Application of the Genetic Algorithm to Calculate the Interaction Parameters for Multiphase and Multicomponent Systems

Rashtchian, Davood*[+]; Ovaysi, Saeed; Taghikhani, Vahid; Ghotbi, Cirous

Department of Chemical and Petroleum Engineering, Sharif University of Technology, Tehran, I.R. IRAN

ABSTRACT: *A method based on the Genetic Algorithm (GA) was developed to study the phase behavior of multicomponent and multiphase systems. Upon application of the GA to the thermodynamic models which are commonly used to study the VLE, VLLE and LLE phase equilibria, the physically meaningful values for the Binary Interaction Parameters (BIP) of the models were obtained. Using the method proposed in this work the activity coefficients for components at infinite dilution, obtained from the local composition based models, can be accurately predicted comparing to the experimental data available in the literature. In this work, a Global Optimization Procedure (GOP) based on the GA was developed to obtain the binary interaction parameters of the Wilson, NRTL and the UNIQUAC models for a number of systems at various temperatures. The VLE, VLLE and LLE values of the binary interaction parameters for the activity coefficients models were compared with those reported in the literature for the systems studied in this work. The results showed that the values reported for the binary interaction parameters can predict the activity coefficients at infinite dilutions for the components in the VLE, VLLE and LLE systems studied in this work. In order to confirm the accuracy of the results, the values for the activity coefficients at infinite dilution in binary solutions were compared with those reported in the literature. The comparison showed that the models studied using the proposed method can predict the physically meaningful binary interaction parameters among the species present in solutions. In addition to the accuracy, simplicity, generality and short CPU computation time of the proposed method should be mentioned as some of its clearest advantages.*

KEY WORDS: *VLE, LLE, VLLE, Multi-component phase behavior, Genetic algorithm.*

INTRODUCTION

Application of the activity coefficient models in order to study the phase behavior of the multiphase and multicomponent systems at low pressure have been commonly used in engineering practice. For instance, modeling of the Liquid-Liquid Equilibria (LLE) systems using the activity coefficient models is important in

simulation of extraction processes. Since the equations for modeling of such systems are nonlinear, usually the estimation of parameters for these systems leads to a non-convex optimization problem. Thus, the multiple optima and the multiple root problems are the main pitfall in the calculations.

Many different optimization techniques have been proposed to estimate the adjustable parameters for the activity coefficient models [1-9] used to study the phase behavior of VLE and LLE systems. But most of them suffer from solution multiplicity and existence of multiple roots for BIP's when applied to correlate the experimental data.

The techniques that use the gradient-based methods, such as *Gauss-Newton* [1], *Gauss-Marquardt* [2], and sequential quadratic programming (SQP) methods [3], are conventional optimization techniques in the phase equilibrium calculations. Since the evaluation of gradients is one of the erroneous parts of these techniques, the simplex pattern search methods have been used to obtain the best values for the BIP of the activity coefficient models [4]. However, these methods cannot provide a theoretical guarantee of global optimality.

In the recent years, new optimization techniques have been proposed based on the global optimization methods [5-9]. Among them the interval [5-7], branch and bound methods [8] can be mentioned. Due to the rounding-errors produced in the computations using these methods, despite the strong theoretical basis, the results obtained are not definitely globally optimum results. Another shortcoming is that because of the need to the gradient evaluations, these methods have a higher error potential to produce.

It should be noted that there are a few techniques that use the gradient less random search (RS) methods as the optimization engines [9]. These techniques are rarely used to study the phase behavior of VLE, VLLE and LLE systems.

In this work a new technique that is based on the genetic algorithm, was developed to study the phase behavior for a number of VLE, VLLE and LLE systems at various temperatures. The results for the activity coefficients for components at infinite dilution obtained in this work were compared with those reported in the literature.

GA OVERVIEW

GA is an evolutionary algorithm that was first proposed by *Holland* [10] and widely used in the recent decades. GA's have been also used to solve a large variety of the problems in chemical engineering. Among them the multi-objective optimization of the Fluidized bed Catalytic Cracking Reactors (FCCR) can be mentioned [11]. The GA has been also used in optimization of polymerization reactors [12]. *Geyer et al.,* used an evolutionary algorithm based on the GA for calculation of the group contribution problem in order to predict the thermodynamic properties for a number of systems [13]. It is a global optimization method that searches the solution space of a function through the simulation of the genetic process of biological organisms. Since the GA works with the initial population which is generated randomly. Despite most of the optimization methods, this method does not require the initial guess to the optimum point. Because a wide variety of genetic operators have been used in the GA, there are many versions of the GA which have been developed in solving the specific problems. The GA is a direct optimization technique and has been used to solve difficult problems, even with objective functions that are not differentiable. This algorithm does not guarantee the finding of the global optima, but it is generally good to find "reasonably acceptable" solutions to problems. A good overview of GA has been presented by *Beasley* [14, 15].

Basic Principles of GA

The first stage in the standard GA is the generation of initial population. In the next stages, this population evolves in a cyclic process that consists of selection of parents and recombination of them to produce new offspring. The GA operators, mainly Crossover and Mutation, have the responsibility for producing new generations. Before running the GA, we require a suitable representation (coding) and a fitness function which assigns a figure of merit to each coded solution.

For any GA, a chromosome or individual representation is required. This representation determines the structure of GA and its operators. Each chromosome consists of a number of genes from a certain alphabet. An alphabet could be binary, floating point, matrices, etc. In the first GA (*Holland's* design), binary alphabet was used. *But Michalewicz* [16] showed that the real-valued

coding moves the problem closer to the problem representation which offers higher precision with more consistent results across replications. He also shows that the real-valued GA is an order of magnitude more efficient in terms of CPU time [16].

The selection of individuals to produce successive generations plays an extremely important role in GA. This is a probabilistic process that the better individuals have the higher probability for selection. The most common selection mechanisms are:
Roulette wheel selection and its extensions, Scaling techniques, Tournament, Elitist models, and Ranking methods. Roulette wheel selection, was the first selection mechanism, which was developed by *Holland*. However, a good selection method must prevent filling the next population by the fittest individuals of the current population, and consequently prevent the premature convergence difficulty. These points have been considered in the scaling and ranking methods.

Genetic operators are the basic search mechanisms for GA. They are used for creating new population (solutions) from the selected population. There are two basic types of these operators: crossover and mutation. The responsibility of crossover is the recombination of two current individuals to produce new individuals. These new individuals may be mutated, usually by a small perturbation, for more exploration of the search space. As mentioned above, the application of genetic operators depends on the coding scheme that has been used. For example in the binary coding scheme, single point crossover, two point crossover, or uniform crossover can be chosen. Whereas the most common crossover mechanisms in the floating point coding scheme are: simple crossover, arithmetic crossover, and heuristic crossover. Also there are several mutation mechanisms which have been used in the floating point coding scheme, such as: uniform mutation, non-uniform mutation, and boundary mutation.

Fitness function is independent of GA (e.g. stochastic decision rules). Usually it is an objective function that returns a numeric value according to the figure of merit of each individual. But in some cases, such as combinatorial optimization, the fitness function is not just a simple objective function.

The reproductive cycle of GA continues until a termination criterion is met. The most frequently used termination criterion is a specified maximum number of generations. Another termination criterion is the convergence of the population to a specific individual.

PARAMETER ESTIMATION

In application of thermodynamic models to studying of the phase equilibrium of multicomponent and multiphase systems, it would be quite possible to obtain different sets of adjustable parameters for the studied models. Therefore, it seems that existence of a criterion to judge about the values for the parameters is a must. As a matter of fact, this criterion should compare the results obtained from the model with those reported in the literature as experimental data. In order to obtain the best values for the parameters of the models studied in this work, two following objective functions were minimized using the experimental data. The objective functions are as follows:

$$f(a) = \sqrt{\sum_{i=1}^{N \exp} \sum_{j=1}^{N} \left(\frac{Y_{exp,ij} - Y_{calc,ij}}{Y_{exp,ij}} \right)^2} \qquad (1)$$

or

$$f(a) = \sqrt{\sum_{i=1}^{N \exp} \sum_{j=1}^{N} \left(Y_{exp,ij} - Y_{calc,ij} \right)^2} \qquad (2)$$

where in equation (1), Y stands for the activity coefficients of components in mixtures, notably in equation (2), Y denotes the mole fractions of components. The subscripts exp. and calc. stand for the experimental and calculated values, respectively. Also in the above equations i, j, N and N_{exp} refer to the components, tie line indices, number of components and number of experimental data respectively.

Simple Procedure

The following well-known isoactivity criterion was used to study the phase equilibria for the various vapor-liquid systems at low pressures as:

$$x_{ij}\gamma_{ij}p_j^{sat} = y_j p \qquad (3)$$

with $i = 1,2,...,N_{LP}$ and $j = 1,2,...,N_C$

where x_{ij} and y_j are the mole fractions of the jth component in the ith liquid phase and the jth component in vapor phase respectively. Also γ_{ij} is the activity
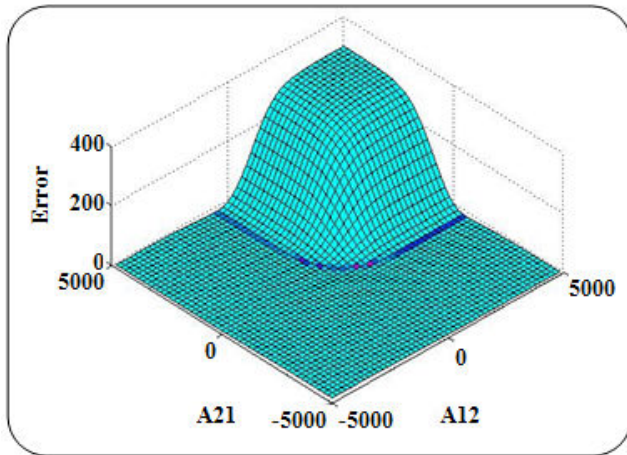
Fig. 1: Form of error function in the VLE system of tert-butanol (1) and 1-butanol (2) at 100 mmHg. Liquid phase is modeled by Wilson activity coefficient model and the vapor phase is ideal. Parameters in Cal/mole.
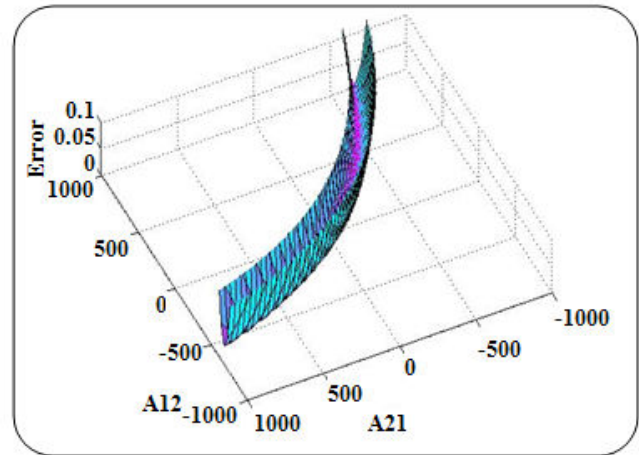

Fig. 2: The valley that the global and local minimum points exist in the error function of the VLE system of tert-butanol (1) and 1-butanol (2) at 100 mmHg. Liquid phase is modeled by Wilson activity coefficient model and the vapor phase is ideal. Parameters in Cal/mole.

coefficient of the jth component in the ith liquid phase which would be determined at the system temperature. Also p_j^{sat} is the vapor pressure of jth pure component at system temperature. According to equation (3) the experimental values for the activity coefficients can be determined as follows:

$$\gamma_{ij,exp} = \frac{y_{j,exp}p}{x_{ij,exp}p_j^{sat}} \qquad (4)$$

where subscript exp stands for the experimental values of γ_{ij}, x_{ij} and y_j. It should be noted that the calculated values of the activity coefficients should be determined at the same conditions as the experimental values reported. In this case the most commonly used form of the objective function is defined as follows:

$$f_1(a) = \sqrt{\sum_{i=1}^{N_{exp}} \sum_{j=1}^{N_{LP}} \sum_{k=1}^{N_C} \left(\frac{\gamma_{ijk,exp} - \gamma_{ijk,calc}}{\gamma_{ijk,exp}} \right)^2} \qquad (5)$$

where $\gamma_{ijk,exp}$ and $\gamma_{ijk,calc}$ are, respectively, the experimental and calculated values of the activity coefficient of kth component in the jth liquid phase at the conditions of ith experiment. The global minimum of this objective function (Error) would give the best sets of adjustable parameters for thermodynamic models. The complexity of the problem and existence of multiple optima is illustrated in the Figs. 1 and 2. In these figures the

error functions have been plotted versus the variations in the binary interaction parameters of the Wilson activity coefficient model for the VLE system of tert-butanol (1) and 1-butanol (2) at 100 mmHg. The heavy strip in Fig. 1 is the area where the minimum points are located. This strip has been zoomed in Fig. 2. As the color gradually shifts from cyan to red the value of the error function decreases. Therefore, the point with the heaviest red color is the global minimum point. It is clear that even for these two variable problems, there are many local minimum points, therefore, it is very important to use a global optimization algorithm to find the correct answer.

Rigorous Procedure

Generally, if the activity and fugacity coefficients cannot be determined directly from the experimental data, the simple procedure is not worthy of use. Parameter estimation in the systems such as VLE, VLLE at high pressure, where the ideal gas law is not applicable any longer, are examples of this case. In such cases the equilibrium conditions take the following forms:

$$x_{ij}\phi_{ij}^L = y_j\phi_j^v \qquad (6)$$

with $i=1,2,\ldots,N_{LP}$ and $j=1,2,\ldots,N_c$

where, ϕ_j^V and ϕ_{ij}^L are, respectively, the fugacity coefficients of jth component in the vapor phase and in the

i^{th} liquid phase. It is clear that the experimental values of the fugacity coefficients cannot be determined directly. Such a difficulty also exists when two or more liquid phases are in equilibrium without existence of vapor phase. It should be noted that the Liquid-Liquid Equilibrium (LLE) systems are common examples for this case. In such conditions the equilibrium criteria can be written as:

$$x_{1j}\gamma_{1j} = x_{2j}\gamma_{2j} \qquad (7)$$

Parameter estimation in all the above cases needs the rigorous procedure. The component mole fractions should be selected as decision variables. The applicable objective functions for the systems that belong to the rigorous procedure can be written as:

$$f_2(a) = \sqrt{\sum_{i=1}^{N_{exp}}\sum_{j=1}^{N_P}\sum_{k=1}^{N_C}\left(y_{ijk,exp} - y_{ijk,calc}\right)^2} \qquad (8)$$

where, $y_{ijk,exp}$ and $y_{ijk,calc}$ are the experimental and calculated values of mole fraction of kth component in j^{th} phase at the conditions of ith experiment. The best sets of model parameters can be obtained from the global minimum of the objective function. Because of the existence of multiple roots, the objective function in the form of equation (8) can further complicate the parameter estimation procedure than that in equation (5).

Implementation of the Proposed Method

Parameter estimation can be done as an unconstrained (simple procedure) or a constrained (rigorous procedure) optimization problem. These can be formulated as forms of equations (9) and (10), respectively. Since the GA can only handle the maximization problems, therefore, in case of minimizing the objective function by the GA it is necessary to maximize the following objective function as:

$$\max : F = -f_1(a) \qquad (9)$$

$$\max : F = -f_2(a)$$
$$\text{subject to} : \text{Flash Calculations} \qquad (10)$$

To increase the efficiency and accuracy of the optimization process, the best individuals of each generation in the GA is considered as the starting point of the Nelder-Mead (NM) method and will progress in a short period optimization by the NM.

Floating point encoding, ranked selection, arithmetic crossover and nonuniform mutation were used to minimize the objective function by the GA. The crossover and mutation operators can be written as the following equations respectively:

$$Child_{i1} = r \times Parent_{i1} + (1-r) \times Parent_{i2} \qquad (11)$$
$$Child_{i2} = (1-r) \times Parent_{i1} + r \times Parent_{i2}$$
$$\text{and} \quad 1 < i < \text{Initial Population}$$

it should be noted that:

$$\text{if } r < 0.5 \Rightarrow Child_{kj} = Child_{kj} - r \times (Child_{kj} - Min_k) \qquad (12)$$
$$\text{if } r > 0.5 \Rightarrow Child_{kj} = Child_{kj} + r \times (Max_k - Child_{kj})$$
$$\text{and} \quad 1 < k < \text{Initial Population}$$
$$1 < j < \text{Number of Variables}$$

where r is a random number that varies between 0 and 1, Child and Parent represent the child and parent chromosomes, respectively. First and second indices of child and parent chromosomes are the variable and chromosome numbers, respectively. Also k is a random number that represents the index of a specified variable for mutation.

In equation (8) the experimental values of component mole fractions at each phase are available, but their corresponding theoretical values cannot be directly calculated and need the conventional flash calculations. In this work for the problems that invoke the rigorous procedure, our suggested two-tier algorithm that alternately solves the flash calculations and parameter estimation problems was used. The flash calculations provide the theoretical values of the component mole fractions at each phase. For such problems, each function evaluation leads to a number of flash calculations. Thus, a reliable and quick flash algorithm is needed to have reasonable results in terms of accuracy and computation time. The flash algorithm proposed by *Lucia et al.*, [17] is an appropriate algorithm and used in this work. However, since the flash problem is a constrained optimization with linear constraints and bounds on the variables, it is more appropriate to use Generalized Reduced Gradient (GRG) method for such a problem. Thus, the GRG method instead of the SQP was used in the modified version of the flash calculation algorithm. As a result of this modification, the same degree of reliability can be

reached in flash calculations with less iteration in comparison with the method proposed by *Lucia et al.*

To determine the upper and lower bound of the binary interaction parameters, the values of activity coefficients at infinite dilution could be useful. According to *Stadtherr et al.*, [7] these values are in the interval [0.03, 109000]. These values were used with a conservative view to determine the boundary values of the binary interaction parameters.

The values of the initial population can be determined through a trial and error procedure. It should be noted that the reported values correspond to 95% efficiency in the result which means 95% of the different runs led to the same result.

RESULTS AND DISCUSSIONS

Six different equilibrium systems have been used to study the performance of the new algorithm developed in this work. These systems are categorized into two groups. The first group consists of the three VLE systems at low pressures. To study these systems, simple procedure was used to minimize the corresponding objective function in the form of equation (5). The second group consists of one VLLE system at high pressures and four LLE systems. For estimating the parameters of the thermodynamic models that have been used for modeling such systems, it is necessary to use the optimization form as equation (10).

The results of the present work have been also assessed by comparing it with the most widely used methods in this field. These methods have been discussed in detail elsewhere [7].

In reference [7] the interval method has been used. Interval method is essentially a global optimization method that searches the whole area of the possible solution by dividing it to several intervals and finally finding the most optimum maximum or minimum solution. As expounded earlier despite its strong theoretical background, the interval method practically suffers from the rounding error problem and, thus, there is no guarantee to find the global optimum point. However for the two-variable example to be discussed later there is an opportunity to see the global minimum point in advance as shown in Figs. 1 and 2. Therefore, the global optimality of the results can be obtained. Another commonly used method in the area of determining the

binary interaction parameters is the Nelder-Mead method which has been widely used [20-23]. The Nelder-Mead method is a direct optimization method which suffers from the need to an initial guess and uses a combination of interpolation and extrapolation functions to locate an optimum point. There is neither a theoretical nor a practical proof for this method to find the global optimum point. Unfortunately, this method has traditionally been used to determine the binary interaction parameters for the variety of systems (probably because of unavailability of a global direct optimization method). Hence, most of the interaction parameters reported in the literature match the local minimum points of the error function.

The results of parameter estimation for the NRTL, UNIQUAC and the Wilson models obtained from study of the VLE phase behavior of 2-butanone and hexadecane at 60ºC and of 2-butanone and octadecane at 80 ºC have been presented in the tables 1 and 2, respectively. Both systems belong to the simple procedure of parameter estimation. As can be seen from tables 1 and 2, the physically meaningful results for the activity coefficients at infinite dilution were obtained and the results are in better consistency with the experimental data, than those reported in the literature [18]. The Wilson, NRTL and the UNIQUAC models are respectively given through equations (12) to (15) as follows:

$$\ln \gamma_i = -\ln(x_i + \Lambda_{ij}x_j) + \tag{12}$$

$$x_j \left[\frac{\Lambda_{ij}}{x_i + \Lambda_{ij}x_j} - \frac{\Lambda_{ji}}{x_j + \Lambda_{ji}x_i} \right]$$

$$\text{with} \quad \Lambda_{ij} = \frac{v_j}{v_i} \exp\left[\frac{-\theta_i}{RT} \right]$$

In equation (12) which expresses the Wilson activity coefficient model, v_i and v_j are the pure liquid molar volumes for components i and j respectively. T is the absolute temperature of the system and θ_i is the energy parameters that need to be estimated.

The NRTL activity coefficient model takes the following form:

$$\ln \gamma_i = \frac{\displaystyle\sum_{j=1}^{Nc} \tau_{ji} x_j G_{ji}}{\displaystyle\sum_{k=1}^{Nc} x_k G_{ki}} \tag{13}$$

Table 1: The results of parameter estimation for the NRTL, UNIQUAC and the Wilson models obtained using the VLE of 2-butanone and Hexadecane at 60 ºC. Experimental data are taken from [18]

with 14 equilibrium data points and $\gamma^{\infty}_{1,Experimental} = 2.79$ ***[19].***

Present Work Initial population=10, Maximum number of iterations=20, CPU time=0.004 s					
	A_{12} (Cal/mol)	A_{21} (Cal/mol)	$\alpha_{12}=\alpha_{21}$	$\gamma^{\infty}_{1,Calculated}$	f_1
NRTL	918.58	321.259	0.66934	2.81	0.01476
UNIQUAC	-161.134	506.084	--	2.49	0.02927
Wilson	904.038	252.682	--	2.66	0.01552
Ref. [18]					
	A_{12} (Cal/mol)	A_{21} (Cal/mol)	$\alpha_{12}=\alpha_{21}$	$\gamma^{\infty}_{1,Calculated}$	f_1
NRTL	919.171	339.288	0.6870	2.85	0.01553
UNIQUAC	-153.126	492.080	--	2.45	0.03113
Wilson	899.317	260.192	--	2.64	0.01571

Table 2: The results of parameter estimation for the NRTL, UNIQUAC and the Wilson models obtained using the VLE of 2-butanone and Octadecane at 80 ºC. Experimental data are taken from [18]

with 9 equilibrium data point and $\gamma^{\infty}_{1,Experimental} = 2.2$ ***[19].***

Present work Initial population=20, Maximum number of iterations=30, CPU=0.01 s					
	A_{12} (Cal/mol)	A_{21} (Cal/mol)	$\alpha_{12}=\alpha_{21}$	$\gamma^{\infty}_{1,Calculated}$	f_1
NRTL	1202.21	-166.1	0.41	1.84	0.01292
UNIQUAC	-139.172	447.676	--	1.88	0.01301
Wilson	892.04	187.881	--	2.16	0.01080
Ref. [18]					
	A_{12} (Cal/mol)	A_{21} (Cal/mol)	$\alpha_{12}=\alpha_{21}$	$\gamma^{\infty}_{1,Calculated}$	f_1
NRTL	1500.313	-416.732	0.2923	1.73	0.01333
UNIQUAC	-134.874	439.714	--	1.86	0.01307
Wilson	882.183	197.264	--	2.14	0.01083

$$\sum_{j=1}^{Nc}\frac{x_j G_{ij}}{\sum\limits_{k=1}^{Nc}x_k G_{kj}}\left(\tau_{ij}-\frac{\sum\limits_{m=1}^{Nc}\tau_{mj}x_m G_{mj}}{\sum\limits_{k=1}^{Nc}x_k G_{kj}}\right)$$

with $\quad G_{ij}=\exp\left(-\tau_{ij}\alpha_{ij}\right)$

and $\quad \tau_{ij}=\dfrac{\Delta g_{ij}}{RT}$

where α_{ij} is the nonrandomness factor and Δg_{ij} are the energy parameters that again need to be estimated.

The UNIQUAC activity coefficient model can be expressed as:

$$\ln\gamma_i=\ln\left(\frac{\Phi_i}{x_i}\right)+\frac{Z}{2}q_i\ln\left(\frac{\theta_i}{\Phi_i}\right)+L_i- \qquad (14)$$

$$\left(\frac{\Phi_i}{x_i}\right)\sum_{j=1}^{Nc}L_j x_j+q_i\left(1.0-\ln\sum_{j=1}^{Nc}\theta_j\tau_{ji}\right)-q_i\sum_{j=1}^{Nc}\left(\frac{\theta_j\tau_{ij}}{\sum\limits_{k=1}^{Nc}\theta_k\tau_{kj}}\right)$$

$$L_j=\frac{Z}{2}(r_j-q_j)-r_j+1$$

$$\theta_i=\frac{q_i x_i}{\sum\limits_{j=1}^{Nc}q_j x_j}$$

$$\Phi_i=\frac{r_i x_i}{\sum\limits_{j=1}^{Nc}r_j x_j}$$

$$\tau_{ij}=\exp\left(-\frac{\Delta u_{ij}}{RT}\right)$$

where Z is the coordination number and is set to be equal to 10. r and q are the structural parameters for the pure components and Δu_{ij} are the energy parameters.

As shown in tables 3 and 4, the estimated parameters for the NRTL and the *Wilson* models using the VLE data of tert-butanol and 1-butanol at different pressures correspond to the global minimum point of the error function. It is clear that the reported values in the literature [20] correspond to the local minimum point. The average trend of minimizing the objective function is shown in Fig 3. Since the best values of each generation are less than unity while the average individuals are greater than 100, the best values of each generation could not be shown in Fig. 3, in fact their representing
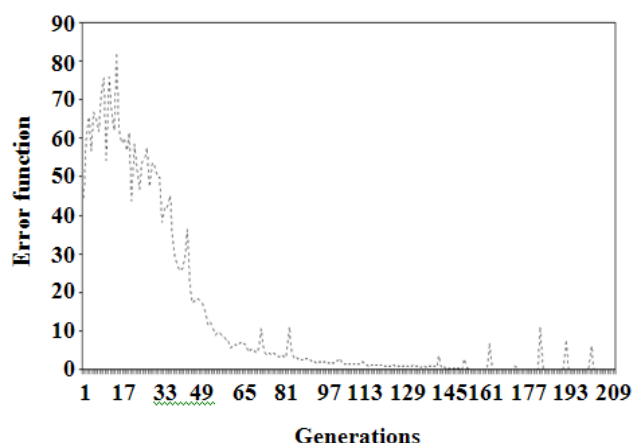

Fig. 3: The average trend of minimizing the error function in the VLE system of tert-butanol (1) and 1-butanol (2) at 700 mmHg. Liquid phase is modeled by Wilson and vapor phase is ideal. Experimental data is taken from [20].

line coincides with the X-axis. Also, these experimental data sets have been used by *Stadtherr et al.,* [7] to estimate the parameters of the *Wilson* model. They have also obtained the same parameter values as those reported in the present work. The calculated values of the activity coefficients at infinite dilution for the *Wilson* model are also reported in table 3.

As can be observed in table 5, estimating the parameters of NRTL model in the LLE system of water, acetic acid and 1-hexanol has been done at three different cases. In the first and second cases the nonrandomness factor (α) is fixed to 0.2 and 0.3, respectively, whereas in the third case these parameters are considered as decision variables. As shown, the minimum values of the objective function are nearly the same for all three cases, but the case II has the most accurate (optimum) results. Thus the nonrandomness factor of NRTL model does not have an important role in the minimum value of the objective function. However, its effect on the optimum values of parameters is considerable. The first case (α=0.2) has also been studied by *Al-Muhtasab et al.,* [21]. From table 5, it is concluded that the results of the present work are more accurate comparing to those reported in the literature [21]. Since the new values of the parameters in this work are different from the ones reported by reference [21] and these parameters lead to a lower value of the objective function relative to results of reference [21], the results of the present work correspond to the new and more optimal solution.

Table 3: The results of parameter estimation for the Wilson model obtained using the VLE data of tert-butanol (1) and 1-butanol (2) at four different pressures. The experimental data are taken from [20].

	Ref. [20]				Ref. [7]			
Press.(mmHg) /Equilibrium points	A_{21} (Cal/mol)	A_{12} (Cal/mol)	f_l	γ_1^∞	A_{21} (Cal/mol)	A_{12} (Cal/mol)	f_l	γ_1^∞
100/9	950.52	-601.50	0.11661	0.93	-568	745	0.10149	0.86
300/9	1067.96	-637.79	0.12670	1.02	-525	626	0.11402	0.88
500/9	901.40	-593.78	0.09849	0.96	-718	1265	0.08307	0.89
700/9	800.56	-561.30	0.13191	0.94	-734	1318	0.11705	0.90
Present work Initial population=100, Maximum number of generations=110								
Press.(mmHg) /Equilibrium points	A_{21} (Cal/mol)		A_{12} (Cal/mol)		f_l	γ_1^∞		CPU (s)
100/9	-567.57		743.00		0.10149	0.86		3.03
300/9	-523.78		622.91		0.11402	0.88		2.94
500/9	-718.22		1264.1		0.08307	0.89		3.41
700/9	-735.05		1321.6		0.11705	0.90		3.80

Table 4: The results of parameter estimation for the NRTL model obtained using the VLE data of tert-butanol (1) and 1-butanol (2) at four different pressures. The experimental data are taken from [20].

Ref. [20]				
Pressure (mmHg)	A_{21} (Cal/mol)	A_{12} (Cal/mol)	$\alpha_{12}=\alpha_{21}$	f_l
100	-744.09	970.77	0.3123	0.10789
300	-826.36	1142.0	0.3035	0.11068
500	931.86	-741.53	0.2968	0.13693
700	1010.5	-873.53	0.2915	0.14117
Present work Initial population=200, Maximum number of generations=210				
Pressure (mmHg)	A_{21} (Cal/mol)	A_{12} (Cal/mol)	$\alpha_{12}=\alpha_{21}$	f_l	CPU (s)
100	-981.71	1231.30	0.2000	0.10602	13.7
300	-1088.7	1441.8	0.2000	0.10784	13.6
500	-965.38	1191.5	0.2000	0.09849	14.6
700	1279.6	-1014.1	0.2000	0.12942	13.5

Table 5: The results of parameter estimation for the NRTL model obtained using the LLE data of water (1), acetic acid(2)
and 1-hexanol (3) at 288 K and 1 atm. The experimental data are taken from [21] with 8 equilibrium data points.

	Present work				Ref. [21]		
	Parameters (Cal/mole)		f_2	CPU (s) (Init.pop., Max.No.of Generations)	Parameters (Cal/mole)		f_2
Case I	A_{12}=-203.188 A_{13}=5003.9 A_{23}=309.045	A_{21}=2593.0 A_{31}=-222.821 A_{32}=716.838	0.16820	165 (100, 110)	A_{12}=1073.4 A_{13}=4203.1 A_{23}=0.6020	A_{21}=-353.25 A_{31}=-192.87 A_{32}=779.34	0.16906
	$\alpha_{12}=\alpha_{21}$=0.2 $\alpha_{13}=\alpha_{31}$=0.2 $\alpha_{23}=\alpha_{32}$=0.2						
Case II	A_{12}=-409.64 A_{13}=4146.8 A_{23}=570.84	A_{21}=-265.02 A_{31}=714.131 A_{32}=-1664.71	0.14744	165 (100, 110)			
	$\alpha_{12}=\alpha_{21}$=0.3 $\alpha_{13}=\alpha_{31}$=0.3 $\alpha_{23}=\alpha_{32}$=0.3						
Case III	A_{12}=-1446.69 A_{13}=4610.96 A_{23}=422.35	A_{21}=2378.91 A_{31}=-226.322 A_{32}=1513.28	0.16220	765 (200, 210)	$\alpha_{12}=\alpha_{21}$=0.2 $\alpha_{13}=\alpha_{31}$=0.2 $\alpha_{23}=\alpha_{32}$=0.2		
	$\alpha_{12}=\alpha_{21}$=0.264363 $\alpha_{13}=\alpha_{31}$=0.205197 $\alpha_{23}=\alpha_{32}$=0.208295						

The values for the parameters of the NRTL and the UNIQUAC models obtained from the LLE data of water, propionic acid and methyl butyl ketone system are shown in the tables 6 and 7. Using the values for the parameters presented in tables 6 and 7 the activity coefficients of the components at infinite dilution can be evaluated.

Tables 8 and 9 present the values for the parameters of the NRTL and the UNIQUAC models along with the Root Mean Square Deviations (RMSD) from the experimental data produced from the models using the LLE data for the system of water, propionic acid and methyl isopropyl ketone at 25, 35 and 45 °C. It should be noted that in tables 6 to 9 the equilibrium data reported in reference [22] have been used to obtain the binary interaction parameters for the new system which has not been studied before.

The results of the parameter estimation for the NRTL and the UNIQUAC models for the quaternary LLE system of 1-octanol (1), tert-amyl methyl ether or TAME (2), Water (3) and Methanol (4) at 25 °C were reported in tables 10 and 11, respectively. As can be seen from tables 10 and 11, the NRTL and UNIQUAC parameters obtained using the method proposed in this work can be compared with those reported in the literature [23]. It should be stated that the nonrandomness factors of the

NRTL model were set to be equal to 0.3 for all the components present in solutions. It goes without saying that, because of the higher dimensionality, the parameter estimation for such system could be much more difficult than the previous systems with lower dimensionality invoking the rigorous procedure of parameter estimation.

Estimating the parameters of the Peng-Robinson (PR) equation of state in the VLLE system of methane (1), water (2) and methylcyclohexane (3) at high pressure has been considered by invoking the rigorous procedure. The results of parameter estimation for this system were reported in table 12. The PR equation of state along with its quadratic mixing rules was given by equation (15).

$$P = \frac{RT}{v-b} - \frac{a}{v(v+b)+b(v-b)} \tag{15}$$

$$a = \sum_{i=1}^{Nc} \sum_{j=1}^{Nc} x_i x_j \sqrt{a_i a_j}$$

$$b = \sum_{i=1}^{Nc} x_i b_i$$

where x_i stands for the mole fraction for component i and a_i and b_i are, respectively, the parameters of pure components calculated using the critical properties and acenteric factor data.

Table 6: The results of parameter estimation for the NRTL model obtained using the LLE system of water (1), propionic acid (2) and methyl butyl ketone (3) at 1 atm. The experimental data are taken from [22].

T(ºC)/ No. of Equilibrium points	Present work			
	Parameters (Cal/mole)		f_2	CPU (s) (Initial.population, Maximum Number of Generations)
25/12	A_{12}=4507.28 A_{13}=5034.96 A_{23}=-541.74	A_{21}=-1493.06 A_{31}=744.738 A_{32}=2371.51	0.14549	290 (100, 110)
	$\alpha_{12}=\alpha_{21}=0.3$ $\alpha_{13}=\alpha_{31}=0.3$ $\alpha_{23}=\alpha_{32}=0.3$			
35/13	A_{12}=612.155 A_{13}=5020.32 A_{23}=-1510.52	A_{21}=1555.48 A_{31}=613.554 A_{32}=5024.23	0.18198	
	$\alpha_{12}=\alpha_{21}=0.3$ $\alpha_{13}=\alpha_{31}=0.3$ $\alpha_{23}=\alpha_{32}=0.3$			
45/12	A_{12}=2328.75 A_{13}=4999.43 A_{23}=-1801.39	A_{21}=-1020.47 A_{31}=484.742 A_{32}=4149.7	0.15894	
	$\alpha_{12}=\alpha_{21}=0.3$ $\alpha_{13}=\alpha_{31}=0.3$ $\alpha_{23}=\alpha_{32}=0.3$			

Table 7: The results of parameter estimation for the UNIQUAC model obtained using the LLE data of water (1), propionic acid (2) and methyl butyl ketone (3) at 1 atm. The experimental data are taken from [22].

T (ºC)/ No. of Equilibrium points	Present work			
	Parameters (Cal/mole)		f_2	CPU (s) (Initial.population, Maximum Number of Generations)
25/12	A_{12}=-2795.45 A_{13}=3463.96 A_{23}=-402.898	A_{21}=4889.72 A_{31}=-264.609 A_{32}=4791.18	0.18905	180 (100, 110)
35/13	A_{12}=-3058.24 A_{13}=3219.8 A_{23}=-349.23	A_{21}=4777.01 A_{31}=-308.866 A_{32}=5121.79	0.19845	
45/12	A_{12}=-3121.34 A_{13}=5006.66 A_{23}=-406.956	A_{21}=5031.72 A_{31}=-345.28 A_{32}=5017.23	0.20663	

CONCLUSIONS

In this work the application of the GA for estimation of the binary interaction parameters for the Wilson, NRTL and the UNIQUAC activity coefficient models and the PR equation of state was proposed. To evaluate the performance of this algorithm, eight different systems at various physical conditions were studied. The systems studied here to test the capability of the GA algorithm are 2-butanone and hexadecane, 2-butanone and octadecane, tert-butanol and 1-butanol, water, acetic acid and 1-hexanol, water + propionic acid + methyl butyl ketone, water + propionic acid + methyl isopropyl ketone, 1-octanol + TAME + water + methanol and methane + water + methylcyclohexane at various temperatures and pressures. The results obtained in this work showed that the hybrid GA can produce physically meaningful as well as accurate values for the binary interaction parameters of the activity coefficient models. The values

Table 8: The results of parameter estimation for the NRTL model obtained using the LLE data of water (1), propionic acid (2) and methyl isopropyl ketone (3) at 1 atm. The experimental data are taken from [22].

T(°C)/ No. of Equilibrium points	Present work			
	Parameters (Cal/mole)		f_2	CPU (s) (Initial population, Maximum Number of Generations)
25/7	$A_{12}=-3376.14$ $A_{13}=5027.46$ $A_{23}=-1960.04$	$A_{21}=5035.52$ $A_{31}=206.591$ $A_{32}=2577.08$	0.12837	150 (100, 110)
	$\alpha_{12}=\alpha_{21}=0.3$ $\alpha_{13}=\alpha_{31}=0.3$ $\alpha_{23}=\alpha_{32}=0.3$			
35/7	$A_{12}=-856.767$ $A_{13}=5023.34$ $A_{23}=-2594.04$	$A_{21}=4857.47$ $A_{31}=132.59$ $A_{32}=5029.00$	0.14100	
	$\alpha_{12}=\alpha_{21}=0.3$ $\alpha_{13}=\alpha_{31}=0.3$ $\alpha_{23}=\alpha_{32}=0.3$			
45/7	$A_{12}=-1876.12$ $A_{13}=5112.55$ $A_{23}=-2077.87$	$A_{21}=190.305$ $A_{31}=27.006$ $A_{32}=-1389.22$	0.16588	
	$\alpha_{12}=\alpha_{21}=0.3$ $\alpha_{13}=\alpha_{31}=0.3$ $\alpha_{23}=\alpha_{32}=0.3$			

Table 9: The results of parameter estimation for the UNIQUAC model obtained using the LLE data of water (1), propionic acid (2) and methyl isopropyl ketone (3) at 1 atm. The experimental data are taken from [22].

Temp.(°C) / No. of Equilibrium points	Present work			
	Parameters (Cal/mole)		f_2	CPU (s) (Initial population, Maximum Number of Generations)
25/7	$A_{12}=-4794.27$ $A_{13}=1429.13$ $A_{23}=-269.742$	$A_{21}=4908.41$ $A_{31}=-341.525$ $A_{32}=4845.09$	0.22226	120 (100, 110)
35/7	$A_{12}=4893.65$ $A_{13}=4420.15$ $A_{23}=-3894.88$	$A_{21}=-716.811$ $A_{31}=-415.635$ $A_{32}=-3666.52$	0.18221	
45/7	$A_{12}=5020.18$ $A_{13}=5038.62$ $A_{23}=5066.68$	$A_{21}=-937.402$ $A_{31}=-435.021$ $A_{32}=5039.2$	0.17564	

for the Binary Interaction Parameters (BIP) were used to obtain the activity coefficients for components at infinite dilutions. The results obtained from the models along with this algorithm were favorably compared with those reported in the literature.

One important issue that affects the calculated interaction parameters for the NRTL model is the non-randomness factor. Because of its order of magnitude, this non-randomness factor can create some difficulties during the calculations. Therefore, because of the lack of an efficient optimization problem to handle the variables with different orders of magnitude, in the techniques used in the literature it is assumed that the value for the non-randomness factor is considered to be constant. However, using a robust method like the hybrid GA, used in the present work, there is no need to assign a fixed value to this factor. In addition to the accuracy, simplicity and the generality of the new algorithm, short CPU computation time should be mentioned as its clearest advantage.

Table 10: The results of parameter estimation for the NRTL model obtained using the LLE data of 1-Octanol (1), TAME (2), Water (3), Methanol (4) at 25 °C. Parameters in Cal/mole and α = 0.3 for all the components. The experimental data are taken from [23] with 29 equilibrium data points.

Present work Initial Population = 100 Maximum Number of Generations = 110 CPU = 881 s		f_2	Ref.[23]		f_2
$A_{12}= 10000.8$	$A_{31}=4815.4$		$A_{12}=-282.87$	$A_{31}=1281.17$	
$A_{13}=9999.5$	$A_{32}=3132.27$		$A_{13}=205.6$	$A_{32}=1416.15$	
$A_{14}=9999.7$	$A_{34}=10000.8$	0.521044	$A_{14}=-618.66$	$A_{34}=-235.98$	1.348111
$A_{21}=9947.2$	$A_{41}=9999.16$		$A_{21}=-332.01$	$A_{41}=60.61$	
$A_{23}=13.07$	$A_{42}=10000.00$		$A_{23}=543.35$	$A_{42}=47.05$	
$A_{24}=8425.99$	$A_{43}=462.95$		$A_{24}=-9.61$	$A_{43}=449.09$	

Table 11: The results of parameter estimation for the UNIQUAC model obtained using the LLE data of 1-Octanol (1), TAME (2), Water (3), Methanol (4) at 25 °C. Parameters in Cal/mole. The experimental data are taken from [23] with 29 equilibrium data points.

Present work Initial Population = 100 Maximum Number of Generations = 110 CPU = 950 s		f_2	Ref.[23]		f_2
$A_{12}=357.61$	$A_{31}=1084.33$		$A_{12}=-247.15$	$A_{31}=157.50$	
$A_{13}=257.35$	$A_{32}=794.50$		$A_{13}=193.63$	$A_{32}=65.45$	
$A_{14}=1093.27$	$A_{34}=103.17$	0.371223	$A_{14}=-327.61$	$A_{34}=-47.18$	0.382536
$A_{21}=607.29$	$A_{41}=-290.70$		$A_{21}=374.23$	$A_{41}=486.91$	
$A_{23}=1108.57$	$A_{42}=1078.32$		$A_{23}=612.75$	$A_{42}=606.48$	
$A_{24}=-285.13$	$A_{43}=-391.88$		$A_{24}=-290.60$	$A_{43}=-225.31$	

Table 12: The results of parameter estimation for the PR model obtained using the VLLE data of Methane (1), Water (2), Methylcyclohexane (3) at 275.5 K. The experimental data are taken from [24] with 4 equilibrium data points. Initial Population = 30. Maximum Number of Generations = 40. CPU = 26 s.

$K_{12} = K_{21}$	$K_{13} = K_{31}$	$K_{23} = K_{32}$	f_2
-0.07865	0.13691	-0.19686	0.166501

REFERENCES

[1] Anderson, T.F., Abrams, D.S. and Grens, E. A. II, Evaluation of Parameters for Nonlinear Thermo-dynamic Models, *AIChE J.*, **24,** p. 20 (1978).

[2] Valko, P. and Vajda, S., An Extended Marquardt-type Procedure for Fitting Error-in-Variables Models, *Comput. Chem. Eng.*, **11**, p. 37 (1987).

[3] Tjoa, I.B., Biegler, L.T., Reduced Successive Quadratic Programming Strategy for Error-in-Variables Models, *Comput. & Chem. Eng.* **16**, p. 523 (1992).

[4] Vamos, R. J. and Hass, C. N., Reduction of Ion-Exchange Equilibria Data Using an Error-in-Variables Approach, *AIChE J.*, **40**, p. 556 (1994).

[5] Moore, R.E., Hansen, E., Leclerc, A., "Recent Advances in Global Optimization", Floudas, C.A. and Pardolos, P.M., Eds., Princeton Univ. Press, Princeton, NJ, p. 321 (1992).

[6] Ratz, D., Csendes, T., On the Selection of Subdivision Directions in Interval Branch-and-Bound Methods for Global Optimization, *J. Global Optim.*, **7**, p. 183 (1995).

[7] Gau, C. Y., Brennecke, J. F. and Stadtherr, M. A., Reliable Nonlinear Parameter Estimation in VLE Modeling, *J. Fluid Phase Equilibria*, **168**, p. 1 (2000).

[8] Esposito, W. R. and Floudas, C. A., Global Optimi-zation in Parameter Estimation of Non-Linear Algebraic Models Via the Error-in-Variable Approach, *Ind. Eng. Chem. Res.*, **37,** p. 1841 (1998).

[9] Panaitescu, G., *Int. Chem. Eng. J.*, **25**, p. 692 (1985).

[10] Holland, J.H., "Adaption in Natural and Artificial Systems", MIT Press, (1975).

[11] Rahul, B., Kasat, Sanatosh, Gupta, K., Multi-Objective Optimization of an Industrial Fluidized-

Bed Catalytic Cracking Unit (FCCU) Using Genetic Algorithm (GA) with the Jumping Genes Operator, *Comput. & Chem. Eng.*, **27**, p. 1785 (2003).

[12] Silva, C. M. and Biscaia, Jr, E. C., Genetic Algorithm Development for Multi-Objective Optimization of Batch Free-Radical Polymerization Reactors, *Comput. & Chem. Eng.*, **27**, p. 1329 (2003).

[13] Geyer, H., Ulbig, P. and Schulz, S., Use of Evolutionary Algorithms for the Calculation of Group Contribution Parameters in Order to Predict Thermodynamic Properties: Part 2: Encapsulated Evolution Strategies, *Computers & Chemical Engineering*, **23**, p. 955 (1999).

[14] Beasley, D., Bull, D. R. and Martin, R. R., An Overview of Genetic Algorithms: Part 1, Fundamentals, *University Computing*, **15**, p. 58 (1993).

[15] Beasley, D., Bull, D. R. and Martin, R. R, An Overview of Genetic Algorithms: Part 2, Research Topics, *J. Univ. Computing*, **15**, p. 170 (1993).

[16] Michalewicz, Z., "Genetic Algorithms + Data Structures = Evolution Programs", AI Series, Springer-Verlag, NewYork, (1994).

[17] Lucia, A., Padmanabhan, L., Venkataraman, S., Multiphase Equilibrium Flash Calculations, *Comput. & Chem. Eng.*, **24**, p. 2557 (2000).

[18] Gmehling, J., Onken, U. and Arlt, W., "Vapor-Liquid Equilibrium Data Collection: Vol. I, Part 3.b", Chemistry Data Series, DECHEMA, Frankfurtr-Main, Germany, (1983).

[19] Tiegs, D., Gmehling, J., Medina, A., Soares, M., Bastos, J., Alessi, P. and Kikic, I., "Activity Coefficients at Infinite Dilution: Vol. IX, Part 2", Chemistry Data Series, DECHEMA, Frankfurtr-Main, Germany, (1986).

[20] Gmehling, J., Onken, U. and Arlt, W., "Vapor-Liquid Equilibrium Data Collection: Vol. I, Part 2.a", Chemistry Data Series, DECHEMA, Frankfurtr-Main, Germany, (1983).

[21] Al-Muhtasab, Sh. A. and Al-Nashef, I., M., Liquid-Liquid Equilibria of the Ternary System Water + AceticAcid + 1-Hexanol, *J. Chem. Eng. Data*, **42**, p. 183 (1997).

[22] Taghikhani, V., Vakili-Nezhaad, G.R., Khoshkbarchi, M. K. and Shariaty-Niassar, M., Liquid-Liquid Equilibria of Water + Propionic Acid + Methyl Butyl Ketone and Water + Propionic Acid + Methyl Isopropyl Ketone, *J. Chem. Eng. Data*, **46**, p. 1107 (2001).

[23] Arce, A., Blanco, M. and Soto, A., Determination and Correlation of Liquid-Liquid Equilibrium Data for the Quaternary System 1-octanol + 2-methoxy-2-methylbutane + water + methanol at 25 °C, *J. Fluid Phase Equilibria*, **158**, p. 949 (1999).

[24] Susilo, R., Lee, J. D. and Englezos, P., Liquid-Liquid Equilibrium Data of Water with Neohexane, Methylcyclohexane, Tert-butyl methyl ether, n-Heptane and Vapor-Liquid-Liquid Equilibrium with Methane, *J. Fluid Phase Equilibria*, **231**, p. 20 (2005).