Journal of the Taiwan Institute of Chemical Engineers 113 (2020) 165–177

Contents lists available at ScienceDirect

Journal of the Taiwan Institute of Chemical Engineers

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

Predicting thermal conductivity of carbon dioxide using group of data-driven models

Menad Nait Amar[a,b], Ashkan Jahanbani Ghahfarokhi[c,*], Noureddine Zeraibi[b]

[a] Département Etudes Thermodynamiques, Division Laboratoires, Sonatrach, Avenue 1er Novembre, 35000, Boumerdes, Algeria
[b] Laboratory of Hydrocarbons Physical Engineering, Faculty of Hydrocarbons and Chemistry, University of M'Hamed Bougara Boumerdes, Avenue de l'Indépendance, 35000, Boumerdes Algeria
[c] Department of Geoscience and Petroleum, Norwegian University of Science and Technology, Trondheim (NTNU), Norway

ARTICLE INFO

Article History:
Received 9 April 2020
Revised 20 July 2020
Accepted 2 August 2020
Available online 18 August 2020

Keywords:
CO_2
CO_2 transportation
Thermal conductivity
CCS
Data-driven

ABSTRACT

Thermal conductivity of carbon dioxide (CO_2) is a vital thermophysical parameter that significantly affects the heat transfer modeling related to CO_2 transportation, pipelines design and associated process industries. The current study lays emphasis on implementing powerful soft computing approaches to develop novel paradigms for estimation of CO_2 thermal conductivity. To achieve this, a massive database including 5893 experimental datapoints was acquired from the experimental investigations. The collected data, covering pressure values from 0.097 to 209.763 MPa and temperature between 217.931 and 961.05 K, were employed for establishing various models based on multilayer perceptron (MLP) optimized by different back-propagation algorithms, and radial basis function neural network (RBFNN) coupled with particle swarm optimization (PSO). Then, the two best found models were linked under two committee machine intelligent systems (CMIS) using weighted averaging and group method of data handling (GMDH). The obtained results showed that CMIS-GMDH is the most accurate paradigm with an overall AARD% and R^2 values of 0.8379% and 0.9997, respectively. In addition, CMIS-GMDH outperforms the best prior explicit models. Finally, the leverage technique confirmed the validity of the model and more than 96% of the data are within its applicability realm.

1. Introduction

The increased level of CO_2 emissions in the atmosphere is regarded as the primary issue of climate change [1–3]. As a result, the environmental and industrial researches are often confronted with the famous dichotomy: on one hand, the technological advancements must cater the lifestyle and meet the practical needs, and on the other hand, the emission of CO_2 must be kept at low levels [4]. With the intention of reducing greenhouse gas emissions, enormous efforts have been made to adequately establish proper techniques for Carbon Capture and Sequestration (CCS) in geological formations [5–9]. In addition, some other technologies that aim at capturing CO_2 at various stages from processes such as combustion or gasification are increasingly getting attention from technical and environmental perspectives [4].

Being an attractive strategy for dealing with the industrial CO_2 emissions, CCS is generally performed through chronological cutting-edge approaches and technologies attributed to three classes, viz. post-combustion capture; pre-combustion capture and oxy-fuel combustion capture; CO_2 transport and storage [10–12]. These classes include various industrial units, land transportation using pipeline, maritime transport by ship, and lastly injection or storage in geological formations [13,14]. In the context of injecting CO_2 in geological formations, enhanced oil recovery (EOR) techniques by injecting miscible CO_2 (with oil) have shown increased ultimate recovery factors in various oil reservoirs. The CO_2-oil system can easily reach the miscibility condition during CO_2 injection, leading to the improvement of the microscopic displacement efficiency [15–20].

In addition, CO_2 mainly in its supercritical (SC) state is gaining great attention in numerous industrial sectors such as refrigeration, dry cleaning, pharmaceuticals and food [21–23]. This noticeable trend towards multi-application of CO_2 is due to various factors, namely its non-flammability, accessible simplicity at high purity, and being inexpensive and non-toxic.

While performing simulations and design tasks related to different fluids utilization such as CO_2, nano-fluids and ionic liquids, various vital thermophysical properties that describe the behavior and the state of the fluids under different thermodynamic conditions should be determined properly [10,24–33]. The proper knowledge and assessment of these parameters are required for various advanced thermal engineering applications such as the fabrication and design of new energy systems [34], flow studies and modeling

* Corresponding author.
E-mail address: ashkan.jahanbani@ntnu.no (A. Jahanbani Ghahfarokhi).

https://doi.org/10.1016/j.jtice.2020.08.001

[31–33], transportation and lubrication processes [35]. In this context, relevant experimental techniques and modeling approaches exist in the literature regarding the investigation of thermal proprieties of fluids [36–40].

Among thermophysical properties of a fluid, thermal conductivity has a noticeable effect on heat transfer behavior [41,42]. For CO_2, this parameter is considered paramount in the mathematical formulation of heat transfer modeling associated with the design of CO_2 transportation pipelines [43]. It is worth noting that CO_2 thermal conductivity is needed under extensive pressure and temperature conductions due to the particularity of CO_2 transportation and its concerns [43,44]. Due to this fact, accurate knowledge of this parameter is necessary to ensure the proper design and correct flow conditions of CO_2 during transportation.

Due to the cost and considerable time needed to determine CO_2 thermal conductivity using lab measurements based on sophisticated procedures such as steady-state hot-wire measurements of the dilute gas, single-wire steady-state measurements and steady-state low-density measurements [45,46], many researches have been conducted to establish predictive models under different operational conditions. Bahadori and Vuthaluru [43] established an explicit model based on third order polynomial for correlating CO_2 thermal conductivity with pressure and temperature. Their correlation was developed using experimental data covering temperature values from 260 to 450 K and pressure values within the interval 10–70 MPa. Jarrahian and Heidaryan [47] applied multiple regression analysis method to develop a temperature and pressure dependent correlation for estimating CO_2 thermal conductivity. Their proposed model was developed using 688 real datapoints gathered from prior experimental studies, covering pressures from 7.41 to 209.68 MPa and temperatures from 311.25 to 960.68 K. Amooey [48] formulated a minimization of the sum of square of errors to establish a predictive model of CO_2 thermal conductivity by considering density and temperature as the input parameters. The model was developed using 600 experimental datapoints covering temperatures from 290 to 800 K and densities between 1 and 1200 kg/m^3. Recently, Rostami et al. [49] applied genetic programming (GP) to provide a predictive correlation using 752 datapoints. Their correlation depends on temperature, density and pressure. Another explicit correlation was proposed by Rostamian and Lotfollahi [50] for predicting thermal conductivity of vapor, liquid, and supercritical CO_2 as a function of density and temperature. The proposed correlation can be applied for temperature and density values in the ranges of 250 to 1100 K and 0.3 to 1400 kg/m^3, respectively. Table 1 illustrates the main mathematical formulas of the aforementioned correlations, the number of considered datapoints as well as the validity in terms of the covered operational conditions. According to this table, with the simple form of the correlations, some of them are developed with limited number of datapoints. Besides, it can be seen from the applicability domains that the lowest temperature value is equal to 250 K while very low temperature values are not covered.

Data-driven techniques are increasingly exhibiting high performance in modeling many complicated systems [35,51]. For CO_2 related topics, these methods showed very interesting results when modeling various parameters and phenomena such as CO_2 solubility in water [52], interfacial tension in impure CO_2-brine systems [53], minimum miscibility pressure of the CO_2-crude oil system [54–56], CO_2 viscosity [57], CO_2 solubility in ionic liquids [58] and CO_2 injection-based EOR [16,20,59]. In recent years, some researchers have investigated the ability of these data-driven techniques to model CO_2 thermal conductivity. Shams et al. [60] modeled CO_2 thermal conductivity by applying Least-Square Support Vector Machine (LSSVM). To improve the performance of their ML technique, the authors applied Coupled Simulated Annealing (CSA) for auto-tuning of the LSSVM hyper-parameters. Their proposed hybridization was established using 558 experimental datapoints. Tatar et al. [61] implemented another hybridization by coupling Adaptive Neuro-Fuzzy Inference System (ANFIS) with Particle Swarm Optimization (PSO) to establish a predictive paradigm of CO_2 thermal conductivity using 1042 experimental data from the literature. Ahmadi and Baghban [62] proposed a paradigm for estimating CO_2 thermal conductivity based on a coupled LSSVM and genetic algorithm (GA). In their investigation, the authors considered 746 experimental datapoints for establishing the predictive model. Rostamian and Lotfollahi [63] applied surface response methodology to generate a CO_2 thermal conductivity predictive model by considering density, pressure and temperature as the input parameters. However, in their study, an important part of the employed data was not experimental but generated from the equation of Vesovic et al. [64]. A deep survey on these performed studies based on soft computing techniques reveals that some of the proposed models are limited in terms of the operational conditions and the experimental datapoints utilized in their establishment. The literature survey reveals that modeling CO_2 thermal conductivity is still an open area of research.

The major aim of the present research is to establish accurate and robust paradigms using advanced intelligent schemes for estimating CO_2 thermal conductivity under widespread temperature and pressure conditions. To achieve this goal, multilayer perceptron (MLP) optimized by different back-propagation techniques, and radial basis function neural network (RBFNN) coupled with Particle Swarm Optimization (PSO) were implemented as preliminary predictors. Afterwards, the two best found models were linked under two committee machine intelligent systems (CMIS) using weighted averaging and group method of data handling (GMDH). A massive experimental database encompassing more than 5890 experimental datapoints was collected from literature to develop these models. The experimental data cover noticeable ranges of operational conditions, namely pressure range of 0.097 to 209.763 MPa, and temperature between 217.931 and 961.05 K. The present work contributes significantly to the fields of CO_2 utilization and application of soft computing techniques in engineering: (1) the size of the employed database is highly important, and this makes the validity of the established model more pertinent; (2) the employed experimental datapoints cover extensive operational conditions; and (3) a new CMIS scheme was proposed for the first time by combining some MLP models with GMDH.

The remainder of the included sections in this investigation are outlined as follows. Section 2 presents the theoretical background related to the implemented soft computing techniques as well as the proposed CMIS models. Section 3 describes the collected experimental data employed in building the models. In Section 4, the results are highlighted and discussed. The paper ends with Section 5, where the main findings are illustrated.

2. Models

2.1. Multilayer perceptron neural network

Artificial neural network (ANN) is a well-developed and applied soft computing technique. ANN is inspired from the neurological system and its basic elements for handling the information [65]. This technique is recognized to be very robust in emulating the outputs of various complex systems after its learning phase based on prior knowledge of a set of representative data [52]. An ANN model is associated with a topology that encompasses neurons distributed in three kinds of layers, namely input, hidden and output layers [65]. The data are received from the input layer and transformed into higher-dimension features in the hidden layer using predefined processing steps, while the results of the model are delivered from the output layer [65]. The neurons of a given layer are linked to their counterparts of the succeeding layer by means of weights. Moreover, bias terms are frequently considered in the neurons of hidden and output layers. The structure of an ANN allows a flexibility for learning the

Table 1
Summary of the prior correlations for estimating CO_2 thermal conductivity.

Correlation	Mathematical formula	Number of datapoints used in the development	Operational conditions	Reported accuracy	Applied modeling method	Type of fluid
Bahadori and Vuthaluru [43]	$Ln(\lambda) = a + \frac{b}{P} + \frac{c}{P^2} + \frac{d}{P^3}$ where: $a = 2.51177 - \frac{(4.61299E+3)}{T} + \frac{(1.5604E+6)}{T^2} - \frac{(1.64868E+8)}{T^3}$ $b = -(6.78436E+2) + \frac{(5.94729E+5)}{T} - \frac{(1.81369E+8)}{T^2} + \frac{(1.86064E+10)}{T^3}$ $c = (2.064898E+4) - \frac{(1.99667E+7)}{T} + \frac{(6.42367E+9)}{T^2} - \frac{(6.8022E+11)}{T^3}$ $b = -(1.09504E+5) + \frac{(1.08783E+8)}{T} - \frac{(3.57549E+10)}{T^2} + \frac{(3.855E+12)}{T^3}$ λ is expressed in W/m.K	Not mentioned	260 $K \le T \le 450$ K 10 $MPa \le P \le 70$ MPa	AAD*=1.30%	Third order polynomial	CO_2
Jarrahian and Heidaryan [47]	$\lambda = \frac{A_1 + A_2 P + A_3 P^3 + A_4 Ln(T) + A_5 Ln(T)^2}{1 + A_6 P + A_7 Ln(T) + A_8 Ln(T)^2 + A_9 Ln(T)^3}$ where $A_1 = 1.49288E+1$; $A_2 = 2.62541E-3$; $A_3 = 8.77805E-6$; $A_4 = -5.11425$; $A_5 = 4.37711E-1$; $A_6 = 2.11405E-5$; $A_7 = -4.73036E-1$; $A_8 = 7.36636E-2$; and $A_9 = -3.76340E-3$. λ is expressed in mW/m.K	688	311.25 $K \le T \le 960.68$ K 7.41 $MPa \le P \le 209.68$ MPa	AAD=2.70% R^2 **=0.994	Multiple regression analysis	$SC-CO_2$
Amooey [48]	$\lambda = \frac{A_1 + A_2\rho + A_3\rho^2 + A_4\rho^3 T^3 + A_5\rho^4 + A_6 T + A_7 T^2}{\sqrt{T}}$ where $A_1 = -1.05161E+2$; $A_2 = 9.00700E-1$; $A_3 = 7.00000E-4$; $A_4 = 3.50000E-15$; $A_5 = 3.76000E-10$; $A_6 = 7.50000E-1$; and $A_7 = 1.70000E-3\lambda$ is expressed in mW/m.K	600	290 $K \le T \le 800$ K 1 $\frac{kg}{m^3} \le \rho \le 1200$ $\frac{kg}{m^3}$	AAD=2.74%	Minimization of the sum of square of errors	$SC-CO_2$
Rostami et al. [49]	$P < 20$ MPa $\lambda = 0.0936 \times T - 0.448 \times P + 0.0739 \times \rho - 0.244 \times Ln(\rho) - 10.8 \frac{\times P}{T+0.00753} \times P^2 + 1.85 \times 10^{-5} \times \rho^2 + 94.7 \frac{\times \rho}{T \times P - 16.7} \lambda$ is expressed in mW/m.K $P \ge 20$ MPa $\lambda = 0.0575 \times T + 0.0151 \times P + 0.0372 \times \rho + 4.96 \times 10^{-5} \times Ln(Ln(\rho)) - 0.00695 \times Ln(P) + 1.41 \times 10^{-5} \times T^2 + 7.2 \times 10^{-8} \times \rho^3 + 1.78\lambda$ is expressed in mW/m.K	752	293.65 $K \le T \le 961.05$ K 0.1 $MPa \le P \le 127.8$ MPa 0.7 $\frac{kg}{m^3} \le \rho \le 1145$ $\frac{kg}{m^3}$	AAD=2.31% R^2=0.997	Genetic programming	$SC-CO_2$
Rostamian and Lotfollahi [50]	$\lambda = \frac{A_1 + \frac{A_2}{T} + \frac{A_3}{T^{2.5}} + A_4\rho + A_5\rho^{2.5}T^{2.5} + A_6\rho^3}{1 + \frac{A_7}{T} + \frac{A_8}{T^{2.5}} + \frac{A_9}{T^3} + A_{10}\sqrt{\rho T}}$ where $A_1 = -29.9717451505165E+15$; $A_2 = 9.65637447009372E+18$; $A_3 = -13.8288944829492E+21$; $A_4 = -21.1152877719961E+12$; $A_5 = 9.26006733304733$; $A_6 = -30.7171646680127E+6$; $A_7 = -408.256276723566E+15$; $A_8 = 130.491020289031E+18$; $A_9 = -13.4237924607890E+21$; and $A_{10} = 60.9547298940653E+9$. λ is expressed in mW/m.K	2319	250 $K \le T \le 1100K$ 0.3 $\frac{kg}{m^3} \le \rho \le 1400$ $\frac{kg}{m^3}$	AAD=1.98% R^2=0.995	Error minimization using Nedler-Mead optimization method	Liquid, vapor and $SC-CO_2$

* Average Absolute Deviation.
** Coefficient of Determination.

information and a pertinence for identifying the relationship describing the system. Multilayer perceptron (MLP) and radial basis function neural networks (RBFNN) are among the well-formulated ANN types for the modeling tasks. These two types differ in their topologies and learning procedures [35,66].

An MLP model can include one or more hidden layers according to the system's complexity. In this context, one hidden layer is generally deemed adequate for modeling low-complexity systems, whereas more than one hidden layer is mandatory for complex systems [54]. Trial and error is the commonly considered technique to investigate the proper number of hidden layers and their covered neurons [65]. For each of the hidden layers, a nonlinear activation function is implemented to map the data in higher dimension space and capture the complexity of the system accordingly. The frequently applied activation functions in the hidden layers are shown below:

$$\text{Logsig}: f(x) = \frac{1}{e^x + 1} \tag{1}$$

$$\text{Tansig}: f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \tag{2}$$

The output of the MLP is gained from the output layer by applying generally a linear function known as Pureline. The latter is expressed as:

$$g(x) = x \tag{3}$$

To build an accurate predictive MLP model, it should go through the training phase. The objective of this step is to investigate the appropriate bias and weights that generate the best performances [67]. The so-called back-propagation approaches are applied for this purpose. Levenberg-Marquardt Algorithm (LMA), Scaled Conjugate gradient (SCG), Bayesian Regularization (BR), and Resilient Backpropagation (RB) are among the well-known optimization techniques for improving the MLP learning procedure. Therefore, the aforementioned algorithms were applied in the present work. Optimization tasks related to MLP and the learning algorithms were performed using MATLAB®. More details about these algorithms can be found in previous works [16,35,68].

2.2. Radial basis function neural network (RBFNN)

Radial basis function neural network (RBFNN) is another rigorous and reliable ANN type. The topology of RBFNN contains only one hidden layer [69,70], represented by its nodes (N_h) which include bias terms (N_b). Besides, a radial basis function (RBF) is applied in each hidden node to transform the data and recognize the system. The applied RBFs are characterized by their centers and width. Gaussian function is the most implemented RBF in this kind of ANN [65]. It is formulated based on its center (c_k) and the spread coefficient (σ^2).

During the processing steps, the Euclidian norm is applied to assess the distance of a vector ($\boldsymbol{x}$) from the center (c_k) of the Gaussian function. Accordingly, the following expression is obtained:

$$r_k = \sqrt{\sum_{i=1}^{D} (x_i - c_{ik})^2} \tag{4}$$

where D denotes the dimension of the system and c_{ik} represents the centers. Combining the above equation with the Gaussian RBF, the following formula is gained:

$$\varphi(r_k) = \exp\left[\frac{\sum_{i=1}^{D} (x_i - c_{ik})^2}{2\sigma^2}\right] \tag{5}$$

where the parameter σ^2 represents the spread coefficient of the ϕ Gaussian RBF.

The outcome from the RBFNN output layer is obtained as shown below:

$$y = \sum_{i=1}^{N_h} w_i \varphi_i(r) + b_i, i = 1, \ldots, N_h \tag{6}$$

where b_i represents the bias term, N_n is the number of the hidden layer's neurons, and w_i is the connecting weight between the hidden node i and the output layer.

To improve the reliability of RBFNN, we have applied particle swarm optimization (PSO) for investigating the RBFNN control parameters, namely the spread coefficient and the number of nodes. More details about PSO can be found in the published literature [71,72].

2.3. Committee machine intelligent system (CMIS)

In most of the published works related to modeling tasks based on soft computing techniques, an evaluation of the prediction performance is done, and the most reliable prediction is kept as the main representative paradigm for the studied phenomenon. Nevertheless, for the purpose of improving the prediction accuracy, the most reliable implemented soft computing models can be combined under a unified paradigm by means of the so-called Committee Machine Intelligent System (CMIS). The first combination scheme was developed by Nilsson [73]. In recent years, the linear linking technique is widely applied for developing CMIS models [35,74]. However, in the present work, a new scheme based on group method of data handling (GMDH) is proposed for building CMIS model.

2.3.1. Linear linking (CMIS-Linear)

A CMIS based on the linear combination of the models is established by means of weighted average method with an appended bias factor [75].

2.3.2. Group method of data handling linking (CMIS-GMDH)

Group Method of Data Handling (GMDH) is an ANN type which processes the information in a polynomial form [76]. GMDH encompasses nodes configurated in the form of layers, allowing the interaction between the input parameters in a progressive manner. The earliest version of GMDH that was proposed by Ivakhnenko [77] is based on a quadratic polynomial scheme applied to the nodes of the previous layer. However, this version presented some accuracy shortcomings, and hence, an improved form called hybrid GMDH was proposed [78]. The hybrid version can consider higher order polynomials and enables interactions between nodes pertaining to various layers [79]. The main formula that describes the hybrid version is shown below:

$$y = a + \sum_{i=1}^{D} \sum_{j=1}^{D} \cdots \sum_{k=1}^{D} d_{ij\ldots k} x_i^n x_j^n \ldots x_k^n n = 1, 2, \ldots, 2^l \tag{7}$$

In this equation, a is the bias term, y, $x_i, x_j, \ldots x_k$ represent the output and the inputs of the model, respectively; $d_{ij\ldots k}$ are the polynomial coefficients; l and D correspond to the numbers of layers and input parameters, respectively.

In addition, partial polynomials with specified orders are applied to link the nodes and to form other intermediate variables (O_1, $O_{2,\ldots}$). The following equation illustrates an example based on quadratic polynomial:

$$O^{GMDH} = c_0 + c_1 x_k + c_2 x_k + c_3 x_k x_j + c_4 x_k^2 + c_5 x_j^2 \tag{8}$$

The coefficients in the above equation are calculated using the least square method (LSM). To this end, the following equation is obtained:

$$\Delta_j^2 = \sum_{z=1}^{N_t} \left(t_z - y_z^{GMDH}\right)^2 j = 1, 2, \ldots, \binom{D}{2} \tag{9}$$

where D and N_t refer to the number of input parameters and training datapoints, respectively, and t_z and y_z denote the targets and the GMDH predictions, respectively. The subscript z is the index of the output.

Afterwards, the matrix form is considered to resolve the problem [80]:

$$y = C^T X \tag{10}$$

The LSM results in the final solution as shown below:

$$C^T = yX^T\left(XX^T\right)^{-1} \tag{11}$$

where $y = \{y_1, y_2, \ldots, y_D\}$, $C = \{c_0, c_1, c_2, c_3, c_4, c_5\}$, and D denotes the number of input parameters.

In the present work, the hybrid version of GMDH was applied to build the CMIS-GMDH. It should be mentioned that the initial nodes of GMDH were the best found paradigms.

3. Data acquisition and preparation

To establish accurate predictive models for estimating CO_2 thermal conductivity, an extensive experimental database including 5893 datapoints was collected from various literature [45,46,81–89]. A detailed description of the collected data is illustrated in Tables 2 and 3. Temperature (T) expressed in K and pressure (P) expressed in MPa are the considered input parameters for establishing the models, while the output is the CO_2 thermal conductivity (λ_{CO_2}) expressed in (mW/m.K). According to the information in Table 2, the collected experimental data covered wide operational conditions with temperature values from 217.93 K to 961.05 K and pressures varying from 0.097 MPa to 209.763 MPa. It is worth mentioning that this reported literature data was converted into readily available data by unifying the units of the different variables and conversion into one csv file. The latter was imported automatically into the codes of the considered data-driven techniques.

Before proceeding to the learning phase of the proposed soft computing approaches, the collected data were normalized between −1 and 1 using the expression below:

$$x_{ni} = \frac{2(x_i - x_{min})}{(x_{max} - x_{min})} - 1 \tag{12}$$

where x_{ni} is the normalized value of x_i, x_{max} and x_{min} are the maximum and minimum values of the variable x (corresponding to λ_{CO_2}, T and P), respectively.

After the normalization, the database was divided randomly into training (80% of the experimental datapoints) and testing (the remaining 20%) sets. The training set was utilized for building the models and the testing set was used for checking their accuracy with unseen real data. It is worth mentioning that this dataset apportioning shows typically extremely good outcomes [35,76,78,90–92]. The models were run several times with this partitioning of the data, and the best models are exhibited in the following section.

Table 2
Summary of the gathered data.

Data	Max	Avg.	Min	Standard deviation (SD)
T (K)	961.05	421.60	217.93	150.15
P (MPa)	209.763	20.0119	0.097	22.9104
Thermal Conductivity (mW/m.K)	198.79	59.8571	10.82	41.9814

4. Results and discussion

4.1. Implementation procedure

As mentioned in the previous sections, the aim of the proposed soft computing schemes is to build accurate models for estimating CO_2 thermal conductivity (λ_{CO_2}) under various operational conditions. Therefore, the following equation is considered:

$$\lambda_{CO_2} = f(T, P) \tag{13}$$

During the training phase of different soft computing methods, mean square error (MSE) was the considered fitness function to assess the reliability of the approaches. MSE is expressed as shown below:

$$MSE = \frac{\sum_{i=1}^{N}\left(\lambda_{iexp} - \lambda_{ipre}\right)^2}{N} \tag{14}$$

where λ_i denotes CO_2 thermal conductivity, N is the number of training datapoints and the subscripts *pre* and *exp* refer to the predicted values and the experimental measurements, respectively.

For the case of MLP, trial and error technique was applied to investigate the appropriate structures for each considered algorithm. The inputs as well as the outputs are provided during the training phase. The network then processes the inputs and matches them up to the specified outputs. Errors will then be propagated back across the system, and the weights and bias of the network will be regulated and adjusted iteratively. This process is constantly repeated until a stopping condition is fulfilled. The resulting paradigms are denoted MLP-LMA, MLP-BR, MLP-SCG and MLP-RB, respectively. The best found structure in each of the aforementioned MLP models includes three hidden layers with *Tansig* as activation function in each of them. Moreover, *Pureline* was the considered transfer function in the output layer of these models. The number of neurons in these hidden layers are 12, 11 and 9, respectively.

For RBFNN, PSO was applied to optimize its control parameters. The optimization results revealed that the appropriate number of neurons and spread coefficient are 162 and 0.1936, respectively. The resulting hybridization is symbolized RBFNN-PSO.

After accomplishing the establishment of the ANN models, a performance comparison was done, and the two fittest models were used for implementing two other paradigms, namely CMIS-Linear and CMIS-GMDH. It is worth noting that the procedure described in Section 2 was adapted for developing these schemes.

The workflow illustrated in Fig. 1 briefly summarizes the steps in the model implementation.

Table 3
Details about the gathered data.

Reference	Thermal Conductivity (mW/m.K)	P (MPa)	T (K)	Number of datapoints
[83]	16.5–177.5	0.1–127.8	293.65–960.85	536
[85]	16.4–179	0.1–120	298.15–951.15	193
[89]	16.32–194.68	0.1–209.763	298.166–348.29	248
[87]	17.12–41.75	0.497–15.1	298.89–428.05	77
[82]	10.82–198.79	0.097–68.702	217.931–757.014	4389
[86]	17.44–99.9	0.3–24.56	301.25–349.52	92
[46]	24.22–61.57	1.83–30.6	380.92–474.31	46
[84]	17.73–28.74	0.67–6.66	305.32–426.19	90
[81]	17.38–29.2	0.601–5.92	301.34–303.59	22
[88]	15.30–190.60	3.11–48.7043	298.125–308.07	200

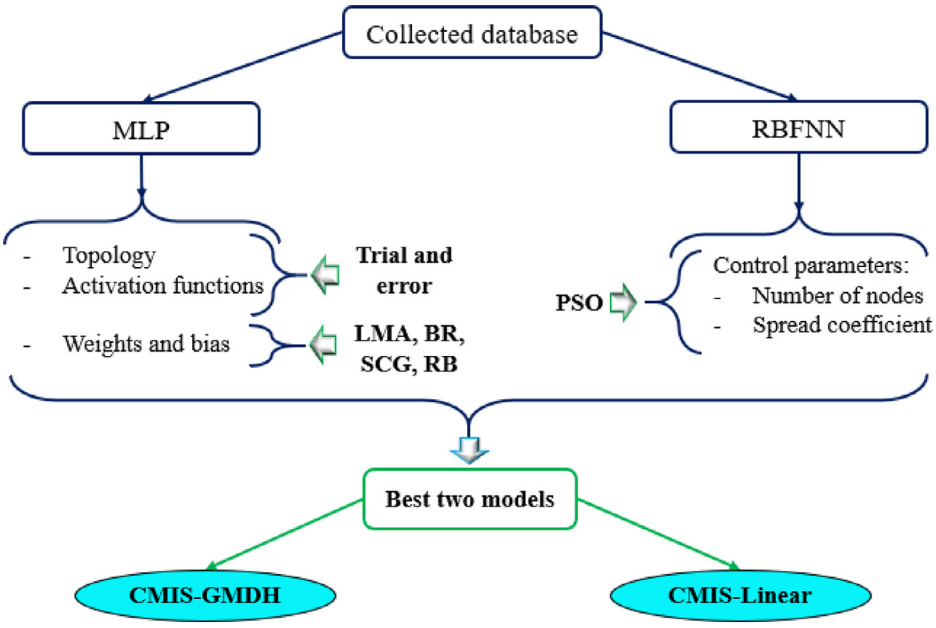

Fig. 1. The flowchart of the implementation procedure.

4.2. Statistical and graphical evaluation

Various statistical criteria were applied for assessing the reliability of the proposed models to predict CO_2 thermal conductivity. These statistical indices are expressed as follows:

1. Average Absolute Relative Deviation (AARD)

$$AARD\% = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{\lambda_{iexp}-\lambda_{ipred}}{\lambda_{iexp}}\right| \times 100 \tag{15}$$

2. Root Mean Square Error (RMSE)

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\lambda_{iexp}-\lambda_{ipred}\right)^2} \tag{16}$$

3. Standard Deviation (SD)

$$SD = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}\left(\frac{\lambda_{iexp}-\lambda_{ipred}}{\lambda_{iexp}}\right)^2} \tag{17}$$

4. Coefficient of Determination (R^2)

$$R^2 = 1-\frac{\sum_{i=1}^{n}\left(\lambda_{i\ exp}-\lambda_{i\ pred}\right)^2}{\sum_{i=1}^{n}\left(\lambda_{i\ exp}-\overline{\lambda}_{i\ pred}\right)^2} \tag{18}$$

where λ_{iexp} and λ_{ipred} represent the experimental and predicted CO_2 thermal conductivity, respectively. $\overline{\lambda}$ is the average thermal conductivity of CO_2, and N is the number of samples.

The evaluation of the proposed ANN models including MLP optimized with LMA, BR, SCG and RB, and RBFNN-PSO using these indices is depicted in Table 4. In this table, the evaluation is reported separately for training and testing sets and also for the whole utilized database. As seen in Table 4, MLP-LMA outperforms the other MLP models as well as RBFNN-PSO. MLP-LMA exhibited AARD% values of 1.2213%, 1.5759% and 1.2922 for the training set, testing set and overall data, respectively. According to Table 4, the established ANN models can be ranked in the following descending order: $MLP-LMA > MLP-BR > MLP-SCG > RBFNN-PSO > MLP-RB$. Therefore, the two best models, i.e. MLP-LMA and MLP-BR were kept for developing the two CMIS schemes.

Fig. 2 illustrates the structures of the proposed CMIS-Linear and CMIS-GMDH models. Furthermore, the mathematical formulations of these CMIS are shown below:

- CMIS-Linear:

$$\lambda_{CO_2} = a_0 + a_1 \times M_1 + a_2 \times M_2 \tag{19}$$

where the coefficients $a_0 = -0.0290$, $a_1 = 0.7835$, *and* $a_2 = 0.2169$ were obtained by regression, and M_1 and M_2 refer to the MLP-LMA and MLP-BR models, respectively.

- CMIS-GMDH:

$$\begin{aligned}\lambda_{CO_2} = {} & -0.016363 + 2.187283 \times N - 1.186711 \times M_1 \\ & -0.817671 \times M_1 \times N + 0.396172 \times N^2 \\ & + 0.421492 \times M_1^2\end{aligned} \tag{20}$$

Table 4
Performance evaluation of the implemented models.

		AARD (%)	R^2	RMSE	SD
Training data	MLP-LMA	1.2213	0.9991	1.0608	0.0003
	MLP-BR	1.2845	0.9987	1.2495	0.0004
	MLP-SCG	2.3963	0.9897	3.4930	0.0018
	MLP-RB	2.7941	0.9915	3.1677	0.0021
	RBF-PSO	2.6195	0.9896	3.5534	0.0019
	CMIS-LINEAR	1.2347	0.9986	1.1779	0.0003
	CMIS-GMDH	0.8372	0.9997	0.6966	0.0002
Test data	MLP-LMA	1.5759	0.9987	1.2025	0.0005
	MLP-BR	1.4581	0.9979	1.5636	0.0004
	MLP-SCG	2.7468	0.9918	3.2603	0.0023
	MLP-RB	2.8965	0.9789	5.2664	0.0026
	RBF-PSO	2.9461	0.9881	3.7283	0.0025
	CMIS-LINEAR	1.1537	0.9881	0.9127	0.0002
	CMIS-GMDH	0.8407	0.9997	0.6836	0.0002
All data	MLP-LMA	1.2922	0.9990	1.0906	0.0004
	MLP-BR	1.3193	0.9986	1.3183	0.0004
	MLP-SCG	2.4664	0.9900	3.4477	0.0019
	MLP-RB	2.8146	0.9887	3.6843	0.0022
	RBF-PSO	2.6848	0.9893	3.5891	0.0021
	CMIS-LINEAR	1.2185	0.9991	1.1248	0.0003
	CMIS-GMDH	0.8379	0.9997	0.6941	0.0002

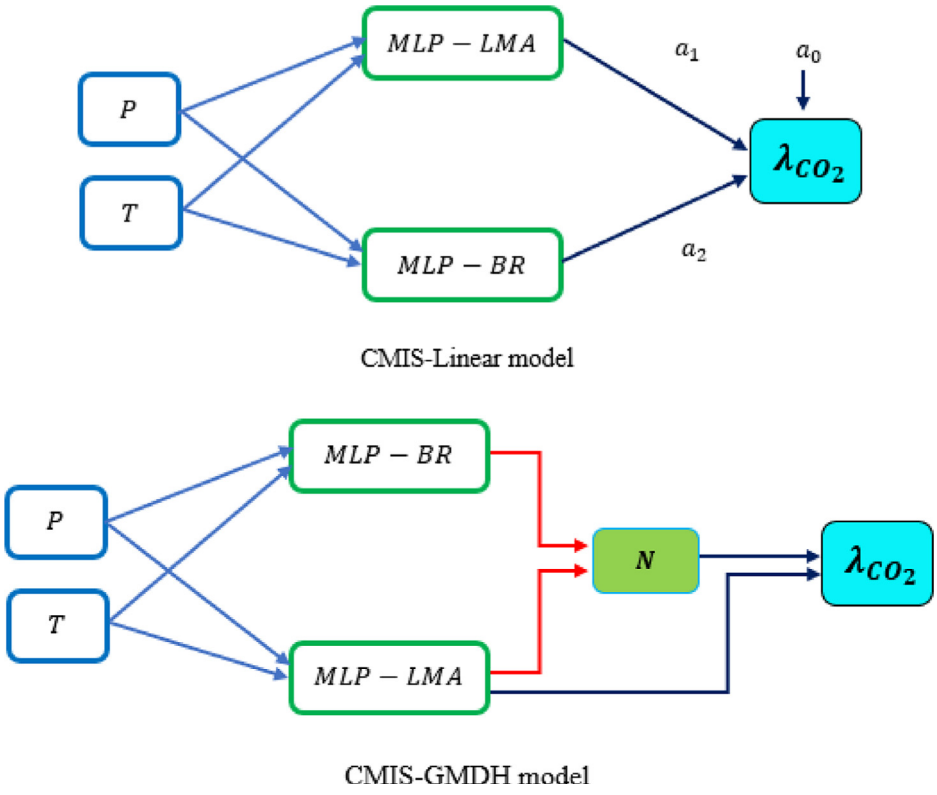

Fig. 2. CMIS-Linear and CMIS-GMDH models proposed in this study.

where M_1 refers to the MLP-LMA model and N is expressed as follows:

$$N = 0.015045 + 0.443216 \times M_2 + 0.5565212 \times M_1 - 0.029334 \times M_1 \times M_2 + 0.013144 \times M_2^2 + 0.016189 \times M_1^2 \quad (21)$$

M_2 represents the MLP-BR model.

The statistical evaluation of these CMIS models is demonstrated in Table 4. By comparing the performance of these paradigms against various ANN models, it can be deduced that the proposed CMIS schemes provide better prediction abilities. In addition, a detailed analysis of the reported information in Table 4 reveals that the newly proposed CMIS-GMDH outperforms CMIS-Linear. CMIS-GMDH achieved an overall AARD% value of 0.8379% and a total coefficient of determination (R^2) equal to 0.9997. Therefore, CMIS-GMDH is the model further used in this paper.

To extend the performance analysis of CMIS-GMDH, various graphical plots such as cross plot, error distribution diagrams and cumulative frequency of AARD% were illustrated for the visual evaluation. Cross plots show the distribution of model's prediction nearby the line $X = Y$ that emulates the perfect case. Existence of a great portion of the predictions nearby this line demonstrates the reliability of the paradigm. Error distribution diagrams exhibit the distribution of the relative error denoted in the predictions of the model versus the real output or as a function of the independent variables. Low relative errors in these distributions represent an indication of the high accuracy of the model. Cumulative frequency diagrams of AARD% depict different portions of the data predicted according to a specified AARD% value.

Fig. 3 illustrates the scatter plot of the outputs of CMIS-GMDH model versus the corresponding measured λ_{CO_2} for the entire datapoints. In addition, subplots of Fig. 4 show zoom-in into the cross plot of Fig. 3 to display the data overlap for various intervals of λ_{CO_2}.The predictions of CMIS-GMDH for both training and testing sets, and for the whole intervals of λ_{CO_2}, lie close to the unit-slope line, showing the approach's accuracy. Moreover, the relative deviations of the outcomes obtained via CMIS-GMDH versus corresponding real CO_2 thermal conductivity values are shown in Fig. 5. As illustrated, the outcomes of the newly implemented paradigm (CMIS-GMDH) have maximum deviation of 9.60% from real λ_{CO_2} data.

The cumulative frequency of CMIS-GMDH versus AARD% for the prediction of λ_{CO_2} is displayed in Fig. 6. This figure reveals that the proposed model predicts λ_{CO_2} of more than 70% of the datapoints with absolute relative deviation of less than 1%, and for 90% of the data, the absolute relative deviation is less than 2%.

The 2D and 3D contour map depictions of relative error of the proposed CMIS-GMDH model are shown, respectively, in Figs. 7 and 8 versus the independent input parameters, namely pressure and temperature. Based on the color changes, it can be deduced that CMIS-GMDH model has very satisfactory performance for different ranges (i.e. low, medium and high) of pressure and temperature. Furthermore, these figures verify that a huge portion of the model's predictions has very low absolute relative error.

The performance evaluation reported in Table 4 and Figs. 3–8 confirms that the implemented approach remarkably exhibits excellent prediction abilities.

4.3. Comparison of CMIS model with prior correlations

The statistical assessment is conducted to examine the accuracy of the implemented CMIS-GMDH, compared to the best prior correlations, namely those of Bahadori and Vuthaluru [43], Jarrahian and Heidaryan [47] and Rostami et al. [49]. The value of statistical indices of these correlations and CMIS-GMDH model are stated in Table 5. It is worth mentioning that in order to ensure the comparison fairness, only the datapoints respecting the correlations' applicability conditions were considered from the utilized data. Accordingly, it can be seen from Table 5 that the prior models are somewhat limited from

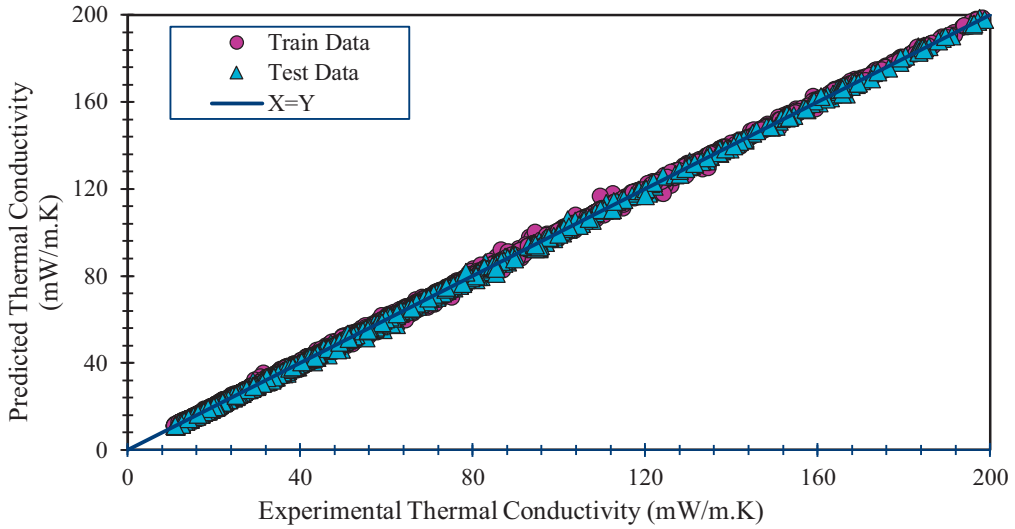

Fig. 3. Scatter plot of the outputs of CMIS-GMDH model versus the corresponding measured CO_2 thermal conductivity for training data set (4714 points) and testing data set (1180 points).

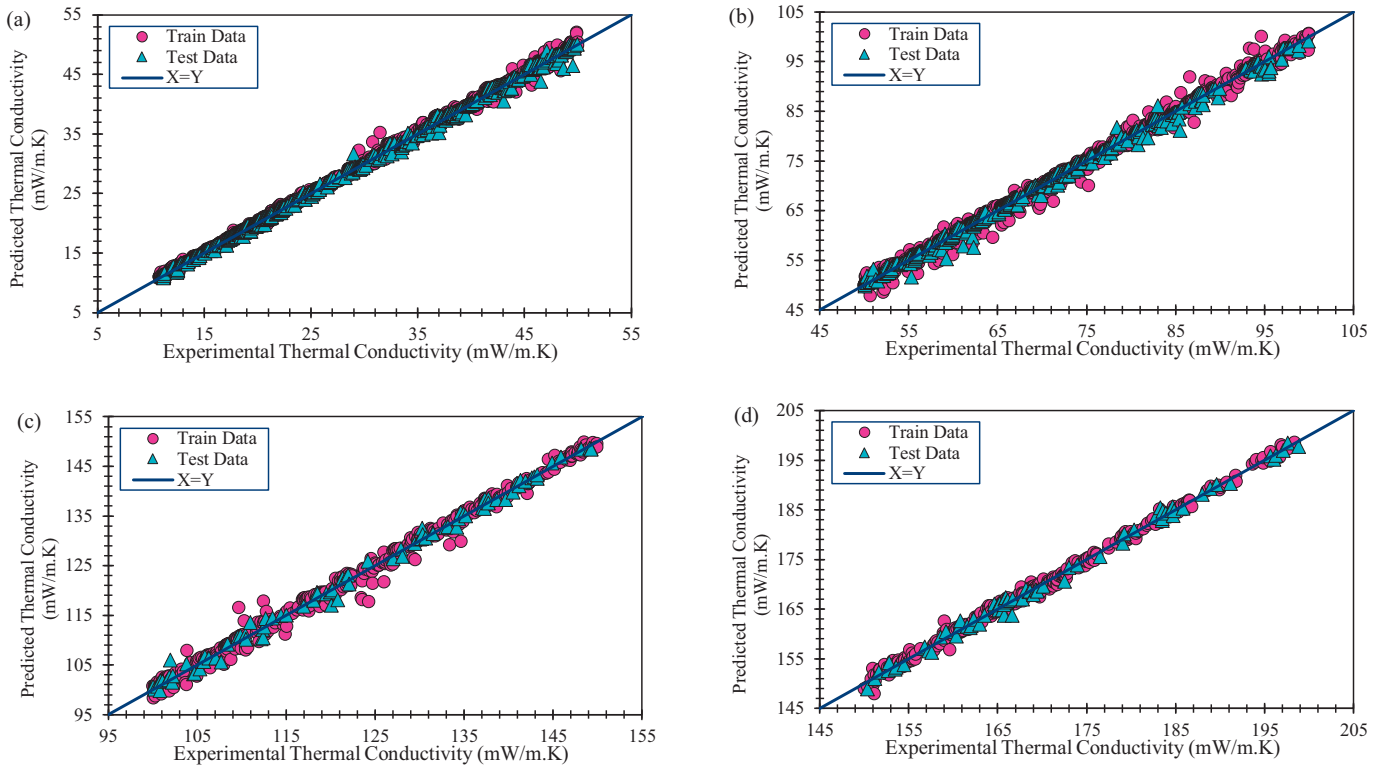

Fig. 4. Detailed scatter plots of the outputs of CMIS-GMDH model versus the corresponding measurements for different intervals of CO_2 thermal conductivity values: (a) 10.82–50 mW/m.K; (b) 50–100 mW/m.K; (c) 100–150 mW/m.K; (d) 150–198.79 mW/m.K.

the applicability perspective, where Rostami et al. [49] correlation covers the greatest number of datapoints, namely 4543 datapoints, with applicable operational conditions, while Bahadori and Vuthaluru [43] correlation is deemed appropriate in only 1219 experimental datapoints. As demonstrated in Table 5, the prior approaches showed moderate to good performance, where the worst AARD% value was achieved by Rostami et al. [49] correlation. It is evident from Table 5, that CMIS-GMDH surprisingly exhibits higher accuracy and outperforms all the existing explicit paradigms for the estimation of CO_2 thermal conductivity. It is worth mentioning that the improvement in the predictions of CO_2 thermal conductivity shown by the implemented CMIS-GMDH model can be explained by the concept of this approach which ensures the proper interactions between the best models and their linking in a single and robust scheme, and hence, it provides an efficient strategy for dealing with big data. Besides, it can be added that the proposed CMIS-GMDH model covers wider range of operational conditions compared with the existing correlations and also the prior intelligent schemes such as those proposed by Shams et al. [60] (LSSVM-CSA), Tatar et al. [61] (ANFIS-PSO) and Ahmadi and Baghban [62] (LSSVM-GA). In this context, although these existing soft computing-based paradigms achieved satisfactory prediction performance (the worst AARD value was 2.63448%, exhibited by Tatar et al. [61] model), all of them were resulted from a limited number of experimental datapoints (558, 1042 and 746 points for Shams et al. [60], Tatar et al. [61]

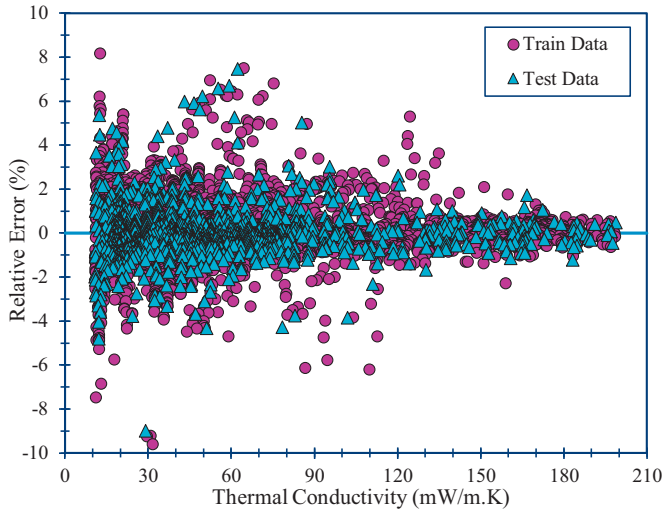

Fig. 5. Relative error of the predictions of CMIS-GMDH model against corresponding real values of CO_2 thermal conductivity for training data set (4714 points) and testing data set (1180 points).

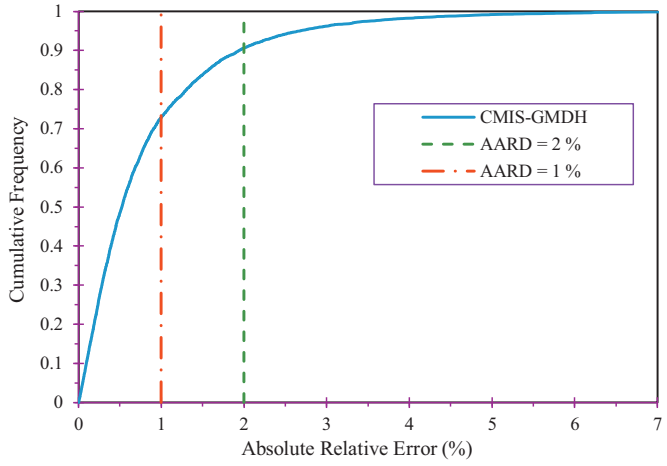

Fig. 6. Cumulative frequency vs. absolute relative deviation of CMIS-GMDH.

and Ahmadi and Baghban [62], respectively) compared to our proposed CMIS-GMDH model (5893 points).

4.4. Validity of CMIS-GMDH model in terms of the operational conditions

In this step of evaluation, the trend of outcomes of CMIS-GMDH with respect to pressure and temperature is investigated based on various experimental samples. Subplots (a), (b) and (c) of Fig. 9 display the emulated CO_2 thermal conductivity by CMIS-GMDH and the real experimental values versus pressure at constant (or almost constant) temperature values. Clearly, the proposed CMIS-GMDH demonstrates an excellent fit with the experimental data under different operational conditions, implying the integrity and the appropriateness of the implemented CMIS-GMDH in terms of physical interpretation.

4.5. Applicability domain of the CMIS-GMDH model and the quality of the experimental data

In the last step of this work, the well-known Leverage approach was performed to define the applicability realm of the proposed CMIS-GMDH and to distinguish the suspected experimental measurements from the employed database. The method provides a measure of the distance between the datapoints and the data used in the training phase [93]. The Leverage technique is based on the so-called Hat matrix, which is gained using [94,95]:

$$H = X(X^T X)^{-1} X^T \tag{22}$$

where X is a $N \times D$ matrix. N and D are the number of samples and variables, respectively; and X^T is the transpose of the matrix. The diagonal elements of H represent the hat (h_i) known as leverages, which show the distance of the data from the centroid of X [93].

In addition, the standard residual (SR) is a useful parameter for implementing this technique. It corresponds to simple residual "R" (i.e. the difference between the real and the model predicted values) divided by their estimated standard deviation. The parameters R and SR are calculated as follows [96]:

$$R_i = \lambda_{iexp} - \lambda_{ipred} \tag{23}$$

$$SR_i = \frac{R_i}{\sqrt{MSE(R_i) \times (1 - h_i)}} \tag{24}$$

where MSE denotes Mean Squared Error.

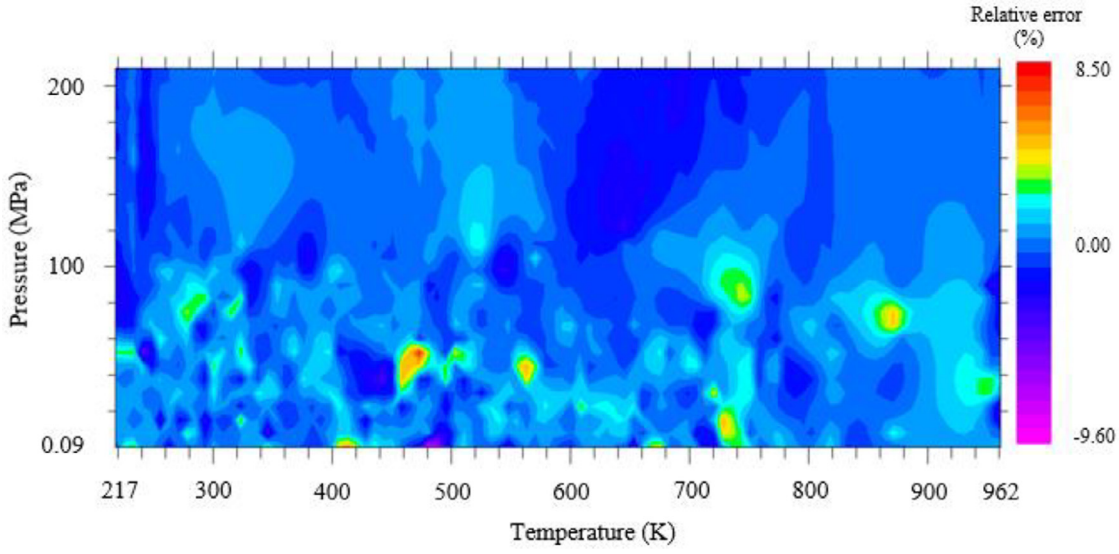

Fig. 7. Contour map representation of relative error in the 2D continuous map for CMIS-GMDH: variation of relative error versus pressure and temperature.

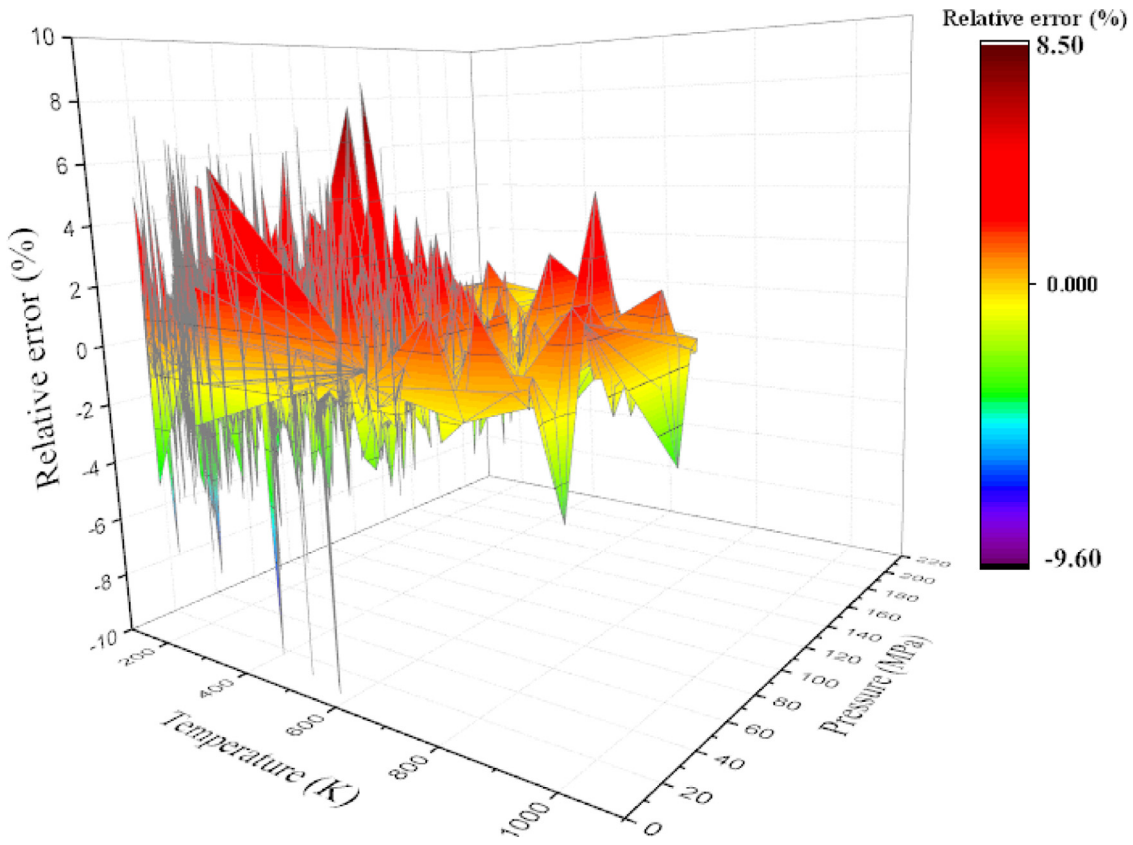

Fig. 8. Contour map representation of relative error in the 3D continuous map for CMIS-GMDH: variation of relative error versus pressure and temperature.

The two main parameters in the Leverage approach, namely SR and the hat are exploited in graphical form which is known as William's plot that depicts the applicability domain of the model. To this end, a warning Leverage limit (H^*) is obtained as: $\frac{3(D+1)}{N}$; and for the CMIS-GMDH model, H^* is equal to 0.0015. It can be stated that the model is statistically valid if a significant part of the predictions is located within the ranges of $0 \leq H \leq H^*$ and $-3 \leq R \leq 3$. In addition, some statistical terms can be obtained from this approach and the resulting plot: suspected data are located within the section with R greater than 3 or less than −3 regardless of the associated H value; out of Leverage data are within the domain $H^* \leq H$ and $-3 \leq R \leq 3$. William's plot is sketched in Fig. 10. Obviously, having a great part of the datapoints within the ranges $0 \leq H \leq 0.0015$ and $-3 \leq R \leq 3$ demonstrates that the evolved approach is statistically confirmed. Furthermore, only small portion of the experimental datapoints (3.46%) are noticed out of the applicability domain.

5. Conclusions

Throughout this investigation, enormous attempts have been put forward to implement new accurate paradigms for predicting CO_2 thermal conductivity under various temperature and pressure conditions. To achieve this, a massive representative experimental database was gathered from the published researches and was fed into various soft computing schemes to evolve and test their reliability. The database includes 5893 experimental measurements of CO_2 thermal conductivity. The included soft computing techniques were MLP optimized with LMA, BR, SCG and RB; RBFNN coupled with PSO, and finally two committees of models were used to combine the two best found ANN paradigms into two single models, namely CMIS-Linear and CMIS-GMDH. The following conclusions can be drawn on the basis of the results achieved:

1. The proposed models cover extensive operational conditions and can be applied to temperature and pressure values in the ranges of 217.931 to 961.05 K and 0.097 to 209.763 MPa, respectively.
2. Based on the statistical evaluations, the values of CO_2 thermal conductivity obtained by the implemented approaches yield very satisfactory agreement with the real measurements.
3. Among the proposed ANN models, MLP-LMA and MLP-BR presented the best agreement with the experimental data, and hence, they were combined under CMIS-Linear and CMIS-GMDH.
4. The proposed CMIS paradigms exhibited better performance compared with the ANN models.

Table 5
Comparison of the model performance with recent models.

	CMIS-GMDH	Bahadori and Vuthaluru [43]	Jarrahian and Heidaryan [47]	Rostami et al. [49]
AARD (%)	0.8379	3.286	3.5409	3.8918
Max ARD (%)	9.60	56.66	61.1083	76.6777
R^2	0.9997	0.9624	0.9527	0.9670
RMSE	0.6941	3.0565	2.9	8.1735
SD	0.0002	0.0017	0.0042	0.0071
Number of applicable points	5893	1219	2621	4543

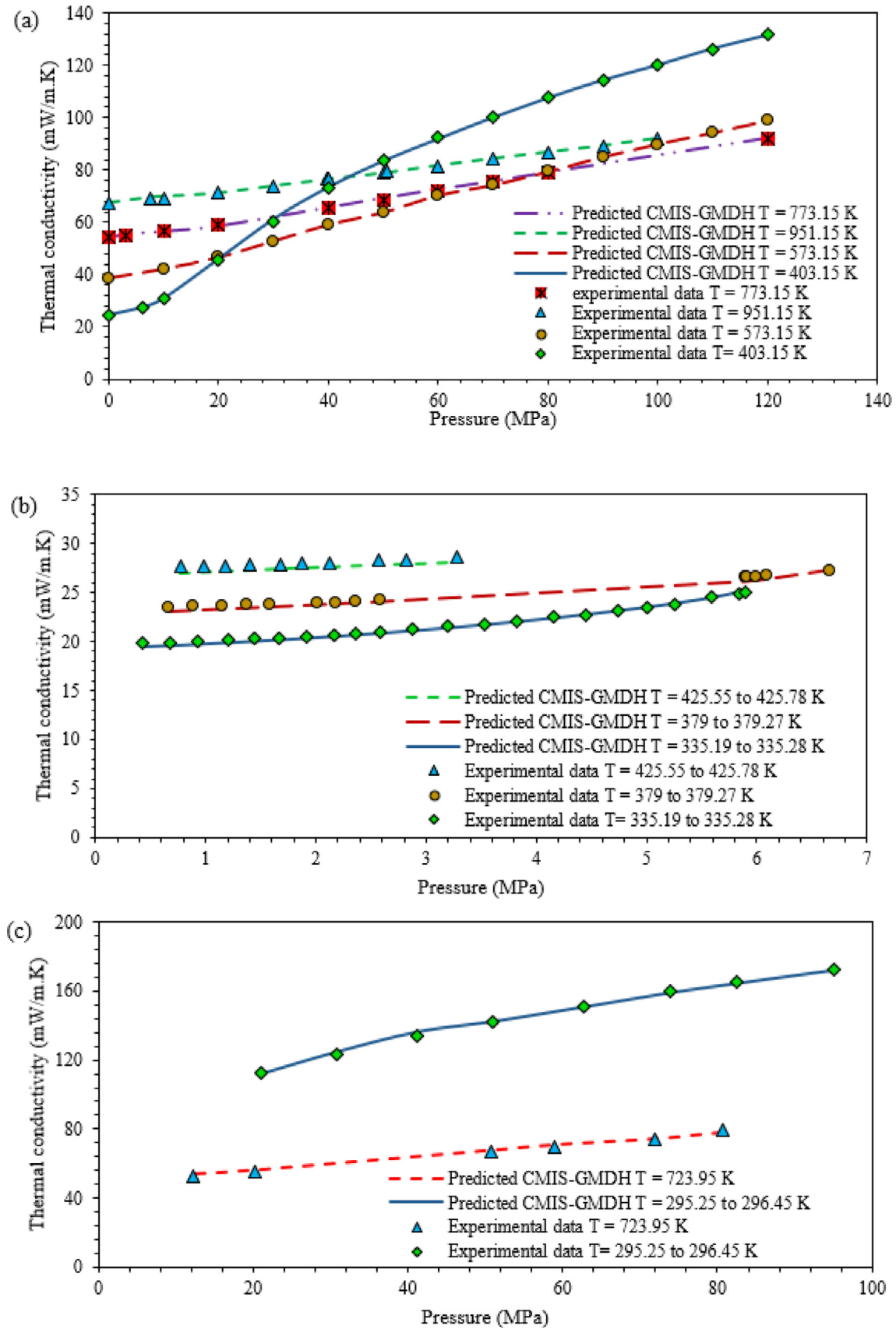

Fig. 9. Experimental CO_2 thermal conductivity and predictions of CMIS-GMDH versus pressure at constant (or almost constant) temperature values: (a) data from Le Neindre (1972) [85]; (b) data from Millat et al. (1987) [84]; and (c) data from Leneindre et al. (1973) [83].

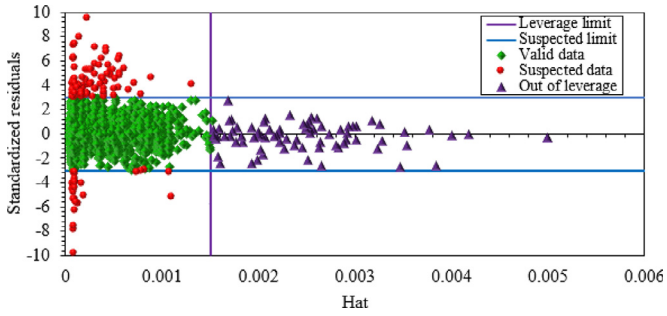

Fig. 10. The William's plot of CO_2 thermal conductivity dataset for CMIS-GMDH model (5689 points are valid; 124 points are suspected, and 80 points are out of leverage).

5. CMIS-GMDH was deemed the best developed model with an overall AARD% value of 0.8379% and a coefficient of determination (R^2) equal to 0.9997.
6. The existing correlations and models fail to cover the whole considered database of CO_2 thermal conductivity.
7. The implemented CMIS-GMDH outperforms all the existing explicit correlations in the prediction of CO_2 thermal conductivity.
8. The trends of the resulting outcomes from the CMIS-GMDH model are logical in terms of pressure and temperature.
9. The Leverage approach revealed the statistical validity of the model and the employed data, where more than 96% of the considered experimental datapoints lie in the applicability domain of CMIS-GMDH.
10. The best-found predictive model can be utilized by means of the provided Excel macro to estimate the thermal conductivity of CO_2 (Appendix A).

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.

Supplementary materials

Supplementary material associated with this article can be found in the online version at doi:10.1016/j.jtice.2020.08.001.

Appendix A. CMIS-GMDH implemented model

To apply the established CMIS-GMDH paradigm, please open the macros titled "CO2_TC_calculator.xlsm".

Note: The macros must be activated in your Excel. The program predicts the CO_2 thermal conductivity by specifying the values of pressure (MPa) and temperature (K) and clicking on the calculate button.

References

[1] Knight J, Harrison S. The impacts of climate change on terrestrial Earth surface systems. Nat Clim Chang 2013;3:24–9.
[2] Abas N, Khan, Carbon N. conundrum, climate change, CO2 capture and consumptions. J CO_2 Util 2014;8:39–48.
[3] AghaKouchak A, Feldman D, Hoerling M, Huxman T, Lund J. Water and climate: recognize anthropogenic drought. Nat News 2015;524:409.
[4] Boot-Handford ME, Abanades JC, Anthony EJ, Blunt MJ, Brandani S, Mac Dowell N, et al. Carbon capture and storage update. Energy Environ Sci 2014;7:130–89.
[5] Singh J, Basu S, Bhunia H. CO2 capture by modified porous carbon adsorbents: effect of various activating agents. J Taiwan Inst Chem Eng 2019;102:438–47.
[6] Muromachi S, Shijima A, Miyamoto H, Ohmura R. Experimental measurements of carbon dioxide solubility in aqueous tetra-n-butylammonium bromide solutions. J Chem Thermodyn 2015;85:94–100.
[7] Anastas PT, Kirchhoff MM. Origins, current status, and future challenges of green chemistry. Acc Chem Res 2002;35:686–94.
[8] Chiang Y-C, Juang R-S. Surface modifications of carbonaceous materials for carbon dioxide adsorption: a review. J Taiwan Inst Chem Eng 2017;71:214–34.
[9] Montazerolghaem M, Aghamiri SF, Talaie MR, Tangestaninejad S. A comparative investigation of CO_2 adsorption on powder and pellet forms of MIL-101. J Taiwan Inst Chem Eng 2017;72:45–52.
[10] Li H, Dong B, Yu Z, Yan J, Zhu K. Thermo-physical properties of CO2 mixtures and their impacts on CO_2 capture, transport and storage: progress since 2011. Appl Energy 2019;255:113789.
[11] Yan J, Zhang Z. Carbon capture, utilization and storage (CCUS). Appl Energy 2019;235:1289–99.
[12] Yang L, Xu M, Yang Y, Fan J, Zhang X. Comparison of subsidy schemes for carbon capture utilization and storage (CCUS) investment based on real option approach: evidence from China. Appl Energy 2019;255:113828.
[13] Wood DA. Carbon dioxide (CO2) handling and carbon capture utilization and sequestration (CCUS) research relevant to natural gas: a collection of published research (2009–2015). J Nat Gas Sci Eng 2015;100:A1–9.
[14] Metz B, Davidson O, de Coninck H, Loos M, Meyer L. Carbon dioxide capture and storage. IPCC Spec Rep 2005:342.
[15] Ayirala SC, Rao DN. Comparative evaluation of a new MMP determination technique. SPE/DOE Symp Improv Oil Recov 2006.
[16] Nait Amar M, Zeraibi N. An efficient methodology for multi-objective optimization of water alternating CO_2 EOR process. J Taiwan Inst Chem Eng 2019;99:154–65.
[17] Eshraghi SE, Rasaei MR, Zendehboudi S. Optimization of miscible CO_2 EOR and storage using heuristic methods combined with capacitance/resistance and Gentil fractional flow models. J Nat Gas Sci Eng 2016;32:304–18.
[18] Grude S, Landrø M, Dvorkin J. Pressure effects caused by CO2 injection in the Tubàen Fm., the Snøhvit field. Int J Greenh Gas Control 2014;27:178–87.
[19] Liu B, Shi J, Wang M, Zhang J, Sun B, Shen Y, et al. Reduction in interfacial tension of water–oil interface by supercritical CO2 in enhanced oil recovery processes studied with molecular dynamics simulation. J Supercrit Fluids 2016;111:171–8.
[20] Nait Amar M, Zeraibi N, Jahanbani Ghahfarokhi A. Applying hybrid support vector regression and genetic algorithm to water alternating CO_2 gas EOR. Greenh Gases Sci Technol 2020;10:613–30. doi: 10.1002/ghg.1982.
[21] Rozzi NL, Singh RK. Supercritical fluids and the food industry. Compr Rev Food Sci Food Saf 2002;1:33–44.
[22] Lang Q, Wai CM. Supercritical fluid extraction in herbal and natural product studies—A practical review. Talanta 2001;53:771–82.
[23] Srinivasan K. Identification of optimum inter-stage pressure for two-stage transcritical carbon dioxide refrigeration cycles. J Supercrit Fluids 2011;58:26–30.
[24] Hendriks E, Kontogeorgis GM, Dohrn R, de Hemptinne J-C, Economou IG, Zilnik LF, et al. Industrial requirements for thermodynamics and transport properties. Ind Eng Chem Res 2010;49:11131–41.
[25] Wilkes JS. Properties of ionic liquid solvents for catalysis. J Mol Catal A Chem 2004;214:11–7.
[26] Jiang X. A review of physical modelling and numerical simulation of long-term geological storage of CO_2. Appl Energy 2011;88:3557–66.
[27] Li H, Jakobsen JP, Ø Wilhelmsen, Yan J. PVTxy properties of CO_2 mixtures relevant for CO_2 capture, transport and storage: review of available experimental data and theoretical models. Appl Energy 2011;88:3567–79.
[28] Li H, Wilhelmsen Ø, Lv Y, Wang W, Yan J. Viscosities, thermal conductivities and diffusion coefficients of CO_2 mixtures: review of experimental data and theoretical models. Int J Greenh Gas Control 2011;5:1119–39.
[29] Hemmat Esfe M, Tatar A, Ahangar MRH, Rostamian H. A comparison of performance of several artificial intelligence methods for predicting the dynamic viscosity of TiO2/SAE 50 nano-lubricant. Phys E Low-Dimensional Syst Nanostruct 2018;96:85–93. doi: 10.1016/j.physe.2017.08.019.
[30] Sarafraz MM, Arjomandi M. Demonstration of plausible application of gallium nano-suspension in microchannel solar thermal receiver: experimental assessment of thermo-hydraulic performance of microchannel. Int Commun Heat Mass Transf 2018;94:39–46. doi: 10.1016/j.icheatmasstransfer.2018.03.013.
[31] Sarafraz MM, Pourmehran O, Yang B, Arjomandi M. Assessment of the thermal performance of a thermosyphon heat pipe using zirconia-acetone nanofluids. Renew Energy 2019;136:884–95. doi: 10.1016/j.renene.2019.01.035.
[32] Pourmehran O, Sarafraz MM, Rahimi-Gorji M, Ganji DD. Rheological behaviour of various metal-based nano-fluids between rotating discs: a new insight. J Taiwan Inst Chem Eng 2018;88:37–48. doi: 10.1016/j.jtice.2018.04.004.
[33] Sarafraz MM, Arya H, Saeedi M, Ahmadi D. Flow boiling heat transfer to MgO-therminol 66 heat transfer fluid: experimental assessment and correlation development. Appl Therm Eng 2018;138:552–62. doi: 10.1016/j.applthermaleng.2018.04.075.
[34] Sarafraz MM, Arya H, Arjomandi M. Thermal and hydraulic analysis of a rectangular microchannel with gallium-copper oxide nano-suspension. J Mol Liq 2018;263:382–9. doi: 10.1016/j.molliq.2018.05.026.
[35] Hemmati-Sarapardeh A, Varamesh A, Husein MM, Karan K. On the evaluation of the viscosity of nanofluid systems: modeling and data assessment. Renew Sustain Energy Rev 2018;81:313–29.
[36] Hemmat Esfe M, Rostamian H, Shabani-samghabadi A, Abbasian Arani AA. Application of three-level general factorial design approach for thermal conductivity of MgO/water nanofluids. Appl Therm Eng 2017;127:1194–9. doi: 10.1016/j.applthermaleng.2017.07.211.
[37] Sarafraz MM, Arjomandi M. Thermal performance analysis of a microchannel heat sink cooling with copper oxide-indium (CuO/In) nano-suspensions at high-temperatures. Appl Therm Eng 2018;137:700–9. doi: 10.1016/j.applthermaleng.2018.04.024.

[38] Sarafraz MM, Pourmehran O, Yang B, Arjomandi M, Ellahi R. Pool boiling heat transfer characteristics of iron oxide nano-suspension under constant magnetic field. Int J Therm Sci 2020;147:106131. doi: 10.1016/j.ijthermalsci.2019.106131.

[39] Sarafraz MM, Safaei MR. Diurnal thermal evaluation of an evacuated tube solar collector (ETSC) charged with graphene nanoplatelets-methanol nano-suspension. Renew Energy 2019;142:364–72. doi: 10.1016/j.renene.2019.04.091.

[40] Sarafraz MM, Hart J, Shrestha E, Arya H, Arjomandi M. Experimental thermal energy assessment of a liquid metal eutectic in a microchannel heat exchanger equipped with a (10 Hz/50 Hz) resonator. Appl Therm Eng 2019;148:578–90. doi: 10.1016/j.applthermaleng.2018.11.073.

[41] Ahmadi MH, Nazari MA, Ghasempour R, Madah H, Shafii MB, Ahmadi MA. Thermal conductivity ratio prediction of Al2O3/water nanofluid by applying connectionist methods. Colloids Surfaces A Physicochem Eng Asp 2018;541:154–64.

[42] Ahmadi MH, Mirlohi A, Nazari MA, Ghasempour R. A review of thermal conductivity of various nanofluids. J Mol Liq 2018;265:181–8.

[43] Bahadori A, Vuthaluru HB. Predictive tool for an accurate estimation of carbon dioxide transport properties. Int J Greenh Gas Control 2010;4:532–6.

[44] Liu Z-B, He Y-L, Li Y-S, Qu Z-G, Tao W-Q. Heat transfer characteristics of supercritical CO_2 flow in metal foam tubes. J Supercrit Fluids 2015;101:36–47.

[45] Huber ML, Sykioti EA, Assael MJ, Perkins RA. Reference Correlation of the Thermal Conductivity of Carbon Dioxide from the Triple Point to 1100K and up to 200MPa. J Phys Chem Ref Data 2016;45:13102.

[46] Johns AI, Rashid S, Watson JTR, Clifford AA. Thermal conductivity of argon, nitrogen and carbon dioxide at elevated temperatures and pressures. J Chem Soc Faraday Trans 1 Phys Chem Condens Phases 1986;82:2235–46.

[47] Jarrahian A, Heidaryan E. A novel correlation approach to estimate thermal conductivity of pure carbon dioxide in the supercritical region. J Supercrit Fluids 2012;64:39–45.

[48] Amooey AA. A simple correlation to predict thermal conductivity of supercritical carbon dioxide. J Supercrit Fluids 2014;86:1–3.

[49] Rostami A, Arabloo M, Ebadi H. Genetic programming (GP) approach for prediction of supercritical CO2 thermal conductivity. Chem Eng Res Des 2017;122:164–75.

[50] Rostamian H, Lotfollahi MN. A new correlation method for estimating thermal conductivity of carbon dioxide in liquid, vapor and supercritical phases. Period Polytech Chem Eng 2020;64:146–52.

[51] Haji-Savameri M., Nait Amar M., Norouzi-Apourvari S., Hemmati-Sarapardeh A. Modeling dew point pressure of gas condensate reservoirs: comparison of hybrid soft computing approaches, correlations, and thermodynamic models. J Pet Sci Eng2019:106558.

[52] Hemmati-Sarapardeh A, Nait Amar M, Soltanian MR, Dai Z, Zhang X. Modeling CO 2 solubility in water at high pressure and temperature conditions. Energy Fuels 2020;34:4761–76. doi: 10.1021/acs.energyfuels.0c00114.

[53] 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.

[54] Hemmati-Sarapardeh A, Ghazanfari M-H, Ayatollahi S, Masihi M. Accurate determination of the CO2-crude oil minimum miscibility pressure of pure and impure CO2 streams: a robust modelling approach. Can J Chem Eng 2016;94:253–61.

[55] Karkevandi-Talkhooncheh A, Hajirezaie S, Hemmati-Sarapardeh A, Husein MM, Karan K, Sharifi M. Application of adaptive neuro fuzzy interface system optimized with evolutionary algorithms for modeling CO2-crude oil minimum miscibility pressure. Fuel 2017;205:34–45.

[56] Nait Amar M, Zeraibi N. Application of hybrid support vector regression artificial bee colony for prediction of MMP in CO_2-EOR process. Petroleum 2018. doi: 10.1016/j.petlm.2018.08.001.

[57] Nait Amar M, Ghriga MA, Ouaer H, El Amine Ben Seghier M, Pham BT, Andersen PØ. Modeling viscosity of CO_2 at high temperature and pressure conditions. J Nat Gas Sci Eng 2020:77. doi: 10.1016/j.jngse.2020.103271.

[58] Baghban A, Ahmadi MA, Shahraki BH. Prediction carbon dioxide solubility in presence of various ionic liquids using computational intelligence approaches. J Supercrit Fluids 2015;98:50–64.

[59] Ahmadi MA, Zendehboudi S, James LA. Developing a robust proxy model of CO 2 injection: coupling Box–Behnken design and a connectionist method. Fuel 2018;215:904–14.

[60] Shams R, Esmaili S, Rashid S, Suleymani M. An intelligent modeling approach for prediction of thermal conductivity of CO2. J Nat Gas Sci Eng 2015;27:138–50.

[61] Tatar A, Barati-Harooni A, Najafi-Marghmaleki A, Norouzi-Farimani B, Mohammadi AH. Predictive model based on ANFIS for estimation of thermal conductivity of carbon dioxide. J Mol Liq 2016;224:1266–74.

[62] Ahmadi MA, Baghban A. Evolving simple-to-apply models for estimating thermal conductivity of supercritical CO2. Int J Ambient Energy 2017;38:300–7.

[63] Rostamian H, Lotfollahi MN. A novel statistical approach for prediction of thermal conductivity of CO2 by Response Surface Methodology. Phys A Stat Mech Its Appl 2019;527:121175.

[64] Vesovic V, Wakeham WA, Olchowy GA, Sengers JV, Watson JTR, Millat J. The transport properties of carbon dioxide. J Phys Chem Ref Data 1990;19:763–808.

[65] Haykin S. 3rd Edition Neural networks and learning machines, 40. NJ, USA: Pearson Upper Saddle River; 2001.

[66] Tatar A, Shokrollahi A, Mesbah M, Rashid S, Arabloo M, Bahadori A. Implementing radial basis function networks for modeling CO2-reservoir oil minimum miscibility pressure. J Nat Gas Sci Eng 2013;15:82–92.

[67] Haykin S. Neural networks a comprehensive foundation. Englewood Cliffs: Pretice Hall International. Inc; 1999.

[68] Benamara C, Gharbi K, Nait Amar M, Hamada B. Prediction of wax appearance temperature using artificial intelligent techniques. Arab J Sci Eng 2020;45:1319–30.

[69] Halali MA, Azari V, Arabloo M, Mohammadi AH, Bahadori A. Application of a radial basis function neural network to estimate pressure gradient in water–oil pipelines. J Taiwan Inst Chem Eng 2016;58:189–202.

[70] Tatar A, Naseri S, Bahadori M, Hezave AZ, Kashiwao T, Bahadori A, et al. Prediction of carbon dioxide solubility in ionic liquids using MLP and radial basis function (RBF) neural networks. J Taiwan Inst Chem Eng 2016;60:151–64.

[71] Kennedy J, Eberhart R. Particle swarm optimization. Proc. ICNN'95-International Conf. Neural Networks 1995;4:1942–8.

[72] Shi Y. Particle swarm optimization: developments, applications and resources. In: Proceedings of the 2001 Congress on Evolutionary Computation (IEEE Cat. No. 01TH8546), 1; 2001. p. 81–6.

[73] Nilsson N.J. Learning machines. 1965.

[74] Rostami A, Baghban A, Mohammadi AH, Hemmati-Sarapardeh A, Habibzadeh S. Rigorous prognostication of permeability of heterogeneous carbonate oil reservoirs: smart modeling and correlation development. Fuel 2019;236:110–23.

[75] Shokrollahi A, Tatar A, Safari H. On accurate determination of PVT properties in crude oil systems: committee machine intelligent system modeling approach. J Taiwan Inst Chem Eng 2015;55:17–26.

[76] Benamara C, Nait Amar M, Gharbi K, Hamada B. Modeling wax disappearance temperature using advanced intelligent frameworks. Energy Fuels 2019;33:10959–68. doi: 10.1021/acs.energyfuels.9b03296.

[77] Ivakhnenko A.G.Polynomial theory of complex systems. IEEE Trans Syst Man Cybern1971:364–78.

[78] Dargahi-Zarandi A, Hemmati-Sarapardeh A, Hajirezaie S, Dabir B, Atashrouz S. Modeling gas/vapor viscosity of hydrocarbon fluids using a hybrid GMDH-type neural network system. J Mol Liq 2017;236:162–71.

[79] Nait Amar M, Jahanbani Ghahfarokhi A. Prediction of CO2 diffusivity in brine using white-box machine learning. J Pet Sci Eng 2020;190:107037. doi: 10.1016/j.petrol.2020.107037.

[80] Hemmati-Sarapardeh A, Mohagheghian E. Modeling interfacial tension and minimum miscibility pressure in paraffin-nitrogen systems: application to gas injection processes. Fuel 2017;205:80–9.

[81] Clifford AA, Kestin J, Wakeham WA. Thermal conductivity of N2, CH4 and CO2 at room temperature and at pressures up to 35MPa. Phys A Stat Mech Its Appl 1979;97:287–95.

[82] Perkins Richard A. Thermal conductivity measurements of carbon dioxide at temperatures from the triple point to 756K with pressures up to 68MPa. NIST 2019 Unpubl.

[83] Leneindre B, Tufeu R, Bury P, Sengers JV. Thermal conductivity of carbon dioxide and steam in the supercritical region. Berichte Der Bunsengesellschaft Für Phys Chemie 1973;77:262–75.

[84] Millat J, Mustafa M, Ross M, Wakeham WA, Zalaf M. The thermal conductivity of argon, carbon dioxide and nitrous oxide. Phys A Stat Mech Its Appl 1987;145:461–97.

[85] Le Neindre B. Contribution à l'étude expérimentale de la conductivité thermique de quelques fluides à haute température et à haute pression. Int J Heat Mass Transf 1972;15:1–24.

[86] Scott AC, Johns AI, Watson JTR, Clifford AA. Thermal conductivity of carbon dioxide in the temperature range 300–348K and pressures up to 25MPa. J Chem Soc Faraday Trans 1 Phys Chem Condens Phases 1983;79:733–40.

[87] Pátek J, Klomfar J, Čapla L, Buryan P. Thermal conductivity of carbon dioxide–methane mixtures at temperatures between 300 and 425K and at pressures up to 12MPa. Int J Thermophys 2005;26:577–92.

[88] Becker H, Grigull U. Messung der Temperatur-und der Wärmeleitfähigkeit von kohlendioxid im kritischen gebiet mittels holographischer interferometrie. Wärme-Und Stoffübertragung 1978;11:9–28.

[89] Michels A, Sengers JV, der Gulik PS. The thermal conductivity of carbon dioxide in the critical region: II. Measurements Concl. Phys 1962;28:1216–37.

[90] Yan Y, Xu L, Lee P. Mass flow measurement of fine particles in a pneumatic suspension using electrostatic sensing and neural network techniques. IEEE Trans Instrum Meas 2006;55:2330–4.

[91] Aminu KT, McGlinchey D, Chen Y. Optimal design for real-time quantitative monitoring of sand in gas flowline using computational intelligence assisted design framework. J Pet Sci Eng 2019;177:1059–71.

[92] Mirjalili S. How effective is the Grey Wolf optimizer in training multi-layer perceptrons. Appl Intell 2015;43:150–61.

[93] Netzeva TI, Worth AP, Aldenberg T, Benigni R, Cronin MTD, Gramatica P, et al. Current status of methods for defining the applicability domain of (quantitative) structure-activity relationships: the report and recommendations of ecvam workshop 52. Altern to Lab Anim 2005;33:155–73.

[94] Gramatica P. Principles of QSAR models validation: internal and external. Mol Inform 2007;26:694–701.

[95] Leroy AM, Rousseeuw PJ. Robust regression and outlier detection. New York: J Wiley&Sons; 1987.

[96] Tatar A, Moghtadaei GM, Manafi A, Cachadiña I, Mulero Á. Determination of pure alcohols surface tension using Artificial Intelligence methods. Chemom Intell Lab Syst 2020;201. doi: 10.1016/j.chemolab.2020.104008.