We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

6,900
Open access books available

186,000
International authors and editors

200M
Downloads

154
Countries delivered to

Our authors are among the

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS
BOOK CITATION INDEX
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Prediction of Solubility of Active Pharmaceutical Ingredients in Single Solvents and Their Mixtures — Solvent Screening

Ehsan Sheikholeslamzadeh and Sohrab Rohani

Additional information is available at the end of the chapter

Abstract

In this chapter, the applicability of two predictive activity coefficient-based models will be examined. The experimental data from five different types of VLE (vapor-liquid equilibrium) and VLLE (vapor-liquid-liquid equilibrium) systems that are common in industry are used for the evaluation. The nonrandom two-liquid segment activity coefficient (NRTL-SAC) and universal functional activity coefficient (UNI-FAC) were selected to model the systems. The various thermodynamic relations existing in the open literature will be discussed and used to predict the solubility of active pharmaceutical ingredients and other small organic molecules in a single or a mixture of solvents. Equations of states, the activity coefficient, and predictive models will be discussed and used for this purpose. We shall also present some of our results on solvent screening using a single and a mixture of solvents.

Keywords: Solubility prediction, Pharmaceuticals, , NRTL-SAC Thermodynamic model, Activity coefficient, Solvent screening, Single solvent, Solvent mixture

1. Introduction

The study of solutions and their properties is one of the important branches of thermodynamics. It is important to know the behavior of a mixture of components for a change in temperature, pressure, and composition. Knowledge in this area of thermodynamics helps engineers

and scientists to better design, optimize, and operate the process units in the oil, gas, petro-chemicals, pharmaceuticals, agricultural, and other chemical-related industries. Knowledge of the thermodynamics is essential to improve process performance and product quality. For example, the use of phase behavior calculations to understand and estimate the production rate of solids from a crystallization process in a pharmaceutical industry is of paramount importance in modeling, optimization, and control of product quality.

A solution is called ideal if its mixture property is a linear combination of the properties of each of its constituents at the given temperature and pressure [1]:

$$m_{solution} = \sum_{i=1}^{N} x_i m_i \tag{1}$$

where $m_{solution}$ and m_i represent the molar property of the mixture and pure component i, respectively, and x_i denotes the mole fraction of each constituent i in the solution. Not all solutions can be considered ideal. The interactions between the solute and solvent molecules in the solution renders a solution nonideal. The fundamental relationship which relates the thermodynamic properties is the Gibbs free energy [1]:

$$dG = VdP - SdT + \sum_{i=1}^{N} \frac{\partial G}{\partial n_i}\bigg|_{P,T,n_j} dx_i \tag{2}$$

where G, V, P, S, and T are, respectively, Gibbs free energy, volume, pressure, entropy, and temperature of the solution. The $\frac{\partial G}{\partial n_i}\big|_{P,T,n_j}$ represents the change in the Gibbs energy as a result of changes in the concentration of species i while the pressure, temperature, and the molar content of other species are kept constant. This is known as the chemical potential and in some textbooks is shown by μ_i. For a real solution, the chemical potential for each species is expressed by

$$\mu_i = \mu_i^o + RT ln a_i \tag{3}$$

In which μ_i^o and a_i are the chemical potential of species i in its standard conditions and activity of the species i. The activity is defined as

$$a_i = \gamma_i x_i \tag{4}$$

where γ_i is the activity coefficient which is 1 for ideal solutions. Inserting Equation (4) into Equation (3), results in

$$\mu_i = \mu_i^o + RTln\gamma_i + RTlnx_i \tag{5}$$

The term $RTln\gamma_i$ accounts for the nonideality of the solution (it is also referred to as the partial molar energy). As it can be seen from Equation (5), the terms on the right side, except $Tln\gamma_i$, are calculated from the pure properties. However, in order to have an accurate prediction of the chemical potential of any species in the solution, $RTln\gamma_i$ should be known as well. In general, the activity coefficient is a function of temperature and composition and to a much less extent, the pressure. Because the activity coefficient is defined for a liquid solution, the pressure has very little effect on it. However, temperature and mole fractions of the species have significant effects on the activity coefficient of each species in a solution.

This chapter provides the reader with different and the most up-to-date thermodynamic models to estimate the activity coefficients of active pharmaceutical ingredients (API) in a solution.

1.1. Thermodynamics of the solutions containing dissolved solids

For a solid in equilibrium with itself in a solution, there is a thermodynamic relation [2]:

$$\hat{f}_i^s(T,P) = \hat{f}_i^L(T,P,x_i) \tag{6}$$

where $\hat{f}_i^s(T, P)$ denotes fugacity of component i in solid phase and $\hat{f}_i^L(T, P, x_i)$ represents the fugacity in the liquid phase. Equation (7) below correlates the fugacity of a pure component i in solid and liquid state (f_i^s and f_i^L, respectively).

$$f_i^s = x_i\gamma_i f_i^L \tag{7}$$

In order to find the ratio of the two fugacities, we need to consider all the thermodynamic processes that are involved from a solid to the liquid state. The three processes are shown in Figure 1. Path A in this figure shows the transformation from process conditions to the state where the solid starts melting. Path B shows the melting process at constant temperature and pressure. Path C indicates the change from melting to the process state. The sum of the three paths will give us the whole change from solid to liquid state of a pure component.

From the fundamental rules in thermodynamics, we have:

$$\Delta G_{s \to L} = RTln\frac{f_i^L}{f_i^S} \tag{8}$$

And the Gibbs energy change is related to change of enthalpy and entropy:

$$\Delta G_{s \to L} = \Delta H_{s \to L} - T\Delta S_{s \to L} \tag{9}$$
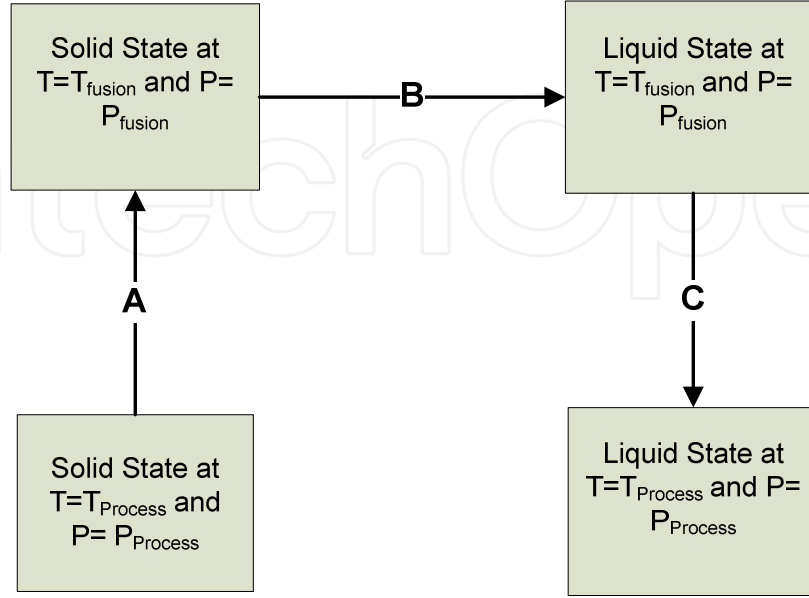

Figure 1. Schematic diagram for finding the fugacity change from solid to liquid state of a pure substance.

From Figure 1, the whole change in enthalpy is:

$$C_{p,solid}dT + \Delta H_{fusion} + \int_{T_f}^{T} C_{p,liquid}dT = \int_{T_f}^{T} \Delta C_p dT + \Delta H_{fusion}$$

$$\Delta H_{s \to L} = \Delta H_A + \Delta H_B + \Delta H_C = \int_{T}^{T_f} \tag{10}$$

where $\Delta C_p = C_{p,liquid} - C_{p,solid}$. In the same manner, the entropy change from solid to liquid state can be found from

$$\Delta S_{s \to L} = \int_{T_f}^{T} \frac{\Delta C_p}{T} dT + \Delta S_{fusion} \tag{11}$$

Also, $\Delta S_{fusion} = \frac{\Delta H_{fusion}}{T_{fusion}}$. If we substitute Equations (10) and (11) into Equation (9), then:

$$RT \ln \frac{f_i^L}{f_i^S} = \int_{T_f}^{T} \Delta C_p dT + \Delta H_{fusion} + \int_{T_f}^{T} \frac{\Delta C_p}{T} dT + \frac{\Delta H_{fusion}}{T_{fusion}} \tag{12}$$

If we neglect the terms including change in the heat capacity (because of the large value of heat of fusion compared to the heat capacities), then we will get the following important equation:

$$ln x_s = \frac{-\Delta H_{fus}}{R}\left(\frac{1}{T_m} - \frac{1}{T}\right) - ln\gamma_s \tag{13}$$

where x_s is the equilibrium mole fraction of the solute (dissolved solid) in a solution at temperature T. Equation (13) is the starting point in every solid-liquid equilibrium calculation procedure that relates the three important variables x_s, T, and γ_s. However, we already know that γ_s is a function of solution temperature and the mole fraction of the species. Therefore, one can fix all the thermodynamic properties of a solution if the mole fraction of the species and the solution temperature are known. We recall that although the pressure has to be fixed as well, but due to its minimal effect on the activity coefficient and other properties of the system, we don't need to have it for calculations.

1.2. Classification of the thermodynamic models for solutions

In order to predict the solubility of solids in pure and mixed solvents, the use of noncorrelative (predictive) thermodynamic models is of high importance. Since several years ago, there have been many thermodynamic models introduced to help scientists and engineers in the predic‐ tion of phase behaviors of various liquid-liquid and vapor-liquid systems, such as the Margules equation [3], the Wilson equation [4], the Van Laar equation [5], the nonrandom two-liquid (NRTL) equation [6], and the UNIQUAC equation [7]. As these models are applicable in the estimation of activity coefficients, they can also be utilized to predict solid-liquid equilibrium systems. In general, the models which govern the activity coefficients of the species in the solutions are grouped into two categories:

1. The models which need experimental data at various temperatures, pressures, and compositions of the mixture, such as UNIQUAC equation.

2. The models which only need some fundamental properties of the chemical molecule and a very few experimental data to predict the phase behavior of the solids within various solvents, such as the UNIFAC [8] and NRTL-SAC models [9].

As it is apparent from the above two categories, the first one is not useful for the prediction of the phase behavior of mixtures, specifically for mixtures of more than two species. Because of the time-consuming and expensive nature of binary interaction parameter evaluation of various chemical components, the first group of thermodynamic models (above) is not practical. As an example, the activity coefficient from Wilson's equation of state is found from [4]:

$$ln\gamma_i = -ln\left[\sum_{j=1}^{n} x_j \Lambda_{ij}\right] + 1 - \sum_{z=1}^{n} \frac{x_z \Lambda_{ki}}{\sum_{k=1}^{n} x_k \Lambda_{zk}} \tag{14}$$

The binary interaction parameter is:

$$\Lambda_{ij} = \frac{V_j^L}{V_i^L} exp\left[\frac{-\lambda_{ij} - \lambda_{ii}}{RT}\right] \tag{15}$$

where i and j refer to the compounds present in the solution. From Equation (14), it can be seen that for obtaining the activity coefficient of a component 1 in a pure solvent 2, we need four interaction parameters (Λ_{12}, Λ_{21}, $\Lambda_{11} \wedge \Lambda_{22}$, which are temperature dependent. In addition, from Equation (15), it is evident that for calculating the value of the binary interaction parameters, additional experimental data, such as molar volume is needed.

From the predictive category (second group), the universal functional activity coefficient (UNIFAC) model is a well-known example. The main application of the UNIFAC model is in systems showing nonelectrolytic and nonideal behaviors. Fredenslund et al. [8] developed the UNIFAC model. The NRTL segment activity coefficient (NRTL-SAC) model was first introduced in 2004 by Chen et al. [9]. This model was proposed in order to compensate for the weakness of the UNIFAC in predicting the solubility of complex chemical molecules with functional groups that had not been studied for the UNIFAC parameters. Also, in some cases, the UNIFAC group addition rule becomes invalid [2]. One of the main advantages of the NRTL-SAC model in comparison to the other predictive methods is its ability to predict organic electrolyte systems. The UNIFAC method identifies the molecule in terms of its functional groups, while the NRTL-SAC model divides the whole surface of the molecule to four segments. These segments' values are found from their interactions with other molecules in a solution. Based on Chen et al., three purely hydrophilic, hydrophobic, and polar segments have been selected as water, hexane, and acetonitrile, respectively.

2. Ideal solutions

An ideal solution is simply defined as a mixture of chemical components with its thermodynamic property related to the linear sum of each pure species thermodynamic property (Equation (1)). A common example is a solution which obeys Raoult's law. This law states that the total pressure of a system is a linear combination of the component's vapors pressure at the system's temperature, provided that the total pressure is less than 5 atm. In order to derive Raoult's law, we start from Equation (5) and assume that the liquid solution is ideal:

$$\mu_i = \mu_i^o + RTlnx_i \tag{16}$$

If Equation (16) is written for component i in both the liquid and vapor phases, we have:

$$\mu_i^L = \mu_i^{o,L} + RTlnx_i \tag{17}$$

$$\mu_i^V = \mu_i^{o,V} + RTln\frac{\hat{f}_i}{P} \tag{18}$$

where $\hat{f}_i$ and P are the fugacity of species i in the vapor mixture and the total pressure, respectively. From the thermodynamic equilibrium criteria, if the two phases of liquid and vapor are in a state of equilibrium, then:

$$\mu_i^L = \mu_i^V \tag{19}$$

After substitution of Equations (17) and (18) into Equation (19), we get:

$$\mu_i^{o,L} + RTlnx_i = \mu_i^{o,V} + RTln\frac{\hat{f}_i}{P} \tag{20}$$

If Equation (19) is written for the case of pure component i in liquid phase at equilibrium with its vapor phase, then:

$$\mu_i^{o,L} = \mu_i^{o,V} + RTln\frac{f_i}{P} \tag{21}$$

By comparing Equations (20) and (21), it yields:

$$RTlnx_i = RTln\frac{\hat{f}_i}{f_i} \tag{22}$$

Equation (22) is simplified to:

$$\hat{f}_i = f_i x_i \tag{23}$$

Now, if the liquid and vapor mixtures are assumed to be ideal, then the fugacity values can be substituted by their corresponding pressure and thus:

$$P_i = P_i^{sat} x_i \tag{24}$$

P_i is called the partial pressure of species i in the vapor mixture and $P_i{}^{sat}$ denotes the pure vapor pressure of the component i at the solution temperature.

2.1. Ideal solution mixtures: VLE phase behavior

In order to calculate the properties of a system of components that obey Raoult's law, Equation (24) is written for each of the species present in the solution,

$$\sum_{i=1}^{N} y_i P_{tot} = \sum_{i=1}^{N} P_i^{sat} x_i \tag{25}$$

and by simplification:

$$P_{tot} = \sum_{i=1}^{N} P_i^{sat} x_i \tag{26}$$

by which the solution equilibrium temperature can be found. As we know, the vapor pressure of each species is temperature-dependent and by having the mole fraction of the components in the liquid phase, one can find the temperature which corresponds to the total pressure of the system. This problem can be solved quickly using a spreadsheet and by changing the solution temperature until Equation (26) is satisfied. Sometimes, the temperature of the solution is known and the total pressure needs to be calculated, which is a straightforward calculation from Equation (26). This way, there is no need to use the nonlinear solvers to find the temperature.

There are two examples of the binary systems which exhibit negative deviation from Raoult's law.
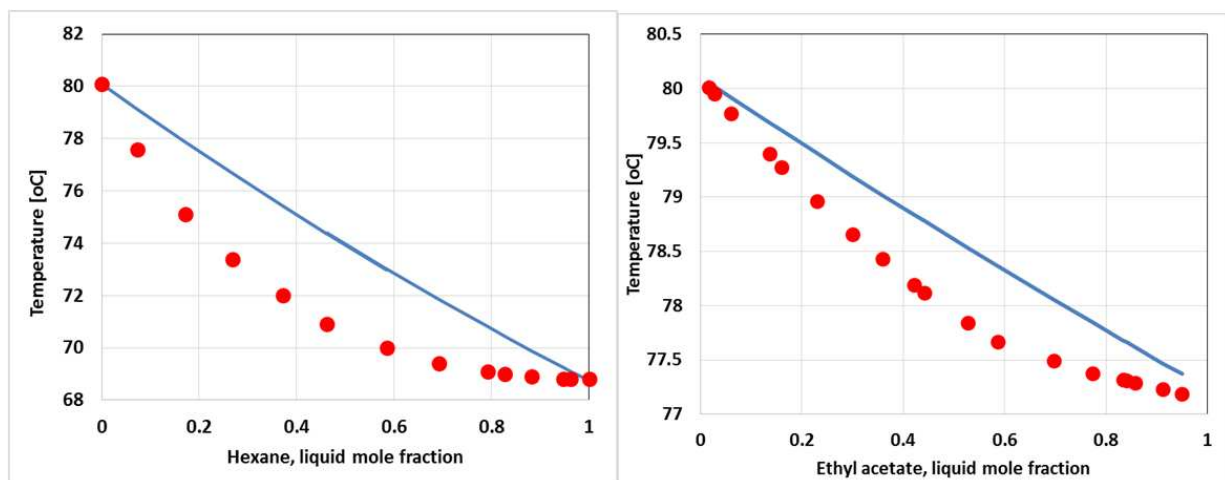

Figure 2. T-x-y diagram for the binary systems of hexane-benzene (left) and ethyl acetate-benzene (right) at a pressure of 101.33 kPa.

In order to better demonstrate whether a system follows Raoult's law, a diagram of the phase equilibrium called T-x-y should be plotted. This plot (Figure 2) shows the equilibrium temperatures at which either a liquid solution will start bubbling (bubble curve) or a vapor mixture starts condensing (dew curve). The two systems with their experimental data and the calculation curve of the ideal solution is shown in Figure 2. In Figure 2, the system of hexane-benzene at the pressure of 101.33 kPa [10] and the system of ethylacetate-benzene [11] show negative deviations from Raoult's law.

If one is interested in finding the bubble and dew curves for an ideal solution at low to moderate pressures, then:

• **Bubble point**

From Equation (26), by knowing the mole fraction of each component in the liquid phase, two cases are possible:

1. **Pressure known.** If the total pressure of the system is given, then Equation (26) should be solved for the temperature (if the vapor pressure follows Antoine's law):

$$P_i^{sat} = A - \frac{B}{T+C}, P_{tot} = \sum_{i=1}^{N} P_i^{sat}(T) x_i \tag{27}$$

2. **Temperature known.** For this case, the problem is straightforward. The vapor pressure of each species can be found readily from Antoine's equation and then inserted into Equation (27) to find the total pressure.

• **Dew point**

1. **Pressure known.** If Raoult's law is isolated for the mole fraction of the species in the vapor phase, then:

$$y_i = \frac{P_i^{sat} x_i}{P_{tot}} \tag{28}$$

by taking the sum for all the components in the liquid phase and knowing that the sum of the mole fractions in a phase is 1, we get:

$$1 = \sum_{i=1}^{N} x_i = \frac{y_i P_{tot}}{P_i^{sat}(T)} \tag{29}$$

If the total pressure is known, then Equation (29) should be solved for the temperature which determines the vapor pressure of each of the species in the mixture.

2.   **Temperature known.** In this case, Equation (29) is isolated for the P_{tot} to find the total pressure of the system:

$$P_{tot} = \frac{1}{\sum\limits_{i=}^{} N \dfrac{P_i^{sat}(T)}{y_i}} \tag{30}$$

which has a straightforward solution.

2.2. Ideal solution mixtures: SLE phase behavior

For a solid-liquid equilibrium behavior, Equation (13) for the ideal solution (in which $\gamma_s = 1$) is simplified to:

$$lnx_s = \frac{-\Delta H_{fus}}{R}\left(\frac{1}{T_m} - \frac{1}{T}\right) \tag{31}$$

From Equation (31), one can find the mole fraction (or what is called *solubility* in the case of SLE) by inserting the appropriate values for the enthalpy of fusion, ΔH_{fus}, melting temperature, T_m, and the solution temperature, T. It can be found from this equation that for any solvent, the solubility of a solid will be the same regardless of the nature of the solvent. For very few cases, this assumption might be correct; however, for most of the practical and common applications, the above formula does not work well. Therefore, a term accounting for the nonideality of the system should be added to Equation (31).

3. Nonideal solutions

Almost all real VLE or SLE systems show nonideality. Equations that predict the behavior of different phase equilibria are divided into two:

- Equations of state (EOS) that are useful to predict the nonideal behavior of the vapor phase

- Equations which are applied to the liquid phase to predict the nonideal behavior of the liquid solutions

We focus on the thermodynamic models that deal with the liquid mixtures in this chapter. From the two categories of activity coefficient models, the correlative one is not very useful for solubility prediction and solvent screening purposes. The main reason for this is the lack of experimental data for the binary interaction parameters of the solute-solvent, solute-antisolvent, and solvent-antisolvent systems. As an example, the activity coefficient from

Wilson's equation of state is found from Equations (14) and (15). It can be seen that for obtaining the activity coefficient of a component 1 in a pure solvent 2, we need four interaction parameters (Λ_{12}, Λ_{21}, $\Lambda_{11} \wedge \Lambda_{22}$, which are temperature dependent. It is evident that for calculating the value of the binary interaction parameters, additional experimental data, such as molar volume is needed. Other models which belong to the first category have the same limitations as Wilson's method. The Wilson model was used in the prediction of various hydrocarbons in water in pure form and mixed with other solvents by Matsuda et al. [11]. In order to estimate the pure properties of the species, the Tassios method [12] with DECHEMA VLE handbook [13] were used. Matsuda et al. also took some assumptions in the estimation of binary interactions (because of the lack of data) that resulted in some deviations from the experimental data.

From the predictive category, we bring some examples of the application of the UNIFAC model. In one study, this model has been used to predict the solubility of naphtalene, anthracene, and phenanthrene in various solvents and their mixtures [8]. They showed the applicability of the UNIFAC model in prediction of the phase behavior of solutes in solvents. There have been efforts to make the UNIFAC model more robust and powerful in the prediction of phase behaviors [14]. In one study, the solubility of buspirone-hydrochloride in isopropyl alcohol was measured and evaluated by the modified UNIFAC model [15]. It was concluded that for highly soluble pharmaceutics, the modified form of the UNIFAC model was not suitable. In another study, the solubility of some chemical species in water and some organic solvents was predicted by the UNIFAC model [16]. For some unknown functional groups, they used other known groups which had chemical structures that were similar to unknown ones. In conclusion, it was stated that the UNIFAC model is not a proper model for use in crystallization and related processes. The UNIFAC model also has been utilized to predict the solubility of some aromatic components as well as long-chain hydrocarbons [17]. The results showed that the predictions for the linear hydrocarbons are not as good as the ones for the aromatics.

Chen et al. developed the NRTL-SAC model in 2004 [9]. One of the main reasons for developing this model was to enhance the predicting capability of the solubility of complex chemical molecules with functional groups that were not included in the UNIFAC model. Also, in some cases, the UNIFAC group addition rule becomes invalid [2]. One of the main advantages of the NRTL-SAC model in comparison to the other predictive methods is its ability to predict organic electrolyte systems [18]. The UNIFAC method identifies the molecule in terms of its functional groups, while the NRTL-SAC model divides the whole surface of the molecule to four segments. The hydrophobic segment (X) denotes parts of the molecule surface which don't participate in forming a hydrogen bond, such as hexane. The polar segments (Y- and Y+) do not belong to either hydrophobic or hydrophilic segment. The polar attractive segment (Y-) shows attractive interaction with hydrophilic segment, while the polar repulsive segment (Y +) has repulsive characteristic with hydrophilic segment. Hydrophilic segment (Z) contributes to the part of the molecule which tends to form a hydrogen bond, such as water.

Based on the interaction forces between the molecule surfaces, the four segments of the NRLT-SAC model have been identified. The three reference solvents mentioned before were used to

determine the rest of chemical components' segment numbers [9]. As an example, the VLE and LLE data containing the component of interest in binary mixtures with hexane, water, and acetonitrile were collected and then used in order to find the segment numbers of the component. Sheikholeslamzadeh et al. have investigated various industrial case studies to show the applicability of the NRTL-SAC in processes consisting of solvent mixtures [19]. The segment numbers for each of the nonreference solvents are found from the nonlinear optimization methods which minimize the deviation from the model output and the experimental data for the whole range of data (VLE and LLE). Most of the solvents have some segment numbers that are high in value compared to other segment numbers. The reason is that it is less probable that a chemical component collects all four segments at the same level of value. Currently, more than a hundred solvent segment numbers have been evaluated and optimized which can be found in some commercial simulation packages. The two important predictive equations of UNIFAC and NRTL-SAC, as representatives of the activity coefficient models, are presented here.

3.1. UNIFAC model

In general, the activity coefficient models of the predictive category split the activity coefficients into two segments:

- A part that includes the contribution of the chemical structure and the size of a compound (combinatorial part)

- The second part that includes the contribution of the functional sizes and binary interaction between pairs of the functional groups (residual part)

With the above definition, the total activity of a component in the solution is the sum of the two parts:

$$ln\gamma_i = ln\gamma_i^C + ln\gamma_i^R \tag{32}$$

In which γ_i is the activity coefficient of component i in the solution, γ_i^C is the combinatorial part and γ_i^R is the residual part. Up to this point, all of the group contribution and activity coefficient methods (i.e. NRTL-SAC) have been the same, but the methods in which the activities have been calculated are different. In the UNIFAC model, the combinatorial part for component i is found from the following equation [8]:

$$ln\gamma_i^c = ln\left[\frac{\phi_i}{x_i}\right] + \frac{z}{2}q_i ln\left[\frac{\theta_i}{\phi_i}\right] + L_i - \frac{\phi_i}{x_i}\sum_{j=1}^{n}x_j L_j \tag{33}$$

where

$$L_i = \frac{z}{2}(r_i - q_i) - (r_i - 1) \tag{34}$$

z is the coordination number and is taken to be 10. In Equation (33), ϕ_i is the segment fraction and θ_i is the area fraction of component i and is related to the mole fraction of species i in the mixture:

$$\phi_i = \frac{r_i x_i}{\sum_{j=1}^{n} r_j x_j} \tag{35}$$

$$\theta_i = \frac{q_i x_i}{\sum_{j=1}^{n} q_j x_j} \tag{36}$$

q_i and r_i are the pure component surface area and volumes (van der Waals), respectively. These parameters are not temperature-dependent and are only functions of the chemical structure of a functional group. In the UNIFAC model, for every functional group, there is a unique value for surface area and volume that can be found in common texts and handbooks [1]. The first step in modeling the UNIFAC for a specific binary or ternary system is to break down the chemical structure of a molecule into the basic functional groups. As it is suggested in thermodynamic textbooks [19], the optimum way of breaking it down is the one which results in the minimum number of subgroups, and with each subgroup having the maximum replicates. The q_i and r_i can be found from Equations (37) and (38):

$$r_i = \sum_k v_k^{\,i} R_k \tag{37}$$

$$q_i = \sum_k v_k^{\,i} Q_k \tag{38}$$

$v_k^{\,i}$ is the number of subgroup k in component i. The residual part of the UNIFAC is found from the equation below:

$$ln\gamma_i^R = \sum_{k=1}^{n_k} v_k^{\,i} \left[ln\Gamma_k - ln\Gamma_k^{\,i} \right] \tag{39}$$

In Equation (39), Γ_k is the residual activity coefficient of subgroup k in the mixture and $\Gamma_k^{\,i}$ is that value in a pure solution of the component i. This term is added so when the mole fraction approaches unity, the term $ln\gamma_i^R$ tends to zero ($\gamma_i^R \rightarrow 1$). The residual activity coefficient of subgroup k in a solution is given by

$$ln\Gamma_k = Q_k\left[1 - ln\sum_m\theta_m\psi_{mk} - \sum_m\left[\frac{\theta_m\psi_{mk}}{\sum_n\theta_n\psi_{nm}}\right]\right] \tag{40}$$

Equation (40) is also applicable to the case of $\Gamma_k{}^i$, in which the parameters of the right-hand side of the equation are written based on the pure component i. θ_m is the area fraction of the functional group m in the mixture:

$$\theta_m = \frac{Q_mX_m}{\sum_n Q_nX_n} \tag{41}$$

X_m is the mole fraction of subgroup m in the mixture. ψ_{nm} is the group interaction parameter between groups n and m and is dependent on the temperature:

$$\psi_{nm} = exp\left[\frac{-u_{nm} - u_{mm}}{RT}\right] = exp\left[\frac{-a_{nm}}{T}\right] \tag{42}$$

The group interaction parameter a_{nm} is found from the large sets of VLE and LLE data in the literature, which are tabulated for many subgroups. It is worth noting that $a_{nm} \neq a_{mn}$. There are some modifications to the original UNIFAC equation in order to make the model robust for some complex systems. In the UNIFAC-DM method, the modification is made on the combinatorial part:

$$ln\gamma_i^c = ln\left[\frac{\phi_i'}{x_i}\right] + \frac{z}{2}q_i ln\left[\frac{\theta_i}{\phi_i}\right] + L_i - \frac{\phi_i}{x_i}\sum_{j=1}^n x_j L_j \tag{43}$$

In which the term $\frac{\phi_i'}{x_i}$ is defined as

$$\frac{\phi_i'}{x_i} = \frac{r^{3/4}}{\sum_j x_j r_j^{3/4}} \tag{44}$$

The algorithm for finding the solute solubility in the mixture of ternary solution (solvent/cosolvent/solute) is shown in Figure 3. The algorithm starts with known values, such as the physical properties of the solute. After making an initial guess for the solubility, the program obtains the activity coefficients and the new solubility is found and is compared with the old

value and the calculations are repeated to converge to a unique value for solubility. This procedure is done for all the experimental data points.
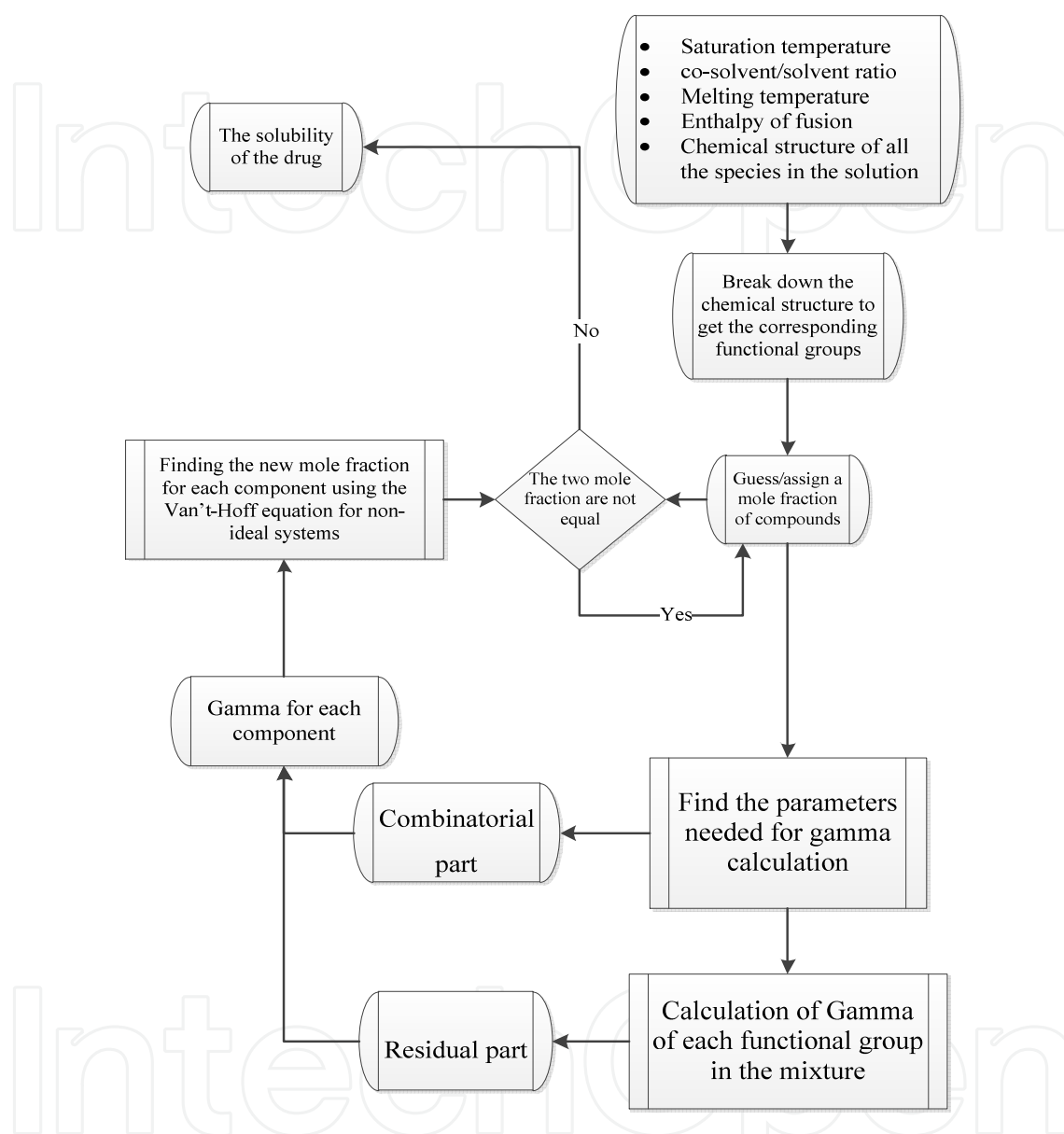

Figure 3. The algorithm of converging to the solubility of a ternary system using the UNIFAC model.

3.2. Nonrandom two-liquid segment activity coefficient (NRTL-SAC)

According to Chen et al. [9], the NRTL-SAC model is based on the derivation of the original NRTL model for polymers. From Equation (32), the activity coefficient is made up of two terms, combinatorial and residual. Like the UNIFAC model, the activity coefficients must be generated in order to obtain solubility. In the NRTL-SAC model, the combinatorial part is calculated by Equation (45):

$$ln\gamma_i^{\ C} = ln\frac{\emptyset_i}{x_i} + 1 - r_i\sum_j\frac{\emptyset_j}{x_j} \tag{45}$$

With the definitions:

$$r_i = \sum_j r_{j,i} \tag{46}$$

$$\emptyset_i = \frac{r_i x_i}{\sum_j r_j x_j} \tag{47}$$

where x_i is the mole fraction of component i, $r_{m,i}$ is the number of segment m, r_i is the total segment number in component i, and $\emptyset_i$ is the segment mole fraction in the mixture.

The residual term is defined as

$$ln\gamma_i^{\ R} = ln\gamma_i^{\ lc} = \sum_m r_{m,i}\left(ln\Gamma_m^{\ lc} - ln\Gamma_m^{\ lc,i}\right) \tag{48}$$

In Equation (48), there are two terms, $ln\Gamma_m^{\ lc}$ and $ln\Gamma_m^{\ lc,i}$, which are the activity coefficients of segment m in solution and component i, respectively.

The two mentioned terms are found using Equations (49) and (50):

$$ln\Gamma_m^{\ lc} = \frac{\sum_j x_j G_{j,m}\tau_{j,m}}{\sum_k x_k G_{k,m}} + \sum_{m'}\frac{x_{m'}G_{m,m'}}{\sum_k x_k G_{k,m'}}\left(\tau_{m,m'} - \frac{\sum_j x_j G_{j,m'}\tau_{j,m'}}{\sum_k x_k G_{k,m'}}\right) \tag{49}$$

$$ln\Gamma_m^{\ lc,l} = \frac{\sum_j x_{j,l} G_{j,m}\tau_{j,m}}{\sum_k x_{k,l} G_{k,m}} + \sum_{m'}\frac{x_{m',l}G_{m,m'}}{\sum_k x_{k,l} G_{k,m'}}\left(\tau_{m,m'} - \frac{\sum_j x_{j,l} G_{j,m'}\tau_{j,m'}}{\sum_k x_{k,l} G_{k,m'}}\right) \tag{50}$$

In the two above equations, l is referred to as the component and j, k, m, and m' are referred to the segments in each component. $x_{j,l}$ is the segment-based mole fraction of segment species j in component l only. The mole fractions of segments in the whole solution and in components are defined as below:

$$x_j = \frac{\sum_l x_l r_{j,l}}{\sum_z\sum_i x_z r_{j,z}} \tag{51}$$

$$x_{j,l} = \frac{r_{j,l}}{\sum_i r_{j,l}} \tag{52}$$

$G_{i,j}$ and $\tau_{i,j}$ are the local binary values which can be related to each other based on NRTL nonrandom parameter α_{ij}, and are shown by their values in Table 1. $G_{i,j}$ and $\tau_{i,j}$ have the following relation:

$$G_{i,j} = e^{-\alpha_{i,j}\tau_{i,j}} \tag{53}$$

Therefore, from fixed values of $\tau_{i,j}$ and $\alpha_{i,j}$ one can find $G_{i,j}$. The segment numbers for the common solvents can be found from the literature [20]. After putting the values of segments for solvents and initial guess values for the solute segments, the written code for NRTL-SAC starts solving for the mole fractions at saturation for all of the species in the solution (see Figure 3).
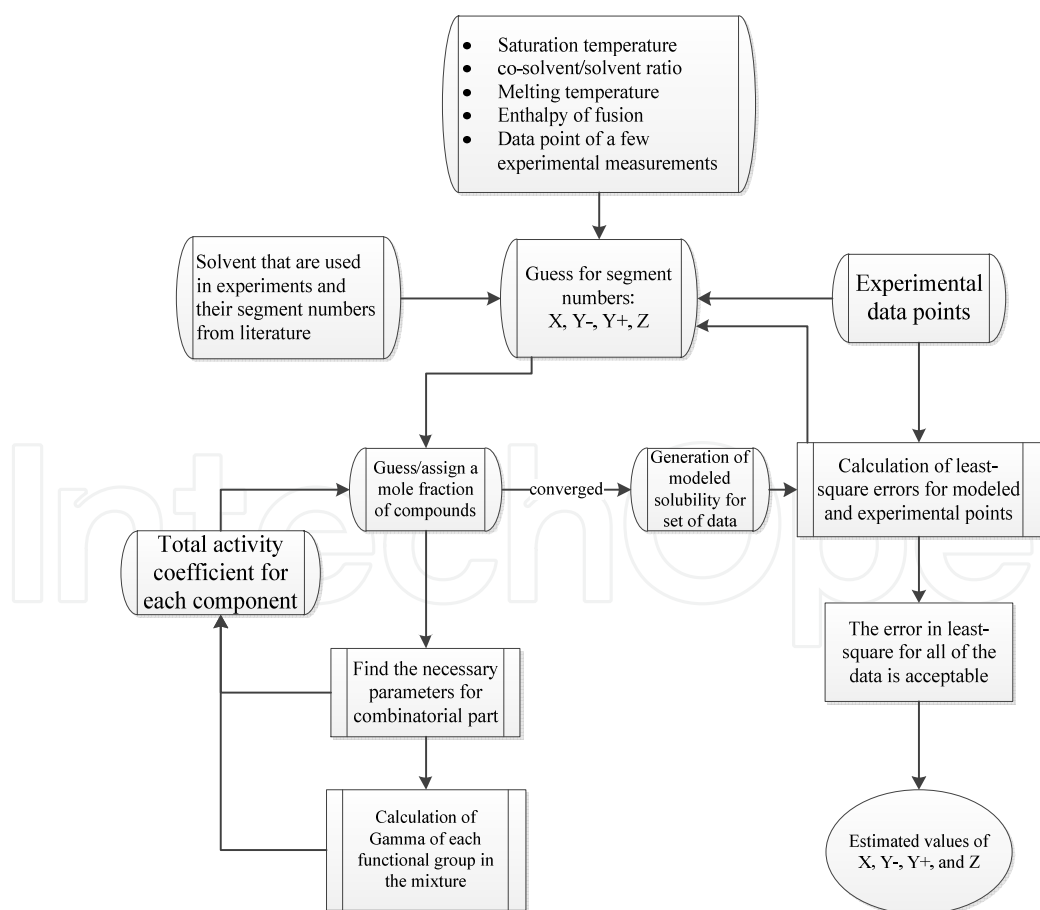

Figure 4. Algorithm flowchart for parameter estimation using NRTL-SAC model.

It is worth noting that the main difference in Figures 3 and 4 is the use of parameter estimation method for the calculation of the NRTL-SAC parameters, while for the UNIFAC model, the calculation is straightforward. Once the parameters (here, the segment numbers) are found, then they could be used for validation against other experimental data.

4. Application of solution thermodynamics in industry

One of the main applications of the thermodynamic models is in the chemical industries which use solvent (or their mixtures) [19-22]. Two cases of the vapor-liquid equilibrium of common industrial solvent systems are discussed here.

4.1. VLE study of two binary azeotropic systems

There are some solvents within the chemical and pharmaceutical industries which are of importance to study, such as ethanol, isopropyl alcohol, and water in addition to some aromatic components, such as benzene and toluene. These solvents are sometimes used as additives to other valuable chemicals to maintain some performances or enhance their physical or chemical properties. The systems that form azeotropes make the distillation process calculations complex. As a result, having an accurate knowledge about the phase behaviors of such systems are important (such as toluene-ethanol or toluene-isopropyl alcohol). The four case studies of the VLE data for the mentioned azeotropic binary systems were used in a study by Chen et al. [24]. The operating pressures were at four distinct levels for the calculations. They have used three equations of state (from the correlative group) to fit the experimental data with the model outputs and found the necessary binary interaction parameters. Based on their work, the estimated parameters were found; however, they were not used to further validate the models for other operating conditions. In a study by Sheikholeslamzadeh et al., they used the NRTL-SAC and UNIFAC models to predict these azeotropic systems [20]. Table 1 (from their work) shows average relative deviation for the compositions and temperature for the two systems. It is seen from the table that the deviations from the experimental data in the vapor phase mole fractions are almost one-third relative to the NRTL-SAC model. On the other hand, the relative deviation for the saturation temperature is higher using the NRTL-SAC model. Another fact from this finding is that the deviations for both binary systems when the NRTL-SAC model is used are nearly the same. This is not the case when utilizing the UNIFAC model to predict the systems' behaviors. The final conclusion from Table 1 would be the better predictive capability of the UNIFAC model for light alcohols than the heavier ones.

| | | %ARD (equilibrium temperature and vapor-phase mole fractions) | | | | | |
| | | System 1 | | | System 2 | | |
Thermodynamic model	Pressure, (kPa)	Temperature, (K)	Ethanol	Toluene	Temperature, (K)	2-Propanol	Toluene
NRTL-SAC	101.3	0.46	5.14	9.81	0.41	4.91	7.19

Thermodynamic model	Pressure, (kPa)	%ARD (equilibrium temperature and vapor-phase mole fractions)					
		System 1			System 2		
		Temperature, (K)	Ethanol	Toluene	Temperature, (K)	2-Propanol	Toluene
	201.3	0.64	7.63	9.72	0.66	7.76	9.77
UNIFAC	101.3	0.10	1.08	3.15	0.22	2.04	3.87
	201.3	0.27	2.03	3.47	0.36	3.43	5.40

Table 1. The results of the prediction using the NRTL-SAC and UNIFAC models with the experimental values of Chen et al. [20,24]
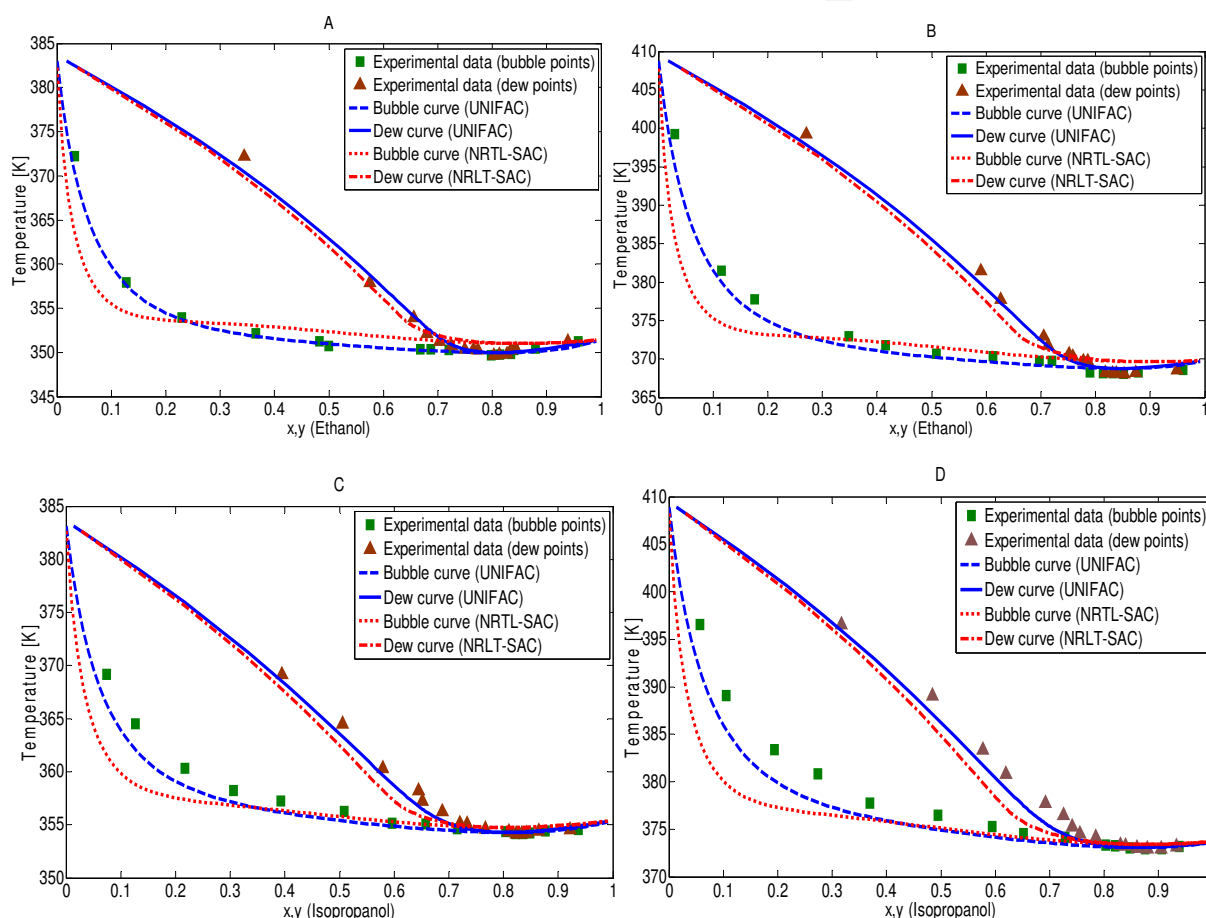

Figure 5. The results for the systems (A) ethanol-toluene at 101.3 kPa, (B) ethanol-toluene at 201.3 kPa, (c) isopropyl alcohol-toluene at 101.3 kPa, and (D) isopropyl alcohol-toluene at 201.3 kPa [20].

4.2. VLE study of a ternary system

With the increasing prices of fossil fuels, the demand to obtain alternatives has received much attention. One of the candidates for this purpose is ethanol, as it decreases the air pollution when blended to the conventional fossil fuels and thereby, increasing the performance of burning fuels within the vehicle engines. It would also be cost-effective when adding other

low-price additives to the fuel. The higher the purity of the alcohol being used as additive, the better the performance of the fuel. It was found that the addition of glycols to the mixture containing alcohols and water can improve the separation processes and utilize less energy to perform the process [25,26]. The experimental data containing the ternary system of ethylene glycol-water-ethanol and the performance of separation by varying the glycol concentration were performed by Kamihama et al. [27].
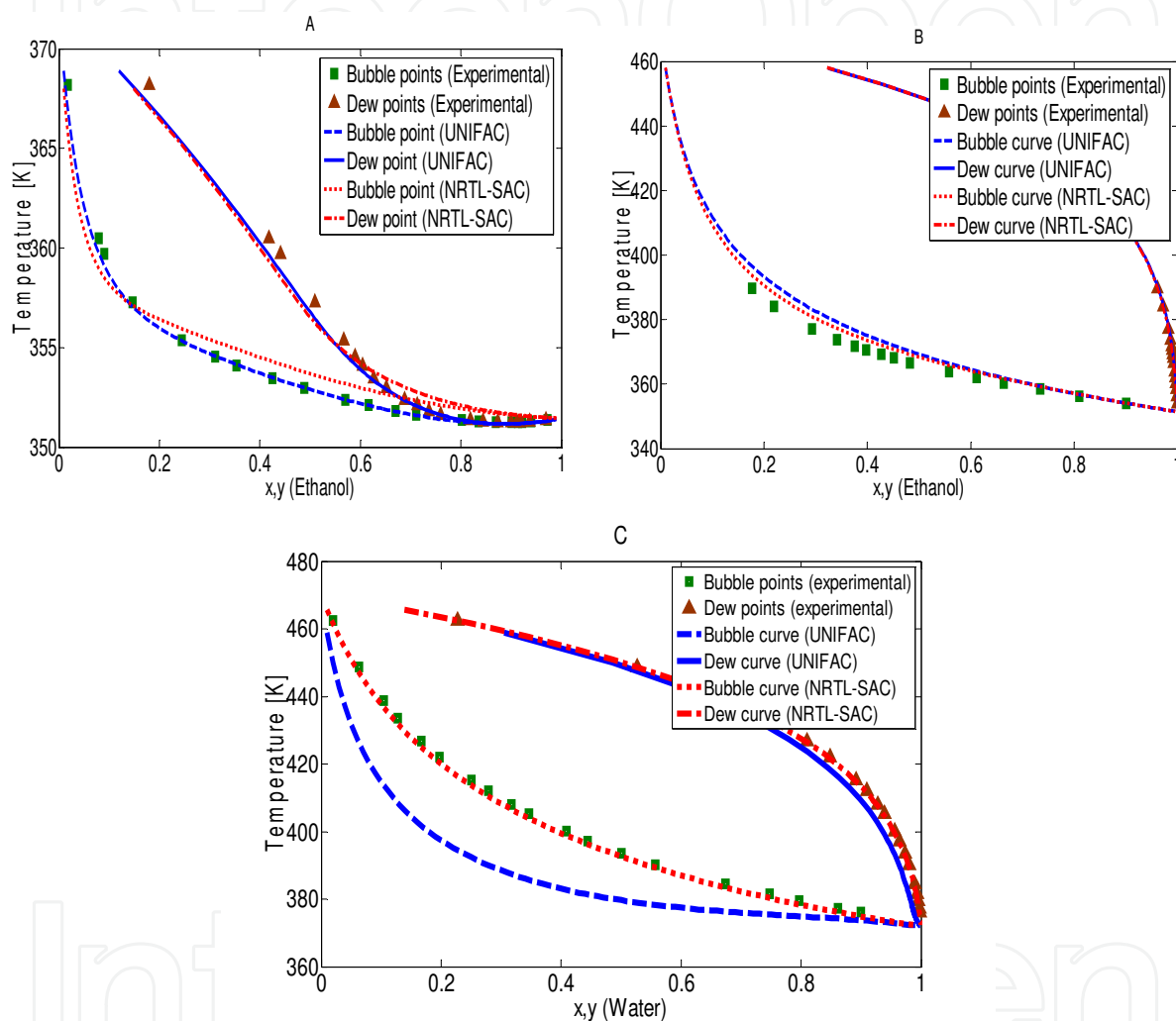

Figure 6. VLE diagrams of the systems (A) ethanol-water, (B) ethanol-ethylene glycol, and (C) water-ethylene glycol at 101.3 kPa [20].

They performed binary system experiments for each of the pairs in the ternary system. It was found that glycol can move the azeotrope point and therefore, enhance the separation process. Sheikholeslamzadeh et al. have used both the NRTL-SAC and UNIFAC models to perform phase calculations and assess the capacity of the mentioned models in the prediction of binary and ternary systems containing glycol and alcohols [20].

From Figure 6, the UNIFAC model has the capability of predicting the system of ethanol-water, perfectly. However, this is not the case for the systems using ethylene glycol-water and

ethylene glycol-ethanol. On the other side, the NRTL-SAC model gave satisfactory results for all three binary system, specifically, the systems that contain ethylene glycol. Also from Figures 6B and 6C, it can be seen that for nonazeotropic systems, the NRTL-SAC model could best show capability in prediction comparable to the UNIFAC model. The UNIFAC model could best locate the azeotrope point and the VLE behavior of the ethanol-water system.

For the ternary system of solvents consisting of ethanol, water, and ethylene glycol, Kamihama et al. [27] conducted the vapor-liquid measurements at a pressure 101.3 kPa. In order to use the correlative models (such as Wilson) for the ternary system, the binary interaction parameters should be known for each pair of components in the mixture at that temperature and pressure. They used this method to find the ternary behavior of the system. Sheikholeslamzadeh et al. used the NRTL-SAC model with the four conceptual segments of each solvent, which were already accessible in the literature [9, 18]. The results showed the high performance of the NRTL-SAC model in the prediction of this ternary system. The bubble point temperature as well as the vapor and liquid compositions could be estimated fairly well with the NRTL-SAC model. The predicted results are shown in Figure 7, giving the vapor phase mole fractions of the three species as well as the mixture temperature. It is apparent that the deviation from the experimental data for the case of water and ethanol in this ternary mixture is almost zero. The UNIFAC model could not match the experimental data as well as the NRTL-SAC model. For the ethylene glycol, as the experimental concentrations of the vapor phase are very small, with the inclusion of errors of the experimentation, the NRTL-SAC model could also give satisfactory results. Finally, for the bubble point temperature, Figure 7 illustrates the perfect predictions of the NRTL-SAC model compared with the UNIFAC model. The average relative deviation for the whole set of experimental data on the saturated temperature from the NRTL-SAC model is one-third that from the UNIFAC model.
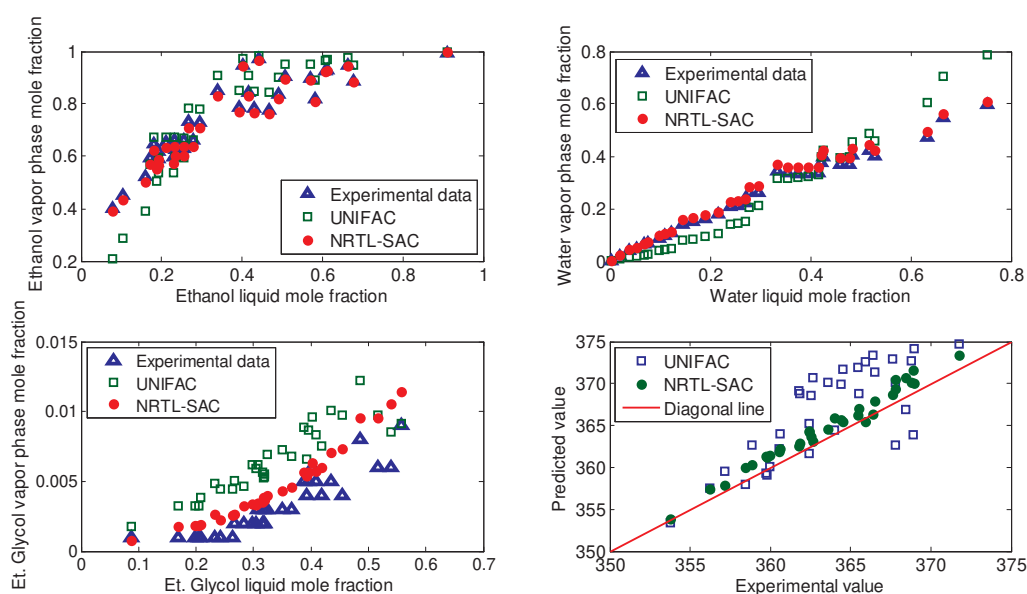

Figure 7. Vapor phase mole fractions of ethylene glycol (bottom-left), ethanol (top-left), water (top-right), and the bubble temperature (bottom-right) prediction from the NRTL-SAC model [20].

5. Conclusion

There are several equations of state that describe the phase behavior of chemical components of a system at various temperatures, pressures, and compositions. From these models, the first group which needs various experimental data to predict the system behavior at other conditions is not very attractive. Instead, the second group (predictive models) is based on the activity coefficients that are found from the molecular structures with a few experimental data. In this chapter, the capacity of handling two binary and ternary systems of solvents using those predictive models was assessed. The NRTL-SAC and UNIFAC models were chosen for the modeling of those systems. The NRTL-SAC model showed relative advantage over the UNIFAC model in almost all cases, except for the systems containing light alcohols with water. The preference of using NRTL-SAC is due to its simplicity compared to the UNIFAC. If the four segment parameters of a specific component are known, then they can be set as a unique value for that component irrespective of the mixture conditions. This way, various operating conditions can be defined and the phase behavior of the components can be predicted accurately and rapidly. One of the main advantages of the NRTL-SAC model is that it can be written in various computer programming languages to be used in process simulation analysis. One good example of such work can be found in Sheikholeslamzadeh et al. [19,21]. They used the NRTL-SAC model to find the solid-liquid equilibrium for three pharmaceuticals. Then, the parameters they optimized for the given pharmaceuticals were used further to perform a solvent screening method. This method is really costly and time-consuming in industry. The authors [19,20] have developed an algorithm to find the best combination of solvents, temperatures, and pressures for the best yield of pharmaceutical production using the NRTL-SAC model. In conclusion, the NRTL-SAC model and other similar ones open a new window for the engineers and scientists to have a wider, more accurate, and rapid predictions of the solubility of active pharmaceuticals in mixtures of solvents.

Author details

Ehsan Sheikholeslamzadeh and Sohrab Rohani[*]

*Address all correspondence to: Srohani@uwo.ca

Department of Chemical and Biochemical Engineering, Western University, Canada

References

[1] J. M. Smith, Hendrick C Van Ness, Michael Abbott. Introduction to Chemical Engineering Thermodynamics, McGraw-Hill Science, Sixth Edition, 2004.

[2] J. M. Prausnitz, R. N. Lichtenthaler, E. G. Azevedo. Molecular Thermodynamics of Fluid-Phase Equilibria, Prentice-Hall International Series, Third Edition, 1999.

[3] M. Margules, Uber die Zusammensetzung der gesattigten Dampfe von Mischungen, Sitzungsber Akad. Wiss. Wien. Math. Naturw. Klasse, II, vol. 104, pp. 1243-12778, 1985.

[4] G. M. Wilson. Vapor-liquid equilibirium. XI. A new expression for the excess free energy of mixing. J. Am. Chem. Soc., vol. 86, pp.127-130, January 1964.

[5] J. J. van Laar. Sechs Vortragen uber das thermodynamische potential (Six Lectures on the Thermodynamic Potential). Braunschweig, Fried. Vieweg & Sohn, 1906.

[6] H. Renon and J. M. Prausnitz. Local compositions in thermodynamic excess functions for liquid mixtures. AIChE J., vol. 14, pp. 135-144, January 1968.

[7] D. S. Abrams and J. M. Prausnitz. Statistical thermodynamics of liquid mixtures: A new expression for the excess Gibbs energy of partly or completely miscible systems. AIChE J., vol. 21, pp. 116-128, January 1975.

[8] A. Fredenslund, R. L. Jones, J. M. Prausnitz. Group contribution estimation of activity coefficients in nonideal liquid mixtures. AIChE J., 1975, 21 (6), 1086-1099.

[9] C. C. Chen, and Y. Song. Solubility modeling with a non-random two-liquid segment activity coefficient model. Ind. Eng. Chem. Res. 2004, 43, 8354-8362.

[10] Van Winkle M. Effect of polar components on the relative volatility of the binary system n-hexane-benzene. J. Chem. Eng. Data 8 (1963) 210-214.

[11] H. Matsuda, K. Kaburagi, K. Kurihara, K. Tochigi, K. Tomono. Prediction of solubilities of pharmaceutical compounds in water + co-solvent systems using an activity coefficient model. Fluid Phase Equilibria 290 (2010) 153-157.

[12] D. Tassios. 62nd Annual Meeting, American Institute of Chemical Engineers, Washington, DC, November 1969.

[13] J. Gmehling and U. Onken. Vapor-Liquid Equilibrium Data Collection. DECHEMA, Franckfurt/Main, 1977.

[14] U. Weidlich, J. Gmehling. A modified UNIFAC model. 1. Prediction of VLE, hE. Ind. Eng. Chem. Res., 1987, 26 (7), pp 1372-1381.

[15] M. Sheikhzadeh, S. Rohani, M. Taffish, S. Murad. Solubility analysis of buspirone hydrochloride polymorphs: Measurements and prediction. Int. J. Pharm. 338 (2007) 55-63.

[16] S. Gracin, T. Brinck, and A. C. Rasmuson. Prediction of solid organic compounds in solvents by UNIFAC. Ind. Eng. Chem. Res. 2002, 41, 5114-5124.

[17] A. T. Kan and M. B. Tomson. UNIFAC prediction of aqueous and nonaqueous solu-
      bilities of chemicals with environmental interest. Environ. Sci. Technol. 1996, 30,
      1369-1376.

[18] C. C. Chen and Y. Song. Extension of nonrandom two-liquid segment activity coeffi-
      cient model for electrolytes. Ind. Eng. Chem. Res. 2005, 44, 8909-8921.

[19] Ehsan Sheikholeslamzadeh and Sohrab Rohani. Solubility prediction of pharmaceuti-
      cal and chemical compounds in pure and mixed solvents using predictive models.
      Ind. Eng. Chem. Res. (2011) 51, 464-473.

[20] Ehsan Sheikholeslamzadeh and Sohrab Rohani. Vapour-liquid and vapour-liquid-
      liquid equilibrium modeling for binary, ternary, and quaternary systems of solvents.
      Fluid Phase Equilibria, 333, 97-105.

[21] Ehsan Sheikholeslamzadeh, Chau-Chyun Chen, and Sohrab Rohani. Optimal solvent
      screening for the crystallization of pharmaceutical compounds from multisolvent
      systems. Ind. Eng. Chem. Res. (2012) 51, 13792-13802.

[22] Ehsan Sheikholeslamzadeh and Sohrab Rohani. Modeling and optimal control of sol-
      ution mediated polymorphic transformation of l-glutamic acid. Ind. Eng. Chem. Res.
      (2013) 52, 2633-2641.

[23] Stichlmair, J., Fair, J., Bravo, J. L. Separation of azeotropic mixtures via enhanced dis-
      tillation. Chem. Eng. Prog. B 1989, 85, 63.

[24] Chen, R., Zhong, L., Xu, C. Isobaric vapor-liquid equilibrium for binary systems of
      toluene + ethanol and toluene + isopropanol at (1013.3, 121.3, 161.3, and 201.3) kPa. J.
      Chem. Eng. Data. 2011, 57, 155.

[25] Pequenín, A., Carlos Asensi, J., Gomis V. Isobaric vapor–liquid–liquid equilibrium
      and vapor–liquid equilibrium for the quaternary system water–ethanol–cyclohexane
      –isooctane at 101.3 kPa. J. Chem. Eng. Data, 2010, 55, 1227.

[26] Pequenín, A., Carlos Asensi, J., Gomis V. Experimental determination of quaternary
      and ternary isobaric vapor-liquid-liquid equilibrium and vapor-liquid equilibrium
      for the systems water-ethanol-hexane-toluene and water-hexane-toluene at 101.3
      kPa. J. Chem. Eng. Data, 2010, 56, 3991.

[27] Kamihama, N., Matsuda, H., Kurihara, K., Tochigi, K., Oba, S. Isobaric vapor-liquid
      equilibria for ethanol + water + ethylene glycol and its constituent three binary sys-
      tems. J. Chem. Eng. Data. dx.doi.org/10.1021/je2008704.