Heliyon 9 (2023) e22050

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Development of experimental error-Driven model for prediction of corrosion rates of amines based on their chemical structures

Jessica Narku-Tetteh [a], Ebenezer Mensah [a], Pailin Muchan [a], Teeradet Supap [a], Supranee Lisawadi [b], Raphael Idem [a,*]

[a] Clean Energy Technologies Research Institute, University of Regina, 3737 Wascana Parkway, SK S4S 0A2, Regina, Saskatchewan, Canada
[b] Department of Mathematics and Statistics, Faculty of Science and Technology, Thammasat University, Rangsit Centre, Khlong Nueng, Khlong Luang, Pathum Thani, Thailand, 12120

ARTICLE INFO

Keywords:
Graphical user interface
Amine structures
corrosion rate
Hydrophobicity
Steric effect
Alkyl length

ABSTRACT

This work investigated the relationships between amine corrosion rates and their chemical structural properties for application in the development of a Gaussian Process Regression (GPR) model for chemical structure-based prediction of corrosion rate of any amine. The GPR model accounted for experimental errors, which widened its scope to accurately predict the true corrosion rates, being restricted only to error associated with the trained model. The Average Absolute Deviation (AAD) between experimental corrosion rates and model predicted rates was 4.26 % for the test data, and 5.32 % for two test data unknown to the model. This showed that the model is generalizable and its predictions are accurate. This work also developed a user-friendly Graphical-User Interface, which allows a user to define any amine's structure to provide needed information to calculate its surface tension and steric effects for use as input variables to the model in predicting the corrosion rate of the amine.

1. Introduction

As the world moves towards a cleaner energy ecosystem, a lot of emphasis is put on the reduction of greenhouse gases, specifically, CO_2, the main culprit of climate change and global warming. Carbon capture utilisation and storage (CCUS) is one of the anchor technologies that is employed to help abate this serious environmental menace. In the field of CCUS, post combustion capture using amine solvents is the most mature and effective strategy that can help mitigate GHG emissions. Amines are reactive chemical solvents that are derivatives of ammonia. Amines have different properties like surface tension, steric effect, alkalinity, etc. These chemical properties influence the amine compatibility with the material of construction and the ability of amine to remove acid gases like CO_2 from the exhaust flue gases from different industrial processes. It is well known that amine solvents are used to remove acid gases like CO_2 from exhaust flue gases in the amine-based post combustion CO_2 capture process. This amine scrubbing process can be applied to a wide variety of CO_2 emitters including coal-fired powered plants, natural gas plants, cement plants, etc. For carbon capture plants, carbon steel and/or stainless steel is used as the material of construction. Carbon steel is susceptible to corrosion, which is a severe operational problem in process plants that can lead to production losses, unplanned downtime, reduced equipment life or even cause injury or death [1]. The severity of corrosion depends on the type (i.e. chemical structure) and concentration of the amine solution,

* Corresponding author.
E-mail address: raphael.idem@uregina.ca (R. Idem).

https://doi.org/10.1016/j.heliyon.2023.e22050
Received 8 September 2023; Received in revised form 19 October 2023; Accepted 2 November 2023
Available online 4 November 2023

oxygen ingression, temperature, CO_2 loading, amine degradation products and other contamination in the solution [2].

Studies done by some researchers [3] have indicated that the rate of corrosion of carbon steel increases as partial pressure of CO_2 (p_{CO2}) increases. Similar relations of the rate of corrosion and p_{CO2} are reported in different studies [4]. Further studies done in the literature [5] revealed the corrosion inhibition potential of CO_2 at certain CO_2 loadings. At elevated temperatures, amines along with their degradation products act as corrosion agents [6]. It has also been established that solutions with higher CO_2 concentration have higher corrosion potential than that of lean solutions [7]. Different researchers [8–10] have found $HCO_3-$and H_2O to be the major oxidizing agents for corrosion for aqueous amine-based CO_2 systems. All the aforementioned factors that affect the rate of corrosion become relevant only when we fundamentally understand the corrosion behaviour of the amine itself. The inherent structural property of amines defines their chemical behavior. Since the corrosiveness of amines is contingent on their structural property, it becomes vitally important to investigate the corrosiveness of amines in relation to their chemical structural features. Corrosion can be controlled by using the appropriate equipment design which involves highly corrosion resistant materials of construction such as stainless steel as well as the use of chemical treatment [11,12] (e.g the addition of corrosion inhibitors to the amine solvent).

Considering the time-consuming nature and laborious nature of carrying out experiments to determine the rate of amine corrosion, it becomes prudent to find an effective alternative that can help determine the corrosion rates of amine solvents. In addition, the development of innovative solutions from laboratory experiments to an industrial scale can be costly, restricting the breadth of experimentation. Hence the need for mathematical models to predict corrosion rates of amines accurately. Various researchers have developed models to predict the corrosion behavior of amines, including empirical, semi-empirical and mechanistic models [13–17]. All models mentioned use process parameters like amine concentration, $[O_2]$, $[SO_2]$ $[O_2]$, $[CO_2]$, solution pH, temperature, CO_2 product species like HCO_3^- and H_2CO_3 as variables to predict the corrosion rates of amines. The rate of corrosion in the capture plant increases in areas of high acid gases like CO_2, SO_2 and NO_2 as well as regions of high amine temperatures like the regeneration unit. Clearly, all these factors only determine the severity and extent of the corrosion behaviour of the amines. Additionally, different amines respond to these factors to different extents based on their unique structural differences. Therefore, a fundamental knowledge of the inherent corrosion behavior of amines based on their structures is crucial to the ability to select an amine solution for use in the design of the CO_2 capture process plant, for the purpose of providing a good understanding of how to manage corrosion issues under the typical CO_2 capture process conditions. It is noteworthy to mention that, it is necessary to fix the process conditions like CO_2 partial pressure, and temperature so as to see the true intrinsic effect of the amine structure on the amine corrosion behaviour. A literature survey on corrosion of conventional amines shows that there are limited corrosion studies that relate the corrosiveness of amines with their inherent structural differences.

On the other hand, considering the complexity associated with using amine structure to predict the corrosion rate of amines, it is pertinent to use a machine learning tool that is not only capable to handle the complexity of the system, but most importantly, is suitable for handling the non-linearity in the system. Machine learning (ML) has been used for diverse applications in the CO_2 capture process and environmental problems. Different ML tools such as Artificial Neural Networks (ANN), Function Networks (FN), Gaussian Process Regression (GPR), Random Forest (RF) and Support Vector Machine (SVM) can be applied to predict specific parameters from readily available data without additional cost [18–22]. Although most of these tools can be applied to complex nonlinear systems, the Gaussian Process Regression (GPR) model is well known to handle such a high level of complexity and non-linearity in different systems, hence its selection for use in this work. As a probabilistic regression model, it gives an advantage of increasing resilience against learning errors [23]. Therefore, the main objective of this work is to investigate the relationships between amine corrosiveness and their unique structural properties and to use this to develop a GPR model, using machine learning to predict the corrosion rates of different amines based only on their structures. The amines used for this work will encompass a wide variety of amines with varying structural functional group features in order to account for a large variety of different structures. For this work, one condition where amine corrosion is affected, specifically absorber conditions, where the amine temperature is around 40 °C has been selected. The CO_2 partial pressure for this work is fixed at 15 % CO_2. The amines were loaded to their respective rich loadings at this fixed condition of temperature and partial pressure. These fixed conditions will allow one to correctly evaluate the effect of the different structural functional groups on the amine corrosion rate. This GPR model will be used to predict the corrosion rates of amines at the typical conditions of absorption temperature and rich CO_2 loading without resorting to experimental measurements. These results are presented and discussed in this paper. The innovation and contribution that this study bring are: 1. To establish the relationships between amine corrosiveness and their unique structural properties 2. To develop a user-friendly Graphical User Interface using the GPR model for making corrosion rate predictions to address the issue of complexity, tediousness and the time-consuming nature of using machine learning models for making predictions and 3. A new approach to replicating data which inherently eliminates the experimental margin of error of the model, widening its scope of accurate prediction and is only restricted to the error from training the model.

2. Theory

2.1. Gaussian Process Regression

According to Guo et al. [24], Gaussian Process Regression is a supervised machine learning method capable of generating probabilistic predictions in Gaussian distribution [24]. It describes a probability distribution over possible functions that fit a dataset, by assuming that these functions are jointly Gaussian distributed [25]. Compared with other machine learning methods, Gaussian process has the ability to compute the empirical confidence interval for each prediction using kernels. Research on this machine learning method has shown that it is capable of handling complex issues in predictive modeling straightforwardly [26]. An Exponential GPR model was developed for this work through the process of training the model with a training data set, plotting the behavior of the

algorithm and determining the root mean squared error (RMSE), the R-Squared value and the average absolute deviation (AAD). The Exponential GPR is a function space expression of a radial basis function regression model with infinitely many basis functions. It replaces inner products of basis functions with kernels. The Exponential GPR handles smooth functions well with minimal errors. The algorithm of the Exponential GPR is illustrated below [27].

Input.

1. A training data set of the form:

$$\{(x_i, y_i); i = 1, 2, \ldots, n\} \tag{1}$$

Where $x_i \in R^d and y_i \in R$.

2. A linear regression model of the form:

$$y = x^T\beta + \varepsilon \tag{2}$$

Procedure.

1. Let the given training data set of n points be in the form of:

$$\{(x_i, y_i); i = 1, 2, \ldots, n\} \tag{3}$$

Where $x_i \in R^d and y_i \in R$.

2. A linear regression model of the form:

$$y = x^T\beta + \varepsilon \tag{4}$$

3. The linear regression model, where $K(X, X)$ is parametrized looks as follows:

$$K(X,X) = \begin{pmatrix} k(x_1,x_1) & k(x_1,x_2) & \cdots & k(x_1,x_n) \\ k(x_2,x_1) & k(x_2,x_2) & \cdots & k(x_2,x_n) \\ \vdots & \vdots & \vdots & \vdots \\ k(x_n,x_1) & k(x_n,x_2) & \cdots & k(x_n,x_n) \end{pmatrix} \tag{5}$$

3. The Exponential GPR becomes

$$k(x_i, x_j|\theta) = {\sigma_f}^2 \, exp\left(-\frac{r}{\sigma_l}\right) \tag{6}$$

$$\text{Where} r = \sqrt{(x_i - x_j)^T (x_i - x_j)} \tag{7}$$

θ is the maximum a posteriori estimates. σ_f is the signal standard deviation.

3.1. Calculations

3.1.1. Surface tension

The surface tension of each amine solution was calculated using the equation:

$$\sigma = \gamma_{amine}\sigma_{amine} + \gamma_{water}\sigma_{water} \tag{8}$$

where γ_{amine} is the mole fraction of amine, γ_{water} is the mole fraction of water, σ_{amine} (mN/m^2) is the surface tension of the pure amine, σ_{water} (mN/m^2) is the surface tension of water and σ (mN/m^2) is the surface tension of the amine solution. σ_{amine} is calculated using equation (9) [28]

$$\sigma_{amine} = 0.0146\left(2.28\delta_D^2 + \delta_P^2 + \delta_H^2\right) M_{vol}^2 \tag{9}$$

where δ_D is the dispersion partial solubility parameter (MPa$^{0.25}$), δ_P is the polar partial solubility parameter (MPa$^{0.25}$), δ_H is the hydrogen bonding partial solubility parameter (MPa$^{0.25}$), and M_{vol} is the molar volume (m3/mol). δ_D, δ_P and δ_H are calculated using equations (10)–(12) [29]

$$\delta_D = \frac{\sum F_{d_i}}{\sum V_i} \tag{10}$$

$$\delta_P = \frac{\sqrt{\sum F_{p_i}^2}}{\sum V_i} \tag{11}$$

$$\delta_H = \sqrt{\frac{\sum E_{h_i}}{\sum V_i}} \tag{12}$$

where F_{d_i} is in $((Jm^3)^{0.5}/mol)$, F_{p_i} is in $((Jm^3)^{0.5}/mol)$, and E_{h_i} is in (J/mol).

F_{d_i}, F_{p_i} and E_{h_i} are contributions to cohesive energy from dispersion forces, polarity and hydrogen bonds of a structural group *i* [29]. However, to determine these parameters the structural features of the amine must be classified. Using MEA as an example, the structural classifications are:

1. Number of $–NH_2$, n $(NH_2) = 1$
2. Number of –OH = 1
3. Alkyl length between $–NH_2$ and –OH (that is the number of $–CH_2$ groups) = 2

Therefore, F_{d_i}, F_{p_i} and E_{h_i} of MEA are obtained using equations (13)–(16) as follows:

$$\sum F_d = n(OH) \times F_d(OH) + n(NH_2) \times F_d(NH_2) + n(CH_2) \times F_d(CH_2) \tag{13}$$

$$\sum F_p = n(OH) \times F_p(OH) + n(NH_2) \times F_p(NH_2) + n(CH_2) \times F_p(CH_2) \tag{14}$$

$$\sum E_h = n(OH) \times E_h(OH) + n(NH_2) \times E_h(NH_2) + n(CH_2) \times E_h(CH_2) \tag{15}$$

$$\sum V = n(OH) \times V(OH) + n(NH_2) \times V(NH_2) + n(CH_2) \times V(CH_2) \tag{16}$$

The calculations are summarised in Tables 1 and 2 below.

3.1.2. Steric effect

(i) Steric effect for straight chain alkanolamines

The steric effect was calculated using the A-values (Table 3) of the adjacent functional groups attached to the amino group. Using ammonia NH_3 as the base structure (Fig. 1) to illustrate the steric effect calculations, it was established that NH_3 has the least steric effect based on the fact that there are three hydrogen atoms that make the –N atom easily accessible. When one or more hydrogen atoms are substituted with any other functional group, the –N atom becomes less accessible, meaning that this new substituent group creates a steric effect on the central –N atom. Using MEA as an example, MEA has two hydrogen atoms attached to the central –N atom and an ethanol group replacing the third hydrogen group of the NH_3 structure as shown in Fig. 1. Therefore, the steric effect is calculated using the A values of the ethanol group. In the same manner, for the other straight chain alkanolamines, if the amine had a primary amino group, it meant that it had two hydrogen atoms attached to the central nitrogen atom and therefore, the steric effect is determined based on the substituent functional groups attached to the other end of the –N atom.

The steric effect for MEA is therefore calculated using the structural classification described earlier in section 2.2a

$$Steric(MEA) = n(OH) \times A_{\text{value}}(OH) + n(C_2H_5) \times A_{value}(C_2H_5) \tag{17}$$

$$Steric(MEA) = 1 \times 0.87 + 1 \times 1.75 = 2.62 \tag{18}$$

Note: The number of C atoms of the substituent group was used as basis for using the A-values described in Table 3. For example, if the number of Cs for the substituent group is two, we use the A-value of ethyl group, if the -Cs are three we use the A-value of propyl and so forth. For example, in the case of MEA substituent group, the A-value for the ethyl group (C_2H_5) was used even though the

Table 1
Calculated parameters.

	Fd	Fp	Fh	V
OH	76.5	1225	6060	10
CH_2	469.2	0	0	32.2
NH_2	132.6	1176	11541.8	17.5
Total	678.3	2401	17601.8	59.6

Table 2
Summary of results.

	δ_D	δ_H	δ_P	M_{vol}	σ_{amine}	σ
MEA	11.36	17.17	28.44	60.36	46.35	71.96

Table 3
A_{values} of different functional groups [30–33].

Functional group	A_{value}
CH_3	1.7
CH_2CH_3	1.75
CH_2CH_2CH3	1.8[a]
$CH_2CH_2CH_2CH_3$	1.85[a]
$CH_2CH_2CH_2CH_2CH_3$	1.9[a]
OH	0.87[b]
NH_2	1.6[c]

[a] extrapolated values.
[b] the value of OH was used for O (for morpholine).
[c] the value of NH_2 was used for both NH and N.

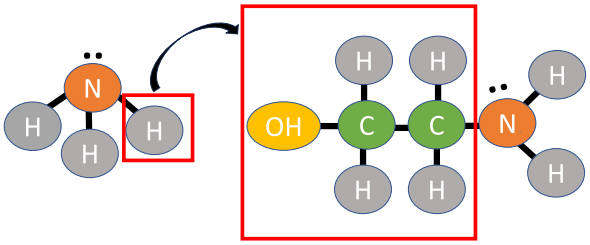

Fig. 1. Illustration on steric effect using NH_3 as the reference structure.

adjacent group is **(**C_2H_4).

(ii) Steric effect for non-straight chain alkanolamines

The steric effect for non-straight chain alkanol amines also followed the same procedure as that used for straight chain alkanol amines. We first of all had to describe the structural features and then used that with the corresponding A values to get the overall steric effect of the substituent group as shown in Fig. 2. Using AMP as an example, and using AMP's primary amine group structure as the reference point, the steric effect of the substituted group is calculated based on the classifications described below.

1. Number of $–NH_2$, n (NH_2) = 1
2. Number of –OH = 1
3. Alkyl length between $–NH_2$ and –OH (that is the number of $–CH_2$ groups) = 2 (C_2H_5)
4. Number of alkyl attached to the alpha C in between the amino group and –OH group = 2
5. Length of alkyl attached to the alpha C in between the amino group and –OH group = 1(CH_3)

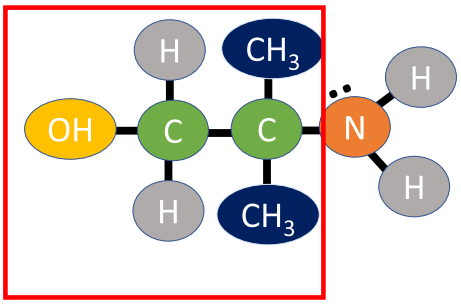

Fig. 2. Illustration on steric effect of substituent group in the AMP structure.

Therefore, the steric of AMP is calculated as:

$$Steric(AMP) = n(OH) \times A_{value}(OH) + n(C_2H_5) \times A_{value}(C_2H_5) + n(CH_3) \times A_{value}(CH_3) \quad (19)$$

$$Steric(AMP) = 1 \times 0.87 + 1 \times 1.75 + 2 \times 1.7 = 6.02 \quad (20)$$

(iii) Steric effect for multiamines

The steric effect for multiamines also followed the same procedure as that used for straight and non-chain alkanol amines. We first of all had to describe the structural features and then used that with the corresponding A values to get the overall steric effect. In the case of the multiamines, because they had two terminal primary amino groups, we used one of the terminal primary amino groups as the reference point and then found the steric effect of the entire functional group including the other terminal primary amino group on the reference group (see Fig. 3). Using EDA as an example the classification is shown below.

1. Number of primary amino groups = 2
2. Number of alkyl groups in between the two terminal primary amino groups = 1
3. Length of alkyl (number of Cs) in between the two terminal primary amino groups = 2 (C_2H_5)

Therefore, the steric of EDA is calculated as:

$$Steric(EDA) = n(NH_2) \times A_{value}(NH_2) + n(C_2H_5) \times A_{value}(C_2H_5) \quad (21)$$

$$Steric(EDA) = 1 \times 1.6 + 1 \times 1.75 = 3.35 \quad (22)$$

A multi amine like DETA has an additional secondary amino group and more than one alkyl group in between the amino groups. The classification is as ff.

1. Number of primary amino groups = 2
2. Number of secondary amino groups = 1
3. Number of alkyl groups in between the amino groups = 2
4. Length of alkyl (number of Cs) in between the two terminal primary amino groups = 2 (C_2H_5)

Therefore, the steric of DETA is calculated as:

$$Steric(DETA) = n(NH_2) \times A_{value}(NH_2) + n(NH) \times A_{value}(NH) + n(C_2H_5) \times A_{value}(C_2H_5) \quad (23)$$

$$Steric(EDA) = 1 \times 1.6 + 1 \times 1.6 + 2 \times 1.75 = 6.70 \quad (24)$$

(iv) Steric effect for cyclic amines

The steric effect calculation for cyclic amines followed a similar procedure used for multiamines. The secondary amino group was used as the reference point and the steric effect of the substituent group on the reference group (-NH) was calculated (Fig. 4). Using PZ as an example, the classification is as ff.

1. Number of secondary amino groups in cyclic structure = 2
2. Number of alkyl groups in between amino groups in cyclic structure = 2
3. Length of alkyl groups in between amino groups in cyclic structure = 2 (ethyl)

Therefore, the steric of PZ is calculated as:

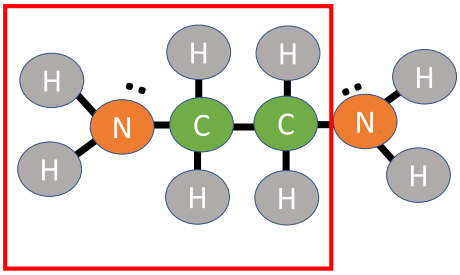

Fig. 3. Illustration on steric effect of substituent group in the EDA structure.

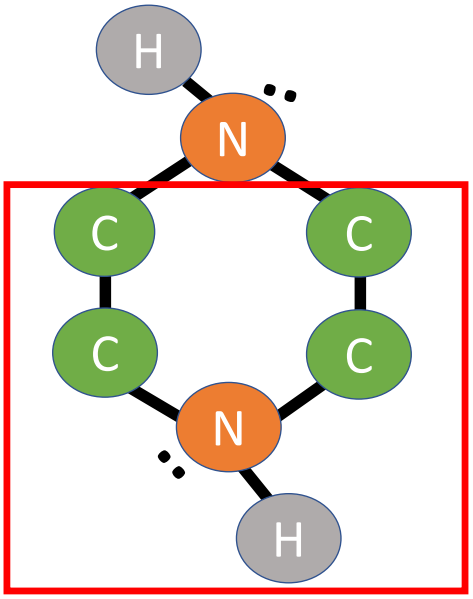

Fig. 4. Illustration on steric effect of substituent group in the PZ structure.

$$Steric(PZ) = n(NH) \times A_{value}(NH) + n(C_2H_5) \times A_{value}(C_2H_5) \tag{25}$$

$$Steric(PZ) = 1 \times 1.6 + 2 \times 1.75 = 5.1 \tag{26}$$

3.2. Corrosion mechanism

Corrosion is an electrochemical process involving oxidation and reduction reactions. Corrosion occurs when water molecules interact with the metal surface to cause ionization of the metal atoms, resulting in iron deposition normally known as rust.

Amines can accelerate corrosion depending on the concentration and amounts of oxidizing species which are typically product species formed from the interaction of CO_2 with amines (such as bicarbonates and carbamates) as well as O_2 itself all of which are present in the amine solution. A common contributor of internal corrosion are water corrosivity and corrosive gases such as CO_2 and H_2S [34] The rate of corrosion can be minimized by certain factors or key properties of amines. This is shown in Fig. 5 which shows how corrosion can be inhibited. These characteristics are: (1) the ability of the amine molecules to adsorb strongly onto the metal surface by their hydrophilic (polar) head group, (2) the bond between the amine molecule and the metal surface and (3) the alkyl tail length (hydrophobic group) of the amine molecules. 4) surface tension of the amine. The aforementioned factors will be used to explain the results obtained for this work. The equations governing the corrosion reaction is shown in equations (27)–(32).

Oxidation reaction:

$$Fe \rightarrow Fe^{2+} + 2e^{-} \tag{27}$$

Reduction reactions:

$$H^{+} + 2e^{-} \rightarrow H_2 \tag{28}$$

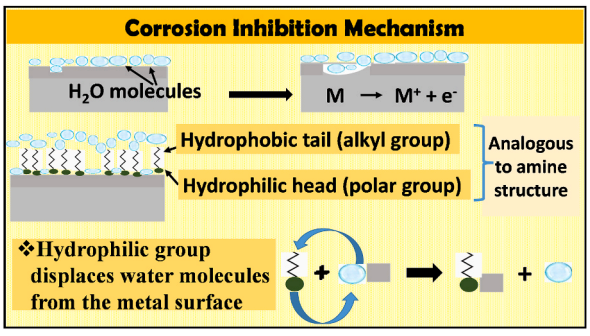

Fig. 5. Mechanism of corrosion inhibition.

$$O_2 + 2H_2O + 4e^- \rightarrow 4OH^- \quad (29)$$

$$2HCO_3^- + 2e\text{-} \rightarrow H_2 + 2CO_3^{2-} \quad (30)$$

Rust Formation:

$$Fe^{2+} + 2CO_3^- \rightarrow FeCO_3 \quad (31)$$

$$Fe^{2+} + 2OH^- \rightarrow Fe(OH)_2 \quad (32)$$

The product of corrosion, specifically ferrous hydroxide $Fe(OH)_2$, is highly unstable and when exposed to O_2 and water, oxidises to a trivalent hydrated ferric oxide ($Fe_2O_3.nH_2O$) or ferric hydroxide ($Fe(OH)_3$. In an oxygen limited environment, iron oxide (Fe_3O_4) is potentially formed instead of the trivalent corrosion products [35].

The surface tension of amines can also be related to the corrosion behavior of amines. Any molecule that contains a hydrophilic portion (polar) and a hydrophobic portion (non-polar) part and has the ability to adsorb to surfaces or interfaces, creating a marked decrease in the surface tension has the potency to have reduced corrosion behaviour. In other words, amines with lower surface tension have lower corrosion rates holding all other factors that impart corrosion constant.

4. Materials and methods

4.1. Materials

For alkanolamines studied, the amines used were: (1) alkanolamines: monoethanolamine (MEA, 99 %, Sigma-Aldrich), 4-amino-1-butanol (4A1B, 99.8 %, Chem Impex Int'l Inc.), 5-amino-1-pentanol (5A1P, 95 %, Sigma –Aldrich), 2-(methylamino)ethanol (MMEA, ≥98.5 %, Sigma-Aldrich), 2-(ethylamino)ethanol (EMEA, 98 %, Sigma-Aldrich), 2-(butylamine)ethanol (BMAE, 99 %, Sigma-Aldrich), diethanolamine (DEA, >98.5 %, Sigma-Aldrich), 2-dimethylaminoethanol (DMAE, >99 %, Sigma-Aldrich), N-methyldiethanolamine (MDEA, >99 %, Sigma-Aldrich), triethanolamine (TEA, 99 %, Sigma-Aldrich), 3-dimethylamino-1-propanol (3DMA1P, 98 %, Sigma-Aldrich), Butyldiethanolamine (BDEA, 98.6 %, Sigma Aldrich) Ethyldiethanolamine (EDEA, 98 %, Sigma-Aldrich), 4-dimethyl-amino-1-butanol (4DMA1B, >98 %, TCI America); (2) sterically hindered amines/branched chain amines: 2-amino-2-methyl-1-propanol (AMP, 99 %, Sigma-Aldrich), 2-amino-2-ethyl-1,3-propanediol (AEPD, 99 %, Sigma-Aldrich), 2-amino-2-hydroxymethyl-1,3-propanediol (AHMPD, >99.8 %, Sigma-Aldrich), tert-butylethanolamine (t-BAE, 98 %, Sigma-Aldrich), and tert-butyldiethanolamine (t-BDEA, 99 %, Sigma-Aldrich), 1-dimethylamino-2-propanol (1DMA2P, >98 %, Sigma-Aldrich); (3) multi-alkylamines: ethylenediamine (EDA, 98 %, Sigma-Aldrich), Methylethylenediamine (MEDA, 98 %, Sigma-Aldrich), 1, 3-diaminopropane (DAP, 99 %, Sigma-Aldrich), hexamethylenediamine (HMDA, 99 %, Sigma-Aldrich), diethylenetriamine (DETA, 98 %, TCI America), triethylenetetramine (TETA, technical grade, Sigma-Aldrich), and tetraethylenepentamine (TEPA, 60 %, TCI America); (4) cyclic-amine; piperazine (PZ, 99 %, Sigma-Aldrich), 1-methyl-piperazine (MPZ, >98 %, Sigma-Aldrich), 1-ethyl-piperazine (EPZ, >98 %, Sigma-Aldrich), 1-(2-aminoethyl)piperazine (EAPZ, >98 %, Sigma-Aldrich), morpholine (MPL, 99 %, Sigma-Aldrich), and 2-Hydroxy ethyl piperazine (HEP, >98 %, Sigma Aldrich). The structures of these amines are in supplementary information section Fig. S1. 1 N HCl (Fisher Scientific,

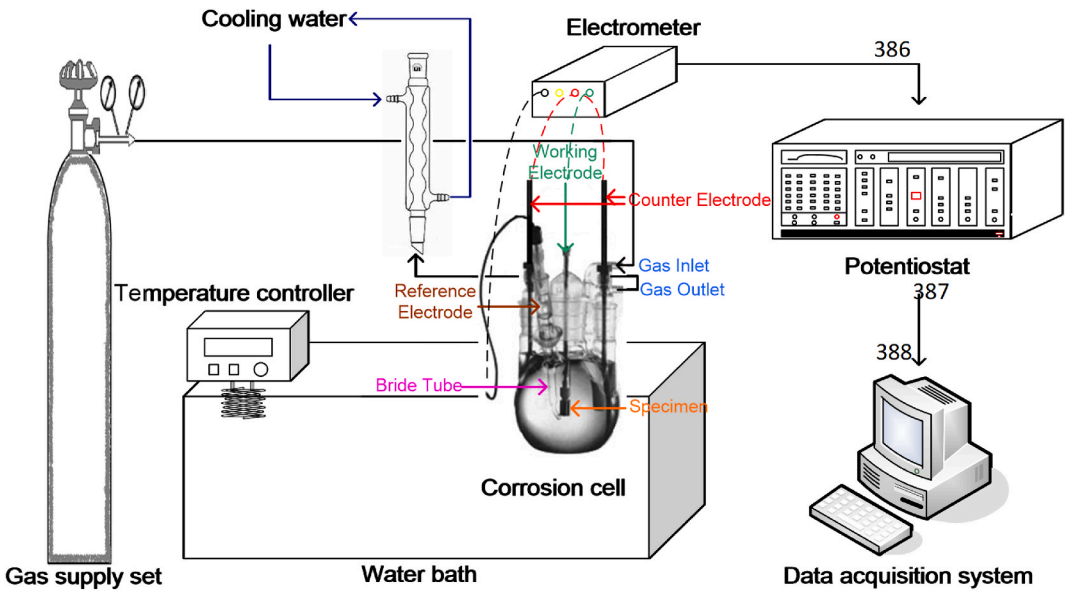

Fig. 6. Experimental set-up for corrosion measurement [36].

Canada) was used for acid titration to determine the actual amine concentration.

The premixed 15 % CO_2 (N_2 balance) used to preload the amine solutions as well as the 6 % O_2 (N_2 balance) gas were supplied by Praxair Inc., Regina, Canada. The Carbon Steel C1020 specimen (with chemical composition in % as follows: C, 0.19; Cr, 0.01; Cu, 0.01; Mn, 0.56; Mo, 0.01; N, 0.0036; Ni, 0.01; P, 0.009; S, 0.007; and Fe, balance) which was used to study the corrosion in the different amines was obtained from Metal Samples Company, AL, United States. The corrosion studies employed the use of the electrochemical technique following ASTM procedures. An ASTM corrosion cell model K47 which is a 1L flat-bottomed five-necked flask, was obtained from London Scientific, Ltd., ON, Canada. A potentiostat model 273A used to control the potential and for accurate current reading, as well as the PowerCORR version 2.47 Software used to acquire and analyze experimental data were also obtained from London Scientific, Ltd., ON, Canada.

4.2. Equipment

In this study, the electrochemical technique was used to study the corrosion behavior of carbon steel 1020 in different amine–CO_2–H_2O–O_2 systems. The corrosion experimental setup consisted of an ASTM corrosion cell, a water bath with a temperature controller, a potentiostat, a 6 % O_2 (N_2 balance) gas tank, a condenser, and a data acquisition system. The corrosion cell is made up of a working electrode mounted with the carbon steel specimen, two carbon graphite counter electrodes, and a mercury mercurous sulfate electrode (MSE), a salt bridge and a gas inlet for the transfer of gas. The potentiostat was used to control the potential and to read the current, whereas the data acquisition system, the PowerCORR version 2.47 was used to analyze the data obtained from the experiment. The carbon steel specimen was cylindrical in shape with a 3–48 threaded hole at one end. The condenser was used to maintain the amine concentration by condensing all water vapor in order to eliminate water evaporation from the amine solution. The schematic of the experimental setup is shown in Fig. 6.

4.3. Corrosion Measurement method

An experimental run started typically by first loading the amine solution (2 M) using 15 % CO_2 (N_2 balance) to its equilibrium CO_2 loading at 40 °C. All amines were fixed at a concentration of 2 M and loaded to their respective CO_2 equilibrium loadings at 40 °C. The CO_2 loaded amine solution contained in the corrosion cell was then immersed in the water bath whose temperature was kept constant at 40 °C. The flask with the amine solution was left in the bath until the desired temperature of 40 °C was achieved. The 6 % O_2 gas was then introduced at a constant flow rate of 150 ml/min for 1 h. During this 1-h period, the specimen was prepared by wet grinding it with 220 grit silicon carbide paper, followed by wet polishing with 600 grit silicon carbide paper. The polished specimen was then washed with distilled water, methanol, and then air dried and stored in a desiccator prior to use. The length and diameter of the specimen was measured with a Vernier caliper and then used to determine the specimen's surface area. At the end of the 1-h gas flow time, the counter electrodes were installed in the corrosion cell. Then the salt bridge was filled with the amine solution and placed in the test cell. The specimen was taken out of the desiccator, washed again with distilled water, followed by methanol. After then, it was mounted on the electrode holder, and then immersed into the corrosion cell. The wires were connected correctly and then the corrosion potential known as the E_{CORR} versus the MSE reference electrode of the amine system was measured for 1 h in order to ensure that the corrosion potential value remained constant. The actual electrochemical experiment was then started, in which the applied potential and the measured current were continuously measured, to generate a Tafel Plot. The Tafel Plot was generated by starting the scan from −250 mV to +250 mV versus the corrosion potential using a scan rate of 0.6 V/s. The final data was plotted as the applied potential versus the logarithm of the measured current. In order to obtain the corrosion current (i_{CORR}), the tangential lines of the anodic and cathodic curves were drawn to obtain the intersection point at the I_{CORR} as illustrated in our previous work [36] shown in Fig. 7. Once Icorr was determined, the corrosion rate was calculated using equation (33) below:

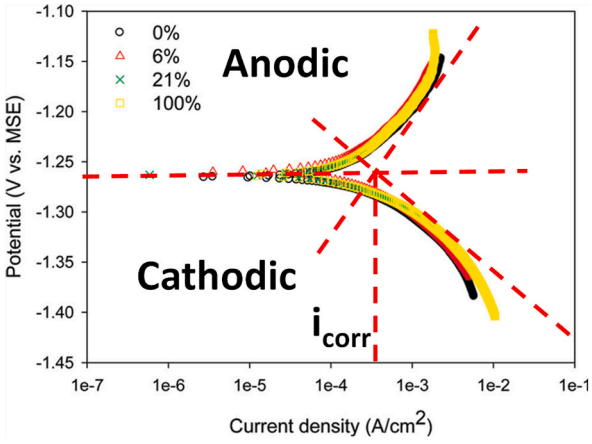

Fig. 7. A Hypothetical Tafel plot [36].

$$\text{Corrosion rate, mpy} = \frac{0.13 * I_{corr} * EW}{A * d} \quad (33)$$

where CR is the corrosion rate in mils per year, Icorr is the corrosion current (μA), EW is the equivalent weight of corroding species (g), A is the surface area of specimen (cm^2) and d is the specimen density (g/cm^2).

4.4. Viscosity measurement method

The viscosities of different amines were determined using the Lovis 2000 M/ME which operates using the rolling ball principle. All amines were loaded to their respective rich loadings at 40 °C and their viscosities were measured at 40 °C with a measurement error of ±2 %.

4.5. Results and discussion

The results are presented and discussed in two parts, namely; Part 1: Effect of structural functional groups for different classes/families or categories of amines on their corrosion behaviour and Part 2: Model development and Graphic User Interface for the prediction of amine based on their structural features.

4.6. Structure-corrosion behaviour of the amines studies

Studies on the effects of the amine structural functional groups for different classes of amines are discussed here.

4.6.1. Effect of alkyl chain length in 2° and 3° alkyl alkanolamines

The amines used to study the effect of the alkyl length of 2° alkyl alkanolamines were: MEA, MMEA, EMEA and BMEA and MDEA, EDEA and BDEA for 3° alkyl alkanolamines. The results obtained are presented in Fig. 8 which shows clearly that as the alkyl chain length increased, the corrosion rate reduced considerably. That is the corrosion rates for the alkanolamines followed the order MEA > MMEA > EMEA > BMEA for 2° alkyl alkanolamines and the order MDEA > EDEA > BDEA in the case of 3° alkyl alkanolamines. Table 4 shows that the surface tension of the amines showed a similar reducing order as MEA > MMEA > EMEA > BMEA and MDEA > EDEA > BDEA. It is important to note that in order to isolate the effect of the alkyl length, the same alkanol group (the ethanol group) was maintained. Therefore, the trend as seen in the results is purely from the variation in the length of the alkyl group. According to the literature [37–41], there exists a linear relationship between the alkyl tail length and the corrosion mitigation efficiency. An increase in the alkyl length results in increased hydrophobicity which repels water molecules more effectively from the metal surface thereby protecting it from molecules that can cause corrosion.

A group of researchers [42] evaluated the effect of the alkyl tail length of four different corrosion inhibitors with the same head group and found out that as the alkyl tail length increased, corrosion inhibition increased. These researchers correlated the alkyl tail length of the corrosion inhibitor molecule to the changes in the activation energies of the electrochemical process associated with the corrosion. Their results revealed that the longer the alkyl tail the higher the activation energy, which suggest that it becomes difficult for metal atoms to cross the energy barrier to form ions in solution. Therefore, using the same analogy, an increase in the alkyl length of the alkanol amines studied in this work, increased the water repelling ability (hydrophobicity) of the amine, which in effect, increased the ionisation energy of the metal, thus making it difficult for metal dissolution to occur. This water repelling characteristics of the amine can also be seen in the surface tension results (Table 4), indicating that the ability to displace water molecules from the surface reduces the surface tension of these amines, which thereby reduces the metal-water surface interactions and consequently reduces the corrosion rates.

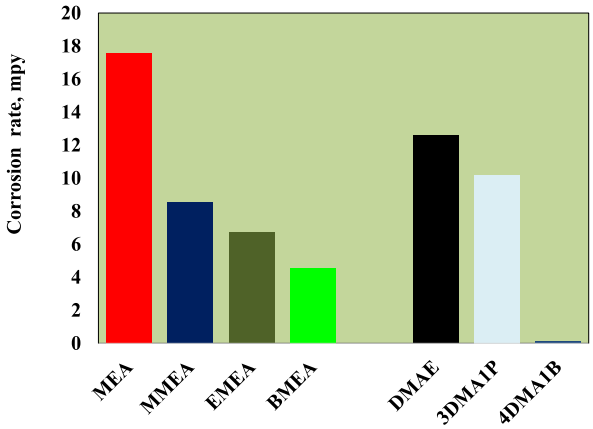

Fig. 8. Alkyl chain length in 2° and 3° alkyl alkanolamines.

Table 4
Summary of corrosion rate, surface tension and equilibrium loadings for amines studied.

Amine	Loading, mol CO_2/mol amine	Surface Tension, mN/m^2	Viscosity, mPas	Steric effect, kcal/mol	Corrosion rate, mpy
MEA	0.55	71.96	1.276	2.62	17.59
MMEA	0.58	71.10	1.524	4.32	8.56
EMEA	0.62	70.91	1.863	4.37	6.71
BMEA	0.57	70.69	2.977	4.47	4.58
tBMEA	0.7	70.45	3.062	7.77	9.03
tBDEA	0.42	71.76	3.518	10.39	4.3
MDEA	0.42	71.47	2.064	6.94	13.81
EDEA	0.45	71.25	2.379	6.99	10.37
BDEA	0.32	70.89	3.048	7.09	5.61
4A1B	0.69	71.28	1.770	2.72	12.19
5A1P	0.64	71.07	2.084	2.77	8.07
DEA	0.51	72.05	1.827	5.24	10.79
TEA	0.32	72.50	2.240	7.86	20.7
AMP	0.65	71.05	2.085	6.02	18.3
AEPD	0.35	71.85	1.856	5.99	13.29
AHMPD	0.2	74.79	1.828	7.86	0.14
DMAE	0.45	70.77	1.642	6.02	12.6
3DMA1P	0.88	70.64	2.025	6.07	10.19
4DMA1B	0.9	70.52	2.510	6.12	0.15
1DMA2P	0.9	70.60	2.097	7.72	21.47
EDA	0.9	72.57	1.605	3.35	40.28
TMDA	1.06	72.04	2.190	3.40	23.54
HMDA	1.32	71.08	6.103	3.55	14.49
PZ	0.74	71.43	1.936	5.10	1.87
1-MPZ	0.57	70.80	2.507	6.80	1.35
1-EPZ	0.65	70.68	2.981	6.85	0.35
MORPHOLINE	0.35	71.36	1.468	4.37	4.77
HEP	0.8	71.15	3.438	7.72	7.13
EAPZ	1.17	71.07	7.231	8.45	4.78
DETA	1.25	71.71	3.839	6.70	7.78
TETA	1.72	71.37	13.035	10.05	7.27
TEPA	1.89	71.13	25.875	13.4	1.65
MEDA	0.98	71.12	1.875	5.05	43.27

4.6.2. *Effect of alkyl chain length in 1° and 3° alkanolamines*

The amines used to study the effect of the alkyl length of 1° alkanolamines were: MEA, 4A1B and 5A1P and DMAE, 3DMA1P and 4DMA1B for 3° alkanolamines. For the primary alkanol group the alkyl length between the primary amino group and the –OH group was the variable of interest. For the case of MEA the alkyl length was 2, and then increased to 4 and 5 in 4A1B and 5A1P respectively. The corrosion rate reduced in the order MEA > 4A1B > 5A1P as the alkyl length increased as shown in Fig. 9. In the case of the tertiary alkanolamines, the number and length of the attached end alkyl group (that is the 2 end methyl groups) remained fixed whilst the alkyl length between the tertiary amino group and the –OH group increased from to 2 (DMAE) to 3 (3DMA1P) to 4 (4DMA1B). Increasing the alkyl length reduced the corrosion rate in the order DMAE > 3DMA1P > 4DMA1B.

A similar reasoning used to explain the role that the alkyl plays in reducing the corrosion rate for section 4.1.1 can be applied here; increasing the alkyl length results in increased hydrophobicity which repels water molecules more effectively from the metal surface thereby protecting it from corrosion. Surface tension also showed a similar trend in the same order as that of the corrosion rates; this

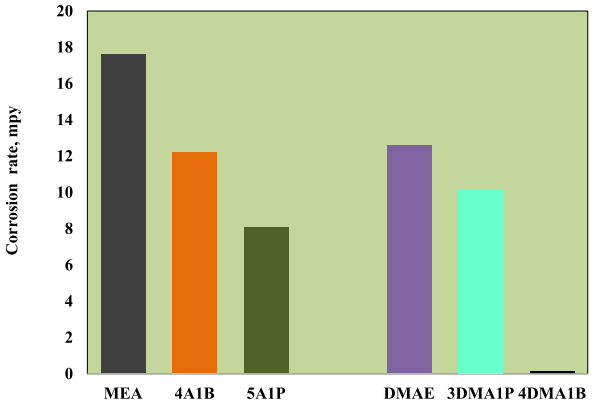

Fig. 9. Alkyl chain length in 1° and 3° alkanolamines.

further confirms that for the same family of amines, the amines with the lower surface tensions have minimal metal surface reactions, thereby having lower corrosion rates. A comparison of the rich loadings for the tertiary alkanolamines showed an increasing trend. Although an increase in CO_2 loading is seen to affect corrosion rates adversely because of high concentrations of oxidizing agents such as HCO_3^-, this effect is masked by the actual structural effect of the amines. This also buttresses the point that it is important to study the inherent corrosion behavior of amines based on their structures.

4.6.3. Effect of alkyl chain length diamines

The amines used to illustrate this structural feature are: EDA, TMDA and HMDA. For this group the two terminal primary amino groups were fixed, while the alkyl length between the end amino groups varied from 2 (EDA) to 3 (TMDA) to 6 (HMDA). Fig. 10 shows that the corrosion rate reduced in the order EDA > TMDA > HMDA which also further confirms the role the increasing alkyl length plays in reducing the corrosion rate. The surface tension results also followed the order: EDA > TMDA > HMDA. The same underlying reasons attributed to the above-mentioned sections with regard to the surface tension and corrosion rate relationship for the same family of amines, is applicable here as well. Based on Table 4, even though the loadings of EDA, TMDA and HMDA increases with increasing alkyl length, the corrosion rate reduces. This reveals that the structural effect supersedes the loading effect.

4.6.4. Effect of alkyl substituent length in cyclic amines

The effect of the alkyl substituent in cyclic amines was studied using PZ, 1-MPZ and 1-EPZ. For this category of amines, the H- atom attached to the secondary amine group in PZ is replaced with a methyl group as seen in the structure of 1-MPZ, and then the methyl is increased to an ethyl group as seen in 1-EPZ. Fig. 11 shows that replacing the H-atom with a methyl group decreased the corrosion rate and as the alkyl length was increased from 1 to 2 the corrosion rate further reduced. The hydrophobic nature of the alkyl group allows it to repel water more effectively to limit water-metal surface interactions leading to corrosion as already mentioned in the earlier sections. The surface tension results as shown in Table 4 further confirms the surface tension -corrosion inter-relationship.

4.6.5. Effect of the number of alkylene amine groups in multiamines

The effect of increasing the number of alkylene secondary amino groups was studied using DETA, TETA and TEPA. These three amines are multi amines with two terminal primary amino groups but varying number of ethylene secondary amine groups. DETA has two ethylene groups with an attached secondary amine group, TETA has three ethylene groups with two attached secondary amine groups, and TEPA has four ethylene groups with three attached secondary amine groups. The results in Fig. 12 showed that increasing the ethylene groups as well as the number of secondary amine groups reduced the corrosion rate. The works done in the literature [43, 44] showed that –N atom is an active adsorption site capable of bonding strongly with the metal using its polar head group through electron transfer to protect the metal from corrosion. In addition, increasing the ethylene groups also introduces more hydrophobicity which further reduces the metal surface interactions. Increasing the active adsorption sites coupled with the increased hydrophobicity provides the synergy in reducing corrosion significantly. The results are further buttressed by the surface tension results shown in Table 4. The trend shown in Fig. 12 is another good example of how the structural effect dominates the loading effect.

4.6.5.1. Effect of the number of –OH and alkyl substituent groups and -in sterically hindered amines. The structural effect of the number of –OH and alkyl groups made use of the following sterically hindered amines AMP, AEPD and AHMPD. AMP has one primary amino group and two methyl groups attached to the alpha carbon and one –OH group. A critical look at the AEPD structure shows that it also has one primary amino group, but instead of two methyl group (as shown in AMP structure), it is replaced with an ethyl group and an extra ethanol (-OH) group attached to the alpha carbon. Through this substitution with a longer alkyl length and an extra ethanol (-OH group), the corrosion rate reduces as shown in Fig. 13. Increasing the length increases the hydrophobicity as already explained earlier and adding an extra –OH group also adds an additional active adsorption site that can adsorb onto the metal surface to reduce the

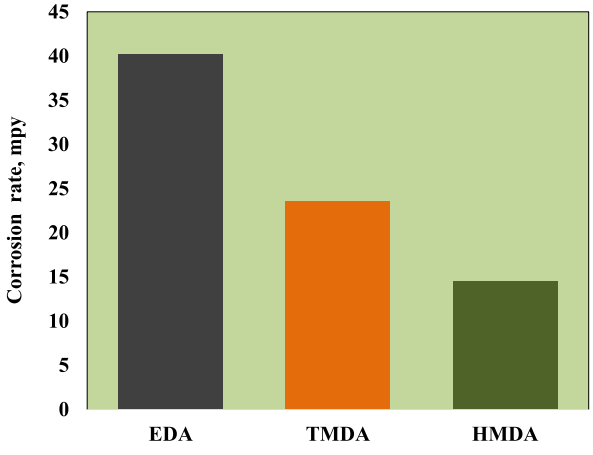

Fig. 10. Alkyl chain length in diamines.

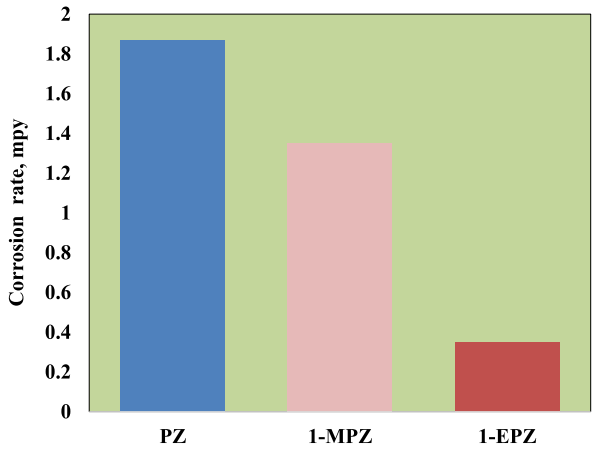

Fig. 11. Alkyl substitution in cyclic amine.

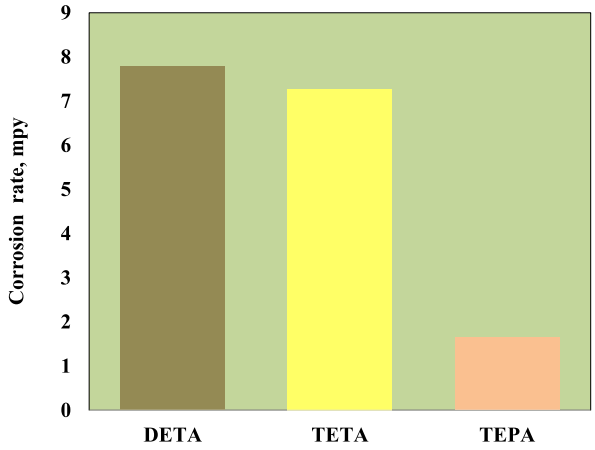

Fig. 12. Number of secondary amine groups and ethylene group effect in multiamines.

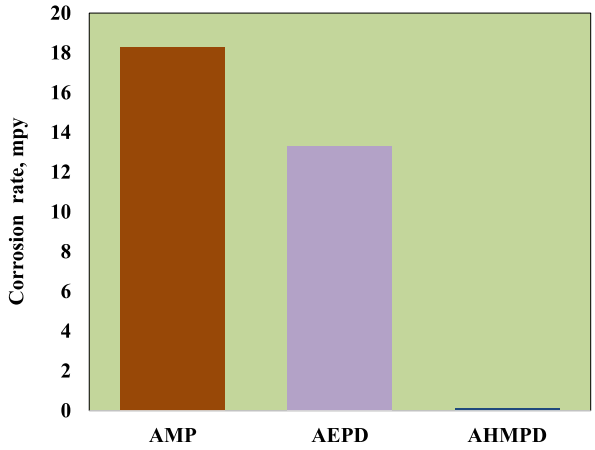

Fig. 13. Effect of –OH and alkyl substitution in sterically hindered amines.

corrosion rate. These coupled effects results in the reduction in corrosion rate. In the case of AHMPD, it has one primary amino group and two ethanol (-OH groups) attached to the alpha carbon. Comparing AHMPD to AEPD, the extra ethanol group in AHMPD is substituted for the ethyl group present in AEPD. This replacement of the ethyl group with the ethanol group drastically reduces the

corrosion rate as seen in Fig. 13. Even though the ethyl group has the potency to minimize corrosion, the results show that the –OH group has a stronger contributory effect in corrosion inhibition than the alkyl group. A comparison of the CO_2 loading for these amines shows that the loading decreases from AMP to AEPD to AHMPD. This may also be a contributing factor to the results. As such, the structural effect coupled with the loading effect both reduce the corrosion rates.

4.6.6. *Effect of substituent groups in cyclic amines*

To study the effect of the –O substituent group, the amines used are PZ and Morpholine. PZ is a cyclic amine with two secondary amino groups. Morpholine has one secondary amine and an –O group. Comparing these two structures as shown in section 2.1, it can be seen that the second secondary amino group present in PZ is replaced with an –O group in the structure of Morpholine. The results show that replacing the amino group with the –O atom increases the corrosion rate (Fig. 14). Even though both groups (-NH and –OH) are polar groups that act as active adsorption sites, the results reveal that the amine group has a stronger inhibition effect than the –O atom. The loadings of these two amines (PZ > Morpholine) also reveals the dominating contribution of the structural effect. A look at the surface tension of these amines shows an opposite trend to the corrosion rate results. This inconsistency can be attributed to other contributing factors.

Similarly, when comparing the structures of EAPZ and HEP, the former has one secondary amino group, one tertiary amino group and one primary amino group, whereas the latter has the same features but the only distinguishing feature is that the primary amino group is substituted for an –OH group. The results (EAPZ < HEP) as shown in Fig. 14 also reveal the stronger corrosion inhibition effect of the $–NH_2$ group over the –OH group.

4.7. *Model development*

Model development involved the combination of all the different structures of amines with their corresponding corrosion data which led to the development of a GPR model that can help to predict the corrosion rates of amines once their structures are known. Following the GPR model development is the development of a Graphical User interface that allows one to specify the structural features of the amine and the resulting corrosion rate of the amine is given as feedback to the user.

4.7.1. *Feature selection and data visualization*

Representing all 33 amines used in this study, the total number of data points obtained for the development of the model was 33 with 19 variables as shown in Table S1 (supplementary information section). However, only 3 of the variables were independent variables namely, amine viscosity, surface tension and steric effect. The remaining 16 variables were classification variables. These 16 classification variables were used to describe the different structural functional groups of the amines. These structural features are listed as variables X1 to X16 in Fig. S1 of the supplementary section. It is important to mention that, these 16 features which describe the structure of the amines were used to calculate surface tension and steric effect. Data visualization was employed to further understand the relationship between the three independent variables and the corrosion rate by plotting corrosion rate against these features. Figs. 15–17 show the relationship between corrosion rate and amine viscosity, surface tension and steric effect, respectively.

4.7.2. *Corrosion vs amine viscosity*

Amine viscosity is a property of amine that shows the resistance of amine to flow. Viscosity poses mass transfer limitation which affects the rate of reaction, especially rate-based systems that involve the movement of material in and out of the amine system. Therefore, corrosion, a rate-based mechanism, although governed by electrochemical reactions can be impacted by mass transfer limitations. It was expedient to visualize the data to investigate the relationship between corrosion rate and amine viscosity. Fig. 15 shows a plot of the corrosion rate versus amine viscosity. A blown-out portion of the data shows an irregular trend in the data. From the

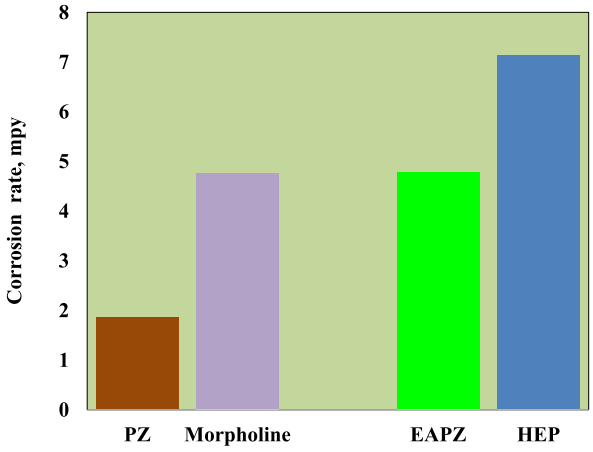

Fig. 14. Effect of –OH and alkyl substitution in sterically hindered amines.

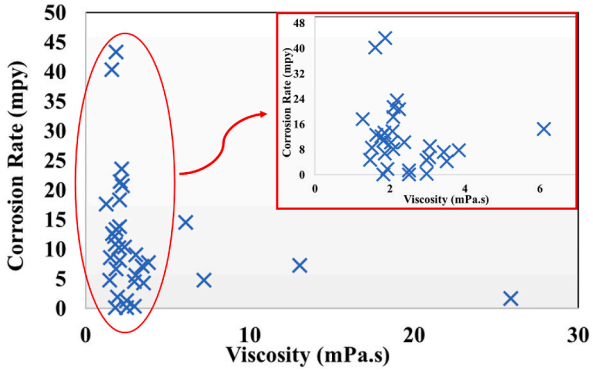

Fig. 15. Corrosion rate vs. Amine viscosity.

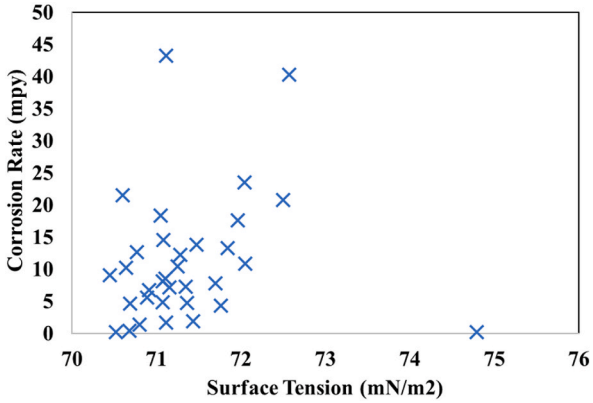

Fig. 16. Corrosion rate vs. Surface tension.

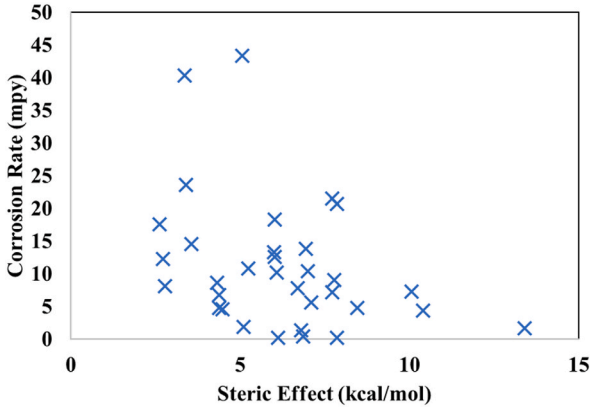

Fig. 17. Corrosion rate vs. Steric effect.

graph, it can be concluded that the relationship between the corrosion behavior and the amine viscosity for different structures of amines is not linear; therefore, a simple linear regression cannot be used. This stresses the need for a model that can handle the complexities of these inter-relationships.

4.7.3. Corrosion vs. surface tension

It was also important to visualize the data to find out the relationship between the surface tension and the rate of corrosion. As already discussed in section 4, based on the structure-corrosion relationship of different families of amines, there existed a direct correlation between the surface tension of amines and their corrosion rates for the same family of amines. Even though there were a

few outliers, for the most part, amines that had low surface tension had low corrosion rates and vice versa when comparing amines belonging to the same family. When different types of amines from different categories, classes or families are combined together and their corrosion data versus surface tension is plotted, the resulting graph shows a non-linear complex relationship as seen in Fig. 16. Therefore, it was necessary to explore the use of a machine learning model to deal with these complexities.

4.7.4. Corrosion vs. steric effect

Fig. 17 shows the plot of corrosion rate against the steric effect. The steric effect shows the impact of the substituent group on with the primary secondary or tertiary amine group. The bulkiness of the substituent group as well as the type significantly affects the steric impact. As already discussed in section 4, based on the structure-corrosion relationship of different families of amines, there existed a direct correlation between the surface tension of amines and their corrosion rates for the same family of amines. However, plotting all the different structures together as seen in Fig. 17 shows an irregular trend which supports the reason for an appropriate and most suitable model for this kind of interrelationships.

4.7.5. Data splitting and replication

The data was split into 3 groups. 22 were used for training/validation of model, 9 for testing the trained model and 2 for additional testing of the model to confirm the accuracy and generalizability of the final model obtained. The training/validation dataset was then replicated to generate enough data to train the model. The replication was done by using the experimental error associated with the reported values for amine viscosity and corrosion rate. The percentage errors associated with the experimentally determined amine viscosity and corrosion rate were 2 % and 11 % respectively. The raw data was then replicated to include amine viscosity values of 1 %, 1.5 % and 2 % above and below the reported values and corrosion rate values of 3 %, 7 % and 11 % above and below the reported values. This innovative method of replication gives flexibility and versatility to the model and allows it to be able to predict the corrosion rate while accounting for the error associated with the experimental measurements of the viscosities and corrosion rates. The experimental error of 11 % and 2 % for the corrosion rates and viscosity respectively implies that any time the corrosion rates are measured experimentally, the true rates lie within 11 % and 2 % respectively of the values in Table 4. For example, the corrosion rate for MEA as reported in Table 4 is 17.59. However, the true corrosion rate can lie between 15.66 and 19.52 as captured in Table S2 (Supplementary information). Similarly, the viscosity as reported in Table 4 is 1.276. However, the true viscosity can lie between 1.2505 and 1.3015 as captured in Table S2. By this method of replicating the data, the model inherently eliminates the experimental margin of error widening its scope of accurate prediction of the true corrosion rates, and is only restricted to the error from training the model. An example of the replication process is shown for MEA in Table S2 (Supplementary information).

The data on the first row of Table S2 shows the actual experimental data obtained for the viscosity and corrosion rate for MEA. The surface tension and steric effect were calculated based on the structure of MEA and data available in literature and therefore had no errors associated with them. The viscosity data however, was experimentally determined with an error margin of 2 %. The data was therefore replicated using values within the viscosity error margin and mapping each value to a corrosion rate which also lied within an error margin of 11 %. This approach allows for the trained model to predict MEA corrosion rate within 11 % error margin if the value of the experimentally determined viscosity lies within an error margin of 2 %. It therefore accounts for the error or deviations associated with experimentally determining the viscosity of the amine as an input to the model for prediction. This means that the error associated with predicted values from the model is only restricted to the error from training the model, eliminating the error that could be linked with determining the viscosity. Replicating the data using this approach resulted in 7 different data points for each of the 22 amines in the training dataset, generating enough data points to train the model.

4.7.6. Feature transformation

Before training the model, function transformers i.e., log transform and reciprocal transform were used to transform the training dataset from one form to another while keeping the essence of the data. This was done to ensure normal distribution of the data before applying the machine learning algorithm to it for better performance.

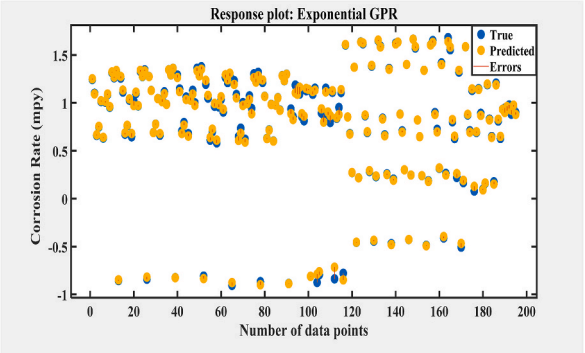

Fig. 18. Response plot for trained GPR model.

4.7.6.1. Model development. A Gaussian Process Regression (GPR) model was trained using the Regression Learner tool in the Statistics and Machine Learning Toolbox in MATLAB software with 10-fold cross-validation and later exported to make predictions on new data. The hyperparameters obtained for the trained model were isotropic exponential kernel function, constant basis function, automatic kernel scale, signal standard deviation and sigma with standardization. The predictive accuracy of the trained model was determined by using the model to predict the response variables from the training dataset and comparing the output of the model to the true values. The response plot generated in Fig. 18 showed a very good performance of the model in terms of predicting the output of the data it was trained with. Most of the predicted output from the model (in yellow) coincided with the true values (in blue), showing very good predictability of the trained model.

With an R^2 value of 0.99 for the validation set and RMSE of 0.028, the trained model showed very good correlation between predictor and response variables and good predictive ability, respectively. The parity plot for the validation set is shown in Fig. 19.

The model also returned an R^2 value of 0.99 and RMSE of 0.016 for the test dataset, indicating a good correlation between predictor and response variables and very good predictive accuracy of the model respectively. The parity plot for the test data is shown in Fig. 20. The Average Absolute Deviation (AAD) between the experimental values and the model predicted values was determined to be 4.26 %. To further confirm the performance of the model, two extra test data points unknown to the model were used to test the model. The AAD between the experimental corrosion rate and the model predicted corrosion rate was found to be 5.32 % indicating the model's ability to generalize well and make accurate predictions on unseen data. The model's incredible performance on the unseen data was also an indication of the absence of overfitting, making it very reliable for a wide range of predictions as overfitted models would only predict the known data accurately but behave very poorly towards unseen data.

It is noteworthy to mention that the resulting output from the developed model will be limited to the conditions under which the variables were obtained. This means that when this model is used, the predicted corrosion rates will be for amines at 2 M concentration, at their typical rich loading at 40 °C. This prediction helps us in the screening and selection of amines to use in the design of a CO_2 capture plant. Nevertheless, it is known that amine related corrosion increases with CO_2 loading, amine concentration and temperature. Therefore, if conditions are higher than those stipulated in this study, the amine corrosion rates will increase correspondingly. This extra step is needed for the amine selected based on this model. This model also required information on the viscosity of the loaded amines at 40 °C. Viscosity of CO_2 loaded amines was determined experimentally (as used in this study). However, viscosity can also be determined from well-known existing correlations [45]. However, once the viscosity is obtained, the other two main variables, surface tension and steric effect, can be easily obtained all of which are used with the user interface to estimate the corrosion rate accurately.

4.7.7. Development of Graphical User Interface for prediction

One of the limitations of regression models developed using machine learning approach is the complexity involved in using these models for predictions. For instance, it is very tedious and time consuming to extract the weights from a trained neural network model to make new predictions. Sometimes, it is also required that one understands how these models are developed and exported to make predictions on new data in order to apply them. On the other hand, conventional regression models developed using Excel and Minitab mostly have equations that the user can easily access and use, thus making predictions on new data very easy. To make predictions easy using the GPR model that was trained, a user-friendly Graphical User Interface was developed in this work. The 16 structural classification variables (that were not used in developing the GPR model) were used in the development of this interface to help calculate the steric effect and surface tension and use them to predict the corrosion rates.

To predict the corrosion rate of an amine, the interface allows a user to define the structure of the amine and use that to calculate

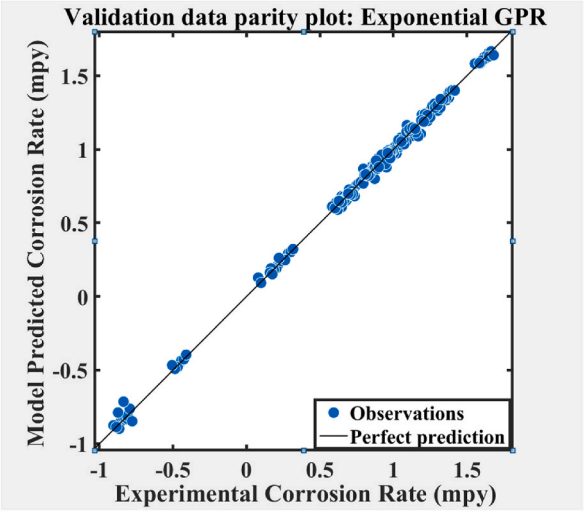

Fig. 19. Parity plot for validation set.

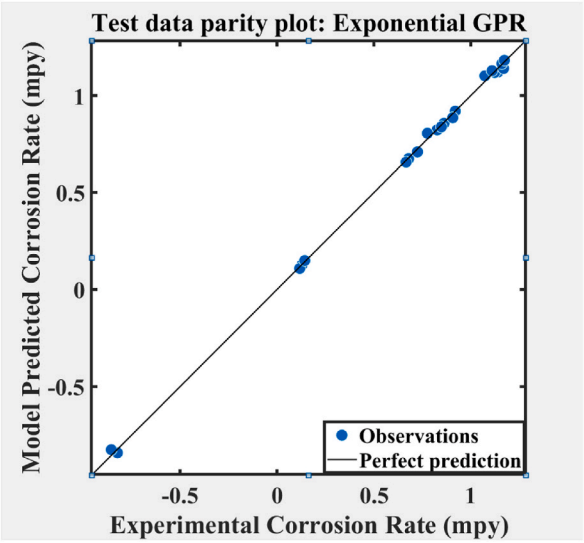

Fig. 20. Parity plot for test data.

the surface tension and steric effect. The user is also required to provide the viscosity, density and molecular weight of the amine. With the surface tension and steric effect calculated, as well as knowledge of viscosity, the corrosion rate of the amine can be predicted. The classification variables for the straight/branched chain amines were different from the variables for the multi/cyclic amines. This necessitated the development of two tabs in the interface, each tab predicting the corrosion rate of the respective amine based on the unique structural features. The use of the interface in predicting the corrosion rate is illustrated in Figs. S2 and S3 respectively for AMP and EAPZ as examples.

5. Conclusions

- Increasing the alkyl length of alkanolamines, multi amines and cyclic amines led to a reduction in the corrosion rate due to an increase in the hydrophobicity which limited the rate of corrosion
- Increasing the number of –OH group minimized the rate of corrosion
- Results showed that the $–NH_2$ group has a stronger mitigation effect than the –OH group
- A new approach to replicating data has been developed which inherently eliminates the experimental margin of error of the model, widening its scope of accurate prediction and is only restricted to the error from training the model.
- The developed model for predicting the corrosion rate of amines (whose concentration is 2 M and their rich CO_2 loading is at a fixed temperature of 40 °C and a fixed CO_2 partial pressure of 15 %) returned an R^2 value of 0.99 for the validation set and RMSE of 0.028; the trained model showed very good correlation between predictor and response variables and good predictive ability respectively.
- The developed model for predicting the corrosion rates of amines (whose concentration is 2 M and their rich CO_2 loading is at a fixed temperature of 40 °C and a fixed CO_2 partial pressure of 15 %) returned an R^2 value of 0.99 and RMSE of 0.016 for the test dataset
- The Average Absolute Deviation (AAD) between the experimental values and the model predicted values was determined to be 4.26 % for the test data.
- The AAD of 5.32 % obtained for the two extra test data points unknown to the model shows the model's ability to generalize well and make accurate predictions on unseen data.
- The corrosion model for the prediction of amines (whose concentration is 2 M and their rich CO_2 loading is at a fixed temperature of 40 °C and a fixed CO_2 partial pressure of 15 %) has good predictability, versatility and flexibility and allows one to easily predict the corrosion rates based on the amine structure and can be used in the screening and selection of amines for use in the design of a CO_2 capture plant.
- To address the issue of complexity, tediousness and the time-consuming nature of using machine learning models for making predictions, this work developed a user-friendly Graphical User Interface using the GPR model that was trained to make predictions
- This model requires information on the viscosity of the loaded amines at 40 °C as an input variable.
- The resulting output from the developed model will be limited to the conditions under which the variables were obtained. That is, when this model is used, the predicted corrosion rates will be for amines at 2 M concentration, at their typical rich loading at 40 °C.
- Having used the model to screen and select the amine with the least corrosion rate, further research should focus on determining the effect of other operational parameters such as increased CO_2 loading, higher amine concentration and higher temperature

Declarations

- Informed consent was not required for this study because the work did not involve humans and biohazards.

Data Availability statement

Data will be made available on request.

CRediT authorship contribution statement

Jessica Narku-Tetteh: Formal analysis, Investigation, Methodology, Validation, Writing – original draft, Writing – review & editing. **Ebenezer Mensah:** Formal analysis, Investigation, Methodology, Validation, Writing – original draft, Writing – review & editing. **Pailin Muchan:** Formal analysis, Investigation, Methodology, Validation, Writing – original draft. **Teeradet Supap:** Formal analysis, Investigation, Methodology, Supervision, Validation, Writing – original draft, Writing – review & editing. **Supranee Lisawadi:** Formal analysis, Funding acquisition, Supervision, Writing – review & editing. **Raphael Idem:** Conceptualization, Funding acquisition, Supervision, Validation, Writing – review & editing.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:Raphael Idem reports financial support was provided by Bualuang ASEAN Chair Professor Fund, Thammasat University, Bangkok, Thailand. Raphael Idem reports financial support was provided by Natural Sciences and Engineering Research Council of Canada. Raphael Idem reports financial support was provided by SaskPower Clean Energy Research Chair.

Acknowledgements

This study was supported by Bualuang ASEAN Chair Professor Fund. Also, the financial support provided by Natural Science & Engineering Research Council of Canada (NSERC), and the SaskPower Clean Energy Research Chair program is gratefully acknowledged.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2023.e22050.

References

[1] A. Veawab, P. Tontiwachwuthikul, A. Chakma, Corrosion behavior of carbon steel in the CO_2 absorption process using aqueous amine solutions, Ind. Eng. Chem. Res. 38 (1999) 3917–3924.
[2] S.A. Mazari, L. Ghalib, A. Sattar, M.M. Bozdar, A. Qayoom, I. Ahmed, N.M. Mubarak, Review of modelling and simulation strategies for evaluating corrosive behavior of aqueous amine systems for CO_2 capture, Int. J. Greenh. Gas Control 96 (2020), 103010.
[3] C. De Waard, D. Milliams, Carbonic acid corrosion of steel, Corrosion 31 (1975) 177–181.
[4] Z. Feng, M. Jing-Wen, Z. Zheng, W. You-Ting, Z. Zhi-Bing, Study on the absorption of carbon dioxide in high concentrated MDEA and ILs solutions, Chem. Eng. J. 181 (2012) 222–228.
[5] L. Zheng, N.S. Matin, J. Thompson, J. Landon, N.E. Holubowitch, K. Liu, Understanding the corrosion of CO_2-loaded 2-amino-2-methyl-1-propanol solutions assisted by thermodynamic modeling, Int. J. Greenh. Gas Control 54 (2016) 211–218.
[6] Z. Feng, M. Jing-Wen, Z. Zheng, W. You-Ting, Z. Zhi-Bing, Study on the absorption of carbon dioxide in high concentrated MDEA and ILs solutions, Chem. Eng. J. 181 (2012) 222–228.
[7] A.L. Kohl, R. Nielsen, Gas Purifi, Gulf Professional Publishing, 1997.
[8] A. Benamor, M.J. Al-Marri, Modeling analysis of corrosion behavior of carbon steel in CO_2 loaded amine solutions, Int. J. Chem. Eng. Appl. 5 (2014) 353–358.
[9] Y.-S. Choi, S. Nešić, D. Duan, S. Jiang, Mechanistic Modeling of Carbon Steel Corrosion in a MDEA-Based CO_2 Capture Process, CORROSION 2012, NACE International, 2012.
[10] D. Duan, Y.-S. Choi, S. Jiang, S. Nešic, Corrosion Mechanism of Carbon Steel in MDEA-Based CO_2 Capture Plants, CORROSION/2013, 2013.
[11] D.A. Jones, Principles and Prevention of Corrosion, MacMillan, New York, NY, 1996, pp. 503–570.
[12] E. McCafferty, Introduction to Corrosion Science, Springer, New York, NY, 2013, pp. 359–400.
[13] A. Najumudeen, Development of a Mechanistic Corrosion Model for Carbon Steel in MEA-Based CO_2 Absorption Process. Faculty of Graduate Studies and Research, University of Regina, 2012.
[14] Y.-S. Choi, S. Nešić, D. Duan, S. Jiang, Mechanistic Modeling of Carbon Steel Corrosion in a MDEA-Based CO_2 Capture Process, Corrosion 2012, NACE International, 2012.
[15] A.F. Ciftja, A. Hartono, E.F. da Silva, H.F. Svendsen, Study on carbamate stability in the AMP/CO_2/H_2O system from 13C-NMR spectroscopy, Energy Proc. 4 (2011) 614–620.
[16] Y.-S. Choi, D. Duan, S. Jiang, S. Nešić, Mechanistic modeling of carbon steel corrosion in a methyldiethanolamine (MDEA)-based carbon dioxide capture process, Corrosion 69 (2013) 551–559.
[17] L. Ghalib, B. Si Ali, S. Mazari, W.M. Ashri, I.M. Saeed, Modeling the effect of piperazine on carbon steel corrosion rate in carbonated activated MDEA solutions, Int. J. Electrochem. Sci. 11 (2016) 4560–4585.

[18] R.J.E. Abrantes, Y.W. Mao, D.D.W. Ren, Rate coefficient function estimation using Gaussian process regression, J. Quant. Spectrosc. Radiat. Transf. 283 (2022), https://doi.org/10.1016/j.jqsrt.2022.108134.
[19] N.A. Menad, A. Hemmati-Sarapardeh, A. Varamesh, S. Shamshirband, Predicting solubility of CO_2 in brine by advanced machine learning systems: application to carbon capture and sequestration, J. of CO_2 Utilization 33 (2019) 83–95, https://doi.org/10.1016/j.jcou.2019.05.009.
[20] Y. Li, S. Li, X. Sun, D. Hao, Prediction of carbon dioxide production from green waste composting and identification of critical factors using machine learning algorithms, Bioresour. Technol. 360 (2022) 111, https://doi.org/10.1016/j.biortech.2022.127587.
[21] H. Jin, V. Andalib, G. Yasin, D.O. Bokov, M. Kamal, M. Alashwal, S. Ghazali, M. Algarni, A. Mamdouh, Computational simulation using machine learning models in prediction of CO_2 absorption in environmental applications, J. of Molecular Liq. 358 (2022), https://doi.org/10.1016/j.molliq.2022.119159.
[22] A.F. Ibrahim, Prediction of coal wettability using machine learning for the application of CO2 sequestration, 2022, Int. J. Greenh. Gas Control (2022) 118, https://doi.org/10.1016/j.ijggc.2022.103670.
[23] A. Shalaby, A. Elkamel, P.L. Douglas, Q. Zhu, Q.P. Zheng, A machine learning approach for modeling and optimization of a CO_2 post-combustion capture unit, 2021, Energy (2021) 215, https://doi.org/10.1016/j.energy.2020.119113.
[24] F. Guo, Z. Chen, F. Xiao, A. Li, J. Shi, Real-time energy performance benchmarking of electric vehicle air conditioning systems using adaptive neural network and Gaussian process regression, Appl. Therm. Eng. (2023) 222.
[25] C. Barile, S. Carone, C. Casavola, G. Pappalettera, Implementation of Gaussian Process Regression to strain data in residual stress measurements by hole drilling, Measurement 211 (2023), 112590.
[26] Q. Lv, A. Rashidi-Khaniabadi, R. Zheng, T. Zhou, M.R. Mohammadi, A. Hemmati-Sarapardeh, Modelling CO_2 diffusion coefficient in heavy crude oils and bitumen using extreme gradient boosting and Gaussian process regression, Energy (2023) 275.
[27] N. Zhang, J. Xiong, J. Zhong, K. Leatham, Gaussian process regression method for classification for high-dimensional data with limited Samples, in: 2018 Eighth International Conference on Information Science and Technology (ICIST), Cordoba, Granada, and Seville, Spain, 2018, pp. 358–363, https://doi.org/10.1109/ICIST.2018.8426077.
[28] J.G.J. Eberhart, Phys. Chem. 70 (1996) 1183–1186.
[29] S. Just, F. Sievert, M. Thommes, J. Breitkreutz, Eur. J. Pharm. Biopharm. 85 (2013) 1191–1199.
[30] E.L. Eliel, S.H. Wilen, L.N. Mander, Stereochemistry of Organic Compounds, Wiley, New York, 1994. ISBN 81-224-0570-3.
[31] E.L. Eliel, N.L. Allinger, S.J. Angyal, G.A. Morrison, Conformational Analysis, Interscience Publishers, New York, 1965.
[32] J.A. Hirsch, Topics in Stereochemistry, John Wiley & Sons, Inc., New York, 1967, p. 199.
[33] C. Romers, C. Altona, H.R. Buys, E. Havinga, Topics in Stereochemistry, John Wiley & Sons,Inc., New York, 1969, p. 40.
[34] M.E.A.B. Seghier, D. Höche, M. Zheludkevich, Prediction of the internal corrosion rate for oil and gas pipeline: implementation of ensemble learning techniques, J. Nat. Gas Sci. Eng. 99 (2022), 104425.
[35] I.R. Soosaiprakam, A. Veawab, Corrosion and polarization behaviour of carbon steel in MEA-based CO_2 capture process, Int. J. Greenh. Gas Control 2 (2008) 553–562.
[36] N. Kladkaew, R. Idem, P. Tontiwachwuthikul, C. Saiwan, Corrosion behavior of carbon steel in the monoethanolamine−H_2O−CO_2−O_2−SO_2 system: products, reaction pathways, and kinetics, Ind. Eng. Chem. Res. 48 (2009) 10169–10179.
[37] A. Edwards, C. Osborne, S. Webster, D. Klenerman, M. Joseph, P. Ostovar, M. Doyle, Mechanistic studies of the corrosion inhibitor oleic imidazoline, Corrosion Sci. 36 (1994) 315–325.
[38] M. Palomar-Pardavé, M. Romero-Romo, H. Herrera-Hernández, M.A. Abreu-Quijano, N.V. Likhanova, J. Uruchurtu, Juárez-García, Influence of the alkyl chain length of 2 amino 5 alkyl 1,3,4 thiadiazole compounds on the corrosion inhibition of steel immersed in sulfuric acid solutions, Corrosion Sci. 54 (2012) 231–243.
[39] D. Asefi, M. Arami, A.A. Sarabi, N.M. Mahmoodi, The chain length influence of cationic surfactant and role of nonionic co-surfactants on controlling the corrosion rate of steel in acidic media, Corrosion Sci. 51 (2009) 1817–1821.
[40] S.A. Ali, A.M. El-Shareef, R.F. Al-Ghamdi, M.T. Saeed, The isoxazolidines: the effects of steric factor and hydrophobic chain length on the corrosion inhibition of mild steel in acidic medium, Corrosion Sci. 47 (2005) 2659–2678.
[41] E. McCafferty, N. Hackerman, Double layer capacitance of iron and corrosion inhibition with polymethylene diamines, J. Electrochem. Soc. 119 (1972) 146–154.
[42] J.D. Olivo, D. Young, B. Brown, S. Nesic, Effect of corrosion inhibitor alkyl tail length on the electrochemical process underlying CO_2 corrosion of mild steel, in: NACE Corrosion Conference, 2018.
[43] A.M. Al-Sabagh, H.M. Abd-El-Bary, R.A. El-Ghazawy, M.R. Mishrif, B.M. Hussein, Corrosion inhibition efficiency of heavy alkyl benzene derivatives for carbon steel pipelines in 1 M HCl, Egyptian J. of Petroleum 21 (2012) 89–100.
[44] S.A.A. El–Maksoud, The effect of organic compounds on the electrochemical behaviour of steel in acidic media. A review, Int. J. Electrochem. Sci. 3 (2008) 528–555.
[45] M.M.H. Rocky, S. Akhtar, Correlations and predictions for viscosity of binary liquid systems: new UNIFAC-VISCO interaction parameters for O, N, and S containing organic liquids, Ind. Eng. Chem. Res. 59 (16) (2020) 8004–8017, https://doi.org/10.1021/acs.iecr.0c00335.