



Chapter Four

METHODOLOGY

The overview of the entire procedure for the parameter estimation and the calculation of VLE using UNIFAC Dortmund model are shown in the figures below. The first step involves the compilation of experimental VLE data of 19 binary, ternary and quaternary systems with and without salt from the literature. For some systems, the liquid compositions from their experimental VLE data were converted in terms of actual liquid mole fraction x_i , which is required for the parameter estimation and application. Next is the application of the existing parameters from the DDBST in the UNIFAC Dortmund model. This is to observe whether these parameters can predict VLE that are close to the experimental VLE of binary and ternary systems without salt. This is followed by series of parameter estimation processes in which new parameters were obtained and some existing parameters in the literature were revised. To obtain the new parameters, specifically the binary interaction parameters for $\text{CNH}_2 - \text{OH}$, initial guesses of the parameters were introduced into the program and iterated until the new parameters are calculated. For the revision of parameters for NaCl and KCl, the same procedure is applied except that the initial guesses are provided by the existing parameters from the literature. The new and revised parameters were applied in the calculation of VLE of systems containing the required binary interactions. For the new parameters, these were applied for ethanol + water + TRIS system to verify if they can predict the VLE with good agreement with the experimental VLE of the system. For the revised parameters, they were used for systems containing NaCl and KCl, which should provide better predictions at VLE than the existing parameters from the literature. If good prediction results are not obtained using the new and revised parameters, then both are further refined and iterated from the parameter estimation process.

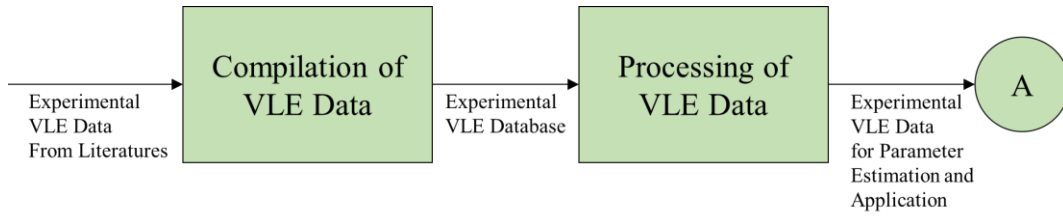


Figure 4.1. Compilation and Processing of VLE Data

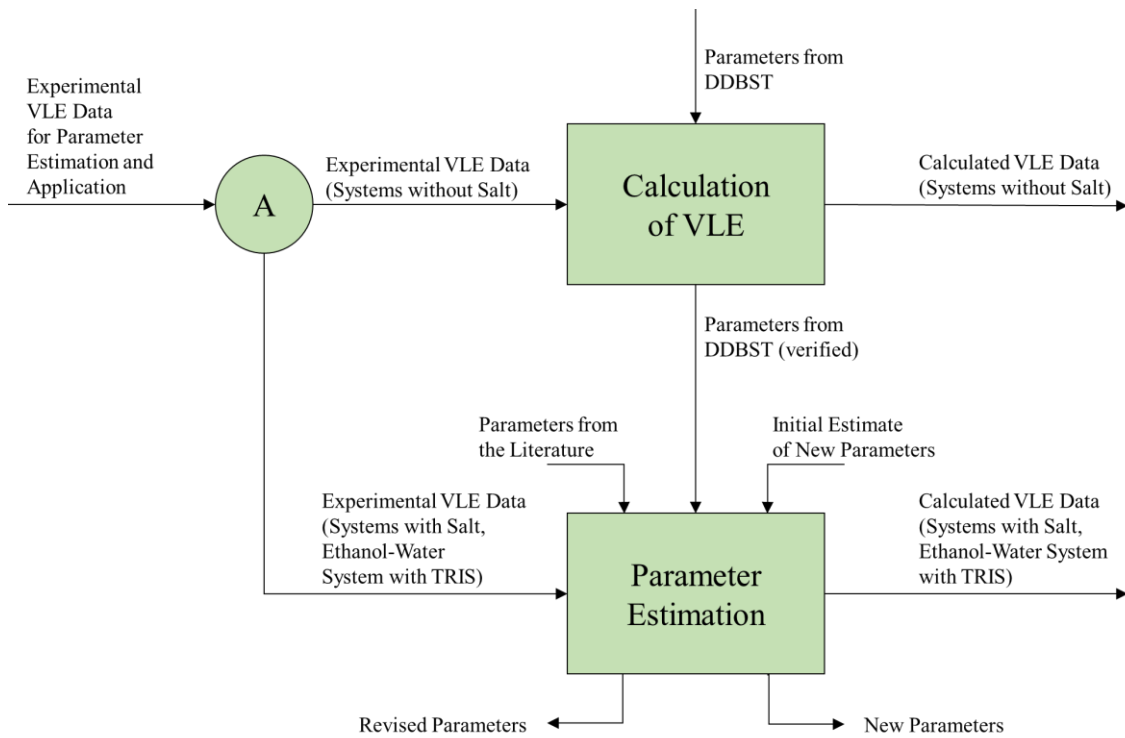


Figure 4.2. Parameter Estimation and Calculation of VLE Data Using the UNIFAC Dortmund Model



4.1 Compilation of VLE Data

The VLE data and the classification of 19 binary, ternary and quaternary systems with and without salt that have been obtained from the literature are as follows:

Table 4.1. VLE Data Collection of Various Binary, Ternary and Quaternary Systems

System No.	System	Total Pressure (kPa)	Reference
1	Water (2) and NaCl (8)	14 - 91.2	Faggion et. al. (2016)
2	Water (2) and KCl (9)	14 - 91.2	
3	Ethanol (1), Glycerol (3) and NaCl (8)	91.2	
4	Water (2), Glycerol (3) and NaCl (8)	91.2	
5	Water (2), Glycerol (3) and KCl (9)	91.2	
6	Ethanol (1), Water (2), Glycerol (3) and NaCl (8)	91.2	
7	Ethanol (1), Water (2), 1-Butanol (5) and NaCl (8)	101.325	Tan & Gan (2005)
8	Ethanol (1), Water (2), 1-Butanol (5) and KCl (9)	101.325	
9	Ethanol (1), Water (2), 1-Propanol (6) and NaCl (8)	101.325	Tan et. al. (2005)
10	Ethanol (1), Water (2), 1-Propanol (6) and KCl (9)	101.325	
11	Ethanol (1), Water (2), 2-Propanol (7) and NaCl (8)	101.325	Tan et. al. (2004)
12	Ethanol (1), Water (2), 2-Propanol (7) and KCl (9)	101.325	
13	Ethanol (1) and Glycerol (3)	14 - 91.2	Faggion et. al. (2016)
14	Ethanol (1), Water (2) and 1-Butanol (5)	101.325	Tan & Gan (2005)
15	Ethanol (1), Water (2) and 1-Propanol (6)	101.325	Tan et. al. (2005)
16	Ethanol (1), Water (2) and 2-Propanol (7)	101.325	Tan et. al. (2004)
17	Ethanol (1), Water (2) and Glycerol (3)	101.3	Pla-Franco et. al. (2013)
18	Ethanol (1), Water (2) and Ethylene Glycol (4)	101.3	
19	Ethanol (1), Water (2) and TRIS (10)	101.325	Bungay et. al. (2011)

Table 4.2. Classification of Systems Considered in This Study

System Classification	Type (Based on Number of Components)	Systems Included
Systems with Salt (12 systems)	Binary	1, 2
	Ternary	3, 4, 5
	Quaternary	6, 7, 8, 9, 10, 11, 12
Systems without Salt (7 systems)	Binary	13
	Ternary	14, 15, 16, 17, 18, 19



4.2 Processing of VLE Data

For systems containing glycerol whose liquid compositions are expressed in mass fractions, these data were converted in terms of the liquid mole fraction, x_i , using the molecular weight values of the components.

For quaternary systems that are saturated with salt, the only concentration data in the liquid phase provided are the mole fractions of three solvents in salt-free basis, hence requiring further calculation to obtain the actual liquid mole fraction of all the components. To obtain the mole fraction of salt in liquid phase for each data, the given liquid mole fraction of water in salt-free basis was converted to mole fraction of salt using the solubility of the salt in water at the specified temperature. In this calculation, the liquid mole fractions of organic solvents were not anymore considered because of the negligible solubility of NaCl and KCl in these solvents. On the other hand, for the determination of the actual liquid mole fraction of three solvents for each data, the following formula was used:

$$x_{ij} = (1 - x_s) \frac{x_{ij}'}{\sum_{i=1}^c x_{ij}'} \quad (4.1)$$

where x_{ij} is the liquid mole fraction in component i of data j , x_s is the mole fraction of the salt, x_{ij}' is the liquid mole fraction in component i of data j in salt-free basis and c is the number of solvent components in a given system.



4.3 Application of Existing Parameters from the DDBST for Binary and Ternary Systems Without Salt

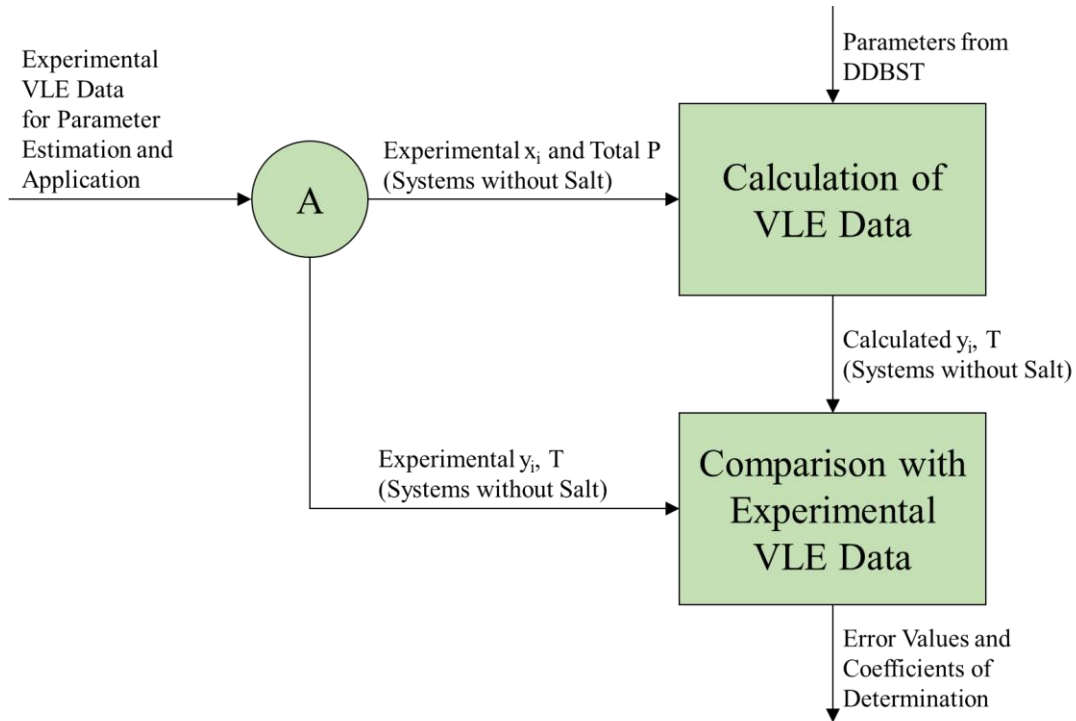


Figure 4.3. Application of Parameters from the DDBST for the Calculation of VLE Data

In this procedure, the parameters from the DDBST were applied first for the calculation of VLE of binary and ternary systems without salt. This is to test whether the parameters from the DDBST can predict VLE data that are close to the experimental VLE data before applying them to other systems. This requires the following important steps: calculation of VLE data using the existing parameters and comparison of the calculated VLE data with the experimental VLE data, which are described in Figure 4.3. The detailed discussion of the following steps are as follows:

4.3.1 Calculation of VLE Data Using the Parameters from the DDBST

The calculation of VLE data for binary and ternary systems using the parameters from the DDBST requires a program that contains the equations for Modified Raoult's Law and the UNIFAC Dortmund model. Thus, this program was created in MATLAB® version 2014a. Then the experimental total pressure, P , and liquid mole fractions, x_i , of binary and



ternary systems without salt were included in the program together with the existing parameters from the DDBST. Since the unknown variables in the calculation of VLE are the temperatures, T , initial guesses for the temperature were defined, where the experimental T values were used as the initial guesses.

The algorithm for the calculation of VLE data is shown in Figure 4.4. Once the program is started, it substitutes the existing parameters from the DDBST, the experimental P and x_i values at VLE, and the initial guesses for the temperatures to obtain the activity coefficient values using the UNIFAC Dortmund model and consequently, the vapor mole fraction values using the Modified Raoult's Law. Then the program calculates the total sum of the vapor mole fractions y_i of the components, which should be equal to 1. If this constraint is not satisfied, then the program updates the initial guesses for the temperature and the previous calculations are repeated. Once the constraint $\sum y = 1$ is satisfied, the resulting T and y_i values at this point are taken as the calculated VLE data for the binary and ternary systems without salt using the existing parameters from the DDBST. The calculated VLE data were verified using Microsoft Excel 2019.

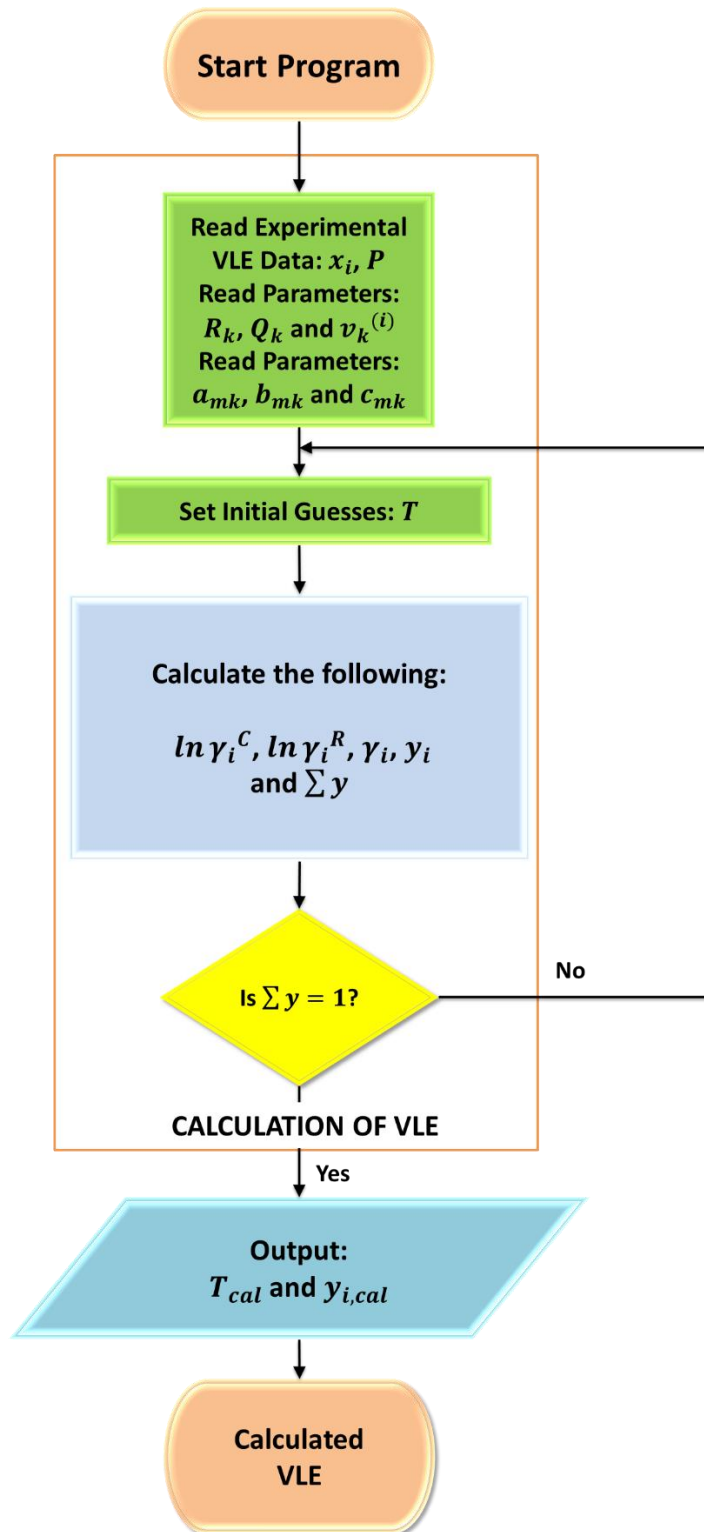


Figure 4.4. Algorithm for the Calculation of VLE Using the UNIFAC Dortmund Model



4.3.2 Comparison of Experimental and Calculated VLE Data of Binary and Ternary Systems Without Salt

The calculated VLE data obtained for the binary and ternary systems without salt was compared to the experimental VLE data of the same systems which are found in the literatures. The comparison was performed by calculating the error values which are the following: standard error of regression (SE), root mean square error (RMSE), average absolute deviation (AAD) and percent average absolute relative deviation (%AARD). The formulas are provided as follows:

$$SE = \sqrt{\frac{(q_j^{exp} - q_j^{cal})^2}{n - p}} \quad (4.2)$$

$$RMSE = \sqrt{\frac{(q_j^{exp} - q_j^{cal})^2}{n}} \quad (4.3)$$

$$\%AARD = \frac{1}{n} \sum_{j=1}^n \left| \frac{q_j^{exp} - q_j^{cal}}{q_j^{exp}} \right| \times 100\% \quad (4.4)$$

$$AAD = \frac{1}{n} \sum_{j=1}^n |q_j^{exp} - q_j^{cal}| \quad (4.5)$$

where the variable q is a thermodynamic property which can either be temperature (T) or vapor mole fraction of a particular component (y_i), q_j^{exp} is the experimental value at VLE, q_j^{cal} is the calculated value from the model, also at VLE, n is the number of data points in a given system and p is the number of parameters calculated for a given system. The calculated VLE data is considered to be in good agreement with the experimental VLE data if low error values are obtained, particularly the SE since it is a metric specifically used for modelling or regression. This will also show the acceptability of the existing parameters from the DDBST to predict VLE data.



Other metrics for the goodness of fit such as mean squared error (MSE), sum of squared errors (SSE), coefficient of determination (r^2) and adjusted coefficient of determination (r^2_{adj}) were calculated as follows:

$$SSE = \sum (q_j^{exp} - q_j^{cal})^2 \quad (4.6)$$

$$MSE = \frac{(q_j^{exp} - q_j^{cal})^2}{n} \quad (4.7)$$

$$r^2 = 1 - \frac{SSE}{\sum (q_j^{exp} - \bar{q}_j^{exp})^2} \quad (4.8)$$

$$r^2_{adj} = 1 - \frac{(1 - r^2)(n - 1)}{n - 2} \quad (4.9)$$

4.4 Estimation of New Parameters for $CNH_2 - OH$ Interaction

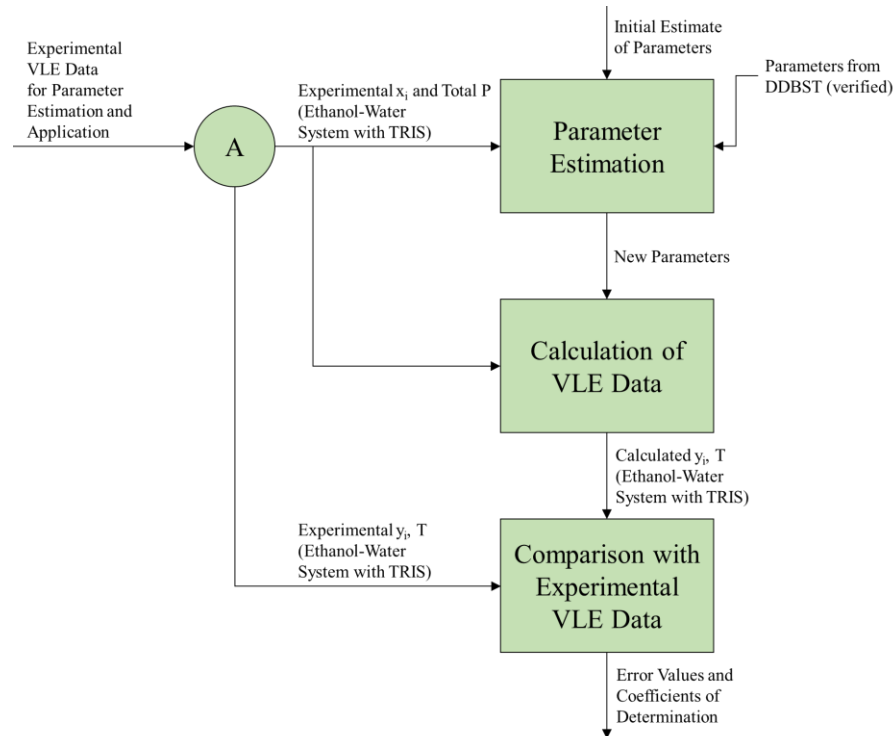


Figure 4.5. Estimation and Application of New UNIFAC Dortmund Parameters for the $CNH_2 - OH$ Interaction



In this procedure, new UNIFAC Dortmund parameters for the $\text{CNH}_2 - \text{OH}$ interactions were obtained, which will be useful for the calculation of VLE of ethanol, water and TRIS system. As shown in Figure 4.5, the calculation of new parameters involves the following steps: parameter estimation, calculation of VLE using the calculated parameters and comparison of the calculated VLE with the experimental VLE. The parameter estimation process involves the minimization of the objective function to obtain the new parameters. The calculation of VLE data and its comparison with the experimental VLE data are similar to the previous section, except that these were performed for the ethanol, water and TRIS system as mentioned. The following steps are discussed as follows:

4.4.1 Parameter Estimation Process for the Calculation of New Parameters

Two subprograms were created: one subprogram contains the equations for the Modified Raoult's Law and the UNIFAC Dortmund model which is similar to that of the previous section, and another contains the error or the objective function which is to be minimized in the parameter estimation. The objective function is defined by the equation:

$$OF = \sum_{j=1}^n \sum_{i=1}^c \left(\frac{y_{ij}^{exp} - y_{ij}^{cal}}{y_{ij}^{exp}} \right)^2 \quad (4.6)$$

where y_{ij}^{exp} is the experimental vapor mole fraction at VLE, y_{ij}^{cal} is the calculated vapor mole fraction using the model, c is the number of solvent components in the given system, and n is the number of data points in that system. This objective function was chosen since the minimization of the difference between experimental the and calculated vapor mole fraction values also minimizes those of the temperatures and activity coefficient values.

An optimization program was created which contains the experimental P and x_i of ethanol, water and TRIS system, the initial guesses for the temperatures, and more importantly, the initial guesses for the parameters which are already the unknown variables in this process. The allowable range for the initial guesses of the temperatures and parameters were also set into the program to avoid possible errors in the iteration process. Specifically, the initial guesses for the temperatures were defined to be within 5 K from the experimental



temperatures, while those for the parameters were constrained within -0.1 and 0.1, which is based on the current minimum and maximum values of c_{mk} in the parameter matrix in the DDBST website. The initial guesses for both the temperatures and parameters were chosen randomly using the *rand* command of MATLAB® 2014a, which were ensured to lie within the allowable range for faster convergence to the solution.

The two subprograms that were previously created were connected into the optimization program in order to begin the parameter estimation process. The parameter estimation process was performed using the command *fmincon* in MATLAB Optimization Toolbox, which uses sequential quadratic programming (SQP) as the optimization method. This is one of the fastest gradient-based optimization methods that applies Lagrangian multipliers and derivatives for every iteration in order to yield the parameters at the minimum value of the objective function.

The parameter estimation process follows the algorithm as presented in Figure 4.6. The process is quite similar to the calculation of VLE data in Figure 4.4; however, an objective function is introduced in order to obtain the best set of parameters that can predict VLE with the least possible deviation from the experimental VLE. For the first step, the experimental P and x_i values and the initial guesses for T and parameters were substituted in the UNIFAC Dortmund and Modified Raoult's Law equations to obtain the activity coefficient and vapor mole fraction values. The same equation $\sum y = 1$ must be satisfied, where the program continues to iterate the initial guesses for the temperature and the parameters until this constraint is satisfied. The resulting temperatures and parameters at $\sum y = 1$ were used for another series of iterations in order to achieve the minimum value of the objective function, which were done in a total of about 2,000 iterations. During the iteration process, the c_{mk} parameters were still constrained to -0.1 to 0.1 as to avoid the parameters from reaching very large values.

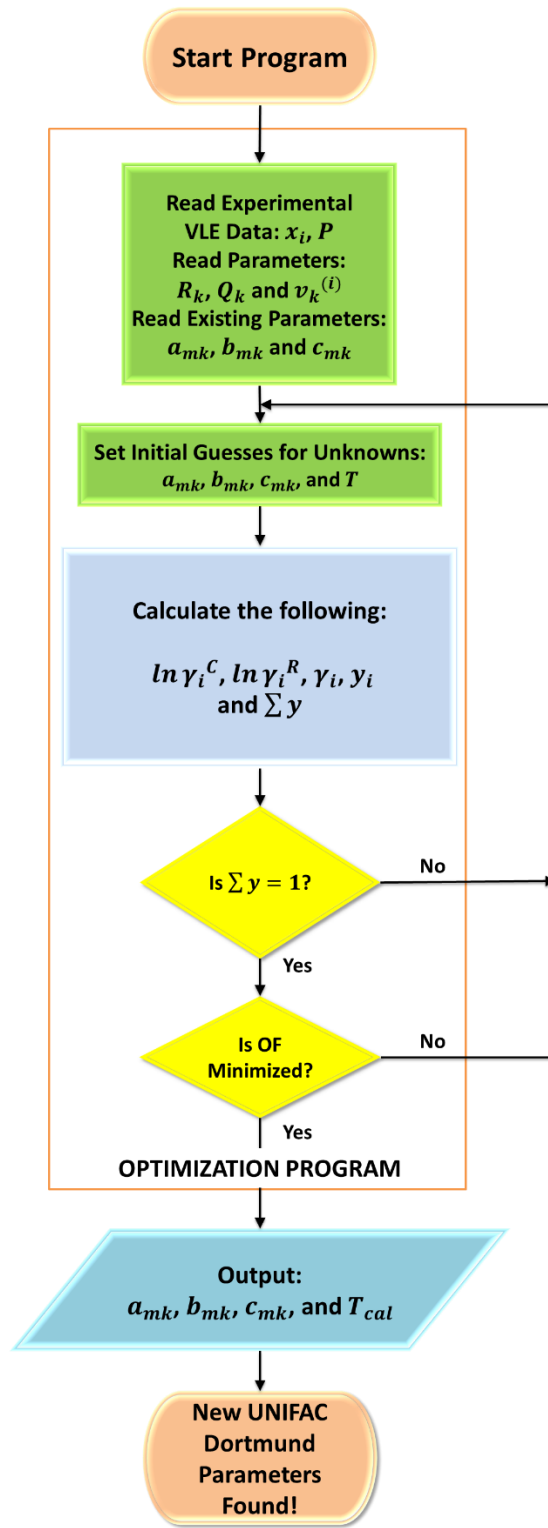


Figure 4.6. Algorithm for the Parameter Estimation



Once the objective function reaches the minimum, the program terminates and the resulting parameters are considered the new UNIFAC Dortmund binary interaction parameters for $\text{CNH}_2 - \text{OH}$ group interaction. For cases where the program terminates due to error, the entire parameter estimation process is repeated for a new set of initial guesses.

4.5 Revision of Parameters for Interactions with NaCl and KCl

For this procedure, the existing binary interaction parameters involving NaCl and KCl from the literature were revised. This is to further minimize the errors in the modelling when these parameters are applied to systems containing NaCl and KCl. The entire procedure also involves parameter estimation, calculation of VLE data, and the comparison with the experimental VLE data as shown in Figure 4.7. However, since both the existing parameters from the literature and revised parameters were applied in the calculation of VLE data of systems with NaCl and KCl, two sets of calculated VLE data were obtained and thus, two sets of error values were calculated. The error values of both the existing and revised parameters were compared in order to see whether the revised parameters are more reliable for the prediction of VLE data than the existing parameters from the literature.

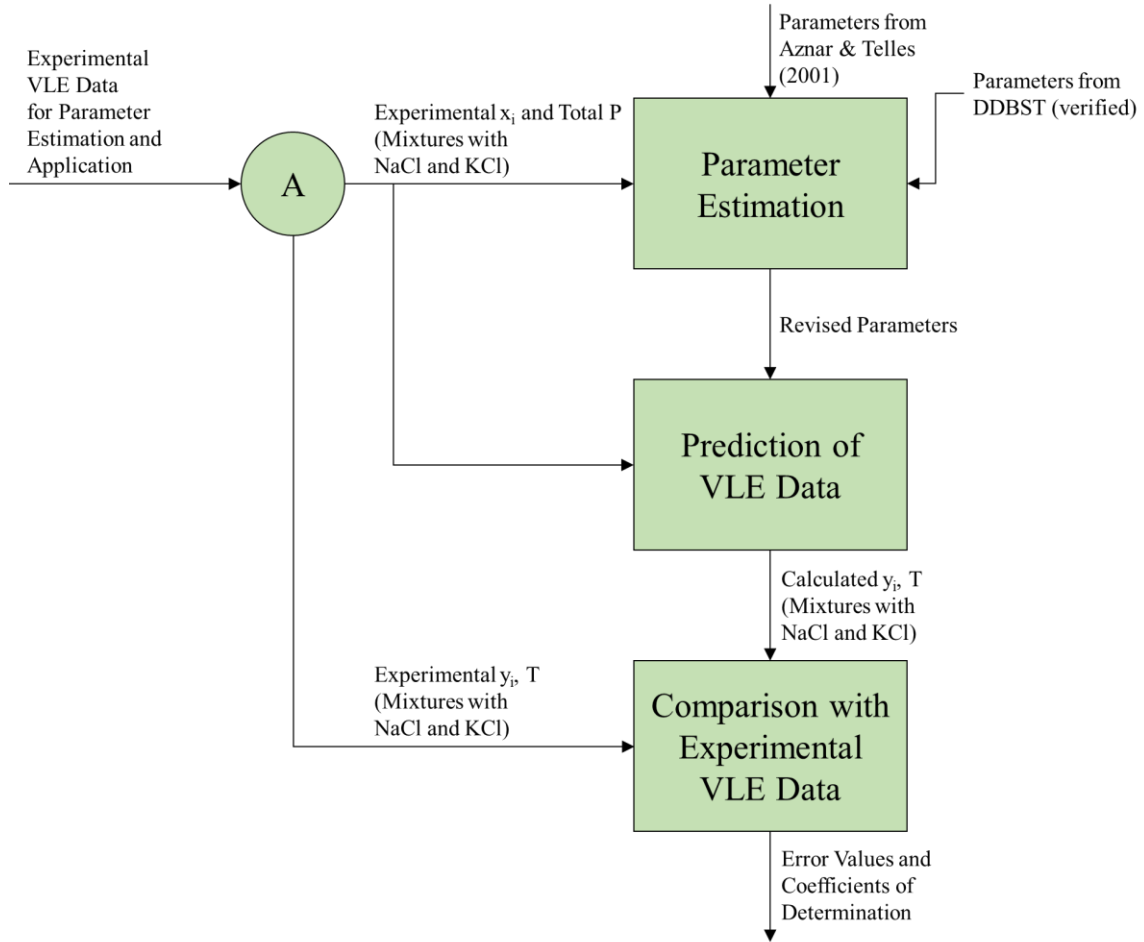


Figure 4.7. Estimation and Application of Revised UNIFAC Dortmund Parameters for the Interactions with NaCl and KCl



4.5.1 Parameter Estimation Process for the Revision of Existing Parameters in the Literature

The same parameter estimation process in the previous discussion was performed, except that the experimental P and x_i values used were from the systems containing NaCl and KCl, and the initial guesses for the parameters were taken from the existing list of parameters of Aznar & Telles (2001) for NaCl and KCl. The initial guesses for the temperature were still obtained by randomization, which are also within 5 K from the experimental temperatures. Since the a_{mk} and b_{mk} are to be revised in this procedure, the initial guesses and the iteration of these parameters were constrained within the bounds provided in Table 4.3, which are also based on the typical minimum and maximum values of the parameters in the parameter matrix. Similarly, once the constraint $\sum y = 1$ is satisfied and the objective function is minimized, the resulting parameters are considered the revised UNIFAC Dortmund binary interaction parameters for NaCl and KCl.

Table 4.3. Lower and Upper Bounds for the Parameter Estimation Process

Parameter	Lower Bound	Upper Bound
a_{mk}	-30,000	30,000
b_{mk}	-100	100

4.5.2 Calculation of VLE Data Using the Revised Parameters

Similar to the previous section, after the parameter estimation process, temperature and vapor mole fraction values were also obtained as outputs of the program. These correspond to the calculated VLE data for systems with NaCl and KCl using the revised parameters, which were verified using Microsoft Excel 2019.

Aside from the calculated VLE data using the revised parameters, the calculated VLE data were also obtained for the systems with salt using the existing parameters of Aznar & Telles (2001). Similar process of calculating the VLE in Figure 4.4 was followed. Hence, two sets of calculated VLE data were calculated for systems with NaCl and KCl.



4.5.3 Comparison of Experimental and Calculated VLE Data of Systems Containing NaCl and KCl

The two sets of calculated VLE data, which were calculated using both the parameters of Aznar & Telles (2001) and the revised parameters, were compared to experimental VLE data of the systems with NaCl and KCl. Hence, the SE, RMSE, AAD and %AARD values in the temperature and vapor mole fraction were calculated for both the existing and revised parameters using the formulas in the previous sections.

4.5.4 Comparison of Error Values of the Existing and Revised Parameters

The error values corresponding to the existing and revised parameters were compared with each other in order to see whether the revised parameters provided lower error values for most cases as compared to the existing parameters of Aznar & Telles (2001). If this is not achieved, then the parameters were repeatedly revised until lower error values than the existing parameters are obtained for most cases.