APPLICATIONS OF MACHINE LEARNING ALGORITHMS IN INTEGRATED CARBON CAPTURE-MINERALIZATION (ICCM) PROCESSES

A Thesis

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Master of Science

by

Jinglin Li

December 2022

ABSTRACT

Regulation of CO_2 emissions is imperative to prevent environmental catastrophes brought on by climate change. Technologies for carbon capture and storage are being developed extensively as environmentally friendly solutions to the issue. Traditional methods divide carbon capture and storage into two phases, leading to inefficient operation and excessive energy usage. To expedite the reactions and enhance productivity, an integrated carbon capture−mineralization process was proposed at a later stage. The classic carbon-water-mineralization system is improved by this combined carbon capture and mineralization technology by removing the constraint of limited carbon dioxide solubility. However, aside from the advantages, the integrated carbon capture−mineralization processes suffer from some limitations, such as instability of amine-based solvents or gel-like formation during the reactions. In addition, competing phenomena exist among amines, such as tradeoffs between reaction rates and viscocities, so it is insufficient to examine a single property of the amine for improving mineralization performance. Thus, for researchers and engineers to increase mineralization efficiency, simulations were used to help optimize process schemes or selecting the best reactants by comprehensive analysis on physical, chemical, and reactions properties of amines involved in carbon capture and mineralization processes.

Aspen Plus was used to simulate the two-stage integrated carbon capture−mineralization processes and to identify reinforcing feedback between mineralization and related parameters. In the simulation, CO_2 was first removed from the flue gas stream using amine-based solvents (monoethanolamine diethanolamine, triethylamine, and blended monoethanolamine+Methyl diethanolamine). After this, an alkaline substance (CaO) was used to regenerate the CO_2-loaded

amines through mineralization processes, and the final desorbed CO_2 was transformed into bicarbonates, which were then mixed with calcium ions (Ca2+) to create calcium carbonate (CaCO3).

The Aspen-generated data in the simulation processes were then used for the next part: implementing machine learning algorithms for enhancing CO_2 mineralization level in the integrated absorption−mineralization processes. Machine learning algorithms use statistical calculations to predict outcomes and find optimizations. The objective of this study is to forecast how different amines will perform on carbon mineralization under various reaction conditions by using well-know supervised machine learning algorithms, such as Artificial Neural Networks (ANN), Support Vector Regression Machine (SVM), K-nearest neighbor (KNN), Decision Tree Regression (DTR), Lasso Regression, and Ridge Regression(RR),

Using multiple statistical indicators, our findings show that KNN outperformed other models in terms of Root Mean Square Error (RMSE) (0.027), Mean Absolute Error (MAE) (0.012), Mean Absolute Percentage Error (MAPE) (0.479%), and correlation coefficient (R^2) (0.9929) values. The accurate predictive modeling is of great importance for potential industrial advancement and solving CO_2 emission problems, as it can find optimal reaction conditions. The combined “machine learning-integrated absorption and mineralization” is beneficial for the development of smart carbon capture and storage technologies and the creation of efficient reaction systems, while bridging the gap between artificial intelligence and environmental protection.

BIOGRAPHICAL SKETCH

Jinglin Li was born in Inner Mongolia (China), a region to the northwest of the Gobi desert. The location's natural circumstances caused regular dust storms and had detrimental effects on her daily life, so she knew from a young age that she desired an uncontaminated environment and pure air. When she grew up, she first left her hometown to Beijing for high school. The three years was the time that Beijing encountered the most severe situations of PM 2.5, and many people got respiratory difficulties during that period. That was the moment she understood the intrinsic connection between the environment and human health. Therefore, she had grown intense interest in nature since then, and she chose environmental science as her undergraduate major upon entering UC Davis without hesitation: her ambitions were to make a difference in the years to come.

The years at UC Davis led her into a broader environmental field, and after graduating from UC Davis, she chose to seek further education to improve her knowledge and skills. She is a Master's degree candidate in the School of Civil and Environmental Engineering at Cornell University. Her current research utilizes Machine Learning (ML) techniques to predict mineralization rates in carbon capture-mineralization systems. Her research focuses on the use of artificial intelligence to environmental studies, which could increase the efficiency of future carbon capture and storage technologies.

To expand her horizons and sharpen her decision-making, execution, and communication

skills, she also participated in a number of extracurricular activities. Her field trip to Hopland Research and Extension Center and volunteer experience in the National Natural Reserve taught her how to use natural resources effectively without exhausting them; taking the position as head of the Publicity Department at the Chinese Students and Scholars Association has fostered her abilities in decision making and leadership. In addition, she organized many environmental-related activities, trying to promote the ideology of protection and preservation of the environment to the public.

Now, standing at the cusp of Master graduation, she plans to pursue a Ph.D. degree to achieve her goals better. Apart from senior professors who will lead her further into environmental fields, the more advanced research experience will bring her closer to her research field and career targets in solving environmental problems.

ACKNOWLEDGMENTS

I would express deepest gratitudes to my advisor Prof Greeshma Gadikota. The pandemic had forced me to take online courses and remote research, but Prof Gadikota helped me work out. She devoted her time and arranged online meetings to assist me with my research as needed. I could not have completed my thesis without her patience, guidance, and feedback not only on the overall content but also on each and every detail.

I would like to extend my sincere thanks to my special committee members Prof Damian Helbling in the School of Civil and Environmental Engineering and Prof Buz Barstow in the School of Biological and Environmental Engineering.Their guidance provided me with new perspectives on how to conduct my research. Taking a course taught by professor Helbling let me feel a perfect blend of wisdom and humor, which was one of my favorite studying times in Ithaca; different angles of ideas from professor Barstow guided me to a new standpoint on research structures.

Very special thanks to every member of this research group: Xu Gao, Zihao Li, Sohaib Mohammed, Hassnain Asgar, Peilong Lu, Prince Ochonma, Ruiyi Zhang, Victor Ruan, Tianhe Yin, Yang Jia, and Minkyeong Kim. I am very grateful for their companionship in the class and the laboratory, and their expertise and advice regarding the execution of experiments have expedited my preparations. My most memorable experiences in Ithaca were with them.

Many thanks to the staff in CCMR. When I was looking for equipment to analyze my samples, they assisted me in connecting with facility personnel. In addition to facilitating the

advancement of my research, their comprehensive training safeguarded me from potential safety issues.

Last but certainly not the least, I am deeply indebted to my family. Their spiritual accompaniments, care, and encouragements flourish my growth. It is impossible for me to experience these novel things and cultures without their support. I love you, mom, dad, and my adorable little-sister.

TABLE OF CONTENT

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

CCS
: Carbon capture and storage

GHG
: Greenhouse Gas

MEA
: Monoethanolamine

DEA
: Diethanolamine

TEA
: Triethylamine

MDEA
: N-methyldiethanolamine

AMP
: 2-amino-2-methyl-1-propanol

NaGly
: Sodium glycinate

PZ
: Piperazine

EDA
: EthylenedICCMine

ICCM
: Integrated absorption and mineralization

DBU
: 1,8-diazabicyclo[5.4.0]undec-7-ene

EDDA
: EthylenedICCMine diacetate

CFA
: Coal fly ash

SFGD
: Dry flue gas desulfurization

UNIFAC
: Universal quasichemical Functional group Activity Coefficients

UNIQUAC
: Universal Quasi-chemical Theory

NRTL
: Nonrandom Two Liquid Theory

ENRTL
: Electrolyte Nonrandom Two Liquid Theory

RK
: Redlich-Kwong

ANN

Artificial neuron network

RSM
: Response surface model

GPR
: Gaussian process regression

SVM
: Support vector machine

DTR
: Decision tree regression

KNN
: K-nearest neighbor

RR
: Ridge Regression

FG
: Flue gas

ML
: Machine learning

RMSE
: Root mean square error

MAE
: Mean absolute error

MAPE
: Mean absolute percentage error

MBE
: Mean bias error

KF
: Kernel function

Minsplit
: Minimal number of observations in a split node

Cp
: Complexity parameter

MLP
: Multi-Layer Perceptron

MSE
: Mean square error

CHAPTER 1

INTRODUCTION

Greenhouse Gas (GHG) are dominant drivers of climate change, and carbon dioxide (CO_2) accounts for 78% of the GHG emissions (Jiang, 2018; Edenhofer, 2014). In order to limit global warming temperature within the 2°C borders and prevent climate-change induced disasters, such as flooding, drought, and air pollution (Tan, 2022), CO_2 emissions should be monitored to zero or even negative in the future. Many countries have joined the group of combating passing the 2°C thresholds, with the UK government being the first to legislate this, followed by Japan, South Korea, and the United States. A total of more than 120 nations have declared their net-zero emissions around the year 2050 (Kang et al., 2021).

Carbon capture and storage (CCS) is one of the most promising and scalable technologies for satisfying the carbon-neutral targets, which can contribute to 94 Gt of cumulative CO_2 reduction by 2050 (Wei et al., 2021). This chapter presented a general overview of carbon capture and storage strategies to better identify the technology gaps and proposed an integrated artificial intelligence and carbon capture-mineralization technique as a potential breakthrough.

1.1 Carbon Capture

Using patent data to illustrate the evolutionary paths of CCS, Kang et al. (2021) determined that CO_2 capture remains the foundation of CCS technology. Carbon capture includes pre-, post-, and oxy-fuel combustion technologies (Wilberforce et al., 2021). Pre-combustion capture involves the production of syngas (H_2 and CO_2) from fuel reforming, whereas

post-combustion capture involves the separation of CO_2 from industrial flue gas (Wilberforce et al., 2021). During oxy-fuel combustion, fuels are burned in an oxygen-purified environment, so the final product will be a gas consisting mainly of CO_2 and water, which can be directly used for sequestration without stripping CO_2 from the gas stream (Buhre et al., 2005). Although the pre-combustion technology is more efficient than post-combustion, it is also more expensive. Furthermore, the requirements for pure oxygen, an energy-intensive and costly process in the air separation unit restrict the use of oxy-fuel combustion routes (Osman et al., 2020). Consequently, post-combustion technology is the most prevalent and developed CCS technology at present (Osman et al., 2020).

Amine-based absorption is the only commercialized method of post-combustion capture with high absorption capacities and good selectivity toward acidic gases (Osman, 2021). Common types of amine solvents include primary, secondary, and tertiary alkaline amines, such as monoethanolamine (MEA), diethanolamine (DEA), N-methyldiethanolamine (MDEA), 2-amino-2-methyl-1-propanol (AMP), Na-glycinate (NaGly) etc. Numerous studies have evaluated the effectiveness of amines in carbon capture. For example, Tan et al (2015). investigated the performance of 20 wt% MEA solution under different pressures. Results showed that approximately 76% of CO_2 was removed when the pressure was low (0.1 MPa), and the removal efficiency increased to 95% with higher pressure (1 MPa). Comparing the working capacity of 10% (w/w) MEA, MDEA, DEA, piperazine (PZ), and ethylenedICCMine (EDA) on CO_2 capture with a CO_2 flow rate of 20 ml/min, Gomes et al. (2014) noted that DEA had the highest working efficiency with a rich loading of 0.492 mol CO_2/mol amine. PZ and MEA had a similar capacity of around 0.409 mol CO_2/mol amine, which could be substituted for one

another, whereas EDA had the lowest loading performance at 0.321mol CO2/mol amine (Gomes et al., 2014).

However, none of the single amines could satisfy requirements on both physical and chemical properties, so recent research has started focusing on using blended amines (Ji et al., 2018). Blended amines provide additional molecules by withdrawing protons from the amino group, increasing molecular interactions and diminishing the activation energy in the reactions, which increases CO_2 absorptions (Kim et al., 2016). Wang et al. (2021a) proposed DMCA-MCA solvents, and the CO_2 loading could reach 0.875–0.985 mol CO_2/mol amine with a flow rate of 120 mL/min at 303 K for 100 mins. Except for the improvements in CO_2 removal, the blended DMCA-MCA system could reduce regeneration energy to only 2.20 GJ/t CO_2, which was 44.9% lower than using 5 M MEA alone (Wang et al., 2021a). However, not all combinations of blended amines improve the efficiency of CO_2 capture. Using MEA as the base solution with additions of PZ, AEEA, DETA, AMP, and DEA separately, Kim et al. (2016) found that compared with a single MEA solution, the groups of MEA + PZ, MEA + AEEA, and MEA + DETA could improve the CO_2 removal, but adding AMP or DEA to the MEA solution will decrease the performance. Therefore, material selection is of the utmost importance for the improvement of CO2 capture that will be compressed, transported, and stored in later stages.

1.2 Carbon Storage

Sedimentary formations and carbon mineralization are typically selected for long-term CO_2 storage, with carbon mineralization having the potential to sequester up to 60,000,000 $GtCO_2$ globally if the resource is commercially viable and fully carbonated (Kelemen et al.,

2019). The reactions involve turning CO_2 into solid carbonates with alkali or alkaline earth metals (Gyudae, 2022). There are three primary types of mineralization: ex-situ, in-situ, and surficial (Kelemen et al., 2019). Ex-situ mineralization stands for the engineering process in which CO_2 is converted to carbonates, and the in-situ technique injects CO_2 and allows the circulation of the gas in the calcium- and magnesium-bearing minerals or rocks (Gadikota, 2016). Surficial mineralization reacts to alkalinity sources at the surface with either dilute or concentrated CO_2 (Kelemen et al., 2019).

Numerous studies have examined the influence of various elements on mineralization efficiency. Dlugogorski and Balucan (2014), for example, concluded that structural disorder is the primary factor controlling carbonation or mineralization in thermally treated serpentine. Therefore, it is essential to reduce the partial pressure of water vapor to prevent recrystallizations of disordered phases. The other direct factor influencing the efficiency of carbon mineralization is the concentration of released Ca^{2+}, Mg^{2+} , Al^{3+} from mafic minerals or rocks, which significantly depends on the PH (Snæbjörnsdóttir et al., 2020). Plagioclase, volcanic glasses, and other Al-based minerals dissolved least at neutral PH, but their dissolution rates increased with higher PH values. On the opposite side, the dissolution rate of olivine, as a magnesium iron silicate, performed well under acidic conditions and decreased with increasing PH (Snæbjörnsdóttir et al., 2020). Hänchen et al. (2008) examined the effects of temperature (25◦C - 120◦C) and pressure (1 - 100 bars) on the precipitation of Mg-bearing systems of mineralization. At low temperature and pressure (at 25 ◦C and P_{CO2}=1 bar), solutions were undersaturated, and no precipitation was formed during 16 hours. However, magnesite formed when temperature and pressure increased to 120 ◦C and 100 bar (Hänchen et al., 2008).

Apart from the physical limitations, mineralization efficiency is restricted by the dissolution kinetics of minerals and CO_2 dissolution in a CO_2–water–solid system (Cain et al., 2021 & Sanna et al., 2014). Therefore, finding a technique that can remove the barrier of limited dissolution of CO_2 in the conventional carbon-water-mineralization system could further enhance the efficiency of CCS.

1.3 Integrated carbon capture and mineralization

Integrated carbon capture and mineralization (ICCM) techniques have been proposed to improve the efficiency of carbon mineralization and encourage the advancement of CCS technologies. ICCM is a spontaneously chemical regeneration process that operates on the same principle as mineralization, but is able to remove CO_2-dissolving barriers, resulting in stable precipitation as the end product (Kang et al., 2018). In the traditional CO_2 capture and storage processes, CO_2 is firstly absorbed by amine-scrubbing processes with an initial temperature of 40-80◦C, and then CO_2 regenerates the amines through thermal-swing desorption with an elevated temperature of 100-140◦C (Arti et al. 2016 & Ji et al., 2018). This thermal-related techniques require lots of energy. Besides, the thermal-desorbed CO_2 needs extra energy for storage and transportation. Using ICCM is a great way to reduce energy penalty and improve the CO_2 capture efficiency in the process by eliminating the heat duty and additional post-treatment steps involved in the CO_2 capture process.

In the initial development stage of the process, highly abundant natural silicates, such as serpentine, olivine and wollastonite, were used (Ji, et al. 2018). In the latter stage, to better solve environmental problems caused by solid waste, alkaline industrial wastes, including fly ash,

carbide slag, and steel slag, are also involved in CO_2 mineralization (Ji et al., 2018). For instance, Yu et al. (2019) used CaO-containing coal fly ash (CFA) as the feedstock and 2M AMP as the solvents. The ICCM experiment was performed at room temperature for 90 mins, and chemically regenerated AMP solutions could be recycled for over 20 loops with only 10.2% loss. In the research of Yan et al. (2021), semi-dry flue gas desulfurization (SFGD) slag, which contains a large amount of $Ca(OH)_2$, was used as the alkaline source for ICCM, resulting in a desorption rate of 68% and CO_2 fixation efficiency of 87.6% under optimized reaction conditions (Yan et al., 2021).

Numerous studies have demonstrated that ICCM requires lower operating temperatures and less energy while mineralizing at a rapid rate. For instance, $CaCl_2$ acted as a calcium source for CO_2 desorption via chemical-based amine regeneration, resulting in 97.3% desorption efficiency at a temperature as low as 40◦C, which was significantly higher than the thermally-generated desorption rate (Kang et al., 2018). In order to investigate the interactions between organic and inorganic additional ions, Chuajiw et al. (2013) used hexamethylene dICCMine, amine salt (hexamethylenedICCMine-adipic acid salt), and amide ("-caprolactam), monitoring the change of Ca^{2+} and percentage of CO_2 in exhaust gas during the reaction processes. The results confirmed that the reaction rate was improved by adding either base amines or neutral amide ("-caprolactam) or amine salt into calcium hydroxide suspension to form precipitation through mineralization (Chuajiw et al., 2013). Then, in 2016, Arti et al. evaluated the performance of MEA, DEA, MDEA, and AMP in ICCM at a concentration of 1.6M per amine solution. The authors demonstrated that this ICCM technique had 1.3- 3 times higher working capacityies than the conventional thermal amine-scrubbing process, and among all

solvents, MEA outperformed other amine solutions in enhancing working capacity and regeneration efficiency (Arti et al., 2016).

However, the ICCM has limitations, the most important of which is the stability of amine-based solvents. (Wang et al., 2019). In addition, highly viscous and gel-like fluids may form during the process at high concentrations of solvents, such as AMP and DBU (Liu et al., 2021). Thus, to make the technique scalable, future researchers and engineers must maximize the working efficiency at a fixed cost or select the optimal solvents and mineral types for reactions. Machine learning is an algorithm that can be applied to these types of problems in order to predict outcomes and discover optimizations.

1.4 Machine Learning (ML) Algorithms

Machine learning is categorized into supervised, unsupervised, and reinforcement learning methods (Ozba et al., 2019). In supervised learning, algorithms are trained using a labeled dataset, meaning that each data has known outcomes. Unsupervised learning is the inverse of supervised learning: there are no labels or answers, and reinforcement learning is commonly used in decision-making processes. Predicting the performance of a particular system has frequently relied on supervised learning, which can be further categorized into classification models and regression models. For example, Jang (2019) used KNN to predict real-time traveling time in Korea with a prediction error of 3.7%; lasso and ridge regression were compared in the study of Ciftsüren and Akkol (2018) for predictions of internal egg quality characteristics; to optimize variables influencing removal rates of NOx in the flue gas, Shin et al.

(2019) applied ANN models and concluded that the concentration of O_2 was the most critical factor for NOx reduction (Shin et al., 2019).

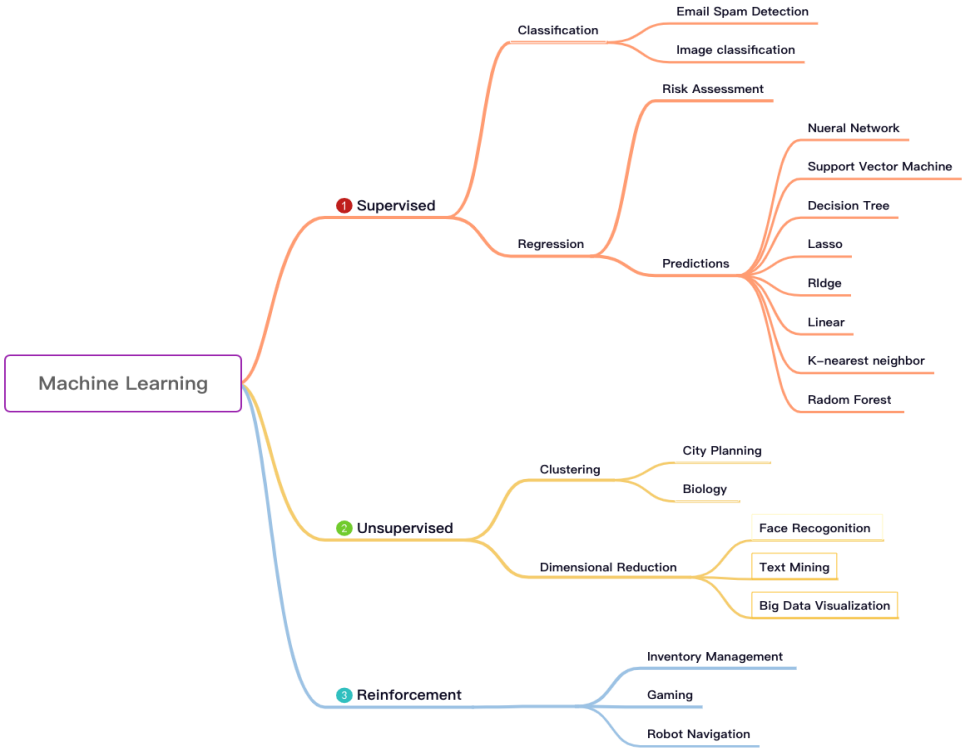


Figure 1.1 Categories of machine learning algorithms

ML techniques have gained popularity not only in the fields of chemistry and transportation, but also in environmental studies, where their ability to solve nonlinear relationships enables the resolution of complex environmental problems (Zhong et al., 2021). Zhong et al. (2021) summarized the current applications of ML in Environmental Science and Engineering into four categories: forecasting, determining the significance of a feature, spotting abnormalities, and discovering new materials or compounds, and these applications involve

water, air, soil, etc. Liu et al. (2020) incorporated anomaly detection techniques based on machine learning into the vertical plant wall system to achieve interior climate control on CO_2 and temperature, targeting on human health protection. In solid waste management, Lu et al. (2022) developed a new mutant FAST-PETase enzyme for PET depolymerization with the aids of ML. The FAST-PETase can almost entirely break down the untreated post-consumer PET in a week. In order to provide evidence-based information on how stressors and biological drivers interact and influence ecological processes and microbial biomass in hyporheic zone (HZ), Mori et al. (2019) employed machine learning (ML) technologies and identify the important features. Results highlighted the significance of temperature and nutrient inputs from anthropogenic activities on structures and processes of HZ. The authors also summarized that climate change will have more pronounced effect on ecosystem functioning .

In order to manage climated change, Sipocz et al. (2011) combined ML algorithms on technologies of CO_2 capture and targest on finding the optimum operation for the example plant. They simulated over 4000 data through CO_2SIM software for an amine-based post-combustion CO_2 capture process, which constructed a highly accurate ANN model with only 0.17% prediction error. Based on the same operational principle, the ANN and response surface model (RSM) model in the research of Nuchitprasittichai and Cremaschi (2013) predicted that diglycolamine-based amine solvents could minimize CO_2 capture cost while maintaining the same capture efficiency. Other than using single or pairwise models, a number of researchers examined the performance of a variety of models for predicting CO2 capture. For instance, fine tree, ANN, Matern Gaussian process regression (GPR), rational quadratic GPR, and squared exponential GPR models were developed in the study of Shalaby et al. (2019), and they used flue

gas flow rates, temperatures, pressures, and the reboiler and condenser duties as inputs. The results demonstrated that ANN contributed to the most accurate predictions. Deng et al. (2020) conducted an extension of the methods. Among four commonly used prediction models, Back Propagation Neural Network, Decision Tree, Random Forest (RF), and SVM, RF had the best prediction results with $R^2 = 0.982$. Considering the carbon storage processes, Vo Thanh and Lee (2022) applied Gaussian Process Regression (GPR), SVM, and RF to predict CO_2 trapping in saline aquifers. Based on the analysis of Root Mean Square Error (RMSE) and R^2, GPR outperformed other models on estimating CO_2 trapping efficiency.

The limitations of these studies are that none of the research attempted to to anticipate capture and sequestration efficiency simultaneously. Instead, they continued to view carbon capture and storage as separate processes. As some research has demonstrated, however, that ICCM processes are more energy-efficient and cost-effective than carbon capture and sequestration alone, it is essential to propose the integrated ML-ICCM methods for better CO_2 control.

1.5 Research Objectives

This study proposed to use an integrated ML-ICCM technique to predict carbon mineralizations under various reaction conditions in amine-based systems, which involves choosing the best amine solvent and finding the most cost-effective process scheme. The first part of the study was about simulations of the ICCM process by using Aspen Plus, and reinforcing feedbacks between mineralization and related parameters, such as flow rates, temperature, and concentrations, were identified. Furthermore, this part helped determining

competing phenomena among amines by comprehensive analysis on physical, chemical, and reactions properties of amines involved in carbon capture and mineralization processes. Part Two used the Aspen-generated data, which is the performance of different amines under various reaction conditions, to build machine learning models, including Artificial Neural Networks (ANN), Support Vector Regression Machine (SVM), K-nearest neighbor (KNN), Decision Tree Regression (DTR), Lasso Regression, and Ridge Regression(RR). Part Three used different statistical indicators to test the prediction performance and reliabilities of the models, and the most accurate model was chosen for future implementations in ICCM.

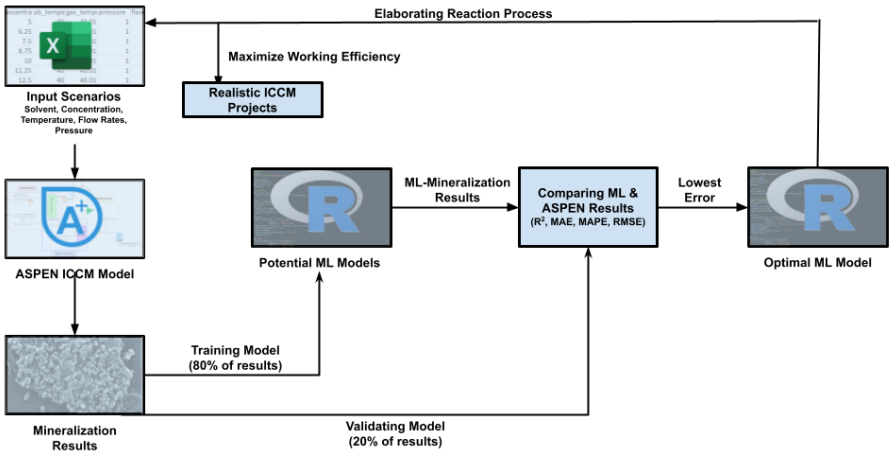


Figure 1.2 Methodology diagram for the proposed integrated ML-ICCM technique

CHAPTER 2

ASPEN SIMULATED INTEGRATED ADSORPTION AND MINERALIZATION (ICCM)

2.1 Aspen Plus

Aspen plus is an engineering simulation software with a comprehensive chemical database, and a general simulation flowsheet in the software consists of core reactors and various operational units (Al-Malah, 2016). It is a potent tool for implementing dynamic models, solid handling, energy saving, sensitivity analysis, and unit operation calculation-based tasks (Al-Malah, 2016). Therefore, this study chose Aspen software to simulate the ICCM process to mimic realistic experimental processes, and the effects of various reaction conditions on mineralization rates were investigated further.

2.2 Simulation setup

The simulated ICCM process is as follows:

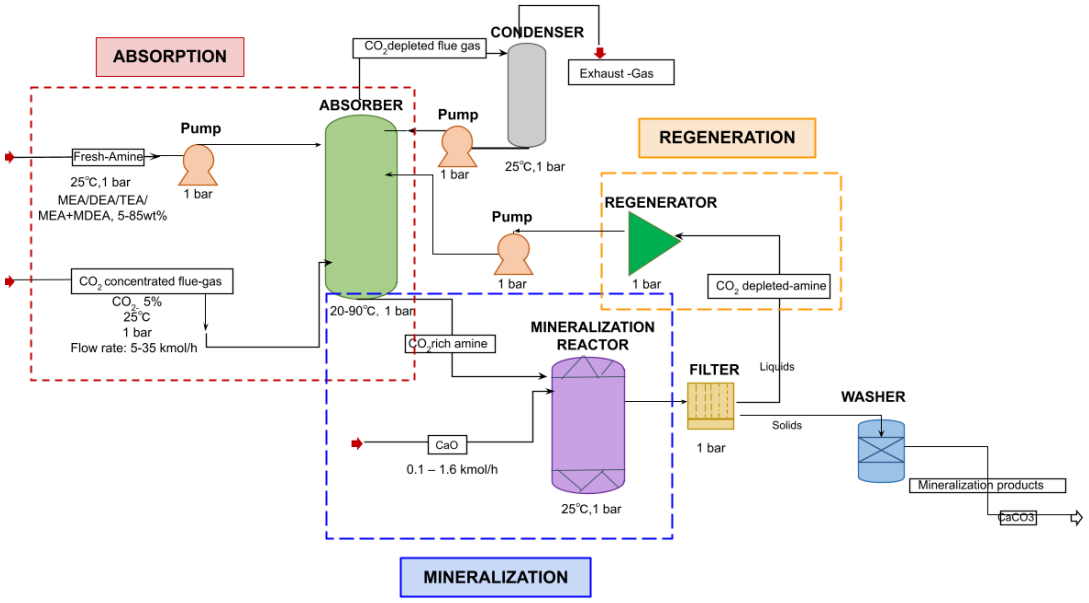


Figure 2.1 Flowsheet of an Aspen simulated ICCM process

The input gas was desulfurized flue gas (DSFG) composed of CO_2 (5%), O_2 (12%), N_2 (76%), and H_2O (7%) (Arachchige & Melaaen, 2012). The DSFG flow rate was set at 25 kmol/h under the pressure of 1 bar and temperature at 25°C, and then the flue gas was blown into the ABSORBER that contains either fresh amine or cycled CO_2-lean solvents. Solvents used in the simulation included primary (MEA), secondary (DEA), tertiary (TEA), and blended (MEA+MDEA) amines. The amine solvents absorbed the FG-contained CO_2, and the CO_2-depleted FG went through a CONDENSER to recuperate the residual amines in the gas phase and pumped the amine into the ABSORBER for re-utilization, and the final tail gas vented out as the exhaust gas. The CO_2-rich amine solution was transported into the MINERALIZATION REACTOR, where it reacted with alkaline solids (CaO) to produce CaCO3 via the mineralization process. CaO was selected as the mineral because it is the major

component of coal fly ash and other solid-waste ashes. Mineralization reactions occurred at 1 bar with a temperature at 25°C, and the mineralized precipitation moved into the FILTER. The FILTER separated $CaCO_3$ solids from liquids. The separated CO_2 - lean solution was transferred to the REGENERATOR and thoroughly mixed before being recycled back into the ABSORBER. Simultaneously, the separated solids were transported to a SOLIDS WASHER and purified $CaCO_3$ was produced.

Table 2.1 outlines the configurations of each block, and ENRTL was selected from a variety of activity coefficient models, including Wilson, Van Laar, Universal quasichemical Functional group Activity Coefficients (UNIFAC), Universal Quasi-chemical Theory (UNIQUAC), Flory Huggins, Nonrandom Two Liquid Theory (NRTL), Electrolyte NRTL (ENRTL), and Scatchard Hilbert models (Carlson, 1996). Based on the illustration in Carlson's (1996) study regarding the selection of physical property methods, the Wilson model is suitable for polar but non-electrolyte simulations. According to Unlu and Hilmioglu (2020), UNIFAC models are frequently used for non-ideal systems at low pressures, and based on the formation of by-products, they utilized the UNIFAC model in Aspen Plus to examine the precess of steam reforming of the glycerol (Unlu and Hilmioglu, 2020). Flory Huggins is a common model for determining the characteristics and physical properties of polymer-solvent systems (Ovejero et al., 2009), whereas Van Laar can be used to predict vapor-liquid behavior in simple and non-polar systems. UNIQUAC models were not originally designed for electrolytes, but with the modifications, numerous researchers have begun to employ the model in CO_2 capture processes. Darde et al. (2012), for instance, stimulated the CO_2–NH_3–H_2O system in Aspen Plus by employing the extended UNIQUAC thermodynamic model. Compared to UNIQUAC, NRTL,

and E-NRTL are more suitable for electrolyte and polar systems (Carlson, 1996). In addition, the study of Simoni et al. (2007) compared prediction performance of UNIQUAC, NRTL and E-NRTL in systems that contain ionic liquids. Results demonstrated that the ENRTL model outperformed the other two models with small sentivity to the length of the alkyl chains involved in the reactions (Simoni et al., 2007).

In this study, Electrolyte NRTL (ENRTL) was initially chosen. ENRTL-RK, which stands for Electrolyte NRTL with Redlich-Kwong physical property, was used as the final thermodynamic model to address the issue of thermodynamic inconsistency (Adams, 2022).

Table 2.1 Setup and configurations of each block in the simulated ICCM process

Blocks	**Model**	**Specification**	**Parameters**
ABSORBER	RadFrac	Calculation type: rate-based Number of stages: 20 Valid phases: Vapor-liquid Convergence: standard Property methods: ENRTL-RK	Reaction temperature: 25°C Pressure: 1 bar Solvents: 30 wt% Flue gas temperature: 25°C Flue gas flow rate: 25 kmol/h
MINERALIZATION REACTOR	RStoic	Flash type: temperature; vapor fraction Valid phases: Vapor-liquid Property methods: ENRTL-RK	Pressure: 1 bar Temperature: 25°C CaO: 1.6 kmol/h
FILTER/ SOLID-LIQUID SEPARATOR	Filter	Model: solids separator Specify phase separation: solid load of liquid outlet Classification characteristic: particle size Property methods: ENRTL-RK	Pressure: 1 bar
SOLID WASHER	Sep	Outlet stream: $CaCO_3$ Valid phase: vapor-liquids	Pressure: 1 bar
CONDENSER	Flash 2	Flash type: temperature-pressure	Temperature: 25°C pressure: 1 bar
REGENERATOR	Mixer	Valid phase: vapor-liquids	Pressure: 1 bar
PUMP	Pump		Discharge pressure: 1bar

2.3 Reaction mechanisms in the Aspen simulated ICCM process

The types of amines used in this study is alkanolamine which is composed of amino and alcohol groups (Muchan et al., 2017). Alkaoamine can be further categorized into primary, secondary, and tertiary amines, depending on the number of carbons attached to the amino nitrogen. Primary and secondary amines involve carbamate formation when reacting with CO_2 (Liu et al., 2021). On the other hand, tertiary amines belong to the direct bicarbonate-formation groups (Ji et al., 2018). For primary amines, the maximum absorption capacity is 0.5 mol CO_2/mol amine, which has a relatively fast reaction rate (Ji et al., 2018). Secondary and tertiary amines have higher loadings of up to 1 mol CO_2/mol amine (Chen et al., 2021). However, the tertiary amines accept protons at a slower rate, compared to both primary and secondary amine, because they have steric hindrance which converts carbamates to HCO_3^-, but this process will release more amines and can capture more CO_2 (Ji et al., 2018).

The reaction mechanisms of the amines are shown in Table 2.2, which occurred after water dissociation and CO_2 hydration (Liu et al., 2021):

$$H_2O_{(l)} \Leftrightarrow H^+_{(aq)} + OH^-_{(aq)} \quad \text{Rx 2.1}$$

$$CO_{2\,(g)} + H_2O_{(l)} \Leftrightarrow H_2CO_3^*{}_{(aq)} \quad \text{Rx 2.2}$$

$$H_2CO_3^*{}_{(aq)} \Leftrightarrow HCO_3^-{}_{(aq)} + H^+_{(aq)} \quad \text{Rx 2.3}$$

$$HCO^{3-}{}_{(aq)} \Leftrightarrow CO_3^{2-}{}_{(aq)} + H^+_{(aq)} \quad \text{Rx 2.4}$$

In the table 2.2, primary and secondary amines were shown in the form of R1R2NH, where R1=H and R2=C_2H_4OH for MEA, and R1=R2=C_2H_4OH for DEA. Tertiary amines (TEA and MDEA) were represented as R1R2R3N, where R1=R2=C_2H4OH and R3=CH_3 (Maakoul et

al. 2021). In the case of mixed amines (MEA/MDEA), the overall reaction mechanisms are the combination reactions of each individual amine (Conway et al., 2015).

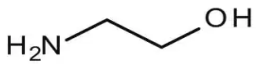
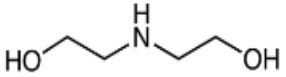
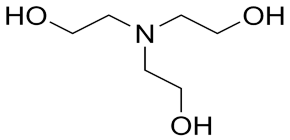
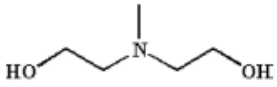
Then the last step in the ICCM process is forming precipitates from Ca-bering solids which were CaO in our study, through carbonation reactions (Ji et al., 2018):

$$CaO_{(s)} + H_2O \rightarrow Ca^{+} + 2OH^{-} \quad \text{Rx 2.5}$$

$$HCO_3^{-} + OH^{-} \leftrightarrow CO^{3-} + H_2O \quad \text{Rx 2.6}$$

$$Ca^{2+} + CO^{-} \leftrightarrow CaCO \downarrow \quad \text{Rx 2.7}$$

Table 2.2 Reaction mechanisms of amine-based solvents in the ICCM simulation process

Name	**Abbreviation**	**Structure**	**Reactions with CO_2** (Muchan et al., 2017; Gao et al., 2017; Osman et al., 2012).
Monoethanolamine	MEA (C_2H_7NO)		1. R1R2NH + CO_2 (aq) ⟷ R1R2N$^+$HCOO- (zwitterion) 2. R1R2NH + R1R2NH$^+$COO$^-$ ⟷ R1R2N$^+$H2 +R1R2NCOO$^-$ (carbamate) 3. R1R2NCOO$^-$ + H_2O ⟷ R1R2NH (free amine) + HCO_3^- (bicarbonate)
Diethanolamine	DEA ($C_4H_{11}NO_2$)		1. R1R2NH + CO_2 (aq) ⟷ R1R2N$^+$HCOO- (zwitterion) 2. R1R2NH + R1R2NH$^+$COO$^-$ ⟷ R1R2N$^+$H2 +R1R2NCOO$^-$ (carbamate) 3. R1R2NCOO$^-$ + H_2O ⟷ R1R2NH (free amine) + HCO_3^- (bicarbonate)
Triethanolamine	TEA ($C_6H_{15}NO_3$)		1. R1R2R3N+ H^+ ⟷ R1R2R3NH$^+$ 2. R1R2R2N + CO_2 + H2O ⟷ R1R2R3NH$^+$ + HCO_3^-
Monoethanolamine/ Methyldiethanolamine	MEA/MDEA MDEA:		1. R1R2NH + CO_2 (aq) ⟷ R1R2N$^+$HCOO$^-$ (zwitterion) 2. R1R2NH + R1R2NH$^+$COO$^-$ ⟷ R1R2N$^+$H2 +R1R2NCOO$^-$ (carbamate) 3. R1R2NCOO$^-$ + H_2O ⟷ R1R2NH (free amine) + HCO3$^-$ (bicarbonate) 4. CO_2 + R1R2R3N + H_2O ↔ R1R2R3NH$^+$ + HCO3$^-$

2.4 Reaction results in the Aspen simulated ICCM process

Table 2.3 Reactions results of various amines in the ICCM process
(Fixed condition: 30wt% amines; 25 kmol/h CO_2; 1.6 kmol/h CaO; 1 bar;25 ℃)

	Parameters	Units	MEA	DEA	TEA	MEA+MDEA
Inputs	CO_2 in FG	kmol/h	1.25	1.25	1.25	1.25
	PH of fresh amine solvents	\	12.29	11.85	11.25	12.13
Outputs	CO_2 in Exhaust Gas	kmol/h	0.41	0.726	0.726	0.33
	PH of CO_2-rich amine solvents	\	10.11	10.36	10.43	9.57
	Absorption efficiency	%	67.2	41.92	41.91	73.6
	Mineralization rates	kmol/h	0.84	0.523	0.521	1.23

The absorption efficiency in Table 2.3 is calculated as follows:

$$\text{Absorption efficiency} = \frac{(CO_2\ in\ FG) - (CO_2\ in\ Exhaust\ Gas)}{CO_2\ in\ Exhaust\ Gas} * 100\% \qquad \text{Eq 2.1}$$

According to the data in Table 2.3, the PH decreased while amines absorbed CO_2. This is because, based on RX 2.2, the dissociation of carbon dioxide forms carbonic acids while combining with water in amine solvents, and the acids will be separated into bicarbonates and hydrogen ions. In addition to the fact that final mineralization rates are dependent on the number of bicarbonates, a low PH value also favors CO_2 desorption and results in higher mineralization

rates (Du, 2017). The results were consistent with the study of Arti et al. (2016) in which the lowest PH value produced the highest mass yield of mineralization products (g of $CaCO_3$/g of $CaCl_2$).

Table 2.4 Physiochemical properties of the amine solvents at ambient conditions

	Units	MEA	DEA	TEA	MEA+MDEA	Reference
Reaction rate	m^3/kmol*s	5939	412	3.060	> 5939	Park et al. (2006) Sema et al. (2013) Critchfield. (1988) Jiang et al. (2017)
Density (ρ)	g/cm^3	1.010	1.034	1.046	1.018	Zhang et al. (2015) Concepción et al. (2017) Karunarathne et al. (2020)
Viscosity (η)	mPa·s	2.540	3.633	3.208	2.780	Zhang et al. (2015) Concepción et al. (2017) Karunarathne et al. (2020)
Corrosion rate	mpy	15.500	5.400	<5.400	6.500	Hasib-ur-Rahman et al. (2012)

Physiochemical properties is the other way to explain the results of Table 2.3. Even though the viscosities of the blended amine are higher than MEA (Table 2.4), the MEA+MDEA solvents contributed to the highest mineralization rates due to the highest reaction rates and relatively small corrosion rates. This is because using a mixture of amine solvents could take

dual benefits of every single amine and enhance capture and loading performance (Aghel et al., 2022). In the combination of MEA and MDEA, MDEA replaces H_2O and can act as a base to release free MEA from $MEAH^+$, and the MEA could be used as CO_2 adsorption solvents again (Zhang et al., 2016). Therefore, compared to using a single MEA which is restricted to the maximum loading of 50%, adding MDEA, which has a CO_2 loading capacity approaching 100%, can theoretically improve CO_2-capture effieciency. Furthermore, as MEA degrades when contact with oxygen (O_2) and impurities, CO_2 absorption capacities tend to decrease in the scrubbing process. In contrast, the study of Lawal et al. (2005) proved that the presence of MDEA in the system of MEA-H_2O-CO_2 protects MEA from degradation and reduces the amount of nonenvironmentally benign by-products. These findings elucidated why the blended amine resulted in the greatest absorption efficiency.

In a single amine-based solution, MEA exhibited the highest carbon capture and mineralization rates compared to DEA and TEA. This could be attributed to the rapid reaction rate of MEA, which is 5939 m^3/kmol*s in the system of MEA-CO_2-H_2O at room temperature and atmospheric pressure (Edali et al., 2009). The higher reaction rates contributed to greater mass transfer rates, deriving higher CO_2 absorption and mineralization rates (Chakma, 1997). On the other hand, DEA, with a reaction rate of 412 m^3/kmol*s, and TEA, with the lowest reaction rate (3.060 m^3/kmol*s) developed relatively lower mineralization rates than MEA (Edali et al., 2009). Due to its extremely slow reaction rates, TEA was predicted to result in much lower mineralization rates. However, the absorption efficiencies of DEA and TEA were comparable, which may be attributed to the relatively lower viscosity of TEA at 30 wt% compared to that of DEA. This mitigates the negative effects caused by its slow reaction times on carbon capture

effectiveness. Table 2.4 also showed that TEA is less corrosive than DEA, and this is because TEA does not produce carbamate ions (DuPart et al., 1993; Mazari et al., 2020). This may also account for the slight difference between DEA and TEA absorption efficiency.

2.5 Univariate analysis

2.5.1 Calcium oxide (CaO) flow rates

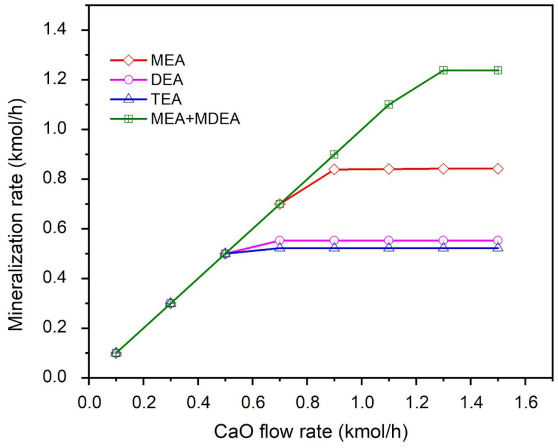


Figure 2.2 Effects of CaO flow rates on final mineralization rates.

Concentrations of solvents = 30 wt% (MEA + MDEA：15 wt% + 15 wt%); flue gas flow rate=25kmol/h;absorption temperature =25℃; flue gas temperature = 25℃.

Figure 2.2 showed that when CaO flow rates were less than 0.5 kmol/h, all types of solvents exhibited an identically linear-shaped increase in mineralization rates. This was because when flow rates were less than 0.5 kmol/h, CaO was undersaturated for the four types of amines, compared to the amount of CO_2 absorbed by the solvents. As a result, the dissolved Ca^{2+} was

completely consumed by the captured CO_2. The rising trend was attributed to the increased availability of reactants for carbonation processes at high CaO flow rates. As the flow rate approached 0.7 kmol/h, DEA and TEA began to stabilize, resulting in mineralization rates of 0.523 kmol/h and 0.521 kmol/h, respectively. After the CaO flow rate reached 0.9 kmol/h and 1.3 kmol/h, respectively, the mineralization rates of MEA and the blended MEA+MDEA remained constant, resulting in much higher mineralization rates than DEA and TEA. Stabilization stages exhibited excessive Ca^{2+} that was far in excess of what was required and could desorb all CO_2 while regenerating protonated amines. Due to the complete depletion of CO_2 in solvents, any additional increase in CaO flow rates would not increase mineralizations. At that time, CO_2 availability was the determining factor for mineralization rates, and the low mineralization rates of DEA and TEA could be attributed to their inferior CO_2 absorption abilities compared to MEA and blended amine.

2.5.2 Concentrations of amine solvents

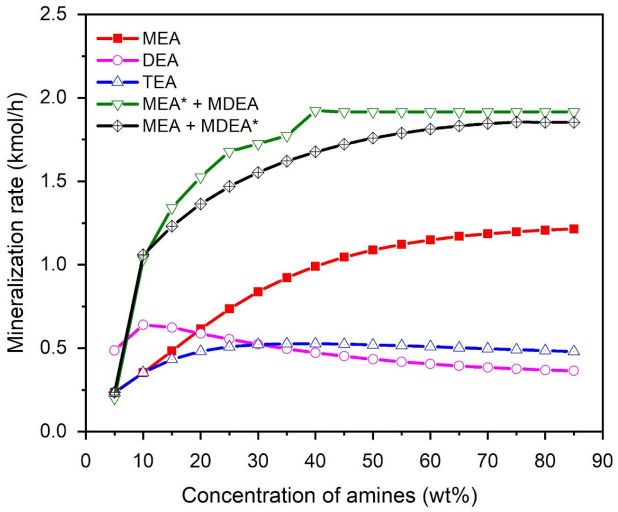


Figure 2.3 Effects of solvents' concentrations on final mineralization rates

MEA* + MDEA: MEA has a fixed value of 5wt% with MDEA changing from 0-80 wt%； MEA + MDEA*: MDEA has a fixed value of 5wt% with MEA changing from 0-80 wt%. Absorption temperature =25℃; flue gas flow rate=25kmol/h; CaO flow rate = 1.6 kmol/h; flue gas temperature = 25℃.

This study examined the effects of amine concentrations ranging from 5 to 85 wt% on the final carbon mineralization behavior, using 1.6 kmol/h CaO and 25 kmol/h DSFG at ambient pressure and temperature. Figure 2.2 showed that CaO at 1.6 kmol/h could support the complete conversion of CO_2. At this time, the only factor limiting the final mineralization rate was the availability of CO_2, and CO_2 availability under fixed DSFG flow rates and absorption temperature would be determined by the concentration of amine solvents. Figure 2.3 showed that the mineralization rates of MEA and blended amines exhibited increasing trends with higher concentrations. This was because the presence of a large number of solvents stimulates not only CO_2 absorption but also the dissolution of CaO. The statement was proved by Ji et al. (2021) who used $Ca(OH)_2$ as minerals to investigate the effects of amines' concentration on final carbon mineralizations. They stated that increasing the concentration enhanced interactions between amino groups and active sites of $Ca(OH)_2$ and could release more Ca^{2+} and OH^-, which resulted in higher mineralization rates (Ji et al., 2021). Compared to $Ca(OH)_2$, CaO contains more active sites and hydroxyl groups (Petitjean et al., 2010). Therefore, the addition of amine solvents could result in more pronounced interactions between amino groups and CaO, thereby enhancing mineralization rates. In contrast, the decreasing trend of TEA and DEA may be attributable to the

formation of gel-like fluids at high concentrations, which inhibits mineralization processes (Liu et al., 2021).

The other important finding in single amine solutions was that when concentrations were less than 30 wt%, both MEA and DEA performed better than TEA. This was because TEA as a tertiary amine does not react directly with CO_2, resulting in low reaction and mineralization rates (Edali et al., 2019). In addition, bulky groups and unstable intermediates of TEA impeded their reactions with CO_2 (Kang et al., 2017). However, after 30 wt% , TEA performed slightly better than that of DEA, resulting in higher mineralization rates. The viscosities of amines explain this phenomenon: viscosities of DEA increase along with the increased concentration of amine solvents, and the viscosities induce difficulties in CO_2 desorption from CaO (Aghel et al., 2020).

Based on the data from Concepción et al. (2017), with a concentration of 10 and 20 wt% at 1 bar and room temperature, the viscosities of DEA are similar to that of TEA. Therefore, viscosities were not taken into account for DEA and TEA at low concentrations, and the fast reaction rates of DEA in the amine-CO_2-H_2O system led to higher mineralization rates under the conditions. However, as concentration increased from 10 wt% to 40 wt%, viscosities of DEA increased from 1.493/mPa·s to 6.2479 /mPa·s (Concepción et al., 2017), and the high viscosities promoted the formation of gel-like fluids and inhibited carbon mineralizations (Liu et al., 2021). Moreover, as the concentration gets higher, the viscosity of DEA increases at a greater rate, esulting in a subsequent decrease in mineralization rate (Concepción et al., 2017). TEA also had increased viscocities under higher concentrations, but the increasing level was insignificant versus DEA. Compared to DEA and TEA, which have viscosities of 3.633 and 3.2075/mPa·s respectively at 30 wt%, MEA has a much lower value of 2.54/mPa·s at the same concentration

(Concepción et al., 2017; Zhang et al., 2015). This is also one of the reasons that MEA performed better than any of the single amines used in this study.

For blended amines, when the concentration of blended amines exceeded 10 wt%, MEA* + MDEA outperformed MEA + MDEA*, which was attributed to the dominance of MEA in the former situation. This result was consistent with Sema et al. (2012) who stated that in the blended MEA-MDEA system, higher ratios of MEA offer more reactive molecules for absorbing CO_2. Moreover, higher MEA concentrations in the blended solutions may also reduce the viscosity of the system at room temperature. For instance, when the blended ratio for MDEA: MEA was 25:5, the viscosity was 3.034 /mPa·s, but it decreased to 2.780 /mPa·s when the ratio was changed to 15:15. (Karunarathne et al., 2020).

2.5.3 Absorption temperature

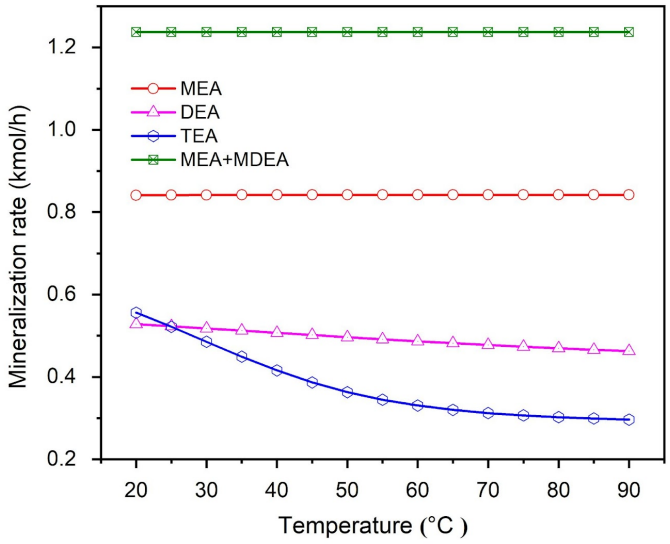


Figure 2.4 Effects of absorption temperatures on final mineralization rates.

Concentrations of solvents = 30 wt% (MEA + MDEA： 15 wt% + 15 wt%); flue gas flow rate=25kmol/h; CaO flow rate = 1.6 kmol/h; flue gas temperature = 25℃.

Figure 2.4 demonstrated that MEA and blended amines maintained nearly constant mineralization rates, which can be attributed to their high reaction rates. Figure 2.5 displayed that both MEA and high MEA blended solvents had a temperature increase at the top stages of the ABSORBER, which means CO_2-capture-induced exothermic reactions occurred at the very beginning of the absorption phase. As the flow rates in the Aspen simulation processes were significantly higher than those typically employed in experimental conditions, rapid reactions of the solvents may reduce detention time to seconds. Consequently, the temperature may not significantly affect the reactions. The result was consistent with Zhang et al. (2020) who used MDEA + MEA as solvents and $Ca(OH)_2$ as minerals in the ICCM process, and they concluded that increasing temperature alone is not effective in improving desorption rates.

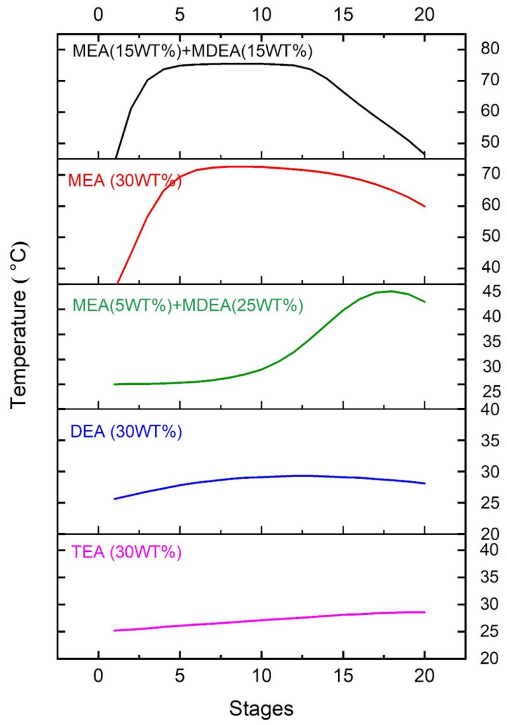


Figure 2.5 Temperature profile for stages in the ABSORBER

However, for DEA, TEA, and high MDEA blended solvents (5wt% MEA + 25 wt% MDEA), the temperature increased in the later stages (Figure 2.5), indicating low reaction rates. Low reaction rates lengthened the time CO_2 and solvents were in contact, making temperature a crucial factor in the absorption process and subsequent mineralization rates. The high absorption temperature decreased the solubility and availability of CO_2 in the absorption and desorption processes, thereby decreasing the final mineralization rates for DEA and TEA. The relatively flat temperature changes for DEA and TEA in Figure 2.5 also supported our statements that less CO_2 was absorbed in the solvents: as an exothermic reaction, the temperature would rise if a large amount of CO_2 was absorbed (Xiao et al., 2018). As a result, DEA and TEA produced fewer

zwitterions and carbamates, reducing their mineralization rates. The other explanation is that relatively low temperature (~30 ℃) in the CO_2-rich DEA and TEA could not provide additional heat for the desorption processes in the MINERALIZATION REACTOR, which caused low mineralization rates.

2.4 Gas flow rate

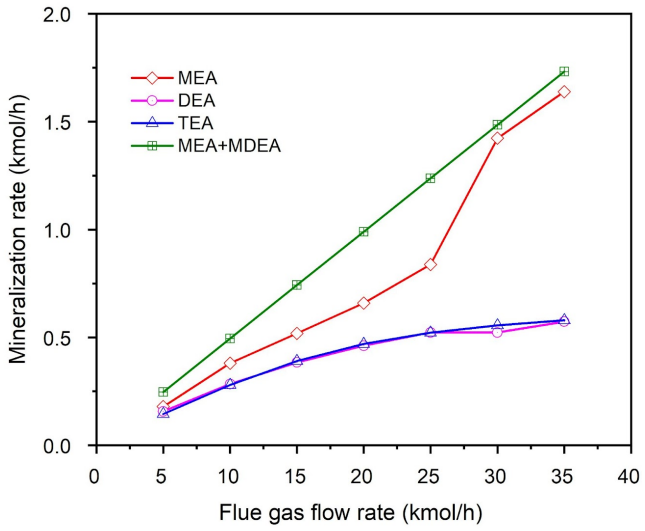


Figure 2.6 Effects of flue-gas flow rates on final mineralization rates.

Concentrations of solvents = 30 wt% (MEA + MDEA： 15 wt% + 15 wt%); CaO flow rate= 1.6 kmol/h; absorption temperature =25℃; flue gas temperature = 25℃.

Figure 2.6 showed that as flue gas flow rates increased, mineralization rates were also improved for all types of amines. With a constant CO_2 concentration of 5% in the DSFG, high flow rates indicate a high molecular concentration of CO_2, which increased the formation of zwitterions and carbamates in CO_2-rich amine solvents, thereby enhancing the formation of

bicarbonate ions and final mineralization rates (Rivera and Gerven, 2020). Moreover, the higher gas flow rate increased bubble velocity and created turbulence in the ABSORBER, and a highly turbulent environment would enhance the effective mass transfer area (Ma et al., 2020). In other words, he relative velocity between the particles and the fluid favored greater mass transfer coefficients in the presence of high gas flow rates (Haugen et al., 2017). The sudden increase in mineralization rates from 0.84 kmol/h to 1.42 kmol/h when flue-gas flow rates increased from 25 to 30 koml/h was attributed to PH swing. When the flue-gas flow rate was 25 kmol/h, the PH value of CO_2-riched MEA was 10.11, while the flow rate further enhanced to 30 kmol/h, the PH value decrease to 9.74. At this time, the PH was in the most sensitive range (8-10) for CO_2 desoprtion in MEA systems, as demosrated by Du (2017), so the mineralization rates increased significantly compared to reactions at other flow rates.

CHAPTER 3

OPTIMIZATION OF THE ICCM PROCESS BY MACHINE LEARNING ALGORITHMS

3. 1 Data Preprocessing

The simulation of various operational conditions in Chapter 2 generated a total of 1290 data. With the gathered data, programming applications (R programming) and machine learning algorithms were used to develop a framework for predicting CO2 mineralization rates. 80% of the generated data were used as training data, while the remaining 20% were testing datasets. Both datasets were preprocessed by scaling the dataset into values between 0-1 with Min-Max normalization. Normalization techniques are frequently employed for the manipulation of data, either to scale down or scale up the range, particularly for forecasting issues (Patro & Sahu, 2015). The normalization strategy prevents the formation of significant variation in prediction by converting independent and dependent variables within the same predetermined boundary. The used normalization equation is as follows:

$$\text{normalized data} = \frac{true\ value - min(x)}{max(x) - min(x)} \qquad \text{Eq 3.1}$$

min and max values are presented in Table 3.1.

Table 3.1 The range of independent variables

Parameters	Units	Min	Max	Increments
Concentration of solvents	wt%	5	85	5
Flue gas flow rates	kmol/h	5	35	5
Pressure	bar	0.1	1	1
Absorption temperature	℃	20	90	5
Gas temperature	℃	20	90	5

3.2 Models

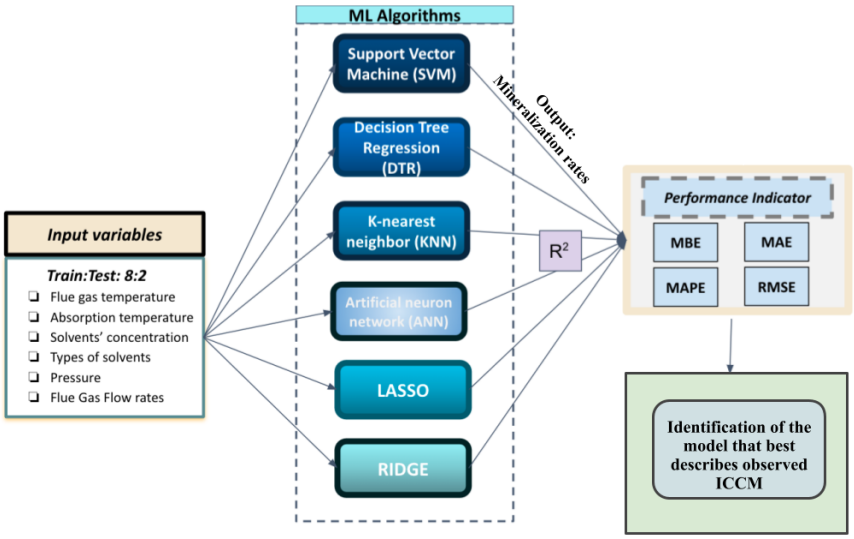


Figure 3.1 The flowchart for machine learning based predictions

As stated previously, Aspen generated data were divided into 8:2 of training : testing with 6 feature variables. Each machine learning model (SVM, DTR, KNN, ANN, Lasso, and Ridge regression) received these data as inputs, and each model predicted different mineralization rates as outputs. R^2 indicates the correlation between predicted and actual values. In addition, Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and Root Mean Square Error (RMSE) were used to evaluate the accuracy of each model's predictions, while Mean Bias Error (MBE) revealed whether the prediction results were negatively or positively biased.

3.3.1 Support Vector Machine (SVM)

SVM was developed by Vapnik who introduced the Vapnik-Chervonenkis theory, and the model is built on the structure risk minimization principle which introduces a trade-off between complexity and errors (Hao et al., 2020; Liu et al., 2019). The basic steps of SVM are to separate input training data into their respective categories and set up boundaries between them. The data used to categorize these classes are called support vectors, and these boundaries are named as decision boundaries or hyperplanes (Ozbas et al., 2019). The main function of the hyperplanes is to distinguish new inputs from test data into one of the separated categories and predict their numeric values. Therefore, the main targets of using SVM should focus on how to draw the decision boundaries.

The area between the support vectors and the decision boundaries is called margins, and the distance of the margin represents the significance of the difference between classes (Cheng and Yang, 2007). A larger marginal distance implies greater difference between classes, which

are easier for models to define the categories of a new input data. Therefore, when SVM is used to seek optimization, the algorithm is trying to find the maximum margin (Saha et al., 2021).

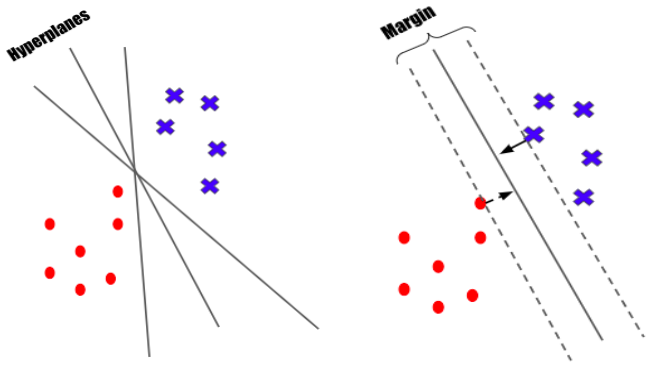


Figure 3.2 Graphical representation for principles of SVM separation

However, in reality, data is usually non-linearly distributed. At this time, the kernel function will be used to model non-linear decision boundaries by mapping low-dimensional vectors into high-dimensional space (Wang et al., 2018). There are four commonly used kernel functions: Linear, Polynomial, Gaussian, and Sigmoid (Wang et al., 2018). Many researchers used Gaussian Kernel Function (KF) in SVM for prediction purposes because it can result in complex boundaries by adjusting only one parameter, making the process easy to handle (Wang et al., 2018). Besides, based on the statement of Wang et al. (2015), Gaussian KF can divide feature space into infinite dimensions, so all data in the feature space can be separated linearly. Therefore, in this study, we chose Gaussian Radical KF. The equation of the KF is as follows (Al-Mejibli et al., 2020):

$$K(X,Y) = \exp(-\gamma \mid X - Y \mid 2) \qquad \text{Eq 3.2}$$

where X and Y act as the vectors in the input space, and γ ($\gamma > 0$) represents the gamma parameter in the kernel function.

The other important parameter in the SVM model is the penalty parameter *C* which is used to keep the balance between margin maximization and training error minimization (Rashidi et al., 2016). The trial and error method was used in this study to choose the best number of γ and C for both the training and test dataset to avoid overfitting, and combinations resulting in the lowest RMSE were chosen (Table 3.2) (Rashidi et al., 2016 & Velasco et al., 2022).

Table 3.2 Tuning parameters for machine learning models

Model	Tuning parameters	Notations	Tuning methods	Tuning range	Optimal value
SVM	spread of the kernel	γ	trial-and-error	{0.01,10}	0.16
	penalty value	*C*	trial-and-error	{1,100}	20
DTR	complexity	cp	trial-and-error	{0.01,0.95}	0.01
	minimum observations	minsplit	trial-and-error	{1, 20}	11
KNN	number of neighbors	k	trial-and-error	{1, 30}	2
ANN	number of hidden layers	size	trial-and-error	{2, 30}	5
Lasso	degree of shrinkage	λ	cross-validation	-	0.001
Ridge	degree of shrinkage	λ	cross-validation	-	0.020

3.3.2 Decision Tree Regression (DTR)

DTR was invented in 1960's, and it is built by root nodes, internal nodes, leaf nodes, and branches (Song and Lu, 2015). In DTR, root nodes represent the entire sample or population. The dataset is originally divided into mutually exclusive groups by internal nodes, and each group will have a specific outcome, either as a value or a subset (Song and Lu, 2015). Nodes of a value or subset that do not further split are called leaf nodes or terminal nodes (Kuhn et al., 2013). If the jth feature variable x^{(j)} and its value s in the training set are used as the splitting node, two regions will form:

$$R_1(j, s) = \{x|x^{(j)} \leq s\} \quad \text{Eq 3.3}$$

$$R_2(j, s) = \{x|x^{(j)} > s\} \quad \text{Eq 3.4}$$

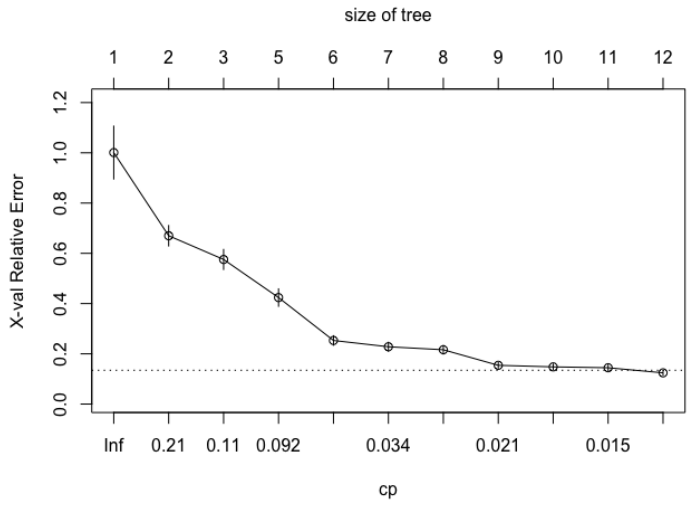


Figure 3.3 Error rates versus complexity/tree size in the DTR model

The target of DTR is to balance the outcome between sensitivity and specificity for predicting the target through the splitting (Kuhn et al., 2013). Therefore, the key to using DTR is to decide how to split nodes, or how to find appropriate $x^{(j)}$ (Eq 3.3 & 3.4) with the value of the "s". This could be represented by the minimal number of observations in a split node (minsplit). Complexity parameter (cp), which controls the size of the regression tree, is also a critical variable in the model. The cp value eliminates splits that add little or no value to the tree, acting like a stopping rule. Therefore, finding the best cp stands for choice of the optimal tree size, which avoids results of overfitting (Kuhn et al., 2013). For regression analysis, models with low RMSE are usually classified as exemplary models, so we used grid-based (trail-and-error) tuning parameters to find the best combinations between cp and minsplit. In our data, when the number of minimum observations= 11 and cp = 0.01 (Figure 3.5), there was the smallest RMSE, and method ANOVA was used in the prediction model because our output value has a numeric feature.

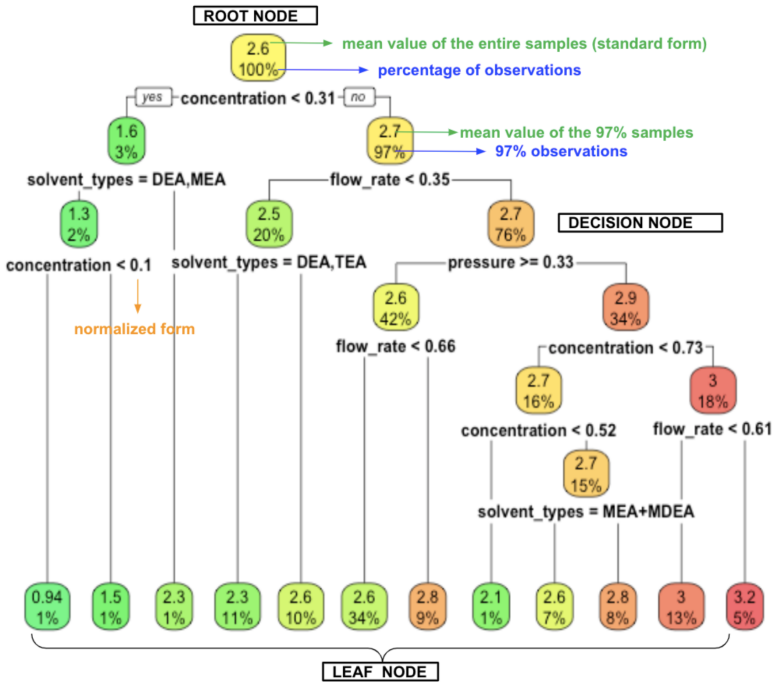


Figure 3.4 Visualization of Decision Tree Regression

Figure 3.4 interpreted the splitting process of the DTR in this study. In the figure, except the mineralization rates and percentages shown in the boxes, all other variables are in a normalized form which can be converted back by using the normaliation function (Eq 3.1). Each box in the figure displays the mean value of the nodes and the percentage of node-containing observations. For example, the predicted "mineralization rates" at the root is the mean for the whole testing dataset, which is 2.6 kmol/h. Then the first split node depended on the concentration of solvents. If the concentration were smaller than 0.31 (29.8 wt%), the data would be processed into the next phase by using "types of solvents" as the decision node. If the solvent's type is either DEA or MEA, the concentration of solvents acted as the second decision node to decide their final desorption rates. On the other hand, if the solvents were TEA or

blended solvents (MEA+MDEA), the DTR would directly give an output with a desorption rate of 2.3 kmol/h. The figure also presented that our optimal tree-structure predicted mineralization rate into eight categories with values of 0.94, 1.5, 2.1, 2.3, 2.6, 2.8, 3, 3.2 kmol/h.

3.3.3 K-nearest neighbor (KNN)

KNN is a supervised machine learning technique that is being widely used for both classification and regression (Cristian, 2018). The model decides the number of nearest neighbors based on some computational rules, and the average value of these selected features will act as the predicted outcome (Liu et al., 2020). Based on the principle of KNN, the value of k is a vital parameter for getting an accurate prediction output. In this study, the K contributing to the smallest RMSE would be chosen as the K-value in the model. The range of K value started from k=1 to k = 30 (the square root of the number of points in the training set) (Hassanat et al., 2014). Results showed that when k equaled to 2, RMSE was the smallest one with a value of 0.0079, so k=2 was used in the model (Table 3.2).

After choosing the number of neighbors, the next important step is to determine which two points should be chosen in the training dataset. In KNN, the points are chosen based on the distance between test points and all other points in the training data, and the two points resulting in the shortest distance should be chosen. In this paper, Minkowski distance was used to calculate the minimum distance between the predicted value and training features. Equation of Minkowski distance of order m (m is integer) is as follows (Yesilbudak et al., 2017):

$$DMink\ (pi, qj) = (|p_{i1} - q_{j1}|^m + |p_{i2} - q_{j2}|^m + \ldots + |p_{i(n-1)} - q_{j(n-1)}|^m + |p_{in} - q_{jn}|^m)^{\frac{1}{m}} \qquad \text{Eq 3.5}$$

where p and q are two points in n dimensions, and m is a value to manipulate types of distance metrics. For example, if P equals 1, the Minkowski distance is equivalent to the Manhattan distance; if P is 2, it approaches the Euclidean distance (Xu et al., 2019).

3.3.4 Artificial Neural Network (ANN)

ANN is a popular and effective method for nonlinear regression problems, especially when relationships between inputs and outputs are unknown (Tso & Yau, 2007). The principle of the ANN working process is similar to human neural networks. The main difference is that linkage between neurons in ANN is shown in numerical weights (Shin et al., 2019). ANN contains three basic layers: one input layer, one output layer, and one or more hidden layers (Yahya et al., 2020). The defined three layers are consistent with the Multi-Layer Perceptron (MLP) concept. This paper used MLP to process data from lower to upper layers (Shin et al., 2019). In MLP, backward propagation of errors could minimize information loss by determining the loss of every node and then reducing the weights of high error nodes, and vice versa (Karlik and Olgac, 2011).

Choosing the number of hidden layers is the first step in ANN. Shin et al. (2019) stated that one hidden layer is sufficient for most problems. Hajek and Henrique (2017) also proposed that more than one hidden layer could result in overfitting under certain conditions. Therefore, this paper also chose to use one hidden layer, and the next step to determine the number of neurons in the hidden layer. The number of neurons was calculated by "trial and error method," and the number resulting in the smallest RMSE was used. The number was tested from 2 to 30, which was in the range of many published studies (Alsina et al., 2016; Syahrullah &Sinaga,

2016; Rahimi-Ajdadi & Abbaspour-Gilandeh, 2011; Esfe et al., 2016). The results showed that when neurons equaled to 5, the smallest RMSE of 0.0239 was produced (Table 3.2). Therefore, the topology of our ANN model is 6-5-1, which stands for 6 input variables, 5 neurons on one layer, and 1 output.

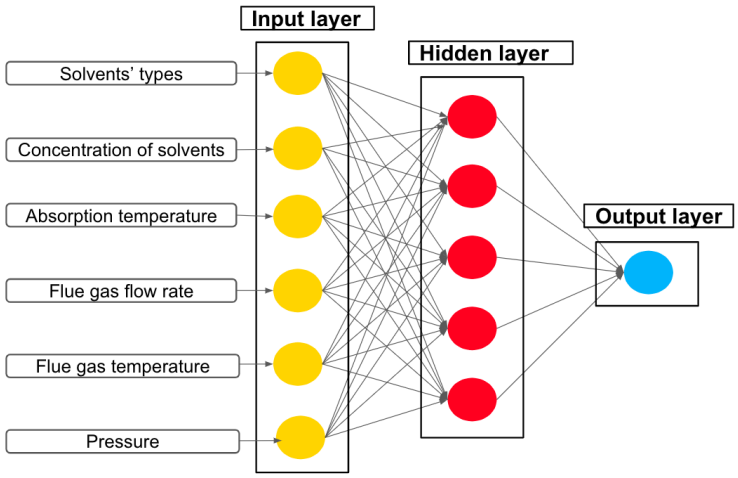


Figure 3.5 Schematic of 6-5-1 artificial neural network

The other unique parameter included in ANN is the activation function. Activation functions describe relationships between nodes, either linear or nonlinear, and transform information in neurons into an output signal (Shin et al., 2019). There are three types of activation functions: sigmoid, tanh and ReLu. The sigmoid function is an S-shaped figure with values ranging from 0 to 1; ReLu stands for Rectified Linear Unit with values from 0 to infinite, and Tanh ranges from -1 to 1 (Karlik, 2011). For regression models, tanh outperforms sigmoid (Shin et al., 2019). The statement was portrayed in the study of Shin et al. (2019) who used ANN to predict the concentration of nitric acid, which solved a similar prediction problem as our study. Furthermore, Karlik and Olgac (2011) tested the effect of using different activation

functions with 500 iterations and 40 hidden neurons. With the same number of iterations and neurons in hidden layers, their results showed that the tanh activation function was the most suitable one for all test parameters (Karlik and Olgac, 2011). Therefore, this study used tanh activation function.

3.3.5 Least Absolute Shrinkage and Selection Operator (Lasso) Regression

Lasso regression is also called the L1 penalty method, which uses penalized or regularization regression techniques to improve prediction performance on extreme observations (Ranstam & Cook, 2018). The other function of Lasso regression is to solve the problem of underlying collinearities which tend to result in inaccurate estimations of the optimal values (Mahata et al., 2020). The basic linear regression equation is as follows:

$$f(x) = wX + b \qquad \text{Eq 3.6}$$

where w is the weight vector (Hao et al., 2020).

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \ldots + \epsilon \qquad \text{Eq 3.7}$$

In multivariate regression, the equation is revised as above, which has p feature space. β_0 is the intercept, and β_1 is called the coefficient of X_1. ϵ is the random error of real values. Based on the equation, it is hard to get accurate β, ($\beta = (\beta_1. \beta_2 \ldots. \beta_p)^T$), as regression parameters when the dimension of p is too large. At this time, Lasso regression helps choose the most related variables and reduce unrelated regression parameters to absolute zero by imposing a constraint.

Variants with a regression coefficient of zero after shrinkage will be excluded from the prediction (Mozafari et al., 2020; Ranstam & Cook, 2018).

The degree of shrinkage is determined by the parameter λ (Roberts & Nowak, 2014). Therefore, whether in feature selection or normal model prediction, λ is a critical value to define before using the model. This study used cross-validation to choose lambda with the least Mean Square Error (MSE), resulting in a Lambda value of 0.001 (Table 3.2).

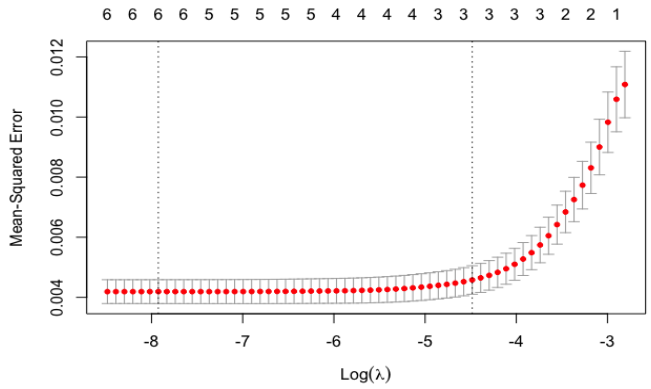


Figure 3.6 MSE values of lambda in the Lasso Regression

3.3.6 Ridge Regression (RR)

Same as Lasso Regression, RR is also a penalized regression method to avoid expansion of the standard errors of the regression coefficients produced by multicollinearity (Çiftsüren & Akkol, 2018). However, compared to Lasso Regression, RR does not turn unrelated regression parameters to absolute zero. Instead, RR reduces the value of the coefficient toward zero, which has the principle of minimizing the sum of the residual squares (Muthukrishnan & Rohini, 2016). Therefore, without eliminating coefficients of some features, RR is a better regression method

when all predictors are relevant and when variables are small. The equation of RR is: (Muthukrishnan & Rohini, 2016))

$$\widehat{\beta} = \text{argmin} \{ \sum_{i=1}^{N} (y_i - \beta_0 - \sum_{j=1}^{p} x_{ij}\beta_j)^2 + \lambda \sum_{j=1}^{p} \beta_j^2 \} \quad \text{Eq 3.8}$$

Based on the equation, it is clear to interpret RR as a penalty equivalent to the square of the magnitude of the coefficients. Argmin means 'Argument of Minimum' which aims to get the minimum value of the function. β stands for coefficient matrix with p dimensions, and β_j is the jth coefficient. y_i is the dependent variable. x_{ij} is the ith feature variable and λ controls the amount of shrinkage (Nanda et al., 2016).

The procedure of lambda selection is the same as the one illustrated in Lasso Regression. K-Fold Cross-Validation was used to find the optimal lambda value which resulted in the smallest MSE, so a lambda of 0.020 was used in RR (Figure 3.7).

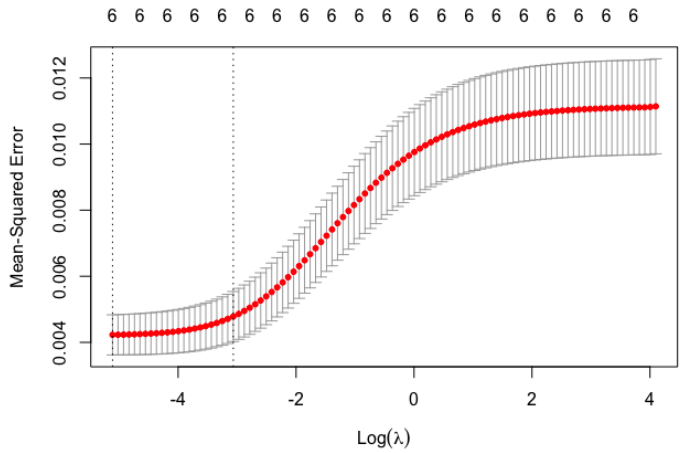


Figure 3.7 MSE values of lambda in the Ridge Regression

3.4 Evaluation criteria

3.4.1 Statistical indicators

Correlation coefficient (R^2) is one of the most prevalent indicators for measuring forecasting accuracy of the models by directly revealing the closeness between actual and predicted values. The equation is as follows (Ağbulut et al., 2020):

$$R^2 = 1 - \frac{\Sigma(y_i - x_i)^2}{\Sigma(y_i - \overline{x_i})^2} \quad \text{Eq 3.9}$$

where y_i is the predicted value; x_i is the true value, and $\overline{x_i}$ is the mean value of the true data. The range of R^2 is between 0 and 1, and values approaching 1 stands for higher agreements between predictions and observations.

The percentage differences between predicted and real mineralization rates are calculated by (Vo Thanh & Lee, 2022):

$$\text{Error Rate} = \frac{|y_i - x_i|}{y_i} * 100\ \% \quad \text{Eq 3.10}$$

However, in regression model predictions, higher R^2 does not always stand for higher accuracy. Therefore, some other statistical indicators, such as Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Mean Bias Error (MBE), and Mean Absolute Percentage Error (MAPE), are further used to determine errors of models and to choose the best model for process optimization.

RMSE is the standard deviation of prediction errors, used as the loss function, and the calculation equation is as follows:

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - x_i)^2} \quad \text{Eq 3.11}$$

where n is the total number of values. The value of RMSE means that for each predicted value, there is a certain value of difference to actual values. Therefore, this value is ideal to be close to zero (Ağbulut et al., 2020)

For MAE, Willmott and Matsurra (2005) stated that MAE is better for testing accuracy of models because compared to RMSE, MAE describes average error alone. The evaluation function of MAE is: (Ozbas et al., 2019)

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left|x_i - y_i\right| \quad \text{Eq 3.12}$$

When MAE approaches zero, model accuracy is improved because the difference between predicted and observed values becomes small.

MAPE is the other popular indicator in the regression prediction analysis because of its intuitive interpretation of errors (De Myttenaere et al., 2016). The value of MAPE is reliable if no significant number of extremes exist in the dataset. In our dataset, all observations were performed within a predefined range, so no extreme outliers existed in our dataset. Therefore, high MAPE resulting from models stands for low prediction accuracy (Agbulut et al., 2020)

$$\text{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{x_i - y_i}{x_i}\right| * 100\% \quad \text{Eq 3.13}$$

MBE was also used in our model evaluation process, but it does not relate to error magnitudes. Instead, it determines whether the model overestimates or underestimates the predicted values (Crowe et al., 2020).

$$\text{MBE} = \frac{1}{n}\sum_{i=1}^{n}\left|x_i - y_i\right| \qquad \text{Eq 3.14}$$

3.4.2 K-fold cross-validation

The overall prediction performance of different ML models were compared after applying K-fold cross-validation. K-fold cross-validation is a resampling technique that splits the dataset into K groups. K-1 groups are training sets, and the rest will be the validation dataset to evaluate generalization errors (Anguita et al., 2009). Consequently, selecting K value is crucial for minimizing prediction errors. Rodriguez et al. (2010) demonstrated that K=5 or K=10 is more appropriate for calculating prediction errors, whereas K=2 is optimal for comparing classifiers with comparable bias. In our study, K=5 was chosen due to the relatively small sample size. Although K=10 may perform better than K=5 at reducing errors, it requires significantly more computational resources than K=5 (Rodriguez et al., 2010).

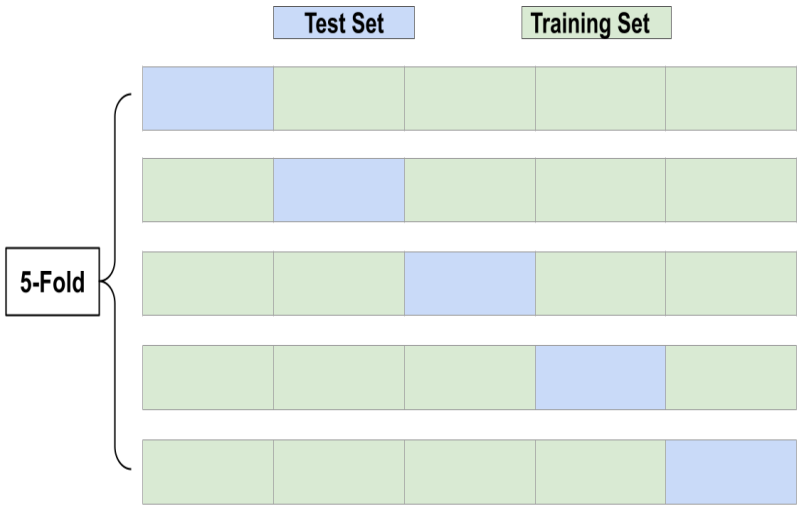


Figure 3.8 Visualization on a 5-fold validation

3.5 Evaluation results

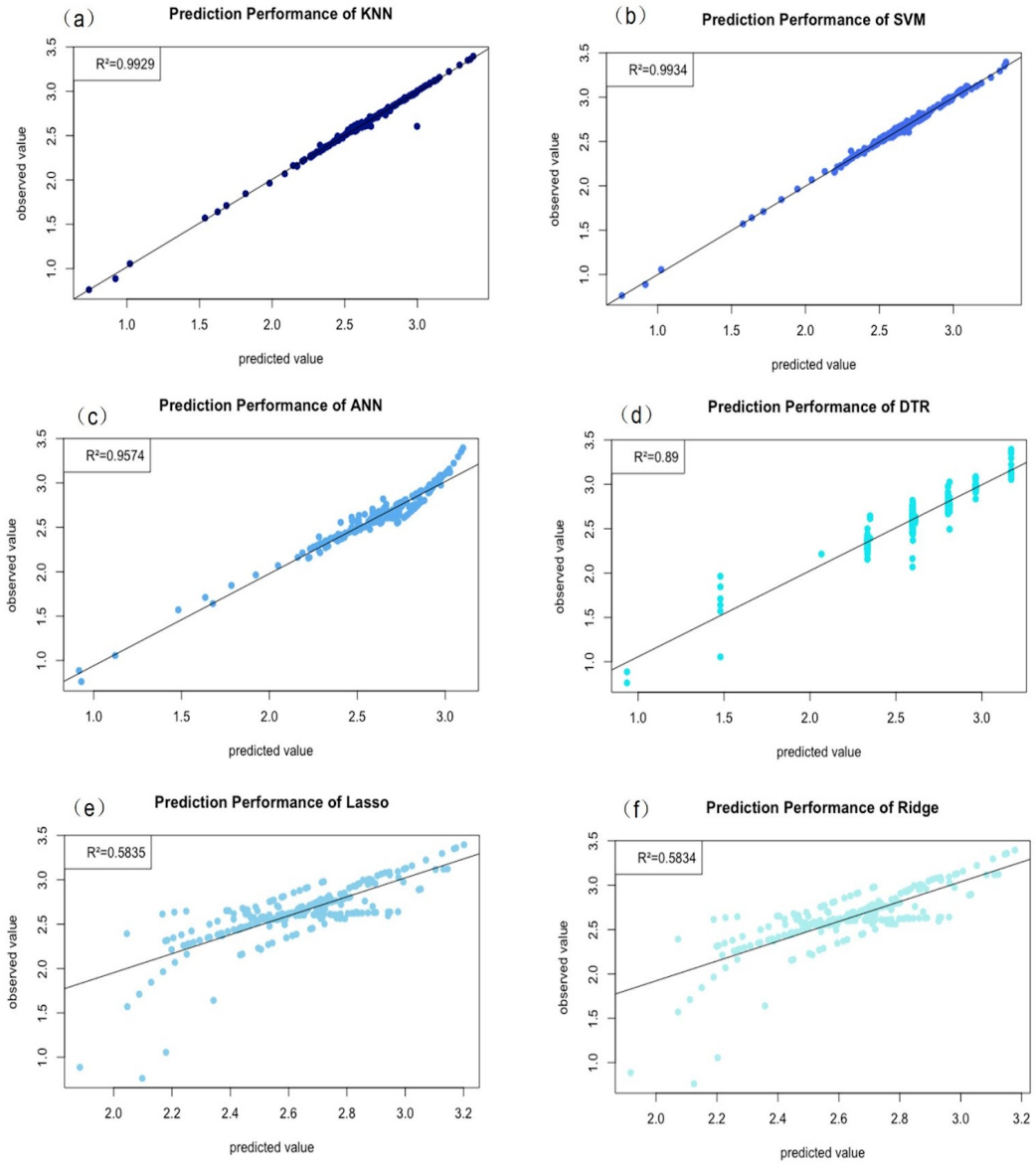


Figure 3.9 Correlation coefficients of the ML models

Figure 3.9 demonstrated that both the KNN and SVM models produced high correlation factors (R2 > 0.99, P 0.05), indicating that their predictions and true values are highly congruent. In addition, Figure 3.10 showed that the predictions of these two models are stable. Excluding

the outlier, approximately 90% of the data in KNN have an error rate of less than 1%, Even when the outlier is included, the average error value is only 0.38%, representing high agreements between actual observations and KNN predicted values. Our model performed more stable and accurate than Amin et al. (2017), whose data showed a minimum error of 1.255% and a maximum error of 44.678%. The average error rates of SVM is 0.76%, and all error rates were below 5% in which half of the points approximated a 0% error line. Furthermore, SVM model has high high accuracy on predicting peak values. Furthermore, the SVM model predicts peak values with a high degree of precision. By locating the optimal mineralization rate of 3.395 kmol/h in the observed data, the SVM predicted the value as 3.356 kmol/h, resulting in a 1.16% error rate. The small percentage error proved that the SVM model is capable of locating the maximum CO_2 mineralization rate in the ICCM process.

Although the correlation coefficients of the ANN and DTR models are lower than those of the KNN and SVM models, R^2=0.9574 and 0.89 for ANN and DTR, respectively, could still indicate a significant agreement between observed and predicted data. Error rates for the ANN model range from 0% to 22.13%, with an average error rate of 2.07%. Compared to the study by Ghiasi et al. (2019), our ANN model produced more accurate forecasts. In their study, the smallest prediction error rate of the ANN model was 0 - 30.48% in the system of CO_2 + TEA + H_2O. The range became larger when predictions were performed in the system of CO_2 + MEA + H_2O, with values ranging from 0.16 - 62.35%. For CO_2 + DEA + Water, and the range of error rates further increased to 0 - 317.94% (Ghiasi et al., 2019). In the DTR, error rates for around 95% of the model predictions were below 10% (Figure 3.10) with an average value of 3.04%. The error values were considerably lower compared to other published works. For example, in

the study of Yildiz et al. (2019) who used DTR to predict CO_2 adsorption capacity of amine-functionalized MCM-41 and SBA-15, their testing error was 16.5%, with 83.5% of the test data correctly classified.

Lasso and Ridge models performed worst among the models with both R^2 smaller than 0.6 (Figure 3.9). RR had a slightly lower average error rate (7.60%) than Lasso (7.78%). To predict the peak value of mineralization rates, the RR model resulted in a 6.41% error rate, which underestimated the value from 3.395 kmol/h to 3.178 kmol/h.

To better confirm the accuracy rank of the models, the other three evaluation metrics, including MAE, MAPE, and RMSE, were used.

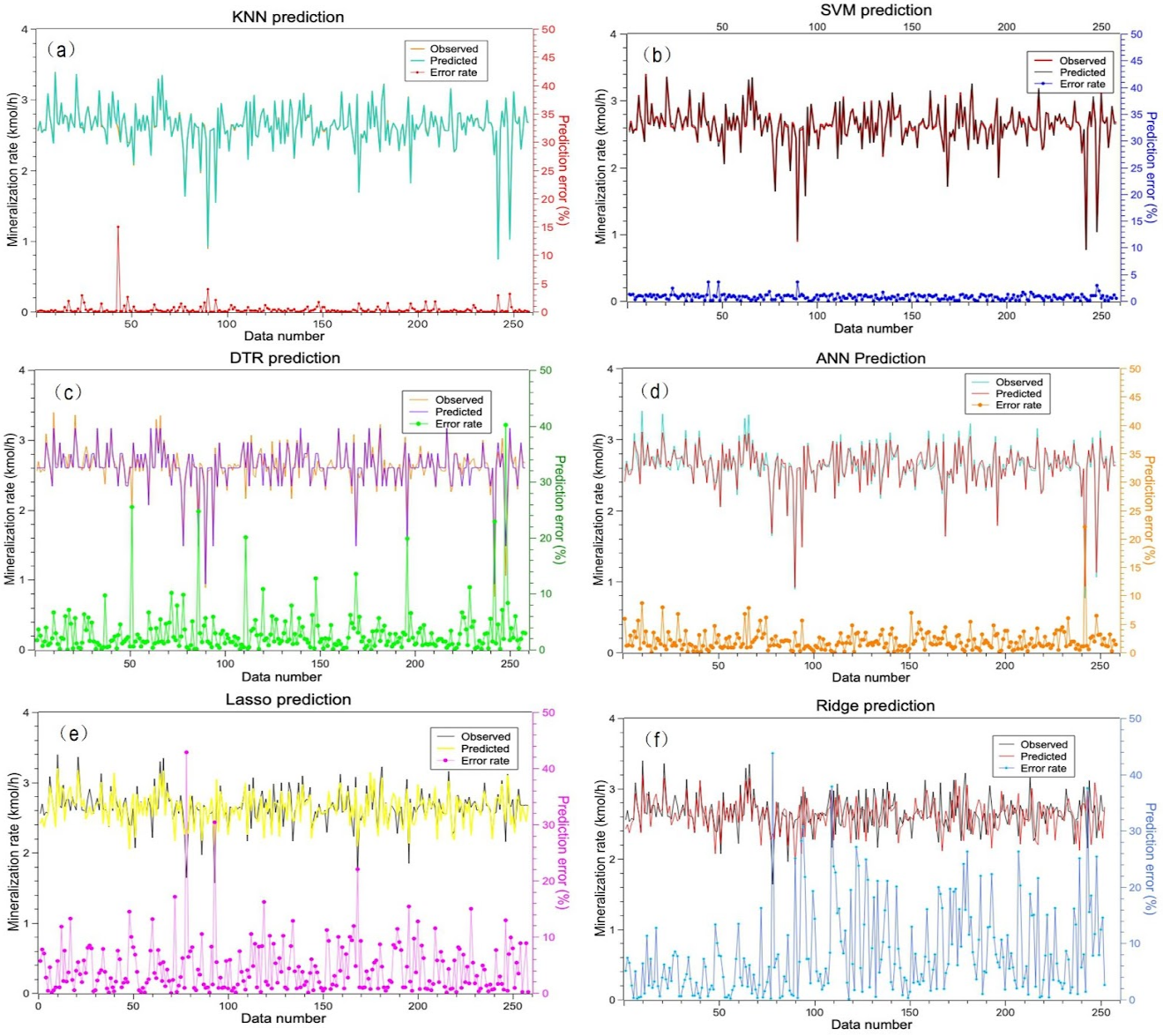


Figure 3.10 Prediction error rates of the ML models

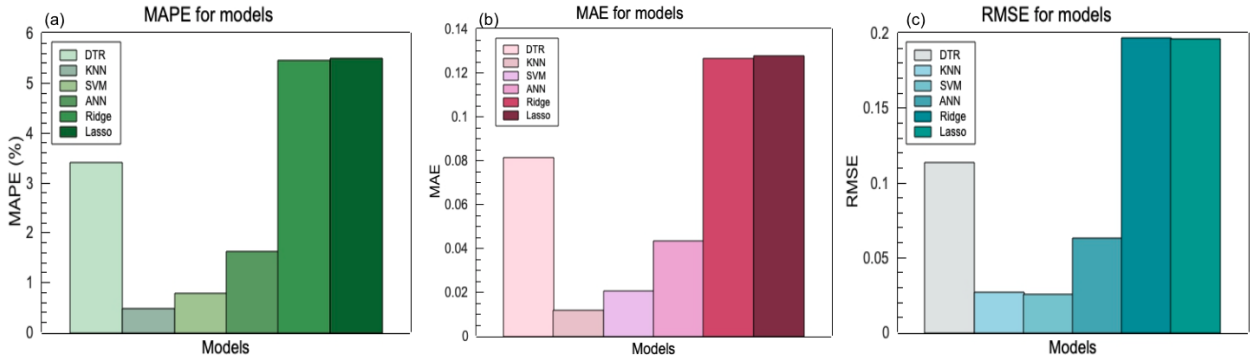


Figure 3.11 Prediction performance of the ML models

(a) MAPE values for ML model predictions (b) MAE values for ML model predictions (c) RMSE values for ML model predictions

Based on the values of MAE and MAPE, the prediction accuracy ranking from highest to lowest is as follows: KNN > SVM > ANN > DTR > > Ridge > Lasso. The MAPE and MAE of KNN with the lowest deviation between predicted and actual values is 0.479%. MAPE values for SVM, ANN, DTR, Ridge, and Lasso were 0.780%, 1.621%, 3.405%, 5.447%, and 5.490%, respectively. The MAE from the lowest to the highest is 0.0091, 0.0197, 0.0534, 0.0728, 0.1250, and 0.1263.

However, the RMSE in our study shows the good precision of the SVM (0.0252) model over the KNN (0.0269) model in estimating mineralization rates, while keeping the orders of other models the same. This could be explained by the dependence of RMSE on three characteristics of a set of errors instead of the average error (Willmott et al., 2005). Besides, RMSE relates to the variance of the dataset, which is sensitive to even a small number of outliers, (Willmott and Matsuura, 2005). Based on Figure 3.10, an outliner existed in KNN predictions, which may explain why KNN had a relatively higher RMSE value than SVM.

However, the difference in RMSE between SVM and KNN was less than 0.002, and the prediction performance based on MAE and MAPE demonstrated that the measured data and KNN outputs provided the best fit and most satisfactory agreement. Therefore, KNN is concluded to be the best model for predicting the carbon mineralization rate in ICCM processes based on the combination of metrics.

MBE was another metric discussed in this paper. The MBE indicator does not represent the magnitude of prediction errors, but rather the estimated bias. DTR, SVM, KNN, ANN, Lasso, and Ridge each have MBE values of 2.00×10^{-2}, 3.15×10^{-3}, 9.85×10^{-5}, -2.83×10^{-3}, 3.99×10^{-3}, and 4.13×10^{-3}, respectively. The only model with a negative MBE value is the ANN model, indicating that the model's predictions are slightly understated. This means that the actual mineralization rate may exceed the predicted ANN value. On the other hand, all remaining models slightly overestimated results, thereby exaggerating actual mineralization rates.

3.6 Discussion

The result provided a new perspective on the application of KNN models to regression problems, particularly when KNN outperforms SVM. This is due to the fact that SVM is a popular regression model, whereas KNN is typically used for classification problems. One of the characteristics of KNN is that it retains all prediction data. In our study with a relatively small number of observations, this is crucial because any elimination may result in a significant change. The alternative explanation is that the reaction conditions in our simulated ICCM process were within a predetermined range, as extreme operating conditions are either impractical or non-functional for the process. As a result, our dataset had a low degree of

sparsity, which prevented the lack of overlapping values and unreliable neighborhoods for KNN models (Grar et al., 2006). Consequently, KNN performed well in our research. Figure 2.6 could be the reason for SVM's relatively inferior performance in comparison to KNN. The figure indicated that the DSFG flow rate was linearly related to the mineralization rates in the MEA+MDEA - H_2O - CO_2 system, but SVM is an attractive algorithm for dealing with nonlinearly related problems.

DTR had a lower prediction accuracy than KNN and SVM because DTR is oversensitive to the training set (Rokach, 2016). Consequently, small changes in the dataset could result in high variance, which explains the high RMSE, MAE, and MAPE values of DTR in this study. Additionally, DTR models exhibit the trait of "divide and conquer." This implies that the models partition the covariates' space recursively into subspaces, with some fragments containing a relatively small number of instances (Rokach, 2016). In our dataset, for instance, some fragmentations contained less than 1% of the data (Figure 3.4), so the prediction confidence of these fragments would be limited.

The high RMSE and MAPE values of ANN are due to the "black box" nature of the model, which provides no hints on the function's structure (Tu, 1996). In addition, the relatively low prediction accuracy of the ANN model in this study could be attributed to the small number of training datasets. ANN models perform better under conditions of training with a large amount of data. (Ağbulut et al., 2020).

The low accuracy of Lasso and RR models could be explained by the relationship between independent variables and dependent variables in our dataset. The relationship between them was nonlinear (Figure 2.3 - 2.6), but Lasso and RR performed better for linear problems. In

addition, our dataset lacked high collinearities, which led to the inaccuracy of the Lasso and RR models. RR performed slightly better than Lasso Regression in terms of MAE and MAPE because RR never shrinks coefficients to zero and retains all variates, which is important when predicting datasets with a small number of features.

In conclusion, KNN performed better than other algorithms in predicting mineralization rates in the ICCM process. SVM performed less accurately than KNN, but was still superior to all other prediction models. Consequently, either KNN or SVM models could be used in practical applications to improve carbon mineralization rates.

CHAPTER 4

CONCLUSION

This study used Aspen Plus to simulate the ICCM processes and determine the effects of various factors on the final mineralization rates. Results demonstrated that the mineral and flue gas flow rates, solvent concentration, and reaction temperature are all crucial considerations during ICCM implementation. However, certain factors, such as temperature, only affect specific amine solvent types. For instance, high temperatures decreased the mineralization rates of DEA and TEA, which have relatively low reaction rates. However, the mineralization rates of MEA and MEA mixed with MDEA remained constant at varying temperatures. This is due to the fact that the fast reaction rates of the two types of amines reduce detention time to seconds, thereby diminishing the impact of external physical conditions. Concerning flue-gas flow rates, the impact patterns for all types of amines were identical and positively correlated between 5 and 35 kmol/h. CaO flow rate also had positive effects on the final mineralization rates, but it reached saturation at certain rates, and after this point, amines entered stabilization stages.

The other important investigation was that the concentration of solvents has dissimilar impacts on final mineralization rates by using different types of amines. Higher concentrations of MEA and MEA+MDEA amine solvents demonstrated rising mineralization rates. In contrast, DEA and TEA enhanced their mineralization rates at low concentrations, but mineralization rates began to decline at high concentrations, with DEA exhibiting a greater decrease. This may be explained by the formation of fluids with a gel-like consistency at high concentrations.

After the univariate analysis for the simulation process, this study employed Aspen-generated data to develop ML models that are capable of predicting CO_2 mineralization rates in ICCM processes. The performance of each model was assessed by using statistical matrices including R^2, RMSE, MAE, MAPE, and MBE. According to the criteria, KNN and SVM had significantly higher prediction accuracy than other employed models. R^2 = 0.9929, RMSE = 0.027, MAE = 0.012, MAPE = 0.479%, and MBE = 9.85×10^{-5} for KNN models; R^2 = 0.993, RMSE = 0.025, MAE = 0.020, MAPE = 0.780%, and MBE = 3.15×10^{-3} for SVM models. In addition, our results demonstrated that the Lasso and RR models performed poorly, indicating that these two models are unnecessary for non-multicollinear problems.

The highly accurate ML models, especially KNN and SVM, in this study, may lead to greater insights into real-world ICCM applications to mitigate CO_2 emissions. The potential benefits of employing the models in real-world processes include cost and energy savings, which allow for increased productivity. Furthermore, this study also demonstrated that combinations of different reaction substances and operating conditions can remarkably boost the final carbon mineralization rates, and using ML models to predict results avoid unnecessary time consumption and help design optimal experimental set-ups. The integrated ML-ICCM technique is beneficial for developing smart CCS technologies, which will improve both capture and storage levels in the future.

REFERENCES

Adams II, T. A. (2022). *Learn Aspen Plus in 24 Hours*. Tutorial 12. McGraw-Hill Education.

Ağbulut, Ü., Gürel, A. E., Ergün, A., & Ceylan, İ. (2020). Performance assessment of a V-trough photovoltaic system and prediction of power output with different machine learning algorithms. Journal of Cleaner Production, 268, 122269. https://doi.org/10.1016/j.jclepro.2020.122269

Aghel, B., Janati, S., Wongwises, S., & Shadloo, M. S. (2022). Review on CO2 capture by blended Amine Solutions. *International Journal of Greenhouse Gas Control*, *119*, 103715. https://doi.org/10.1016/j.ijggc.2022.103715

Aghel, B., Sahraie, S., & Heidaryan, E. (2020). Carbon dioxide desorption from aqueous solutions of monoethanolamine and diethanolamine in a microchannel reactor. Separation and Purification Technology, 237, 116390. https://doi.org/10.1016/j.seppur.2019.116390

Al-Malah, K. I. (2016). Aspen plus: chemical engineering applications. John Wiley & Sons.

Al-Mejibli, I. S., Alwan, J. K., & Abd, D. H. (2020). The effect of gamma value on support vector machine performance with different kernels. International Journal of Electrical and Computer Engineering (IJECE), 10(5), 5497. https://doi.org/10.11591/ijece.v10i5.pp5497-5506

Alsina, E. F., Bortolini, M., Gamberi, M., & Regattieri, A. (2016). Artificial Neural Network optimisation for monthly average daily global solar radiation prediction. Energy Conversion and Management, 120, 320–329. https://doi.org/10.1016/j.enconman.2016.04.101

Anguita, D., Ghio, A., Ridella, S., & Sterpi, D. (2009). K-Fold Cross Validation for Error Rate Estimate in Support Vector Machines. In DMIN (pp. 291-297).

Aqeeli, A. A. (2020). Enhancing carbon dioxide desorption from mono-ethanolamine solutions through in-situ mineralization and generation of calcium carbonate particles (thesis).

Arachchige,P.R., U. S., & Melaaen, M. C. (2012). Aspen plus simulation of CO2 removal from coal and gas fired power plants. Energy Procedia, 23, 391–399. https://doi.org/10.1016/j.egypro.2012.06.060

Arti, M., Youn, M. H., Park, K. T., Kim, H. J., Kim, Y. E., & Jeong, S. K. (2016). Single process for CO2 Capture and mineralization in various alkanolamines using calcium chloride. Energy & Fuels, 31(1), 763–769.

Buhre, B. J. P., Elliott, L. K., Sheng, C. D., Gupta, R. P., & Wall, T. F. (2005). Oxy-fuel combustion technology for coal-fired power generation. Progress in Energy and Combustion Science, 31(4), 283–307. https://doi.org/10.1016/j.pecs.2005.07.001

Cain, M., Jenkins, S., Allen, M. R., Lynch, J., Frame, D. J., Macey, A. H., & Peters, G. P. (2021). Methane and the Paris Agreement Temperature Goals. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 380(2215). https://doi.org/10.1098/rsta.2020.0456

Carlson, E. C. (1996). Don't gamble with physical properties for simulations. *Chemical engineering progress*, *92*(10), 35-46.

Chakma, A. (1997). CO2 Capture Processes — opportunities for improved energy efficiencies. Energy Conversion and Management, 38. https://doi.org/10.1016/s0196-8904(96)00245-2

Chen, S., Xiong, L., Ma, Q., Kim, J.-S., Chen, J., & Xu, C.-Y. (2020). Improving daily spatial precipitation estimates by merging gauge observation with multiple satellite-based precipitation products based on the geographically weighted ridge regression method. Journal of Hydrology, 589, 125156. https://doi.org/10.1016/j.jhydrol.2020.125156

Cheng, J., Yu, D., & Yang, Y. (2007). Application of support vector regression machines to the processing of end effects of Hilbert–Huang Transform. Mechanical Systems and Signal Processing, 21(3), 1197–1211. https://doi.org/10.1016/j.ymssp.2005.09.005

Chuajiw, W., Nakano, M., Takatori, K., Kojima, T., Wakimoto, Y., & Fukushima, Y. (2013). Effects of amine, amine salt and amide on the behaviour of carbon dioxide absorption into calcium hydroxide suspension to precipitate calcium carbonate. Journal of Environmental Sciences, 25(12), 2507–2515.

Çiftsüren, M. N., & Akkol, S. (2018). Prediction of internal egg quality characteristics and variable selection using regularization methods: ridge, LASSO and elastic net. Archives Animal Breeding, 61(3), 279–284. https://doi.org/10.5194/aab-61-279-2018

Concepción, E. I., Gómez-Hernández, Á., Martín, M. C., & Segovia, J. J. (2017). Density and viscosity measurements of aqueous amines at high pressures: Dea-water, Dmae-water and tea-water mixtures. The Journal of Chemical Thermodynamics, 112, 227–239. https://doi.org/10.1016/j.jct.2017.05.001

Conway, W., Bruggink, S., Beyad, Y., Luo, W., Melián-Cabrera, I., Puxty, G., & Feron, P. (2015). CO2 absorption into aqueous amine blended solutions containing monoethanolamine (MEA), N,N-dimethylethanolamine (DMEA), N,N-diethylethanolamine (DEEA) and 2-amino-2-methyl-1-propanol (AMP) for post-combustion Capture processes. Chemical Engineering Science, 126, 446–454. https://doi.org/10.1016/j.ces.2014.12.053

Cristian M, 2018. "AVERAGE MONTHLY RAINFALL FORECAST IN ROMANIA BY USING k-NEAREST NEIGHBORS REGRESSION," Annals - Economy Series, Constantin Brancusi University, Faculty of Economics, vol. 4, pages 5-12, August.

Critchfield JE. (1988) CO2 absorption/desorption methyldiethanolamine solutions promoted with monoethanolamine and diethanolamine: mass transfer and reaction kinetics.

Crowe, D., Pamula, R., Cheung, H. Y., & De Wekker, S. F. (2020). Two Supervised Machine Learning Approaches for Wind Velocity Estimation Using Multi-Rotor Copter Attitude Measurements. Sensors, 20(19), 5638. https://doi.org/10.3390/s20195638

Darde, V., Maribo-Mogensen, B., van Well, W. J. M., Stenby, E. H., & Thomsen, K. (2012). Process simulation of CO2 Capture with aqueous ammonia using the extended UNIQUAC model. *International Journal of Greenhouse Gas Control*, *10*, 74–87. https://doi.org/10.1016/j.ijggc.2012.05.017

De Myttenaere, A., Golden, B., Le Grand, B., & Rossi, F. (2016). Mean Absolute Percentage Error for regression models. Neurocomputing, 192, 38–48. https://doi.org/10.1016/j.neucom.2015.12.114

Dlugogorski, B. Z., & Balucan, R. D. (2014). Dehydroxylation of serpentine minerals: Implications for mineral carbonation. Renewable and Sustainable Energy Reviews, 31, 353–367. https://doi.org/10.1016/j.rser.2013.11.002

DuPart, M. S., Bacon, T. R., & Edwards, D. J. (1993). Understanding corrosion in alkanolamine gas treating plants: Part 1. *Hydrocarbon Processing;(United States)*, *72*(4).

Edali, M., Aboudheir, A., & Idem, R. (2009). Kinetics of carbon dioxide absorption into mixed aqueous solutions of MDEA and mea using a laminar jet apparatus and a numerically solved 2D absorption rate/kinetics model. International Journal of Greenhouse Gas Control, 3(5), 550–560. https://doi.org/10.1016/j.ijggc.2009.04.006

Esfe, M. H., Ahangar, M. R. H., Rejvani, M., Toghraie, D., & Hajmohammad, M. H. (2016). Designing an artificial neural network to predict dynamic viscosity of aqueous nanofluid of TiO2 using experimental data. International communications in heat and mass transfer, 75, 192-196.

Esfe, M. H., Ahangar, M. R. H., Rejvani, M., Toghraie, D., & Hajmohammad, M. H. (2016). Designing an artificial neural network to predict dynamic viscosity of aqueous nanofluid of TiO2 using experimental data. International communications in heat and mass transfer, 75, 192-196.

Gadikota, G. (2016). Commentary: Ex situ aqueous mineral carbonation. Frontiers in Energy Research, 4. https://doi.org/10.3389/fenrg.2016.00021

Gao, H., Wu, Z., Liu, H., Luo, X., & Liang, Z. (2017). Experimental studies on the effect of tertiary amine promoters in aqueous monoethanolamine (MEA) solutions on the absorption/stripping performances in post-combustion co2 capture. Energy & Fuels, 31(12), 13883–13891. https://doi.org/10.1021/acs.energyfuels.7b02390

Ghiasi, M. M., Abedi‐Farizhendi, S., & Mohammadi, A. H. (2019). Modeling Equilibrium Systems of Amine‐Based CO2 Capture by Implementing Machine Learning Approaches. Environmental Progress & Sustainable Energy, 38(5), 13160.

Gomes, J., Santos, S., & Bordado, J. (2014). Choosing amine-based absorbents for co2capture. Environmental Technology, 36(1), 19–25. https://doi.org/10.1080/09593330.2014.934742

Grčar, M., Fortuna, B., Mladenič, D., & Grobelnik, M. (2006). kNN versus SVM in the collaborative filtering framework. In Data Science and Classification (pp. 251-260). Springer, Berlin, Heidelberg.

Guan, J., Huang, T., Liu, W., Feng, F., Japip, S., Li, J., Wang, X., & Zhang, S. (2022). Design and prediction of Metal Organic Framework-based mixed matrix membranes for CO2 capture via machine learning. Cell Reports Physical Science, 3(5), 100864. https://doi.org/10.1016/j.xcrp.2022.100864

Gul, A., & Tezcan Un, U. (2022). Effect of temperature and gas flow rate on CO2 Capture. European Journal of Sustainable Development Research, 6(2). https://doi.org/10.21601/ejosdr/11727

Hajek, P., & Henriques, R. (2017). Modelling innovation performance of European regions using multi-output neural networks. PLOS ONE, 12(10). https://doi.org/10.1371/journal.pone.0185755

Hänchen, M., Prigiobbe, V., Baciocchi, R., & Mazzotti, M. (2008). Precipitation in the MG-carbonate system—effects of temperature and CO2 pressure. Chemical Engineering Science, 63(4), 1012–1028. https://doi.org/10.1016/j.ces.2007.09.052

Hao, D., Mehra, R. K., Luo, S., Nie, Z., Ren, X., & Fanhua, M. (2020). Experimental study of hydrogen enriched compressed natural gas (HCNG) engine and application of support vector machine (SVM) on prediction of engine performance at specific condition. International Journal of Hydrogen Energy, 45(8), 5309–5325. https://doi.org/10.1016/j.ijhydene.2019.04.039

Hao, P.-Y. (2020). Dual possibilistic regression analysis using support Vector Networks. Fuzzy Sets and Systems, 387, 1–34. https://doi.org/10.1016/j.fss.2019.03.012

Hasib-ur-Rahman, M., Bouteldja, H., Fongarland, P., Siaj, M., & Larachi Faïçal. (2012). Corrosion behavior of carbon steel in alkanolamine/room-temperature ionic liquid based

CO2 Capture Systems. *Industrial & Engineering Chemistry Research*, *51*(26), 8711–8718. https://doi.org/10.1021/ie2019849

Hassanat, A. B., Abbadi, M. A., Altarawneh, G. A., & Alhasanat, A. A. (2014). Solving the problem of the K parameter in the KNN classifier using an ensemble learning approach. arXiv preprint arXiv:1409.0919.
Hassanat, A. B., Abbadi, M. A., Altarawneh, G. A., & Alhasanat, A. A. (2014). Solving the problem of the K parameter in the KNN classifier using an ensemble learning approach. arXiv preprint arXiv:1409.0919.

Haugen, N. E., Krüger, J., Mitra, D., & Løvås, T. (2017). The effect of turbulence on mass transfer rates of small inertial particles with surface reactions. Journal of Fluid Mechanics, 836, 932–951. https://doi.org/10.1017/jfm.2017.820

Iqbal, A., & Ahmad, S. A. (2016). Pressure swing distillation of azeotropic mixture – A simulation study. *Perspectives in Science*, *8*, 4–6. https://doi.org/10.1016/j.pisc.2016.01.001

Ji, L., Yu, H., Li, K., Yu, B., Grigore, M., Yang, Q., Wang, X., Chen, Z., Zeng, M., & Zhao, S. (2018). Integrated absorption-mineralisation for low-energy CO2 Capture and sequestration. Applied Energy, 225, 356–366. https://doi.org/10.1016/j.apenergy.2018.04.108

Ji, L., Yu, H., Li, K., Yu, B., Grigore, M., Yang, Q., Wang, X., Chen, Z., Zeng, M., & Zhao, S. (2018). Integrated absorption-mineralisation for low-energy CO2 Capture and sequestration. Applied Energy, 225, 356–366. https://doi.org/10.1016/j.apenergy.2018.04.108

Jiang, J., Ye, B., & Liu, J. (2019). Research on the peak of CO2 emissions in the Developing World: Current Progress and Future Prospect. Applied Energy, 235, 186–203. https://doi.org/10.1016/j.apenergy.2018.10.089

Jiang, W., Luo, X., Gao, H., Liang, Z., Liu, B., Tontiwachwuthikul, P., & Hu, X. (2017). A comparative kinetics study of CO2 Absorption into aqueous DEEA/mea and DMEA/Mea Blended Solutions. *AIChE Journal*, *64*(4), 1350–1358. https://doi.org/10.1002/aic.16024

Kang, D., Lee, M.-G., Jo, H., Yoo, Y., Lee, S.-Y., & Park, J. (2017). Carbon capture and utilization using industrial wastewater under ambient conditions. Chemical Engineering Journal, 308, 1073–1080. https://doi.org/10.1016/j.cej.2016.09.120

Kang, D., Park, S., Jo, H., & Park, J. (2014). Carbon fixation using calcium oxide by an aqueous approach at moderate conditions. Chemical Engineering Journal, 248, 200–207. https://doi.org/10.1016/j.cej.2014.03.045

Kang, J. M., Murnandari, A., Youn, M. H., Lee, W., Park, K. T., Kim, Y. E., Kim, H. J., Kang, S.-P., Lee, J.-H., & Jeong, S. K. (2018). Energy-efficient chemical regeneration of AMP using calcium hydroxide for operating carbon dioxide capture process. Chemical Engineering Journal, 335, 338–344. https://doi.org/10.1016/j.cej.2017.10.136

Kang, J.-N., Wei, Y.-M., Liu, L.-cui, & Wang, J.-W. (2021). Observing technology reserves of carbon capture and storage via patent data: Paving the way for carbon neutral. Technological Forecasting and Social Change, 171, 120933. https://doi.org/10.1016/j.techfore.2021.120933

Kang, K., Azargohar, R., Dalai, A. K., & Wang, H. (2016). Hydrogen production from lignin, cellulose and waste biomass via supercritical water gasification: Catalyst activity and process optimization study. Energy Conversion and Management, 117, 528–537. https://doi.org/10.1016/j.enconman.2016.03.008

Karlik, B., & Olgac, A. V. (2011). Performance analysis of various activation functions in generalized MLP architectures of neural networks. International Journal of Artificial Intelligence and Expert Systems, 1(4), 111-122.

Karunarathne, S. S., Eimer, D. A., Jens, K. J., & Øi, L. E. (2020). Density, viscosity, and excess properties of ternary aqueous mixtures of MDEA + mea, DMEA + mea, and DEEA + mea. Fluids, 5(1), 27. https://doi.org/10.3390/fluids5010027

Kelemen, P., Benson, S. M., Pilorgé, H., Psarras, P., & Wilcox, J. (2019). An overview of the status and challenges of CO2 Storage in Minerals and Geological Formations. Frontiers in Climate, 1. https://doi.org/10.3389/fclim.2019.00009

Kim, S., Shi, H., & Lee, J. Y. (2016). CO 2 absorption mechanism in amine solvents and enhancement of CO 2 capture capability in blended amine solvent. International Journal of Greenhouse Gas Control, 45, 181–188. https://doi.org/10.1016/j.ijggc.2015.12.024

Kim, Y. E., Lim, J. A., Jeong, S. K., Yoon, Y. I., Bae, S. T., & Nam, S. C. (2013). Comparison of carbon dioxide absorption in aqueous mea, DEA, tea, and amp solutions. Bulletin of the Korean Chemical Society, 34(3), 783–787. https://doi.org/10.5012/bkcs.2013.34.3.783

Kuhn, L., Page, K., Ward, J., & Worrall‐Carter, L. (2013). The process and utility of classification and Regression Tree Methodology in nursing research. Journal of Advanced Nursing, 70(6), 1276–1286. https://doi.org/10.1111/jan.12288

Lawal, O., Bello, A., & Idem, R. (2005). The role of methyl diethanolamine (MDEA) in preventing the oxidative degradation of CO2 Loaded and concentrated aqueous monoethanolamine (MEA)−MDEA blends during co2 absorption from flue gases. *Industrial & Engineering Chemistry Research*, *44*(6), 1874–1896. https://doi.org/10.1021/ie049261y

Lepaumier, H., Picq, D., & Carrette, P.-L. (2009). New Amines for CO2 capture. I. Mechanisms of amine degradation in the presence of CO2. Industrial & Engineering Chemistry Research, 48(20), 9061–9067. https://doi.org/10.1021/ie900472x

Li, C., Li, X., & Zhu, X. (2017). Examining economic development and carbon emissions in China's low-carbon pilot provinces. Polish Journal of Environmental Studies, 26(6), 2619–2631. https://doi.org/10.15244/pjoes/70479

Li, Y., Duan, X., Song, W., Ma, L., & Jow, J. (2021). Reaction mechanisms of carbon dioxide capture by amino acid salt and desorption by heat or mineralization. Chemical Engineering Journal, 405, 126938. https://doi.org/10.1016/j.cej.2020.126938

Liu, J., Bai, M., Jiang, N., & Yu, D. (2019). Structural risk minimization of rough set-based classifier. Soft Computing, 24(3), 2049–2066. https://doi.org/10.1007/s00500-019-04038-8

Liu, M., & Gadikota, G. (2020). Single-step, low temperature and integrated CO2 capture and conversion using sodium glycinate to produce calcium carbonate. Fuel, 275, 117887. https://doi.org/10.1016/j.fuel.2020.117887

Liu, M., Hohenshil, A., & Gadikota, G. (2021). Integrated Co2 Capture and removal via carbon mineralization with inherent regeneration of aqueous solvents. Energy & Fuels, 35(9), 8051–8068. https://doi.org/10.1021/acs.energyfuels.0c04346

Liu, W., Wang, P., Meng, Y., Zhao, C., & Zhang, Z. (2020). Cloud spot instance price prediction using kNN regression. Human-Centric Computing and Information Sciences, 10(1). https://doi.org/10.1186/s13673-020-00239-5

Lu, C., Li, M., Xu, L., Fu, D., & Zhang, P. (2020). Study on the absorption-mineralisation for low-energy co2 capture in BDA activated DEEA aqueous solution using calcium chloride. IOP Conference Series: Earth and Environmental Science, 514(5), 052020. https://doi.org/10.1088/1755-1315/514/5/052020

Lu, H., Diaz, D. J., Czarnecki, N. J., Zhu, C., Kim, W., Shroff, R., Acosta, D. J., Alexander, B. R., Cole, H. O., Zhang, Y., Lynd, N. A., Ellington, A. D., & Alper, H. S. (2022). Machine learning-aided engineering of hydrolases for pet depolymerization. *Nature*, *604*(7907), 662–667. https://doi.org/10.1038/s41586-022-04599-z

Lyu, Z., Yu, Y., Samali, B., Rashidi, M., Mohammadi, M., Nguyen, T. N., & Nguyen, A. (2022). Back-propagation neural network optimized by k-fold cross-validation for prediction of torsional strength of reinforced concrete beam. Materials, 15(4), 1477. https://doi.org/10.3390/ma15041477

Ma, D., Zhu, C., Fu, T., Yuan, X., & Ma, Y. (2020). An effective hybrid solvent of MEA/DEEA for CO2 absorption and its mass transfer performance in microreactor. Separation and Purification Technology, 242, 116795. https://doi.org/10.1016/j.seppur.2020.116795

Maakoul, O., Beaulanda, R., El Omari, H., Essabri, E. H., & Abid, A. (2021). Modeling CO2 capture processes by chemical absorption: The case of biogas treatment. E3S Web of Conferences, 234, 00024. https://doi.org/10.1051/e3sconf/202123400024

Mahata, C., Ray, S., & Das, D. (2020). Optimization of dark fermentative hydrogen production from organic wastes using acidogenic mixed consortia. Energy Conversion and Management, 219, 113047. https://doi.org/10.1016/j.enconman.2020.113047

Mazari, S. A., Ghalib, L., Sattar, A., Bozdar, M. M., Qayoom, A., Ahmed, I., Muhammad, A., Abro, R., Abdulkareem, A., Nizamuddin, S., Baloch, H., & Mubarak, N. M. (2020). Review of modelling and simulation strategies for evaluating corrosive behavior of aqueous amine systems for CO2 Capture. *International Journal of Greenhouse Gas Control*, *96*, 103010. https://doi.org/10.1016/j.ijggc.2020.103010

Mohammadian, E., Liu, B., & Riazi, A. (2022). Evaluation of different machine learning frameworks to estimate CO2 solubility in NaCl brines: Implications for CO2 injection into low-salinity formations. Lithosphere, 2022(Special 12). https://doi.org/10.2113/2022/1615832

Mori, N., Debeljak, B., Škerjanec, M., Simčič, T., Kanduč, T., & Brancelj, A. (2019). Modelling the effects of multiple stressors on respiration and microbial biomass in the hyporheic zone using decision trees. *Water Research*, *149*, 9–20. https://doi.org/10.1016/j.watres.2018.10.093

Mozafari, Z., Arab Chamjangali, M., & Arashi, M. (2020). Combination of least absolute shrinkage and selection operator with Bayesian Regularization artificial neural network (LASSO-BR-ANN) for QSAR studies using functional group and molecular docking mixed descriptors. Chemometrics and Intelligent Laboratory Systems, 200, 103998. https://doi.org/10.1016/j.chemolab.2020.103998

Muchan, P., Saiwan, C., Narku-Tetteh, J., Idem, R., Supap, T., & Tontiwachwuthikul, P. (2017). Screening tests of aqueous alkanolamine solutions based on primary, secondary, and tertiary structure for blended aqueous amine solution selection in Post Combustion Co 2 capture. Chemical Engineering Science, 170, 574–582. https://doi.org/10.1016/j.ces.2017.02.031

Muthukrishnan, R., & Rohini, R. (2016). LASSO: A feature selection technique in predictive modeling for machine learning. 2016 IEEE International Conference on Advances in Computer Applications (ICACA). https://doi.org/10.1109/icaca.2016.7887916

Nanda, S., Isen, J., Dalai, A. K., & Kozinski, J. A. (2016). Gasification of fruit wastes and agro-food residues in supercritical water. Energy Conversion and Management, 110, 296–306. https://doi.org/10.1016/j.enconman.2015.11.060

Nuchitprasittichai, A., & Cremaschi, S. (2013). Optimization of CO2 capture process with aqueous amines—a comparison of two simulation–optimization approaches. Industrial & Engineering Chemistry Research, 52(30), 10236–10243. https://doi.org/10.1021/ie3029366

Osman, A. I., Hefny, M., Abdel Maksoud, M. I., Elgarahy, A. M., & Rooney, D. W. (2020). Recent advances in carbon capture storage and Utilisation Technologies: A Review. Environmental Chemistry Letters, 19(2), 797–849. https://doi.org/10.1007/s10311-020-01133-3

Osman, K., Coquelet, C., & Ramjugernath, D. (2012). Absorption data and modeling of carbon dioxide in aqueous blends of Bis(2-hydroxyethyl)methylamine (MDEA) and 2,2-iminodiethanol (DEA): 25 % MDEA + 25 % DEA and 30 % MDEA + 20 % DEA. Journal of Chemical & Engineering Data, 57(5), 1607–1620. https://doi.org/10.1021/je201132d

Ovejero, G., Romero, M. D., Díez, E., Díaz, I., & Pérez, P. (2009). Thermodynamic modeling and simulation of styrene−butadiene rubbers (SBR) solvent equilibrium staged processes. *Industrial & Engineering Chemistry Research*, *48*(16), 7713–7723. https://doi.org/10.1021/ie9006497

Ozbas, E. E., Aksu, D., Ongen, A., Aydin, M. A., & Ozcan, H. K. (2019). Hydrogen production via biomass gasification, and modeling by supervised machine learning algorithms. International Journal of Hydrogen Energy, 44(32), 17260–17268. https://doi.org/10.1016/j.ijhydene.2019.02.108

Ozbas, E. E., Aksu, D., Ongen, A., Aydin, M. A., & Ozcan, H. K. (2019). Hydrogen production via biomass gasification, and modeling by supervised machine learning algorithms. International Journal of Hydrogen Energy, 44(32), 17260–17268. https://doi.org/10.1016/j.ijhydene.2019.02.108

Park, S., Min, J., Lee, M.-G., Jo, H., & Park, J. (2013). Characteristics of CO2 fixation by chemical conversion to carbonate salts. Chemical Engineering Journal, 231, 287–293. https://doi.org/10.1016/j.cej.2013.07.032

Park, S.-W., Choi, B.-S., & Lee, J.-W. (2006). Chemical absorption of carbon dioxide with triethanolamine in non-aqueous solutions. *Korean Journal of Chemical Engineering*, *23*(1), 138–143. https://doi.org/10.1007/bf02705705

Patro, S. G. K., & sahu, K. K. (2015). Normalization: A preprocessing stage. IARJSET, 20–22. https://doi.org/10.17148/iarjset.2015.2305

Petitjean, H., Chizallet, C., Krafft, J.-M., Che, M., Lauron-Pernot, H., & Costentin, G. (2010). Basic reactivity of CAO: Investigating active sites under operating conditions. Physical Chemistry Chemical Physics, 12(44), 14740. https://doi.org/10.1039/c0cp00855a

Rahimi-Ajdadi, F., & Abbaspour-Gilandeh, Y. (2011). Artificial neural network and stepwise multiple range regression methods for prediction of tractor fuel consumption. Measurement, 44(10), 2104-2111.

Ranstam, J., & Cook, J. A. (2018). LASSO regression. Journal of British Surgery, 105(10), 1348-1348.

Rashidi, S., Vafakhah, M., Lafdani, E. K., & Javadi, M. R. (2016). Evaluating the support vector machine for suspended sediment load forecasting based on Gamma Test. Arabian Journal of Geosciences, 9(11). https://doi.org/10.1007/s12517-016-2601-9

Rivera,M. R., & Gerven,V. T. (2020). Production of calcium carbonate with different morphology by simultaneous CO2 capture and mineralisation. Journal of CO2 Utilization, 41, 101241. https://doi.org/10.1016/j.jcou.2020.101241

Roberts, S., & Nowak, G. (2014). Stabilizing the lasso against cross-validation variability. Computational Statistics & Data Analysis, 70, 198–211. https://doi.org/10.1016/j.csda.2013.09.008

Rodriguez, J. D., Perez, A., & Lozano, J. A. (2010). Sensitivity analysis of K-fold cross validation in prediction error estimation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 32(3), 569–575. https://doi.org/10.1109/tpami.2009.187

Rodríguez, L. S., Muñoz, J. M. B., Zambon, A., & Faure, J. P. (2017). Determination of the biomass content of end-of-life tyres. *Biomass Volume Estimation and Valorization for Energy*, 444-462. (79)

Rokach, L. (2016). Decision forest: Twenty years of research. Information Fusion, 27, 111–125. https://doi.org/10.1016/j.inffus.2015.06.005

Rostami, A., Arabloo, M., Lee, M., & Bahadori, A. (2018). Applying SVM framework for modeling of CO2 solubility in oil during CO2 flooding. Fuel, 214, 73–87. https://doi.org/10.1016/j.fuel.2017.10.121

Saha, A., Pal, S., Arabameri, A., Blaschke, T., Panahi, S., Chowdhuri, I., Chakrabortty, R., Costache, R., & Arora, A. (2021). Flood susceptibility assessment using novel ensemble of Hyperpipes and support vector regression algorithms. Water, 13(2), 241. https://doi.org/10.3390/w13020241

Sanna, A., Uibu, M., Caramanna, G., Kuusik, R., & Maroto-Valer, M. M. (2014). A review of mineral carbonation technologies to sequester CO2. Chem. Soc. Rev., 43(23), 8049–8080. https://doi.org/10.1039/c4cs00035h

Sarkar, S., Patel, A., Madaan, S., & Maiti, J. (2016). Prediction of occupational accidents using decision tree approach. 2016 IEEE Annual India Conference (INDICON). https://doi.org/10.1109/indicon.2016.7838969

Sayyad Amin, J., Bahadori, A., Hosseini Nia, B., Rafiee, S., & Kheilnezhad, N. (2017). Prediction of hydrate equilibrium conditions using K-nearest neighbor algorithm to co2 capture. Petroleum Science and Technology, 35(11), 1070–1077. https://doi.org/10.1080/10916466.2017.1302475

Sema, T., Naami, A., Fu, K., Edali, M., Liu, H., Shi, H., Liang, Z., Idem, R., & Tontiwachwuthikul, P. (2012). Comprehensive mass transfer and reaction kinetics studies of CO2 absorption into aqueous solutions of blended MDEA–Mea. Chemical Engineering Journal, 209, 501–512. https://doi.org/10.1016/j.cej.2012.08.016

Sema, T., Naami, A., Liang, Z., Chen, G., Gao, R., Idem, R., & Tontiwachwuthikul, P. (2013). A novel reactive 4-Diethylamino-2-butanol solvent for capturing CO2 in the aspect of absorption capacity, cyclic capacity, mass transfer, and reaction kinetics. *Energy Procedia*, *37*, 477–484. https://doi.org/10.1016/j.egypro.2013.05.133

Sema, T., Naami, A., Liang, Z., Chen, G., Gao, R., Idem, R., & Tontiwachwuthikul, P. (2013). A novel reactive 4-Diethylamino-2-butanol solvent for capturing CO2 in the aspect of absorption capacity, cyclic capacity, mass transfer, and reaction kinetics. Energy Procedia, 37, 477–484. https://doi.org/10.1016/j.egypro.2013.05.133

Shalaby, A., Elkamel, A., Douglas, P. L., Zhu, Q., & Zheng, Q. P. (2021). A machine learning approach for modeling and optimization of a CO2 post-combustion Capture unit. Energy, 215, 119113. https://doi.org/10.1016/j.energy.2020.119113

Shin, Y., Kim, Z., Yu, J., Kim, G., & Hwang, S. (2019). Development of NOx reduction system utilizing artificial neural network (ANN) and genetic algorithm (GA). Journal of Cleaner Production, 232, 1418–1429. https://doi.org/10.1016/j.jclepro.2019.05.276

Shin, Y., Kim, Z., Yu, J., Kim, G., & Hwang, S. (2019). Development of NOx reduction system utilizing artificial neural network (ANN) and genetic algorithm (GA). Journal of Cleaner Production, 232, 1418–1429. https://doi.org/10.1016/j.jclepro.2019.05.276

Simoni, L. D., Lin, Y., Brennecke, J. F., & Stadtherr, M. A. (2007). Modeling liquid−liquid equilibrium of ionic liquid systems with NRTL, electrolyte-NRTL, and UNIQUAC. *Industrial & Engineering Chemistry Research*, *47*(1), 256–272. https://doi.org/10.1021/ie070956j

Sipöcz, N., Tobiesen, F. A., & Assadi, M. (2011). The use of artificial neural network models for CO2 Capture Plants. Applied Energy, 88(7), 2368–2376. https://doi.org/10.1016/j.apenergy.2011.01.013

Smola, A. J., & Schölkopf, B. (2004). A tutorial on support vector regression. Statistics and Computing, 14(3), 199–222. https://doi.org/10.1023/b:stco.0000035301.49549.88

Snæbjörnsdóttir, S. Ó., Sigfússon, B., Marieni, C., Goldberg, D., Gislason, S. R., & Oelkers, E. H. (2020). Carbon dioxide storage through mineral carbonation. Nature Reviews Earth & Environment, 1(2), 90–102. https://doi.org/10.1038/s43017-019-0011-8

Solano Meza, J. K., Orjuela Yepes, D., Rodrigo-Ilarri, J., & Cassiraga, E. (2019). Predictive analysis of Urban Waste Generation for the city of Bogotá, Colombia, through the implementation of decision trees-based machine learning, support vector machines and artificial neural networks. Heliyon, 5(11). https://doi.org/10.1016/j.heliyon.2019.e02810

Song, Y., Liang, J., Lu, J., & Zhao, X. (2017). An efficient instance selection algorithm for k nearest neighbor regression. Neurocomputing, 251, 26–34. https://doi.org/10.1016/j.neucom.2017.04.018

Song, Y. Y., & Lu, Y. (2015). Decision tree methods: applications for classification and prediction. Shanghai archives of psychiatry, 27(2), 130–135. https://doi.org/10.11919/j.issn.1002-0829.215044

Syahrullah, L. O. I., & Sinaga, N. (2016). Optimization and prediction of motorcycle injection system performance with feed-forward back-propagation method Artificial Neural Network (ANN). American Journal of Engineering and Applied Science, 9(2), 222-235

Tan, L. S., Shariff, A. M., Lau, K. K., & Bustam, M. A. (2015). Impact of high pressure on high concentration carbon dioxide capture from natural gas by monoethanolamine/N-methyl-2-pyrrolidone solvent in absorption packed column. International Journal of Greenhouse Gas Control, 34, 25–30. https://doi.org/10.1016/j.ijggc.2014.12.020

Tan, X., Liu, Y., Dong, H., Xiao, Y., & Zhao, Z. (2022). The health consequences of greenhouse gas emissions: A potential pathway. Environmental Geochemistry and Health. https://doi.org/10.1007/s10653-021-01142-3

Tso, G. K. F., & Yau, K. K. W. (2007). Predicting electricity energy consumption: A comparison of regression analysis, decision tree and neural networks. Energy, 32(9), 1761–1768. https://doi.org/10.1016/j.energy.2006.11.010

Tu, J. V. (1996). Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. Journal of Clinical Epidemiology, 49(11), 1225–1231. https://doi.org/10.1016/s0895-4356(96)00002-9

Unlu, D., & Hilmioglu, N. D. (2020). Application of aspen plus to renewable hydrogen production from glycerol by steam reforming. *International Journal of Hydrogen Energy*, *45*(5), 3509–3515. https://doi.org/10.1016/j.ijhydene.2019.02.106

Velasco, L. C., Aca-ac, J. M., Cajes, J. J., Lactuan, N. J., & Chit, S. C. (2022). Rainfall forecasting using support vector regression machines. International Journal of Advanced Computer Science and Applications, 13(3). https://doi.org/10.14569/ijacsa.2022.0130329

Vo Thanh, H., & Lee, K.-K. (2022). Application of machine learning to predict CO2 trapping performance in deep saline aquifers. Energy, 239, 122457. https://doi.org/10.1016/j.energy.2021.122457

Wang, H., Xu, D., & Martinez, A. (2018). Parameter selection method for support vector machine based on adaptive fusion of multiple kernel functions and its application in fault diagnosis. Neural Computing and Applications, 32(1), 183–193. https://doi.org/10.1007/s00521-018-3792-7

Wang, R., Liu, S., Li, Q., Zhang, S., Wang, L., & An, S. (2021a). CO2 capture performance and mechanism of blended amine solvents regulated by N-methylcyclohexyamine. Energy, 215, 119209. https://doi.org/10.1016/j.energy.2020.119209

Wang, S., Ji, B., Zhao, J., Liu, W., & Xu, T. (2018). Predicting ship fuel consumption based on lasso regression. Transportation Research Part D: Transport and Environment, 65, 817–824. https://doi.org/10.1016/j.trd.2017.09.014

Wang, X., An, K., Tang, L., & Chen, X. (2015). Short Term Prediction of Freeway Exiting Volume Based on SVM and KNN. International Journal of Transportation Science and Technology, 4(3), 337–352. https://doi.org/10.1260/2046-0430.4.3.337

Wang, Y., Song, L., Ma, K., Liu, C., Tang, S., Yan, Z., Yue, H., & Liang, B. (2021b). An integrated absorption–mineralization process for CO2 Capture and sequestration: Reaction mechanism, recycling stability, and Energy Evaluation. ACS Sustainable Chemistry & Engineering, 9(49), 16577–16587.

Wei, Y.-M., Kang, J.-N., Liu, L.-C., Li, Q., Wang, P.-T., Hou, J.-J., Liang, Q.-M., Liao, H., Huang, S.-F., & Yu, B. (2021). A proposed global layout of carbon capture and storage in line with a 2 °C climate target. Nature Climate Change, 11(2), 112–118. https://doi.org/10.1038/s41558-020-00960-0

Wilberforce, T., Olabi, A. G., Sayed, E. T., Elsaid, K., & Abdelkareem, M. A. (2021). Progress in carbon capture technologies. Science of The Total Environment, 761, 143203. https://doi.org/10.1016/j.scitotenv.2020.143203

Willmott, C. J., & Matsuura, K. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. Climate Research, 30, 79–82. https://doi.org/10.3354/cr030079

Willmott, C. J., & Matsuura, K. (2005). Advantages of the mean absolute error (mae) over the root mean square error (RMSE) in assessing average model performance. Climate Research, 30, 79–82. https://doi.org/10.3354/cr030079

Witoon, T. (2011). Characterization of calcium oxide derived from waste eggshell and its application as CO2 Sorbent. Ceramics International, 37(8), 3291–3298. https://doi.org/10.1016/j.ceramint.2011.05.125

Xiao, M., Liu, H., Gao, H., & Liang, Z. (2018). CO2 absorption with aqueous tertiary amine solutions: Equilibrium solubility and thermodynamic modeling. The Journal of Chemical Thermodynamics, 122, 170–182. https://doi.org/10.1016/j.jct.2018.03.020

Xu, H., Zeng, W., Zeng, X., & Yen, G. G. (2019). An evolutionary algorithm based on Minkowski distance for many-objective optimization. IEEE Transactions on Cybernetics, 49(11), 3968–3979. https://doi.org/10.1109/tcyb.2018.2856208

Yahya, H. S., Abbas, T., & Amin, N. A. (2020). Optimization of hydrogen production via toluene steam reforming over Ni–Co supported modified-activated carbon using ANN coupled GA and RSM. International Journal of Hydrogen Energy. https://doi.org/10.1016/j.ijhydene.2020.05.033

Yahya, H. S., Abbas, T., & Amin, N. A. (2020). Optimization of hydrogen production via toluene steam reforming over Ni–Co supported modified-activated carbon using ANN coupled GA and RSM. International Journal of Hydrogen Energy. https://doi.org/10.1016/j.ijhydene.2020.05.033

Yan, Z., Wang, Y., Yue, H., Liu, C., Zhong, S., Ma, K., Liao, W., Tang, S., & Liang, B. (2021). Integrated process of monoethanolamine-based CO2 absorption and CO2 mineralization with SFGD Slag: Process Simulation and life-cycle assessment of CO2 Emission. ACS Sustainable Chemistry & Engineering, 9(24), 8238–8248. https://doi.org/10.1021/acssuschemeng.1c02278

Yesilbudak, M., Sagiroglu, S., & Colak, I. (2017). A novel implementation of knn classifier based on multi-tupled meteorological input data for wind power prediction. Energy Conversion and Management, 135, 434–444. https://doi.org/10.1016/j.enconman.2016.12.0

Yu, B., Li, K., Ji, L., Yang, Q., Jiang, K., Megharaj, M., Yu, H., & Chen, Z. (2019). Coupling a sterically hindered amine-based absorption and coal fly ash triggered amine regeneration: A high energy-saving process for CO2 Absorption and Sequestration. International Journal of Greenhouse Gas Control, 87, 58–65. https://doi.org/10.1016/j.ijggc.2019.05.006

Yıldız, M. G., Davran-Candan, T., Günay, M. E., & Yıldırım, R. (2019). CO2 capture over amine-functionalized MCM-41 and SBA-15: exploratory analysis and decision tree classification of past data. Journal of CO2 Utilization, 31, 27-42.

Zhang, J., Fennell, P. S., & Trusler, J. P. (2015). Density and viscosity of partially carbonated aqueous tertiary alkanolamine solutions at temperatures between (298.15 and 353.15) K. Journal of Chemical & Engineering Data, 60(8), 2392–2399. https://doi.org/10.1021/acs.jced.5b00282

Zhang, R., Liang, Z., Liu, H., Rongwong, W., Luo, X., Idem, R., & Yang, Q. (2016). Study of formation of bicarbonate ions in co2-loaded aqueous single 1DMA2P and MDEA tertiary amines and blended mea–1DMA2P and mea–MDEA amines for low heat of regeneration. Industrial & Engineering Chemistry Research, 55(12), 3710–3717. https://doi.org/10.1021/acs.iecr.5b03097

Zhang, W., Li, J., Wang, Q., & Qiu, X. (2020). Desorption and mineralization of CO2 in amine-based solution by $Ca(OH)_2$. International Journal of Greenhouse Gas Control, 97, 103056. https://doi.org/10.1016/j.ijggc.2020.103056

Zhang, W., Xu, Y., & Wang, Q. (2022b). Coupled CO2 absorption and mineralization with low-concentration monoethanolamine. Energy, 241, 122524. https://doi.org/10.1016/j.energy.2021.122524

Zhang, W., Xu, Y., Deng, Z., & Wang, Q. (2022a). Experiments on continuous chemical desorption of CO2-Rich Solutions. Energy, 239, 122354. https://doi.org/10.1016/j.energy.2021.122354