Petroleum 10 (2024) 109–123

Contents lists available at ScienceDirect

Petroleum

journal homepage: www.keaipublishing.com/en/journals/petlm

Full Length Article

Intelligent solubility estimation of gaseous hydrocarbons in ionic liquids

Behnaz Basirat [a], Fariborz Shaahmadi [b], Seyed Sorosh Mirfasihi [c], Abolfazl Jomekian [d], Bahamin Bazooyar [e, *]

[a] Ahvaz Faculty of Petroleum Engineering, Petroleum University of Technology (PUT), Ahvaz, 6198144471, Iran
[b] School of Engineering & Physical Sciences, Heriot Watt University, Edinburgh EH14 4AS, UK
[c] Department of Fluids & Environment, The University of Manchester, Manchester, M13 9PL, UK
[d] Department of Engineering, Chemical Engineering Group, Esfarayen University of Technology, Iran
[e] Mechanical and Aerospace Engineering, Brunel University London, Uxbridge UB8 3PH, UK

ARTICLE INFO

Article history:
Received 23 May 2023
Received in revised form
4 September 2023
Accepted 7 September 2023

Keywords:
Solubility
Gaseous hydrocarbon
Intelligent models
Ionic liquids

ABSTRACT

The research focuses on evaluating how well new solvents attract light hydrocarbons, such as propane, methane, and ethane, in natural gas sweetening units. It is important to accurately determine the solubility of hydrocarbons in these solvents to effectively manage the sweetening process. To address this challenge, the study proposes using advanced empirical models based on artificial intelligence techniques like Multi-Layer Artificial Neural Network (ML-ANN), Support Vector Machines (SVM), and Least Square Support Vector Machine (LSSVM). The parameters for the SVM and LSSVM models are estimated using optimization methods like Genetic Algorithm (GA), Particle Swarm Optimization (PSO), and Shuffled Complex Evolution (SCE). Data on the solubility of propane, methane, and ethane in various ionic liquids are collected from reliable literature sources to create a comprehensive database. The proposed artificial intelligence models show great accuracy in predicting hydrocarbon solubility in ionic liquids. Among these, the hybrid SVM models perform exceptionally well, with the PSO-SVM hybrid model being particularly efficient computationally. To ensure a comprehensive analysis, different examples of hydrocarbons and their order are included. Additionally, a comparative analysis is conducted to compare the AI models with the thermodynamic COSMO-RS model for solubility analysis. The results demonstrate the superiority of the AI models, as they outperform traditional thermodynamic models across a wide range of data. In conclusion, this study introduces advanced artificial intelligence algorithms such as ML-ANN, SVM, and LSSVM in accurately estimating the solubility of hydrocarbons in ionic liquids. The incorporation of optimization techniques and variations in hydrocarbon examples improves the accuracy, precision, and reliability of these intelligent models. These findings highlight the significant potential of AI-based approaches in solubility analysis and emphasize their superiority over traditional thermodynamic models.

1. Introduction

One of the primary issues associated with the usage and processing of neat gaseous hydrocarbons is the considerable amount of contaminants and impurities that accompany these substances during oil extraction [1–3]. Natural gas as one of the main extracted oil products usually comprises carbon dioxide (CO_2) and hydrogen sulfide (H_2S), namely acid gases [4,5]. The existence of these accompanying agents is problematic because they reduce the heating value of the hydrocarbons and cause corrosion in

* Corresponding author.
E-mail address: B.Bazooyar@Brunel.ac.uk (B. Bazooyar).
Peer review under responsibility of Southwest Petroleum University.

Production and Hosting by Elsevier on behalf of KeAi

https://doi.org/10.1016/j.petlm.2023.09.002

equipment and transmission pipelines. It is therefore necessary to separate the carbon dioxide (CO_2) and hydrogen sulfide (H_2S) from the natural gas [6–9]. Conventional methods of separation of acid gases from natural gas include amines or another physical absorbents. The use of these solvents is not only economically challenging but also presents many obstacles. Volatility, solvent degradation, regeneration requirements, high energy cost, and equipment corrosion are the main disadvantages of conventional solvents for the removal of acid gases [10].

Ionic liquids (ILs) are seen as alternatives to traditional solvents in various absorption applications due to their ability to overcome several challenges associated with conventional solvents. These unique solvents possess distinct characteristics, characterized by their non-flammability, chemical stability, and versatile molecular structures. Typically, ionic liquids comprise bulky asymmetrical organic molecular cations that could be combined with both organic and inorganic molecular anions. Their molecular structures can be tailored to meet specific requirements, making them customisable for desired applications [11]. They have extensive applications in petroleum industry such as clay swelling inhibitors [12,13] and clay stabilizer additives in fracturing fluids [14]. ILs are also potent materials for the separation of acid gases from natural gas as they selectively absorb carbon dioxide (CO_2) and hydrogen sulfide (H_2S) from natural gas [15,16]. Ionic liquids (ILs) have the potential to absorb hydrocarbon constituents, which can negatively impact the properties of processed gas. ILs may remove beneficial components of natural gases, such as ethane and propane, which are important for the quality and price of natural gas. Therefore, it is crucial to analyze the solubility of hydrocarbons in ILs. However, there are limited studies available to predict and model hydrocarbon solubilities in ILs. The intricate structure of these molecules presents a significant challenge when attempting to mathematically tackle the problem of solubility. Various mathematical models have been developed to analyze solubility in ionic liquids, including molecular dynamic simulation, activity models, Gibbs energy models like NRTL and UNQIUAC, and cubic equations of state (EoS). These models have been used to represent the solubility of acid gases [17–22] and hydrocarbons in ILs. However, these analysis methods are limited, as they can only predict the solubility data for a specific hydrocarbon in a specific IL category. Recently, the group contribution method of analysis has been expanded to provide a more generalized prediction of thermodynamics and solubility for a specific category of ILs [23]. By employing this approach, it is possible to accurately predict the phase behavior of any type of ionic liquid by fitting the properties associated with the functional group of the ILs. This method has the potential to estimate the properties of ILs when data is unavailable [24]. However, the conjugated EoS used in this analysis method is only applicable for pure component data and lacks the necessary functionality across the entire thermodynamic range. The science of thermodynamics in the presentation of the phase behavior of ionic liquids faces several challenges as these green solvents have an extremely complex structure [25]. The study of gaseous solubility in the liquid phase is a matter of thermodynamics as it is a content of vapor a liquid can hold without any macroscopic inter-phase movements of absorbent. Thermodynamics, however, takes numerious parameters into account making it a convoluted analysis and computationally challenging. The fact that molecular structures and properties of materials are considered in all thermodynamics models makes it a specialized mathematical tool for the study of solubility without any generalization. For simple property solubility, the mathematical objects go in-depth into the microscopic scales representing those intricate effects such as intermolecular forces between the vapor and liquid. This makes the mathematical approaches of thermodynamics workable only under narrow specific conditions. Artificial intelligence has begun to solve many engineering problems. They can successfully approach any empirically defined, convoluted, and chaotic problems the way the human mind gains experience and extrapolates any heretofore unknown information [26–28]. Artificial intelligence provides a subset of intelligent models for viscosity calculation [29] and oil recovery performance prediction and vapor-liquid equilibria and solubility analysis. ANN was used by Kalam et al. [30] to predict the oil recovery in heterogeneous oil reservoir from its 5-spot water flood. Their presented models fitted very well with real field data and it saves computational time in prediction of the waterflood performance compared with a reservoir simulator. Kalam et al. [31] in another work developed an artificial intelligence based empirical correlation for prediction of waterflooding performance. They concluded that their model can be a primary asset for a quick estimation of the waterflood oil recovery before a more comprehensive simulation model is built and ran. In another work by Kalam et al. [32] they successfully developed a Data-Driven Model to Predict the Recovery Performance of Low-Salinity Waterfloods. The solubility of carbon dioxide (CO_2), hydrogen sulfide (H_2S), and nitrous oxide (N_2O) has been successfully modeled using intelligent models [33–43]. Using the artificial neural network (ANN) in conjunction with particle swarm optimization (PSO) and backpropagation (BP) techniques, Shafie et al. [42] have modeled the solubility of H_2S in different 11 ionic liquids and reported that PSO-ANN optimization tool is more predictive than BP-ANN. In a report by Baghban et al. [42] the performance of MLP-ANN optimized by ANFIS was compared with the Soave-Redlick-Kwong (SRK) and Peng-Robinson (PR) EoSs. The findings of the study indicated that MLP-ANN exhibited superior precision. In Ahmadi et al. [44], the performance of the GA-LSSVM was compared with the SRK and PR equations of state in predicting the solubility of H_2S in 11 different ionic liquids. While these studies were successful in accurately predicting the solubility of contaminants such as carbon dioxide (CO_2) and hydrogen sulfide (H_2S) in ILs, they were unable to provide selective solubility information for hydrocarbons in ionic liquids. This paper provides an intelligent analysis and estimation for solubilities of pure hydrocarbons in ionic liquids in a way beyond thermodynamics. A wide range of hybrid intelligent models has been tested for 973 data of 16 ionic liquids. The ability and precision of these models are also compared to the COSMO-RS in terms of precision.

2. Data acquisition

The solubility of propane, methane, butane and ethane in 16 various ILs at different conditions was selected for modelling. Table 1 includes the IUPAC names and structures of these 16 ionic liquids. The solubility dataset was obtained from the works of other researchers [45–63]. The solubility is considered the output dataset of intelligent models. For input datasets, the phase thermochemical state parameters such as pressure (P), temperature (T), molecular weight (MW) of gas, and type of ionic liquids are considered given their expectedcrucial role in solubility. Table 2 details the complete characteristics of the solubility data set including the type of hydrocarbons, solubility range, data reference and the number of data points used in each reference, along with experiment conditions like pressure and temperature.

In cases where the properties of a certain ionic liquid cannot be found from the literature, the group contribution method [64,65] was utilized to predict these properties. Specifically, the critical pressure (Pc), temperature (Tc), and acentric factor (ω) are considered as input variables for intelligent models, enabling them to recognize and differentiate each dataset for a specific IL. This approach allows for comprehensive modelling and analysis of ILs, even when specific properties are not directly available. Combined with the molecular weight (Mw) of hydrocarbon gas and

Table 1
List of experimental data points used in this study.

Number of data points	Hydrocarbon	Ionic liquid index	Solubility	Reference
10	C_3H_8	12	0.096 _ 0.241	(Kim et al. 2007) [21]
31	CH_4	8	0.052_0.155	(Althuluth et al. 2012) [44]
25	C_2H_6, CH_4	1	0.0004513 _ 0.003883	(Jacquemin et al. 2006) [47]
22	C_2H_6, CH_4	4	0.001905_0.1138	(Jacquemin, Husson et al. 2006) [48]
24	CH_4	3	0.009123 _ 0.0460531	(Kumeıan, Pérez- Salado Kamps et al. 2007) [49]
24	CH_4	12	0.027517_0.185967	(Kumeıan, Pérez- Salado Kamps et al.2007) [49]
39	C_3H_8, C_4H_{10}	5	0.0284 _ 0.2946	(Lee and Outcalt, 2006) [50]
82	CH_4	5	0.0298 _ 0.2245	(Raeissi and Peters 2010) [51]
22	C_2H_6	6 _ 15	0.008 _ 0.0414	(Stevanovic and Gomes 2013) [52]
30	CH_4	16	0.0009 _ 0.0431	(Yuan, Zhang et al.2006) [53]
21	C_2H_6, CH_4	9	0.0013 _ 0.0405	(Bermejo, Fieback et al. 2013) [55]
159	C_2H_6	4 _ 5	0.000000335 _ 0.126	(Anthony, Anderson et al. 2005) [54]
28	C_2H_6	2 _ 11	0.00907 _ 0.0173	(Almantariotis, Stevanovic et al. 2012) [56]
155	C_2H_6, CH_4	12	0.000679 _ 0.132	(Anderson, Dixon et al. 2007) [57]
160	CH_4	4	0.00000154 _0.043	(Anthony, Maginn et al. 2002) [58]
10	C_2H_6	12	0.008542 _ 0.01802	(Gomes, 2007) [60]
108	C_2H_6	12	0.0501_0.4016	(Florusse, Raeissi et al. 2008) [61]

experimental conditions of solubility analysis, these properties form 6 input data set points for the intelligent models. Table 3 lists the input and output data with their maximum and minimum values.

To ensure a comprehensive overview of the data variability and error in the generated models, the solubility data, inputs, and outputs are scaled and normalized within the [0, 1] range. This normalization process allows for a more effective comparison and analysis of the data, enabling a clearer understanding of the variations and trends within the dataset. By scaling, the error of modelling for large solubility values becomes comparable to those for low solubility values. The formula below is implemented to scale the data into the range of zero to one:

$$x_n = \left(\frac{x - x_{\min}}{x_{\min} - x_{\max}}\right) \quad (1)$$

In this formula, x_n signifies the data in a normalized form. $x_{\min}$ and $x_{\max}$ denotes the minimum and maximum data values, respectively, and x signifies the actual data value. Normalization is crucial in standardizing the data range without compromising the integrity of the model's outcomes. Scaling the data to a common range enables a more effective comparison and analysis of the data. It is vital to note that after completing the analysis or modelling, the data values are transformed back to their original scale to maintain their original interpretation and units. This ensures that the final results are presented in a meaningful and comprehensible manner.

3. Modelling development

3.1. Optimization tools

3.1.1. Particle swarm optimization (PSO)

This optimization method mimics the movement of flocks of birds or swarms of insects in the air. This method is one of the most computationally efficient methods as it requires minimum memory and CPU speed. This optimization tool operates in two steps according to accidental population-based procedure. The initial step in this procedure is to allocate arbitrary velocities and locations to initiate a primary population. The subsequent step is to implement a statistical equation and regression analysis for each population.

The PSO optimization algorithm deals with the problems by labelling each individual as a particle. These particles are considered to move in a multidimensional space corresponding to the belief space. The particles have memory, making it possible for them to store a part of their previous state. Movements of each particle are contingent upon the initial conditions (i.e., velocity and location) and two random parameters: weights of individual and social effects. Individual influences prompt the particle to return to its own best last position, while social influences guide them toward the best prior position in their vicinity.

The optimization of the parameters is achieved through the second phase of calculation using the PSO method of analysis. The optimization process is converged by the attainment of the desired fitness rate by the best particle. In case that there is a discrepancy between the fitness rate of the particles and the predefined convergence criteria, the particles velocity and position are adjusted and updated accordingly. This refinement process occurs in two stages. During these stages, the global or particle best fitness value might be updated depending on the performance of the adjustable parameters. The goal is to continuously enhance fitness and convergence until the desired stopping criteria are met.

More description about the PSO algorithm is given in [66].

3.1.2. Genetic Algorithm (GA)

The GA, similar to other intelligent models has an evolutionary root trying to optimize parameters based on what is happening in nature. Evolution can be considered as one of the optimization processes in nature as adaptability and survival are its inseparable parts. Using Darwin's theory [67] and genetic theory of Mendel, evolutionary computing techniques were developed to create an optimization tool for the variety of arguments.

The Genetic Algorithm (GA) operates by exploring different routes within a specific zone in order to locate a promising region of interest. This exploration is achieved using crossover, mutation, and inversion operators, which help to maintain a balance between exploitation and exploration, allowing for the discovery of new areas. To ensure the efficient functioning of the GA, an initial population is formed by defining the problem at hand. Subsequently, an objective function is formulated to evaluate the fitness of each solution within the population. Based on their fitness, individuals are classified and ranked. The next generation is then formed through a combination of mutation and crossover operations, selecting the most suitable individuals. This process continues iteratively until the stoppage criteria are met, resulting in the identification of the optimal solution. Further details and a comprehensive description of the Genetic Algorithm (GA) can be found in reference [68].

3.1.3. Shuffled complex evolution (SCE)

The SCE algorithm introduces a novel approach to perform optimization by incorporating four key points: 1) a conjunction of

Table 2
Names and structure of ionic liquids used in this study.

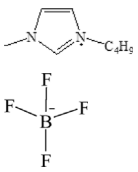
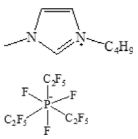
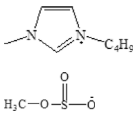
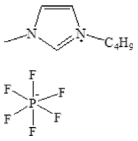
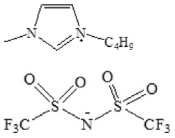
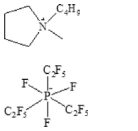
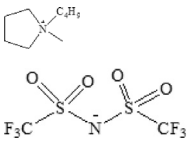
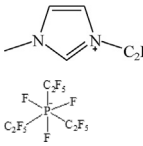
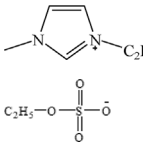
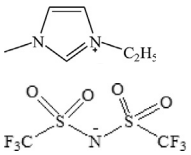
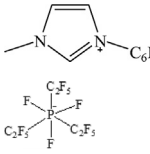
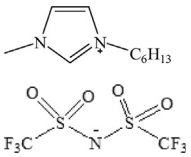
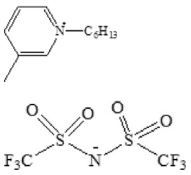
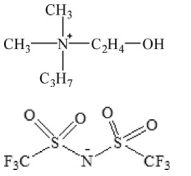
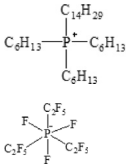
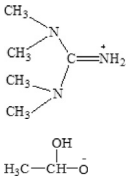
IUPAC name	Index number	Abbreviation	Structure
1-butyl-3-methyl-imidazolium tetrafluoroborate	IL-1	[BMIM][BF_4]	
1-butyl-3-methyl-imidazolium tris(nonafluoroethyl)trifluoro-phosphate	IL-2	[BMIM][eFAP]	
1-butyl-3-methyl-imidazolium methyl sulfate	IL-3	[BMIM][$MeSO_4$]	
1-butyl-3-methyl-imidazolium hexafluorophosphate	IL-4	[BMIM][PF6]	
1-butyl-3-methyl-imidazolium bis(trifluoromethylsulfonyl)-amide	IL-5	[BMIM][Tf2N]	
1-butyl-1-methyl-pyrrolidinium tris(nonafluoroethyl)trifluoro-phosphate	IL-6	[BMPYR][eFAP]	
1-butyl-1-methyl-pyrrolidinium bis(trifluoromethylsulfonyl)-amide	IL-7	[BMPYR][Tf2N]	
1-ethyl-3-methyl-imidazolium tris(nonafluoroethyl)trifluoro-phosphate	IL-8	[EMIM][eFAP]	
1-ethyl-3-methyl-imidazolium ethyl sulfate	IL-9	[EMIM][$EtSO_4$]	
1-ethyl-3-methyl-imidazolium bis(trifluoromethylsulfonyl)-amide	IL-10	[EMIM][Tf2N]	
1-hexyl-3-methyl-imidazolium tris(nonafluoroethyl)trifluoro-phosphate	IL-11	[HMIM][eFAP]	

Table 2 (*continued*)

IUPAC name	Index number	Abbreviation	Structure
1-hexyl-3-methyl-imidazolium bis(trifluoromethylsulfonyl)-amide	IL-12	[HMIM][Tf2N]	
1-hexyl-3-methylpyridinium bis(trifluoromethylsulfonyl)-amide	IL-13	[HMPY][Tf2N]	
Propylcholinium bis(trifluoromethylsulfonyl)-amide	IL-14	[N1,1,3,2-OH][Tf2N]	
trihexyl(tetradecyl)phosphonium tris(nonafluoroethyl)trifluoro-phosphate	IL-15	[P6,6,6,14][eFAP]	
1,1,3,3-tetramethyl-guanidium lactate	IL-16	[TMG][L]	

deterministic with probabilistic procedures, 2) clustering, 3) systematic evolution of spanning points, and 4) competitive evolution [69]. By integrating these elements, the SCE leverages the power of the CRS procedure, competitive evolution, and convoluted shuffling to enhance its effectiveness in finding optimal solutions. The CRS emphasizes the best point of a given sample steering the SCE strategy toward a local region. The SCE does not possess the weakness of CRS in terms of the appearance of repeated points, and the permanent termination of the optimization loop. This makes it possible to be efficient and potent in dealing with small space samples. In SCE, the global search within the sample space goes forward the way natural evolution takes place on earth, the information set with the samples is efficiently and thoroughly exploited. This makes the SCE one of the strongest and fastest optimization tools in the objective finding of the global optimum.

Table 3
Type and range of studied data points.

Type of data	Property	Minimum	Maximum
Inputs	Molecular weight of hydrocarbon gas	16.0423	58.1218
	Temperature, K	279.98	453.15
	Pressure, bar	0.0002	161.05
	T_c,K	643.2	1533.988
	P_c, bar	5.13319398	40.5
	ω	0.20456556	1.393344
Output	Solubility	0.000000335	0.4016

3.2. Predictive tools

3.2.1. Multi-layer perception (MLP)

The artificial intelligence models operate based on experience, changing behavior, and adaptability. Gaining experience through supervision of any unknown situation, these models, derived for analysis, enable it to enhance the execution of commands and favorably respond to any change in environment. Artificial intelligence (AI) possesses a remarkable capability to handle large volumes of information, making it one of its key strengths. Within the realm of AI, artificial neural networks (ANNs) play a pivotal role. ANNs comprise interconnected layers of neurons, each contributing to the localization and processing of information. Through the utilization of non-linear mappings, ANNs are able to effectively process data by establishing connections between inputs and outputs within their network framework [70]. To obtain the required accuracy of the network, precise data processing is of utmost importance [71]. One commonly employed type of ANN is the MLP-ANN. In this procedure, the input data points are mapped onto their corresponding data points. In MLP-ANN, the input data will propagate and connect to the output only in one direction. This method of analysis has a minimum of three numbers of layers: 1) input, 2) hidden, and 3) output. Each MLP layer comprises interconnected process units called neurons. The number of input and output data

points are equal to the number of neurons in the input and output layers. In the hidden layer(s), the neurons are typically numbered and counted through a trial-and-error process during network generation. Each input-to-output transmitted signal establishes connections between the input neurons, hidden neurons, and output neurons, allowing for the mapping of inputs to the desired output data points. The signal is processed and modified in the hidden layers and find a relevant connection between the output and input.

The MLP-ANN equipped back propagation can recognize, differentiate and classify linearly separable data, and convert nonlinear data points to a linear form [72]. The generated network is optimized by allocating objective weights for each interconnected node and bias. As a result, the neurons at the output layers approach the actual outputs [73]. The MLP-ANN implements the backpropagation technique, a supervised learning method for network training through synaptic weights and biases. The supervised learning will balance the synaptic weight from feedback it receives from network error [74]. The training procedure employs an iterative method for optimization. This training method adjusts the synaptic weights and biases iteratively, aiming to minimize the errors and optimize the network's performance over multiple iterations [75,76].

Fig. 1 represents the structure of the constructed multilayer perception-artificial neural network. The mean square error (MSE) of the MLP-ANN is calculated from the following:

$$MSE = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{m}[Y_j(i) - T_j(i)]^2 \tag{2}$$

where m and n are the node numbers at the output layer, and the training sample. $Y_j(i)$ represents the output data points. To mitigate the issue of data overfitting and the associated error in non-linear regression, the data setpoints are classified in two groups: training and validation. The neural network is trained by the training dataset, while the validation data is employed to assess the model's accuracy. The number of neurons in the hidden layer can be adjusted to strike a balance and minimize the occurrence of overfitting and the corresponding error.

3.2.2. Support vector machine (SVM)

Vapnik [77] presented a multifunction analytical approach, namely support vector machine (SVM), to recognize a pattern among the data, classify them and provide a regression analysis [78]. The support vector machine (SVM) implements the statistical learning theory (SLT) to interrelate the data and perform structural risk minimization (SRM). Support vector machines were initially employed to recognize the data pattern and were then implemented for classification and data regression [79]. The input data are mapped to a high-dimensional feature space using support vector machines. A hyperplane is found via nonlinear input-output variable mapping [80,81]. From the discovered hyper plane, a linear relationship among the data layers (input-output) is found in a featured space. Therefore, the pre-nonlinear problem is now treated linear in the defined feature space. The accuracy and robustness of support vector machines is contingent upon a constructed function, namely kernel function, and coefficient and SVM internal dataset. The coefficient for the kernel function and SVM internal algorithm also impact the training data in the new space. Unlike other regression procedures, the quadric programming (QP) is employed to obtain the global results from the local optima. The procedure of finding the results in this algorithm is not fast as a set of nonlinear relationships are solved [82].

A train set forms as $\{x_i, y_i, i = 1, 2, ..., N | x_i \in R^n, y_i \in R\}$ where x_k shows ith input data, y_k indicates the output data, and N stands for the number of data.

For the training data, the regression function $f(x)$ is discovered:

$$f(x) = w^T\varphi(x) + b \tag{3}$$

In the regression equation, x demonstrates the input vector. The dimension of the input vector is $N \times n$ (N is the number of training data; n indicates the input data numbers). w^T gives the weight vector. $\varphi(x)$ symbolize the kernel function. b denotes the bias term.

The cost function (CF) is minimized to determine the w^T and bias term (b):

$$\text{Cost function (CF)} = \frac{1}{2}w^T + c\sum_{i=1}^{N}(\xi_i - \xi_i^*) \tag{4}$$

The conditions below are employed to calculate the constraints of real values:

$$\begin{cases} y_i - w^T\varphi(x_i) - b \le \varepsilon + \xi_i & i = 1, 2, ..., N \\ w^T\varphi(x_i) + b - y_i \le \varepsilon + \xi_i & i = 1, 2, ..., N \\ \xi_i, \xi_i^* \ge 0 & i = 1, 2, ..., N \end{cases} \tag{5}$$

In this equation, x_i and y_i are the input train and target samples. ε gives the precision obtained from approximation of function and ξ_i (or ξ_i^*) indicates the slack parameter. c denotes the parameter for tuning the SVM representing the discrepancy in ε from the targeted values. The following equation, Lagrange equation, is minimized the parameters of the cost function:

$$L(a, a^*) = -\frac{1}{2}\sum_{i,j=1}^{N}(a_i - a_i^*)\left(a_j - a_j^*\right)K(x_i, x_j) - \varepsilon\sum_{i=1}^{N}(a_i - a_i^*) + \sum_{i=1}^{N}y_i(a_i - a_i^*) \tag{6}$$

In this equation, a_k and a_k^* indicate the Lagrangian multipliers coefficient. SVM function is obtained from the following after the determination of the coefficient in Lagrange equation:

$$f(x) = \sum_{i=1}^{N}(a_i - a_i^*)K(x, x_i) + b \tag{7}$$

The quadratic programming method can be used here to calculate the function variables including a_k, a_k^*, and b. Hence, the SVM model's tuning parameters include ε, c and a. To tackle the and treat the regression function for the solubility of hydrocarbons, the Gaussian radial basis function (RBF) is employed. This method was based on its proven efficiency and speed in the model training phase.

$$K(u, v) = \exp\left(-r \times |u - v|^2\right) \tag{8}$$

In this framework, r stands for the kernel variable function. The values u and v are independent function parameters. RBF requires careful fine-tuning of the Gaussian width to ensure optimal performance. To optimize the SVM model's functionality and select appropriate parameters, comprehensive evaluation of c, r, and ε is essential. The c parameter strikes a delicate balance between

maximizing the margin and minimizing training error. Smaller c values downplay the significance of fitting the training data, while larger values overemphasize it. The ε value, often unknown and derived from inherent noise within the data, plays a critical role in constructing a robust regression function and determining the number of resulting support vectors. Higher ε values correspond to a reduced number of support vectors, leading to a simpler regression structure.

3.2.3. Least-squares support vector machine (LSSVM)

SVM algorithm, while robust, has been noted for its complexity, primarily due to its establishment of intricate, nonlinear input-output relationships via Quadratic Programming (QP). This complexity can lead to analysis that is both cumbersome and time intensive. Recognizing these issues, Suykens and Vandewalle [82] put forth a refined variant of SVM, namely least squares support vector machines (LSSVM). This alternative method is an intelligent approach, circumventing the complications inherent in the original SVM methodology. The LSSVM not only encompasses the precise advantages of SVM but also employs a simpler and more user-friendly algorithm for data analysis. The fundamental approach of LSSVM in establishing nonlinear relationships among data sets involves discovering a transformative nonlinear function that maps the training data from the input space to a multidimensional expansion. This methodology facilitates a more efficient and effective representation of the data, enabling enhanced modelling and analysis capabilities. Considering the training set $\{x_k, y_k\}, k = 1, 2, ..., N$ where $x_k \in R$ signifies the $t'th$ input data in the input space and $y_k \in R$ represents output value for a specific input parameter magnitude $\{i.e.\ x_k\}$ and N denotes the count of training samples, we can derive the subsequent function:

$$y = w^T \varphi(x) + b \quad \text{with } w \in R, \quad b \in R, \quad \varphi(0) \in R \to R^{n_h}, \quad n_n \to \infty \tag{9}$$

Here the weight vector is represented by w, and the bias term is denoted by b. n and n_n indicates the dimension of the input space and the passive characteristic space, respectively. In context of utilizing LSSVM within a data sequence, a new function is formulated to optimize the function and fine the model parameters.

$$\text{Cost function (CF)} = \frac{1}{2} w^T w + \frac{1}{2} \gamma \sum_{i=1}^{N} e_i^2 \tag{10}$$

In the above cost function, γ represents the regularization parameter forming an adjustable factor while e_k denotes the error arising during the LSSVM training phase. Eq. (10) is minimized through forming the following equality function:

$$y_k = w^T \varphi(x_k) + b + e_k \tag{11}$$

To solve Eq. (10) alongside the accompanying equality function delineated in Eq. (11), we make use of the Lagrangian equation, detailed in the following:

$$L(w, b, e, a) = \frac{1}{2} w^T w + \frac{1}{2} \gamma \sum_{k=1}^{N} e_k^2 - \sum_{k=1}^{N} a_k \left[w^T \varphi(x_k) + b + e_k - y_k \right] \tag{12}$$

In this equation, a_k signifies the Lagrange multipliers. The solution can be found by taking the derivatives of Eq. (12) and setting them to zero, thereby yielding the subsequent equations:

$$\begin{cases} \dfrac{\partial L}{\partial w} = 0 \quad \Rightarrow \quad w = \sum\limits_{i=1}^{N} a_i \varphi(x_i) \\ \dfrac{\partial L}{\partial b} = 0 \quad \Rightarrow \quad \sum\limits_{i=1}^{N} a_i = 0 \\ \dfrac{\partial L}{\partial e_i} = 0 \quad \Rightarrow \quad e_i = \gamma e_i \qquad\qquad i = 1, 2, ..., N \\ \dfrac{\partial L}{\partial a_i} = 0 \quad \Rightarrow \quad w^T \varphi(x_i) + b + e_i - y_i = 0 \quad i = 1, 2, ..., N \end{cases} \tag{13}$$

By satisfying Eq. (13), the parameters for LSSVM can be calculated. This system includes $2N + 2$ equations and an equivalent number of unknowns (a_i, e_i, w, and b). This equation set is reformulated into a matrix shape of the following equation:

$$\begin{bmatrix} 0 & 1_v^T \\ 1_v & \Omega + \gamma^{-1} I \end{bmatrix} \begin{bmatrix} b \\ a \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix} \tag{14}$$

In the above matrix relationship, the matrix y is $[y_1 ... y_N]^T (y = [y_1 ... y_N]^T)$, the matrix 1_N denotes the $[1 ... 1]^T$ $(1_N = [1 ... 1]^T)$, α is the $[\alpha_1 ... \alpha_N]^T$ $(\alpha = [\alpha_1 ... \alpha_N]^T)$, and symbol I denotes the unity matrix, Ω represents $\varphi(x_k)^T$ $(\Omega = \varphi(x_k)^T)$, where $\varphi(x_k)$ is the $K(x_K, x_l) \in k = 1, 2, ..., N$ $(\varphi(x_k) = K(x_K, x_l))$. $K(x_K, x_l)$ represents the kernel function which has a limit, the Mercer limit [83]. Another function, namely RBF, is exploited as follows:

$$K(x, x_k) = \exp\left(\frac{-\|x_k - x\|^2}{\sigma^2} \right) \tag{15}$$

Here, σ^2 stands for the squared bandwidth. Lastly, it is important to note that the LSSVM can be precisely affected by an adjustable parameter σ^2 and a regularization parameter (γ). During the training process, the first parameter is calculated using an external optimization technique (Fig. 2). The ultimate version of the model is represented from:

$$y(x) = \sum_{k=1}^{N} a_k K(x, x_k) + b \tag{16}$$

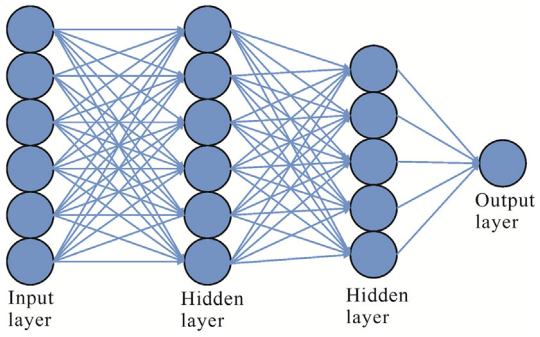


Fig. 1. Structure of studied MLP neural network.

4. Parameter estimation

4.1. Computational procedure

To construct the support vector machines, the dataset is systematically partitioned into two distinct segments: The dataset is

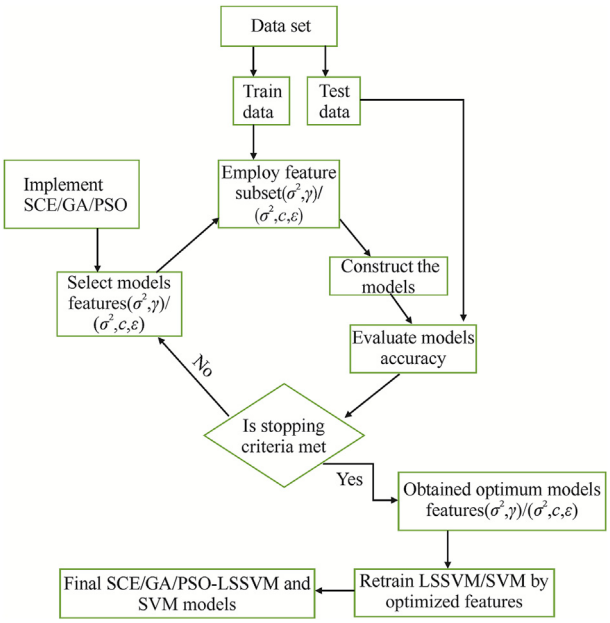


Fig. 2. Schematic presentation of optimization methods for SVM and LSSVM models.

categorized into two groups: the training set and the testing set. The training set is selected randomly and consists of 75% of the total data points. It is employed to create and refine the model's architecture. The remaining 25% subset, known as the testing set, is reserved for evaluating and verifying the precision and effectiveness of the constructed models. This division allows for comprehensive assessment and validation of the models' performance on unseen data.

A total of six input variables - temperature, pressure, critical pressure of IL, critical temperature of IL, acentric factor of IL, and molecular weight of the hydrocarbon gas constitute the model input parameters to investigate the solubility of gaseous hydrocarbon in the IL. This set of inputs forms the foundation for the development of the AI models. The architecture of the MLP-ANN model is then ascertained by examining a variety of networks. To determine the optimal parameters for both the original support vector machine (c, ε, and σ^2) and least square support vector machine (with γ and σ^2) we utilize the experimental training data, employing methodologies such as SCE, GA, and PSO Following this optimization, we proceed with the development of the support vector machines. The test data samples are then integrated into these models, generating a set of predicted data points. These predictions serve as a crucial measure for validating the reliability and effectiveness of the constructed models.

4.2. Accuracy of intelligent models

The analysis of model prediction and precision akin to experimentations, is necessary [84]. The model's precision and accuracy are analyzed using a combination of statistical and graphical accuracy analysis. Several metrics are employed to evaluate the errors, including the determination coefficient (R^2), root-mean-square-error (*RMSE*), average absolute percentage deviation (*AARD*), and standard deviation of error (*STD*). These metrics reflect the model's performance and accuracy. The formulas of metrics are as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^{N}\left[x_{pred}(i) - x_{Exp}(i)\right]^2}{\sum_{i=1}^{N}\left[x_{pred}(i) - \bar{x}_{Exp}(i)\right]^2} \tag{17}$$

$$RMSE = \left\{\frac{\sum_{i=1}^{N}\left[x_{pred}(i) - x_{Exp}(i)\right]^2}{N}\right\}^{0.5} \tag{18}$$

$$AARD = \frac{1}{N}\sum_{i=1}^{N}\frac{\left|x_{pred}(i) - x_{Exp}(i)\right|}{x_{Exp}(i)} \tag{19}$$

$$STD = \sum_{i=1}^{N}\left\{\frac{\left[x_{pred}(i) - \bar{x}_{Exp}(i)\right]^2}{N}\right\}^{0.5} \tag{20}$$

In these equations, $x_{\exp}$ is experimental measurements and x_{pred} correspond to the estimated hydrocarbon solubility, while $\bar{x}_{\exp}$ stands for the solubility-arrhythmic-averaged-values in IL.

To assess the reliability of the intelligent models used in this investigation, we used subplots and graphs to show the trend of error distribution. In the crossplot distribution curves, a graphical comparison is made between the experimentally observed values and those predicted through our models. The perfect model line, represented by a 45° line, helps us ascertain the consistency of our model. The proximity of data points to this line is indicative of the model's accuracy. Concurrently, to scrutinize the error trend, we observe the distribution of errors in relation to the zero-error line. This analysis aids us in understanding the model's precision and reliability.

Table 4
Values of statistical parameters for different ANN structures.

Number of hidden layer/neurons	Train				Test			
	R^2	RMSE	AARD(%)	STD	R^2	RMSE	AARD(%)	STD
1/2	0.9662	0.3706	20.98	1.798	0.9575	0.5875	60.17	3.1953
1/5	0.9761	0.3274	16.74	1.8199	0.9676	0.4952	47.25	3.2008
1/7	0.9924	0.222	10.95	1.8563	0.9857	0.2823	25.70	3.2286
2/5-3	0.9961	0.1976	9.589	1.8654	0.9888	0.2029	19.74	3.2339
2/6-3	0.9948	0.2809	11.61	1.9189	0.9786	0.2339	23.41	3.2586
2/6-5	0.9957	0.1864	8.93	1.8599	0.99	0.2123	18.88	3.2333
1/9	0.995	0.2109	9.62	1.8605	0.9872	0.2281	18.73	3.2316
1/11	0.9957	0.1904	9.03	1.8553	0.9895	0.2118	17.87	3.2336
1/8	0.9917	0.2011	12.26	1.8542	0.9882	0.2948	28.49	3.2265
1/6	0.984	0.2885	18.15	1.8426	0.9755	0.4066	44.58	3.2136
2/7-5	0.996	0.1909	8.99	1.8627	0.9895	0.2052	16.64	3.2336

Table 5
Values of statistical parameters for intelligent models.

Model	Optimization method	Model Parameters				Train				Test			
		γ	σ^2	ε	c	R^2	AARD(%)	RMSE	STD	R^2	AARD(%)	RMSE	STD
ANN	–	–	–	–	–	0.9957	20.854	0.0059	0.0901	0.9900	33.107	0.0090	0.0898
SVM	SCE	–	0.1601	0.00282	62296	0.9952	18.826	0.0063	0.0904	0.9930	19.751	0.0075	0.0902
	PSO	–	0.2014	0.00083	60096	0.9951	10.825	0.0063	0.0903	0.9914	8.324	0.0083	0.0899
	GA	–	0.1585	0.00274	62703	0.9952	17.860	0.0062	0.0904	0.9930	18.667	0.0075	0.0902
LSSVM	SCE	237721	0.7524	–	–	0.9962	18.168	0.0056	0.0900	0.9863	12.693	0.0106	0.0901
	PSO	237787	0.7525	–	–	0.9962	18.168	0.0056	0.0900	0.9863	12.693	0.0106	0.0901
	GA	289621	0.7743	–	–	0.9963	18.361	0.0055	0.0900	0.9862	12.998	0.0106	0.0902

5. Results and discussion

Through a systematic evaluation of various network configurations, the MLP-ANN model was developed. Table 4 presents the outcomes of the trial-and-error process, showcasing the accuracy of different model structures. Notably, the results indicate a neural network with two hidden layers with 6 and 5 neurons respectively, demonstrates optimal performance for the analysis of hydrocarbon solubility. This is evidenced by the lowest *RMSE* value obtained for the test dataset among all the tested network structures. Other error analysis tools such as R^2, *AARD*, and *STD* indicate this network structure to the most accurate MLP-ANN. The structure of the desired MLP-ANN comprising two hidden layers with 6 and 5 neurons is depicted in Fig. 1. This network connects a series of 6 input variables to an output which is the solubility of hydrocarbons.

Fig. 2 gives the schematic of the algorithms used for developing the support vector machines. Three optimization algorithms including SCE, PSO, and GA were utilized to tune the hyperparameters of support vector machines. The optimized values of c, ε and σ^2 for SVM and LSSVM in conjunction with SCE, PSO, and GA are given in Table 5. Our analysis confirms that the GA optimization tool for parameter estimation of SVM and LSSVM are more capable compared to PSO and SCE as the error analysis variables for test sets demonstrates. Hence, the visualized results are presented for the MLP-ANN, GA-SVM, and GA-LSSVM hybrid algorithms in estimation of hydrocarbon solubility in ILs.

The solubility data for the train and test datasets of MLP-ANN, GA-SVM, and GA-LSSVM models can be obtained from Fig. 3 a-c. These figures illustrate the relationship among the predicted solubility data ($x_{\exp}$) and the corresponding real solubility data for each model. It is evident that these intelligent models exhibit a high capability in estimating hydrocarbon solubility, as the predicted and experimental values align closely for both the train and test datasets. However, slight deviations from the ideal $y = x$ line can be observed at higher solubility ranges, indicating that the accuracy of the constructed models in estimating high values of hydrocarbon solubilities in ILs may diminish to some extent. The training set covers the range [0 0.4] of hydrocarbon solubility for

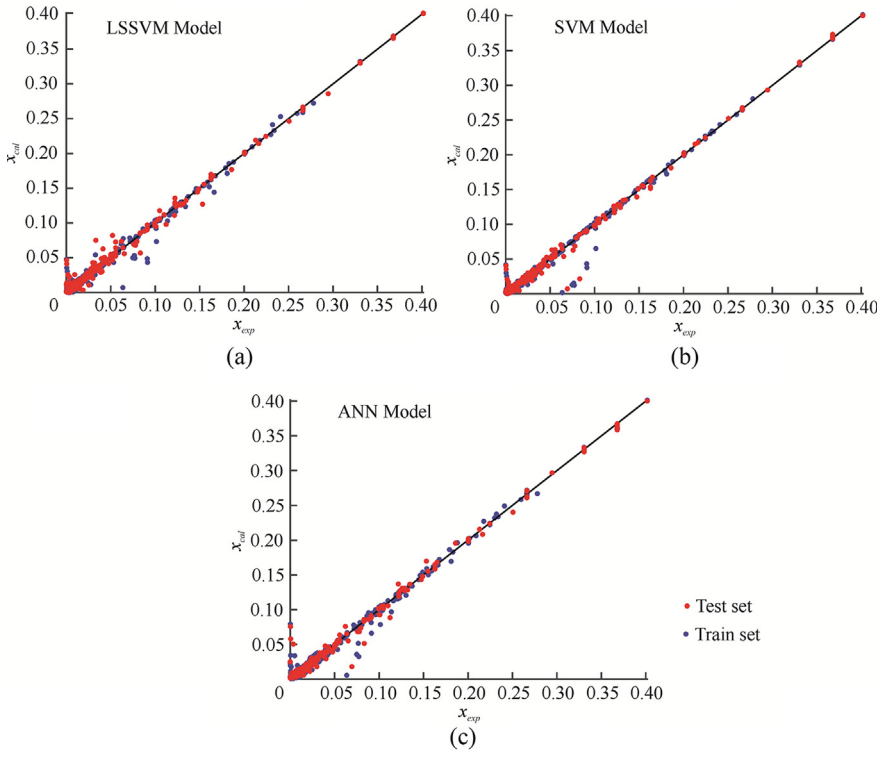


Fig. 3. Crossplot of predicted hydrocarbon solubility data points for intelligent models at both train (blue) and test (red) sets: (a) LSSVM-GA; (b) SVM-GA and (c) ANN.

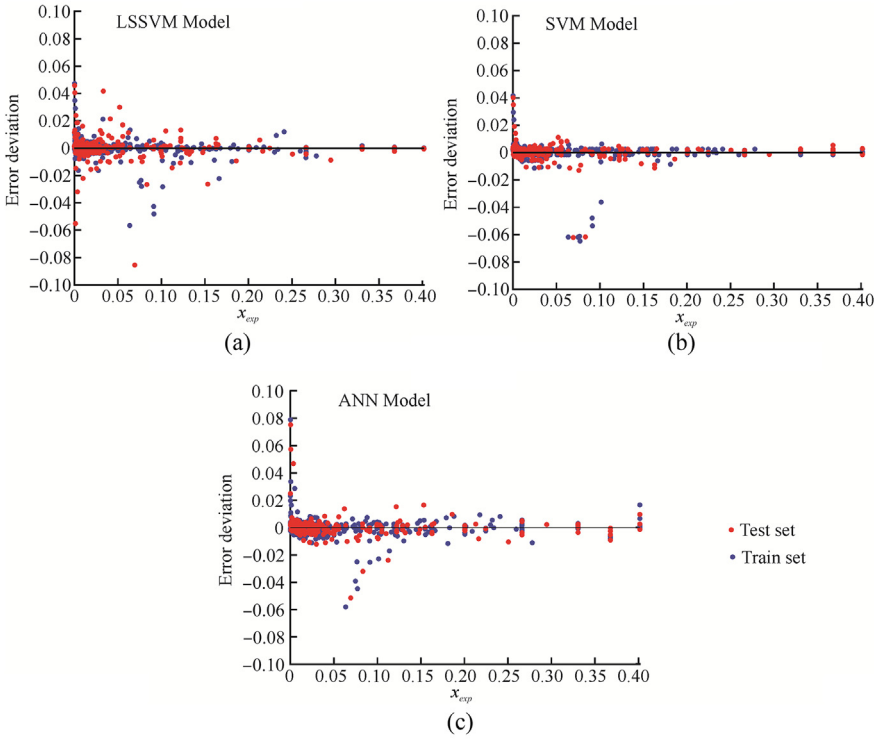


Fig. 4. The error distribution between experimental and predicted hydrocarbon solubility data point for intelligent models at both train, (blue) and test (red) sets: (a) LSSVM-GA; (b) SVM-GA and (c) ANN.

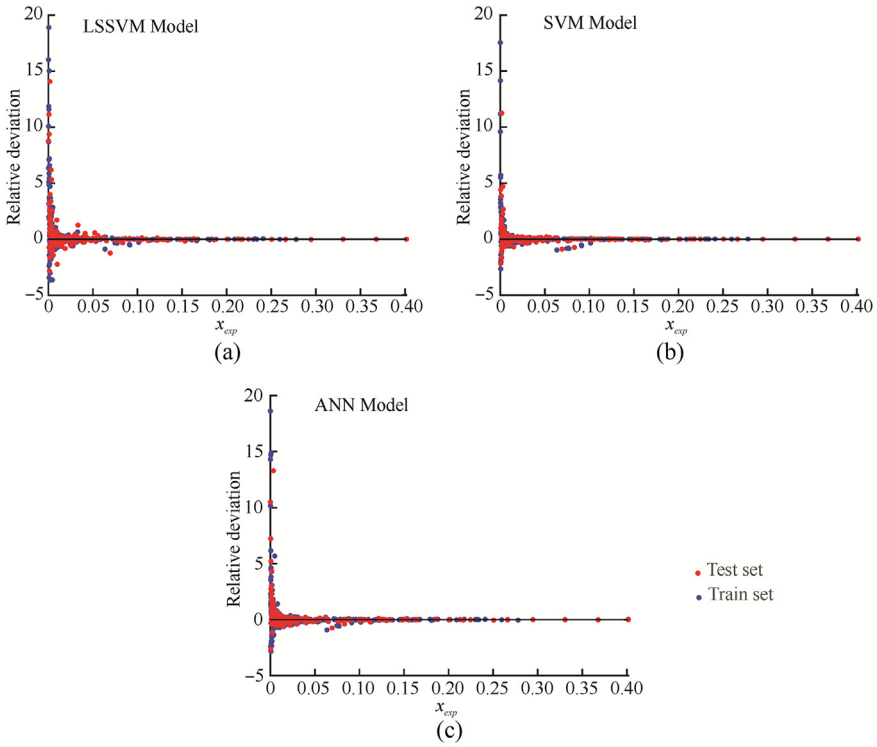


Fig. 5. The relative error distribution between experimental and predicted hydrocarbon solubility data point for intelligent models at both train (blue) and test (red), sets: (a) LSSVM-GA; (b) SVM-GA and (c) ANN.

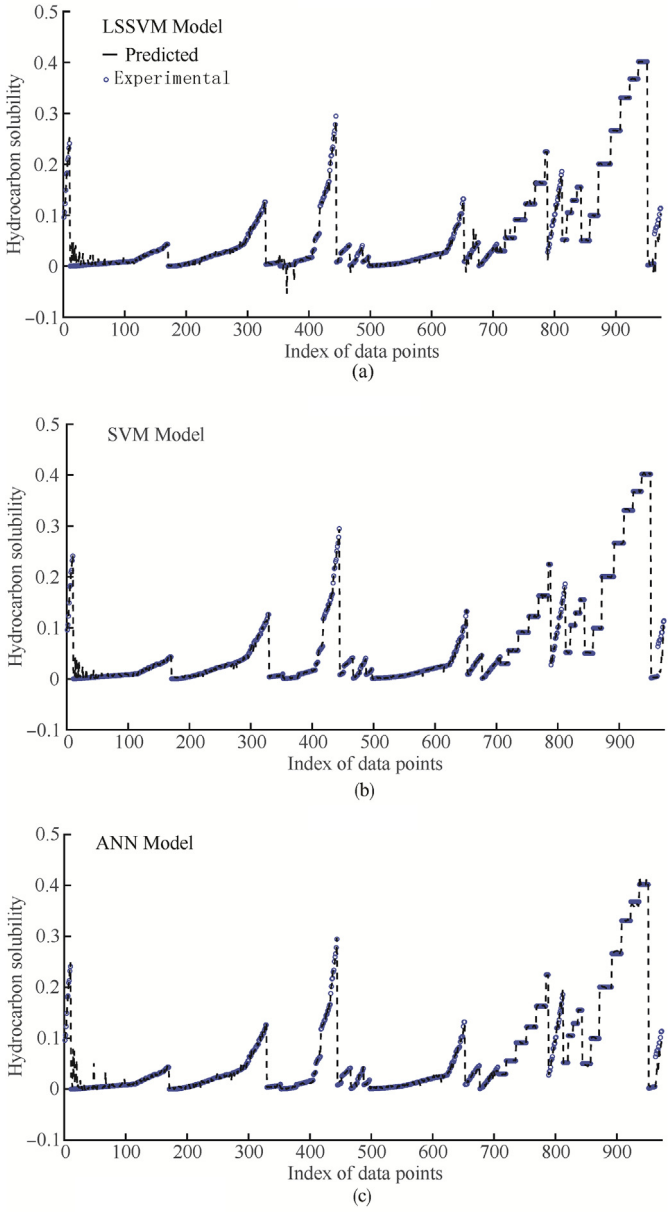


Fig. 6. Comparison between the experimental data and the results of intelligent models for all data points: (a) LSSVM-GA; (b) SVM-GA and (c) ANN.

ILs. However, it is limited in number at large values of solubility because the experimentation data points from the literature are rare for high solubilities. In this research, key differentiator lies in the intelligent models' ability to integrate the impacts of ILs and hydrocarbons in their forecasting processes. This implies that the developed models have a broad applicability, extending to 16 distinct ILs and 4 separate hydrocarbons, showcasing a significant strength of the modeling approach implemented here. The hydrocarbon solubilities in various ILs, demonstrated in Fig. 3 for both trained and tested data sets, further reinforce the reliability and precision of these intelligent models in solubility prediction. The data dispersed and distributed around the bisector in this illustration underscore the models' competence in delivering precise predictions.

Figs. 4 and 5 show the error analysis of the developed models to better highlight the discrepancy among the real and estimated solubilities. To this end, two variables for error analysis (namely error deviation and relative error deviation) were identified. Fig. 4 shows the error deviation over the entire dataset which is the ratio of predicted to the experimental value of solubility in percent. The absolute deviation in error for all investigated models approximately ranges between −0.085 and 0.08. While the error deviation for both training and testing data sets may appear minimal and inconsequential, these small values of deviation error cannot perfectly provide a compelling reason for the accuracy of intelligent models. A new relative deviation error was defined to better give the precision of the intelligent algorithm. The relative deviation error shows the error of models at each solubility in proportion to the it's absolute value thereby providing a better ground for solubility comparison. Fig. 5 gives the relative error deviation of estimated solubility data in relation to the experimental values. Despite the rather low values of error deviation over the entire range of solubility, the relative error does not possess the same behavior. The value of relative error deviation is higher for the extremely low amounts of solubility. This likely occurs because of low absolute amounts of solubility which makes the relative error deviation higher for low solubility values. Fig. 5 also authenticates that the error of modelling is rather low for the entire population of investigated solubility and that the developed algorithms can predicti the same thermodynamic science in an even a better way. Relatively higher precision in the intelligent analysis of solubilities is attributed to the judicious selection of input parameters, a diverse spectrum of data, and suitable identifiers for each of the ionic liquids incorporated in the model. Fig. 6 gives the predicted in relation to the experimental hydrocarbon solubility for all datasets. Evidently, the absolute value of predicted solubility within the entire investigated range is close to the experimental. However, this graph is only valid for ascertaining the discrepancy among the estimated and real values and no realistic comparison can be made with the preciseness of the models. In return, statistical model coefficients including R^2, *RMSE*, *STD*, and *AARD* can better show the value and effectiveness of intelligent algorithms for the solubility analysis over the entire range of solubility. Table 5 enlists the model coefficients of intelligent algorithms. Coefficient of determination

Table 6
Predictive accuracy of hydrocarbon solubility in AARD for some ionic liquids by intelligent models and COSMO-RS model.

Index number	Temperature (K)		Pressure (kPa)		Mole fraction of hydrocarbon		AARD				
	min	max	min	max	min	max	ANN	SVM	LSSVM	COSMO_RS (combi2005)	COSMO-RS (combi1998)
IL-1	302.05	363.45	610	24040	0.119	0.781	0.0171	0.0138	0.0116	0.4163	0.5323
IL-2	302.85	363.15	1020	30100	0.152	0.691	0.0167	0.0135	0.0094	0.6623	0.5927
IL-3	292.45	373.55	1140	13330	0.101	0.371	0.0886	0.0268	0.0245	0.1651	0.1669
IL-4	293.05	373.3	1030	18490	0.029	0.294	0.0682	0.0330	0.0366	1.4374	1.5952
IL-6	283.1	348.2	25.3	2000.6	0.002	0.49	0.2932	0.1466	0.1843	0.4223	0.4197
IL-10	303.15	343.23	100	100	0.017	0.0349	0.2655	0.0570	0.0298	0.6432	0.6506
IL-12	303.15	343.21	100	100	0.0242	0.0421	0.3740	0.0330	0.0256	0.5741	0.5853
IL-13	303.15	343.22	60.493	77.173	0.017	0.0278	0.2626	0.0451	0.0286	0.4691	0.4853

Table 7
Predictive accuracy of hydrocarbon solubility in AARD for some ionic liquids by intelligent models and COSMO-RS model.

Gas	IL index	Reference	AARD				
			GA-LSSVM	GA-SVM	MLP-ANN	COSMO-RS(combi2005)	COSMO-RS(combi1998)
CH_4	IL-1	(Jacquemin, Gomes et al. 2006) [47]	0.1229	0.0205	0.0149	0.7985	0.8099
	IL-3	(Kumeıan, Pérez-Salado Kamps et al. 2007) [49]	0.2520	0.1009	0.1057	0.7118	0.6974
	IL-4	(Jacquemin, Husson et al. 2006) [48]	0.4492	0.6008	0.4039	0.9858	0.9859
	IL-5	(Raeissi and Peters 2010) [51]	0.0280	0.0266	0.0136	0.9212	0.9233
	IL-8	(Althuluth, Kroon et al. 2012) [44]	0.0309	0.0239	0.0132	0.3989	0.3327
	IL-16	(Yuan, Zhang et al. 2006) [53]	0.2128	0.2486	0.4191	0.4103	0.3874
C_2H_6	IL-2	(Almantariotis, Stevanovic et al. 2012) [56]	0.1659	0.1305	0.0985	0.6469	0.5902
	IL-5	(Anthony, Anderson et al. 2005) [54]	0.7710	1.6806	1.1801	1.0772	1.1144
	IL-6	(Hong, Jacquemin et al. 2007) [46]	0.6318	0.2637	0.3746	0.5459	0.5410
	IL-10	(Hong, Jacquemin et al. 2007) [46]	0.3049	0.3473	0.3235	0.5465	0.5360
	IL-12	(Costa Gomes 2007) [60]	0.1686	0.1579	0.1104	0.2537	0.2477
	IL-15	(Stevanovic and Gomes 2013) [52]	0.0246	0.0247	0.0245	0.1544	0.2128
c_3H_8	IL-5	(Lee and Outcalt 2006) [50]	0.0221	0.0166	0.0289	0.4409	0.4205
	IL-12	(Kim, Jang et al. 2007) [21]	0.0220	0.0154	0.0213	0.1257	0.1084
C_4H_{10}	IL-5	(Lee and Outcalt 2006) [50]	0.1092	0.0535	0.0280	0.4303	0.4050
Total average			0.2211	0.2474	0.2107	0.5632	0.5542

(R^2) value for MLP-ANN in train and test sets is 0.9957, 0.9900, respectively. For SVM, the optimization tools SCE, PSO, and GA leads to respectively R^2 values 0.9930, 0.9914, and 0.9930 in test sets and in train sets the R^2 values 0.9952, 0.9951, and 0.9952 are obtained. The hybrid algorithms SCE-LSSVM, PSO-LSSVM, and GA-LSSVM have 0.9863, 0.9863, and 0.9862 for R^2 values in test sets and in train sets the coefficient of determination values are 0.9962, 0.9962, and 0.9963 respectively. Validating the results, it's apparent that the solubility of hydrocarbons is successfully estimated utilizing all of intelligent method of analysis. However, the original support vector machine shows a slightly stronger performance compared to the least square support vector machines and artificial neural network. This improved accuracy of SVM is likely due to its unique approach in dealing with model errors. Instead of solely minimizing training errors, SVM focuses on structural risk minimization, which aims to reduce the upper-bound error. As a result, SVM tends to provide better accuracy in predicting solubility compared to LSSVM and MLP-ANN.

The result of this analysis is also compared with the activity models in terms of *AARD* to demonstrate the effectiveness of intelligent models for estimation of hydrocarbon solubility in ILs. Table 6 and Table 7 include the absolute average relative deviation

Table 8
The comparison of the accuracy of the model presented with other similar computational intelligence-based works.

Gas	Method	AARD(%)	R^2	RSME	Max AD	Reference
CH_4	CSA-LSSVM	2.29	–	–	–	Amirkhani et al. [84]
	PSO-ANFIS	128.62	–	–	–	Amirkhani et al. [84]
	Hybrid-ANFIS	22.64	–	–	–	Amirkhani et al. [84]
	MLP-ANN	26.47	–	–	–	Amirkhani et al. [84]
	ANN	17.33	–	–	0.257	Safamirzaei et al. [89]
	PSO-ANFIS	14.12	0.9397	–	–	Dashti et al. [85]
	CSA-LSSVM	3.04	0.9846	–	–	Dashti et al. [85]
N_2O	DBN	–	0.9958	0.01186	0.0389	Nakhaei-Kohani et al. [88]
	Cat-Boost	–	0.9959	0.0116	0.0737	Nakhaei-Kohani et al. [88]
	ELM	–	0.9931	0.0152	0.1238	Nakhaei-Kohani et al. [88]
	XGB	–	0.9999	0.0016	0.0095	Nakhaei-Kohani et al. [88]
	Voting	–	–	0.01873	0.9405	Feng et al. [87]
	TLFFNN	–	–	0.00423	0.9984	Feng et al. [87]
	CFNNLM	–	0.9994	0.0047	–	Safamirzaei et al. [89]
CH_4, C_2H_6, C_2H_6,	Boosting-SVR	–	0.9907	0.0083	–	Nakhaei-Kohani et al. [88]
C_3H_6, C_3H_8, C_3H_{10}, $C_4H_{\backslash\backslash}$	DBN	–	0.9943	0.0065	–	Nakhaei-Kohani et al. [88]
	ELM	–	0.9938	0.0068	–	Nakhaei-Kohani et al. [88]
	MARS	–	0.9822	0.0115	–	Nakhaei-Kohani et al. [88]
C_2H_6	ANN	21.86	–	–	1.37	Safamirzaei et al. [89]
N_2O	ANN (average),	17.32	–	–	–	This work
	SVM (average),	4.61				
	LSSVM (average)	4.38				
CH_4	ANN (average),	16.17	–	–	–	This work
	SVM (average),	17.02				
	LSSVM (average)	18.26				
C_2H_6	ANN (average),	35.19	–	–	–	This work
	SVM (average),	43.41				
	LSSVM (average)	34.44				
C_3H_8	ANN (average),	25.1	–	–	–	This work
	SVM (average),	16				
	LSSVM (average)	22.1				
C_4H_{10}	ANN (average),	2.8	–	–	–	This work
	SVM (average),	5.35				
	LSSVM (average)	10.92				

(*AARD*) of AI models in comparison with the COSMO-RS (combi1998, combi2005) model 72. The accuracy of the proposed ANN, SVM and LSSVM in this work based on average of *AARD* of all investigated ILs considered for each gas, was presented in Table 8 in comparison with the state-of-the-art intelligence model-based works on gas solubility prediction in ILs. As can be seen in this table, Amirkhani et al. [85] utilized four different intelligence approaches such as multilayer Perceptron-Artificial Neural Networks (MLP-ANN), Hybrid-Adaptive Neuro Fuzzy Inference System (Hybrid-ANFIS), Particle Swarm Optimization-Adaptive Neuro Fuzzy Inference System (PSO-ANFIS) and Coupled Simulated Annealing-Least Squares Support Vector Machine (CSA-LSSVM) to predict CH_4 solubility in forty types of ILs scattered *ARDD%* values from very low to very high while Dashti et al. [86] achieved lower values of *ARDD%* for the similar case of CH_4 solubility prediction in ILs but with PSO-ANFIS and CSA-LSSVM models. In two separate studies on N_2O solubility prediction in ILs, Nakhaei-Kohani et al. [87] and Feng et al. [88] intelligence models such as deep belief network (DBN), categorical boosting algorithm (Cat-Boost), extreme learning machine (ELM), and extreme gradient boosting (XGB) along with two-layer feed forward neural network (TLFFNN) and cascaded forward neural network with Levenberg–Marquardt (CFNNLM) for prediction of N_2O solubility in different ILs. Models such as Boosting support vector regression (Boosting SVR), DBN, Multivariate adaptive regression splines (MARS) and Extreme learning machine (ELM) were used to predict the solubility of various light hydrocarbons in several ILs by Nakhaei-Kohani et al. [89]. Safarmirzaei et al. [90] used ANN to estimate the solubility of C_2H_6 in different ILs. As can be seen, in terms of *AARD* which can be separately compared for CH_4, the three models proposed in this work, show comparable and to some extent superior predictive performance compared with the works of other groups. The same also can be said about *RMSE* and R^2 when related values are compared in Table 5 and Table 8 with each other. Even superior accuracy can be seen for the case of LSSVM presented in this work compared with other works. Therefore, it can be concluded that the models presented here can be comparable and even superior to the state-of-the-art relevant works presented in literature.

These tables clearly demonstrate the better results with intelligent models compared to COSMO-RS model are obtained. Hence, the AI models appear capable of estimating hydrocarbon solubility while they offer generalization over the wide range of conditions and materials. However, the thermodynamic models are limited and should be specified for the investigated material and conditions.

6. Conclusions

This research aimed to estimate the solubility of four hydrocarbons in 16 different ionic liquids using intelligent algorithms such as ANN, SVM, and LSSVM. Optimization tools like SCE, GA, and PSO were used to improve the accuracy of the models by finding the optimal parameters. The input parameters for the models included thermodynamic conditions and properties of the ionic liquids, while the output was the solubility of the hydrocarbons. Different types of ANN algorithms were tested to determine the most effective structure for providing accurate solubility analysis in a timely manner. The intelligent models showed strong alignment between experimental and estimated data, indicating their reliability and effectiveness. The performance of the intelligent models was compared to the COSMO-RS model, and the results showed that the intelligent approaches performed better in estimating hydrocarbon solubility in ionic liquids. Among the optimization tools, GA was found to be more accurate for tuning SVM parameters, but it had higher computational costs compared to other tools. The RMSE for hybrid algorithms SCE-SVM, PSO-SVM, and GA-SVM was 0.0075, 0.0083, and 0.0075, respectively. For LSSVM, SCE and PSO were more accurate than GA for parameter estimations, with R^2 values of 0.9863, 0.9863, and 0.9862 for SCE-LSSVM, PSO-LSSVM, and GA-LSSVM, respectively. The intelligent models also outperformed activity models in terms of accuracy, requiring fewer optimization parameters for solubility estimation. These intelligent models provide a significant advantage in predicting hydrocarbon solubility in ionic liquids and can be valuable tools in designing efficient processes for hydrocarbon capture in these systems.

Data availability

Experimental, predicted, and input data used to build the intelligent framework models are accessible from Brunel University London repository at: https://doi.org/10.17633/rd.brunel.23937918.v1.

Credit author statement

Behnaz Basirat: Conceptualization, Data curation, Investigation, Methodology, Validation, Visualization, Writing – original draft, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Writing – review & editing. **Fariborz Shaahmadi**: Conceptualization, Data curation, Investigation, Methodology, Validation, Visualization, Writing – original draft, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Writing – review & editing. **Seyed Sorosh Mirfasihi**: Conceptualization, Data curation, Investigation, Methodology, Validation, Visualization, Writing – original draft, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Writing – review & editing. **Abolfazl Jomekian**: Conceptualization, Data curation, Investigation, Methodology, Validation, Visualization, Writing – original draft, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Writing – review & editing. **Bahamin Bazooyar**: Conceptualization, Data curation, Investigation, Methodology, Validation, Visualization, Writing – original draft, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Writing – review & 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.

Nomenclature

Abbreviations

AARD	Average absolute relative deviation
ANN	Artificial neural network
IL	Ionic liquid
LSSVM	Least square support vector machine
MLP	Multilayer perceptron
QP	Quadratic programming
RBF	Radial basis function
RMSE	Root mean square error
SCE	Shuffled complex evolution
STD	Standard deviation error
SVM	Support vector machine

Variables

$a_k(a_k^*)$	Lagrangian multiplier
c	Adjustable parameter of SVM model
e_k	Error of LSSVM in training phase
P_c	Critical pressure of ionic liquid
T_c	Critical temperature of ionic liquid
w^T	Weight vector
x_k	Input vector at the train sample k
x_n	Normalized data
y_k	Target vector at the train sample k

Greek symbols

γ	Regularization parameter
	Adjustable parameter of SVM model
ξ_k (ξ_k^*)	Slack variable
σ^2	Squared bandwidth
$\varphi(x)$	Kernel function
ω	Acentric factor of ionic liquid

Subscripts

Exp	Experimental
Max	Maximum value
Pred	Predicted

References

[1] A. Jomekian, S.A.A. Mansoori, B. Bazooyar, A. Moradian, Enhancement in thermal and hydrothermal stabilities of novel mesoporous MCM-41, J. Porous Mater. 19 (2012) 979–988.

[2] B. Bazooyar, S.Y. Hosseini, S. Moradi Ghoje Begloo, A. Shariati, S.H. Hashemabadi, F. Shaahmadi, Mixed modified Fe_2O_3-WO_3 as new fuel borne catalyst (FBC) for biodiesel fuel, Energy 149 (2018) 438–453.

[3] B. Bazooyar, A. Jomekian, A. Shariati, Analysis of the formation and interaction of nitrogen oxides in a rapeseed methyl ester nonpremixed turbulent flame, Energy Fuel. 31 (2017) 8708–8721.

[4] H.G. Darabkhani, H. Varasteh, B. Bazooyar, Carbon Capture Technologies for Gas-Turbine-Based Power Plants, Elsevier, 2022.

[5] B. Bazooyar, M. Zhu, V. Manovic, S.A. Nabavi, Direct numerical simulation (DNS) of packed and monolith syngas catalytic combustors for micro electrical mechanical systems (MEMS), Energy Convers. Manag. X (2023) 100422.

[6] B. Bazooyar, H. Gohari Darabkhani, The design strategy and testing of an efficient microgas turbine combustor for biogas fuel, Fuel (2021) 294.

[7] B. Bazooyar, H. Gohari Darabkhani, Design, manufacture and test of a micro-turbine renewable energy combustor, Energy Convers. Manag. 213 (2020) 112782.

[8] B. Bazooyar, G. Coomson, V. Manovic, S.A. Nabavi, Comparative analysis of ammonia combustion for domestic applications, J. Energy Inst. (2023) 106.

[9] B. Bazooyar, A. Shariati, M. Khosravi-Nikou, S.H. Hashemabadi, Numerical analysis of nitrogen oxides in turbulent lifted H_2/N_2 cabra jet flame issuing into a vitiated coflow, Int. J. Hydrogen Energy 44 (2019) 13932–13952.

[10] X. Kang, C. Liu, S. Zeng, Z. Zhao, J. Qian, Y. Zhao, Prediction of Henry's law constant of CO_2 in ionic liquids based on SEP and Sσ-profile molecular descriptors, J. Mol. Liq. 262 (2018) 139–147.

[11] J.K. Shah, E.J. Maginn, Monte Carlo simulations of gas solubility in the ionic liquid 1 -n-butyl-3-methylimidazolium hexafluorophosphate, J. Phys. Chem. B 109 (2005) 10395–10405.

[12] R. Ahmed Khan, S. Kalam, K. Norrman, M.S. Kamal, M. Mahmoud, A. Abdulraheem, Ionic liquids as clay swelling inhibitors: adsorption study, Energy Fuel. 36 (2022) 3596–3605.

[13] R. Ahmed Khan, M. Murtaza, A. Abdulraheem, M.S. Kamal, M. Mahmoud, Imidazolium-based ionic liquids as clay swelling inhibitors: mechanism, performance evaluation, and effect of different anions, ACS Omega 5 (2020) 26682–26696.

[14] B. Bazooyar, F. Shaahmadi, A. Jomekian, H.G. Darabkhani, Modelling of wax deposition by perturbed hard sphere chain equation of state, J. Pet. Sci. Eng. (2020) 185.

[15] T. Banerjee, M.K. Singh, A. Khanna, Prediction of binary VLE for imidazolium based ionic liquid systems using COSMO-RS, Ind. Eng. Chem. Res. 45 (2006) 3207–3219.

[16] Y. Chen, F. Mutelet, J.N. Jaubert, Modeling the solubility of carbon dioxide in imidazolium-based ionic liquids with the PC-SAFT equation of state, J. Phys. Chem. B 116 (2012) 14375–14388.

[17] M.C. Kroon, E.K. Karakatsani, I.G. Economou, G.J. Witkamp, C.J. Peters, Modeling of the carbon dioxide solubility in imidazolium-based ionic liquids with the tPC-PSAFT equation of state, J. Phys. Chem. B 110 (2006) 9262–9269.

[18] J.W. Qian, R. Privat, J.N. Jaubert, Predicting the phase equilibria, critical phenomena, and mixing enthalpies of binary aqueous systems containing alkanes, cycloalkanes, aromatics, alkenes, and gases (N_2, CO_2, H_2S, H_2) with the PPR78 equation of state, Ind. Eng. Chem. Res. 52 (2013) 16457–16490.

[19] X. Xu, S. Lasala, R. Privat, J.N. Jaubert, E-PPR78: A proper cubic EoS for modelling fluids involved in the design and operation of carbon dioxide capture and storage (CCS) processes, Int. J. Greenh. Gas Control 56 (2017) 126–154.

[20] S. Zhang, Y. Chen, R.X.F. Ren, Y. Zhang, J. Zhang, X. Zhang, Solubility of CO_2 in sulfonate ionic liquids at high pressure, J. Chem. Eng. Data 50 (2005) 230–233.

[21] Y.S. Kim, W.Y. Choi, J.H. Jang, K.P. Yoo, C.S. Lee, Solubility measurement and prediction of carbon dioxide in ionic liquids, Fluid Phase Equil. 228–229 (2005) 439–445.

[22] Y.S. Kim, J.H. Jang, B.D. Lim, J.W. Kang, C.S. Lee, Solubility of mixed gases containing carbon dioxide in ionic liquids: measurements and predictions, Fluid Phase Equil. 256 (2007) 70–74.

[23] T. Wang, C. Peng, H. Liu, Y. Hu, J. Jiang, Equation of state for the vapor-liquid equilibria of binary systems containing imidazolium-based ionic liquids, Ind. Eng. Chem. Res. 46 (2007) 4323–4329.

[24] C.A. Faúndez, F.A. Quiero, J.O. Valderrama, Correlation of solubility data of ammonia in ionic liquids for gas separation processes using artificial neural networks, Compt. Rendus Chem. 17 (2014) 1094–1101.

[25] S.P. Seyyedi Razaz, B. Bazooyar, T. Pirhoushyaran, F. Shaahmadi, Evolving a least square support vector machine using real coded shuffled complex evolution for property estimation of aqueous ionic liquids, Thermochim. Acta 670 (2018) 27–34.

[26] F. Sarlak, T. Pirhoushyaran, F. Shaahmadi, Z. Yaghoubi, B. Bazooyar, The development of intelligent models for liquid–liquid equilibria (LLE) phase behavior of thiophene/alkane/ionic liquid ternary system, Separ. Sci. Technol. 53 (2018) 2935–2951.

[27] B. Bazooyar, F. Shaahmadi, M.A. Anbaz, A. Jomekian, Intelligent modelling and analysis of biodiesel/alcohol/glycerol liquid-liquid equilibria, J. Mol. Liq. (2021) 322.

[28] F. Shaahmadi, M.A. Anbaz, B. Bazooyar, The analysis of liquid–liquid equilibria (LLE) of toluene + heptane + ionic liquid ternary mixture using intelligent models, Chem. Eng. Res. Des. 130 (2018) 184–198.

[29] X. Kang, Z. Zhao, J. Qian, R.M. Afzal, Predicting the viscosity of ionic liquids by the ELM intelligence algorithm, Ind. Eng. Chem. Res. 56 (2017) 11344–11351.

[30] S. Kalam, U. Yousuf, S.A. Abu-Khamsin, U Bin Waheed, R.A. Khan, An ANN model to predict oil recovery from a 5-spot waterflood of a heterogeneous reservoir, J. Pet. Sci. Eng. (2022) 210.

[31] S. Kalam, S.A. Abu-Khamsin, H.Y. Al-Yousef, R. Gajbhiye, A novel empirical correlation for waterflooding performance prediction in stratified reservoirs using artificial intelligence, Neural Comput. Appl. 33 (2021) 2497–2514.

[32] S. Kalam, R.A. Khan, S. Khan, M. Faizan, M. Amin, R. Ajaib, et al., Data-driven modeling approach to predict the recovery performance of low-salinity waterfloods, Nat. Resour. Res. 30 (2021) 1697–1717.

[33] B. Bazooyar, F. Shaahmadi, A. Jomekian, S.S. Mirfasihi, Carbon capture via aqueous ionic liquid green solutions intelligent modelling, Case Stud Chem Environ Eng (2023) 100444.

[34] C.A. Faúndez, E.N. Fierro, J.O. Valderrama, Solubility of hydrogen sulfide in ionic liquids for gas removal processes using artificial neural networks, J. Environ. Chem. Eng. 4 (2016) 211–218.

[35] R.L. Gardas, J.A.P. Coutinho, Estimation of speed of sound of ionic liquids using surface tensions and densities: a volume based approach, Fluid Phase Equil. 267 (2008) 188–192.

[36] A. Shariati, C.J. Peters, High-pressure phase behavior of systems with ionic liquids: II. The binary system carbon dioxide+1-ethyl-3-methylimidazolium hexafluorophosphate, J. Supercrit. Fluids 29 (2004) 43–48.

[37] F. Shaahmadi, M.A. Anbaz, B. Bazooyar, Analysis of intelligent models in prediction nitrous oxide (N2O) solubility in ionic liquids (ILs), J. Mol. Liq. 246 (2017) 48–57.

[38] D. Camper, C. Becker, C. Koval, R. Noble, Diffusion and solubility measurements in room temperature ionic liquids, Ind. Eng. Chem. Res. 45 (2006) 445–450.

[39] L.A. Blanchard, Z. Gu, J.F. Brennecke, High-pressure phase behavior of ionic liquid/CO_2 systems, J. Phys. Chem. B 105 (2001) 2437–2444.

[40] M. Costantini, V.A. Toussaint, A. Shariati, C.J. Peters, I. Kikic, High-pressure phase behavior of systems with ionic liquids: Part IV. Binary system carbon dioxide + 1-hexyl-3-methylimidazolium tetrafluoroborate, J. Chem. Eng. Data 50 (2005) 52–55.

[41] M. Shokouhi, M. Adibi, A.H. Jalili, M. Hosseini-Jenab, A. Mehdizadeh, Solubility and diffusion of H_2S and CO_2 in the ionic liquid 1-(2-Hydroxyethyl)-3-methylimidazolium tetrafluoroborate, J. Chem. Eng. Data 55 (2010) 1663–1668.

[42] A. Shafiei, M.A. Ahmadi, S.H. Zaheri, A. Baghban, A. Amirfakhrian, R. Soleimani, Estimating hydrogen sulfide solubility in ionic liquids using a machine learning approach, J. Supercrit. Fluids 95 (2014) 525–534.

[43] A. Baghban, M.A. Ahmadi, B.H. Shahraki, Prediction carbon dioxide solubility in presence of various ionic liquids using computational intelligence approaches, J. Supercrit. Fluids 98 (2015) 50–64.

[44] M.A. Ahmadi, B. Pouladi, Y. Javvi, S. Alfkhani, R. Soleimani, Connectionist technique estimates H_2S solubility in ionic liquids through a low parameter approach, J. Supercrit. Fluids 97 (2015) 81–87.

[45] M. Althuluth, M.C. Kroon, C.J. Peters, Solubility of methane in the ionic liquid 1-ethyl-3-methylimidazolium tris(pentafluoroethyl)trifluorophosphate, Ind. Eng. Chem. Res. 51 (2012) 16709–16712.

[46] T. Wang, C. Peng, H. Liu, Y. Hu, Description of the pVT behavior of ionic liquids and the solubility of gases in ionic liquids using an equation of state, Fluid Phase Equil. 250 (2006) 150–157.
[47] M.D. Bermejo, T.M. Fieback, Á. Martín, Solubility of gases in 1-alkyl-3methylimidazolium alkyl sulfate ionic liquids: experimental determination and modeling, J. Chem. Thermodyn. 58 (2013) 237–244.
[48] D. Almantariotis, S. Stevanovic, O. Fandiño, A.S. Pensado, A.A.H. Padua, J.Y. Coxam, et al., Absorption of carbon dioxide, nitrous oxide, ethane and nitrogen by 1-alkyl-3-methylimidazolium (Cnmim, n = 2,4,6) tris(pentafluoroethyl) trifluorophosphate ionic Liquids (eFAP), J. Phys. Chem. B 116 (2012) 7728–7738.
[49] J.L. Anderson, J.K. Dixon, J.F. Brennecke, Solubility of CO_2, CH_4, C_2H_6, C_2H_4, O_2, and N_2 in 1-hexyl-3-methylpyridinium bis(trifluoromethylsulfonyl) imide: comparison to other ionic liquids, Acc. Chem. Res. 40 (2007) 1208–1216.
[50] J.L. Anthony, E.J. Maginn, J.F. Brennecke, Solubilities and thermodynamic properties of gases in the ionic liquid 1-n-butyl-3-methylimidazolium hexafluorophosphate, J. Phys. Chem. B 106 (2002) 7315–7320.
[51] C. Cadena, J.L. Anthony, J.K. Shah, T.I. Morrow, J.F. Brennecke, E.J. Maginn, Why is CO_2 so soluble in imidazolium-based ionic liquids? J. Am. Chem. Soc. 126 (2004) 5300–5308.
[52] M.F.C. Gomes, Low-pressure solubility and thermodynamics of solvation of carbon dioxide, ethane, and hydrogen in 1-hexyl-3-methylimidazolium bis(-trifluoromethylsulfonyl) amide between temperatures of 283 K and 343 K, J. Chem. Eng. Data 52 (2007) 472–475.
[53] L.J. Florusse, S. Raeissi, C.J. Peters, High-pressure phase behavior of ethane with 1-hexyl-3-methylimidazolium bis(trifluoromethylsulfonyl)imide, J. Chem. Eng. Data 53 (2008) 1283–1285.
[54] S. Haykin, N. Network, A Comprehensive Foundation, 2004.
[55] G. Hong, J. Jacquemin, M. Deetlefs, C. Hardacre, P. Husson, M.F. Costa Gomes, Solubility of carbon dioxide and ethane in three ionic liquids based on the bis{(trifluoromethyl)sulfonyl}imide anion, Fluid Phase Equil. 257 (2007) 27–34.
[56] J. Jacquemin, M.F. Costa Gomes, P. Husson, V. Majer, Solubility of carbon dioxide, ethane, methane, oxygen, nitrogen, hydrogen, argon, and carbon monoxide in 1-butyl-3-methylimidazolium tetrafluoroborate between temperatures 283 K and 343 K and at pressures close to atmospheric, J. Chem. Thermodyn. 38 (2006) 490–502.
[57] J. Jacquemin, P. Husson, V. Majer, M.F.C. Gomes, Low-pressure solubilities and thermodynamics of solvation of eight gases in 1-butyl-3-methylimidazolium hexafluorophosphate, Fluid Phase Equil. 240 (2006) 87–95.
[58] J. Kumełan, Á.P.S. Kamps, D. Turna, G. Maurer, Solubility of the single gases methane and xenon in the ionic liquid [hmim][Tf2N], Ind. Eng. Chem. Res. 46 (2007) 8236–8240.
[59] B.-C. Lee, S.L. Outcalt, Solubilities of gases in the ionic liquid 1- n -Butyl-3-methylimidazolium bis(trifluoromethylsulfonyl)imide, J. Chem. Eng. Data 51 (2006) 892–897.
[60] S. Raeissi, C.J. Peters, High pressure phase behaviour of methane in 1-butyl-3-methylimidazolium bis(trifluoromethylsulfonyl)imide, Fluid Phase Equil. 294 (2010) 67–71.
[61] S. Stevanovic, M.F. Costa Gomes, Solubility of carbon dioxide, nitrous oxide, ethane, and nitrogen in 1-butyl-1-methylpyrrolidinium and trihexyl(tetradecyl)phosphonium tris(pentafluoroethyl)trifluorophosphate (eFAP) ionic liquids, J. Chem. Thermodyn. 59 (2013) 65–71.
[62] X. Yuan, S. Zhang, Y. Chen, X. Lu, W. Dai, R. Mori, Solubilities of gases in 1,1,3,3-tetramethylguanidium lactate at elevated pressures, J. Chem. Eng. Data 51 (2006) 645–647.
[63] J.L. Anthony, J.L. Anderson, E.J. Maginn, J.F. Brennecke, Anion effects on gas solubility in ionic liquids, J. Phys. Chem. B (2005).
[64] J.O. Valderrama, L.A. Forero, R.E. Rojas, Critical properties and normal boiling temperature of ionic liquids. Update and a new consistency test, Ind. Eng. Chem. Res. 51 (2012) 7838–7844.
[65] J.O. Valderrama, R.E. Rojas, Critical properties of ionic liquids. Revisited, Ind. Eng. Chem. Res. 48 (2009) 6890–6900.
[66] R. Eberhart, J. Kennedy, New optimizer using particle swarm theory, Proc. Int. Symp. Micro Mach. Hum. Sci. (1995) 39–43.
[67] C. Darwin, On the Origins of Species by Means of Natural Selection, 247, Murray, London, 1859, p. 1859.
[68] J.H. Holland, Adaption in Natural and Artificial Systems, 211, MI Univ Michigan Press, Ann Arbor, 1992.
[69] Q.Y. Duan, V.K. Gupta, S. Sorooshian, Shuffled complex evolution approach for effective and efficient global minimization, J. Optim. Theor. Appl. 76 (1993) 501–521.
[70] J.C. Hoskins, D.M. Himmelblau, Artificial neural network models of knowledge representation in chemical engineering, Comput. Chem. Eng. 12 (1988) 881–890.
[71] K. Hornik, M. Stinchcombe, H. White, Multilayer feedforward networks are universal approximators, Neural Network. 2 (1989) 359–366.
[72] M. Mehraban, M.A. Anbaz, F. Shaahmadi, B. Bazooyar, Property estimation of water/alcohol/ionic liquid ternary system: density, J. Mol. Liq. 264 (2018) 88–97.
[73] N. García-Pedrajas, C. Hervás-Martínez, J. Muñoz-Pérez, COVNET: A cooperative coevolutionary model for evolving artificial neural networks, IEEE Trans. Neural Network. 14 (2003) 575–596.
[74] K. Hornik, M. Stinchcombe, H. White, Universal approximation of an unknown mapping and its derivatives using multilayer feedforward networks, Neural Network. 3 (1990) 551–560.
[75] N. Murata, S. Yoshizawa, S.I. Amari, Network information criterion—determining the number of hidden units for an artificial neural network model, IEEE Trans. Neural Network. 5 (1994) 865–872.
[76] T. Hill, L. Marquez, M. O'Connor, W. Remus, Artificial neural network models for forecasting and decision making, Int. J. Forecast. 10 (1994) 5–15.
[77] V. Vapnik, The Nature of Statistical Learning Theory, Springer science & business media, 1999.
[78] C. Cortes, V. Vapnik, Support-vector networks, Mach. Learn. 20 (1995) 273–297.
[79] U. Norinder, Support vector machine models in drug design: applications to drug transport processes and QSAR using simplex optimisations and variable selection, Neurocomputing 55 (2003) 337–346.
[80] A. Baylar, D. Hanbay, M. Batan, Application of least square support vector machines in the prediction of aeration performance of plunging overfall jets from weirs, Expert Syst. Appl. 36 (2009) 8368–8374.
[81] Y. Ren, H. Liu, X. Yao, M. Liu, Prediction of ozone tropospheric degradation rate constants by projection pursuit regression, Anal. Chim. Acta 589 (2007) 150–158.
[82] J.A.K. Suykens, J. Vandewalle, Least squares support vector machine classifiers, Neural Process. Lett. 9 (1999) 293–300.
[83] C.H. Li, X.J. Zhu, G.Y. Cao, S. Sui, M.R. Hu, Identification of the Hammerstein model of a PEMFC stack based on least squares support vector machines, J. Power Sources 175 (2008) 303–316.
[84] B. Bazooyar, A. Shariati, S.H. Hashemabadi, Turbulent non-premixed combustion of rapeseed methyl ester in a free shear swirl air flow, Ind. Eng. Chem. Res. 55 (2016) 11645–11663.
[85] F. Amirkhani, A. Dashti, H. Abedsoltan, A.H. Mohammadi, K.W. Chau, Towards estimating absorption of major air pollutant gasses in ionic liquids using soft computing methods, J. Taiwan Inst. Chem. Eng. 127 (2021) 109–118.
[86] A. Dashti, H. Riasat Harami, M. Rezakazemi, S. Shirazian, Estimating CH_4 and CO_2 solubilities in ionic liquids using computational intelligence approaches, J. Mol. Liq. 271 (2018) 661–669.
[87] R. Nakhaei-Kohani, S. Atashrouz, F. Hadavimoghaddam, A. Abedi, K. Jabbour, A. Hemmati-Sarapardeh, et al., Chemical structure and thermodynamic properties based models for estimating nitrous oxide solubility in ionic Liquids: equations of state and Machine learning approaches, J. Mol. Liq. (2022) 367.
[88] H. Feng, W. Qin, G. Hu, H. Wang, Intelligent prediction of nitrous oxide capture in designable ionic liquids, Appl. Sci. 13 (2023) 6900.
[89] R. Nakhaei-Kohani, S. Atashrouz, F. Hadavimoghaddam, A. Bostani, A. Hemmati-Sarapardeh, A. Mohaddespour, Solubility of gaseous hydrocarbons in ionic liquids using equations of state and machine learning approaches, Sci. Rep. 12 (2022).
[90] M. Safamirzaei, H. Modarress, Correlating and predicting low pressure solubility of gases in [bmim][BF 4] by neural network molecular modeling, Thermochim. Acta 545 (2012) 125–130.