Iranian Journal of Chemistry and Chemical Engineering (IJCCE)

Nonrandom Two-Liquid Model with Differential Evolution for Experimental Liquid-Liquid Equilibrium Calculation

HEBBOUL, Sabrina[1]; BACHA, Oussama[1]; KORICHI, Mourad[1]; HEBBOUL, Amel[2]*

[1]Dynamics, Interaction, and Reactivity of Systems laboratory, *Department of Process Engineering, Faculty of Science Applied, Kasdi Merbah University, Ouargla, ALGERIA*

[2]Automatic and Robotic Laboratory, *Department of Mathematics and Computer science*, Fères Mentouri University, *Constantine, ALGERIA*

* To whom correspondence should be addressed.

E-mail: bacha.oussama@gmail.com / bacha.oussama@univ-ouargla.dz

ABSTRACT: *In this study, the Nonrandom Two-Liquid model (NRTL) was used to calculate the isoactivity equations from the experimental liquid-liquid equilibrium (LLE) data. Additionally, the original Differential Evolution method (DE_rand1) and its modifications involve the self-adaptive control parameters differential evolution (JDE), the adaptive differential evolution with optional external archive (JADE), and the composite differential evolution (CODE) have been used to estimate the binary interaction parameters. Randomization of regression parameters has been used to minimize the fitting objective function. Furthermore, the effectiveness of these optimization methods was tested in a quaternary system of water, acetic acid, 50 % dichloromethane (DCM), and 50 % methyl isobutyl ketone (MIBK) at 301.15 K. Moreover, the optimization process assessment was carried out by a regression analysis using Root Mean Square Deviation (RMSD), mean, standard deviation, and the duration of execution time spent on both the activity and the fractional objective functions. Root Mean Square Deviation (RMSD) results that were less than or equal to 0.0107 demonstrated the effectiveness of Differential Evolution methods in estimating NRTL parameters for this specific system. Finally, the original method (DE_rand1) was found to be the most efficient approach among all its variations.*

KEYWORDS: *Liquid–Liquid Equilibrium, Nonrandom Two-Liquid model, Isoactivity equations, Objective functions, Differential Evolution methods.*

INTRODUCTION

Liquid-liquid extraction is a highly effective technique for separating substances or liquid mixtures and is also called solvent extraction [1–4]. This technique is used in different domains such as industrial processes, pharmaceutical chemistry, biomedical, mineral chemistry, metallurgy fields, and biology [5–9]. Additionally, the ability of this method is evaluated by the distribution coefficients (D_i) and the separation factor (S) [10, 12]. The distribution coefficient assesses the dispersion of elements between equilibrium phases, whereas the separation factor evaluates the effectiveness of this technique [13]. Consequently, these parameters are essential for determining the exact elements that influence solvent extraction, such as the properties of the solvent and the solute, etc [13, 14].

Liquid-Liquid Equilibrium (LLE) is a fundamental for extraction applications and plays a significant role in refining the modeling and optimization of industrial processes. Therefore, determining it precisely is essential to developing reliable and predictive thermodynamic approaches [15, 16]. Thus, it is imperative to conduct these calculations properly and accurately to prevent shortcomings, mistakes, and ambiguities during process development. There are two techniques commonly employed for calculating LLE. They are the Gibbs energy minimization technique and the equation-solving isoactivity technique [17, 18]. Furthermore, the isoactivity method is widely used in LLE calculations involving activity coefficient models due to its simplicity [19].

The activity coefficient models are mainly based on the development of excess Gibbs-energy equations [19, 20]. The activity coefficient is a fundamental parameter in thermodynamics. It is defined to describe the behavior of real solutions deviating from an ideal state in a chemical mixture [21, 22]. The capacity of these models can be divided into two categories, namely predictive models and semi-predictive models [22]. The second approach necessitates experimental data, such as Wilson [20], Universal Quasi-Chemical (UNIQUAC) [19] and the Nonrandom Two-Liquid model (NRTL) [23], which was widely employed for estimating and predicting LLE, as reported in previous studies [24, 25].

The calculations of liquid-liquid equilibrium data (LLE) and studying phase behaviors use both deterministic and stochastic optimization methods [26]. On the one hand, the deterministic optimization techniques [26] like Nelder-Mead (NM) and Quasi-Newton (QN) are non-random and use given initial points when executing their algorithms to determine solutions. These methods are accurate for local optima and computationally efficient. However, because of their deterministic character, these methods may not be adequate to obtain global optimization for multi-modal problems [26].

On the other hand, stochastic optimization methods, which are also called metaheuristics, are a type of single-objective optimization (SOO) method. They use random numbers, sequences and probabilistic elements in their search strategies to find the global optimal [26]. In each iteration or objective function evaluation, this approach manipulates either a single solution or a group of solutions [26]. The Genetic Algorithms (GA) [27, 28], Particle Swarm Optimization (PSO) [29, 31], Differential Evolution (DE) [32], and its derivatives (the self-adaptive control parameters differential evolution (JDE), the adaptive differential evolution with optional external archive (JADE), and the composite differential evolution (CODE)) are widely used due to their ability to tackle a wide range of complex optimization problems efficiently and reliably identify global optima despite their time-consuming execution [33–36].

Each method has preferable conditions for application, e.g. Particle Swarm Optimization (PSO) is known for its fast convergence rate and simplicity. Meanwhile, the Genetic Algorithm (GA) is often used for problems with a large search space and can deal with multiple objectives. Differential Evolution (DE) is known for handling difficult and high-dimensional problems. Selecting the optimization method depends on several factors, such as the specific problem being solved, the execution time [37], the calculated Root Mean Square Deviation (RMSD) [38], and the dependency on initial population positions [39, 40]. Several factors can influence the efficiency of these algorithms in terms of execution time [37], the nature of the problem, the number of iterations, and the hardware used [39, 40]. Initially, the execution time can also vary depending on the implementation of the algorithm. The DE generally outperformed the other algorithms in terms of execution time [41, 42].

The Root Mean Square Deviation (RMSD) is a mathematical metric that was used in several works [43] to assess the fit between the experimental data and the calculated data from the NRTL model. Bacha et al. [43], Zheng Li et al. [44], and Timedjeghdine et al. [45] used the PSO and GA methods for LLE calculation of ternary and quaternary systems. As a result, both the GA and PSO methods have been able to accurately estimate the NRTL model parameters. According to the literature [17, 18, 46], a comparative study was conducted between these two methods, and the results indicated that PSO outperformed GA. Then, Warisa Wisittipanich et al. [47] and Hasan Ozcan [37] utilized the DE and GA methods. The DE approach performs better than the GA method. Because it is less affected by parameter choices, converges more quickly, and performs well when dealing with continuous variables. In accordance with the literature [48], this approach facilitates successful search space exploration, which in many cases yields better optimization results. Zhang et al. [49] presented a study that evaluated and compared the performance of DE and PSO algorithms in solving optimization problems and concluded that DE outperformed PSO in most of the tested functions.

Moreover, the common issue with optimization algorithms is their dependence on the initial population positions because the quality of the solutions that are found relies heavily on this initial configuration [42]. Genetic algorithms (GA) and particle swarm optimization (PSO) are particularly vulnerable since they use the initial population to generate new solutions [50]. Conversely, differential evolution (DE) is the least sensitive method because it uses two population initialization methods including, pre-regressed and random methods (randomization procedure) [44, 51]. The first procedure executes the optimization algorithm and its initialization is random [52]. The second procedure executes the objective function in terms of activity (OF_a) in the first step and its initialization is random. Then, it uses the objective function in terms of mole fraction (OF_x) in the second step, along with the initialization from the previous step [44]. In summary, DE method performance significantly surpasses the other algorithms evaluated in prior studies [42, 48].

One of the most significant issues in improving the prediction of LLE is accurately representing the system's phase behavior. In light of that, this work aims to study the behavior of the quaternary system, which includes water, acetic acid, and a mixture of dichloromethane (DCM) and methyl isobutyl ketone (MIBK) at a temperature of 301.5 K [53] by using the DE standard (DE_rand1) and its variants (JDE, JADE, and CODE). Additionally, the phase diagrams will be represented to illustrate the feasibility of the calculated data in relation to the experimental results.

Furthermore, the NRTL model was employed to model and correlate the studied system based on experimental LLE data. The selection of the DE method was based on a prior comparison demonstrating its superiority over PSO and GA in terms of dependency on initial population positions, Root Mean Square Deviation (RMSD), and execution time. Notably, previous studies have not utilized DE techniques for a quaternary system in LLE calculation and modeling, highlighting the novelty of this approach in studying quaternary system behavior and obtaining optimal NRTL parameters. The DE methods calculations rely on the performance of isoactivity equations as numerical solutions. The purpose of this concept is to make improvements in the optimization process to generate high-quality data.

Moreover, the achieved outcomes underwent optimization by carefully selecting an appropriate objective function, commonly expressed in terms of both the activity objective function (OF_a) and the fractional objective function (OF_x). This work was conducted based on the observation that previous studies did not indicate the significance of the objective function selection in calculating and modeling Liquid-Liquid Equilibrium (LLE). The primary focus of the study is on examining how the population size, range of interaction parameters, and number of iterations influence the value of the objective function (fitness). Besides, these regression parameters were determined using the randomization procedure. The purpose is to find the optimal values for executing the DE algorithm.

Finally, the calculated results for the DE approaches were evaluated according to the regression analysis of estimated binary interaction parameters such as the Root Mean Square Deviation (RMSD), the mean (m), and the standard deviation (std) [54]. This comparative analysis aimed to identify the most effective optimization method for calculating and modeling the Liquid-Liquid Equilibrium quaternary system.

THERMODYNAMICS OF LIQUID-LIQUID EQUILIBRIUM CONDITIONS

The isoactivity equations of elements for both phases are used for the calculation and modeling of the LLE of quaternary systems containing water, acetic acid, and a mixture of dichloromethane (DCM) and methyl isobutyl ketone (MIBK) at a temperature of 301.5 K. It is represented by the following set of Eqs from (1) to (7).

At equilibrium, the Eq. 1 illustrated an equality of pressures, temperatures, and chemical potentials in each phase, as follows: [55]:

$$\begin{cases} P^{P1} = P^{P2} \\ T^{P1} = T^{P2} \\ \mu_i^{P1} = \mu_i^{P2} \end{cases} \quad (1)$$

Where: P, T, and μ represent the pressures (atm), the temperature (K), and the chemical potentials of component i (J/mol), respectively.

The chemical potential plays a crucial role in studying LLE equilibrium conditions [17, 56]. The chemical potential of a component is assumed to be constant during the equilibrium process according to the isoactivity approach [18]. Furthermore, there exists a correlation between the chemical potential and the chemical activity, which is represented in the Eq. 2 [57]:

$$\begin{cases} \mu_i^{P1} = R.T.\ln a_i^{P1} \\ \mu_i^{P2} = R.T.\ln a_i^{P2} \end{cases} \quad (2)$$

Where: a_i is the chemical activity of component *i* (a dimensionless quantity). The superscripts refer to two phases. R indicates the universal gas constant (cal K^{-1} mol^{-1}).

The following Eq. (3) can be concluded from the two Eqs. (1) and (2), as follows [55]:

$$a_i^{P1} = a_i^{P2} \quad (3)$$

Eq. 3 presents the basis for phase equilibrium computations when component activities in both phases are equal. In addition, the concept of chemical activity is also indicated by the relation between the activity coefficient and mole fraction, as shown in the Eq. 4 [55]:

$$a_i = x_i.\gamma_i \quad (4)$$

According to the Eqs. from (1) to (4) [58]:

$$x_i^{P1}\,\gamma_i^{P1} = x_i^{P2}\,\gamma_i^{P2} \quad (5)$$

Moreover, the set of Eqs. from (1) to (5) requires material balance constraints by adding total element quantities, hence [44]:

$$n_i^{P1} + n_i^{P2} = n_i \quad (6)$$

Where: n_i^{P1} *and* n_i^{P2} refer to the mole number of constituent *i*. n_i represent the total quantity of constituents *i* in the mixture (mole). The superscripts refer to two phases respectively.

In a system that contains N elements and two phases, there are a total of 2N variables (the mole fractions of each element in each phase). The mass balance constraints for both phases are determined after adjusting the concentrations of (N - 2) elements in one phase, as follows [3]:

$$\begin{cases} \sum_{i=1}^{N} x_i^{P1} = 1 \\ \sum_{i=1}^{N} x_i^{P2} = 1 \end{cases} \quad (7)$$

The normalization of the mole fraction requires two formulas for each element in the equilibrium of the system according to Eq. (7). As a result, there are (N + 2) variables with (N + 2) equations. These equations involve N isoactivity equations and two constraint equations [58]. Then, the unknown amounts can be determined according to the solving of isoactivity equations when the number of unknowns is similar to the number of equations [56].

Finally, these isoactivity equations are used for the calculation and prediction of LLE when using the NRTL model. This numerical solution performs when an optimization method such as DE is applied to find optimal solutions.

NRTL MODEL

In 1964, Wilson [20] proposed a model showing that local concentrations around molecules differ from the overall concentration. This model has two adjustable parameters (A_{ji} *and* A_{ij}) [56]. However, the Wilson model was unable to predict LLE [59]. After four years, Renon et al. developed another model to be able to predict LLE and called it the non-random two-liquid (NRTL) model [23]. This model is based on Wilson's local composition theory (as represented in Fig. 1) and Scott's two-liquid solution theory [20].

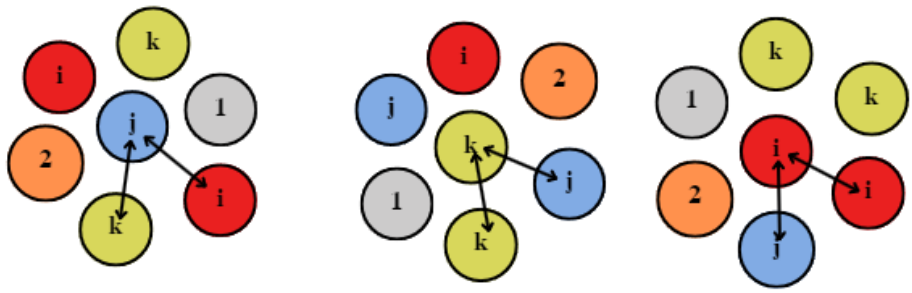


Fig. 1: Local composition theory of Wilson.

The relation between the composition (x_i) and activity coefficient (γ_i) of element *i* is expressed in the following formula [60]:

$$\ln\gamma_i = \frac{\sum_{j=1}^{N} \tau_{ji}.G_{ji}.x_j}{\sum_{k=1}^{N} G_{ki}.x_k} + \sum_{j=1}^{N} \frac{G_{ij}.x_j}{\sum_{k=1}^{N} G_{kj}.x_k}\left(\tau_{ij} - \frac{\sum_{l=1}^{N} \tau_{lj}.G_{lj}.x_l}{\sum_{k=1}^{N} G_{kj}.x_k}\right) \quad (8)$$

The interaction parameters can be defined as follows [4], [23]:

$$\tau_{ij} = \frac{A_{ij}}{R.T} \quad (9)$$

$$G_{ji} = exp\left(-\alpha_{ji}.\tau_{ji}\right) \quad (10)$$

Where: τ_{ij} , G_{ij} , and A_{ij} are the energy interaction parameters ($A \neq A_{ij}$). α_{ji} refers to non-random solution dispersion parameter. R is the gas constant (cal $K^{-1}\,mol^{-1}$). T denotes the temperature of the mixture liquid (K).

This model has three parameters for each pair of constituents $i-j$. The two binary molecular energy interaction parameters are the same, according to Wilson's model. They are determined by using the isoactivity equations after the estimated data and the coefficient activity model are defined for both phases. Additionally, a new parameter presents the non-random distribution of molecules according to the second theory (theory of Scott) on which it is based, which is named the non-randomness factor or the non-random solution dispersion parameter α_{ji} and α_{ij}. Its value ranges between 0.20 and 0.47 [17], with $\alpha_{ji} = \alpha_{ij}$ [3]. The non-randomness factor value of 0.20 is commonly assigned to LLE [18]. In comparison to other activity models [23], this factor gives it more flexibility, making it possible to show the phase behavior of more mixtures in liquid phases.

Although, the primary drawback of the NRTL model is the significant correlation that exists between its two parameters [61]. However, the NRTL model has the ability to reflect experimental equilibrium features with a comparable level of accuracy [62]. In this study, this model was used for LLE calculations and modeling of the quaternary system involving water, acetic acid, and a mixed solvent (50% MIBK with 50% DCM) at a temperature of 301.15 K.

DIFFERENTIAL EVOLUTION METHODS

Differential Evolution (DE) is a stochastic optimization technique within evolutionary calculation that uses a population of potential solutions. In this study, it was used to calculate LLE for the selected quaternary system.

The population is guided towards better solutions in the search range using three main operators, which are mutation, crossover, and selection. Storn and Price introduced this technique in 1995 [32] to minimize non-differentiable and potentially nonlinear continuous functions.

The performance of the DE standard algorithm was based on two steps (as shown in Fig. 2). The first step is initialization, in which the population is formed at random. In the second step, called evolution, the created population continues through three operators such as mutation, crossover, and selection (as represented in Fig. 3), which are repeated until a termination criterion is satisfied [32].

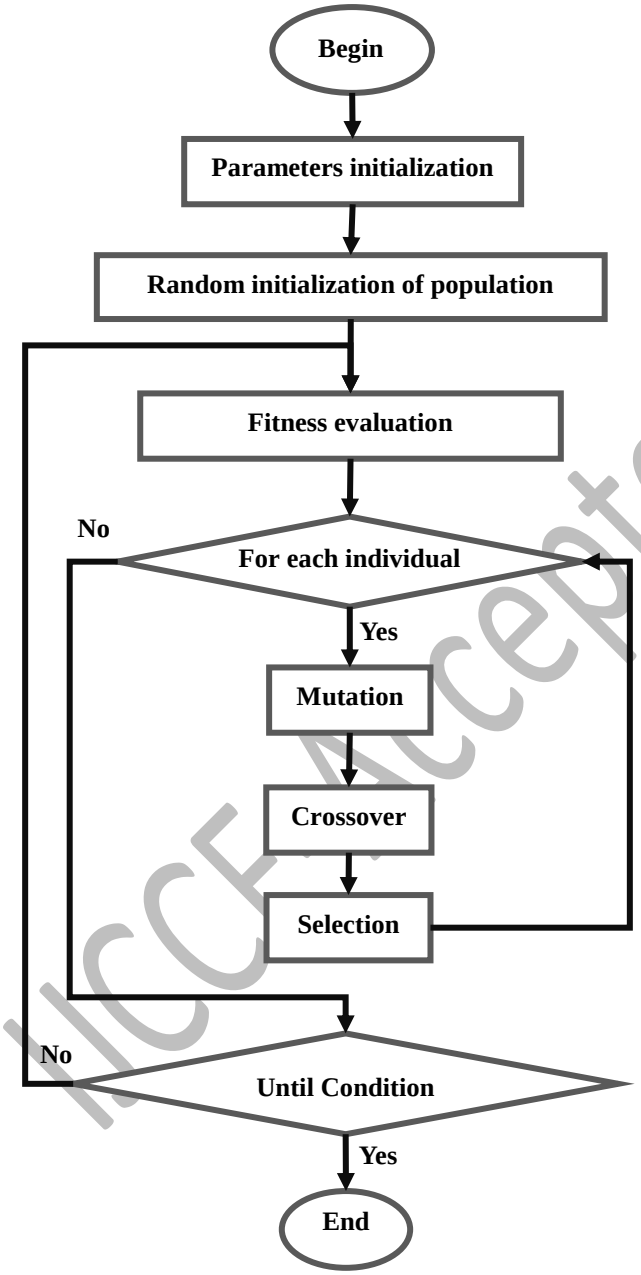


Fig. 2: Flowchart of the Differential Evolution (DE) Algorithm.

Mutation operation

In mutation operation of the DE standard, a new individual is created through three randomly selected and is named a mutant vector as shown in Fig 3 (a) and using Eq. 11 [32]:

$$DE\_rand1: V = X_{r_1} + F\left(X_{r_2} - X_{r_3}\right) \quad (11)$$

In the literature, the DE standard has other mutation equations including DE/best/1, DE/rand/2, DE/best/2, DE/current-to-best/1, and DE/current-to-rand/1, such as:

1- $\mathrm{DE/best/1}: V = X_{best} + F\left(X_{r_1} - X_{r_2}\right);$

2- $\mathrm{DE/rand/2}: V = X_{r_1} + F\left(X_{r_2} - X_{r_3}\right) + F\left(X_{r_4} - X_{r_5}\right);$

3- $\mathrm{DE/best/2}: V = X_{best} + F\left(X_{r_1} - X_{r_2}\right) + F\left(X_{r_3} - X_{r_4}\right);$

4- $\mathrm{DE/current-to-best/1}: V = X_i + F(X_{best} - X_i) + F\left(X_{r_1} - X_{r_2}\right);$

5- $\mathrm{DE/current-to-rand/1}: \mathrm{V} = X_i + rand\left(X_{r_1} - X_i\right) + F\left(X_{r_2} - X_{r_3}\right).$

Where: V indicates the mutant vector. X_i refers to the current vector. $r_1, r_2, r_3, r_4\ and\ r_5$ denote the random integers, $r_1 \neq r_2 \neq r_3 \neq r_4 \neq r_5 \neq i$, with $i \in \{1,2,\dots,Np\}$ and Np represents the population size. X_{best} refers to the best individual vector, or each individual that has the best objective function value at the present generation in the population. F means the mutation factor $\in [0,2]$ which controls the differential vector $(X_{r_2} - X_{r_3})$.

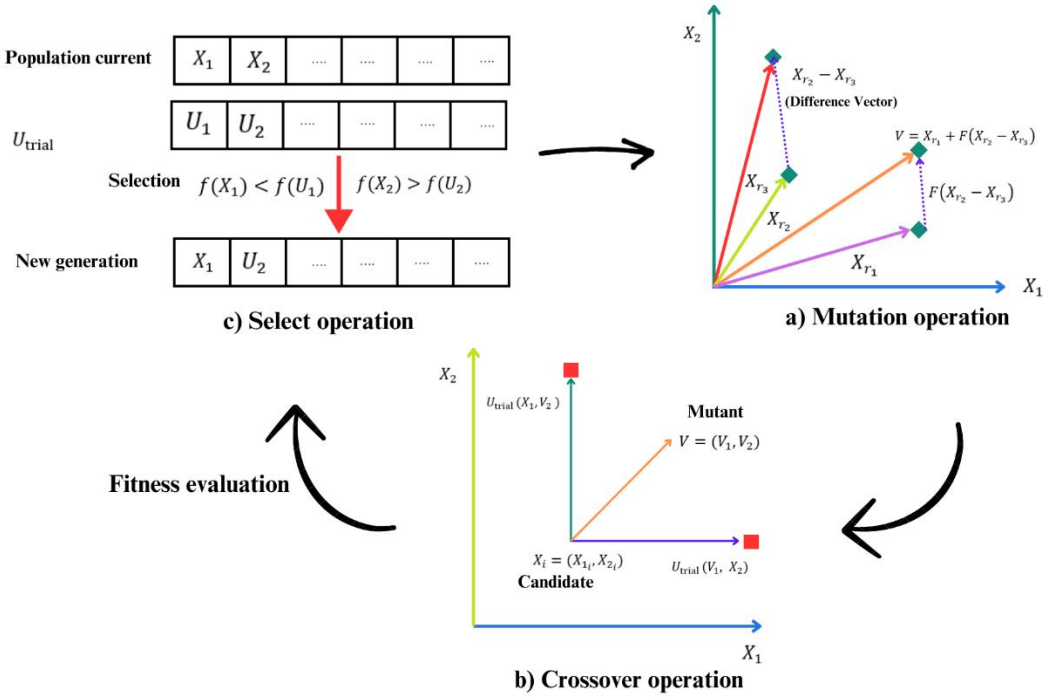


Fig. 3: illustration of the essential operators (selection, crossover, and mutation) used in the DE standard algorithm.

Crossover operation

The crossover operation aims to increase the population variety and is depicted in Fig 3 (b). Additionally, its formula refers to the combination between the mutated vector $V = (V_1, V_2, V_3, \dots V_D)$ with the current vector $X_i = (X_{1i}, X_{2i}, X_{3i}, \dots X_{Di})$, to provide the trial individual (trial vector) presented as follows [32]:

$$U_j = \begin{cases} V_j & if\,(rand < CR)\ or\ (j = irand(D)) \\ X_{ji} & if\,(rand \geq CR) \end{cases} \quad (12)$$

Where: U denotes the trial individual. rand refers to a function that generates random numbers between 0 and 1. CR indicates the crossover factor $\in [0,1]$. D is the dimension of the search space. X_{ji} represente the j^{th} position

in the current vector X_i with $j \in \{1,2,\ldots,D\}$. irand (D) means a function that generates random numbers between 1 and D.

Select operation

The purpose of select operation is to choose the best option from the current vector (X_i) and trial individual (U) [32]. This operation was represented in Fig 3 (c). Moreover, its concept states the X_i new value is set to the U, if U provides the objective function value that is lower than the prior value; Else, its value remains the same. Furthermore, it is defined as:

$$X_i^{t+1} = \begin{cases} U & if\ f(U) < f(X_i^t) \\ X_i^t \end{cases} \quad (13)$$

Where: X_i^t refers to the current vector in the iteration t.

DE's control parameters sitting

The next important point in the DE algorithm is the control parameters. These parameters include population size (Np), the mutation factor (F), Furthermore, selecting these parameters to get better outcomes is relatively easy and greatly affects optimization performance that one needs to adjust. The fundamental instructions of parameters allow for dealing with DE, which is extremely simple and considered one of DE's key advantages. In addition, the user has to choose a maximum number of generations or iterations [32].

Population size

A population is a group of individuals that represent the solution to the issue that the DE algorithm is attempting to solve. Each individual contains a set of genes as the space search dimensionality. In this study, the number of genes equals 2× the number of compounds (as illustrated in Fig. 4). The initial population is randomly selected and should include the entire area of search [32]. The population size (Np) is a fixed value during the minimization operation. The optimum range is between 5 to 20 times the total number of decision variables [26]. Therefore, the Np strongly influences the algorithm's capacity for exploration. Also, it must also be considered to enable the algorithm to search in the multi-dimensional range if there are problems with an abundance of dimensions.

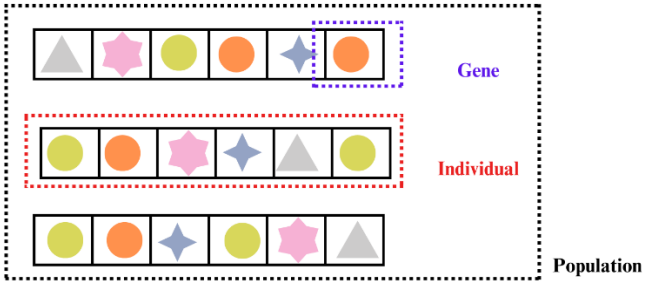


Fig. 4: Structure of Population.

Mutation factor

The mutation factor (F) is also called the scaling factor and the differential weight due to its function which represents scaling the difference vector and manages its amplification. In addition, the desirable range is between 0 and 2. The lower values of F will result in lower mutation step values, which will delay the algorithm's

convergence. While the high F values enable exploration less difficult, they could also lead to a local optimum. As a result, the value must be both sufficiently small to encourage local exploration and big enough to preserve variety. For instance, choosing $F = 0.5$ is often the first perfect option. As well as, the values of F are rarely effective when they are less than 0.4 or bigger than 1 [26, 32]. Increasing the values of F and Np, or one of them, is necessary in cases where the population converges before the usual time.

Crossover probability

The crossover probability (CR) is also known as the crossover control parameter. It has an impact on the variety of DE because it determines how many components will alter. In addition, its value ranges between $]0,1]$. Greater CR values will result in more variance being introduced into the new population, which will promote exploration. Nevertheless, an appropriate value must be determined to guarantee local and global search capabilities. The value of 0.1 is an excellent starting point. However, a large CR frequently accelerates convergence. Therefore, it is desirable to attempt $CR = 0.9\ or\ CR = 1.0$ first to determine whether a speedy solution is feasible [26, 32].

Variations of Differential Evolution

The DE implementation fundamentally relies on two parts. The first is its trial individual (U) generation strategy (the three operators), and the second is its control factors [35]. Then, there is indeed a rising concern about creating new differential evolution (DE) modifications to improve optimization outcomes. This study examined the execution of the DE standard algorithm (DE_rand1) and its variants (JDE, JADE, and CODE) and addressed the challenges associated with constrained structural optimization problems in the LLE calculations and regression using the NRTL model. There are as follows:

- Rainer Storn et al. proposed the standard differential evolution (DE_rand1) in 1997 [32].
- Janez Brest et al. suggested the self-adaptive control parameter differential evolution (JDE) in 2006 [33].
- Jingqiao Zhang et al. presented the adaptive differential evolution with an optional external archive (JADE) in 2009 [34].
- Yong Wang et al. introduced composite differential evolution (CODE) in 2011 [35].

Table 1 compares DE_rand1, JDE, JADE, and CODE methods based on the structure of the algorithm, the mutation operation, the mutation factor, and the chance of a crossover.

Table 1: A comparison between the DE standard (*DE_rand1*) and its variants (JDE, JADE, and CODE)

Method	***Description***	***Refrence***
DE_rand1	***Algorithm: The standard DE algorithm.*** ***Mutation operation: DE_rand1:*** $V = X_{r_1} + F(X_{r_2} - X_{r_3})$ ***Mutation factor (F) and the crossover probability (CR):*** $CR = [0, 1]$, $F = [0, 2]$	**[32]**
JDE	***Algorithm: Adapting the control parameters*** F_i ***and*** CR_i ***associated with each individual.*** ***Mutation operation : DE_rand1:*** $V = X_{r_1} + F(X_{r_2} - X_{r_3})$ ***Mutation factor (F) and Crossover probability (CR):*** $F_{i,G+1} = \begin{cases} F_l + rand_1 * F_u, & if\ rand_2 < \tau_1 \\ F_{i,G}, & otherwise \end{cases}$ $CR_{i,G+1} = \begin{cases} rand_3, & if\ rand_4 < \tau_2 \\ CR_{i,G}, & otherwise \end{cases}$ ***Where:*** $F_l = 0.1, F_u = 0.9$. $rand_i$ ***represents uniform random values.*** τ_1 ***and*** τ_2 ***indicate probabilities to adjust factors F and CR, respectively.***	**[33]**
JADE	***Algorithm: A novel mutation technique called "DE/current-to-pbest" with an optional external archive and automatically changing control parameters.*** ***The original "DE/current-to-best" is generalized into the DE/current-to-pbest, and the optional archive operation uses past data to identify the direction of advancement. Both strategies increase convergence efficiency and variety in the population.*** ***Mutation operation: DE/current-to-pbest:*** $V = X_i + F_i(X_{best}^p - X_i) + F_i(X_{r_1} - \tilde{X}_{r_2})$ ***Where:*** X_{best}^p ***refers to randomly selected as one of the top 100 p% individuals in the current population with*** $p \in (0, 1]$ ***while*** $\tilde{X}_{r_2}$ ***indicates randomly selected from the union,*** $P \cup A$***, P is as the current population and A is as the set of archived inferior solutions.*** ***Mutation factor (F) and the crossover probability (CR):*** $\mu_{CR} = 0.5, \quad \mu_F = 0.5$ $CR_i = randn_i(\mu_{CR}, 0.1), \quad F_i = randc_i(\mu_F, 0.1)$ $\mu_F = (1 - c) \cdot \mu_F + c \cdot mean_L(S_F), \quad \mu_{CR} = (1 - c) \cdot \mu_{CR} + c \cdot mean_A(S_{CR})$ $mean_L(S_F) = \frac{\sum_{F \in S_F} F^2}{\sum_{F \in S_F} F}$ ***Where: c represents a positive constant between 0 and 1. meanL (·) is the Lehmer mean. meanA(·) is the usual arithmetic mean.*** S_F ***notes the set of all successful mutation factors*** F_i ***in generation g.*** S_{CR} ***indicates the set of all successful crossover probabilities*** CR_i ***at generation g.*** μ_F and μ_{CR} ***are location parameter***	**[34]**
CODE	***Algorithm: This method utilizes the three control parameters sitting and three trial vector-generating operations. It creates trial vectors by combining them at random.*** ***Mutation operation*** ***1- DE_rand1:*** $V = X_{r_1} + F(X_{r_2} - X_{r_3})$***;*** ***2- DE/rand/2:*** $V = X_{r_1} + F(X_{r_2} - X_{r_3}) + F(X_{r_4} - X_{r_5})$***;*** ***3- DE/current-to-rand/1:*** $V = X_i + rand(X_{r_1} - X_i) + F(X_{r_2} - X_{r_3})$. ***Mutation factor (F) and the crossover probability (CR):*** ***1- [F = 1.0, CR = 0.1]*** ***2- [F = 1.0, CR = 0.9]*** ***3- [F = 0.8, CR = 0.2]***	**[35]**

Table 1 presented a comparison between the DE standard and its variants. The modification in the DE algorithm uses adaptive and self-adaptive control factors, and some DE variants employ the optimal vector for generating mutant vectors, whereas others utilize several vectors [63]. Then, the determination of the correct control parameter values can be time-consuming and difficult, especially for certain complex tasks. These DE variants have proven advantageous in enhancing vector positioning, mitigating stagnation issues, and accelerating convergence [63].

The DE standard method (DE_rand1) and its variants (JDE, JADE, and CODE) were applied to find optimal solutions based on the performance of numerical solutions (isoactivity equations) in the calculation and modeling of LLE for the studied system when using the NRTL.

NRTL ESTIMATION PARAMETERS FROM LLE

In order to determine the estimated data and the coefficient activity for both phases ($P1\ and\ P2$) in LLE calculations and modeling, the isoactivity equations and DE methods were used to find the best fit for the objective function. Therefore, the NRTL parameters depend on the type of optimization method and objective function chosen [17, 55].

Basically, the objective functions have two essential strategies including the minimization of activity differences and the minimization of the distances between experimental data and calculated data [55]. Furthermore, the last strategy aims to minimize the sum of squared errors between the experimental and model-predicted values. Then, it has two different types, namely mole fraction objective function (or fractional) and activity objective function [17, 55].

Activity objective function

The activity objective function (OF_a) is a measure used to assess the quality of correlation for various applications of phase equilibrium calculations and modeling. It is expressed as [44, 55]:

$$OF_a = \sum_{i=1}^{3} \sum_{k=1}^{M} \sigma_{ik} \qquad (14)$$

With σ_{ik} denotes the intermediate variable or the residual of an equation and can be expressed as [44]:

$$\sigma_{ik} = \begin{cases} \frac{(x_{ik}^{P1}\ \gamma_{ik}^{P1})}{(x_{ik}^{P2}\ \gamma_{ik}^{P2})} - 1, & if\ \frac{(x_{ik}^{P1}\ \gamma_{ik}^{P1})}{(x_{ik}^{P2}\ \gamma_{ik}^{P2})} \geq 1 \\ \frac{(x_{ik}^{P2}\ \gamma_{ik}^{P2})}{(x_{ik}^{P1}\ \gamma_{ik}^{P1})} - 1, & else \end{cases} \qquad (15)$$

Mole fraction objective function

The minimization of the activity objective function does not ensure a minimization of the distances between the experimental data and the estimated data. This limitation is a severe weakness for any objective function based on the isoactivity equations. The mole fraction objective functions (OF_x) describe the goal and accurately predict the experimental tie lines for phase equilibria in a short amount of time [55], which is written as follows [15, 64]:

$$OF_x = \sum_{i=1}^{3} \sum_{p=1}^{2} \sum_{k=1}^{M} (x_{ikexp}^{p} - x_{ikcal}^{p})^2 \qquad (16)$$

Where: $i,\ p,\ and\ k$ represent the constituent, phase, and the number of tie lines, respectively. The experimental and estimated mole fractions of constituents are $x_{ikexp}^{p}\ and\ x_{ikcal}^{p}$, respectively.

The weakness of OF_x is that the execution costs more time than the first type due to solving isoactivity equations for the calculation of the fraction mole x_{ikcal}^{p} [44].

Root Mean Square Deviation

The Root Mean Square Deviation ($RMSD$) [38] is a statistical measure used in various fields, including data analysis, modeling, and optimization. In this context, RMSD is utilized to examine the quality of the correlation or goodness of fit between two data sets in the studied system. These data represent the experimental and estimated LLE data obtained from the NRTL model [65]. It is also applied to evaluate how well the NRTL model and the DE methods perform in the LLE calculation. It is typically expressed as follows [4, 64]:

$$RMSD = \left[\frac{F}{np.\ k.\ nc}\right]^{1/2} \qquad (17)$$

Where: F represents OF_x and OF_a . np, k, nc denote the phases' number, the number of tie lines, and the elements' number, respectively.

According to the outcomes of this metric, a lower RMSD value indicates a closer match. Therefore, the NRTL model and the DE methods were successfully employed to describe the phase behavior for LLE prediction. On the other hand, a higher RMSD suggests that the regression may need further refinement or limitations in its ability.

Mean and Standard deviation

The mean and standard deviation [66] are statistical methods. In this work, they are used as benchmarks for evaluating the outcomes of LLE calculation and modeling.

Firstly, the mean (m) represents the average magnitude of errors or deviations between the estimated values obtained from the NRTL model and the experimental values. Therefore, it is the average of the RMSD absolute value. It is written as [66]:

$$m = \frac{1}{NR}\sum_{it=1}^{NR} RMSD_{it} \qquad (18)$$

Regarding the evaluation of this measure, a lower mean value indicates that the LLE calculation and modeling have a smaller average error in predicting the desired outcomes. Therefore, there is a good fit between the experimental data and the predicted values of the NRTL model.

Then, the standard deviation (std) is a valuable metric and refers to the square root of the variance between RMSD values as follows [66]:

$$std = \sqrt{\frac{1}{NR-1}\sum_{it=1}^{NR}[RMSD_{it} - m]^2} \qquad (19)$$

Where: NR is the repetition's number, $RMSD_{it}$ indicates the value of RMSD in the it^{th} repetition, and it refers to the iteration rank.

Moreover, the standard deviation was used for assessing the reliability and robustness of the NRTL model performance with the DE methods, especially when analyzing the consistency of predictions. A lower standard deviation value implies that the RMSD values are tightly clustered around the mean, indicating excellent stability and consistency in the modeling approach. In contrast, a higher standard deviation value suggests more significant variability and potentially less stable results.

REGRESSION RESULTS AND DISCUSSION

Timedjeghdine *et al.* [53] studied the LLE experimental aspect of several thermodynamic systems consisting of water, acetic acid, and two kinds of solvent at a temperature of 301.15 K. The first is individual solvents (DCM and MIBK). The second is mixed solvents (25% MIBK with 75% DCM), (50% MIBK with 50% DCM), and (75% MIBK with 25% DCM). Then, the separation factor values obtained by utilizing the mixed solvent (50% MIBK with 50% DCM) are the highest. Consequently, this research selected this quaternary system as the studied system for LLE regression and modeling.

The procedures of minimization are the fractional objective function optimization process (DE methods with OF_x) and the activity objective function optimization process (DE methods with OF_a), which are used this regression. In addition, it aims to estimate the binary interaction parameters obtained from the NRTL model for the quaternary system containing water, acetic acid, and mixed solvent (50% MIBK with 50% DCM) at a temperature of 301.15 K.

Therefore, the efficacy of the DE standard method (DE_rand1) and its variants (JDE, JADE, and CODE) will be tested and assessed in the LLE calculation and modeling utilizing the studied quaternary system. Furthermore, theses algorithms were executed in MATLAB toolbox according to the preceding explanation (Fig. 2, Fig. 3, and Table 1). This execution should determine the parameters of regression to obtain good outcomes.

The outcomes were analyzed after executing the four algorithms, which rely on three parts: dependency on initial population positions, the calculated RMSD, and the execution time. Firstly, this research used the randomization procedure for the initiation of the population because it is simple and carried out in one step.

Effect of regression's parameters

In this section, the impact of significant regression parameters of regression on the value of the activity objective function (fitness) will be checked:

- The size of the population (number of individuals),
- The range of interaction parameters (search space limitation),
- The number of iterations.

Effect of population size

In previous literature, the impact of population size on the objective function value has been examined by optimizing the NRTL model for predicting LLE data. For example, Sahoo et al. [67] found that a population size of 100 particles was enough to get a good set of parameters for the ternary system containing cyclohexane, xylene, and sulfolane when the AG method was used.

Additionally, in 2015, Zheng Li et al. [44] employed the PSO method for LLE modeling with 300 particles. This number of individuals was adequate to produce a suitable set of parameters for the ternary system involving ethene tetrachloro, 2-propanol, and water.

In this study, the DE standard method (DE_rand1) and its variants (JDE, JADE, and CODE) were tested with the effect of population size in LLE calculation, regression, and modeling for the studied system. This work used a population size that ranges between [12, 400]. The exact number should be determined to employ it in the algorithm by examination of its effect on the value of the activity objective function, which is shown in Fig. 5.

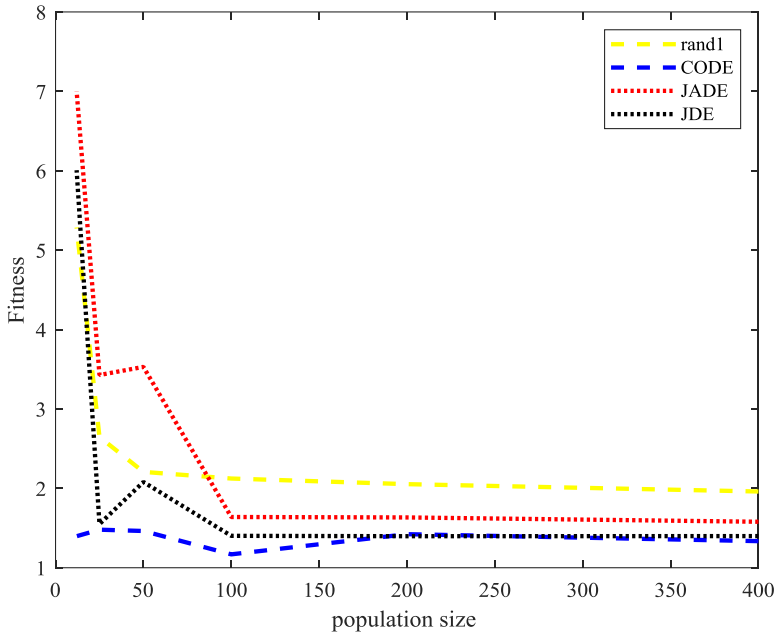


Fig. 5: Effect of population size (individuals' number) on the fitness values

It has been shown that a population size of 200 is adequate to provide the accurate set of parameters for the DE standard method (DE_rand1) and its variants (JDE, JADE, and CODE) since the fitness value becomes roughly steady after a population size of 150. As a result, 200 is the population size value that will be used in DE algorithms to perform calculations.

Effect of range of interaction parameters

In prior literature, the effect of a range of interaction parameters on the objective function value was evaluated in LLE calculation and modeling using the NRTL model. For example, Zheng Li et al. utilized the PSO method in 2015 [44], and the values of the parameters were in the range of [-15, 15], which is necessary for generating better values of RMSD for the studied system.

In this work, the DE standard method (DE_rand1) and its variants (JDE, JADE, and CODE) were evaluated with several intervals of interaction parameters spanning [-10, 10] and [-100, 100] ([lower, upper limits]) of minimum and maximum values to determine the best range.

After minimizing the objective function using DE methods for each range (interval), the effect of interaction parameters' search space size (range or limitation) on the activity objective function is shown in Fig. 6.

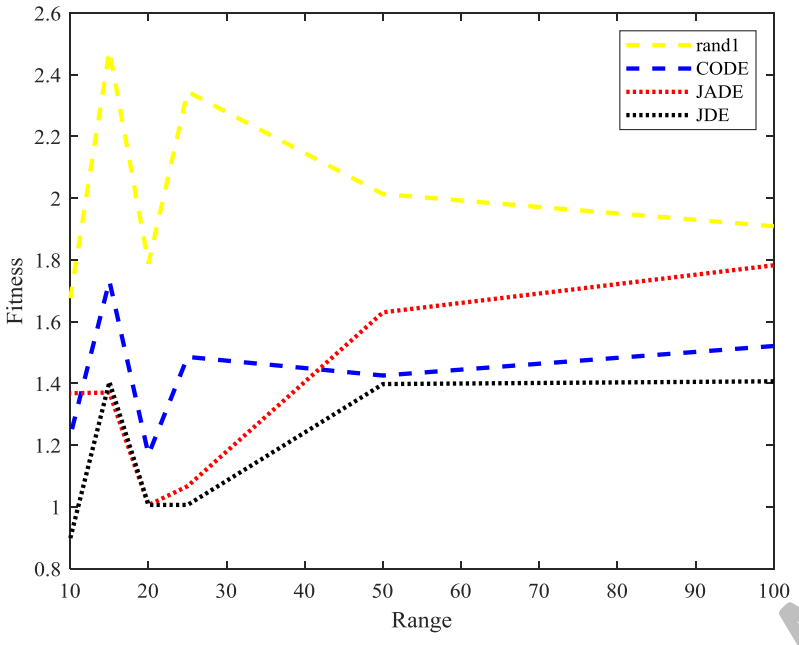


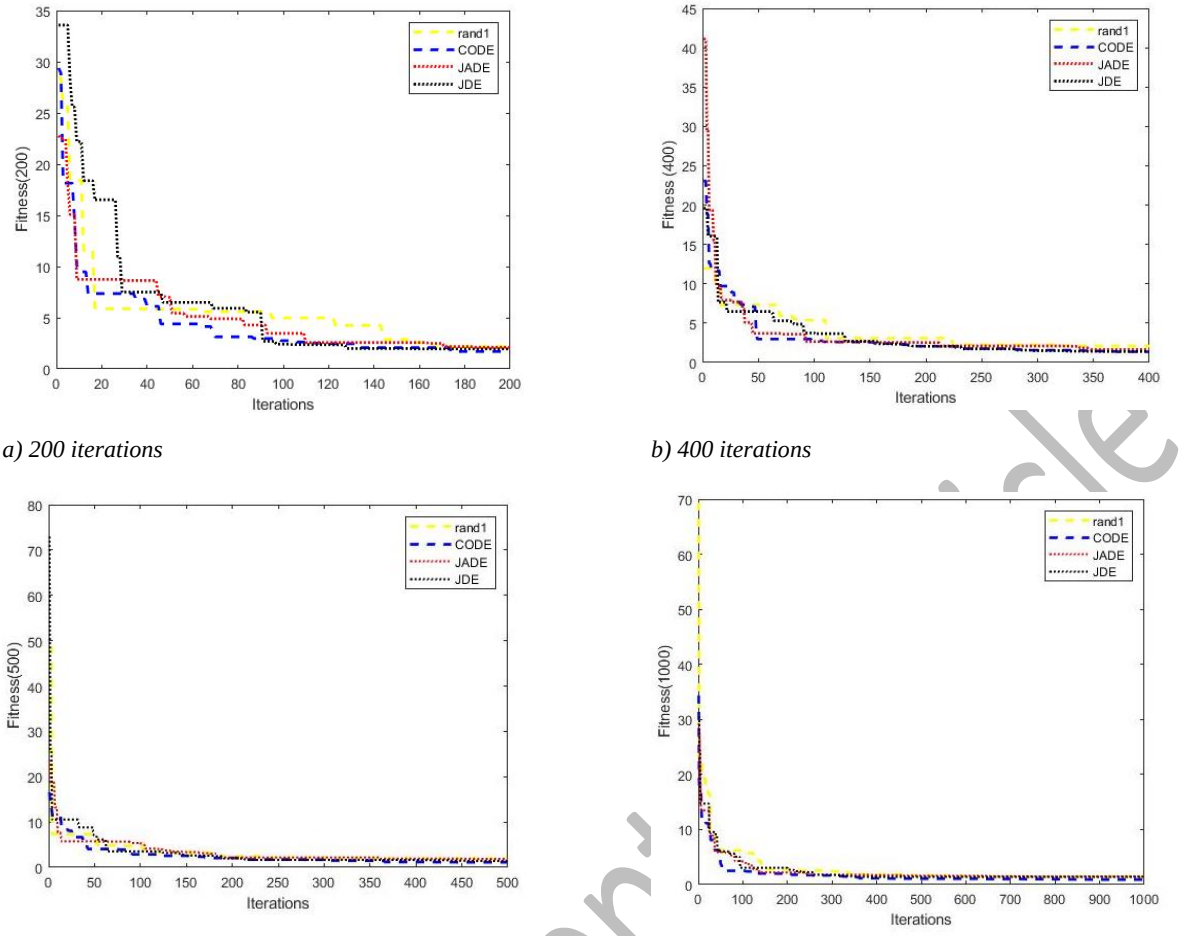
Fig. 6: Effect of parameters' search space size (range or limitation) on the fitness values

According to Fig.6, for example, the value 10 indicates min = -10 and max = + 10. In addition, larger fitness values are in the ranges from [-40, 40] to [-100, 100] than those that exist in the two ranges [-20, 20] and [-30, 30]. However, when compared between the two last ranges, the range [-20, 20] had the lowest fitness values. Regarding Eq. 19, there is a direct relationship between fitness and RMSD. Therefore, their RMSD values were the smallest, which provided the optimal interaction parameters (τ_{ij}). As a result, the algorithm of DE selected this range as an interval of interaction parameters.

Effect of iteration number

In an earlier study, the NRTL model was used to look into how the number of iterations affected the value of the objective function in LLE calculations and models. For instance, Rama Octavian et al. in 2019 [68] employed a fixed iteration number of 250 using the PSO method to model vapor-liquid equilibrium data. This was acceptable for obtaining a suitable RMSD for the water, methanol, and ethanol systems.

In this study, the impact of iterations number 200, 400, 500, and 1000 iterations on the activity objective function is illustrated in Fig. 7 (a), (b), (c), and (d) respectively.



a) 200 iterations

b) 400 iterations

c) 500 iterations

d) 10000 iterations

Fig. 7: Effect of iteration's number ((a) 200, (b) 400, (c) 500, and (d) 1000) on the fitness values

The objective function gets to a minimum point after approximately 180 iterations for 200 iteration. Furthermore, after performing several calculations for 400, 500, and 1000 iterations. It is noticed that the range spanning from 350 to 1000 iterations is adequate to obtain perfect outcomes.

Moreover, for each iteration case for 200, 400, 500, and 1000 iterations, the precise values of the global minimum of the RMSD and their corresponding points were presented in Table 1. In addition, this table demonstrates the impact of the number of iterations on the activity objective function in LLE calculation data and the result of the optimization using the DE standard method (DE_rand1) and its variants.

Table 2: Effect of iteration number in calculation of LLE data for the DE standard method (DE_rand1) and its variants (CODE, JADE, and JDE).

Number of iterations	***DE_rand1***		***CODE***		***JADE***		***JDE***	
	RMSD	***it***	***RMSD***	***it***	***RMSD***	***it***	***RMSD***	***it***
200	***2.1692***	***175***	***1.7054***	***174***	***2.1180***	***184***	***1.9575***	***157***
400	***1.3991***	***317***	***1.3666***	***373***	***1.6412***	***376***	***1.3991***	***393***
500	***1.8057***	***413***	***1.1317***	***444***	***1.8606***	***467***	***1.4145***	***499***
1000	***1.3975***	***943***	***0.9075***	***996***	***1.4653***	***959***	***1.3943***	***998***

According to Table 2, the highest values of the RMSD were for the number of iterations from 157 to 184. On the other hand, the smallest values of the RMSD were within the range of 943 to 998 iterations. In these two cases, this assessment was done according to all the DE methods.

Based on Fig. 7 and Table 2, it is better to set the number of iterations at 1000 to guarantee the attainment of the global minimum (RMSD). As a result, accurate calculations and modeling of LLE will be achieved for the quaternary system containing water, acetic acid, and a mixed solvent (50% MIBK with 50% DCM) at a temperature of 301.15 K.

Random initial particles

This part aims to evaluate the efficacy of the DE standard method and its variants (JDE, JADE, and CADE) and identify the most effective approach for the quaternary system containing water, acetic acid, and mixed solvent (50% MIBK with 50% DCM). Therefore, the exact values of the regression parameters required to execute the DE algorithms according to the results of the previous section are 200 particles ($Np = 200$) for the population size and 1000 iterations.

Furthermore, the non-random solution dispersion parameters (α_{ij}) are fixed at 0.2. Besides, the range of the energy parameters (τ_{ij}) is [-20, 20]. In addition, the randomization procedure was used for the regression of the studied system. Thus, the algorithm was repeated five times to confirm the consistency and reliability of the results.

The phase diagram illustrates the phase behavior of the studied system at equilibrium [69]. Figures 8 and 9 present the phase diagram of water, acetic acid, and a mixed solvent (50% MIBK with 50% DCM) at a temperature of 301.15 K using fractional and activity objective functions, respectively, for the NRTL calculation of the experimental tie-line data.

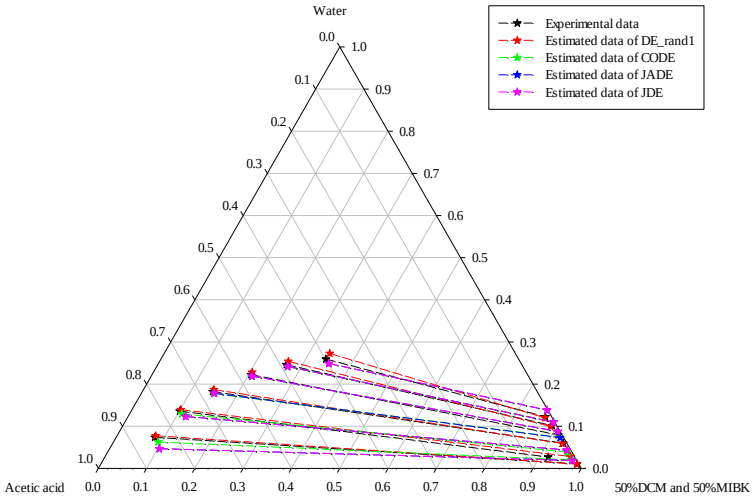


Fig. 8: Comparison of experimental and NRTL correlated Tie- Line data for the studied system when using the fractional objective function optimization process.

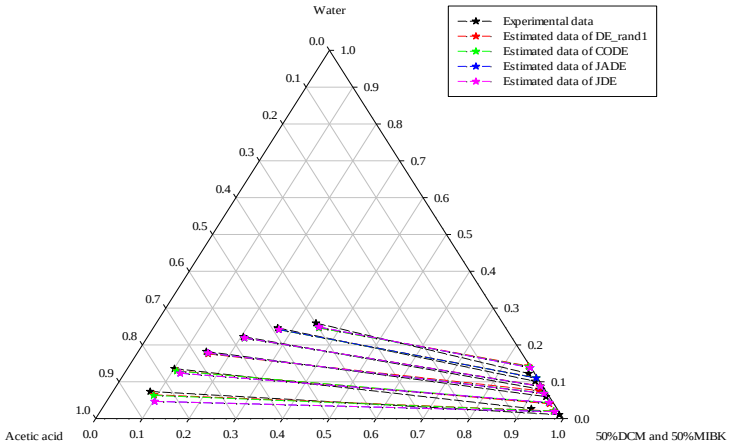


Fig. 9: Comparison of experimental and NRTL correlated Tie- Line data for the studied system when using the activity objective function optimization process.

As a result, the experimental and estimated LLE data exhibited excellent agreement. This alignment between the experimental and calculated data indicates the reliability of the calculations. Additionally, the thermodynamic model parameters [17, 55] have an impact on the computed mole f ractions.

Tables 3 and 4 present the values of the NRTL parameters determined when using the DE standard method and its modifications. These new values in this work differ from those published in Timedjeghdine's study [53] when employing the GA method with minimization by the mole fraction objective function. Also, these parameter values matched the first iteration minimization by the activity objective function and the mole fraction objective function for the quaternary systems that were studied.

Table 3: Calculated interaction parameters for the studied system when using mole fraction objective function minimization with DE standard method (DE_rand1) and its variants (CODE, JADE, and JDE).

DE_rand1		***CODE***		***JADE***		***JDE***	
A_{ij}	A_{ji}	A_{ij}	A_{ji}	A_{ij}	A_{ji}	A_{ij}	A_{ji}
-4.248	*-9.161*	*0.748*	*-9.675*	*1.386*	*-10.274*	*1.383*	*-10.286*
-4.165	*-19.782*	*20*	*-4.284*	*20*	*18.722*	*20*	*18.573*
-12.659	*-19.868*	*13.265*	*-19.89*	*20*	*-20*	*20*	*-20*

Table 4: Calculated interaction parameters for the studied system when using activity objective function minimization with DE standard method (DE_rand1) and its variants (CODE, JADE, and JDE).

DE_rand1		***CODE***		***JADE***		***JDE***	
A_{ij}	A_{ji}	A_{ij}	A_{ji}	A_{ij}	A_{ji}	A_{ij}	A_{ji}
1.0181	*-9.791*	*0.708*	*-9.71*	*1.384*	*-10.275*	*1.383*	*-10.276*
19.926	*-5.254*	*20*	*-4.42*	*20*	*18.573*	*20*	*18.585*
18.416	*-19.706*	*14.144*	*-19.925*	*20*	*-20*	*20*	*-20*

The analysis of RMSD values will evaluate the quality of determined parameters. Therefore, the RMSD values were calculated and listed in Tables 5 and 6 when using the fractional objective function optimization process and the activity objective function optimization process, respectively. Moreover, the time consumed to

execute the four algorithms was taken into account and represented in the precedent tables (Tables 5 and 6). Then, it was measured in minutes.

Table 5: Comparison of regression values and its execution time consumed using the DE standard method (DE_rand1) and its variants (CODE, JADE, and JDE) with the mole fraction objective function minimization.

NR	*DE_rand1*			*CODE*			*JADE*			*JDE*		
	$RMSD_i$	*it*	*Time (min)*	$RMSD_i$	*it*	*Time (min)*	$RMSD_i$	*it*	*Time (min)*	$RMSD_i$	*it*	*Time (min)*
1	*0.0047*	*878*	*460.053*	*0.0090*	*647*	*1226.25*	*0.0107*	*992*	*610.917*	*0.0107*	*999*	*622.35*
2	*0.0047*	*878*	*430.367*	*0.0090*	*964*	*1232.8*	*0.0107*	*945*	*429.1*	*0.0107*	*952*	*429.1*
3	*0.0047*	*878*	*428.1*	*0.0089*	*759*	*1208.1*	*0.0107*	*1000*	*610.167*	*0.0107*	*984*	*444.217*
4	*0.0047*	*878*	*428.283*	*0.0089*	*759*	*1207.23*	*0.0107*	*945*	*424.107*	*0.0107*	*999*	*612.1*
5	*0.0047*	*878*	*429.483*	*0.0089*	*759*	*1209.25*	*0.0107*	*955*	*615.083*	*0.0107*	*991*	*658.15*
m	*0.0047*			*0.0090*			*0.0107*			*0.0107*		
std	*0*			1.1940×10^{-5}			1.1761×10^{-7}			5.3449×10^{-9}		

According to the data shown in Table 5, the results from the five repetitions of the executed algorithms showed that the RMSD values for the DE_rand1 method were constant at 0.0047. On the other hand, the RMSD values for the CODE method ranged from 0.0090 to 0.0089. In addition, a constant value of 0.0107 was determined by using both the JADE and the JDE methods. Hence, it can be demonstrated that the DE_rand1 method presented the lowest RMSD values. Furthermore, the DE_rand1 method provides the highest level of stability, as seen by its minimal variance equal to zero based on the standard deviation calculation (*std*).

Moreover, the values of the consumed time were extremely similar for the five repetitions of the DE standard algorithm and its variants in the domain of [424.1, 1232.8] minutes. In addition, the highest values of the time were for executing the CODE method, which ranged between [1207.23, 1232.8] minutes. Whereas, the smallest values of the time were for executing the DE_rand1 method, which range between [428.1, 460.053] minutes.

As a result, the original method (DE_rand1) outperforms its modifications when comparing the regression values and execution times using the DE methods with the minimization by the mole fraction objective function. This is primarily due to its shorter execution time and superior RMSD results.

Table 6: Comparison of regression values and its execution time consumed using the DE standard method (DE_rand1) and its variants (CODE, JADE, and JDE) with the activity objective function minimization.

NR	*DE_rand1*			*CODE*			*JADE*			*JDE*		
	$RMSD_i$	*it*	*Time (min)*	$RMSD_i$	*it*	*Time (min)*	$RMSD_i$	*it*	*Time (min)*	$RMSD_i$	*it*	*Time (min)*
1	*0.0099*	*915*	*453.2*	*0.0090*	*931*	*1235.98*	*0.0107*	*909*	*422.95*	*0.0107*	*989*	*426.083*
2	*0.0099*	*878*	*430.367*	*0.0090*	*964*	*1232.8*	*0.0107*	*945*	*429.1*	*0.0107*	*952*	*429.1*
3	*0.0099*	*915*	*426.383*	*0.0090*	*971*	*1214.86*	*0.0107*	*909*	*420.9*	*0.0107*	*989*	*448.633*
4	*0.0099*	*878*	*428.283*	*0.0089*	*759*	*1207.23*	*0.0107*	*945*	*424.1*	*0.0107*	*999*	*612.1*
5	*0.0099*	*915*	*432.217*	*0.0090*	*971*	*1203.37*	*0.0107*	*990*	*427.92*	*0.0107*	*989*	*427.717*
m	*0.0099*			*0.0090*			*0.0107*			*0.0107*		
std	*0*			9.8437×10^{-6}			1.0202×10^{-7}			*0*		

According to the data presented in Table 6, the results of the RMSD values for the two methods DE_rand1 and CODE were determined to be 0.0099 and 0.0090, respectively. Meanwhile, the JADE and JDE methods have the same constant value of 0.0107.Regarding the CODE method the RMSD values were found to be the lowest. On the other hand, the DE_rand1 and JDE methods are the most stable when the deviation calculation (std) values are taken into account since their variances are the least, which is zero.

The values of the time consumed for each method are similar for the five repetitions of the DE standard algorithm and its variants. Besides, the range of these values is between [420.9 ,1235.98] minutes. The values of CODE have the greatest values, which are within [1203.37 ,1235.98] minutes. The results of DE_rand1 have the lowest values, which are within [426.383 ,453.2] minutes.

Although the DE_rand1 method produces values of RMSD slightly higher than the results of the CODE method by 0.0009. The stability and execution time were taken into account for the five iterations, the DE_rand1 method was considered the best according to the comparison between the four optimization methods using the activity objective function.

As a result, improvement in the parameters determined in this work provided lower RMSD values (inferior to 10^{-2}) compared to the previous study introduced by Timedjeghdine [53], where its value was higher than 10^{-1} by using the GA method with the fractional objective function. Which confirmed the development and enhancement of the LLE regression parameters using several recent optimization methods (DE, JDE, JADE, and CODE).

CONCLUSION

This study aimed to develop and enhance the regression of LLE by using several recent optimization methods (DE methods). The experimental liquid-liquid equilibrium (LLE) data of a quaternary system containing water, acetic acid, and a mixed solvent (50% dichloromethane (DCM) and 50% methyl isobutyl ketone (MIBK)) at a temperature of 301.15 K were studied.

The calculations were performed using the isoactivity equations to determine the activity coefficients. The NRTL interaction parameters were determined using LLE experimental data regression to enhance the effectiveness of the used model. In addition, this study employed several optimization approaches, such as the DE standard method (DE_rand1) and its variants (JDE, JADE, and CODE) using the randomization process. These approaches used two distinct objective functions, namely the activity objective function and the mole fraction objective function to optimize the NRTL model. The efficiency of the DE method and its variations were evaluated using several regression parameters.

The objective functions are important in influencing the calculation of interaction parameter values. Based on that, three statistical metrics including Root Mean Square Deviation (RMSD), mean (m), and standard deviation (std) were used in order to evaluate the accuracy and reliability of the LLE calculations. Furthermore, the results from the optimisation processes for the two objective functions also show that the DE method and its variations accurately calculated the LLE quaternary mixtures and well described the behavior of this system.

Finally, a comparison was made between the results of the four methods using the two objective functions. The DE standard method (DE_rand1) had better results than its variants (JDE, JADE, and CODE). When the mole

fraction and activity objective functions minimized the RMSD values of the DE_rand1 method, the lowest values were constant at 0.0047 and 0.0099, respectively. In addition, the DE_rand1 approach is the most stable as its variances are the lowest at zero according to the standard deviation calculation. The DE_rand1 has the lowest value of the execution time consumed.

NOMENCLATURES

a	*Chemical activity*
A_{ij}	*Energy interaction between the molecules i and j*
CR	*Crossover probability*
D	**Dimension of the search space**
D_i	*Distribution coefficient*
f	*Fugacity*
F	*Mutation factor*
f_i^{Ri}	*Fugacity of the element i in a real solution*
G_{ij}	*Energy interaction between the molecules i and j*
i	*Element*
$irand\ (D)$	*A function generates random integers between* **1 and** *the dimension of the search space* (**D**)
it	*Iteration rank*
j	*Dimension of the search space*
k	*Tie lines number*
K_i	*Phase equilibrium constant for the element i*
m	*Mean*
M	*Tie line' total number*
n	*Mole number (mol)*
N	*Constituents' number in the studied system*
n_c	*Elements' number*
n_P	*Phases' number*
Np	*Population size*
NR	*Repetition's number*
OF_a	*Objective function in terms of activity*
OF_x	*Objective function in terms of mole fraction*

P	*Pressure (atm)*
$P1, P2, and\ p$	*Phase*
R	*Ideal gas constant (cal K^{-1} mol^{-1})*
$rand$	*A function generates random numbers between* **0 and 1**
$r_1, r_2, r_3, r_4, and\ r_5$	*Random integers*
S	*Separation factor*
T	*Temperature of the mixture (K)*
U	*Trial individual*
V	*Mutant vector*
x	*Mole fraction*
X_{best}	*The best individual vector*
X_i	*Current vector*
x^p_{ikexp}	*Calculated fraction of element i along tie-line k in phase p*
x^p_{ikcal}	*Experimental fraction of elements i along tie-line k in phase p*

Abbreviations

$CODE$	*Composite differential evolution*
DCM	*Dichloromethane*
DE	*Differential evolution*
DE_rand1	*Standard differential evolution*
$JADE$	*Adaptive differential evolution*
JDE	*Self-adaptive control parameters differential evolution*
LLE	*Liquid-liquid equilibrium*
MDE	*Modified differential evolution*
$MIBK$	*Methyl isobutyl ketone*
$NRTL$	*Nonrandom two-liquid*
PSO	*Particle swarm optimization*
$RMSD$	*Root mean square deviation*
$RMSD_{it}$	*RMSD value in the* $\mathbf{it^{th}}$ *Repetition*
SOO	*Single-objective optimization method*
std	*Standard deviation*

Greek letters

α_{ij}	*Non-randomness factor in the mixture*
γ_i	*Activity coefficient of element i*
μ_i	*Chemical potential of element i*
σ_{ik}	*Intermediate variable*
τ_{ij}	*Energy interaction between the molecules i and j*

REFERENCE

[1] Ye M., Chen Z., Jia J., Chen Y., Liquid-liquid equilibrium data for the ternary (nonan-2-one + dihydroxy benzenes + water) systems at 298.15, 313.15, 323.15 and 333.15 K, *J Chem Thermodyn*, **188**:107165 (2024).

[2] Mohadesi M., Aghel B., Gouran A., Liquid-Liquid extraction in a microextractor: a laboratory examination and thermodynamic modeling of N-hexane+ benzene+ sulfolane system, *Iran. J. Chem. Chem. Eng. (IJCCE)*, **40**(2):657–666 (2021).

[3] Regabe S., Merzougui A., Hasseine A., Laiadi D., Bouredji H., Experimental Data and Modeling of Salt Effect on Liquid-Liquid Equilibrium of the Ternary (Water+ 1-Propanol+ Hexane) System at 298K, *Iran. J. Chem. Chem. Eng. (IJCCE)*, **39**(6):199-210 (2020).

[4] Boumaza N., Merzougui A., Nadjar Z., Liquid+ Liquid Equilibria of the Ternary System Water+ Glycerol+ 1-Butanol System: Experimental Data and Modeling, *Iran. J. Chem. Chem. Eng. (IJCCE)*, **41**(1):253- 265 (2022).

[5] Amininia A., Pourshamsian K., Sadeghi B., Nano-ZnO Impregnated on Starch—A Highly Efficient Heterogeneous Bio-Based Catalyst for One-Pot Synthesis of Pyranopyrimidinone and Xanthene Derivatives as Potential Antibacterial Agents, *Russian Journal of Organic Chemistry*, **56**(7): 1279–1288 (2020).

[6] Prateeksha, Rao C. V., Das A. K., Barik S. K., Singh B. N., ZnO/Curcumin Nanocomposites for Enhanced Inhibition of Pseudomonas aeruginosa Virulence via LasR-RhlR Quorum Sensing Systems, *Mol Pharm*, **16**(8):3399–3413 (2019).

[7] Ahmadian E., Kargar Razi M., Sadeghi B., Nakhaei M., Synthesis, characteristics, and photocatalyst effect of nAF and ZnAl2O3/AlF3 nanocomposite with sol-gel method, *International Journal of Nano Dimension*, **14**(4):348–355 (2023).

[8] Ahmadian I., Kargar Razi M., Sadeghi B., Nakhaei M., Sol-Gel Synthesis and Catalytic Properties of PVC/NiAl 2 O 3 /AlF 3 nanocomposite, (2022).

[9] Rahimi M., Sadeghi B., Kargar Razi M., Influence of Al2O3 Additive on Mechanical Properties of Wollastonite Glass-Ceramics, *ADMT Journal*, **14**(3):25–33 (2021).

[10] Dong S., Jia B., Chen X., Jiang X., Liquid–Liquid Equilibrium for Ternary System of Water + 2-Methyl-3-buten-2-ol + (Methyl tert-butyl ether/Butyl acetate/4-Methyl-2-pentanone/2-Ethyl-1-hexanol) at 308.2 K, *J Solution Chem*, **50**(4):544–557 (2021).

[11] Al-Jimaz A. S., Alkhaldi K. H. A. E., Aljasmi A., AlTuwaim M. S., Liquid Extraction of Thiophene from n-Paraffins (C8, C10 or C12) Using Two Ionic Solvents [Mebupy][BF4] or [Emim][CH3SO4], *J Solution Chem*, **52**(8):951–966 (2023).

[12] Pazuki G., Zare Khafri H., Naderifar A., Liquid-Liquid Extraction of Toluene from Heptane Using [EMIM][NTF2] Ionic Liquid: Experimental and Extensive Thermodynamics Study, *Iran. J. Chem. Chem. Eng. (IJCCE)*, **41**(3):913–924 (2022).

[13] Hebboul S., Korichi M., Extraction liquide-liquide : types, facteur influençant, *Algerian Journal of Engineering Architecture and Urbanism*, **5**(5):390–403 (2021).

[14] Yanli Z., Dongguang L., Phase Diagrams for Liquid-Liquid Equilibrium of Neopentyl Glycol+ Sodium Formate+ Water, *Iran. J. Chem. Chem. Eng. (IJCCE)*, **41**(12) (2022).

[15] Ozmen D., Bekri S., Phase diagrams for the aqueous solutions of carboxylic acid with dipropyl ether: Experimental and correlated data, *Iran. J. Chem. Chem. Eng. (IJCCE)*, **39**(6):173–183 (2020).

[16] Mohammadian-Abriz A., Majdan-Cegincara R., Modeling the transport and volumetric properties of solutions containing polymer and electrolyte with new model, *Iran. J. Chem. Chem. Eng. (IJCCE)*, **37**(4):235–252 (2018).

[17] Hebboul S., Korichi M., Hebboul A., Interaction parameters for ternary and quaternary liquid systems using particle swarm optimization, *Algerian Journal of Engineering & Research*, **6**(1):44–54 (2022).

[18] Hebboul S., Korichi M., Hebboul A., Regression of NRTL Parameters from liquid–liquid equilibria for water + ethanol + solvent (dichloromethane, diethyl ether and chloroform) using particle swarm optimization and discussions at T =293.15 K, *Algerian Journal of Engineering & Research*, **5**(2):12–25 (2021).

[19] Abrams D. S., Prausnitz J. M., Statistical thermodynamics of liquid mixtures: a new expression for the excess Gibbs energy of partly or completely miscible systems, *AIChE Journal*, **21**(1):116–128 (1975).

[20] Wilson G. M., Vapor-liquid equilibrium. XI. A new expression for the excess free energy of mixing, *J Am Chem Soc*, **86**(2):127–130 (1964).

[21] Hemond H. F., Fechner E. J., "Chapter 1-Basic Concepts", *in Chemical Fate and Transport in the Environment (fourth edition), Elsevier,* 1–79 (2023).

[22] Bagherinia M. A., Yousefnia S., Measurement and Modeling of Mean Activity Coefficients in Ternary Electrolyte System (Nicl2/Triton X-100/H2O) at T= 298.15±0.1 K, , *Iran. J. Chem. Chem. Eng. (IJCCE)*, **41**(5):1705–1714 (2022).

[23] Renon H., Prausnitz J. M., Local compositions in thermodynamic excess functions for liquid mixtures, *AIChE journal*, **14**(1):135–144 (1968).

[24] Dong S., Sun W., Jiang Y., Jia B., Liquid-liquid equilibrium study for ternary systems of (water + furfuryl alcohol + solvents) at 298.2 K: Measurement and thermodynamic modelling, *J Chem Thermodyn*, **148**:106136 (2020).

[25] Fakhari M. A., Rahimpour F., Measurements and thermodynamic modelling of liquid–liquid equilibrium of PEG 6000–phosphate affinity aqueous two-phase systems at various ligand concentrations and pH, *Phys Chem Liquids*, **58**(4): 483–499 (2020).

[26] Rangaiah G. P., Sharma S., "*Differential Evolution in Chemical Engineering: Developments and Applications*", World Scientific, (2017).

[27] Holland J. H., "Adaptation in natural and artificial systems ", University of Michigan Press, Ann Arbor, MI, USA (1975).

[28] De Jong K., Adaptive System Design: A Genetic Approach, *IEEE Trans Syst Man Cybern*, **10**(9):566–574 (1980).

[29] Eberhart R., Kennedy J., A new optimizer using particle swarm theory, in *MHS'95. Proceedings of the Sixth International Symposium on Micro Machine and Human Science*, IEEE (1995).

[30] Kennedy J., Eberhart R., Particle swarm optimization, in *Proceedings of ICNN'95 - International Conference on Neural Networks*, IEEE (1995).

[31] Hebboul A., Hachouf F., Improving Kernel soft Subspace Clustering Algorithms using a Particle Swarm Optimization, *in Proceedings of the Mediterranean Conference on Pattern Recognition and Artificial Intelligence - MedPRAI-2016, ACM Press, New York, USA* (2016).

[32] Storn R., Price K., Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces, *Journal of global optimization*, **11**(4): 341–359 (1997).

[33] Brest J., Zumer V., Maucec M. S., Self-adaptive differential evolution algorithm in constrained real-parameter optimization, in *2006 IEEE international conference on evolutionary computation*, 215–222 (2006).

[34] Zhang J., Sanderson A. C., JADE: adaptive differential evolution with optional external archive, *IEEE Transactions on evolutionary computation*, **13**(5):945–958 (2009).

[35] Wang Y., Cai Z., Zhang Q., Differential evolution with composite trial vector generation strategies and control parameters, *IEEE transactions on evolutionary computation*, **15**(1):55–66 (2011).

[36] Babu B. V., Angira R., Modified differential evolution (MDE) for optimization of non-linear chemical processes, *Comput Chem Eng*, **30**(6–7):989–1002 (2006).

[37] Özcan H., Comparison Of Particle Swarm And Differential Evolution Optimization Algorithms Considering Various Benchmark Functions, *Journal of Polytechnic*, 899–905 (2017).

[38] Dong S., Sun X., Wang L., Li Y., Zhao W., Xia L., Xiang S., Prediction, Application, and Mechanism Exploration of Liquid–Liquid Equilibrium Data in the Extraction of Aromatics Using Sulfolane, *Processes*, **11**(4):(2023).

[39] Khan A. et al., Adaptive Filtering: Issues, Challenges, and Best-Fit Solutions Using Particle Swarm Optimization Variants, *Sensors*, **23**(18):7710 (2023).

[40] Ab Wahab M. N., Nefti-Meziani S., Atyabi A., A Comprehensive Review of Swarm Optimization Algorithms, *PLoS One*, **10**(5): (2015).

[41] Hsieh F.-S., Development and Comparison of Ten Differential-Evolution and Particle Swarm-Optimization Based Algorithms for Discount-Guaranteed Ridesharing Systems, *Applied Sciences*, **12**(19):9544 (2022).

[42] Greco R., Vanzi I., New few parameters differential evolution algorithm with application to structural identification, *Journal of Traffic and Transportation Engineering (English Edition)*, **6**(1):1–14 (2019).

[43] Bacha O., Hasseine A., Attarakih M., Measurement and correlation of liquid–liquid equilibria for water + ethanol + mixed solvents (dichloromethane or chloroform + diethyl ether) at T = 293.15 K, *Phys Chem Liquids*, **54**(2):245 – 257 (2016).

[44] Li Z., Smith K. H., Mumford K. A., Wang Y., Stevens G. W., Regression of NRTL parameters from ternary liquid–liquid equilibria using particle swarm optimization and discussions, *Fluid Phase Equilib*, **398**:36–45 (2015).

[45] Timedjeghdine M., Hasseine A., Binous H., Bacha O., Attarakih M., Liquid–liquid equilibrium data for water + formic acid + solvent (butyl acetate, ethyl acetate, and isoamyl alcohol) at T = 291.15 K, *Fluid Phase Equilib*, **415** (2016).

[46] Alam S., Zhao X., Niazi I. K., Ayub M. S., Khan M. A., A comparative analysis of global optimization algorithms for surface electromyographic signal onset detection, *Decision Analytics Journal*, **8**:100294 (2023).

[47] Wisittipanich W., Phoungthong K., Srisuwannapa C., Baisukhan A., Wisittipanit N., Performance Comparison between Particle Swarm Optimization and Differential Evolution Algorithms for Postman Delivery Routing Problem, *Applied Sciences*, **11**(6):2703 (2021).

[48] Iwan M., Akmeliawati R., Faisal T., Al-Assadi H. M. A. A., Performance Comparison of Differential Evolution and Particle Swarm Optimization in Constrained Optimization, *Procedia Eng*, **41**:1323–1328 (2012).

[49] Zhang D., Wei B., Comparison between differential evolution and particle swarm optimization algorithms, *in 2014 IEEE International Conference on Mechatronics and Automation, IEEE*, 239–244 (2014).

[50] Vesterstrom J., Thomsen R., A comparative study of differential evolution, particle swarm optimization, and evolutionary algorithms on numerical benchmark problems, *in Proceedings of the 2004 Congress on Evolutionary Computation (IEEE Cat. No.04TH8753), IEEE*, 1980–1987 (2004).

[51] Tharwat A., Schenck W., Population initialization techniques for evolutionary algorithms for single-objective constrained optimization problems: Deterministic vs. stochastic techniques, *Swarm Evol Comput*, **67**:100952 (2021).

[52] Soleimani Lashkenari M., Mehdizadeh B., Movagharnejad K., Application of genetic algorithm based support vector machine model in second virial coefficient prediction of pure compounds, *Iran. J. Chem. Chem. Eng. (IJCCE)*, **37**(5) (2018).

[53] Timedjeghdine M., "Experimental study and modeling of systems (water + carboxylic acid + solvents) ", Doctoral dissertation, Department of Industrial Chemistry, University of Mohamed Khider, Biskra, Algeria (2016).

[54] Hong G.-B., Lee M.-J., Lin H., Liquid–liquid equilibria of ternary mixtures of water + 2-propanol with ethyl acetate, isopropyl acetate, or ethyl caproate, *Fluid Phase Equilib*, **202**(2):239–252 (2002).

[55] Sorensen J. M., Magnussen T., Rasmussen P., Fredenslund A., Liquid-liquid equilibrium data: Their retrieval, correlation and prediction Part II: Correlation, *Fluid Phase Equilib*, **3**(1):47–82 (1979).

[56] Prausnitz J. M., Lichtenthaler R. N., De Azevedo E. G., "*Molecular thermodynamics of fluid-phase equilibria*", Pearson Education (1998).

[57] Malakhov D. V., A linkage between the activity and the chemical potential, *Canadian Metallurgical Quarterly*, **59**(1):36-40 (2020).

[58] Li Z., Mumford K. A., Shang Y., Smith K. H., Chen J., Wang Y., Stevens G. W., Analysis of the nonrandom two-liquid model for prediction of liquid–liquid equilibria, *J Chem Eng Data*, **59**(8):2485–2489 (2014).

[59] Tsuboka T., Katayama T., Modified Wilson equation for vapor-liquid and liquid-liquid equilibria, *Journal of chemical Engineering of Japan*, **8**(3):181–187 (1975).

[60] Fan W., Han Y., Yan H., Xu D., Huang H., Wang Y., Liquid-liquid equilibrium data for ternary mixtures (water + isopropanol + 1-pentanol/1-hexanol/1-heptanol) at 298.15 K: Measurement, correlation and prediction, *J Chem Thermodyn*, **174**:106871 (2022).

[61] Labarta J. A., Olaya M. M., Marcilla A. F., What does the NRTL model look like? Determination of boundaries for different fluid phase equilibrium regions, *AIChE Journal*, **68**(10) (2022).

[62] Gebreyohannes S., Neely B. J., Gasem K. A. M., One-parameter modified nonrandom two-liquid (NRTL) activity coefficient model, *Fluid Phase Equilib*, **379** (2014).

[63] Price K. V., Storn R. M., Lampinen J. A., "Differential Evolution: A practical approach to global optimization", Springer-Verlag, Berlin Heidelberg (2005).

[64] Yu Y., Wei J., Zhao D., Yi L., Han M., Cao J., Zhang F., Zhang X., Liquid–Liquid Equilibrium Measurement and Thermodynamic Modeling for Ternary Systems of 2-Ethyl-1-hexanol + 1,2-Propanediol + (Water and *n* -Heptane) at Different Temperatures," *J Chem Eng Data*, **68**(7):1695–1705 (2023).

[65] Yu Y., Yi L., Wei J., Guo H., Li M., Liquid–Liquid Equilibrium for Ternary Systems of Ethylene Glycol, 2-Ethyl-1-Hexanol and Different Extractants at 298.2 K and 101.3 kPa, *J Solution Chem*, **52**(1):91–104 (2023).

[66] Lee D. K., In J., Lee S., Standard deviation and standard error of the mean, *Korean J Anesthesiol*, **68**(3):220 (2015).

[67] Sahoo R. K., Banerjee T., Ahmad S. A., Khanna A., Improved binary parameters using GA for multi-component aromatic extraction: NRTL model without and with closure equations, *Fluid Phase Equilib*, **239**(1):107–119 (2006).

[68] Oktavian R., Wibowo A. A., Fitriah Z., Study on Particle Swarm Optimization Variant and Simulated Annealing in Vapor Liquid Equilibrium Calculation, *Reaktor*, **19**(2):77–83 (2019).

[69] Gharesheikhloo A. A., Moayyedi M., Experimental and Simulation Constructing of Phase Diagram and Quality Lines for a Lean Gas Condensate Reservoir Fluid, *Iran. J. Chem. Chem. Eng. (IJCCE)*, **33**(3):83–88 (2014).