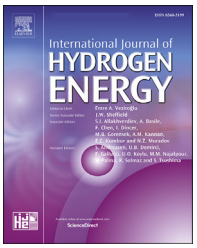
Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/he



Modeling interfacial tension of the hydrogen-brine system using robust machine learning techniques: Implication for underground hydrogen storage

Cuthbert Shang Wui Ng [a], *Hakim Djema* [b], *Menad Nait Amar* [b,**], *Ashkan Jahanbani Ghahfarokhi* [a,*]

[a] *Department of Geoscience and Petroleum, Norwegian University of Science and Technology, Trondheim, Norway*
[b] *Département Etudes Thermodynamiques, Division Laboratoires, Sonatrach, Boumerdes, Algeria*

HIGHLIGHTS

- The IFT of H_2-brine was modeled using different machine learning techniques.
- The suggested ML-based paradigms showed excellent predictions of the IFT values.
- The MLP-LMA based model outperformed the other intelligent and prior paradigms.
- The conservation of physical tendency of IFT of H_2-brine was demonstrated based on the trend analysis.

ARTICLE INFO

Article history:
Received 11 May 2022
Received in revised form
23 August 2022
Accepted 12 September 2022
Available online 12 October 2022

Keywords:
Hydrogen storage
Hydrogen-brine
Interfacial tension
Machine learning
Artificial neural network
Genetic programming

ABSTRACT

During the last years, there has been a surge of interest in cleaner ways for producing energy in order to successfully handle the climate issues caused by the consumption of fossil fuels. The production of hydrogen (H_2) is among the techniques which have grown up as attractive strategies towards energy transition. In this context, underground hydrogen storage (UHS) in saline aquifers has turned into one of the greatest challenges in the context of conserving energy for later use. The interfacial tension (IFT) of the H_2-brine system is a paramount parameter which affects greatly the successful design and implementation of UHS. In this study, robust machine learning (ML) techniques, viz., genetic programming (GP), gradient boosting regressor (GBR), and multilayer perceptron (MLP) optimized with Levenberg-Marquardt (LMA) and Adaptive Moment Estimation (Adam) algorithms were implemented for establishing accurate paradigms to predict the IFT of the H_2-brine system. The obtained results exhibited that the proposed models and correlation provide excellent estimations of the IFT. In addition, it was deduced that MLP-LMA outperforms the other models and the existing correlation in the literature. MLP-LMA yielded R^2 and AAPRE values of 0.9997 and 0.1907%, respectively. Lastly, the trend analysis demonstrated the physical coherence and tendency of the predictions of MLP-LMA.

* *Corresponding author.*
** *Corresponding author.*
E-mail addresses: m.naitamar@univ-boumerdes.dz (M. Nait Amar), ashkan.jahanbani@ntnu.no (A. Jahanbani Ghahfarokhi).
https://doi.org/10.1016/j.ijhydene.2022.09.120

Introduction

Recently, there has been a dramatic increase in the attention put into the development of new clean energy resources for mitigating the different climate issues, such as the increase in average temperatures of the globe [1,2]. Indeed, these worrying climate issues are mainly caused by the consumption and the burning of fossil fuels which lead to large amounts of emitted greenhouse gases (GHG) [3]. Due to this fact, there has been growing interest in the reduction of the amounts of GHG by considering new resources with less impact on the environment and climate such as hydrogen [4]. The production of hydrogen from renewable resources such as solar thermochemistry, biomass gasification, and electrolytic is among the pathways which are increasingly getting interest in various fields of science and engineering [5]. This surge of interest is due to the importance of this source of energy from various perspectives and versatile utilizations such as in agriculture, food processing, petrochemistry, and also in the production of electricity [6]. This importance of hydrogen has triggered a huge number of innovative inquiries and procedures to store it in the medium and long term for later use in a managed manner [7]. Recently, underground hydrogen storage (UHS) has gained prominence as an effective way for preserving hydrogen at various time scales, firstly for medium- and long-term storage [8], in different geological formations, such as saline aquifers, basaltic formations, coal seams, depleted oil/gas reservoirs, and salt caverns [9–14]. In this context, many studies have proved the feasibility of this attractive solution in different worldwide sites [6,15,16]. Besides, real pilot tests with promising results were carried out in Texas, US and Teeside, UK [17]. Furthermore, there are other worldwide successful hydrogen storage sites, such as Kiel in Germany as a salt cavern site, and Lobodice in the Czech Republic, Beynes in France, and Ketzin Germany as aquifer sites [18,19].

As in the case of carbon capture and storage (CCS) in saline aquifers [20–22], key parameters related to UHS should be investigated and determined accurately in order to ensure the successful implementation of this process [5,23]. The interfacial tension of the H_2-brine system is one of the vital parameters known by the relevant impact on UHS [5,24]. Despite the essential role of IFT in the UHS, only a few studies shed light on the experimental determination of this paramount parameter [5,24]. Chow et al. [24] performed experimental measurements of the IFT of the water-H_2 and water-H_2-CO_2 systems using the pendant-drop method. The tests were carried out under comprehensive intervals of pressure (0.5–45 MPa) and temperature (298.15–448.15 K). Recently, Hosseini et al. [5] enriched the knowledge about the IFT of the water/brine-H_2 systems by providing new measurements of this parameter under other pressures (2.76–34.47 MPa), temperatures (298.15–423.15 K), and molalities (0–4.95 mol kg^{-1} of [0.864 NaCl + 0.136 KCl] (aq)). Besides, the authors provided an empirical correlation for estimating the IFT of the water/brine-H_2 systems based on the available experimental tests. Their developed correlation yielded satisfactory predictions of IFT. It is worth mentioning that despite the precision of the experimental procedures employed for determining the IFT, these approaches are recognized to be expensive and time-consuming [20,22].

Recently, the use of new modeling techniques based on machine learning (ML) techniques has been at the center of attention in the context of establishing cheap, user-friendly, and accurate predictive paradigms [25–27]. This family of techniques exhibited excellent prediction performance when they were deployed in different studies related to CO_2 [28–30] and during the modeling of the IFT of brine-some specific pure/mixture gases, mainly CO_2/CH_4/N_2 [20,22,26,31–34]. To the best of our knowledge, no previous study has examined ML methods for establishing accurate paradigms that can estimate the IFT of the H_2-brine systems.

In this study, rigorous machine learning approaches, including gradient boosting regressor (GBR) and multilayer perceptron (MLP) optimized with two robust algorithms, namely Levenberg-Marquardt (LMA) and Adaptive Moment Estimation (Adam) were implemented for establishing ML-based paradigms to predict the interfacial tension of the H_2-brine system. Besides, another type of ML technique, viz., genetic programming (GP) was considered in this study in order to develop an explicit correlation for IFT of H_2-brine system. The proposed ML-based models were trained and tested using comprehensive experimental measurements of the IFT of the H_2-brine system. This experimental database was gathered from the experimental studies performed by Hosseini et al. [5] and Chow et al. [24,35]. The gained intelligent models and explicit correlation were assessed and compared using a variety of statistical and graphical error analysis. Besides, our best ML-based paradigm and our proposed explicit correlation were compared with the empirical correlation suggested by Hosseini et al. [5].

Modeling techniques

Gradient boosting regressor (GBR)

Boosting method is a variant of ensemble approach that was proposed by Schapire [36]. This method fundamentally aims at combining underperforming predictors, which are also termed as "learner", to establish another predictor with better performance [37,38]. In a simpler term, these weak predictors would undergo sequential training phase in which each predictor emphasizes on rectifying the previous predictors. In this aspect, gradient boosting regression is one of the examples of boosting method. Additionally, GBR can be considered as an enhanced version of decision tree (DT) method when the mechanism of boosting is implemented to DT. Hence, some hyperparameters of GBR are relevant with those of decision tree methods, which comprise (but not limited to) the number of estimators, loss function, subsample, maximum number of features, maximum depth, minimum number of samples split, and minimum number of leaf nodes.

GBR also acts as a type of functional gradient descent and hence, an additional hyperparameter of learning rate is required. This denotes that a predefined loss function is minimized by adding a new predictor at each iteration of gradient descent to attain better training outcome [39]. Basically, residual errors (the difference between actual output

and predicted output) produced by the previous predictor would apply to the new predictor [38]. For a more vivid explanation, the algorithm of GBR is summarized as follows:

1. Considering data samples $\{(x_n, y_n),\ n = 1, 2, \ldots,\ N\}$ and a loss function $L(y_n, f(x))$ in which f(x) is the prediction function, perform initialization with constant value to have $f_o(x)$. γ denotes the predicted output value when the loss function is minimized:

$$f_o(x) = \text{argmin}_\gamma \sum_{n=1}^{N} L(y_n, \gamma) \quad (1)$$

2. Conduct the iteration for k = 1: K, in which K represents the number of iterations (also known as number of learners):
   a. Determine the negative gradient (also termed as pseudo-residuals), z_n where n = 1, 2, …N:

$$z_{nk} = -\left[\frac{\partial L(y_n, f(x_n))}{\partial f(x_n)}\right]_{f=f_{k-1}} \quad (2)$$

   b. Regress a base learner (predictor) $G_k(x)$ onto the target $\{z_n,\ n = 1, 2, \ldots,\ N\}$ to conduct the training.
   c. Compute the step size for gradient descent:

$$r = \text{argmin}_\gamma \sum_{n=1}^{N} L(y_n, f_{k-1}(x_n) + \gamma G_k(x_n)) \quad (3)$$

   d. Update the model:

$$f_k(x) = f_{k-1}(x) + r \cdot G_k(x) \quad (4)$$

3. After finishing K iterations, the final output is $f_K(x)$.

Multilayer perceptron (MLP)

Multilayer perceptron (MLP) is a variant of artificial neural network (ANN) applied in different tasks including supervised classification and the development of predictive paradigms [40]. This technique is inspired from the human neurological system which is considered in the learning process [41].

The design of MLP is based on a multitude of perceptrons or neurons connected to each other (an example is shown in Fig. 1). These neurons are distributed into three types of layers as follows [42]:

- Input layer: it represents data with actual input values.
- Hidden layer: it is the most important layer for data transformation by using an activation function. Examples of activation function comprise (but are not limited to) Tansig (Tangent sigmoid, alternatively known as hyperbolic tangent), ReLU (Rectified Linear Unit), Linear, and Sigmoid. Mathematically, given x is the input data with f (x) being the output, tansig is f (x) = tanh (x), ReLU is f (x) = max (0, x), linear is f (x) = x, and sigmoid is f (x) = $(1 + e^{-x})^{-1}$. The aim is to capture and recognize the relationship describing the studied data. The number of hidden layers can be one or more according to the degree of complexity of the phenomenon.
- Output layer: it is the last layer from which the results are delivered. In this layer, there is a possibility to apply different kinds of activation functions such as Pure line or other types.

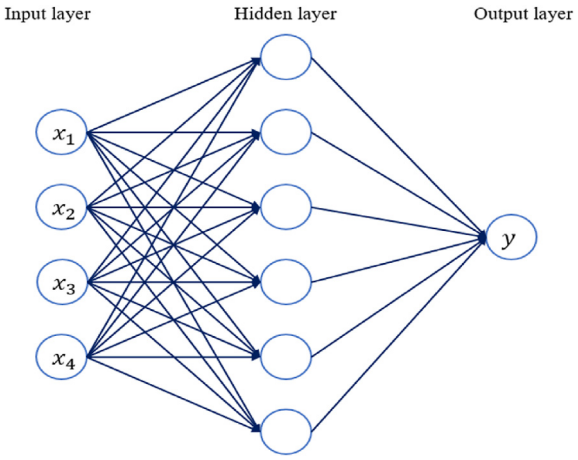


Fig. 1 – Illustration of the architecture of an MLP model.

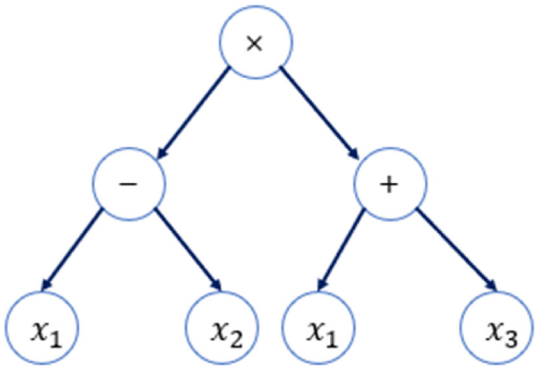
It is worth mentioning that each neural connection between different layers is associated with a set of weights. Additionally, the neurons of the output and hidden layers can contain a bias term.

During the learning phase of MLP model, it is essential to improve its performance by minimizing the calculated errors between the targets and model predictions. The learning phase is perceived as a process of finding the best weights and biases. Many algorithms have been proposed to optimize the weights and bias of MLP. In this study, we have applied Adaptive Movement Estimation (Adam) and Levenberg-Marquardt (LMA) during the training phase of MLP. The details of these two algorithms are reported in the prior published works [43,44].

Genetic programming (GP)

Genetic programming (GP) is another machine learning technique. It is formulated based upon the evolutionary theories inspired by the mechanism of natural selection [45]. This approach can be considered as a special version of genetic algorithm (GA) as it employs the same genetic operators except in the codification type of the chromosomes considered which is different [45]. The chromosomes of the population in GP are represented in a tree representation, and this means that each chromosome corresponds to a possible correlation related to the problems to be modeled [46]. Fig. 2 shows an example of a tree representation in GP.

The process of investigating the best correlation using the GP begins with creating an initial population and assessing the fitness of each individual in the population. Then, the population is subject to genetic operations including [47]:



$$(x_1 - x_2) \times (x_1 + x_3)$$

Fig. 2 – Illustration of the tree representation in GP.

- Elitism which is used to keep the best correlations alive for a new generation.
- Crossover that allows an exchange of parts between two correlations to generate two new ones.
- Mutation that consists of modifying certain parts of correlations to generate others.

Note that the correlations, which are subject to crossover and mutation operators, are chosen through selection techniques, viz. roulette wheel, tournament, ranking or others. Before using the operators, the individuals must satisfy the crossover and mutation probabilities. This means that before applying the crossover and mutation operators, we attribute random values to the concerned individuals and if these values are less than the crossover and mutation probabilities, the operators are applied to get new offspring. If this condition is not satisfied, we keep the same individuals to the following generation.

This set of operators is repeated iteratively until the stopping condition is satisfied.

Results and discussion

Prior to explaining the results from the development of the considered models, it is essential to understand the database applied to establish these models. The database was collected from several literatures, namely the works of Hosseini et al. [5] and Chow et al. [24,35] and it consisted of 107 datapoints in total. The summary of its statistics is illustrated in Table 1. These datapoints were then normalized between −1 and 1 by using the formula below:

Table 1 – Summary of database.

	Molality, m ($mol.kg^{-1}$)	Pressure, P (MPa)	Temperature, T (K)	IFT ($mN.m^{-1}$)
Minimum	0.00	0.50	298.03	42.90
Maximum	4.95	45.20	448.35	80.77
Average	1.37	17.87	353.94	63.21

Table 2 – Performance evaluation of the implemented paradigms.

		AAPRE	RMSE	R^2
MLP-LMA	Training	0.1557	0.1593	0.9998
	Testing	0.3344	0.2625	0.9995
	Total	0.1907	0.1842	0.9997
MLP-Adam	Training	0.3353	0.3116	0.9990
	Testing	1.0622	0.9509	0.9935
	Total	0.4848	0.5129	0.9979
GBR	Training	0.0940	0.0722	1.0000
	Testing	0.6559	0.4917	0.9979
	Total	0.2095	0.2320	0.9996
GP	Training	0.7862	0.5295	0.9964
	Testing	0.6353	0.2137	0.9975
	Total	0.7560	0.4663	0.9966

$$x_{norm,\,i} = 2 \times \left(\frac{x_i - x_{min}}{x_{max} - x_{min}} \right) - 1 \quad (5)$$

Based on the expression above, $x_{norm,\,i}$ indicates the normalized value of datapoint i, x_i refers to the value of datapoint i, x_{min} and x_{max} correspondingly mean the minimum and maximum values of all the datapoints. This method of normalization was preferred due to insights and positive results derived from our previous studies [28,48]. Thereafter, the normalized database was divided into two different sets, namely training and testing, according to the ratio of 8:2. Training dataset was mainly utilized to perform the data-driven modeling whereas testing set was used to ensure the generalization ability of the models.

After the preparation of the database, the training set was fed into the models including MLP with LMA (MLP-LMA), MLP with Adam (MLP-Adam), Gradient Boosting Regressor (GBR), and Genetic Programming (GP). It is worth mentioning that in the case of GP, the original raw database was employed in lieu of applying the normalized database.

To find out the proper architecture of the models, trial and error approach was implemented. Regarding the topology of MLP-Adam, it was made up of one input layer, three hidden layers, and one output layer. The first and last hidden layers had 15 hidden nodes whereas the second hidden layer had 20 hidden nodes. ReLU was selected as the activation function for all the three hidden layers while tansig was the activation function for the output layer. For MLP-LMA, the best-found topology consisted of a network with one hidden layer involving 12 neurons, while the suitable activation functions were tansig and pure line for the hidden and output layers of MLP-LMA model, respectively.

About the hyperparameters of GBR used, the number of estimators (boosting phases) was 50, the learning rate was 0.1 (used to reduce the contribution of each regressor), the maximum depth was 5 (used to limit the number of nodes in decision tree), and the minimum number of samples split was 2 (minimum number of samples to split an internal node in decision tree). The loss function used is the squared error. The other hyperparameters were set as default as recommended in Ref. [49].

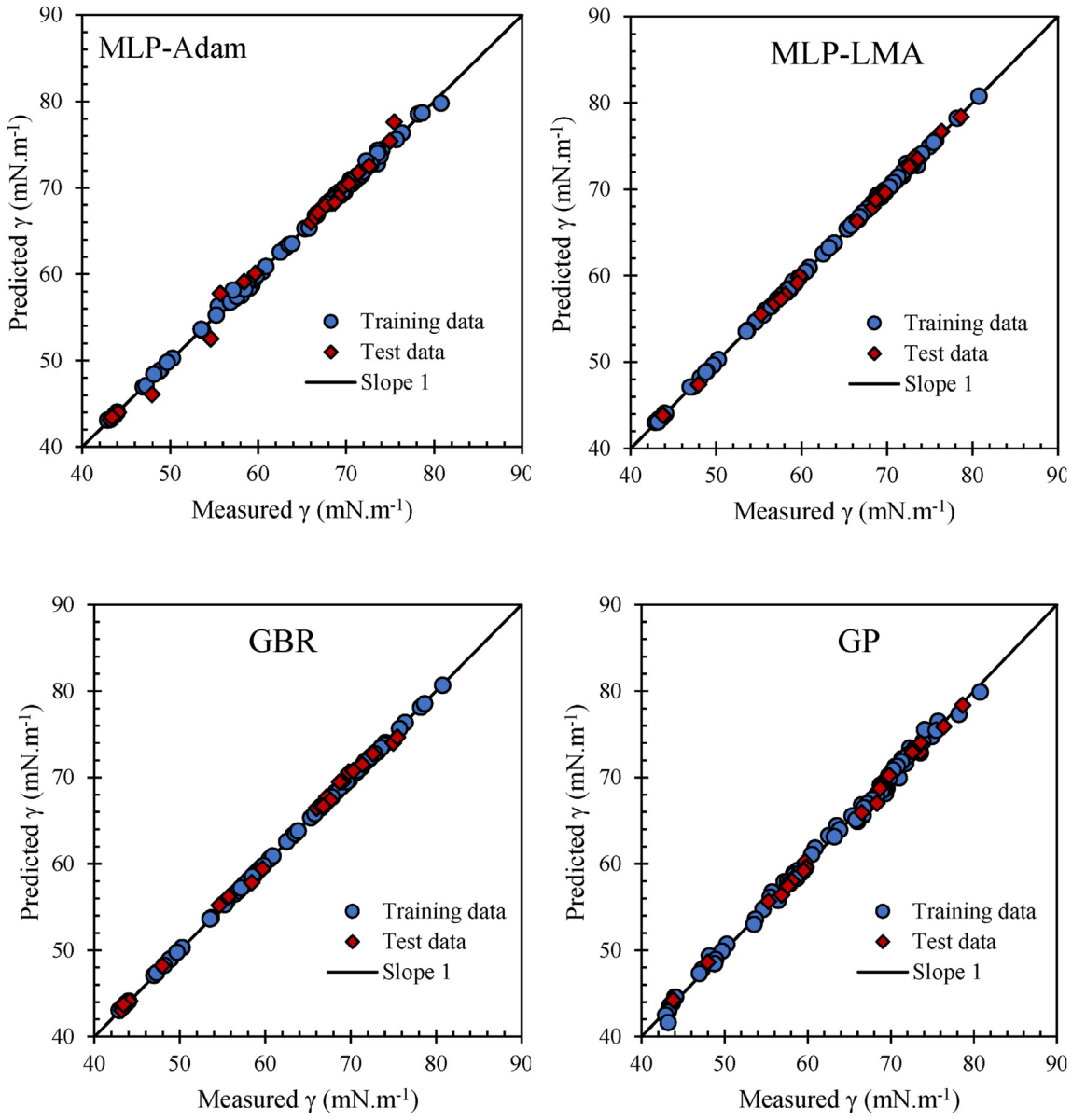


Fig. 3 – Cross plots of the implemented models for predicting IFT of the H_2-brine system.

The GP correlation was determined to be as follows:

$$IFT = 133.796737 - \frac{0.001195 \times (P \times A_1 + 4 \times T^2 \times A_2)}{T} \quad (6)$$

with

$$A_1 = 41 \times m + T \times (5 \times m - 94 \times \cos(T))$$

$$A_2 = 42 - m - \cos(T - \cos(\cos(m - 2 \times T)))$$

To evaluate the prediction performance of the resulting models, we provided the statistical analysis by considering three different criteria, viz. Average Absolute Percentage Relative Error (AAPRE), Root Mean Square Error (RMSE), and Coefficient of Determination (R^2). The formula of each criterion is displayed below:

$$AAPRE = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{\gamma_i^{exp} - \gamma_i^{pred}}{\gamma_i^{pred}}\right| \times 100 \quad (7)$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left[\gamma_i^{exp} - \gamma_i^{pred}\right]^2} \quad (8)$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N}\left[\gamma_i^{exp} - \gamma_i^{pred}\right]^2}{\sum_{i=1}^{N}\left[\gamma_i^{pred} - \overline{\gamma}\right]^2} \quad (9)$$

γ_i implies the value of IFT in which the superscripts of exp and pred correspondingly refer to "experimental" and "prediction". Also, $\overline{\gamma}$ indicates the average value of IFT and N is the

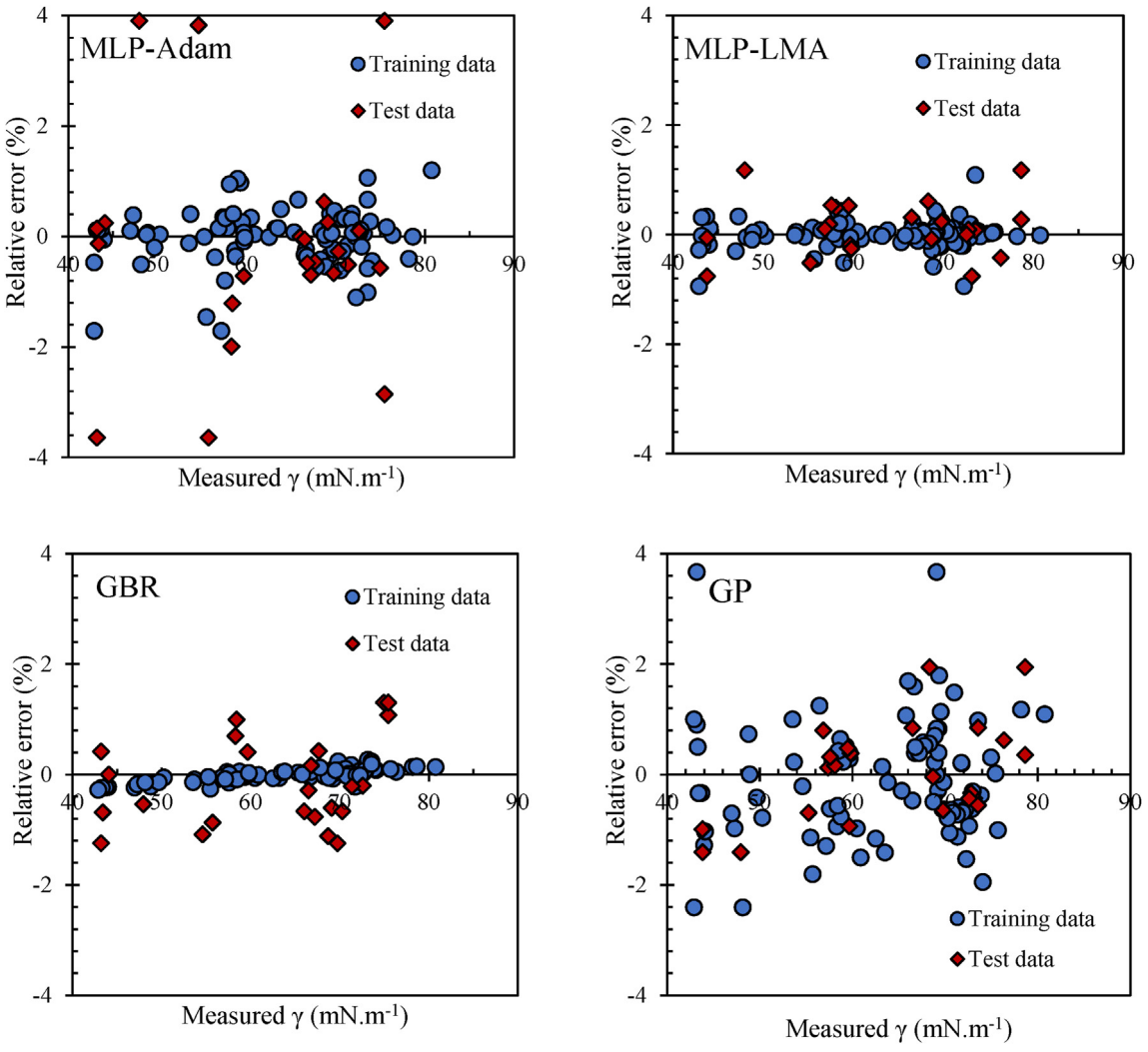


Fig. 4 – Error distribution of the implemented models for predicting IFT of the H_2-brine system.

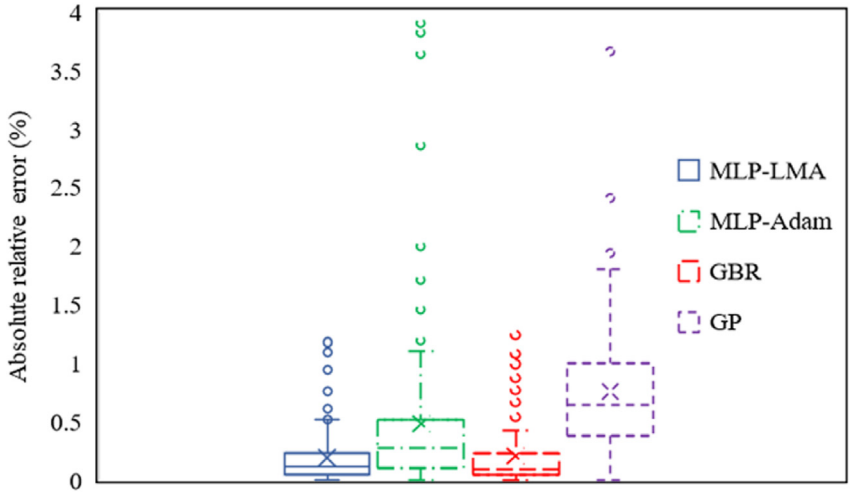


Fig. 5 – Box plot of the distribution of the absolute relative error associated with the model predictions.

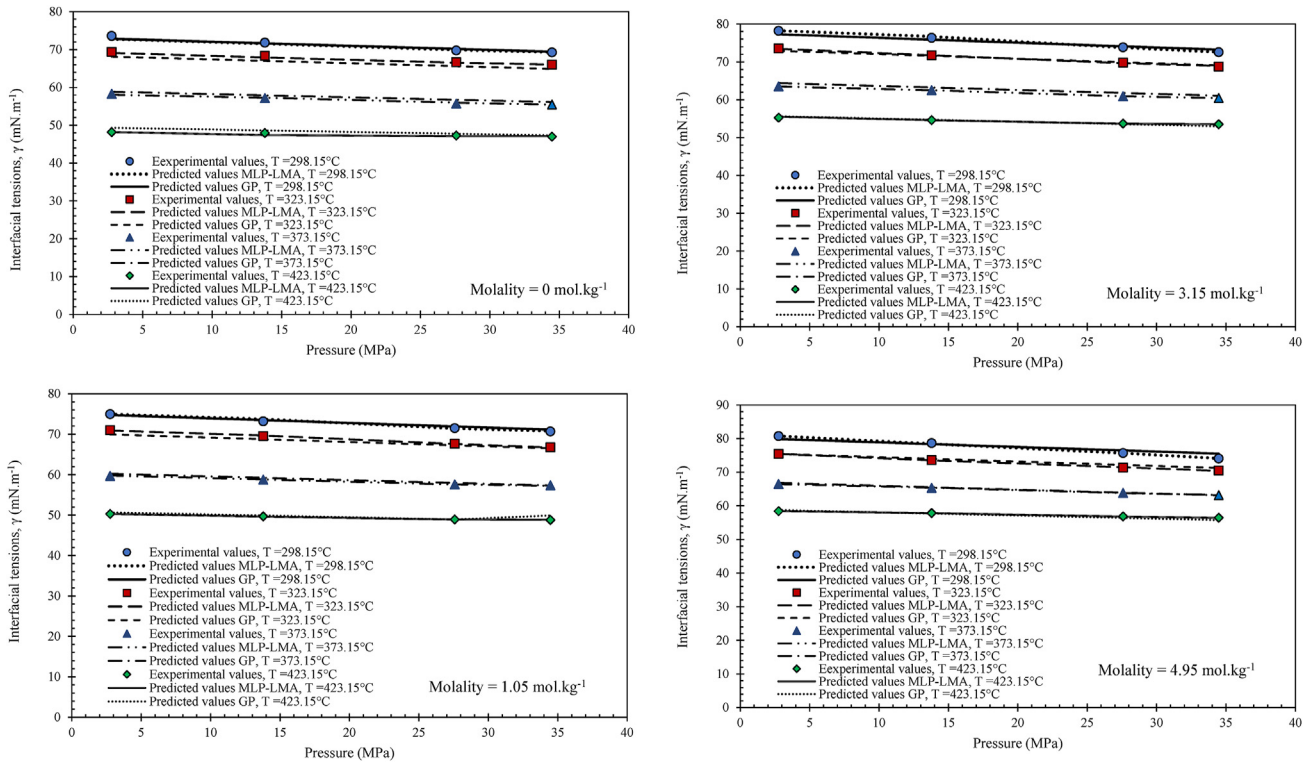


Fig. 6 – Trend analysis of the MLP-LMA model and GP-based correlation.

total number of datapoints. Table 2 summarizes the assessment of performance exhibited by each employed paradigm. Fig. 3 illustrates the cross plots of the developed predictive models of IFT. This figure demonstrates that all the models generally have yielded the predicted IFTs which are in proximity with the actual IFTs. On a closer scrutiny, as portrayed in Table 2, albeit GBR has shown the best training performance in terms of AAPRE, RMSE, and R^2, it is outperformed by MLP-LMA in testing performance. Furthermore, MLP-LMA outperforms the other models in the prediction of all the (total number of) datapoints.

Pertaining to the worst performing model in terms of prediction, GP is comparable to MLP-Adam. Considering all the three criteria, although MLP-Adam outperforms GP in training, it has exhibited a worse testing performance. When all the datapoints are considered, MLP-Adam shows a better prediction performance than GP in terms of AAPRE and R^2. To further evaluate the accuracy of all the four established models, additional analysis has been conducted. Fig. 4 displays the distribution of prediction errors produced by each model in both phases of training and testing. Good distribution of errors is reflected by having as many datapoint as possible lying on the horizontal line with relative error of 0%. Based on this figure, most of the prediction errors of MLP-LMA and GBR (considering both training and testing data) lie on the horizontal line. It can be inferred that MLP-LMA and GBR have outperformed MLP-Adam and GP in terms of distribution of prediction errors. As compared with GBR, fewer training data errors yielded by MLP-LMA lie on the horizontal line. Prediction errors of testing data of GBR deviate more from the horizontal line compared with MLP-LMA. Both MLP-LMA and GBR illustrate similarly good distribution of prediction errors. For MLP-Adam, as shown in Fig. 4, there are a few outliers of the predicted testing datapoints. However, GP demonstrates much more scattered prediction errors than MLP-Adam. Hence, MLP-Adam portrays better prediction performance than GP in this context.

Fig. 5 shows the box plot of the absolute relative error distribution corresponding to each model. This box plot illustrates that the absolute relative error associated with MLP-LMA was distributed within a range of 0% and 0.5% (without considering the outliers). For GBR, the range is between 0% and about 0.49%. Five outliers are noted in the case of MLP-LMA, whereas seven outliers for GBR. Hence, Figs. 4 and 5 have further validated that MLP-LMA and GBR outperform MLP-Adam and GP. It is challenging to say that MLP-LMA performs better than GBR from Figs. 4 and 5. Nevertheless, based on Table 2, it is deduced that MLP-LMA has a better prediction ability than GBR in this case. For the box plots of MLP-Adam and GP, the distribution of absolute relative error achieved by GP is more widespread than that of MLP-Adam. It is inferred that MLP-Adam slightly outperforms GP. Thereafter, to investigate the efficiency of the established models for prediction of IFT values, the trend analysis of the MLP-LMA (best performing) and GP-based correlation (worst performing) is displayed in Fig. 6. It is presented that at four different molality values, the trend of IFT values could be generally captured by both MLP-LMA and GP as temperature and pressure vary. This once again highlights the reliability of these models. Nevertheless, if Fig. 6 is examined in depth, the trend yielded by MLP-LMA in general matched slightly more closely the experimental values as compared with GP. As exhibited by

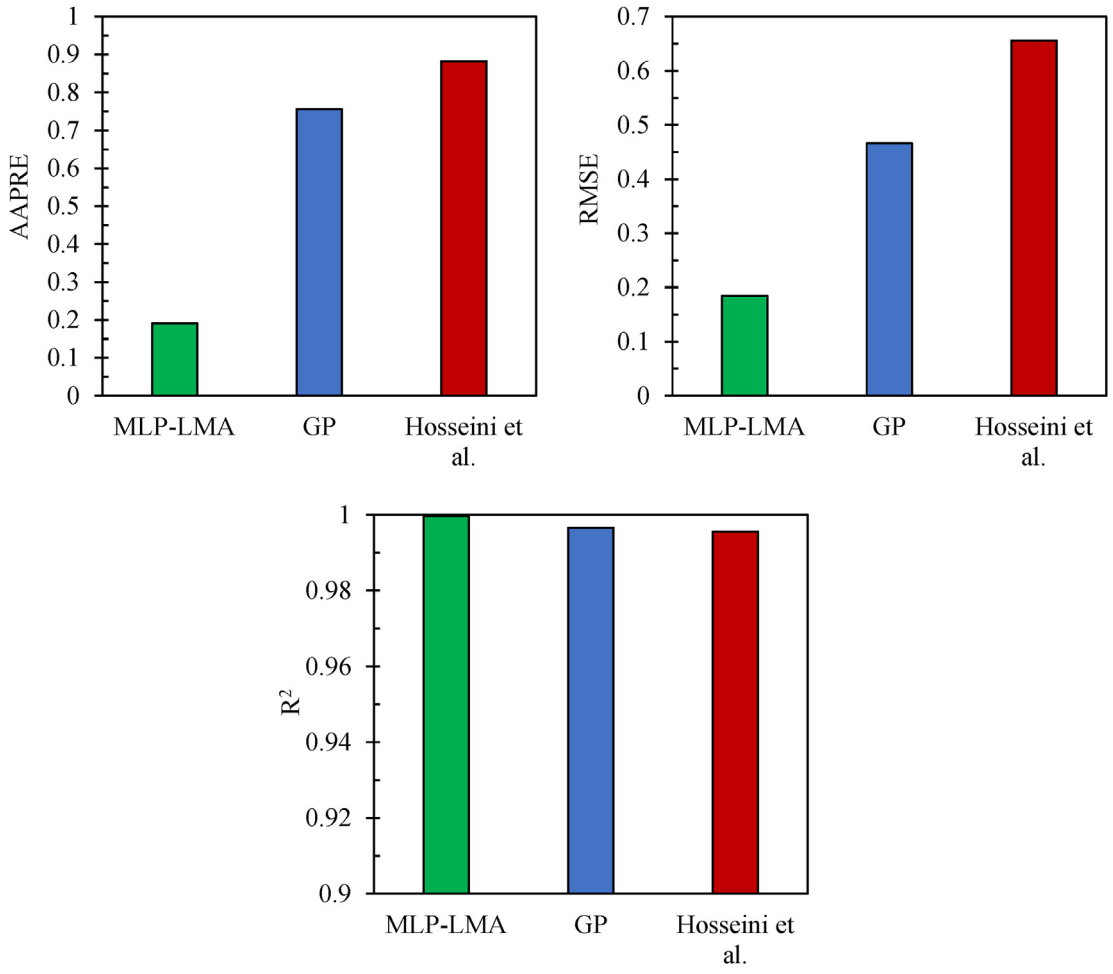


Fig. 7 – Comparison of the performance of MLP-LMA model and GP-based correlation with the correlation proposed by Hosseini et al.

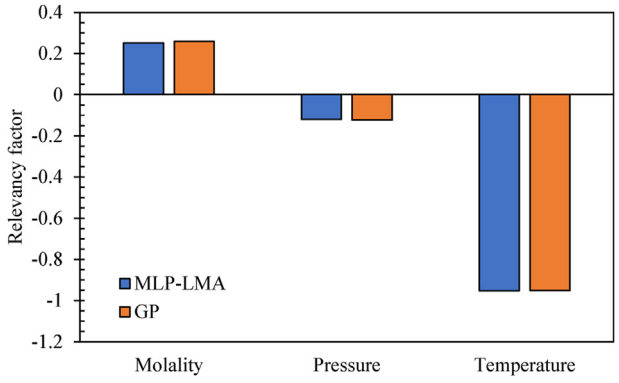


Fig. 8 – Relevancy factor.

the two predictive models and the experimental data, it can be observed that given the same molality and pressure, the IFT would decrease when the temperature increases. Besides that, at constant pressure and temperature, a lower molality would induce a lower IFT. When the temperature and molality remain constant, the IFT would decrease as the pressure increases. Nevertheless, the decrease is not very drastic as shown. The predictions of MLP-LMA and GP mimic perfectly these changes in the trend of IFT with respect to pressure, temperature, and molality. In general, it is shown that these two models are able to capture the physics (behavior of IFT) through the data provided.

To demonstrate the significant predictability of both MLP-LMA and GP, we performed another comparative analysis between these models and another correlation proposed by Hosseini et al. [5]. According to the results of the statistical analysis as displayed in Fig. 7, although MLP-LMA and GP only illustrate slight enhancement of predictions in terms of R^2, the improvement is significant looking at AAPRE and RMSE values, especially for MLP-LMA.

Lastly, relevancy factor study was also conducted as a sensitivity analysis to determine or "quantify" the relative importance of the input parameters on the IFT. Relevancy factor provides insightful ideas about how impactful input parameters are on the predicted outputs [50,51]. In general, the value of relevancy factor of an input parameter is either positive or negative. Positive (Negative) relevancy factor indicates that as the value of input parameters increases, the value of output increases (decreases). Moreover, the higher the magnitude of relevancy factor, the larger the impact of the input parameter will be on the output of models. Mathematically, relevancy factor is shown as:

$$r(I_j, \gamma) = \frac{\sum_{i=1}^{N}(I_{j,i} - \bar{I}_j)(\gamma_i - \bar{\gamma})}{\sqrt{\sum_{i=1}^{N}(I_{j,i} - \bar{I}_j)^2 \sum_{i=1}^{N}(\gamma_i - \bar{\gamma})^2}} \quad (10)$$

$I_{j,i}$ refers to the input parameter j at ith datapoint and $\bar{I}_j$ indicates the average value of I_j. γ_i is the predicted output of ith datapoint whereas $\bar{\gamma}$ represents its average value. N is the total number of datapoints. Considering only MLP-LMA and GP, Fig. 8 presents the bar charts of the relevancy factor of all the three input parameters. It is observed that temperature has the biggest impact on IFT, and it is followed by molality and pressure. Also, temperature and pressure both have an inverse relationship with IFT whereas the relationship between molality and IFT is direct. The result of relevancy factor coincides with the results of the trend analysis performed previously. Finally, it is worth mentioning that these trends and the impact of the input parameters are in accordance with previous published works such as Hosseini et al. [5] and Chow et al. [24,35].

To end with, it is needed to mention that the implemented ML-based schemes for modeling the IFT of the hydrogen-brine system are of great importance from various perspectives and they can be applied without having specific experimental values as they are derived and taught from prior experimental tests. The main application and implication of the established ML-based models pertain to the simulation of different downstream underground hydrogen storage related processes as the provided models in this study can be easily integrated in the associated calculation steps. Moreover, it is common to consider commercial software or some specific paradigms to estimate various properties of different systems such as hydrogen-brine; however, the fact is that these tools do not always deliver accurate predictions. For this purpose, the newly proposed ML-based approaches in this study are of great interest for the other related applications and the exhibited prediction accuracy of the models will contribute to the upgrading of such calculations.

Conclusion

In this work, we have employed four different data-driven techniques to establish a predictive model of IFT considering the input variables of molality, pressure, and temperature. The database used to develop these models was collected from the experimental works presented in several literatures. To assess the predictability of the models created, namely MLP-LMA, MLP-Adam, GBR, and GP-based correlation, we have implemented three metrics for statistical analysis. These metrics included AAPRE, RMSE, and R^2. According to the results of the analysis, it can be inferred that all the models illustrated excellent performance during the training and testing phases. MLP-LMA was identified as the best model because it yielded the highest R^2 in tandem with the lowest AAPRE and RMSE considering all datapoints. We also included some graphical analyses to have a better insight about the prediction ability of the models. Pertaining to this, a trend analysis was performed in which the predictive models were able to capture the changes in the behavior of IFT under the varying conditions of molality, pressure, and temperature. This further confirmed the robustness of the models developed in this study. Analysis of relevancy factor was conducted to identify the impact of each input variable on the IFT. It can be deduced that temperature has the most significant effect on the IFT. The reliability and integrity of the models were also justified when their performance surpassed one of the empirical correlations of IFT.

Author contributions

Cuthbert Shang Wui Ng: Data curation, Formal analysis, Methodology, Investigation and Modeling, Software, Writing.

Hakim Djema: Data curation, Formal analysis, Methodology, Investigation and Modeling, Software, Writing.

Menad Nait Amar: Data curation, Formal analysis, Methodology, Investigation and Modeling, Software, Writing.

Ashkan Jahanbani Ghahfarokhi: Supervision, Methodology, Writing, Reviewing and Editing.

Declaration of competing interest

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

REFERENCES

[1] Hassanpouryouzband A, Joonaki E, Edlmann K, Haszeldine RS. Offshore geological storage of hydrogen: is this our best option to achieve net-zero? ACS Energy Lett 2021;6:2181–6. https://doi.org/10.1021/acsenergylett.1c00845.

[2] Zivar D, Kumar S, Foroozesh J. Underground hydrogen storage: a comprehensive review. Int J Hydrogen Energy 2021;46:23436–62. https://doi.org/10.1016/j.ijhydene.2020.08.138.

[3] Ali M, Yekeen N, Pal N, Keshavarz A, Iglauer S, Hoteit H. Influence of pressure, temperature and organic surface concentration on hydrogen wettability of caprock; implications for hydrogen geo-storage. Energy Rep 2021;7:5988–96. https://doi.org/10.1016/j.egyr.2021.09.016.

[4] Hosseini M, Fahimpour J, Ali M, Keshavarz A, Iglauer S. Hydrogen wettability of carbonate formations: implications for hydrogen geo-storage. J Colloid Interface Sci 2022;614:256–66. https://doi.org/10.1016/j.jcis.2022.01.068.

[5] Hosseini M, Fahimpour J, Ali M, Keshavarz A, Iglauer S. H 2 – brine interfacial tension as a function of salinity , temperature , and pressure ; implications for hydrogen geo-storage. J Petrol Sci Eng 2022;213:110441. https://doi.org/10.1016/j.petrol.2022.110441.

[6] Lemieux A, Shkarupin A, Sharp K. Geologic feasibility of underground hydrogen storage in Canada. Int J Hydrogen Energy 2020;45:32243–59.

[7] Tarkowski R. Perspectives of using the geological subsurface for hydrogen storage in Poland. Int J Hydrogen Energy 2017;42:347–55.

[8] Tarkowski R. Underground hydrogen storage: characteristics and prospects. Renew Sustain Energy Rev 2019;105:86–94. https://doi.org/10.1016/j.rser.2019.01.051.

[9] Pan B, Yin X, Iglauer S. Rock-fluid interfacial tension at subsurface conditions: implications for H2, CO2 and natural

gas geo-storage. Int J Hydrogen Energy 2021;46:25578–85. https://doi.org/10.1016/j.ijhydene.2021.05.067.
[10] Hosseini M, Fahimpour J, Ali M, Keshavarz A, Iglauer S. Capillary sealing efficiency analysis of caprocks: implication for hydrogen geological storage. Energy Fuels 2022;36:4065–75. https://doi.org/10.1021/acs.energyfuels.2c00281.
[11] Hosseini M, Ali M, Fahimpour J, Keshavarz A, Iglauer S. Basalt-H2-brine wettability at geo-storage conditions: implication for hydrogen storage in basaltic formations. J Energy Storage 2022;52:104745. https://doi.org/10.1016/j.est.2022.104745.
[12] Yekta AE, Manceau J-C, Gaboreau S, Pichavant M, Audigane P. Determination of hydrogen–water relative permeability and capillary pressure in sandstone: application to underground hydrogen injection in sedimentary formations. Transport Porous Media 2018;122:333–56. https://doi.org/10.1007/s11242-018-1004-7.
[13] Ozarslan A. Large-scale hydrogen energy storage in salt caverns. Int J Hydrogen Energy 2012;37:14265–77. https://doi.org/10.1016/j.ijhydene.2012.07.111.
[14] Ali M, Jha NK, Al-Yaseri A, Zhang Y, Iglauer S, Sarmadivaleh M. Hydrogen wettability of quartz substrates exposed to organic acids; Implications for hydrogen geo-storage in sandstone reservoirs. J Petrol Sci Eng 2021;207:109081. https://doi.org/10.1016/j.petrol.2021.109081.
[15] Le Duigou A, Bader A-G, Lanoix J-C, Nadau L. Relevance and costs of large scale underground hydrogen storage in France. Int J Hydrogen Energy 2017;42:22987–3003.
[16] Bai M, Song K, Sun Y, He M, Li Y, Sun J. An overview of hydrogen underground storage technology and prospects in China. J Petrol Sci Eng 2014;124:132–6.
[17] Heinemann N, Alcalde J, Miocic JM, Hangx SJT, Kallmeyer J, Ostertag-Henning C, et al. Enabling large-scale hydrogen storage in porous media–the scientific challenges. Energy Environ Sci 2021;14:853–64.
[18] Liebscher A, Wackerl J, Streibel M. Geologic storage of hydrogen - fundamentals, processing, and projects. Hydrogen science and engineering : materials, processes, systems and technology. Weinheim, Germany: Wiley-VCH Verlag GmbH & Co. KGaA; 2016. p. 629–58. https://doi.org/10.1002/9783527674268.ch26.
[19] Kruck O, Crotogino F, Prelicz R, Rudolph T. Overview on all known underground storage technologies for hydrogen. HyUnder 2013;93.
[20] Nait Amar M. Towards improved genetic programming based-correlations for predicting the interfacial tension of the systems pure/impure CO2-brine. J Taiwan Inst Chem Eng 2021;127:186–96.
[21] Chalbaud C, Robin M, Lombard JM, Martin F, Egermann P, Bertin H. Interfacial tension measurements and wettability evaluation for geological CO2 storage. Adv Water Resour 2009;32:98–109.
[22] Amooie MA, Hemmati-Sarapardeh A, Karan K, Husein MM, Soltanian MR, Dabir B. Data-driven modeling of interfacial tension in impure CO2-brine systems with implications for geological carbon storage. Int J Greenh Gas Control 2019;90:102811.
[23] Pan B, Yin X, Iglauer S. Rock-fluid interfacial tension at subsurface conditions: implications for H2, CO2 and natural gas geo-storage. Int J Hydrogen Energy 2021;46(50):25578–85. https://doi.org/10.1016/j.ijhydene.2021.05.067.
[24] Chow YTF, Maitland GC, Trusler JPM. Fluid Phase Equilibria Interfacial tensions of (H2O + H2) and (H2O + CO2 + H2) systems at temperatures of (298 e 448) K and pressures up to 45 MPa. Fluid Phase Equilibria 2018;475. https://doi.org/10.1016/j.fluid.2018.07.022.
[25] Daryayehsalameh B, Nabavi M, Vaferi B. Modeling of CO2 capture ability of [Bmim][BF4] ionic liquid using connectionist smart paradigms. Environ Technol Innov 2021;22:101484.
[26] Mehrjoo H, Riazi M, Nait Amar M, Hemmati-Sarapardeh A. Modeling interfacial tension of methane-brine systems at high pressure and high salinity conditions. J Taiwan Inst Chem Eng 2020;114:125–41. https://doi.org/10.1016/j.jtice.2020.09.014.
[27] Nait Amar M, Ouaer H, Ghriga MA. Robust smart schemes for modeling carbon dioxide uptake in metal- organic frameworks. Fuel 2022;311:122545.
[28] Nait Amar M, Jahanbani Ghahfarokhi A, Zeraibi N. Predicting thermal conductivity of carbon dioxide using group of data-driven models. J Taiwan Inst Chem Eng 2020;113:165–77. https://doi.org/10.1016/j.jtice.2020.08.001.
[29] Nait Amar M, Jahanbani Ghahfarokhi A. Prediction of CO2 diffusivity in brine using white-box machine learning. J Petroleum Sci Eng 2020;190:107037. https://doi.org/10.1016/j.petrol.2020.107037.
[30] Nait Amar M, Zeraibi N, Jahanbani Ghahfarokhi A. Applying hybrid support vector regression and genetic algorithm to water alternating CO2 gas EOR. Greenh Gases: Sci Technol 2020;10:613–30. https://doi.org/10.1002/ghg.1982.
[31] Zhang J, Feng Q, Wang S, Zhang X, Wang S. Estimation of CO2–brine interfacial tension using an artificial neural network. J Supercrit Fluids 2016;107:31–7.
[32] Madani M, Abbasi P, Baghban A, Zargar G, Abbasi P. Modeling of CO2-brine interfacial tension: application to enhanced oil recovery. Petrol Sci Technol 2017;35:2179–86.
[33] Rashid S, Harimi B, Hamidpour E. Prediction of CO2-Brine interfacial tension using a rigorous approach. J Nat Gas Sci Eng 2017;45:108–17.
[34] Kamari A, Pournik M, Rostami A, Amirlatifi A, Mohammadi AH. Characterizing the CO2-brine interfacial tension (IFT) using robust modeling approaches: a comparative study. J Mol Liq 2017;246:32–8.
[35] Chow YTF, Maitland GC, Trusler JPM. Erratum to "interfacial tensions of (H2O + H2) and (H2O + CO2 + H2) systems at temperatures of (298 to 448) K and pressures up to 45 MPa" [Fluid Phase Equil. 475 (2018) 37–44]. Fluid Phase Equil 2020;503:112315. https://doi.org/10.1016/J.FLUID.2019.112315.
[36] Schapire RE. The strength of weak learnability. Machine Learning 1990;5. https://doi.org/10.1007/bf00116037.
[37] Hastie T, Tibshirani R, Friedman J. Springer series in statistics the elements of statistical learning - data mining, inference, and prediction, 2nd; 2009.
[38] Nait Amar M, Shateri M, Hemmati-Sarapardeh A, Alamatsaz A. Modeling oil-brine interfacial tension at high pressure and high salinity conditions. J Petrol Sci Eng 2019;183:106413. https://doi.org/10.1016/j.petrol.2019.106413.
[39] Friedman JH. Stochastic gradient boosting. Comput Stat Data Anal 2002;38. https://doi.org/10.1016/S0167-9473(01)00065-2.
[40] Hosseini M, Rahimi R, Ghaedi M. Hydrogen sulfide solubility in different ionic liquids: an updated database and intelligent modeling. J Mol Liq 2020;317:113984.
[41] Karimi M, Vaferi B, Hosseini SH, Olazar M, Rashidi S. Smart computing approach for design and scale-up of conical spouted beds with open-sided draft tubes. Particuology 2021;55:179–90.
[42] Haykin S. Neural networks and learning machines third edition, vol. 40. Saddle River, NJ, USA: Pearson Upper; 2001. https://doi.org/10.1002/1521-3773 (20010316)40:6<9823::AID-ANIE9823>3.3.CO;2-C.

[43] Ng CSW, Jahanbani Ghahfarokhi A, Nait Amar M, Torsæter O. Smart proxy modeling of a fractured reservoir model for production optimization: implementation of metaheuristic algorithm and probabilistic application. Nat Res Res 2021;30:2431–62.
[44] Kingma DP, Ba J. Adam: a method for stochastic optimization. ArXiv 2014. Preprint ArXiv:14126980.
[45] Kar AK. Bio inspired computing–a review of algorithms and scope of applications. Expert Systems with Applications 2016;59:20–32.
[46] Sette S, Boullart L. Genetic programming: principles and applications. Eng Appl Artif Intell 2001;14:727–36.
[47] Rostami A, Arabloo M, Ebadi H. Genetic programming (GP) approach for prediction of supercritical CO2 thermal conductivity. Chem Eng Res Design 2017;122:164–75.
[48] Nait Amar M, Jahanbani Ghahfarokhi A, Ng CSW. Predicting wax deposition using robust machine learning techniques. Petroleum 2022;8:167–73. https://doi.org/10.1016/j.petlm.2021.07.005.
[49] Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: machine learning in Python. J Machine Learning Res 2011;12.
[50] Ng CSW, Jahanbani Ghahfarokhi A, Nait Amar M. Well production forecast in Volve field: application of rigorous machine learning techniques and metaheuristic algorithm. J Petrol Sci Eng 2022;208:109468.
[51] Chen G, Fu K, Liang Z, Sema T, Li C, Tontiwachwuthikul P, et al. The genetic algorithm based back propagation neural network for MMP prediction in CO2-EOR process. Fuel 2014;126:202–12.