CO_2-water-rock systems depends heavily on pressure as an essential factor. Research findings prove that elevated pressure enhances CO_2-wettability, which shows in larger water-rock contact angles (θ) according to Chiquet et al. (2007). The high compressibility of CO_2 induces a significant density rise under pressure, which becomes the main mechanism explaining wettability change. Intense intermolecular interactions between denser CO_2 and minerals create stronger CO_2-solid affinities, thus leading to rock surface dewetting (Iglauer et al. 2012). Pressure increases cause the density difference between CO_2 and water (Δρ) to decrease simultaneously with CO_2-water interfacial tension (γwCO_2) reduction. The wettability changes stem directly from modifications in fluid properties (AlYaseri et al., 2016). Lab research has established precise data that demonstrates this pressure-related phenomenon. The contact angle measurement of a Weyburn limestone specimen expanded from 122° to 128° while pressure increased from 15 MPa to 20 MPa during tests at 300 K. Research by Arif et al. (2015) demonstrated that when the pressure of mica in contact with CO_2 and brine went from 0.1 to 20 MPa at 323 K the contact angle rose from 0° to 83°. The dissolution of additional CO_2 into the brine droplet under elevated pressure changes the brine-rock interface properties, which leads to increased hydrophobicity of the rock surface. A result of this change, the droplet contact angle increases. The contact angle rises to a point before it reaches equilibrium. Higher brine pressures lead to the solubility of CO_2 to achieve maximum/saturation levels, likely because of decreased brine capacity to dissolve more CO_2. When brine achieves total CO_2 dissolution saturation, the rising pressure fails to modify the interfacial characteristics. At elevated pressures, the rock surface evolves from being water-wet to displaying both intermediate CO_2-wetness and strong CO_2-wetness characteristics. The adsorption of CO_2 onto the rock surface at higher pressures occurs because of CO_2-rock cohesive energy reduction and enhanced CO_2-rock adhesive forces (Iglauer et al., 2015). The decline in rock and CO_2 interfacial tension aids water de-wetting processes while raising the rock surface hydrophobicity. The recorded increase in contact angle under elevated pressures shows that both structural and capillary trapping systems for CO_2 become substantially weaker (Al-Anssari et al., 2016; Iglauer, 2017). The capacity for CO_2 leakage rises significantly in the context of shale formations with low permeability, according to Arif et al. (2016). Figure 6 shows the advancing contact angles as a function of pressure and temperature in the brine–CO_2–quartz system.

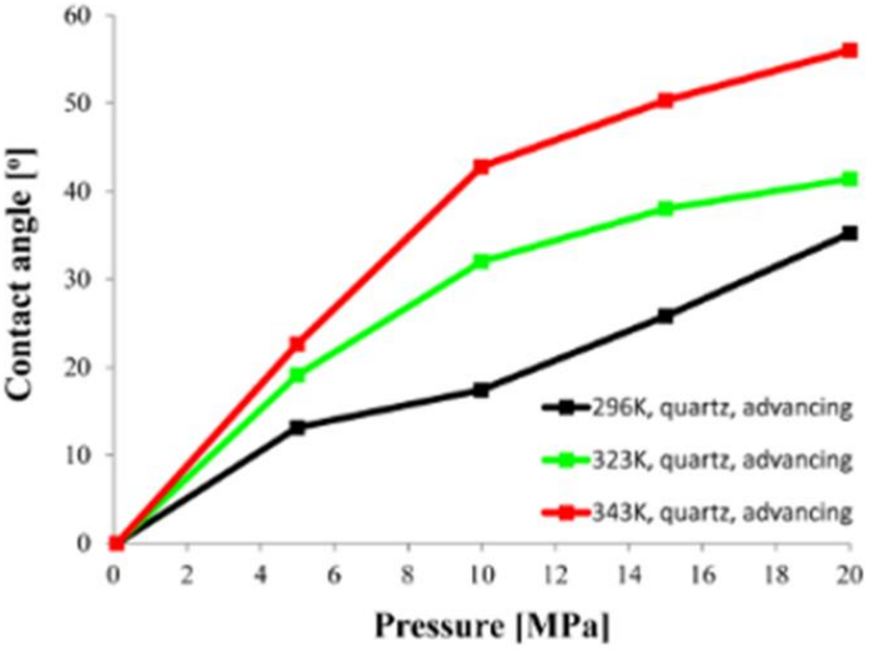


Figure 6. Advancing contact angles as a function of P and T in the brine-CO_2-quartz system (Iglauer, 2017).

2.6.2 Temperature effect

Knowledge about temperature effects on CO_2-water-rock wettability remains uncertain and complex across various studies. Various investigations study how temperature affects the wettability properties of carbonate and sandstone reservoir rocks, as well as representative minerals, quartz and calcite. The inconsistent findings produced by these studies generated continuous scientific disagreement about the issue. Multiple studies indicate that reservoir temperature rise produces decreasing contact angles, which results in better CO_2 storage capabilities alongside enhanced sealing properties (Arif et al., 2017; Broseta et al., 2012; Iglauer et al., 2012). According to Arif et al. (2017), the contact angle diminished from 91° to 66° as the temperature spanned from 308 K to 343 K under a consistent pressure of 15 MPa. Bikkina (2011) reported that quartz substrates showed their contact angle dropped from 45° to 20° during temperature elevation from 298 K to 323 K. Temperature causes wettability changes through multiple factors, which act as mechanisms for these changes. The alteration in the cohesive energies of rock-CO_2 and interfacial tension between solidCO_2 explains this phenomenon (Arif et al., 2017). Analysis through the sharp-kink approximation indicates that the interfacial tension between liquid and gas (γlg), the density difference ($\Delta\rho$), together with the van der Waals potential (I), represent the most vital factors in determining contact angles (Iglauer et al., 2015). Multiple studies conducted contradictory tests because they recorded growing contact angles at higher temperatures and also found that temperature independently created no meaningful change on contact angle readings (Roshan et al., 2016; Al-Yaseri et al., 2016). According to Al-Khdheeawi et al (2017), increased temperatures speed up the upward movement of CO_2 molecules, while lower temperatures improve the trapping efficiency. The rise in contact angles with increasing temperature occurs because the net $\Delta\rho/\gamma lg$ diminishes due to declining van der Waals forces and simultaneously increased γlg and $\Delta\rho$ values (Roshan et al., 2016). The researchers from Roshan et al. (2016) established a wider mathematical model to resolve inconsistencies in reported contact angle measurements through the integration of relevant wettability parameters. The latest mathematical model incorporates dielectric constant and electric charge measurements while tracking ionic concentrations to develop a complex explanation of temperature effects on surface wettability. The contradictory contact angle measurements stem from experimental obstacles, which include surface roughness, together with sample contamination and substrate chemical heterogeneity, and variant cleaning protocols (Arif et al., 2017). Reservoir conditions involving CO_2 and brine-enriched systems react with rock surfaces, and CO_2 droplet dissolution in brine or water and oil affects the experimental outcome (Iglauer et al., 2015).

2.6.3 Salinity effect

Formation brine salinity functions as a major factor that determines how CO_2-water-rock systems interact with each other. Multiple research studies have proven that formations with higher salinity levels tend to become CO_2-wet, but the precise extent, along with key contributing elements, can differ. Salinity modifies rock surface electrical double layers, which constitutes the main process through which wettability changes occur. The rise in brine concentration initiates an ion movement toward the charged rock surface, which forms an electrical double layer. The decay of electrical potential above a solid surface occurs swiftly after the double layer formation, which leads to an overall reduction in surface polarity (Chiquet et al., 2009). An increase in contact angle between rock and water phases results from the decreased hydrophilicity that surfaces experience due to their reduced polarity levels. Evidence from experimental tests shows that greater salinities produce enlarged contact angles, thus demonstrating CO_2-wet behavior. The research by Arif et al. (2016) demonstrated that raising NaCl concentration from 0 wt.% deionized water to 30 wt.% at 323 K and 15 MPa produced a mica contact angle expansion from 59° to 83°. The research findings about salinity effects on contact angles extend to calcite and quartz as well as other minerals, according to Espinoza and Santamarina (2010). Saline solutions affect wettability through both salt concentration and also through the ion qualities in the brine solution. Contact angles show greater increases from divalent cations Mg2+ and Ca2+ compared to monovalent cations Na+ and K+, according to Roshan et al. (2016). Higher ionic strength found in divalent cations creates better protection from electrostatic repulsion between the brine and the rock surface. Numerous research studies demonstrate a proven correlation between CO_2- CO_2-wettability enhancement and water salinity, yet investigators present conflicting results in the literature domain. Several investigations documented slight contact angle reductions at extreme salt concentrations while others detected no substantial variations according to Broseta et al (2012), Wang et al (2013) and Farokhpoor et al (2013). Experimental procedures combined with surface roughness and temperature variations and mineral composition, as well as electrostatic double layer screening mechanisms, contribute to the observed discrepancies. Wettability alteration happens in salt solutions through diverse processes, which include modified ionic distributions and surface ion adsorption and zeta potential changes alongside double layer screening by electrolytes. The increased salinity availability neutralizes the rock surface charges through available counter-ions, which creates rising zeta potential together with deteriorating surface polarity (Arif et al., 2017). Figure 7 demonstrates advancing contact angles as a function of salinity and salt type in the brine–CO_2–quartz system.

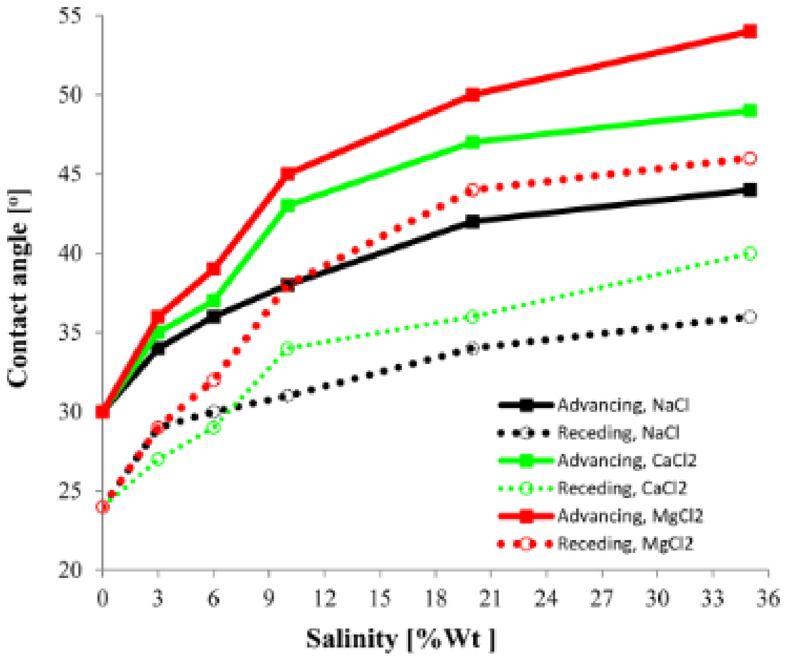


Figure 7. Advancing contact angles as a function of salinity and salt type in the brine-CO_2-quartz system (Iglauer, 2017).

2.6.4 *Rock surface effect*

Rock type significantly influences CO_2-water-rock wettability. Rock surface chemistry establishes the main variable for measuring CO_2 wettability (Iglauer, 2017). The hydrophilic character of quartz as well as calcite and feldspar, and mica make them clean mineral surfaces that become water-wet showing contact angles within 0° to 60° range (Iglauer, 2017). Clean rock surfaces are extremely difficult to find in underground environments. The wettability behavior of alkylated surfaces alongside those which experienced aging from crude oil or coal transforms them into either CO_2-wet or intermediate-wet surfaces with measured contact angles spanning between 70° and 170° according to Iglauer (2017). The wettability change happens because hydrophobic surfaces present lower rock CO_2 interfacial energy and higher rock-water interfacial energy according to Iglauer (2017). Rock surface wettability experiences major changes when organic material appears in rock samples. Small quantities (510-4400 mg/kg rock) of organic matter transition a rock between water-wet and intermediate-wet when placed in the reservoir environment according to Iglauer (2017). According to these observations subsurface reservoirs demonstrate intermediate levels of water saturation along with CO_2-wetness because they deviate from typical water-wet composition assumptions. Reservoir fluid wettability depends on the roughness of rock surfaces. Research findings demonstrate that surface textures with higher roughness levels show reduced contact angles which indicates that water will stick better to such surfaces (Al-Yaseri et al., 2016; Arif et al., 2017). Higher-quality research is required to explain the intricate relationship between surface roughness and surface chemistry changes and rock dissolution dynamics in CO_2-water-rock systems. Figure 8 illustrates contact angles in a brine–CO_2–rock system on hydrophilic (green) and hydrophobic (red) surfaces.

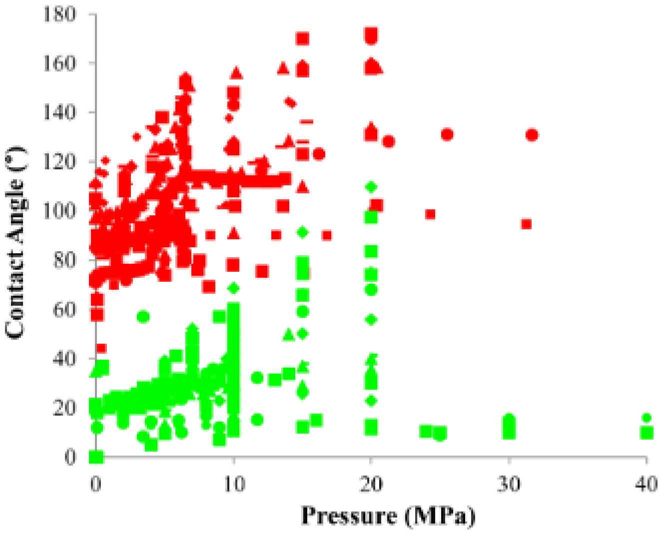


Figure 8. Contact angles in a system of the brine-CO_2-rock type on the hydrophilic (green) and hydrophobic (red) surface (Iglauer, 2017).

2.7 Machine learning tools

2.7.1 Decision Tree Regressor

The robust machine learning model Decision Tree Regression (DTR) operates for classification and regression tasks. DTR expresses both decisions and their results through a hierarchical structure where nodes stand for specific events and edges show the decision-making rules. The tree utilizes nodes and branches to divide the classification process, where each node represents a set of attributes, but each branch indicates particular node values. Tree construction through training data marks the starting point of the process. The basic dataset is split into two parts through binary split execution. The splitting process progresses through new tree branches until no additional splitting is possible, making the adjacent node become a terminal ending (Loh, 2011). DTR employs Mean Squared Error to partition nodes. A binary tree structure allows the algorithm to make a selection from available values for dataset partitioning. When calculating Mean Squared Error for any subsection, the tree chooses values that produce the lowest possible MSE results. Equation 2 can be used to calculate Mean Squared Error.

$$\mathrm{E(n)} = \frac{1}{\mathrm{N}}\sum_{i=1}^{N}[y_i(n) - \bar{y}(n)]^2 \qquad (2)$$

The split selecting process determines the greatest separation between node E(n) MSE and the MSE values of its child nodes E(nl) and E(nr). To make predictions, the model uses the average data values from the current node (James et al., 2013). Decision trees need several hyperparameters that regulate their complexity level and performance development. Setting the maximum depth parameter restricts trees from overgrowing since it controls the number of allowed levels, as explained in Rokach and Maimon (2008). A split in an internal node requires a minimum number of samples according to 17 the minimum samples split parameter, which stops the creation of undersized partitions. Every leaf in the decision tree contains a minimum of the specified number of samples through the minimum samples leaf parameter, which ensures reliable prediction accuracy (Hastie et al., 2009). These design features enable users to fine-tune how complex the model becomes. The maximum depth determines a tree's complexity by finding intricate patterns, but it tends to overfit as trees grow deeper, while shallower trees tend to underfit. According to (James et al., 2013), increasing minimum samples split along with minimum samples leaf produces generalized models, yet decreasing these values creates detailed overfitted models. Figure 9 shows the illustration of the tree structure where the root node sits at the top level within the model. The root node covers the entire sample and enables the creation of successive tree branches (Rokach & Maimon, 2008).

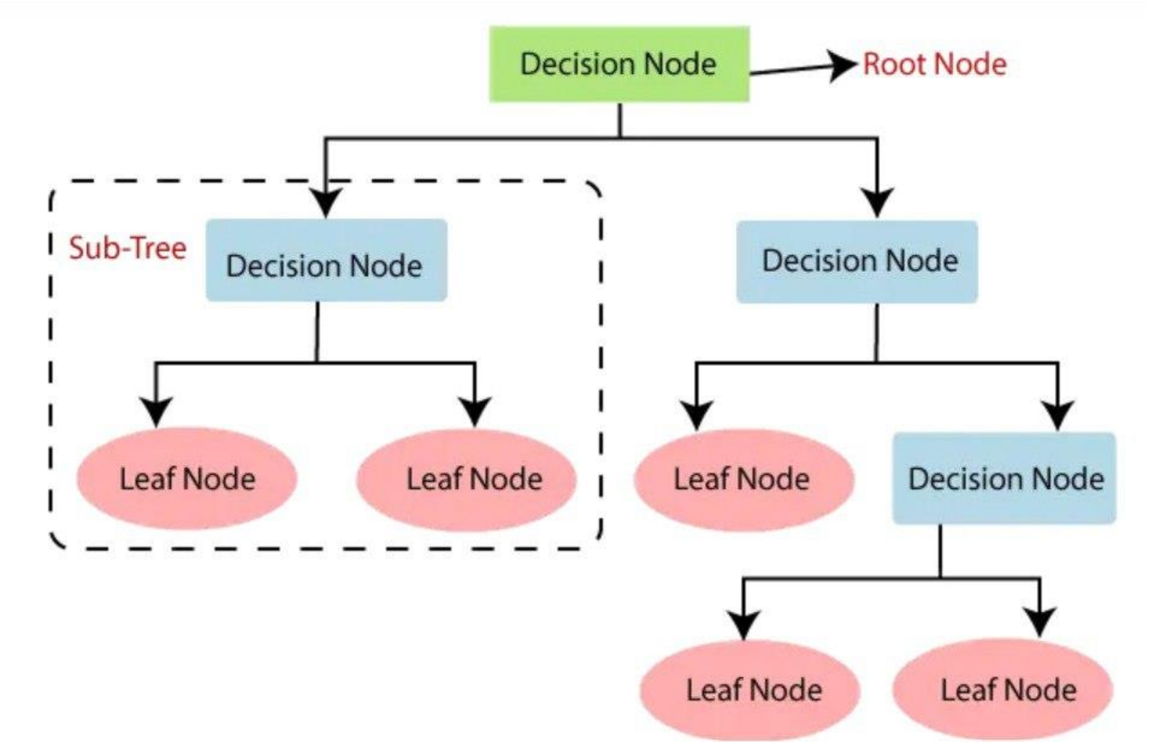


Figure 9. The Decision Tree Regressor's architecture (Sharma et al. 2022).

2.7.2 Random Forest Regressor

The Random Forest algorithm shows effective solutions for both classification and regression problems. In its operation, the method uses initial input data for training many models to gather predictions from each model, followed by a voting process to select the optimal solution. A forecast prediction emerges from multiple Decision Tree Regressors (DTR) through the averaging of their results under this method, which uses decision trees. Evaluating the prediction requires taking the mathematical average of all tree-generated outputs (Biau & Scornet, 2016). Random Forest uses an input partitioning technique to obtain numerous samples according to tree count and creates basic predictive models in each partition until the final forecast emerges from combining all model results by bagging methods (Hastie et al., 2009). Each mature decision tree in the Random Forest requires 18 no adjustment to processing speed because all trees reach completion. A higher number of decision trees enables the model to reduce data overfitting that leads to better results (Cutler et al., 2012). The Random Forest Regressor (RFR) model requires proper adjustment of multiple important hyperparameters to achieve optimal results. The number of trees in the forest depends on the n_estimators parameter while high execution time becomes inevitable to achieve better performance (Probst & Boulesteix, 2017). The selection of features through max_features depend on user-input options between 'auto', 'sqrt', and 'log2' parameters (Genuer et al., 2010). Max_depth controls the deepest part of trees to prevent overfitting. The min_samples_split parameter identifies the necessary minimum node splitting requirement whereas min_samples_leaf establishes minimum sample requirements for leaf node acceptance. Figure 10 shows the Random Forest architecture. The RFR model requires essential parameters encompassing N as the expected tree quantity as well as the number of predictor variables per splitting node alongside the node size that equals the minimum number of samples allowed at leaf nodes.

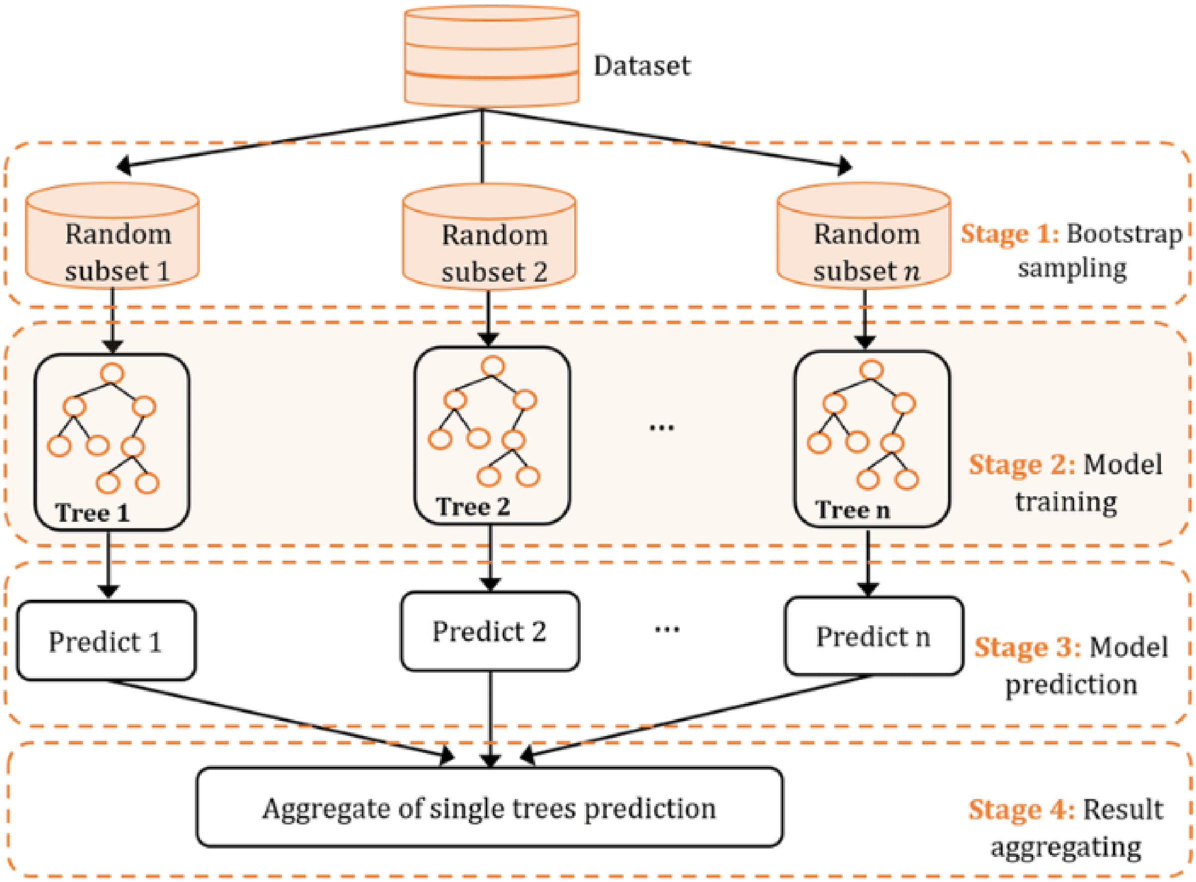


Figure 10. Random Forest architecture (Tao et al. 2022).

A regression issue dependent variable gets estimated through a simple random forest of numerous trees. The method produces K separate regression trees $h(x)$ which depend on the input variable x. During model predictions the algorithm uses the average forecast output from each tree to predict the inputs (x), (k = 1,…, K). Based on Equation 3, the RFR prediction can be determined.

$$\text{RFR prediction} = \frac{1}{\text{K}}\sum_{\text{K}=1}^{\text{K}} \text{h}_{\text{k}}(\text{x}) \qquad (3)$$

2.7.3 Extreme Gradient Boosting (XGBoost)

XGBoost operates as an ML regression and classification technique which generates prediction 19 models through an interpretative decision tree structure (Chen & Guestrin, 2016). XGBoost belongs to supervised machine learning through its roles within the boosting framework. The method applies gradient-boosted trees in ensemble fashion (Younis et al., 2024). The boosting method enables successive learners to use training specimens that received modifications according to the training effects of earlier learners. The combined voting process of all analysts produces the model's final forecast which successfully reduces the prediction errors against actual values. The Gradient Boosting Model tracks negative gradients identifying previous errors until gradient descent performs correction steps using the method. The XGBoost framework consists of streams that feed into multiple nodes including root points and endpoint leaves as shown in Figure 11 CARTs use the x_i parameter to produce initial decisions. The internal nodes execute new decisions between the different branch points that determine possible future actions. The XGBoost model prediction derives from combined CART model predictions found in leaf nodes.

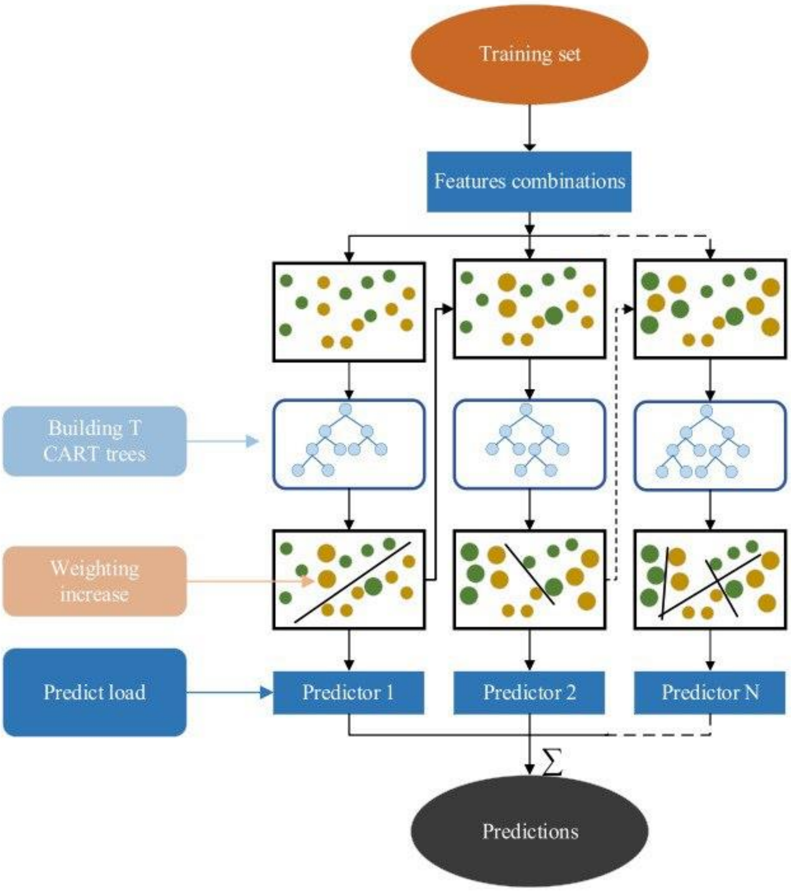


Figure 11. XGBoost architecture (Yao et al. 2022).

Multiple important hyperparameters must be adjusted to optimize XGBoost model performance. The n_estimators parameter controls the boosting rounds that compose an ensemble along with trees. The max_depth parameter functions by limiting the depth of trees in order to prevent overfitting. The learning_rate serves as the shrinkage parameter which controls the contribution of each tree while providing generalization improvements through lower values. The min_child_weight parameter controls over-fitting of high-dimensional data by specifying the required minimum sum of instance weight needed in a child (Chen & Guestrin, 2016). The actual experimental value is represented by the mathematical equation yi within XGBoost. The estimated score $\hat{yi}$ results from the arithmetic calculation of all fk parameters according to Equation 4.

$$\hat{y}_i = \alpha \sum_{k=1}^{k} \mathrm{f}_{\mathrm{k}}(\mathrm{x}_{\mathrm{i}}) \qquad (4)$$

2.7.4 Gradient Boosting Regressor (GBR)

The machine learning method known as Gradient Boosting operates through multiple weak learner combination to generate robust predictive models. Based on Equation 5, the general form of a gradient boosting model can be expressed as:

$$\mathrm{F}(\mathrm{x}) = \sum_{\mathrm{i}=1}^{\mathrm{N}} \mathrm{h}_{\mathrm{i}}(\mathrm{x}) \qquad (5)$$

In this model F(x) serves as the combined model that consists of N weak learners whose functions are $h_i(x)$ (Chen & Guestrin, 2016). Several important hyperparameters exist for Gradient Boosting models that allow users to optimize their performance during implementation. The ensemble learning process depends on the n_estimators parameter to specify the number of decision trees or weak learners which makes up the ensemble. The Maximum Features parameter determines how many features get evaluated for their suitability to become the best split in each tree. The maximum depth parameter defines the tallest permitted level of decision trees in order to limit overfitting. A split requires at least Minimum samples split to proceed while the training stops when reaching Minimum samples leaf. The hyperparameters operate together by managing model complexity level while improving its ability to predict new data which lets data scientists achieve peak model performance for their unique problem. Figure 12 shows a flow diagram of the GBR model, illustrating the splitting of the dataset into training and testing sets.

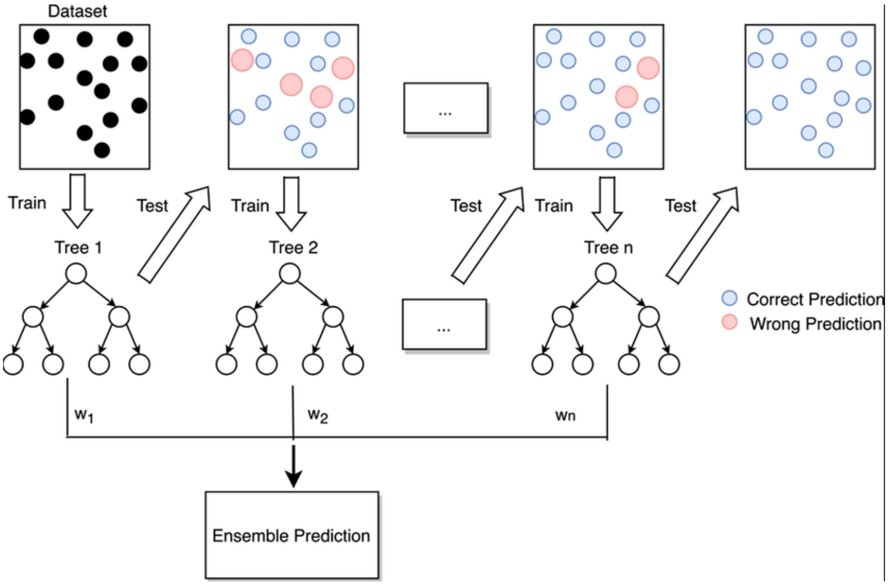


Figure 12. Flow diagram of GBR model (Zhang et al. 2021).

2.7.5 Support vector machine (SVM)

Based on Equation 6, the standard expression for an SVM classifier appears as following:

$$f(x) = \text{sign}(\textstyle\sum_i \alpha_i y_i K(x_i, x) + b) \tag{6}$$

The decision function f(x) contains the Lagrange multipliers α_i with class labels y_i and kernel function $K(x_i, x)$ and bias term b. Multiple influential hyperparameters determine the optimal functioning of SVMs in performance optimization. The model complexity stands as the main function of the regularization parameter C which allows users to strike a balance between minimizing training and testing errors. When applied to decision making boundaries the kernel function selects decision boundaries from linear, polynomial or radial basis function (RBF) methods. The width of the insensitive area for regression tasks is determined by the ε parameter during training which controls the number of support vectors effective in classification. The gamma (γ) parameter, relevant for non-linear kernels like RBF, influences the flexibility of the decision boundary. Predictable SVM results require hyperparameter optimization because these settings influentially affect both generalization of new data and computational speed. Figure 13 illustrates the overall workflow of the Support Vector Machine (SVM) model development process, encompassing cross-validation for hyperparameter optimization, training with feature extraction, and testing with performance evaluation using predicted and actual labels.

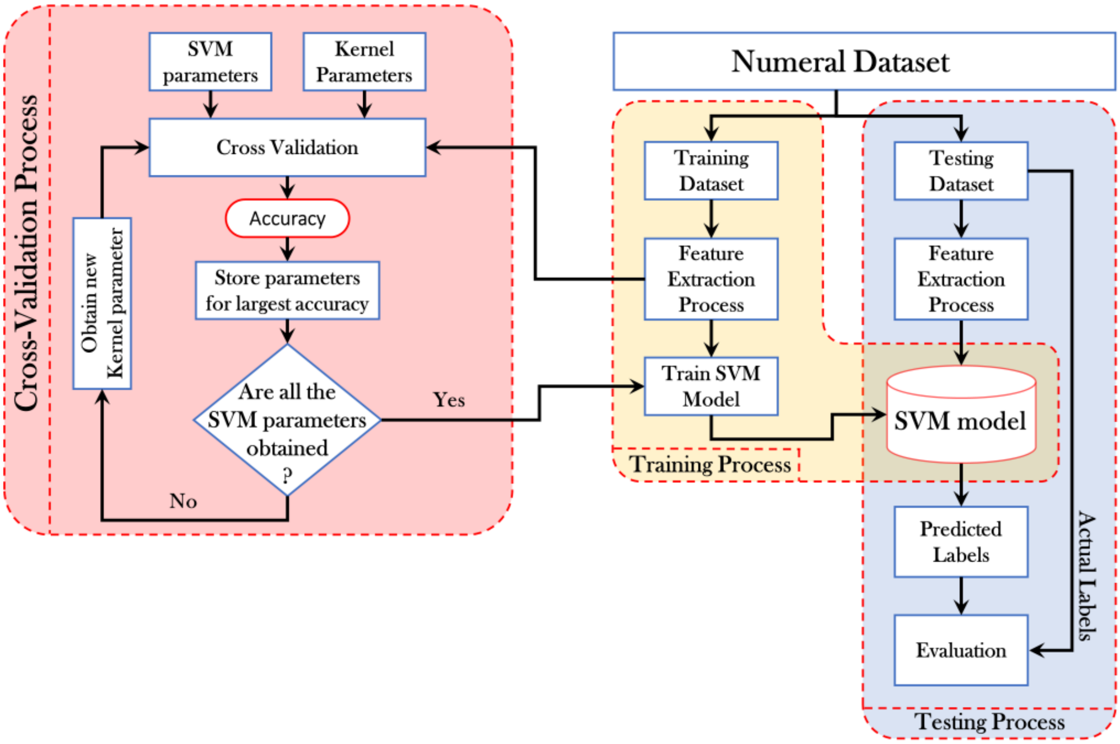


Figure 13. A schematic diagram for SVM model (Abdulhussain, 2021).

2.7.6 *Artificial Neural Network (ANN)*

ANNs function as flexible machine learning algorithms needing many hyperparameter adjustments to control their operational and structural features (LeCun et al., 2015). Based on Equation 7, the overall mathematical representation of ANN architecture appears as follows:

$$y = f(\sum_i w_i x_i + b) \tag{7}$$

The ANN calculation consists of y the output together with the activation function f supported by weights w_i and inputs x_i in addition to the b bias term (Goodfellow et al., 2016). The structure of the network largely depends on the number of hidden layers and the number of neurons located in each layer, since these parameters determine the depth and learning ability of the network. The network gains non-linear behavior through activation functions which include Rectified Linear Unit (ReLU) and sigmoid together with hyperbolic tangent (tanh). The functions mapped by ReLU as $f(x) = \max(0, x)$ while sigmoid defined $f(x) = 1 / (1 + e^{-x})$ and tanh uses f(x) = (e^x - e^(-x)) / (e^x + e^(-x)) according to Nwankpa et al. (2018). The training process requires selection of appropriate optimization methods among popular options like Adam, Nadam, and RMSprop. The Adam optimizer brings together features from momentum and RMSprop while Nadam extends Nesterov-accelerated gradient to Adam and RMSprop performs parameter-specific adaptive learning rate adaptation (Ruder, 2016). The combination of hyperparameters determines the ANN's performance aspects including learning dynamics and both speed and ability to generalize in various machine learning applications. Figure 14 depicts a neural network architecture consisting of three layers: the input layer, hidden layers, and output layer. Each input node connects to all nodes in the hidden layers, which in turn are connected to the output layer. The network processes input data through these layers to produce the corresponding outputs.

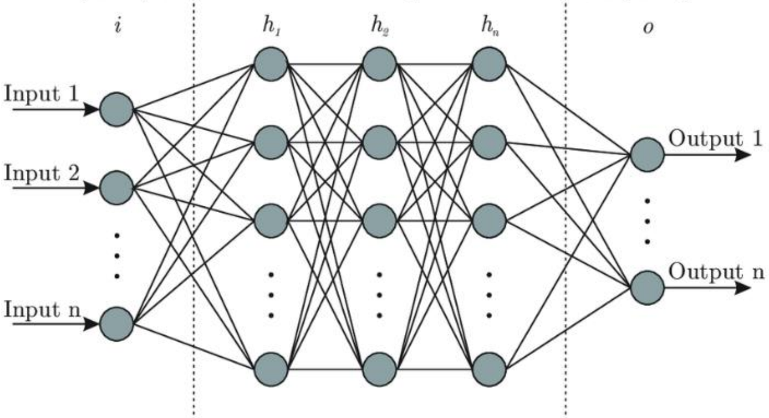


Figure 14. Artificial Neural Network architecture (LeCun et al., 2015).

2.8 Predicting wettability of mineral/CO_2/brine systems via data-driven machine learning modeling

The investigation performed an extensive exploratory analysis on 120 experimental measurements that included both categorical and numerical variables which described contact angles in mineral/CO_2/brine systems. Different visualization techniques along with statistical methods analyzed the relationships between variables while describing the database structure. The study presented data visualization with pair plots, violin plots and conducted correlation analyses and used heatmaps to display patterns and distributions throughout the analysis. Laboratory determinations showed that contact angles reached their minimum with lignoceric acid among all organic acids studied. The researchers determined the entire scope of key input parameters that ranged from pressue (0.1-25 MPa) to organic acid concentration (10^-12 to 10^-5 mol/m3) while contact angle measurements included advancing contact angle (33°-132.96°) together with receding contact angle (26.45°-124.88°). Predictive modeling involved six machine learning techniques including fully connected feedforward neural networks (FCFNNs) together with extreme gradient boosting and k-nearest neighbors and decision trees and adaptive boosting and random forest. Cross-validation and hyperparameter tuning optimized model performance on a training dataset consisting of 70% of the whole dataset while the remaining 30% served for model testing. The FCFNN model generated the best results in general with R^2 values exceeding 0.98 and errors reaching under 3% while evaluating advancing and receding contact angles. Besides FCFNN the XGBoost and random forest models demonstrated adequate performance metrics. The analysis of feature importance showed that both pressure conditions and organic acid measurement values contributed most to contact angle changes. Testing of the models included validation on mica as caprock and quartz as reservoir rock samples. When pressure levels rose together with organic acid strength and acid carbon chain length the contact angles tended to become more elevated. Table 1 summarizes the optimized hyperparameter configurations for various machine learning techniques, revealing distinct preferences in model architectures, tree depths, and learning rates. These differences reflect how each technique was fine-tuned to achieve the best performance for its specific application. Figure 15 compares the performance of machine learning models (DT, RF, KNN, AdaBoost, XGBoost) in predicting receding contact angles, with XGBoost achieving the highest R^2 scores (0.99 training, 0.976 testing). It also highlights feature importance, showing acid concentration as the most influential factor, followed by pressure and specific acid types.

Table 1. Optimized hyperparameters for the ML tools (Tariq, 2023).

Techniques	Hyperparameters	Range	Optimized Values for θ_a	Optimized Values for θ_r
Fully connected feedforward neural network (FCFNN)	Hidden layers	1-3	1	1
	Neurons per layer	10-30	30	30
	Learning rate	$1*10^{-2}$, $1*10^{-3}$, $1*10^{-4}$	$1*10^{-4}$	$1*10^{-3}$
	Activation function	ReLu, Sigmoid, Tanh	Tanh	Tanh
	Dropout rate	0.2, 0.5, 0.8	0.5	0.2
Decision trees (DT)	Maximum depth	5, 10, None	10	5
	Minimum samples split	2, 3, 5, 7, 10	5	2
	Minimum samples leaf	2, 3, 4, 5	2	2
Random forest (RF)	Number of trees	10-120	100	120
	Maximum depth	5, 10, None	10	None
	Samples split minimum	2, 5, 10	5	5
	Samples leaf minimum	1, 2, 5	2	2
	Number of Features	0.5, “sqrt”, “log2.”	0.5	0.5
K-nearest neighbors (KNN)	Number of Neighbors	5, 10, 20	10	20
	Distance metric	Euclidean, Manhattan, Chebyshev	Euclidean	Euclidean
XGBoost (extreme gradient boosting)	Number of trees	10-120	100	110
	Maximum depth	5, 10, None	10	None
	Learning rate	0.01, 0.1, 1	0.01	0.01
	Subsample	0.5, 0.8, 1	0.8	0.8
	Column sample by tree	0.5, 0.8, 1	0.5	0.5
AdaBoost (adaptive boosting)	Number of trees	10, 50, 100	100	100
	Learning rate	0.01, 0.1, 1	0.01	0.01

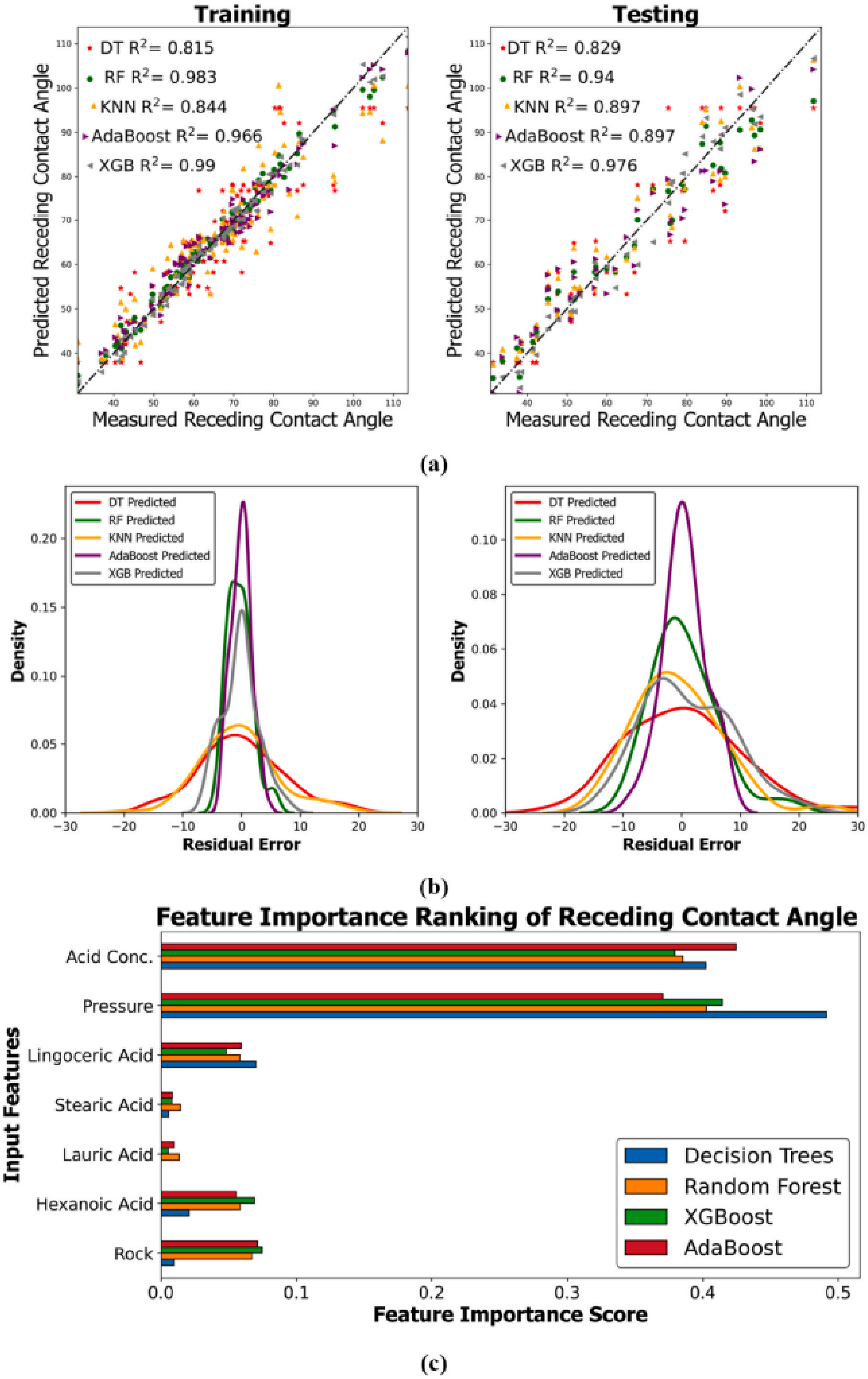


Figure 15. Prediction Performance of ML models for receding contact angle: (a) training and testing datasets, (b) residual error vs. density, (c) feature importance (Tariq, 2023).

3. METHODOLOGY

3.1 Data gathering

The development of data-driven models needs a thorough dataset which includes all necessary parameters. The main purpose of this research entailed measuring the contact angle characteristics of various rock types across brine and deionized water and CO_2 systems during advancing and receding and static conditions. The research used a systematic review process to find experimental reports about contact angles published in peer-reviewed academic sources. Academic papers from 2010 to 2024 contained static contact angle measurements for CO_2-brine-rock systems that amounted to 577 input points regarding pressure, temperature, and salinity data. A total number of 672 data points were collected from peer-reviewed papers published between 2011 and 2023 which studied advancing and receding contact angles in CO_2-brine-rock systems. Each reference provided detailed information regarding lead author name and publication year together with contact angle measurement methods and overall wettability patterns and studied rock types as well as experimental pressures and temperatures and brine salinity levels (including DI water) and the number of data points obtained from contact angle measurements. All papers reviewed displayed contact angle measurements and pressure results and temperature results and salinity results through tables and figures. A Plot Digitizer was employed to extract numerical values from figures. All measurements used standard units throughout the dataset which included MPa for pressure, °C for temperature, parts per billion (ppb) for brine salinity and degrees for contact angle measurement. Machine learning model development required individual Excel sheets for advancing, receding and static contact angles. Data entries consisted of pressure (MPa), temperature (°C), brine salinity (ppb), the experimental contact angle (°) and rock type. A new column entitled "CA type" was added to the combined advancing and receding contact angle table since their data points resided together in one single dataset. The dataset required this approach because it contained only 300 data points per contact angle type so individual models would not achieve reliable results. Each rock type has an Encoded Value for ML model training. The following tables present data for advancing and receding contact angles with descriptions of rock groups accompanied by their respective rock types and their experimental data points. Table 2, 3 and 4 provide a comparative overview of experimental studies measuring static and dynamic contact angles in CO_2/brine/rock systems, detailing variations in rock types, pressure ranges, temperatures, and brine compositions. The data highlights methodological differences and the number of measurements collected across studies, offering insights into the experimental conditions used in these investigations.

Table 2. Summary of experimental static contact angle data for the rock type/CO_2/brine system.

Reference	Method	Rock type	Pressure (MPa)	T (°C)	Brine Salinity	Data Points for θs
Shojai Kaveh et al. (2014)	Captive bubble	Sandstone	0.2-15	45	DI-water	81
Shojai Kaveh et al. (2016)	Pendant-drop	Shale	0.1-15	45	DI-water and 1M brine	134
Botto et al. (2017)	Sessile drop	Quartz	0.1-25	40	149 g/L brine	9
Tudek et al. (2017)	Sessile drop	Sandstone	10	22	10 wt.% brine	3
Mutailipu et al. (2019)	Sessile drop	Quartz, Sandstone, Limestone	3-15	25-100	1.98 mol/kg brine	131
Jung and Wan (2012)	Sessile drop	Silica	7-25	45	DI-water, 1-5M brine	55
Farokhpoor et al. (2013)	Captive bubble	Quartz, Calcite	0.1-40	36-66	DI-water, 0.2-0.8	44

					M brine	
Arif et al. (2017)	Sessile drop	Quartz	0.1-20	35-70	DI-water	15
Yang et al. (2007)	Sessile drop	Limestone	0.1-31	27-58	6.6 wt.% brine	18
Espinoza et al. (2010)	Sessile drop	Calcite, Quartz	0.1-10	22	DI-water, 200 g/kg brine	37
Wang et al. (2012)	Sessile drop	Calcite	7-20	30-50	DI-water, 1.1 g/L brine	4
Dai et al. (2024)	Kruss Drop Shape Analyzer	Shale	5-15	30-120	DI-water, 10 wt.% brine	16
Tianyi et al. (2024)	JY-PHa	Shale	0.1-16	50	DI-water, 50 g/L brine	30

Table 3. Summary of experimental advancing contact angle data for the rock type/CO_2/brine system.

Reference	Method	Rock type	P (MPa)	T (°C)	Brine Salinity	Data Points for θ
Iglauer et al. (2020)	Tilted plate	Basalt	4.48-17.23	35-60	100ppb	10
Alnili et al. (2018).	Tilted plate	Sandstone	5-20	23-50	DI-water, 20wt.%	16
Al-Yaseri et al. (2017)	Tilted plate	Quartz, Dolomite	0.1-20	35-70	DI-water, 20wt.%	24
Arif et al. (2017)	Tilted plate	Quartz, Mica, calcite	0.1-20	23-70	DI-water	45
Broseta et al. (2012)	Captive-bubble	Quartz, Calcite	0.5-15	35	0.08 M brine, 7M brine	6
Gholami et al. (2021)	Kruss Drop Shape Analyser	Shale	15	50	1000-10000ppm	110
Liu et al. (2015)	Sessile drop	Quartz	3	27-40	0.102 mol/l	2
Muhammad et al. (2023)	Tilted plate	Basalt	0.1-20	25-50	0.3 M brine	37

Saraji et al. (2013)	Axisymmetric Drop Shape Analysis	Quartz	3.5-11.5	35-60	DI-water	12
Umar et al. (2020)	Kruss Drop Shape Analyzer	Quartz, Feldspar, Kaolinite	0.1-17	20-65	0 wt.%-36 wt.%	54
Yekeen et al. (2021)	Pendant drop	Shale	0.1-25	40-180	0 wt.%-7 wt.%	68

Table 4. Summary of experimental receding contact angle data for the rock type/CO_2/brine system.

Reference	Method	Rock type	Pressure (MPa)	Temperature (°C)	Brine Salinity	Data Points for θ
Iglauer et al. (2020)	Tilted plate	Basalt	4.48-17.23	35-60	100ppb	10
Alnili et al. (2018).	Tilted plate	Sandstone	5-20	23-50	DI-water, 20wt.%	16
Al-Yaseri et al. (2017)	Tilted plate	Quartz, Dolomite	0.1-20	35-70	DI-water, 20wt.%	26
Arif et al. (2017)	Tilted plate	Quartz, Mica, calcite	0.1-20	23-70	DI-water	49
Broseta et al. (2012)	captive-bubble	Quartz, Calcite	0.5-15	35	0.08 M brine, 7M brine	8

Liu et al. (2015)	Sessile drop	Quartz	3	27-40	0.102 mol/l	2
Muhammad et al. (2023)	Tilted plate	Basalt	0.1-20	25-50	0.3 M brine	37
Saraji et al. (2013)	Axisymmetric Drop Shape Analysis	Quartz	3.5-11.5	35-60	DI-water	12
Umar et al. (2020)	Kruss Drop Shape Analyzer	Quartz, Feldspar, Kaolinite	0.1-17	20-65	0 wt.%-36 wt.%	54
Yekeen et al. (2021)	Pendant drop	Shale	0.1-25	40-180	0 wt.%-7 wt.%	68
Chiquet et al (2007)	Captive bubble	Quartz	0.1–10	35	0.01–1 M NaCl	6

Tables 5, 6, and 7 classify rock types based on experimental contact angle datasets (static, advancing, and receding), assigning encoded values for machine learning applications.

Table 5. Classification of Rock Types Based on Experimental Static Contact Angle Dataset

Rock Type	Encoded Value for ML model	Count of Static CA
Calcite	6	38
Limestone	4	44
Quartz	3	151
Sandstone	1	109
Shale	2	180
Silica	5	55
Grand Total		**577**

Table 6. Classification of Rock Types Based on Experimental Advancing Contact Angle Dataset

Rock Type	Encoded Value for ML model	Count of Advancing CA
Basalt	1	47
Calcite	2	17
Dolomite	3	15
Shale	4	178
Quartz	5	60
Sandstone	6	16
Silicate Minerals	7	33
Clay Minerals	8	18
Grand Total		**384**

Table 7. Classification of Rock Types Based on Experimental Receding Contact Angle Dataset

Row Labels	Encoded Value for ML model	Count of Contact Angle
Basalt	1	47
Calcite	2	21
Dolomite	3	15
Shale	4	68
Quartz	5	70
Sandstone	6	16
Silicate Minerals	7	33
Clay Minerals	8	18
Grand Total		**290**

3.2 Data preprocessing

Designing a proper preprocessing and implementing different visualization approaches constitutes fundamental requirements in the use of machine learning algorithms to explore CO_2-Water-Rock system contact angles. Data preparation steps through cleaning and missing value handling constitute a basic requirement for complex datasets to remove errors and inconsistencies that would cause prediction inaccuracies of contact angles. Statistical analysis reveals quantitative associations between pressure and temperature and salinity and rock type and contact angle which enables researchers to select important features for modeling CO_2-Water-Rock interactions. Box plots together with pair plots and heat maps help create essential visualizations to display how different variables distribute across different rock types as well as their connectedness and associated correlations between variables. Data analysis includes contact angle measurements together with other parameters that present their results through box plots that display median values and quartiles along with outlier points that indicate special contact angle behavior. The pair plots help researchers determine the non-linear pressure-related effects that different rock materials have on contact angles. Preparation methods along with comprehensive data exploration stand as essential elements to extract viable information that machine learning models use for predicting CO_2-Water-Rock. The implementation of Jupiter (Python) facilitated effective information processing which revealed key characteristics of the analyzed data.

3.2.1 Statistical analysis

Predictive modeling statistics with advancing/receding and static contact angles demonstrate how data cleaning affects primary variables such as pressure, temperature, salinity and contact angle. The cleaning procedure decreased the amount of measurement points and consistently lowered mean values and standard deviations especially in salinity measurements and contact angle results as well as removing outlier measurements. The mean values along with the temperature ranges for advancing/receding contact angles exceed those of static contact angles. The cleaned dataset exhibits better consistency and reliability because its maximum parameter values have been reduced for variables such as salinity and contact angle measurement. The preprocessing operation reduces measurement errors which leads to improved reliability of predictive models for contact angle prediction. Tables 8 and 9 present statistical analyses of the static contact angle dataset, comparing raw and cleaned versions by detailing key parameters like pressure, temperature, salinity, and contact angle measurements. Tables 10 and 11 provide statistical summaries of the advancing and receding contact angle datasets, comparing raw and cleaned versions. During the data cleaning process, contact angle values that exceeded 90 degrees were excluded from the static and dynamic contact angle dataset. Hence, the primary focus was on understanding the transition from water-wet to intermediate-wet conditions, which are more relevant for CO2 storage applications

Table 8. Statistical analysis for static contact angle raw dataset

Statistical parameters	Pressure (MPa)	Temperature (C)	Salinity (ppb)	Contact Angle
Count	577	577	577	577
Mean	8.7	50.8	47.2	33.3
Std	6.4	21	70.1	20.1
Min	0.1	22	2.8	2.8
25%	4.4	44.8	2.8	20.4
50%	7.9	45	2.8	30.3
75%	11.8	50	100	40
Max	40	120	292.2	131

Table 9. Statistical analysis for static contact angle cleaned dataset

Statistical parameters	Pressure (MPa)	Temperature (C)	Salinity (ppb)	Contact Angle
Count	567	567	567	567
Mean	8.5	51.1	47.9	31.9
Std	5.9	21.6	70.5	17.1
Min	0.1	22	0	0
25%	4.3	44.8	0	20
50%	7.6	45	0	30
75%	11.5	50	100	39.1
Max	40	120	292.2	90

Table 10. Statistical analysis for advancing and receding contact angle raw dataset

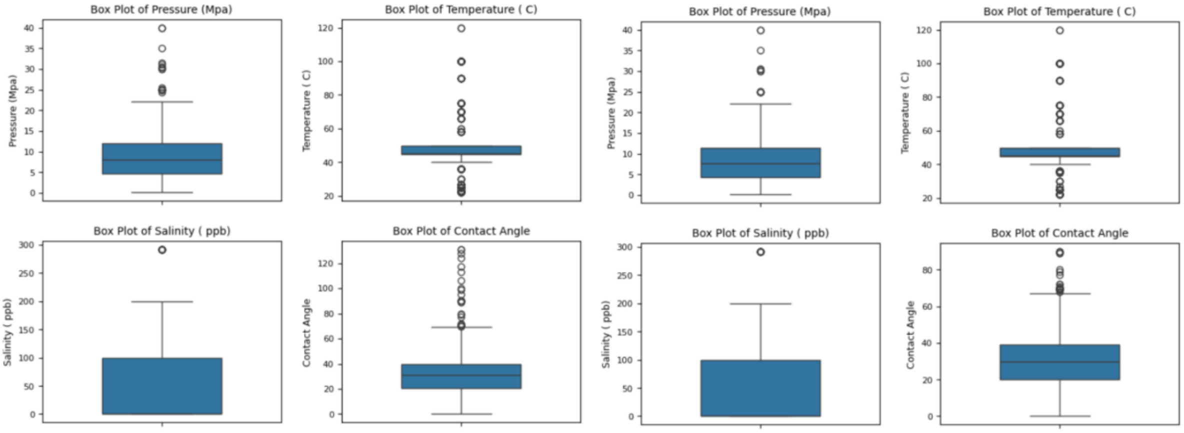
Statistical parameters	Pressure (MPa)	Temperature (C)	Salinity (ppb)	Contact Angle
Count	*672*	672	672	672
Mean	11.6	55.7	14.3	51.9
Std	5.8	28.6	41.3	28.5
Min	0.1	20.0	0.0	0.0
25%	8.0	35.0	0.0	30.0
50%	12.0	50.0	3.0	50.0
75%	15.0	70.0	17.5	72.0
Max	25.0	180.0	409.1	162.0

Table 11. Statistical analysis for advancing and receding contact angle cleaned dataset

Statistical parameters	Pressure (MPa)	Temperature (C)	Salinity (ppb)	Contact Angle
Count	*627.0*	*627.0*	*627.0*	*627.0*
Mean	11.3	54.2	14.2	47.6
Std	5.8	27.9	39.6	24.0
Min	0.1	20.0	0.0	0.0
25%	6.9	35.0	0.0	27.5
50%	11.5	50.0	3.0	47.0
75%	15.0	65.0	17.5	68.5
Max	25.0	180.0	409.1	90.0

3.2.2 Box plots

The preprocessing process depends heavily on box plots because they display detailed visual presentations of data patterns. The figures display major patterns throughout the variables Pressure, Temperature, Salinity, and Contact Angle for both contacting states, Advancing/Receding and Static Contact Angle measurements. The examined data shows stable pressure distribution patterns and broader temperature spans in the Advancing/Receding dataset, while salinity variables present substantial differences, and contact angle outliers are significantly reduced after cleaning procedures. The box plots allow simple data evaluation between original measurements and cleaned results to display distribution alterations emerging from cleaning procedures. The boxplots contribute to detecting anomalies and measuring how data cleaning works on distribution ranges while identifying systematic effects and confirming that preprocessing adjustments did not generate data structure modifications. A visual analysis toolbox proves essential for establishing reliable data quality, thus it helps develop better machine learning models which predict CO_2-Water-Rock system behaviors. Figure 16 compares the distribution of input parameters between raw and cleaned static contact angle datasets, showing how data cleaning reduced outliers and tightened value ranges. Figure 17 similarly contrasts raw versus cleaned data for advancing/receding contact angles, demonstrating the refinement process's impact on parameter distributions for more reliable analysis. Both figures visually highlight the improved data quality after preprocessing.



Raw Dataset Cleaned Dataset

Figure 16. Boxplots for input parameters of static contact angle raw and cleaned datasets

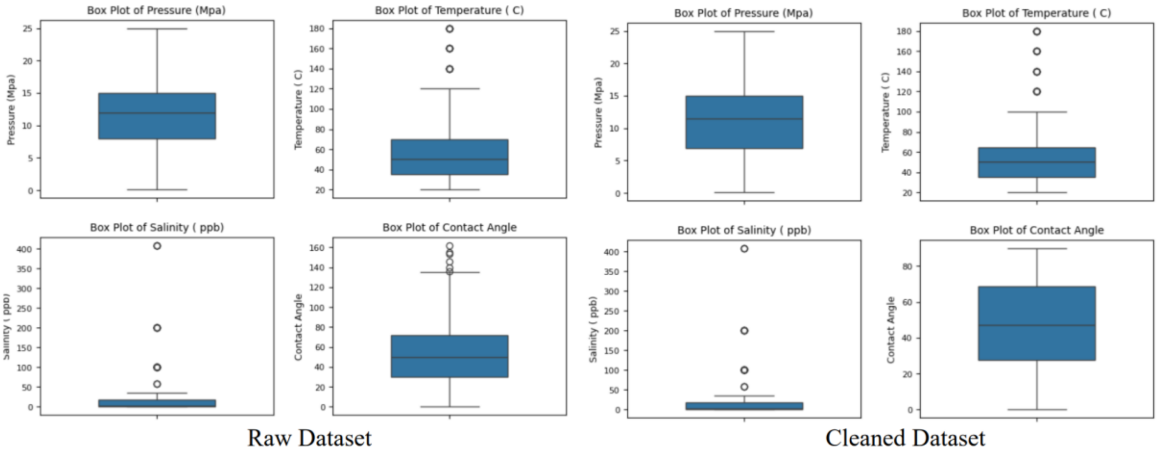


Figure 17. Boxplots for input parameters of advancing and receding contact angle raw and cleaned datasets

3.2.3 Pair plots

Pair plots define the complete visual presentation of pressure and temperature and salinity and contact angle interactions within CO_2-Brine-Rock systems for various rock types. Different rock types display separate patterns and distribution data points for wettability behavior in the provided plots. The static contact angle measurement indicates that rising pressure values create higher contact angles while temperature and salinity values show a weaker link. The rock types show varying patterns in data distribution since some samples appear in greater numbers than others. In the static contact angle dataset, Shale, Quartz, and Sandstone rocks occupy most of the data points, but Shale and Quartz rocks take control of the advancing contact angle dataset. Midway through the plot viewers can observe how contact angle functions in a non-linear manner with respect to pressure. The distribution of Salinity values distorts in a particular way with specific outcomes that occur exclusively for certain rock types. The visuals underline how rock type stands as a vital component for modeling CO_2-Brine Rock wettability since different rock types show distinct reactions to various conditions. Figures 18 and 19 present a pair plot visualization of the cleaned static and dynamic contact angle dataset, illustrating the relationships between input parameters while incorporating rock type as a distinguishing factor. This multivariate analysis helps identify potential correlations and patterns among variables, with rock type serving as a key categorical feature in the dataset's structure.

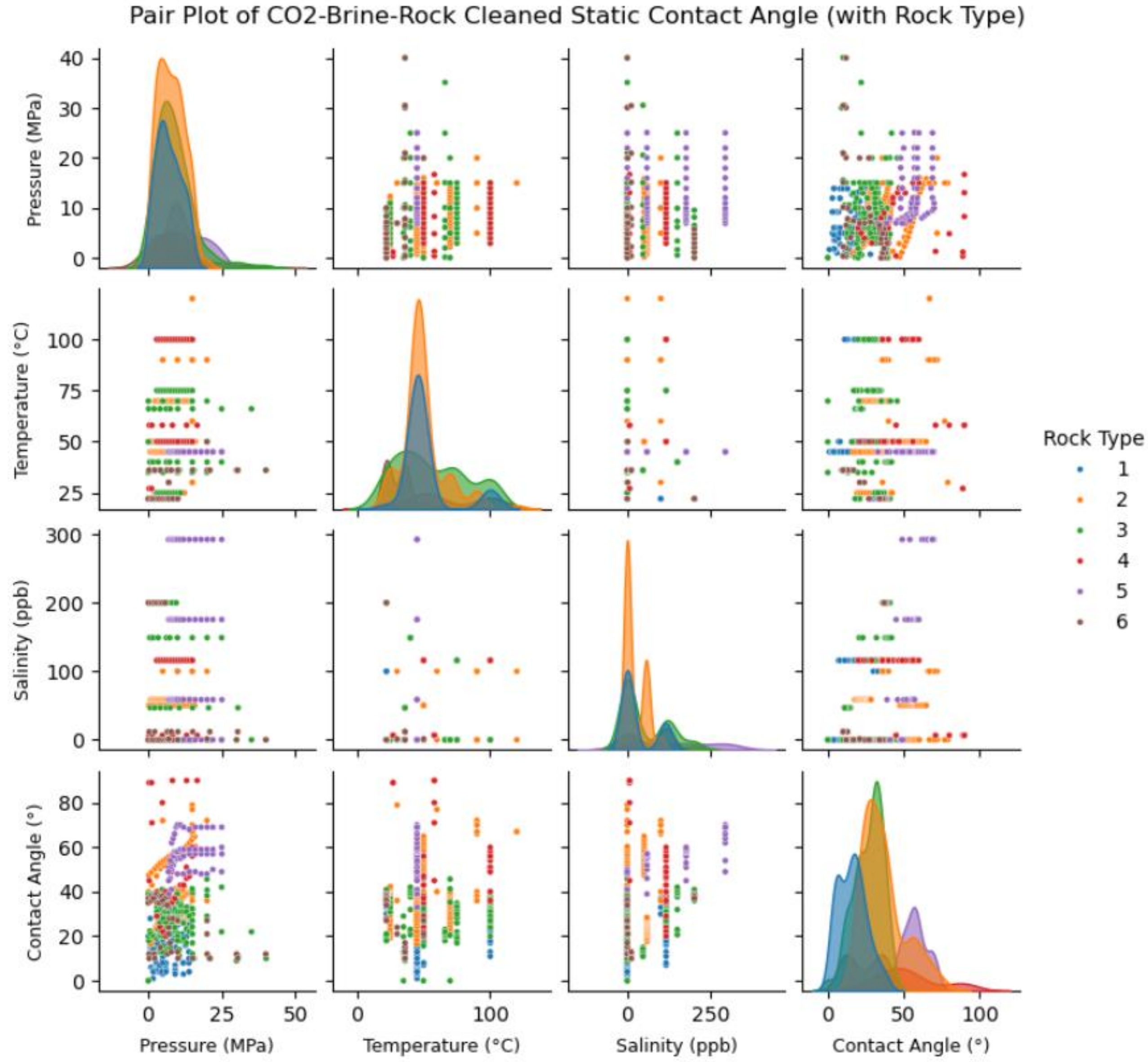


Figure 18. Pair plot for input parameters of static contact angle cleaned dataset (with Rock Type)

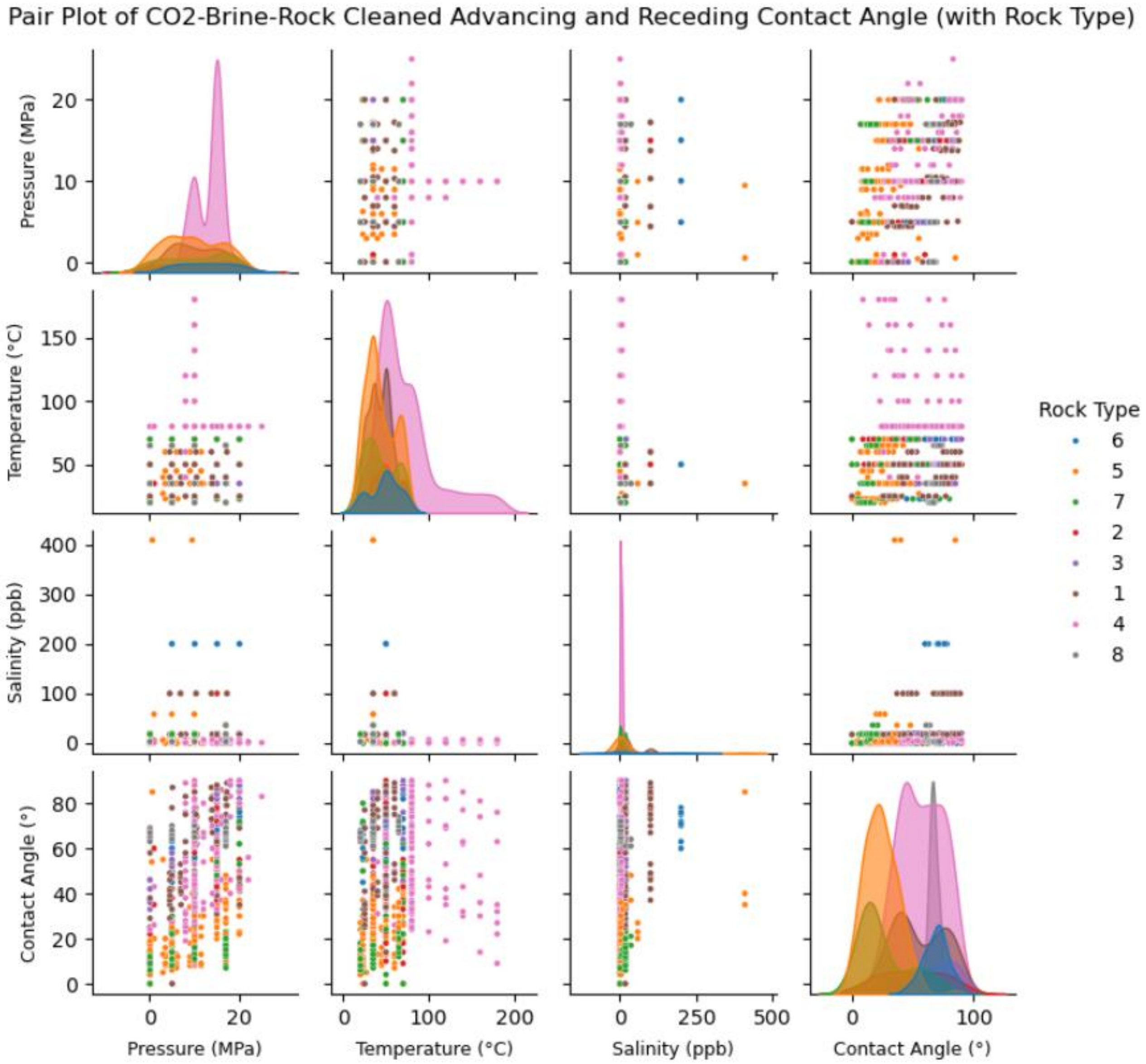


Figure 19. Pair plot for input parameters of advancing and receding contact angles, cleaned dataset (with Rock Type)

3.2.4 Heat maps

Figures 20 and 21 present heatmap visualizations comparing parameter correlations in raw versus cleaned datasets for static (Figure 20) and advancing/receding contact angles (Figure 21). These heatmaps highlight how data cleaning affected variable relationships, with the cleaned datasets likely showing more refined and reliable correlation patterns. The visualizations help identify key parameter interactions while demonstrating the impact of preprocessing on data quality. In the original and refined data sets, pressure established the top positive correlation with contact angle at 0.46, which became 0.44 in the refined data. The research indicates that temperature shows a small positive relationship with contact angle (0.21

raw, 0.13 cleaned). Salinity shows weak positive linkages (0.11 raw-0.15 cleaned) with contact angle measurements. The static contact angle measurements possess stronger correlations compared to other data points. The data demonstrates that an increasing pressure value leads to higher contact angle measurements (0.34 raw and 0.26 cleaned data). Static contact angle presents its greatest positive association with rock type (0.37 for both raw and cleaned samples). The data shows a significant positive connection between salinity measurements (raw 0.30, cleaned 0.40). The relationship between contact angle measurements in static conditions and temperature produces the lowest correlation value (0.05 raw and 0.13 cleaned).

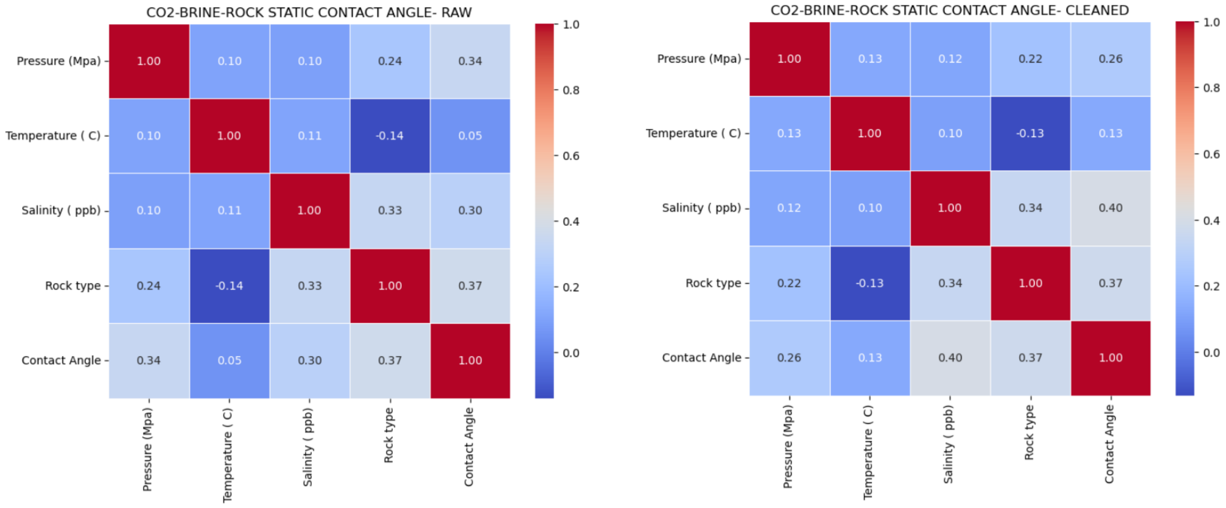


Figure 20. Heat maps for input parameters of static contact angle raw and cleaned datasets

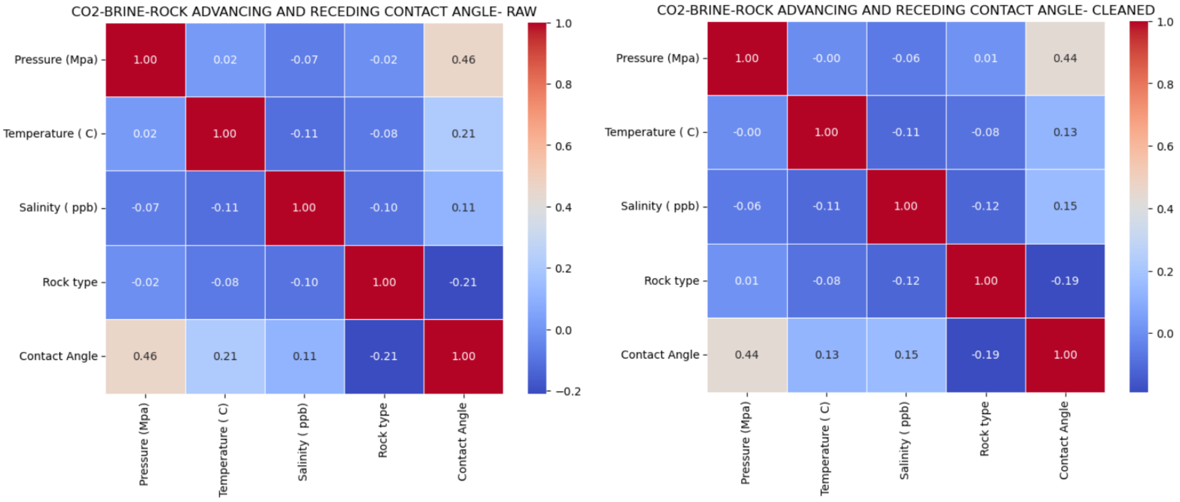


Figure 21. Heat maps for input parameters of advancing and receding contact angles, raw and cleaned datasets

3.3 Data processing

The analysis of CO_2-Water-Rock systems requires a complete data processing methodology, which includes determining models and performing training and testing, followed by hyperparameter adjustments and performance evaluation. Decision Trees, together with Random Forests and XGBoost alongside Gradient Boosting Regressor and Support Vector Machines and Artificial Neural Networks make up the set of machine learning methods considered for their ability to recognize complex patterns in the provided data. Petitions for training and testing are allocated at 70/30 ratios for building trustworthy models while maintaining objective performance measurements. The prediction accuracy demands the use of advanced model types for advancing and receding contact angles, while static angles require simpler models. Various statistical metrics including RMSE and R^2 and AIC and BIC join forces to perform complete model evaluations and to determine generalization potential. Artificial Neural Networks use different network designs to achieve maximum predictive accuracy. A systematic data processing system allows scientists to create accurate models which predict wettability behaviors along with contact angles in CO_2-Water-Rock systems.

3.3.1 Model Selection

Decision Trees provide straightforward understanding regarding modeling capabilities of wettability and contact angle behavior in CO_2-Water-Rock systems. The method succeeds at detecting non-linear patterns yet experiences problems with continuous transitions and tends to overfit data patterns. Decision Trees serve as valuable tools during the initial stages of data analysis because they show feature importance alongside providing insights about data relationships (Breiman et al., 1984). Random Forests have superior predictive capabilities over Decision Trees because they restore both robustness and minimize overfitting occurrences. XGBoost demonstrates ability to detect complex feature interactions while concurrently generating assessments of feature rankings which are integrated automatically. The interpretability of Decision Trees becomes limited when using multiple Decision Trees across a Random Forest (Breiman, 2001). In the field of machine learning XGBoost represents an optimized gradient boosting framework which delivers superior performance in multiple applications. The tool performs automatic handling of missing values while its built-in regularization feature helps avoid problems with overfitting. XGBoost offers potential high accuracy contact angle behavior predictions when users tune its parameters properly according to Chen and Guestrin (2016). The Gradient Boosting Regressor tracks sophisticated non-linear patterns because it possesses strong resistance to outliers in dependent variables. XGBoost constructs its trees in a sequential order to fix previous incorrect trees resulting in potentially accurate outcomes. Data irregularities can affect this model while improper parameter adjustment leads to overfitting of the algorithm (Friedman, 2001). Support Vector Machines (SVMs) function well with large number of dimensions while utilizing different kernel functions to achieve versatile performance. These models function most effectively when the number of dimensions exceeds sample count thus making them useful for particular models of the CO_2-Water-Rock system. Large datasets handling by SVMs demands considerable computational resources and users need to select proper kernel functions along with tuning parameters (Cortes & Vapnik, 1995). The capabilities of Artificial Neural Networks (ANNs) to detect complex patterns along with self-learning functions make these systems ideal for understanding wettability complexity. These models need extensive datasets for proper functioning while their performance may become unstable when regularization techniques are not applied. ANNs provide adaptable architecture options yet they do not provide the same interpretability as other models according to Goodfellow et al. (2016). The selection of preferred model depends on the particular characteristics of the dataset combined with balance needs between prediction precision and model interpretability and computational speed.

3.3.2 Model Training and Testing

A 70/30 separation of data into training and testing groups represents a vital method in machine learning modeling when dealing with complex CO_2-Water-Rock systems. The model training process benefits from this split, which provides complete data efficiency by utilizing 70% of the dataset to build accurate models and then occupies 30% of the data for fair, unbiased testing. The splitting method supports overfitting prevention because models receive unseen data during testing which controls their true generalization potentials. The data separation method secures both a big training sample base and a valid selection of data for testing purposes. The proposed approach provides a method to detect both underfitting and overfitting errors through dual dataset assessment while helping choose models and optimize hyperparameters in addition to delivering standardized evaluation results. This method proves essential for wettability and contact angle behavior modeling because it optimizes model strength across different measurement conditions for CO_2-Water-Rock systems.

3.3.3 Hyperparameter Tuning

An examination of hyperparameter optimization for different machine learning methods applied to CO_2-Water-Rock system wettability, as well as contact angle modeling is shown in this table. The analytical work examines six machine learning methods including Decision Trees (DT), Random Forest (RF), XGBoost, GradientBoostingRegressor (GBR), and Support Vector Machine (SVM) with Artificial Neural Network (ANN). The maximum depth parameter for Decision Trees reaches 15 levels for θa and r rather than 10 levels for θs because advancing and receding contact angles require more complex relationships. Random Forest achieves optimal results with 100 estimators for θa and r compared to θs, which needs 50 estimators, thus demonstrating that additional trees help explain advanced and receding contact angles better. XGBoost: Similar settings for both angle types, with slight variations in maximum depth and number of estimators. GradientBoostingRegressor selects 200 estimators for θa and r while using only 50 estimators for θs which demonstrates that the model applies increased complexity for these contact angles. The SVM model uses RBF kernel with equivalent parameters for advancing and receding angles, together with C, along with Epsilon and Gamma values set to similar levels. The ANN requires varying activation functions because θa and r need a sigmoid function, and θs needs tanh to capture unique underlying data patterns. The tests show that dynamic contact angles θa and θr need models with greater complexity through deeper trees and more estimators than static contact angles θs. The connection between input variables and contact angles demonstrates higher complexity during advancing and receding motions than static conditions within the CO_2-Water-Rock system. Table 12 summarizes the optimized hyperparameters for various machine learning techniques used in predicting static contact angles. The table reveals distinct configurations across models, with Decision Trees favoring deeper structures (max depth 15 for static contact angle), Random Forest using sqrt for feature selection, and XGBoost employing moderate learning rates (0.1). Notably, ANN architectures differed in hidden layers (4 vs. 3) and activation functions (sigmoid vs. tanh), while SVM consistently used RBF kernels with varying epsilon values, reflecting tailored optimization for each predictive task.

Table 12. Optimized hyperparameters of the implemented ML techniques.

Techniques	Hyperparameters	Range	Optimized Values for $\theta_{a\ and\ r}$	Optimized Values for θ_s
Decision trees (DT)	*Maximum depth*	5-15	15	10
	Minimum samples split	1-3	2	3
	Minimum samples leaf	1-3	3	1
Random forest (RF)	N estimators	50-150	100	50
	Maximum features	auto, sqrt, log2	Sqrt	Sqrt
	Maximum depth	10-14	10	10
	Samples split minimum	3-5	5	5
	Samples leaf minimum	1	1	1
XGBoost (extreme gradient boosting)	N estimators	100-110	105	100
	Maximum depth	5-7, none	5	6
	Learning rate	0.01-1	0.1	0.1
	Minimum child weight	1-2	1	1
GradientBoostingRegressor (GBR)	N estimators	50-200	200	50
	Maximum features	sqrt	sqrt	sqrt
	Maximum depth	2-10	4	6
	Minimum samples split	2-6	6	2
	Minimum samples leaf	1	1	1
SupportVectorMachine (SVM)	C	10-10000	1000	1000
	Kernel	rbf, sigmoid	rbf	rbf
	Epsilon	$10^{-4} - 10^{-6}$	10^{-6}	10^{-5}

	Gamma	0.0001-0.01, auto, scale	0.01	0.01
ANN	Hidden layer	1-5	4	3
	Neuron	256	256	256
	Activation function	Relu, sigmoid, tanh	sigmoid	tanh
	Optimizer	Adam, nadam, rmsprop	nadam	nadam

3.3.4 Model evaluation

Various statistical metrics served in the model evaluation process of machine learning models designed to assess wettability and contact angle behavior in CO_2-Water-Rock systems. The assessment used Root Mean Square Error (RMSE) together with coefficient of determination (R^2) as well as Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) metrics. Root Mean Square Error (RMSE) determines the typical difference of forecasted values from their observed counterparts. Based on Equation 8, The calculation for determining RMSE involves taking the square root of a mean operation on squared differences between actual and predicted values (Chai & Draxler, 2014).

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(X_{Ai}-X_{Pi})^2}{n}} \tag{8}$$

The prediction variables are represented by X_{Pi} and actual variables by X_{Ai} while n denotes the count of sample points. A model performs better when it displays lower RMSE values. The Mean Squared Error (MSE) calculates the blank-sized predicted-experimental value difference to generate a summary of model precision performance. Based on Equation 9, The mean of squared residuals calculates RMSE through dividing the sum of observed and predicted value difference squares by the total number of data points. Loss function optimization takes place through mathematical model assessment within machine learning applications.

$$\mathrm{MSE} = \frac{\sum_{i=1}^{n}(X_{Ai}-X_{Pi})^2}{n} \tag{9}$$

The calculation involves predicted values X_{Pi} combined with actual values X_{Ai} using the number of data points n. Model performance strength increases as RMSE values decrease. The coefficient of determination (R^2) shows how much independent variables can explain the dependent variable variability (Nagelkerke, 1991). Model performance outcome values extend from 0 to 1 while improved fit corresponds to higher values. R^2 can be calculated by Equation 10.

$$R^2 = 1 - \sum_{i=1}^{n}\left(\frac{(X_{Ai}-X_{Pi})^2}{(X_{Ai}-\overline{X_{Ai}})^2}\right) \qquad (10)$$

The computational formula for R2 calculation considers X_{Pi} predicted values together with X_{Ai} actual values and $\overline{X_{Ai}}$ as the average actual value and n as the number of points. The model selection criteria include Akaike Information Criterion (AIC) along with Bayesian Information Criterion (BIC) which assess both model complexity along with goodness of fit according to Burnham & Anderson (2004). Based on Equation 11 and 12, AIC and BIC are calculated as:

$$\mathrm{AIC} = -2\ln(\mathrm{likelihood}) + 2\mathrm{P_N} \qquad (11)$$

$$\mathrm{BIC} = -2\ln(\mathrm{likelihood}) + [\ln(\mathrm{n})]\mathrm{P_N} \qquad (12)$$

The calculation includes n data points together with P_N model parameters and likelihood as the maximum likelihood estimate. Several factors contribute to lower AIC and BIC scores that indicate better model selection because they show the right combination of model fit with complexity balance.

The evaluation metrics fulfill specific functions for model assessment purposes. The direct measure of prediction accuracy comes from RMSE whereas R^2 measures how well the model explains the data. According to Kuha (2004) these two methods serve to assess different parameterized models by applying penalties which help in avoiding overfitting issues. The metrics serve in the CO_2-Water-Rock system modeling by enabling thorough evaluation of machine learning approaches (SVM, DT, RF, XGBoost, GBR, ANN) and alternative ANN structure implementations. Using these metrics finds application for both training and testing data sets to determine the models' accuracy in fitting training examples as well as their ability to generalize newly observed data points.

4. RESULTS AND DISCUSSION

4.1 Implementation of ANN model for static contact angle dataset.

Different network architectures used for Artificial Neural Network modeling generate varying outcomes when analyzing static contact angles. Figure 22 and Figure 23 demonstrate that the model generates reliable predictions regarding contact angle values for testing and training scenarios when using models equipped with multiple hidden layers. Table 13 presents complete statistics about error metrics which apply to models with 1 to 5 hidden layers. The 4-hidden layer model shows superior performance by delivering the maximum training and testing R² values at 0.97 and 0.90 and producing the minimal RMSE and MSE numbers. This model demonstrates the highest AIC and BIC scores that indicate optimal combinations of complexity and performance achievement. The number of hidden layers in the model shows an improvement in prediction power until it reaches maximum capability at which point a slight reduction in performance occurs specifically displayed by the 5-hidden layer model results.

Table 13. Error metrics showing the performance comparison of ANN model for static contact angle dataset

		1 hidden layer	2 hidden layers	3 hidden layer	4 hidden layer	5 hidden layer
Training Data	RMSE	10.44	5.63	5.77	3.64	4.28
	MSE	108.99	31.70	33.29	13.25	18.32
	R^2, fraction	0.72	0.92	0.91	0.97	0.95
	AIC	1924.37	1420.39	1440.38	1064.1	1195.53
	BIC	1944.43	1440.45	1460.43	1084.16	1215.59
Testing Data	RMSE	13.56	7.68	8.06	6.5	7.06
	MSE	183.87	58.98	64.96	42.25	49.84
	R^2, fraction	0.56	0.86	0.85	0.9	0.88
	AIC	927.79	727.54	744.64	668.61	697.75
	BIC	943.64	743.39	760.49	684.46	713.61

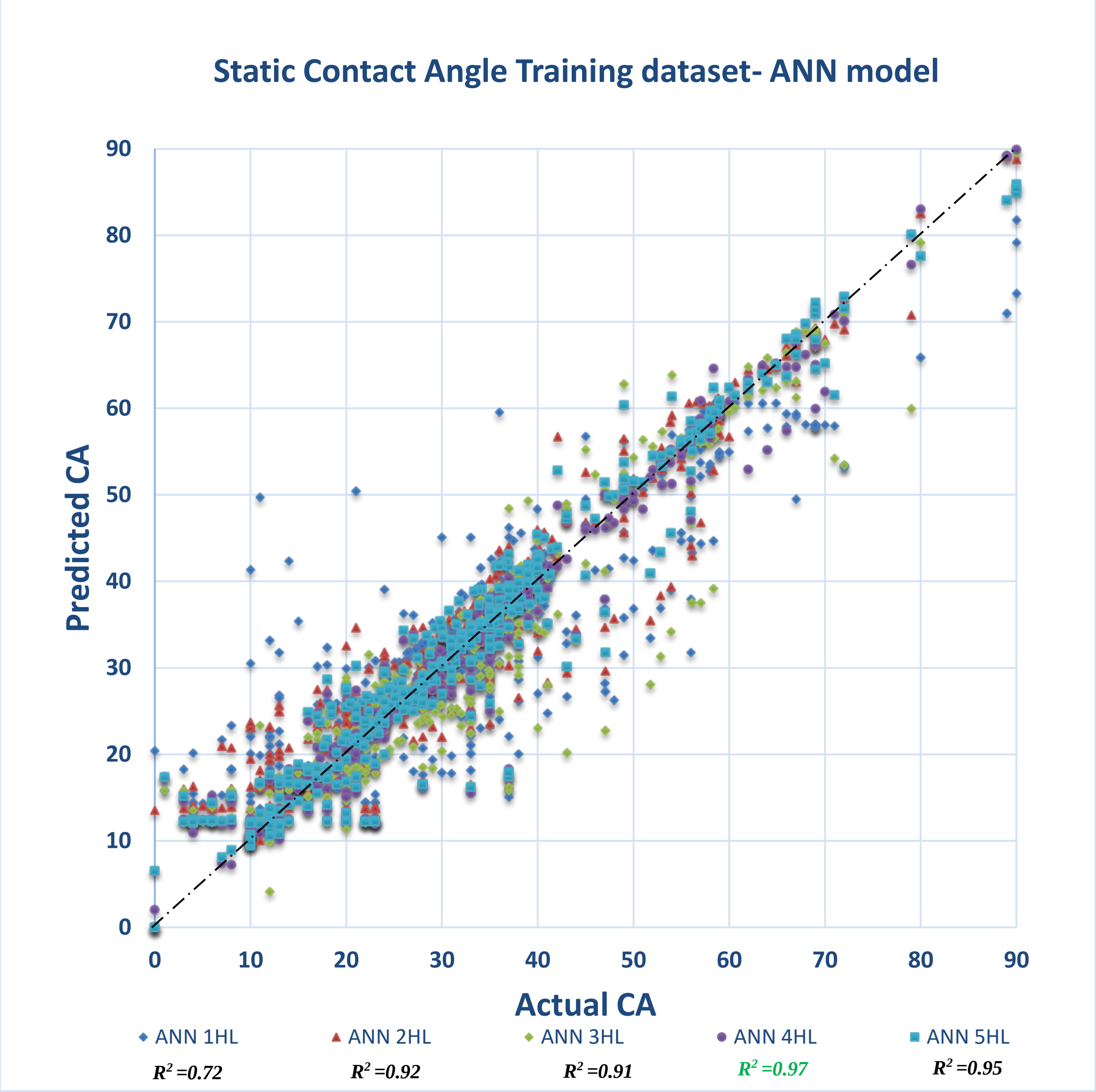


Figure 22. Static Contact Angle Training dataset- ANN model

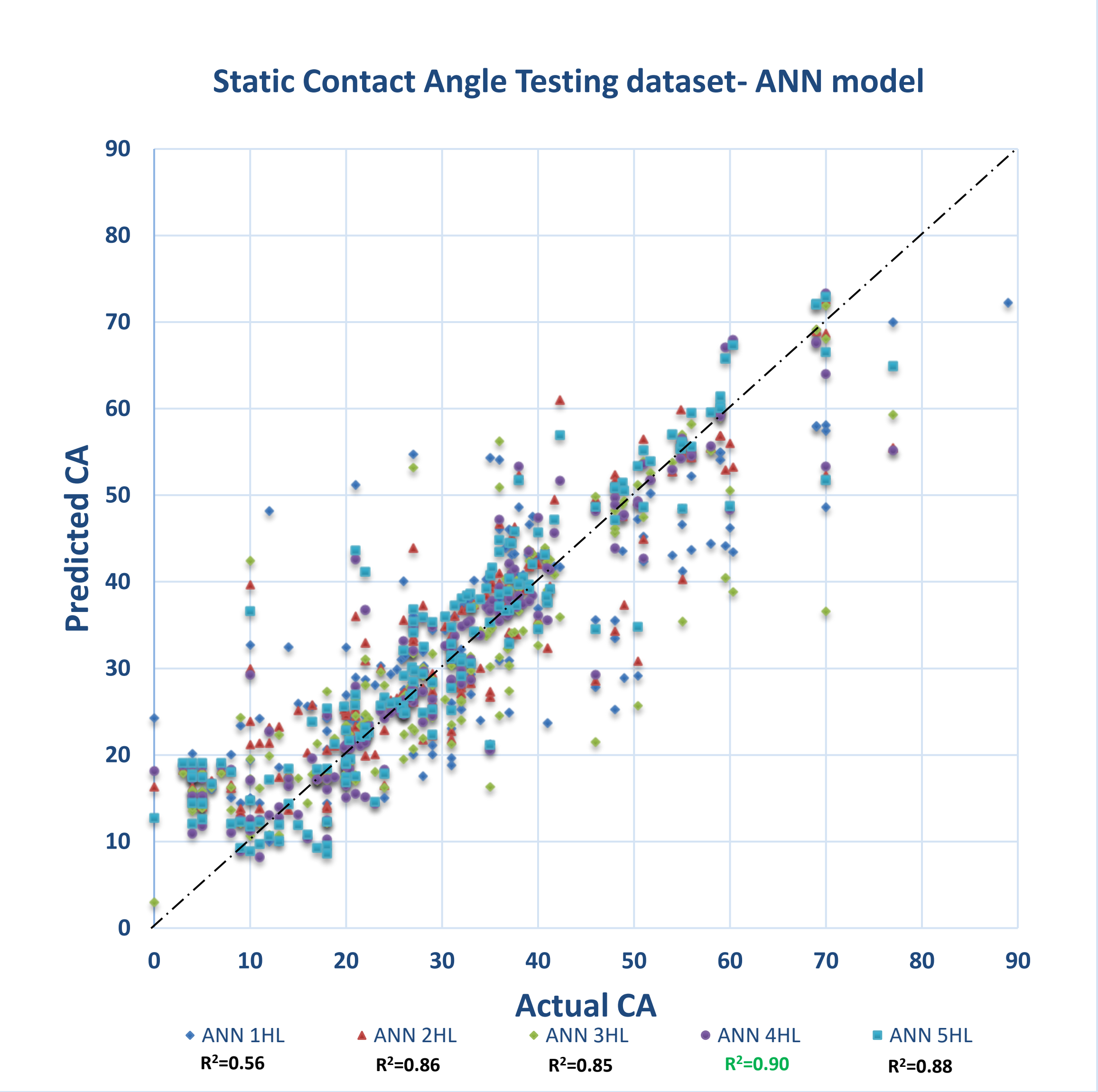


Figure 23. Static Contact Angle Testing dataset- ANN model

4.2 Implementation of 6 ML models for static contact angle dataset.

The static contact angle prediction within a CO_2-Water-Rock system used six machine learning algorithms and tested their performance with SVM and DT and RF and XGBoost and GBR and ANN. Multiple error metrics evaluated the assessment of models through training data and testing datasets. Table 14 compares the performance of six machine learning models (SVM, DT, RF, XGBoost, GBR, ANN-4HL) for predicting static contact angles. XGBoost produces the minimum RMSE of 1.25 and MSE of 1.56 during training which GBR closely follows with RMSE of 1.3 and MSE of 1.69. The data exhibits high excellence in fit reaching R^2 values of 0.99 for XGBoost and GBR as well as for RF. The Support Vector Machine (SVM) presents the worst performance compared to other six models across training and testing data due to its poor error metrics together with low R^2 values. The R^2 values for testing data show quality prediction performance ranging from 0.84 for SVM up to 0.92 for GBR and RF but especially for XGBoost. Besides testing phase performance the GBR model displays competitive metrics which demonstrates its suitability for this kind of data. Figure 24 compares the coefficient of determination for static CA training and testing datasets across 6 modes. Figure 25 compares the predictive performance of 6 ML models on static CA and training. Datasets, with GBR and XGB showing near-perfect fits (R=0.99). Figure 26 shows their testing performance, where RF and GBR maintain strong predictive capability (R2=0.92). Both figures highlight the performance of ensemble methods.

Table 14. Error metrics showing the performance comparison of all applied ML tools for static contact angle

		SVM	DT	RF	XGBoost	GBR	ANN-4HL
Training Data	RMSE	4.8	3.76	2.73	1.25	1.3	3.64
	MSE	23.04	14.14	7.45	1.56	1.69	13.25
	R^2, fraction	0.94	0.96	0.98	0.99	0.99	0.97
	AIC	1290.28	1091.76	830.97	190.02	221.63	1064.1
	BIC	1310.33	1111.82	851.02	210.08	241.69	1084.16
Testing Data	RMSE	8.11	7.9	5.6	6.46	5.73	6.5
	MSE	65.77	62.41	31.36	41.73	32.83	42.25
	R^2, fraction	0.84	0.85	0.93	0.90	0.92	0.90
	AIC	746.67	737.51	616.64	666.74	624.6	668.61
	BIC	762.52	753.36	632.5	682.59	640.45	684.46

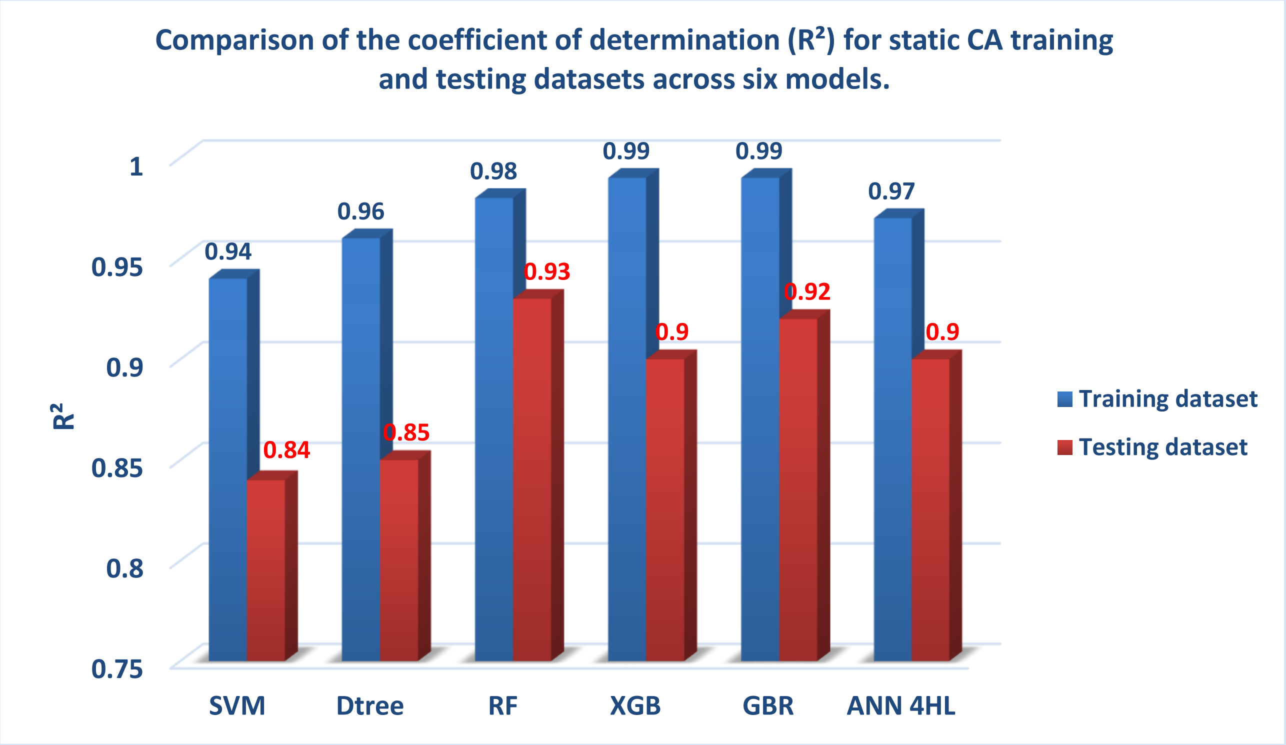


Figure 24. Comparison of the coefficient of determination (R^2) for static CA training and testing datasets across six models.

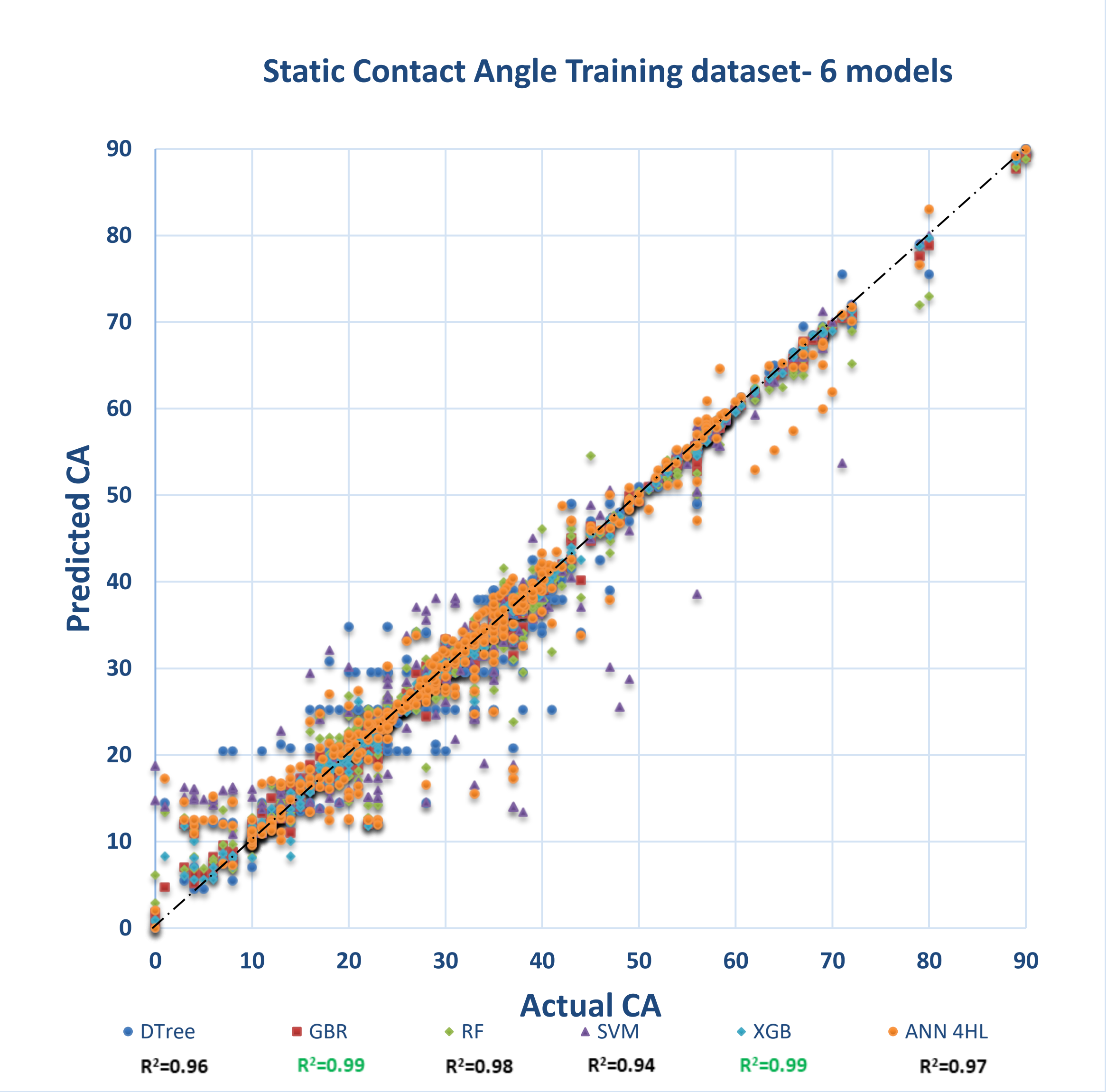


Figure 25. A comparison of scatter plots for all applied ML tools using static contact angle training dataset.

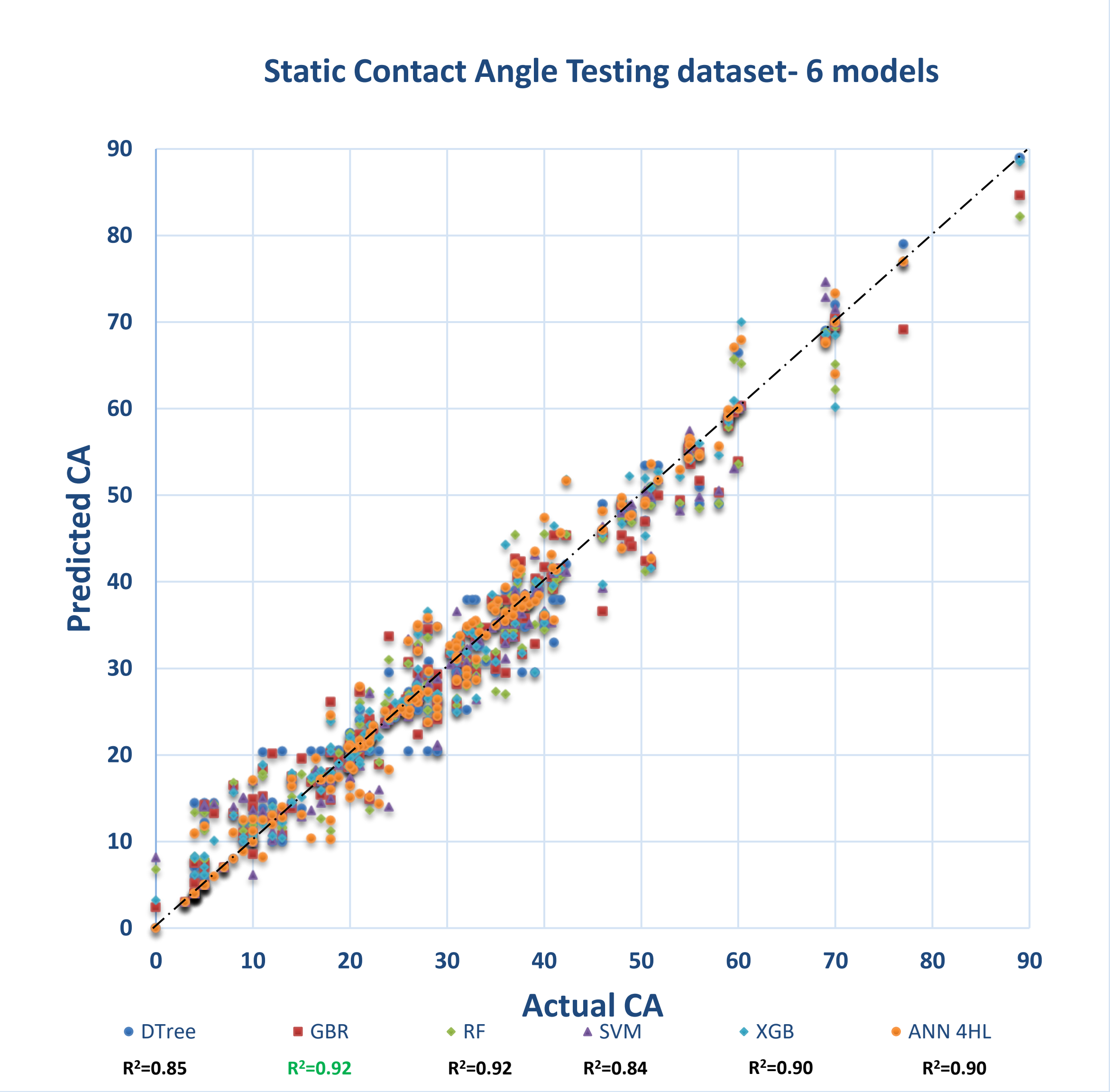


Figure 26. A comparison of scatter plots for all applied ML tools using static contact angle testing dataset.

4.3 Implementation of 5 ML models for advancing and receding contact angle dataset.

The advancing and receding contact angle prediction within a CO_2-Water-Rock system used six machine learning algorithms and tested their performance with DT, RF, XGBoost, GBR and ANN. Multiple error metrics evaluated the assessment of models through training data and testing datasets. This graphic demonstrates how well the models maintain prediction capabilities for information which has not been encountered during training. Analysis of all models through their data points indicates prediction accuracy because they follow the diagonal line pattern. The data points across the testing data deviate slightly from the diagonal line while maintaining greater scatter when compared to the training data points. The study shows R^2 values falling between 0.72 from ANN 3HL and 0.88 from GBR which indicates GBR provides the best performance on new data sets. Figure 27 shows comparison of the coefficient of determination (R^2) for advancing and receding CA training and testing datasets across 5 models. In Figure 28 the training dataset depictions are organized through a scatter plot representation. The testing data evaluation appears in Figure 29. Training data points of every model tightly align with the diagonal line to show excellent match with the training data points. All statistical models in this analysis exhibit high R^2 values which span between 0.86 for ANN 3HL up to 0.93 for the other models. The training data patterns seem well understood by most of the models being evaluated. The error metrics table (Table 15) shows detailed performance analysis of each constructed model. The Random Forest (RF) model demonstrates the lowest RMSE value of 6.23 along with MSE value of 38.81 while evaluating training data points. The Gradient Boosting Regressor (GBR) model achieves the top performance on testing data through its minimal RMSE value of 8.1 and maximum R^2 value of 0.88. The training data shows that RF provides better fit than GBR but GBR demonstrates superior capacity to predict new data points. The ANN-3HL model demonstrates the most erroneous results alongside the weakest R^2 values which indicates it would not perform as well on this problem compared to its tested counterparts.

Table 15. Error metrics showing the performance comparison of all applied ML tools for advancing and receding contact angle

		DT	RF	XGBoost	GBR	ANN-3HL
Training Data	RMSE	6.44	6.23	6.69	6.54	9.19
	MSE	41.47	38.81	44.76	42.77	84.46
	R^2, fraction	0.93	0.93	0.92	0.93	0.86
	AIC	1877.78	1844.75	1916.37	1892.94	2234.35
	BIC	1903.08	1870.05	1941.67	1918.24	2259.65
Testing Data	RMSE	11.37	8.64	8.5	8.1	12.34
	MSE	129.28	74.65	72.25	65.61	152.28
	R^2, fraction	0.76	0.86	0.87	0.88	0.72
	AIC	624.57	555.54	551.39	539.06	645.19
	BIC	641.59	572.56	568.41	556.08	662.21

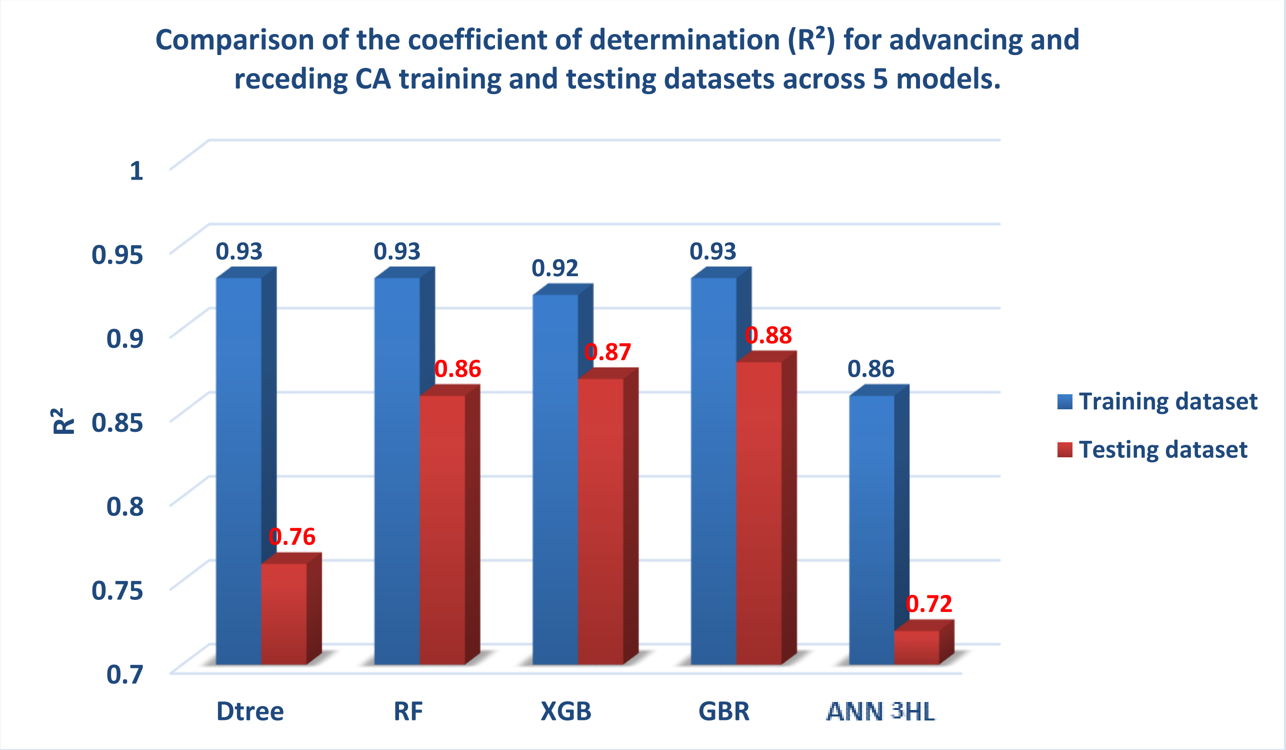


Figure 27. Comparison of the coefficient of determination (R^2) for advancing and receding CA training and testing datasets across 5 models.

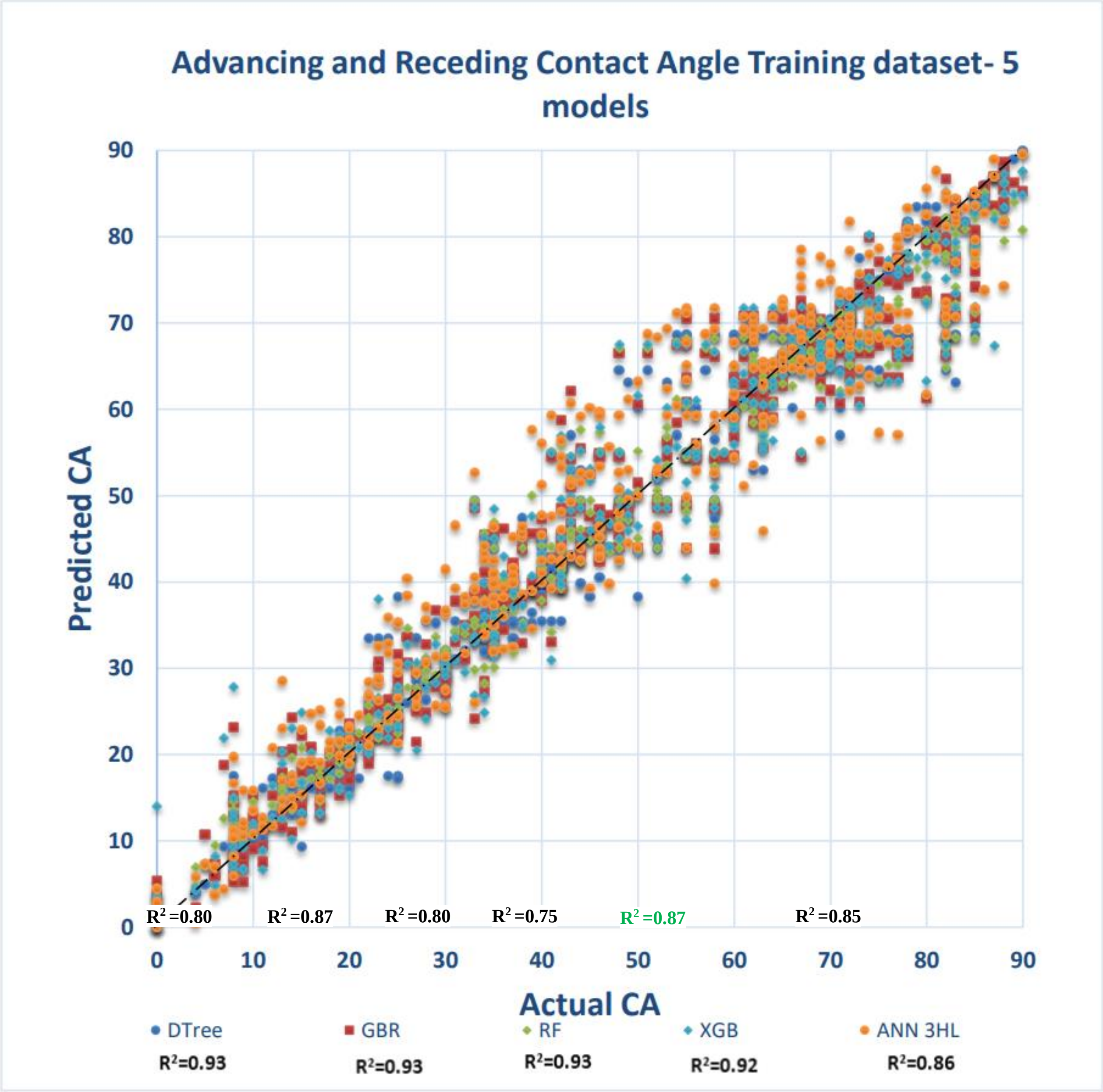


Figure 28. A comparison of scatter plots for all applied ML tools using advancing and receding contact angle training dataset.

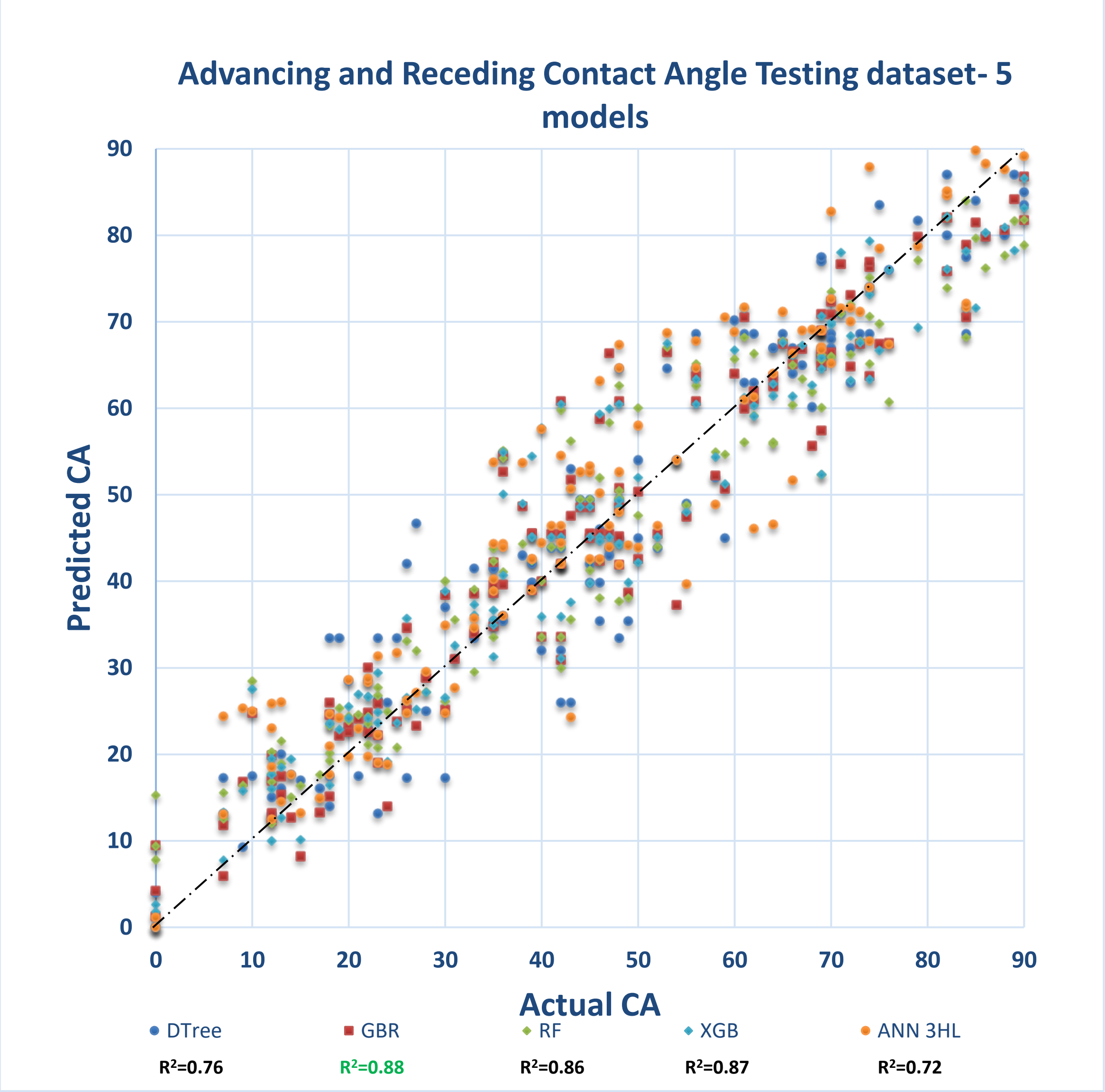


Figure 29. A comparison of scatter plots for all applied ML tools using advancing and receding contact angle testing dataset.

4.4 Sensitivity Analysis

4.4.1 Sensitivity Analysis for Static Contact Angle dataset by GBR model

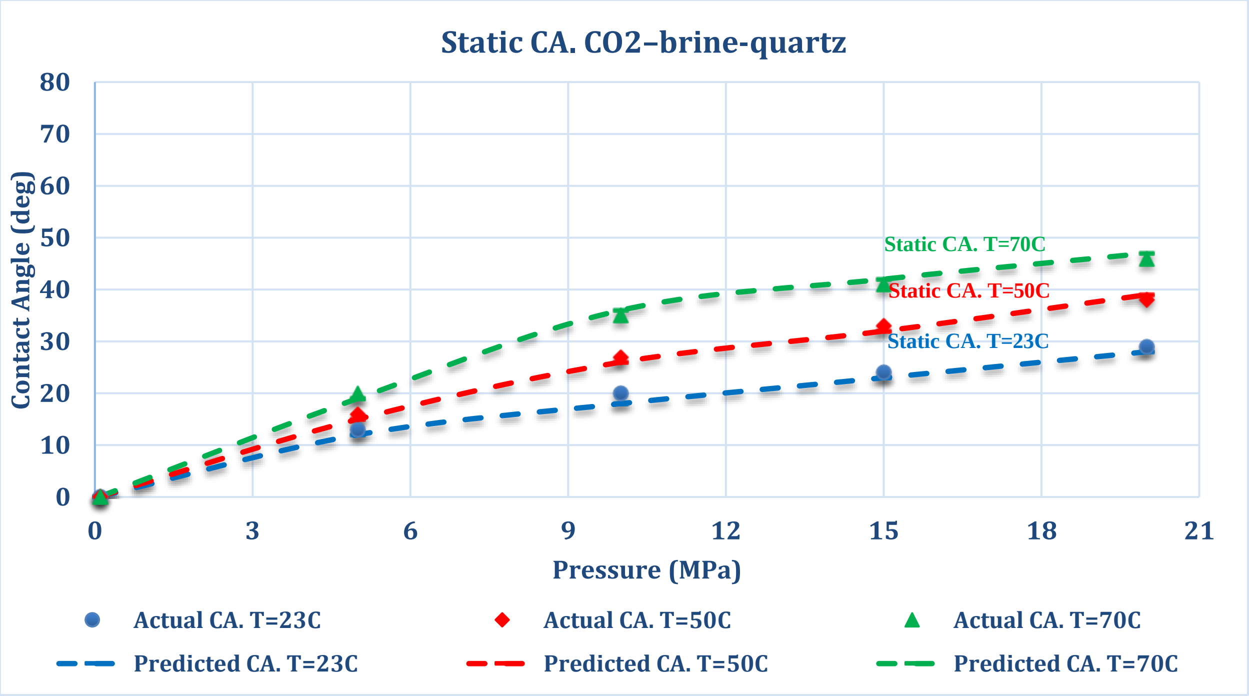


Figure 30. Sensitivity analysis on temperature for static contact angle dataset. CO_2-brine quartz system.

Figure 30 shows how static contact angle (CA) measurements vary when pressure increases within the CO_2-brine-quartz system at three different temperatures testing 23°C, 50°C and 70°C. Experimental trends from the data match the model predictions (dashed lines) that show both increased CA responses under pressure raising conditions and show the influence of temperature variations on surface wettability. The model exhibits reliable prediction capabilities throughout most of the temperature range although small prediction inaccuracies emerge at pressure settings between the lowest and highest pressures. Pressure increases the contact angle because it enhances CO_2 density which affects the interfacial characteristics. The wettability of CO_2 to quartz material increases with rising pressure which manifests as a rising trend in contact angle. The temperature conditions significantly impact this phenomenon. The CO_2 solubility in brine decreases at higher temperatures which results in an intensified CO_2-quartz interaction and ultimately raises the contact angle value. The measurement data in the generated figure demonstrates the pattern of higher contact angles recorded at 70°C when compared to 50°C and 23°C. Temperature effects on wettability theory (Anderson, 1986) produce such expected changes because they modify balances between surface energies and interfacial tensions between CO_2, brine and the mineral surface.

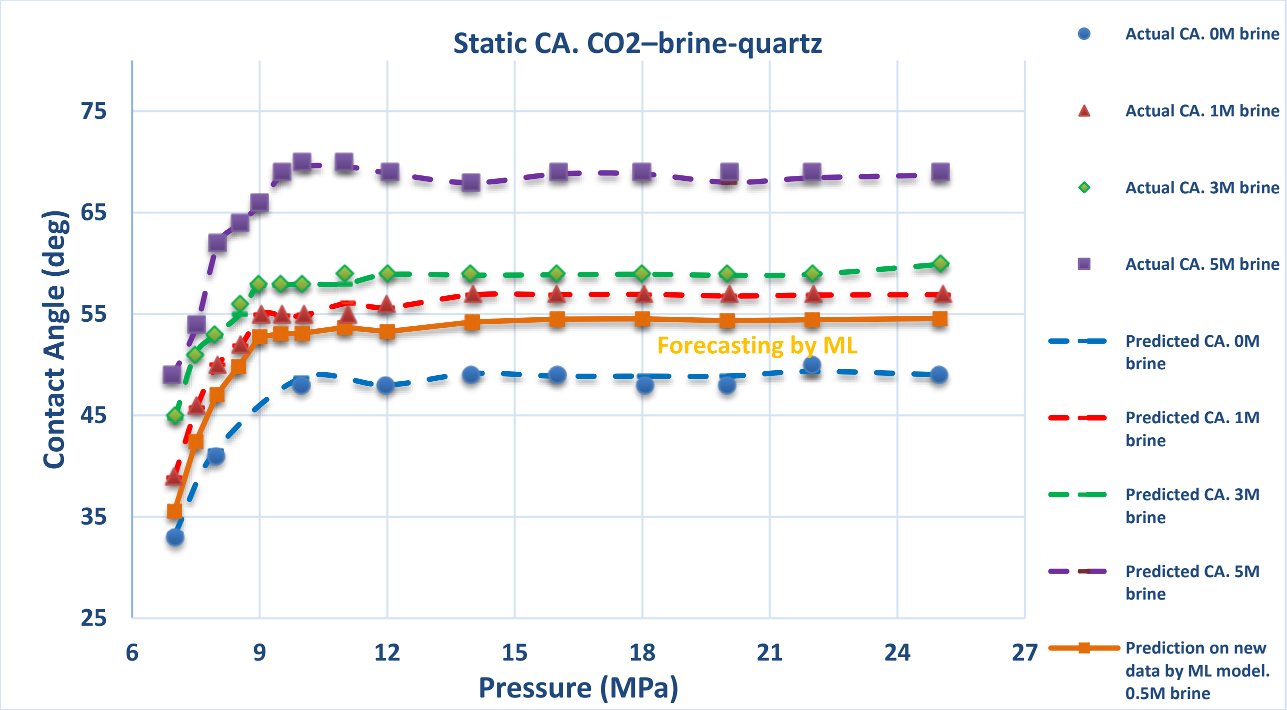


Figure 31. Sensitivity analysis on brine salinity for static contact angle dataset. CO_2-brine quartz system.

As shown in Figure 31, the dashed lines in the provided graph illustrate GBR model predictions which show a good representation of actual data points across different brine molarities. The model showcases high effectiveness at illustrating how elevated pressure leads to increased contact angle within the 3M and 5M brine concentration conditions. The predicted data slightly falls below actual measurements at lower pressure conditions in specific measurements involving brine solutions. Experimental results (Jung & Wan, 2012) support the pattern which shows contact angle growth until pressure reaches 10 MPa then achieves stabilization. The dissolution of CO_2 into brine solutions creates changes in the surface interactions that exist between the quartz interface. Lower $scCO_2$ pressures result in surface modification of silica by interacting with hydroxyl groups to achieve enhanced hydrophobicity (Jung & Wan, 2012). The wettability change achieves equilibrium through surface interaction saturation when pressure surpasses 10 MPa. Previous findings by Jung & Wan (2012) established that contact angle measurement data aligns with observations because increased brine molarity weakens electrostatic repulsion between CO_2 and mineral surfaces thus facilitating better CO_2 adhesion. The static CA for 5M brine maintains higher values than 0M and 1M brine tests. While the GBR model replicates the observed trends it shows small inaccuracies in the lower pressure range because of multiple factor interactions.

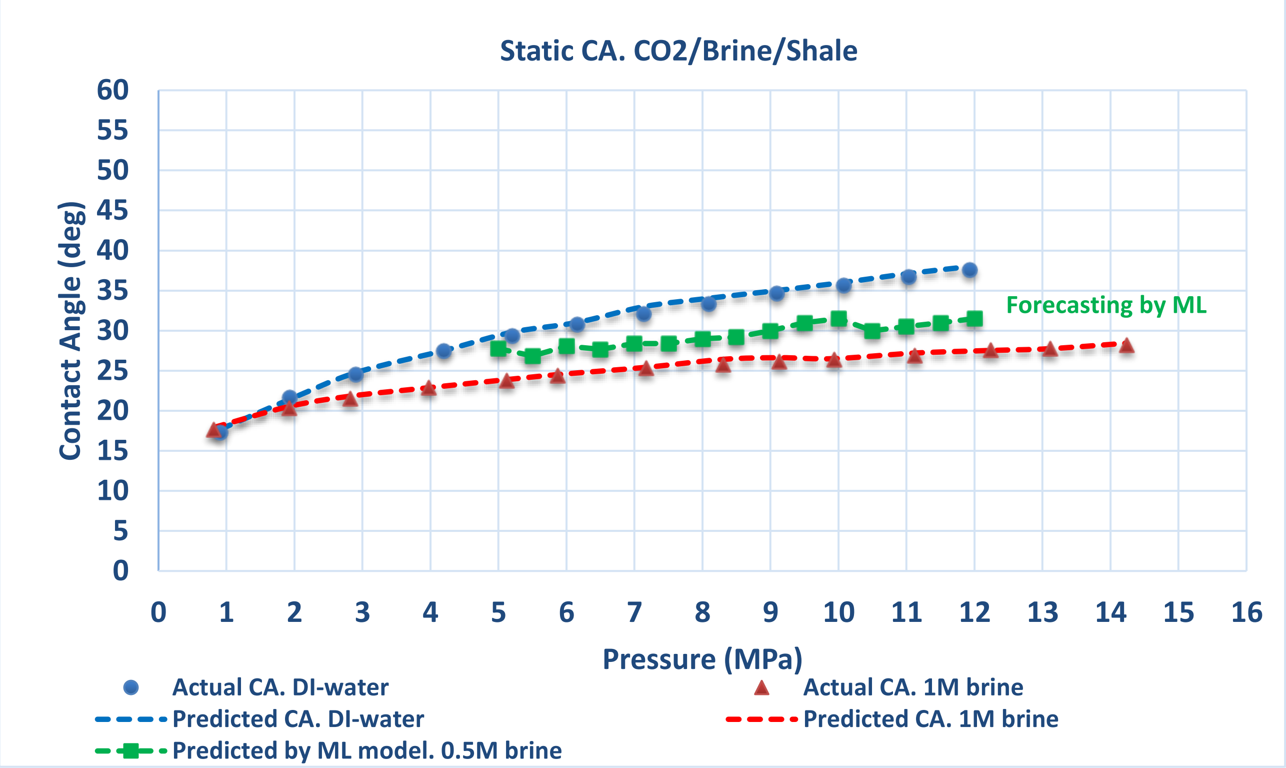


Figure 32. Sensitivity analysis on brine salinity for static contact angle dataset. CO_2-brine shale system.

As can be seen from Figure 32, experimental observations confirm that GBR prediction shows contact angles rise progressively when $ScCO_2$ pressure increases. The data indicates that shale becomes less water-wet as the pressure of $scCo_2$ increases. Changes in mineral composition particularly the loss of calcite play a significant role in water contact angle behavior. The trend develops from CO_2-brine-rock interactions that modify shale surface chemistry along with fluid transport behavior. The wettability changes from waterrustling to weak affinity when the reservoir encounters elevated $ScCO_2$ pressures because these conditions reduce capillary forces while affecting the movement of fluids.

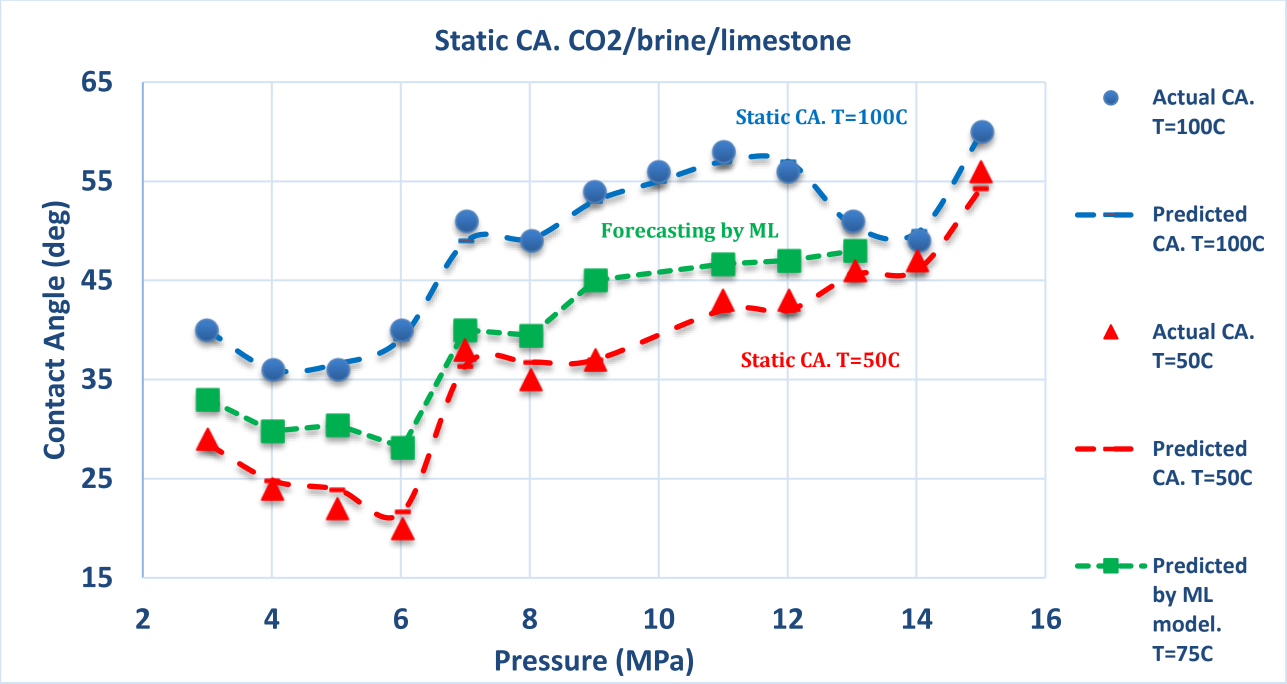


Figure 33. Sensitivity analysis on temperature for static contact angle dataset. CO_2-brine-limestone system.

Based on Figure 33, there was a minimal elevation of CAs at each temperature increase across subcritical and supercritical CO_2 regions. Pressure data showed rapid initial decreases within subcritical CO_2 scope followed by moderate general increases within supercritical CO_2 range. A significant drop in CAs occurred when the pressure reached from 11 MPa to 14 MPa at 100 °C and then pressure measurements showed increasing CAs at 15 MPa. The alteration of wettability properties shifted surfaces of limestone to become less water-wet during the supercritical CO_2 region. Outside the supercritical CO_2 region the media complexity caused the system to exhibit matching jumps and changes. Research by Arif et al. (2017) revealed how pressure along with salinity enhanced CO_2– brine–mica CAs while temperature caused them to decrease. Broseta et al. (2017) found that advancing and receding CAs of CO_2–brine–mica systems followed a temperature-dependent pattern which exhibited initial decrease from 282 to 308 K then started increasing at 393 K. GBR model model output with dashed lines tracks actual contact angle (CA) measurements very closely for conditions at both 50°C and 100°C. Salinity of brine is 1.98mol*kg-1 . The prediction model reproduces the typical pressure-related changes in contact angle measurements which show good correlation with experimental data. The model demonstrates minor deviations when sharp changes take place because it possibly cannot fully represent nonlinear pattern shifts. The GBR model showed accurate results indicating its ability to understand the connection between CA and pressure and temperature measurements.

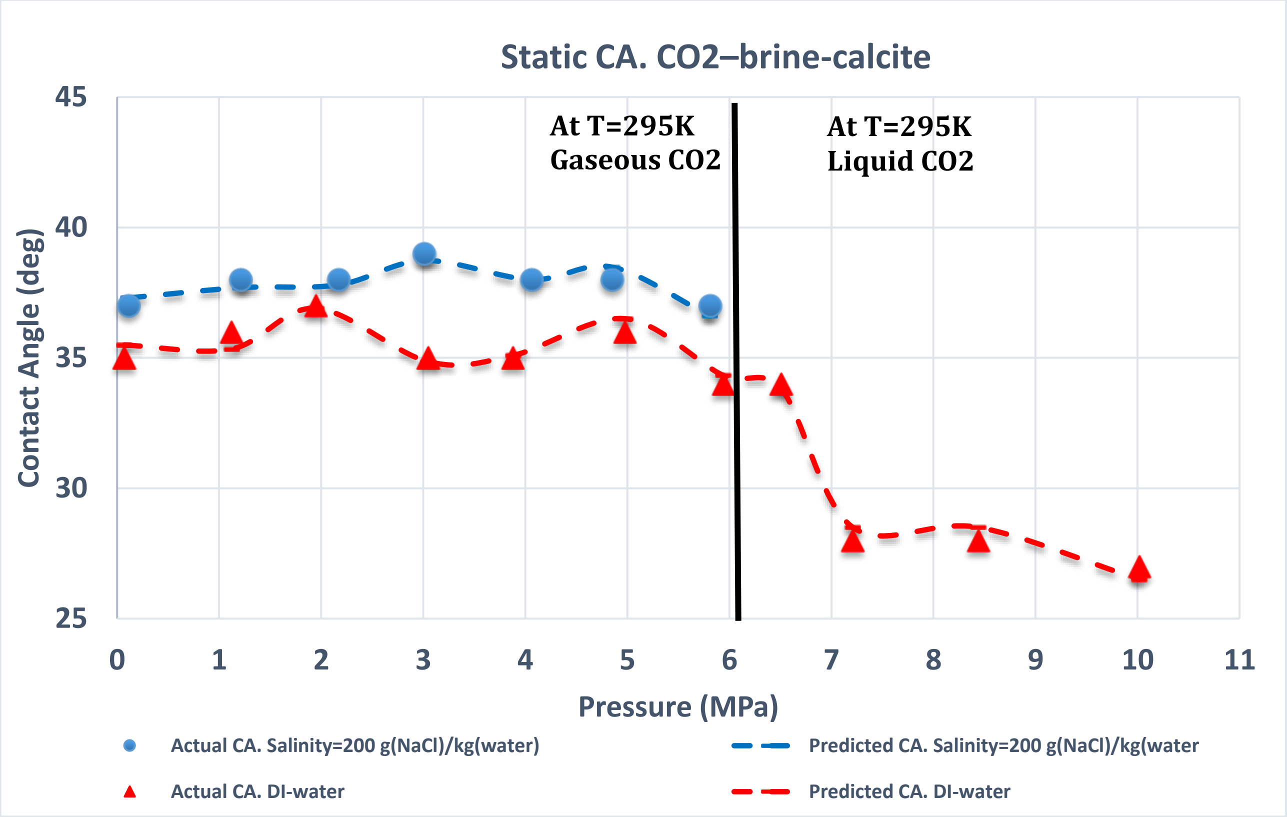


Figure 34. Sensitivity analysis on brine salinity for static contact angle dataset. CO_2-brine-calcite system.

As shown in Figure 34, GBR model accurately tracks the overall patterns at both salinity levels but shows certain deviation points which materialize when sharp fluctuations emerge. At pressures below 6 MPa the accuracy appears higher while the measurements for DI water show greater discrepancies beyond that critical pressure. Both CA and pressure and salinity show an effective relationship within the model which fails to reliably handle abrupt events such as the gaseous to liquid CO_2 state change. Contact angle variations depend on the phase characteristics of CO_2 at 295 K. The contact angle exhibits minimal variability in the gaseous CO_2 area although pressures stay under approximately 6 MPa. The weaker CO_2-water bonding produces a constant wettability state. The contact angle suddenly decreases as pressure exceeds 6 MPa when CO_2 evolves into its liquid state. Liquid CO_2 shows increased water solubility which alters the chemical properties at brine-calcite interfaces thus developing a more water-wet state. The measurement shows greater changes for DI water compared to high-salinity brine because salinity acts as a stabilizing factor for contact angles.

4.4.2 *Sensitivity Analysis for Advancing and Receding Contact Angle dataset by GBR model*

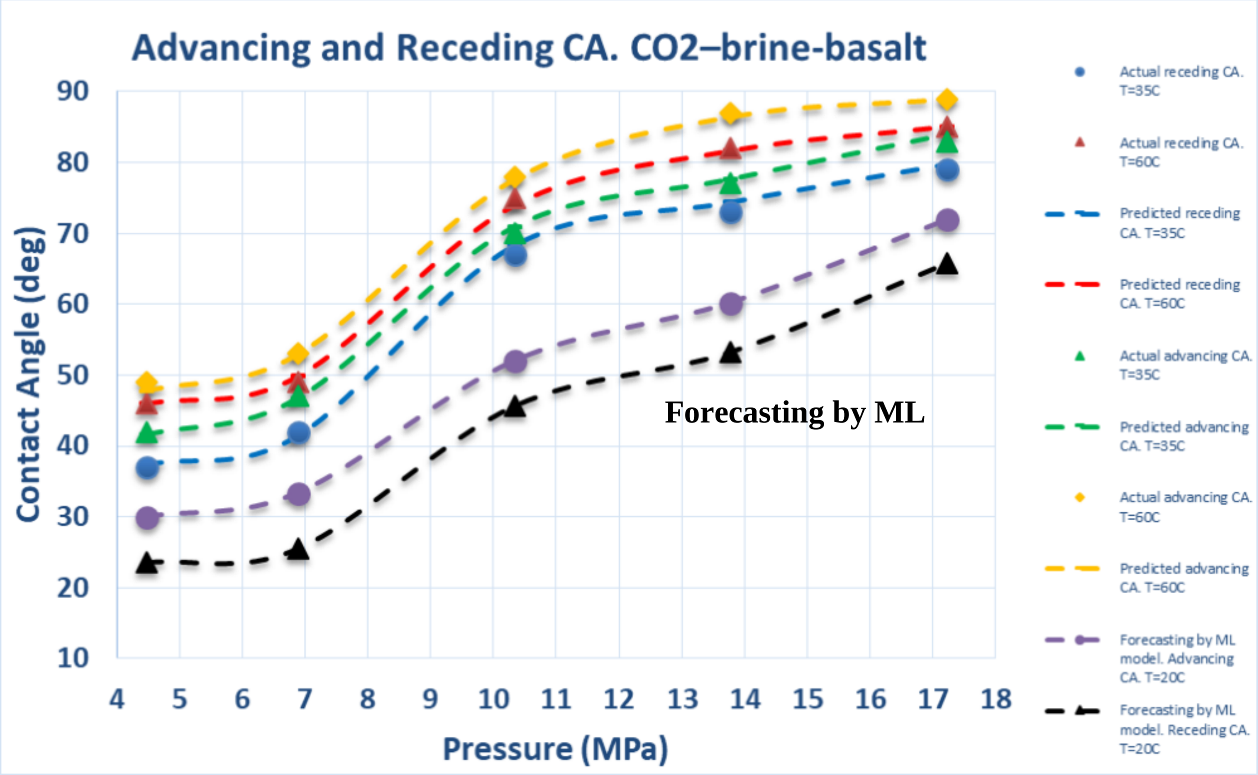


Figure 35. Sensitivity analysis on temperature for advancing and receding contact angle dataset. CO_2-brine-basalt system.

The obtained advancing and receding water contact angles by employing the tilted plate method on typical storage pressures (4.48–17.23 MPa) at two temperatures (35 °C and 60 °C) using synthetic reservoir brine with 10 wt.% salinity. Figure 35 illustrates that the contact angle between brine and solid gradually increased progressively with rising pressure. Higher pressures strengthen the intermolecular forces between CO_2 and basalt which leads to this effect. The pressure increase above 9 MPa transformed the basalt into a medium-wet state with contact angles exceeding 70◦. The upward movement of CO_2 is predicted to occur rapidly through intermediate-wet systems (Al-Khdheeawi et al., 2017) resulting in diminished storage capacities together with reduced containment security. Intermediate-wet conditions affect how CO_2 fills rock pores because important portions of CO_2 now reach basalt surface contacts (Blunt, 2017). High partial pressure of CO_2 leads to significant acidification of brine which affects the interfacial areas between brine and basalt and therefore modifies mineralization rates. At approximately 7 MPa the rocks experienced a shift from strong water-wetness to weak water-wetness ($\theta > 50$◦) with comparable impacts as previously described. The change in temperature produced minimal variations which resulted in contact angles that were slightly increased at higher temperatures. The strong dependence of CO_2

mineralization on wettability change becomes clear since changing interfacial areas between CO_2 and basalt along with brine and basalt would impact 60 this reaction process Blunt (2017). The dewetting effect of high CO_2-rock wettability would in turn reduce dissolution-precipitation reactions because intermediate-wet rock has minimal brine-basalt contact areas. The GBR model shows reliable performance for predicting overall trends but it does not capture data points accurately at lower pressures. The measured advancing angles exceed the receding angles because of the typical hysteresis effect in wetting behavior. The model demonstrates effective prediction of contact angle changes caused by pressure and temperature conditions although it shows minor prediction inaccuracies at reduced pressure values.

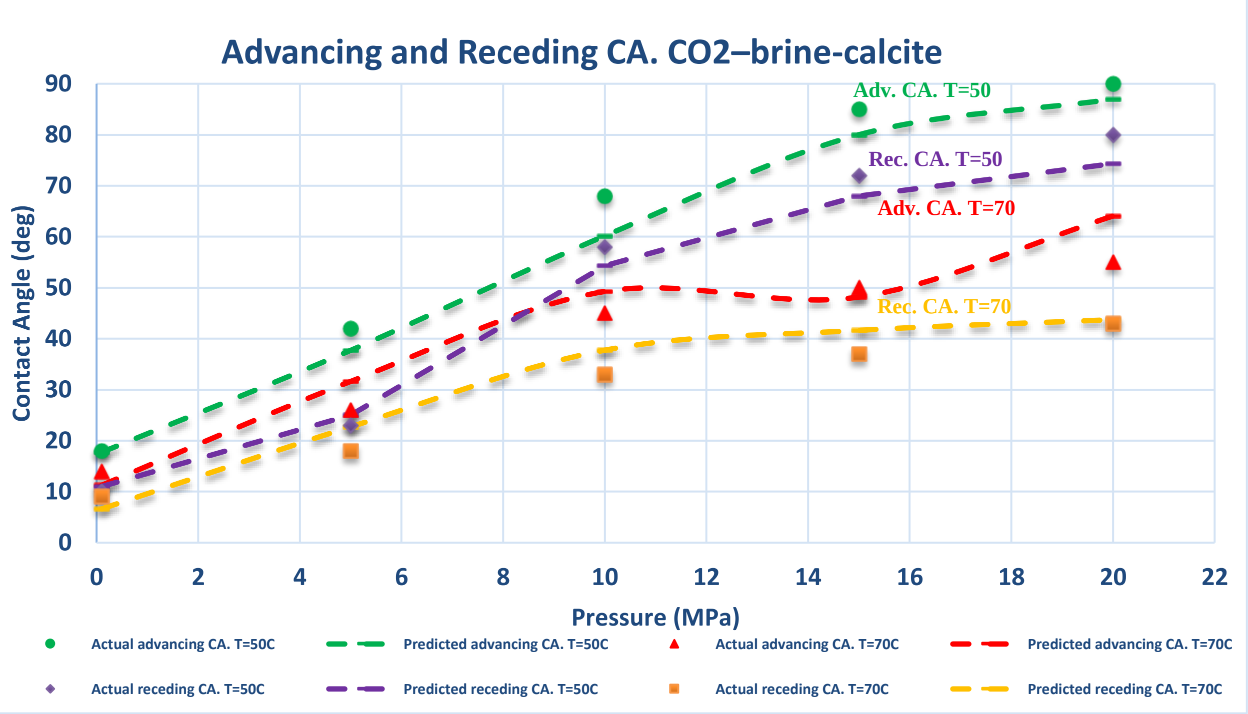


Figure 36. Sensitivity analysis on temperature for advancing and receding contact angle dataset. CO_2-brine-calcite system.

A study of pressure effects examined values from 0.1 MPa to 20 MPa at the temperatures 50 °C and 70 °C for CO_2-brine-calcite. Brine salinity-20 wt.% NaCl. Each of θa and θr values rose across the entire experimental pressure range. Calcite maintained a strong water-wet condition across all temperatures at an ambient pressure setting of 0.1 MPa (e.g. at 50 °C and 0.1 MPa θa was 18° and θr was 10° indicating the surface was highly water-wet). Al-Anssari et al. (2016) confirmed the strong water-wet behavior of calcite at standard conditions which matches observations of other significant minerals like quartz and mica and shale (quartz with θ = 0°; Al-Yaseri et al., 2016, mica with θ = ~10°; Arif et al., 2016b, and shale with θa = 18° and θr = 13°; Roshan et al., 2016). As can be seen from Figure 36, the angle of water contact with calcite (θa) rose from 25° to 42° as pressure raised from 0.1 MPa to 5 MPa at 50 °C and the contact angle (θr) rose from 12° to 22° indicating decreased water-wetness in calcite. Pressure continuing to rise to 10 MPa shifted the calcite/CO_2/water system to become intermediate-wet with contact angles of θa = 68° and θr = 58°. Wettability analysis indicates that the system underwent a transition to weak CO_2-wet character when pressure exceeded 20 MPa which reveals a high pressure-induced hydrophobicity of calcite surfaces. Under these prevailing conditions CO_2 shows upward migration capacity because of an upward suction force that works along with the buoyancy force. Upward migration of CO_2 can occur because of an upward suction force (in combination with buoyancy force) which can result in potential CO_2 leakage through. GBR modeling system makes accurate contact angle predictions for the CO_2–brine–calcite system that supports experimental measurement data. The model detects both how CA increases due to pressure changes and how temperature affects CA behavior. Areas for model enhancement exist at high pressures but the model demonstrates effective tracking of wettability modification throughout.

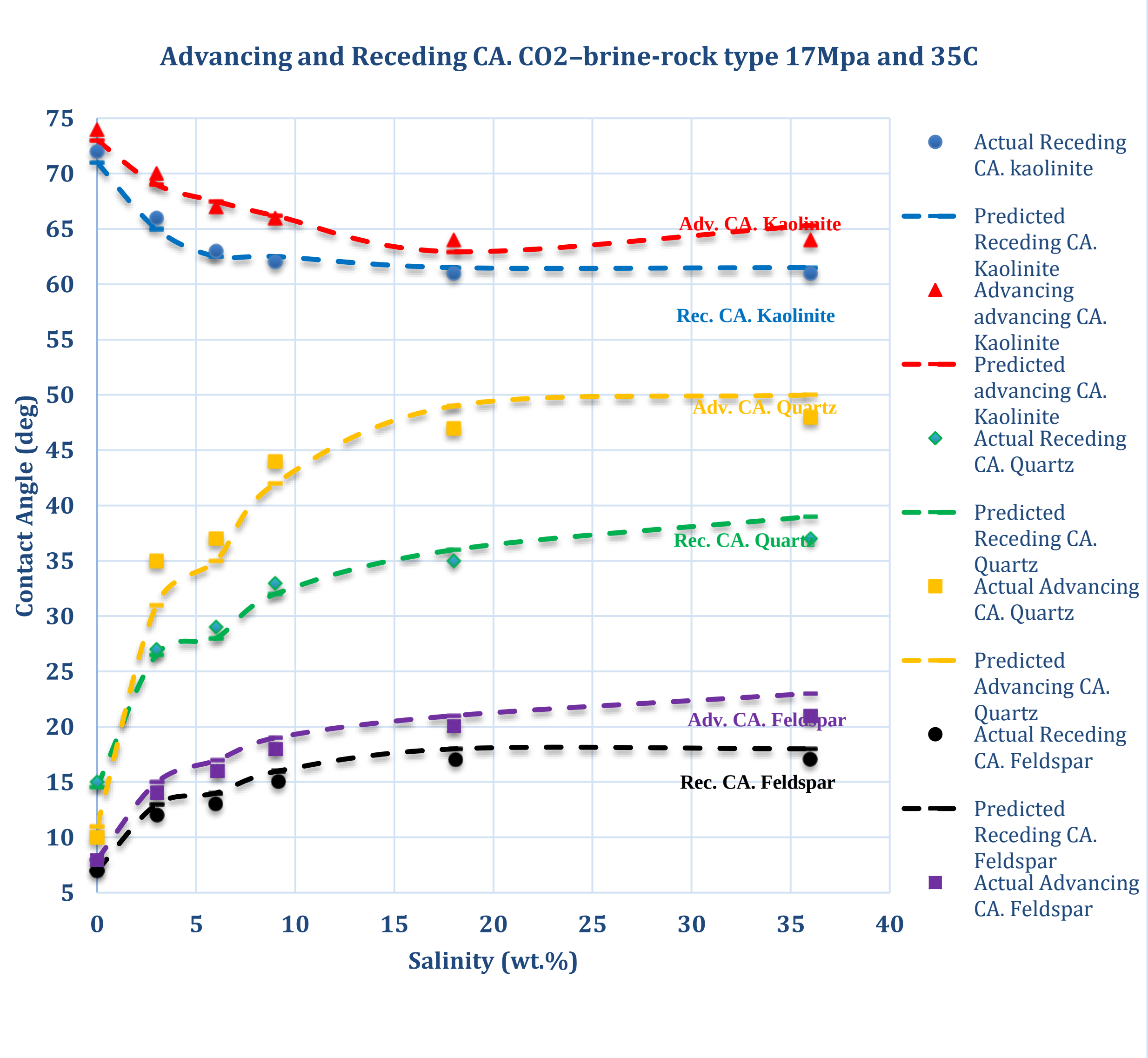


Figure 37. Sensitivity analysis on brine salinity for advancing and receding contact angle dataset. CO_2-brine- kaolinite, quartz, feldspar rock types system.

Figure 37 illustrates a study of pressure effects examined values from 0.1 MPa to 17 MPa at the temperatures 20 °C-65 °C for CO_2-brine-rock type (kaolinite, quartz, feldspar). Brine salinity-0-36 wt.% NaCl. The measurements of quartz-dominated sandstone show that rising brine concentrations result in greater contact angles with brine obtaining larger contact angles than deionized Water. The contact angle increases as per Al-Yaseri et al. (2016) when the salt type changes since the cation charge-to-volume ratio leads to effective shielding. A completely water wet surface delivers efficient CO_2 migration protection yet requires thorough examination of the pore throat dimensions. The findings by Iglauer et al. (2015) support these results because they found feldspar contact angles remain stable throughout variations of temperature and pressure. The study by Farokhpoor et al. (2013a) found that feldspar contact angle demonstrated resistance to solution salinity modifications while experiencing only small fluctuation. The accurate measurement of capillary pressure depends on an evaluation of brine ionic strength. Under reservoir conditions kaolinite-dominated shale demonstrates poor performance as a barrier to CO_2 migration since it fails to fulfill standard assumptions regarding reservoir protection. Introducing water with salt resulted in a decrease of CO_2/water/ kaolinite contact angle contrary to the observed behavior with quartz and feldspar. The research results demonstrate that water droplets containing deionized water form a contact angle that extends beyond the contact angle produced from brine solution drops. The possible cause of contact angle reduction is either known as changes in zeta potential values or escalated ionic strength effects. High negative zeta potential values generate high polar surfaces which improve the water wettability. Under a 36% wt NaCl salinity concentration the surface wettability remained below the strongly water wet condition. The reason for this observation stems from clays' active surface charges which make NaCl an ineffective inhibitor of clay–water interactions. The contact angle difference between advancing and receding for quartz becomes pronounced at increasing salinity values. Surface heterogeneity and roughness together generate high contact angle hysteresis rates according to these results. The rock-fluid interface bond strength boosts as salt concentration elevates therefore advancing contact angle increases but the receding contact angle maintains a reduced value. Impacts between quartz surfaces and CO_2 affect adhesion forces that result in significant hysteresis effects. The GBR model demonstrates acceptable accuracy for identifying kaolinite quartz and feldspar trends through it shows minor deviations from the true data points. The predictive model accurately represents the kaolinite data showing diminishing advancing and receding CA measurements at elevated salinity concentrations. The GBR prediction for quartz shows correct advancement of contact angles but underestimates receding contact angles leading to higher than actual hysteresis calculation. The model follows the advancing and receding contact angle trends for feldspar although it shows minor deviation during lower salinity ranges. The GBR system successfully demonstrates how changes in salinity affect contact angles but needs better calibration for quartz since its fixed angle hysteresis values stand out.

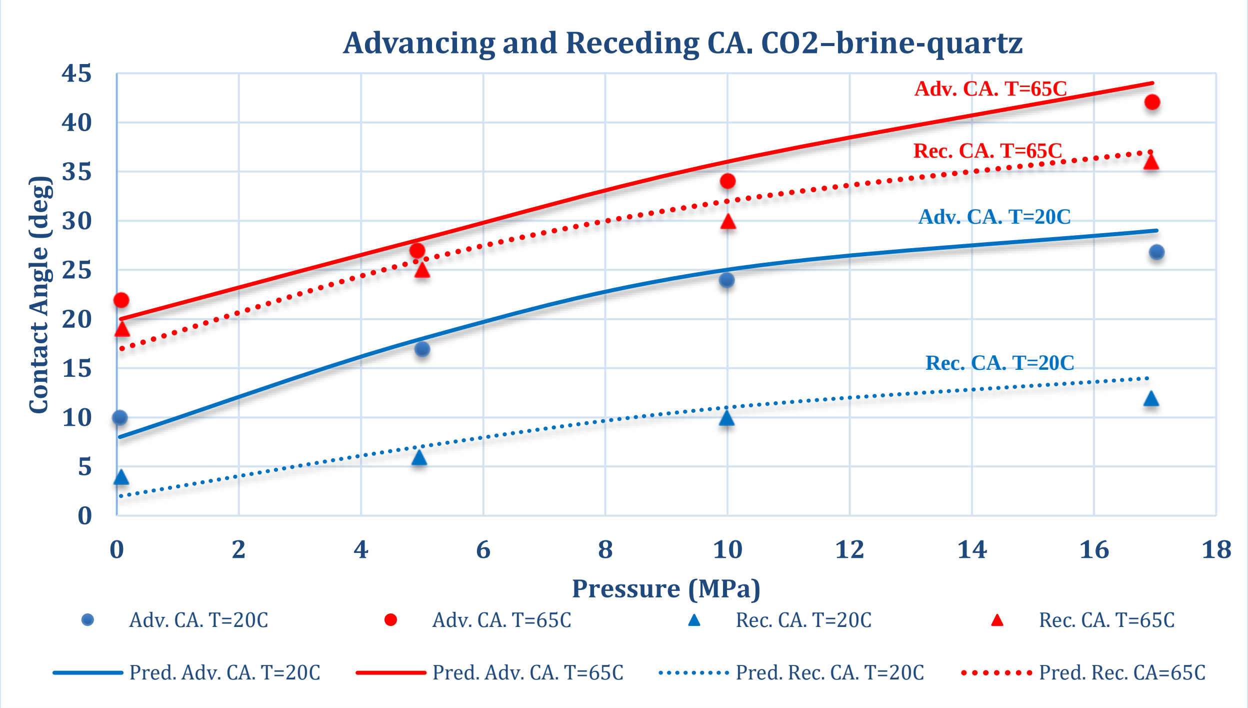


Figure 38. Sensitivity analysis on temperature for advancing and receding contact angle dataset. CO_2-brine-quartz system.

An increase in advancing CA produces substantial effects for quartz which leads to increased hysteresis as shown in Figure 38 and 39. The model provides accurate predictions for advancing CA yet displays slight underestimation of receding CA when pressure reaches high values. The GBR model demonstrates superior predictive ability for kaolinite materials than quartz samples especially when hysteresis phenomena remain minimal.

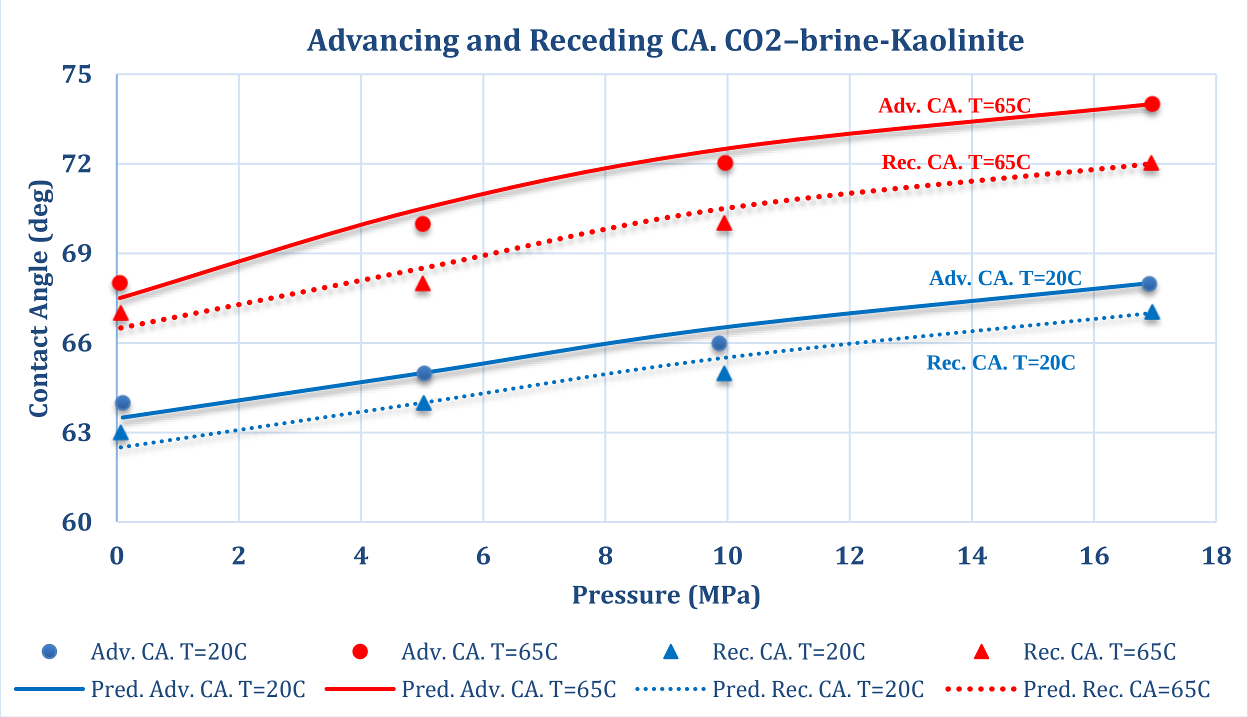


Figure 39. Sensitivity analysis on temperature for advancing and receding contact angle dataset. CO_2-brine-kaolinite system.

4.5 KDE plots

Static CA training datasets receive model implementation of XGBoost, SVR and DT, RF, GBR and ANN by the analysis. Figure 40 and 41 data displays the values between 0 and 90 at the x-axis point and density levels at the y-axis point. The actual distribution in Figure 40 and 41 has matching prediction patterns from the machine learning models that peak near the 25-30 range. According to model prediction data the DT model reaches its maximum density value over other models while delivering a slight overestimation in this particular area. The ANN model displays an underestimated peak compared to the actual data while it shows stronger accuracy when depicting the distribution's trailing section. Both peaks in the actual distribution stand out clearly with the major peak occurring between 30-35 and a secondary peak concentrated between 55. Most models display dual distribution patterns and the XGBoost with GBR models demonstrate the best match against the actual data distribution. The SVR model produces an output that blurs the second peak among other models.

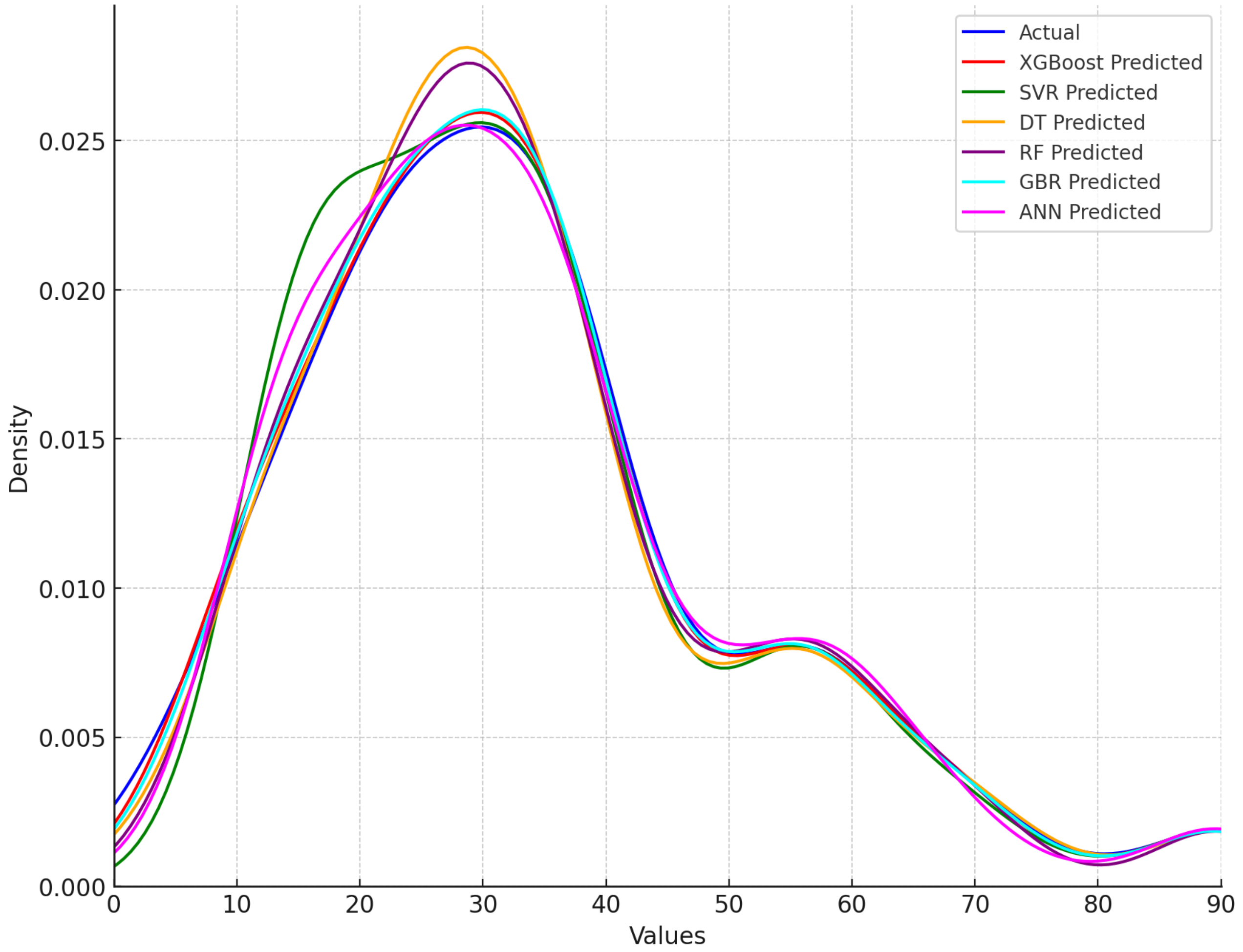


Figure 40. KDE plots by implementing static CA training dataset.

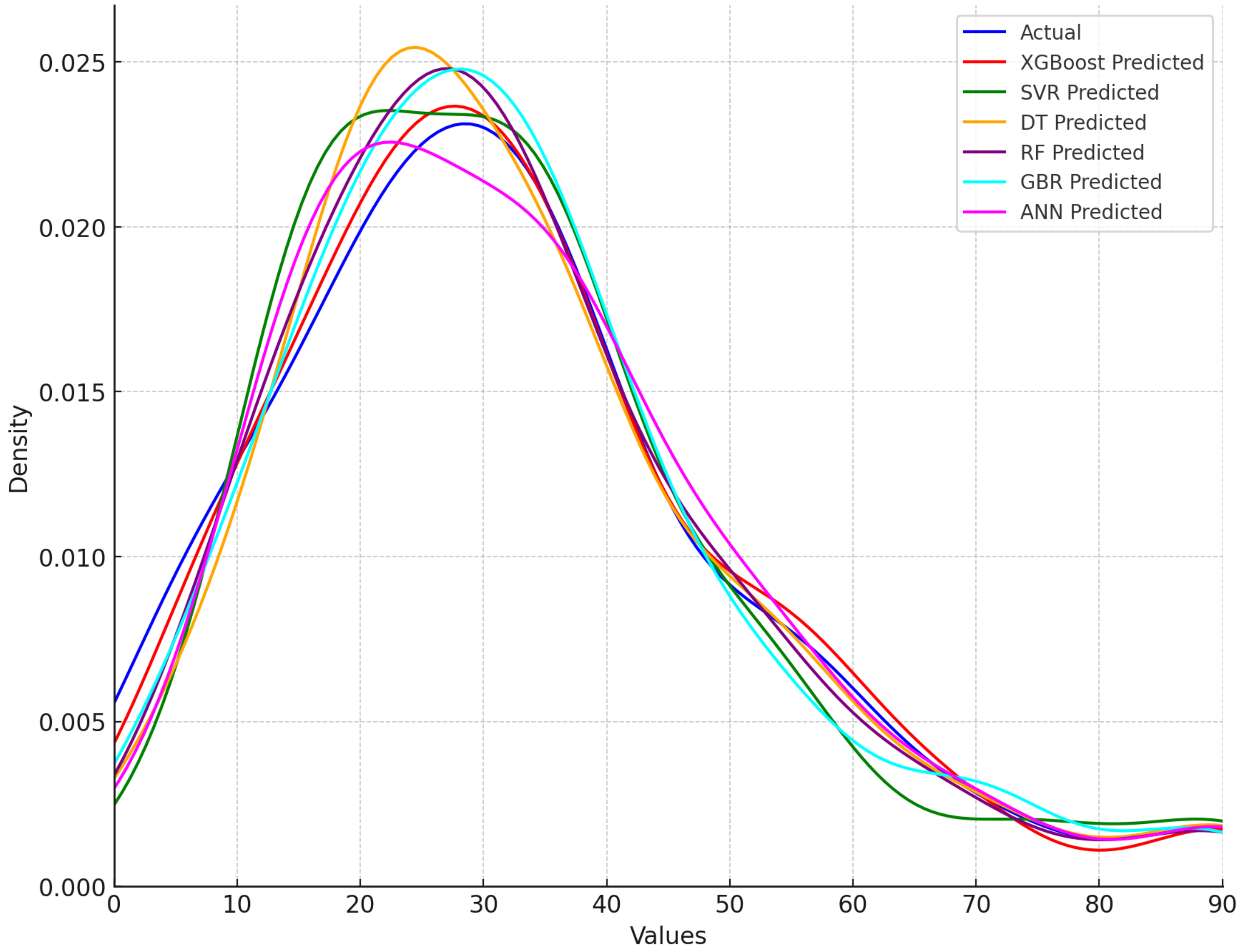


Figure 41. KDE plots by implementing static CA testing dataset.

5. CONCLUSIONS AND RECOMMENDATIONS

1. The GBR model delivered optimum performance in static contact angle prediction for CO_2- water-rock systems by reaching R^2 =0.99 value during training and R^2 = 0.92 value during testing. The high accuracy value shows GBR effectively identifies intricate non-linear patterns in wettability patterns.
2. The GBR model produced the most accurate results for testing dataset advancing and receding contact angles with a calculation error rate of R^2 =0.93 training and R^2 =0.88 testing. GBR model demonstrate excellent proficiency for understanding the changing behavior of liquid wetting in such systems.
3. Higher pressure forces caused the contact angle to increase, generally reducing the water-wetness of rock surfaces and moving them towards intermediate-wet conditions.
4. The contact angle responded differently to temperature fluctuations across various rock types and testing methodologies. Experimental results showed minimal changes to contact angle measurements, though the specific causes behind these variations depended on the particular experimental setup.
5. Quartz minerals exhibited larger contact angles with increased brine salinity, while results for kaolinite clay minerals were more ambiguous. This highlights the importance of specific mineral composition analysis in understanding wettability.
6. The machine learning models successfully identified wettability patterns in silicate, clay, carbonate and basalt rock types which proves their suitability for different geologic contexts related to CO_2 storage.

Recommendations:

1. The experimental dataset needs expansion with additional samples from carbonate and basalt rock types because it will enhance the model performance and its ability to operate in different geological environments.
2. Additional sensitivity analyses should be conducted to demonstrate better understanding of contact angle behaviors associated with pressure factors and temperature variations and salinity conditions and rock types especially when multiple minerals exist in the system.
3. Design a convenient interface or software application with integrated best-performing models to enable both researchers and industry professionals to conduct wettability predictions during different CO_2 storage operations.
4. The model requires validation using practical field-scale CO_2 storage site data in order to determine its operational effectiveness so additional adjustments can be made.
5. Model training techniques which enable the reuse of knowledge acquired from studying one rock

type to analyze similar geological materials through small samples need further investigation.

6. Further research should investigate the impact of different ion types on rock surface wettability in CO2-brine-rock systems. Due to the limited publicly available data on this specific aspect, laboratory experiments should be conducted to study how various ionic species in brine affect wettability, which could enhance the model's predictive capabilities and provide a more comprehensive understanding of wettability behavior in diverse geological settings.

6. REFERENCES

Al-Khdheeawi, E. A., Vialle, S., Barifcani, A., Sarmadivaleh, M., & Iglauer, S. (2017). Impact of reservoir wettability and heterogeneity on CO2-plume migration and trapping capacity. *International Journal of Greenhouse Gas Control*, *58*, 142-158.

Al-Menhali, A. S., & Krevor, S. (2016). Capillary trapping of CO2 in oil reservoirs: Observations in a mixed-wet carbonate rock. *Environmental science & technology*, *50*(5), 2727-2734.

Al-Yaseri, A. Z., Roshan, H., Zhang, Y., Rahman, T., Lebedev, M., Barifcani, A., & Iglauer, S. (2017). Effect of the temperature on CO2/brine/dolomite wettability: Hydrophilic versus hydrophobic surfaces. *Energy & Fuels*, *31*(6), 6329-6333.

Alashari, M., El-Rayes, K., & Helaly, H. (2023). Machine-learning model for predicting maintenance costs of EPDM roofing systems. *Journal of Performance of Constructed Facilities*, *37*(6), 04023048.

Alnili, F., Al-Yaseri, A., Roshan, H., Rahman, T., Verall, M., Lebedev, M., ... & Barifcani, A. (2018). Carbon dioxide/brine wettability of porous sandstone versus solid quartz: An experimental and theoretical investigation. *Journal of colloid and interface science*, *524*, 188-194.

Ali, M., Yekeen, N., Alanazi, A., Keshavarz, A., Iglauer, S., Finkbeiner, T., & Hoteit, H. (2023). Saudi Arabian basalt/CO2/brine wettability: Implications for CO2 geo-storage. *Journal of Energy Storage*, *62*, 106921.

Arif, M., Barifcani, A., Lebedev, M., & Iglauer, S. (2017). Impact of solid surface energy on wettability of CO2-brine-Mineral systems as a function of pressure, temperature and salinity. *Energy Procedia*, *114*, 4832-4842.

Arif, M., Jones, F., Barifcani, A., & Iglauer, S. (2016). Electrochemical investigation of the effect of temperature, salinity and salt type on brine/mineral interfacial properties. *International Journal of Greenhouse Gas Control*, *59*, 136-147.

Ali, M., Al-Anssari, S., Arif, M., Barifcani, A., Sarmadivaleh, M., Stalker, L., ... & Iglauer, S. (2019).

Implications for CO2 trapping/storage. *Journal of colloid and interface science, 534*, 88- 94.

Bikkina, P. K. (2011). Contact angle measurements of CO2–water–quartz/calcite systems in the perspective of carbon sequestration. *International Journal of Greenhouse Gas Control*, *5*(5), 1259-1271.

Botto, J., Fuchs, S. J., Fouke, B. W., Clarens, A. F., Freiburg, J. T., Berger, P. M., & Werth, C. J. (2017). Effects of mineral surface properties on supercritical CO2 wettability in a siliciclastic reservoir. *Energy & Fuels, 31*(5), 5275-5285.

Broseta, D., Tonnet, N., & Shah, V. (2012). Are rocks still water-wet in the presence of dense CO2 or H2S?. *Geofluids, 12*(4), 280-294.

Chen, T., & Guestrin, C. (2016, August). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining* (pp. 785-794).

Chiquet, P., Broseta, D., & Thibeau, S. (2007). Wettability alteration of caprock minerals by carbon dioxide. *Geofluids*, *7*(2), 112-122.

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine learning*, *20*, 273-297.

Dai, J., Raza, A., Wang, T., Tian, S., Li, G., & Zhang, T. (2024). Impact of CO2-Brine-Shale Interaction on Wettability Change at Reservoir Temperature and Pressure via AFM Characterization. *Energy & Fuels*, *38*(21), 21042-21051.

Espinoza, D. N., & Santamarina, J. C. (2010). Water-CO2-mineral systems: Interfacial tension, contact angle, and diffusion—Implications to CO2 geological storage. *Water resources research*, *46*(7).

Farokhpoor, R., Bjørkvik, B. J., Lindeberg, E., & Torsæter, O. (2013). Wettability behaviour of CO2 at storage conditions. *International Journal of Greenhouse Gas Control*, *12*, 18-25.

Gholami, R., Raza, A., Andersen, P., Escalona, A., Cardozo, N., Marín, D., & Sarmadivaleh,

M. (2021). Long-term integrity of shaly seals in CO2 geo-sequestration sites: An experimental study. *International Journal of Greenhouse Gas Control*, *109*, 103370

Goodfellow, I., Bengio, Y., Courville, A., & Bengio, Y. (2016). *Deep learning* (Vol. 1, No. 2).

Cambridge: MIT press.

Gong, T., Tang, J., Cheng, Q., Lu, Y., Liu, C., Zhou, J., & Zhao, G. (2024). Effect of CO2/Brine/Shale Interaction on Shale Water Wettability and Spontaneous

Imbibition. *Energy & Fuels*, *38*(21), 21018-21027.

Gershenzon, N. I., Ritzi Jr, R. W., Dominic, D. F., Mehnert, E., & Okwen, R. T. (2017). Capillary trapping of CO2 in heterogeneous reservoirs during the injection period. *International Journal of greenhouse gas control*, *59*, 13-23.

Guan, Q., Alnili, A., & Iglauer, S. (2019). Wettability and interfacial tension alteration of oil– brine–rock systems by nanoparticles: A review. *Journal of Petroleum Science and Engineering*, 181, 106206.

Iglauer, S., Al-Yaseri, A. Z., & Wolff-Boenisch, D. (2020). Basalt-CO2-brine wettability at storage conditions in basaltic formations. *International Journal of Greenhouse Gas Control*, *102*, 103148.

Iglauer, S., Pentland, C. H., & Busch, A. (2015). CO2 wettability of seal and reservoir rocks and the implications for carbon geo-sequestration. *Water Resources Research*, *51*(1), 729- 774.

Iglauer, S. (2017). CO2–water–rock wettability: variability, influencing factors, and implications for CO2 geostorage. *Accounts of chemical research*, *50*(5), 1134-1142.

Jung, J. W., & Wan, J. (2012). Supercritical CO2 and ionic strength effects on wettability of silica surfaces: Equilibrium contact angle measurements. *Energy & Fuels*, *26*(9), 6053- 6059.

Kaveh, N. S., Rudolph, E. S. J., Van Hemert, P., Rossen, W. R., & Wolf, K. H. (2014). Wettability evaluation of a CO2/water/bentheimer sandstone system: contact angle, dissolution, and bubble size. *Energy & Fuels*, *28*(6), 4002-4020.

Kaveh, N. S., Barnhoorn, A., & Wolf, K. H. (2016). Wettability evaluation of silty shale caprocks for CO2 storage. *International Journal of Greenhouse Gas Control*, *49*, 425-435.

Kus, T., Gegechkori, T., & Shariati, A. (2019). Wettability evaluation using the tilting plate method: Procedures and perspectives. *Journal of Colloid and Interface Science*, 546, 124– 134.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Transfer learning for deep learning on graph- structured data. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 31, No. 1).

Latif, S. D., & Ahmed, A. N. (2023). A review of deep learning and machine learning techniques for hydrological inflow forecasting. *Environment, Development and*

Sustainability, *25*(11), 12189-12216.

Liu, Y., Mutailipu, M., Jiang, L., Zhao, J., Song, Y., & Chen, L. (2015). Interfacial tension and contact angle measurements for the evaluation of CO 2-brine two-phase flow characteristics in porous media. *Environmental Progress & Sustainable Energy*, *34*(6), 1756-1762.

Marmur, A., Della Volpe, C., Siboni, S., Amirfazli, A., & Drelich, J. W. (2017). Contact angles and wettability: towards common and accurate terminology. *Surface Innovations*, *5*(1), 3- 8.

Mutailipu, M., Liu, Y., Jiang, L., & Zhang, Y. (2019). Measurement and estimation of CO2– brine interfacial tension and rock wettability under CO2 sub-and super-critical conditions. *Journal of colloid and interface science*, *534*, 605-617.

Roshan, H., Al-Yaseri, A. Z., Sarmadivaleh, M., & Iglauer, S. (2016). On wettability of shale rocks. *Journal of colloid and interface science*, *475*, 104-111.

Ruder, S. (2016). An overview of gradient descent optimization algorithms. *arXiv preprint arXiv:1609.04747*.

Sarmadivaleh, M., Al-Yaseri, A. Z., & Iglauer, S. (2015). Influence of temperature and pressure on quartz–water–CO2 contact angle and CO2–water interfacial tension. *Journal of colloid and interface science*, *441*, 59-64.

Saraji, S., Goual, L., Piri, M., & Plancher, H. (2013). Wettability of supercritical carbon dioxide/water/quartz systems: Simultaneous measurement of contact angle and interfacial tension at reservoir conditions. *Langmuir*, *29*(23), 6856-6866.

Sedaghatinasab, R., Kord, S., Moghadasi, J., & Soleymanzadeh, A. (2021). Relative permeability hysteresis and capillary trapping during CO2 EOR and sequestration. *International Journal of Greenhouse Gas Control*, *106*, 103262.

Tao, Hai, et al. "Development of new computational machine learning models for longitudinal dispersion coefficient determination: case study of natural streams, United States." Environmental Science and Pollution Research 29.24 (2022): 35841-35861.

Tudek, J., Crandall, D., Fuchs, S., Werth, C. J., Valocchi, A. J., Chen, Y., & Goodman, A. (2017). In situ contact angle measurements of liquid CO2, brine, and Mount Simon sandstone core using micro X-ray CT imaging, sessile drop, and Lattice Boltzmann modeling. *Journal of Petroleum Science and Engineering*, *155*, 3-10.

Umar, B. A., Gholami, R., Raza, A., Downey, W. S., Sarmadivaleh, M., Shah, A. A., & Nayak,

P. (2020). A Study on the Surface Wettability of Clastic Rocks with Potential Application for CO 2 Storage Sites. *Natural Resources Research*, *29*, 2051-2061.

Yao, Xiaotong, Xiaoli Fu, and Chaofei Zong. "Short-term load forecasting method based on feature preference strategy and LightGBM-XGboost." *IEEE Access* 10 (2022): 75257-75268.

Wang, S., Edwards, I. M., & Clarens, A. F. (2013). Wettability phenomena at the CO2–brine–mineral interface: implications for geologic carbon sequestration. Environmental science & technology, 47(1), 234-241.

Yang, D., Gu, Y., & Tontiwachwuthikul, P. (2008). Wettability determination of the reservoir brine−reservoir rock system with dissolution of CO2 at high pressures and elevated temperatures. *Energy & Fuels*, *22*(1), 504-509.

Yekeen, N., Padmanabhan, E., Abdulelah, H., Irfan, S. A., Okunade, O. A., Khan, J. A., & Negash, B. M. (2021). CO2/brine interfacial tension and rock wettability at reservoir conditions: A critical review of previous studies and case study of black shale from Malaysian formation. *Journal of Petroleum Science and Engineering*, *196*, 107673.

Younis, E. M., Mohsen, S., Houssein, E. H., & Ibrahim, O. A. S. (2024). Machine learning for human emotion recognition: a comprehensive review. *Neural Computing and Applications*, *36*(16), 8901-8947.

Yuan, Y., & Lee, T. R. (2013). Contact angle and wetting properties. In Surface science techniques (pp. 3-34). Berlin, Heidelberg: Springer Berlin Heidelber