



Chapter Two

REVIEW OF RELATED LITERATURE

2.1 Development of the UNIFAC Dortmund Model

The development of the UNIFAC Dortmund model started from the invention of local composition models to predict the VLE of several systems, followed by the development of the group contribution models to predict the VLE of systems with no available experimental data. The UNIFAC Dortmund model so far is the most comprehensive among the group contribution models, where the list of its parameters found in the UNIFAC Dortmund Parameter Matrix are constantly being revised and extended in order to provide good prediction results for a wide range of systems at VLE. This historical development of the UNIFAC Dortmund model will be shown by the following sections.

2.1.1 Activity Coefficient Models Before UNIFAC

Activity coefficient models are important for the prediction of VLE of a given system. Thus, various researches have developed several models in order to provide good predictions for systems with the minimum possible deviation from the experimental VLE data. One of the first activity coefficient models that have been developed are called local composition models. These models require experimental data in order to estimate the parameters that can be used to predict the VLE of a given system. It is assumed by these models that the parameters vary from one mixture to another since the interactions are unique for each mixture. Currently, the most widely-used local composition models are Wilson, Nonrandom Two-Liquid (NRTL), and Universal Quasi-Chemical (UNIQUAC) models (Wilson, 1964; Renon & Prausnitz, 1968; Abrams & Prausnitz, 1975).

The Wilson model is used for completely miscible mixtures and is used preferably for polar components that are present in nonpolar solvents. This model has two adjustable parameters that corresponds to the specific binary interaction of components. Because of the few parameters required to predict the VLE of a given system and the good prediction results produced by this model for various systems, it is one of the most widely-used activity



coefficient models as of today. However, the model is found not to be suitable for immiscible and partially immiscible mixtures because of its inability to predict the phase split for these mixtures. Also, it provides poor prediction results for strongly nonideal mixtures or mixtures with strong deviations from Raoult's Law. Hence, the application of Wilson model to these mixtures is not recommended.

In order to provide good predictions at VLE for partially miscible mixtures, the NRTL model was developed. This model can also be applied for strongly nonideal mixtures. This model requires three adjustable parameters for the prediction, where the first two are binary interaction parameters, while the remaining parameter describes the nonrandomness of the components in the mixture. The increase in the required parameters to three can be one of the disadvantages of this model, along with its limited applicability for multicomponent systems. Nevertheless, it is still one of the most widely-used models because of its wide range of applicability.

It has been the objective to develop a model that only requires two parameters, which is similar to Wilson equation, while also being applicable to partially immiscible mixtures, which similar to NRTL model. Thus, the UNIQUAC model was developed. This model is comprised of two main components. First is a combinatorial term, which accounts for the size and shape of the components in the mixture, then a residual term, which considers the energy interactions of the components. It is known to be widely applicable for nonelectrolyte systems containing various organic compounds such as hydrocarbons, alcohols, ketones, etc. Since it only requires a few parameters for the prediction of VLE, and it is more applicable to a greater number of systems because of its more realistic assumptions, then it is one of the widely-used activity coefficient models.

In spite of the advantages of these models, they are only applicable for systems with adequate experimental data and not for systems with few or no experimental data available. This gives rise to the development of group contribution models, where the widely-used models are the UNIFAC models.



2.1.2 *From Original UNIFAC to UNIFAC Dortmund Model*

Since the UNIFAC models are classified under the group contribution models, they assume that the interacting species in the mixture are not the whole components themselves but the functional constituent of these components, which are called subgroups. It is also assumed that subgroups behave similarly and consistently regardless of the mixture in which they are present. Thus, the assumptions allow the prediction of VLE of even greater number of systems, even if these systems do not have available experimental data, as long as these systems contain the desired subgroups. This is the advantage of group contribution models as compared to the local composition models previously mentioned.

The original UNIFAC model was first developed by Fredenslund et al. (1975) which, similar to the UNIQUAC model, also contains the combinatorial and residual terms. Like the UNIQUAC model, it is also applicable for wide range of nonelectrolyte systems with organic components, where systems that can be described by this model may have positive or negative deviations from Raoult's Law. This model requires only two parameters in order to predict the VLE of a given system, which is also the advantage of some local composition models discussed. It can also be used for the prediction of LLE, where separate UNIFAC parameters are provided by Magnussen et al. (1981), thus further extending its applicability to partially immiscible mixtures.

However, some limitations and weaknesses are found in using this model. First is its limited application to certain conditions which are only for temperatures 300-400 K and pressures that are low to moderate. Furthermore, it was also observed that the model cannot provide good predictions for asymmetric systems and for systems whose components have totally different size as the components considered in the original paper. Thus, further development of this model is required to extend its application to these kinds of systems.

Hence, two modified UNIFAC models were developed in the same year by Larsen et al. (1987) and Weidlich & Gmehling (1987). Larsen et al. (1987) developed the UNIFAC Lyngby model by eliminating a single term in the combinatorial term of the original UNIFAC and by introducing new temperature-dependent parameters, resulting to a total of six



parameters required in the model. As a result, the applicability of the model increased to other temperature and pressure conditions, as well as to components of different sizes. On the other hand, Weidlich & Gmehling (1987) developed the UNIFAC Dortmund model by changing the formula for the J_i' parameter in the combinatorial term and introducing new temperature-dependent parameters that are less nonlinear as compared to UNIFAC Lyngby model. Using this model, significantly good prediction results were achieved, where the calculated activity coefficients have lower deviations from the experimental values by a factor of 2 as compared to those of the original UNIFAC model (Gmehling et al., 2015). Recently, the application of this model was even extended to ionic liquids, which have been the most promising compounds for the improvement in the distillation process. This shows the superiority of the UNIFAC Dortmund model compared to other activity coefficient models, which, in addition to its wide range of applicability, it can also be used for systems even without experimental data. This is also the reason why the study focuses on the UNIFAC Dortmund model rather than with the other known activity coefficient models.

2.1.3 Development of the UNIFAC Dortmund Parameter Matrix

In order for the UNIFAC Dortmund model to predict the thermodynamic properties of various of systems with great accuracy, acceptable and adequate number of UNIFAC Dortmund binary interaction parameters are required. The list of these parameters is now accessible in The UNIFAC Consortium (2022), where the parameters are summarized in the UNIFAC Dortmund parameter matrix. Currently, the parameter matrix contains 114 main groups, more than 200 subgroups and hundreds of thousands of parameters that have been published and updated by past researchers. This shows the rapid growth of the UNIFAC Dortmund parameters over the years, where the parameter matrix still continues to update and expand as of today.

The first set of parameters have been calculated from the development of the UNIFAC Dortmund model (Weidlich & Gmehling, 1987). The first set contains 6 main groups, 20 subgroups, and 6^2 or 36 binary interaction parameter sets a_{mk} , b_{mk} and c_{mk} provided for all possible main group pairs. This is different from the original UNIFAC model



which can only provide a_{mk} parameters. Furthermore, the volume and surface area parameters R_k and Q_k corresponding for the individual subgroups were also calculated. All the parameters were used for the prediction of thermodynamic properties of various binary systems containing alkanes, alkenes, alcohols, ketones or aromatics. As reported by the authors, the parameters at that time improved the prediction of VLE, activity coefficients at infinite dilution (γ^∞) and binary excess enthalpies (h^E) by at least 70% as compared to the original UNIFAC (Weidlich and Gmehling, 1987). Therefore, the use of parameter set a_{mk} , b_{mk} and c_{mk} instead of just a single parameter a_{mk} by the original UNIFAC model improved the prediction of thermodynamic properties of binary systems. However, predictions for binary systems containing a mixture primary, secondary or tertiary alcohols still require improvement, even though these alcohols are already assigned with separate R_k and Q_k values. To improve the predictions, the idea of assigning distinct binary interaction parameters for each type of alcohol was proposed which will classify them into different main groups. However, this was opposed by the researchers because it is intended that the UNIFAC Dortmund model should only require few parameter inputs to predict the thermodynamic properties of a given system.

The parameter matrix was further updated by Gmehling et al. (1993), where the matrix already contains 45 main groups and 85 subgroups along with hundreds of binary interaction parameter sets for some main group pairs. It was found out that the expansion and revision of the matrix contributes to greater accuracy of the UNIFAC Dortmund model, with improvements of at least 30% for the vapor mole fraction, temperature, and pressure values as compared the original UNIFAC model. The h^E and γ^∞ also improved by a factor of about 2-3. On the other hand, problems are still encountered for systems containing primary, secondary and tertiary alcohols, especially in predicting the liquid-liquid equilibrium (LLE) of partially immiscible systems. Overall, the number of systems that can be described by the UNIFAC Dortmund model have expanded due to the increase in the number of binary interaction parameters, while revisions are still proposed for some parameters.



The further development of the parameter matrix can still be observed in the study of Gmehling (1995) in which 65 main groups and hundreds of parameter sets are already available at that time. In the matrix, new parameters were obtained while some parameters were revised. The new parameters were determined from 2,200 data sets and were used to predict the thermodynamic properties of various systems. The findings show that the error values of VLE, h^E and γ^∞ are reduced by a factor of at least 2 as compared to those of the original UNIFAC. Therefore, the prediction errors in the modelling have been significantly reduced and the UNIFAC Dortmund model can already be used for the prediction of data of even greater number of systems.

In the following years, the revision of parameters in the UNIFAC Dortmund parameter matrix, as well as the addition of main groups and their corresponding parameters, are still being continued by various researches. In the study of Lohmann et al. (2001), the parameters for aromatics with perfluorinated compounds were revised and new parameters for the epoxide main group were introduced. The application of new parameters yielded good predictions to binary systems containing components such as propylene oxide and epoxides of butane. Some weaknesses of the UNIFAC Dortmund model were also noted in this paper such as the application of the model to isomeric compounds and to compounds with strongly electronegative atoms.

Another process of parameter revision and addition were reported by Gmehling et al. (2002) in the following year. Some of the existing main groups such as monoalkylated amides were finally provided with new parameter sets, while the parameters for the main groups dialkylated formamides and dialkylated amides were revised. Good predictions have been obtained using the new and revised parameters.

Currently, the weaknesses of the UNIFAC Dortmund still remain, which include the lack of applicability for mixtures with alcohols of different types of OH groups (Weidlich & Gmehling, 1987), mixtures containing isomers, and mixtures with highly polar components of multiple groups. (Lohmann et al., 2001). For mixtures of alcohols, poor prediction results are still obtained even though separate volume and surface area parameters (R_k and Q_k) were



provided to distinguish the primary, secondary and tertiary alcohols. For mixtures containing isomers, although isomers have different properties in reality, the model cannot capture their distinctions because they have the same composition of functional groups. For mixtures with highly polar components that have multiple groups, a given group may be influenced by the nearby groups present within the component, which is called the proximity effect, especially if the nearby groups are too electronegative. Hence, the binary interaction parameters of the affected group should be altered, which is not accounted by the model, which assumes that the groups have the same binary interaction parameters regardless of their location in any molecule. To address the problems mentioned, new main groups with new binary interaction parameters are introduced, where it is ensured at the same time that the prediction of VLE of any given mixture will not require too many constants or binary interaction parameters.

Because of the existing problems, the UNIFAC Dortmund parameter matrix is constantly extended and revised. In the recent years, the development for the model is about the addition of main groups and parameters for ionic liquids, which are one of the most useful compounds in the separation of components via distillation and extraction processes. As shall be seen in the next section, the recently obtained parameters for ionic liquids can also provide good predictions in the thermodynamic properties of various ionic liquid systems.

2.2 Application of UNIFAC Dortmund Model to Different Types of Systems

The UNIFAC Dortmund model can be applied for a wide range of systems such as the mixtures of solvents, mixtures containing salt, and as mentioned, mixtures with ionic liquids. Typically, good predictions are obtained, but for systems with poor prediction results or prediction is not possible, the binary interaction parameters in the UNIFAC Dortmund Parameter Matrix are being revised and extended by various studies as will be discussed in the next sections.



2.2.1 Application of UNIFAC Dortmund Model to Systems Containing Ionic Liquids

One of the first studies where the UNIFAC Dortmund model has been applied to systems with ionic liquids is found in the study of Kato & Gmehling (2005), where the new main groups imidazolium ([RMIM]⁺) and bis(trifluoromethylsulfonyl)imide ([BTI]⁻) were added, and their corresponding R_k , Q_k , and binary interaction parameters were calculated from their interaction with the common organic compounds. The binary interaction parameters were used in the model to predict the VLE, h^E and γ^∞ values. It was observed that the parameters provided good predictions for VLE, h^E and γ^∞ of various binary systems containing either an alkane, alkene, alcohol or cyclic alkane as the first component, and [RR'IM][BTI] based ionic liquids as the second component. On the other hand, poor results are obtained when the original UNIFAC model was used. This shows the greater reliability of UNIFAC Dortmund model over the original UNIFAC model for the systems considered in this study.

This is followed by the research conducted by Nebig et al. (2007) in which the parameters involving the interaction of main groups [RMIM]⁺ and [BTI]⁻ with the organic compounds from the previous study have been revised. Furthermore, new binary interaction parameters for the same main groups but interacting with ketones and aromatic compounds have been determined. Good predictions have been obtained for the binary mixtures containing either an alkane, aromatic, alcohol or ketone, together with an [RMIM][BTI] based ionic liquid. As mentioned by the authors, additional main groups are targeted for future research such as pyridinium, pyrrolidinium and trifluoromethylsulfonate.

The addition of the pyrrolidinium ([RMPYR]⁺) main group was achieved in the study of Nebig et al. (2009), where new R_k , Q_k , and binary interaction parameters corresponding to this cation were also estimated. Using the new parameters, good prediction results have been obtained for binary and ternary mixtures containing either an alkane or alkene together with a pyrrolidinium-based ionic liquid. At this point, the UNIFAC Dortmund parameter matrix continues to expand further as the new parameters for ionic liquids are calculated.



The main group trifluoromethylsulfonate ([OTF]⁻) was finally introduced in the study of Nebig & Gmehling (2010), in which new R_k , Q_k , and binary interaction parameters were also calculated. The binary interaction parameters for the interactions of [RMIM]⁺ with alkanes and alkenes were also revised in this paper. The predicted results using the new and revised parameters are in good agreement with the experimental data of binary mixtures of alkane/alkene and [RMIM][OTF] based ionic liquids.

The overall status of the UNIFAC Dortmund parameter matrix for ionic liquids was finally reported by Nebig & Gmehling (2011), where the main groups of tetrafluoroborate anion ([BF₄]⁻) and pyridinium cation ([RPY]⁺) were finally included. The parameters of the recently added main groups were estimated and some of the parameters for the existing main groups were also revised. Using the new and revised parameters, the prediction of VLE, h^E and γ^∞ are in accordance with the experimental data of various binary mixtures with ionic liquids. It can be observed in this paper that the UNIFAC Dortmund parameter matrix has expanded significantly because of the addition of several parameters for ionic liquids, hence various systems containing different ionic liquids can finally be correlated using the UNIFAC Dortmund model.

The addition of main groups and parameters for ionic liquids still continued throughout the years, where the new sulfate main group ([RSO₄]⁻) was introduced by Hector et al. (2013). The addition of hexafluorophosphate main group ([PF₆]⁻) into the parameter matrix was also reported in this paper. The paper focused on the estimation of new parameters for [RSO₄]⁻, particularly with its interaction with imidazolium, although the parameters for some existing main groups for ionic liquids were also revised. Another good prediction of the thermodynamic properties were obtained using the new parameters when applied to binary mixtures of alcohol and [RMIM][RSO₄] based ionic liquid.

This is followed by the introduction of the piperidinium main group ([Pip]) in the study of Padaszynski & Domanska (2013). Binary interaction parameters for piperidinium anion, particularly with its interactions with [BTf]⁻, were obtained. Using the new parameters, the prediction of LLE is in good agreement with experimental LLE of binary mixtures



containing [Pip][BTI] based ionic liquids. Since the new parameters were estimated at LLE, the authors reminded to observe great care in applying the new parameters in the prediction of VLE of other systems with piperidinium anions.

Another report on the current status of the UNIFAC Dortmund parameter matrix for ionic liquids was made by Hector & Gmehling (2014). The new binary interaction parameters obtained by Hector et al. (2013) for $[\text{RSO}_4]^-$ and $[\text{PF}_6]^-$ anions were reported in this paper. Furthermore, the existing parameters for $[\text{RMIM}]^+$ and $[\text{BTI}]^-$ main groups, as well as the recently obtained parameters for the main groups $[\text{RPY}]^+$, $[\text{RMPYR}]^+$, $[\text{OTF}]^-$ and $[\text{BF}_4]^-$, were revised. All parameters provided good predictions in the VLE, h^E and γ^∞ of various binary systems containing an ionic liquid and either an alkane or alcohol component. It was reported that there are already a total of 81 binary interaction parameter sets that are available for ionic liquids, showing the great development of the UNIFAC Dortmund parameter matrix for ionic liquids that has been achieved so far (Hector & Gmehling, 2014). According to the authors, in order to further expand the parameter matrix, more experimental data for systems containing ionic liquids are required.

Some of the existing parameters for ionic liquids were simply applied to binary systems for the prediction of thermodynamic properties. This was performed by Bahadur et al. (2015), where the existing parameters for $[\text{RPY}]^+$ and $[\text{BTI}]^-$ main groups were applied to binary systems containing an alcohol and an $[\text{RPY}][\text{BTI}]$ based ionic liquid. For most cases, the existing parameters predicted thermodynamic properties that are in accordance with the experimental data. However, in the case of h^E values, less accurate predictions are observed. Therefore, the authors call for the revision of the existing parameters.

In more recent years, parameter revision and addition are still being performed. It is because as observed from the previous studies discussed, this significantly improves the predictions of the UNIFAC Dortmund model, which also extends its application to greater number of systems. This is the reason why the study focuses on the revision and extension of parameter matrix in the improvement of the UNIFAC Dortmund model.



2.2.2 *Application of UNIFAC Dortmund Model to Systems Containing Solvents*

The success of the UNIFAC Dortmund model is much more observed in its application to various systems containing solvents. It is because throughout the history, most of the existing binary interaction parameters for solvents are already applicable to wide range of systems, hence they do not require significant revisions. Thus, previous studies focused their work mostly on the application of the existing parameters for solvents in the prediction of thermodynamic properties of solvent mixtures. For cases of solvents whose main groups have no available parameters, new parameters are estimated. One example of which is the parameter estimation performed by Nolker & Roth (1997), where new parameters corresponding to the isocyanate main group were determined. It was found out that the newly obtained parameters predicted VLE data are in good agreement with the experimental VLE data of the binary mixtures containing an isocyanate compound and hexane or benzene. Thus, the determination of new parameters for isocyanate has been successful in this study.

For the application of the existing binary interaction parameters to solvent mixtures, one of the first studies concerned with this objective is that of Arce et al. (1997). The aim is to predict the VLE data for the system water + ethanol + 2-methoxy-2-methylbutane using various local composition models and the group contribution models of ASOG, original UNIFAC, UNIFAC Dortmund, and UNIFAC-Lyngby. For the group contribution models, accurate predictions in the VLE of the system were obtained, particularly using the UNIFAC Dortmund and UNIFAC-Lyngby models. For these models, the calculated errors with respect to the experimental data were found to be less than 0.03 in the vapor composition and less than 2 K in the temperature. Although the UNIFAC-Lyngby was the most appropriate group contribution model for the system, the UNIFAC Dortmund model also provided a good correlation with the experimental data of the system. Hence, it can also be used for the prediction of the VLE of the system at various conditions.

Another ethanol-water system with tertiary ether has been studied by Arce et al. (2005), in which various thermodynamic models were used for the prediction of VLE of the systems water + ethanol + 2-ethoxy-2-methylbutane and water + methanol + 2-ethoxy-2-



methylbutane. The same group contribution models as stated previously were applied to these systems. It was observed that the group contribution models predicted VLE data of the two systems that are quite far from the experimental VLE. At the same time, it was shown that the UNIFAC Dortmund model is the most accurate group contribution model in the prediction of thermodynamic properties at VLE. Hence, this supports the greater reliability of UNIFAC Dortmund as compared to other group contribution models.

The next set of studies involve the application of UNIFAC Dortmund binary interaction parameters to various systems that are commonly found in the production of biodiesel. Most of these systems contain glycerol, water and ethanol, which are closely related to the systems considered in this study. One of the studies by Kuramochi et al. (2009) involve the application of different UNIFAC models in the prediction of the VLE of the binary systems methanol + glycerol and methanol + biodiesel, and LLE of the systems water + biodiesel, methanol + biodiesel + glycerol, and methanol + water + biodiesel. The biodiesel was assumed to contain a fixed proportion of methyl linoleate, methyl oleate and methyl palmitate. For the prediction of VLE, it was found out that the UNIFAC Dortmund model was the most accurate model, with prediction errors of less than 30% for all pressure values. The model is also accurate in the prediction of LLE, alongside the UNIFAC-LLE model of Magnussen et al. (1981). Therefore, the UNIFAC Dortmund model has been shown to be the most suitable group contribution model for the mixtures of biodiesel production.

On the other hand, another set of mixtures from the biodiesel industry were investigated by Coelho et al. (2011), which involve the binary systems water + glycerol, ethanol + glycerol, ethanol + ethyl stearate and ethanol + ethyl palmitate. Different thermodynamic models, which also include the UNIFAC Dortmund model, were used to predict the VLE of the binary systems. It was found out that the UNIFAC Dortmund model was able to provide a good correlation for most binary systems, hence predicting their VLE with good accuracy. On average, the prediction errors in the temperature were found to be only within 6 K. This further provides support on the reliability of the UNIFAC Dortmund model for the mixtures of biodiesel production.



Another set of binary systems from biodiesel production were studied by Muhammad et al. (2017). The binary systems such as ethanol + ethyl hexanoate, 1-pentanol + ethyl hexanoate and 1-pentanol + ethyl octanoate were considered. The VLE of the systems were determined using the UNIFAC Dortmund and UNIFAC Lyngby as the group contribution models alongside the NRTL model. It has been observed that all the models can provide good predictions of VLE for the binary systems, with NRTL being the most accurate in the prediction of temperature and vapor mole fraction values. Furthermore, it has been shown that the UNIFAC Dortmund model can match the predictions of the NRTL model for the system ethanol + ethyl hexanoate at 53.33 kPa. Hence, the UNIFAC Dortmund model already can be used for the prediction of VLE of this system even it has minimal or no experimental data available. Finally, by considering only the group contribution models, it has been shown that the UNIFAC Dortmund provided better predictions as compared to UNIFAC Lyngby. It can be seen so far that the UNIFAC Dortmund model can be used more advantageously than the other known group contribution models.

On the other hand, the next study focuses not on the biodiesel mixtures but on the mixtures of side products which are composed of glycerol and the excess alcohols. Hartanto et al. (2017) predicted the VLE of different binary systems of alcohols using the original UNIFAC and the UNIFAC Dortmund model. The binary systems contain glycerol as the first component, and any of the C_1 - C_4 alcohols as the second component. It has been shown in most cases that the VLE of the binary systems of alcohols can be predicted more accurately by the UNIFAC Dortmund model than the original UNIFAC. Therefore, this confirms that the UNIFAC Dortmund model has always been more reliable than the original UNIFAC model.

However, for some systems, the UNIFAC Dortmund model cannot provide good prediction of thermodynamic properties. For example, the study of Basso et al. (2017) involves the prediction of LLE of the system glycerol + ethanol + ethyl biodiesel using NRTL and UNIFAC Dortmund models. The ethyl biodiesel present in the system consists



of various fatty acid ethyl esters (FAEE), whose proportions are mentioned in the paper. It was found out that UNIFAC Dortmund model provides poor prediction of the LLE of the system. The average deviation of the liquid mass fraction data with the experimental LLE ranged from 1.3 - 6.2%. More importantly, large deviations are obtained for ethanol liquid mass fractions, where these quantities are overestimated at glycerol-rich phase and underestimated at ethyl biodiesel-rich phase. Hence, the UNIFAC Dortmund model only has a limited application for the given system.

The problems of poor prediction results by the UNIFAC Dortmund model can also be seen from the previous studies mentioned. For example, in the study of Coelho et al. (2011), poor predictions are obtained for the system water + glycerol, which is even worse than the predictions obtained when the system is assumed to be ideal. It was reported by the authors that this is caused by the complex interactions of the OH groups of glycerol with water, which was poorly accounted by the model. Another case can be found in the study of Hartanto et al. (2017) for the methanol and glycerol system, where the revision of the UNIFAC Dortmund parameters was proposed by the authors.

In general, the UNIFAC Dortmund model has shown by the literatures to be useful for the prediction of the thermodynamic properties of wide range of solvent systems. However, the problem of poor prediction results and the lack of required binary interaction parameters is still not completely solved, even if the systems only contain solvents and do not contain any ionic (or charged) components. Thus, the study also participates in addressing these gaps by applying the existing parameters to solvent systems to understand the possible reasons of the deviations from the experimental VLE, and also by estimating new parameters to allow the accurate prediction of VLE of certain solvent systems.

2.2.3 *Application of UNIFAC Dortmund Model to Systems With Salt*

The studies that applied UNIFAC Dortmund model to systems with salt either applied existing parameters involving ions to predict the thermodynamic properties of various systems with salt, or revised or estimated new parameters for some ionic interactions.



Due to lack of officially-accepted parameters for most ions, the latter is usually performed which is provided by the following examples:

One of the first studies focusing on the estimation of parameters is the study of Santos et al. (2000). The study wants to investigate on the salt effect of potassium bromide on the LLE of water + ethanol + pentanol system. One of the models used to predict the LLE of the system is the UNIFAC Dortmund model, where new parameters involving the binary interactions with K^+ and Br^- ions were estimated. The application of new parameters resulted to a global prediction error of 15.26%, which is significantly larger than the error found using NRTL, which is only 0.95%. Hence, the UNIFAC Dortmund model cannot be used to accurately predict the LLE of the given system.

Another study was performed by Se & Aznar (2001), which determined the LLE of water and polyethylene glycol system with salt, where the salt used is either K_3PO_4 or Na_2SO_4 . A modified version of UNIFAC Dortmund was used, where the combinatorial term was revised using the Elbro-Free Volume (Elbro-FV) model. New parameters involving the ions Na^+ , SO_4^{2-} and PO_4^{3-} were calculated. The parameters for K^+ have already been estimated in the study of Santos et al. (2000) which were simply applied in this study. Results showed that the UNIFAC Dortmund model alone is not capable of predicting the phase split for the systems water + polyethylene glycol + K_3PO_4 and water + polyethylene glycol + Na_2SO_4 but requires the incorporation of the Elbro-FV formula for better prediction results. However, the prediction results for the phase split of the binary system with Na_2SO_4 still remains unsatisfactory. Hence, although the UNIFAC Dortmund + Elbro-FV model is already a revised version, it remains to have a limited application for water and polyethylene glycol systems with salt.

This is followed by a more extensive study of Aznar & Telles (2001) about the estimation of new parameters for the commonly-known ions, which can be used for the prediction of VLE of a much wider range of systems with salt. The authors used a total of 85 binary and ternary systems with salt for the estimation of new parameters. The Debye-Huckel interaction of salts were neglected in this study since it was argued that it has no



effect on the VLE of the systems with salt. Thus, the UNIFAC Dortmund model was applied to the given systems with salt without any modification in its terms. For most cases, the application of the newly estimated parameters for the commonly-known ions caused the UNIFAC Dortmund model to accurately predict the VLE of the given systems with salt. The global prediction errors were found to be 0.011 atm for the binary systems with salt and 0.0803 atm and 0.051 vapor mole fraction for the ternary systems with salt. However, there are some systems with prediction errors of >0.1 for the vapor mole fraction, which are the ethanol-water systems containing NaCl and KCl and some propanol-water systems with Ca^{2+} . Although the following issues still required to be addressed, the study in general was able to perform a comprehensive determination of UNIFAC Dortmund parameters of ions that extended the application of the UNIFAC Dortmund model to greater number of systems with salt.

The new parameters obtained by Aznar & Telles (2001) were successfully applied in the study of Matugi & Giordano (2015), together with the other existing models used for the prediction of VLE of various systems with salt. The activity coefficient models were applied specifically for ethanol-water system with salt, where the salt used is either CH_3COOK or CaCl_2 . Some of the models used either have equations that represent the long-range Debye-Huckel interactions, or equations that describe the medium-range interactions of ions with water in the mixture. It has been shown that the UNIFAC Dortmund model using the new parameters of Aznar & Telles (2001) provided the best prediction results at VLE, which is followed by the UNIFAC model developed by Kikic et al. (1991), which contains a Debye-Huckel term. Hence, the new parameters of Aznar & Telles (2001) are also applicable for the systems with salt considered in this study. However, the authors reminded that the new parameters of Aznar & Telles (2001) may only work for a limited range of salt concentrations.

Along with the presence of the several parameters provided by Aznar & Telles (2001), there are also studies that independently estimated new parameters for certain ionic interactions. These studies applied the UNIFAC Dortmund model with the addition of Debye-Huckel terms in the residual term. One of which is the study from Yousefi et al.



(2010), where the Pitzer-Debye-Huckel (PDH) model was used along with the UNIFAC Dortmund model. Instead of estimating three binary interaction parameters which is a characteristic of the UNIFAC Dortmund model, only two parameters a_{mk} and b_{mk} were calculated for each ionic interaction. The combination of two models were applied for the prediction of saturation molality of aqueous electrolyte solutions containing either one or two salts from NaCl, CaSO₄ and Na₂SO₄. It has been shown that the two models when put together provided good predictions for the saturation molality which are in good agreement with the experimental data. Hence, it also follows that the parameters estimated for Na⁺, Cl⁻, Ca²⁺ and SO₄²⁻ ions were acceptable.

The same authors (Yousefi et al., 2011) conducted another study using the same salts, however, the study worked on aqueous electrolyte solutions containing all the three salts instead of only one or two salts. The PDH + UNIFAC Dortmund model was still used but with the objective of predicting the saturation molality of NaCl at varying molality ratios of Na₂SO₄ and CaSO₄. The parameters obtained from the previous study (Yousefi et al., 2010) were also applied in this study to test their applicability to the new aqueous electrolyte solutions. It was observed that the model predicted the saturation molality of NaCl for all Na₂SO₄/CaSO₄ molal ratios with good accuracy. Thus, the parameters that were previously estimated for Na⁺, Cl⁻, Ca²⁺ and SO₄²⁻ ions are also applicable to the aqueous electrolyte solutions involved in this study.

From the assessment of the literatures, it can be observed that the UNIFAC Dortmund model without the inclusion of additional equations for ionic interactions can still provide good predictions for most systems with salt, which has been demonstrated by Aznar & Telles (2001). Hence, this study focuses on the application of the UNIFAC Dortmund model alone as it can provide greater convenience in the calculations than its more complex modified versions. With the given direction chosen for this study, the research gap to be addressed is successfully narrowed down to the search of the appropriate set of UNIFAC Dortmund binary interaction parameters for all ionic interactions that can be applicable for wide range of systems with salt.



2.3 Optimization Methods for Parameter Estimation

The optimization methods for the estimation of binary interaction parameters requires the basic knowledge of the optimization process, followed by the selection of the best optimization algorithm for a given optimization problem and the selection of objective function to be minimized. These considerations will ensure that the parameters will be calculated in such a way that it will provide predictions of VLE that have small deviations with the experimental VLE data.

2.3.1 General Optimization Procedure for VLE

The general optimization procedure for parameter estimation includes the following steps as follows: input of experimental VLE data, setting of initial guesses, evaluation of properties using thermodynamic models, satisfaction of required constraints, and more importantly, the minimization of objective function. Although these steps vary for every study depending on the thermodynamic model used and the desired objective function, these steps have always been essential which can be illustrated in a few studies as follows:

Kaewsichan (1999) used the general optimization procedure for the determination of parameters for various systems containing CO_2 , H_2S , H_2O , monoethanolamine, and methyldiethanolamine using the electrolyte UNIQUAC equation. In the program for optimization, the experimental data and the initial guesses for the temperature-dependent parameters were included, which were substituted to the electrolyte UNIQUAC to solve for the activity coefficients, and to the gamma-phi formulation of VLE to calculate for the vapor mole fractions. The program then calculated the residual sum of squares corresponding to the liquid mole fraction and activity coefficient data, which was constrained to fall within a specific range. At the times when the constraint was not satisfied, the initial guesses for the parameters were updated and the series of calculations mentioned were repeated. When the constraint was already satisfied, another set of iterations were performed to minimize the objective function, which is defined to be the residual sum of squares of pressure, temperature and activity coefficient data, The resulting output values were the true



parameters corresponding to all possible binary interactions between the components. As can be seen, the essential steps for the general optimization are present.

The presence of the essential steps for optimization can also be observed in the study of Kim (2005). In this study, the UNIFAC and Knox models were applied in the prediction of VLE of systems containing various organic compounds. New energy interaction terms were introduced to improve the prediction of the models mentioned, where the new terms contain some parameters that are unknown and hence are subject to estimation. In the optimization process, the VLE data of various systems from the literature, together with the initial guesses for energy interaction parameters, were provided as the input of the program. This is followed by the evaluation of different thermodynamic properties using the two models. The study defined the objective function as the residual sum of squares of pressure, which was first constrained to be within a certain tolerance level before minimizing its value. From the results of this study, the energy interaction parameters were found successfully using the general optimization process.

It was also observed that the results of the parameter estimation using the general optimization process can be affected by the initial guesses, hence various initial guesses must be used first before deciding on the final results. This was found out in the study of Jain et al. (2017), which involved the determination of binary interaction parameters for thiol and olefin main groups using the UNIFAC model. In this study, the entire optimization process was performed for 2,000 trials instead of just a single trial, with initial guesses that are modified in every trial. The new UNIFAC parameters were obtained by selecting the trial that provided the minimum value of the objective function. It can be seen that a single trial for optimization does not guarantee the achievement of the minimum value of the objective function, hence it may provide parameters that are different from the true value.

By considering the literatures mentioned, the general optimization procedure for the parameter estimation will be followed in this study along with the method of manipulation of initial guesses by Jain et al. (2017) to obtain the true UNIFAC Dortmund binary interaction parameters.



2.3.2 Optimization Algorithms from Previous Studies

The optimization procedures used by various studies for parameter estimation also differ in terms of the algorithm used to minimize the objective function. Since the minimization of objective function involves iterations, it is important to select a specific algorithm that will reduce the number of iterations and will speed up the iteration process for a given type of objective function and constraints. Three optimization algorithms are commonly used for the minimization of objective function: Nelder-Mead simplex method, Marquardt method, and Sequential Quadratic Programming (SQP).

The simplex method of Nelder & Mead (1965) minimizes a given objective function through the following steps: creating a simplex with vertices that corresponds to the parameters, evaluation of the value of objective function at the given vertices, and the series of calculations that generates new vertices, forming another simplex. The process is repeated until a convergence to a certain simplex and hence to a certain vertex or parameters is achieved. This method is under the direct search algorithm, which involves only the evaluation and comparison of objective function values to find the best set of parameters. According to Lagarias et al. (1998), the advantage of this method is that the value of the objective function at the first iterations is significantly reduced. Gmehling et al. (2015), who applied this method in the parameter estimation for UNIFAC Dortmund model, also confirms the advantage of this method at the first iterations, where the method was described to be robust. Another advantage of this method is that it only requires few evaluations of objective function for every iteration and do not require derivatives in the solving process. With these advantages mentioned, this method is widely used by authors for the parameter estimation using the UNIFAC Dortmund model.

Some disadvantages can also be found in this method. McKinnon (1998), Lagarias et al. (1998) and Dennis and Torczon (1991) discussed the cases where the method does not converge at the minimum value of objective function, which includes nonconvex objective functions or functions with multiple local minima, and even in some objective functions whose number of variables are only 8 or less (Dennis and Torczon, 1991). Therefore, it is



necessary to check first the nature of the objective functions selected for a given optimization problem before using this method.

Another method that can be used is the method of Marquardt (1963), which is known as the Levenberg-Marquardt method. This method is a combination of gradient descent and Gauss-Newton method. Similar to the gradient descent method, initial guesses for parameters are continuously subtracted with the value of the gradient of the objective function multiplied by a certain damping factor, until the convergence to the solution of parameters are obtained. It is important that the correct values of damping factors are used for every iteration so as not to slow down the convergence to a solution or that the real solution will not be skipped in the iteration. The values of damping factors can be obtained by determining the curvature or second derivative around the parameters using the Gauss-Newton method. In summary, the Marquardt method takes advantage of the certainty of convergence of the gradient descent method and the speed of convergence of the Gauss-Newton method.

Because of the fast convergence to the true solution, Gmehling (1993) used this method for the first-ever calculated binary interaction parameters in the UNIFAC Dortmund model, where the method was applied as the final step for the minimization of objective function. This is because according to Marquardt (1963), if the iteration is already close to the minimum value of the objective function, then the convergence to the parameters will be faster. This method is used even for the determination of the most recent UNIFAC Dortmund parameters along with Nelder-Mead simplex method for the initial iterations (Constantinescu & Gmehling, 2016). Thus, this method has been helpful in the determination of binary interaction parameters for the UNIFAC Dortmund model.

Some disadvantages of this method are mentioned by Transtrum & Sethna (2012) such as the slow convergence for large number of unknown parameters, failure to converge if the initial guesses for parameters are far from the actual solution, and the possibility of parameters having infinite values, which is called parameter evaporation. Therefore,



according to the authors, the number of parameters should be limited to 10 and the approximation of the damping factor values must be ensured to be acceptable.

Finally, one of the most widely used methods is the sequential quadratic programming (SQP) by Powell (1978). Since it is also a gradient-based method similar to the Marquardt method, it requires that both the objective and constraint functions must be continuous and also differentiable (Schittkowski, 1986). This method formulates a quadratic subproblem for every iteration by linearizing the constraint functions and minimizing the quadratic approximation of the Lagrangian function. The Lagrangian function used in this method is the sum of the values of the objective function and also of the constraint functions multiplied by certain Lagrangian multipliers, which is analogous to the damping factor of Marquardt method. The correct selection of the Lagrangian multipliers are also essential in this method which should adhere to the Karush-Kuhn Tucker (KKT) equations, a set of equations that ensures the constrained minimization of the objective function.

Based on the series of calculations discussed, several advantages are obtained in the use of SQP. According to Boggs & Tolle (2000), the method is effective in solving optimization problems with high degree of nonlinearity and is always applicable regardless of the number of constraints and variables. Furthermore, the constraints of a given optimization problem is not required to be satisfied at every iteration, only the KKT equations is enough to be met. The method is also known to capitalize well on the good initial guesses in the iterations (Gill & Wong, 2012). Thus, several studies are using this method, which includes the determination of parameters at phase equilibria and the optimization of design processes.

For example, Bortz et al. (2019) used SQP method as well as the Marquardt method to find NRTL parameters for acetone and methanol system at VLE. It was observed that both the SQP and Marquardt method provided parameters sets whose predictions at VLE are in good agreement with the experimental data. Good results are also obtained in the study of Montastruc et al. (2002), which used SQP in determination of solid-liquid equilibrium (SLE) of calcium phosphate compounds and water, combined with genetic algorithm (GA)



for the determination of good initial guesses for the SQP. The study was able to find solutions using the combination of the two methods. Hence, it has been shown that SQP can be successfully used for phase equilibria calculations.

The SQP method also has some disadvantages which includes the slow convergence to the true solution for large number of variables and constraints, where the latter can also cause errors. Furthermore, the early termination of the iterations of SQP may lead to parameters that do not satisfy the constraints of a given optimization problem and may provide a large value when substituted into the objective function. These situations that prevent the SQP method to reach the minimum value of objective function must be avoided.

Through the comparison of the optimization algorithms discussed, the SQP was found out to be the most suitable algorithm for this study. This is primarily because it allows constraining the value of the parameters to a certain range, hence preventing the parameters from reaching infinity during their iteration process. Moreover, since the UNIFAC Dortmund is a highly nonlinear model with several variables, the SQP can handle the calculations resulting from the model and can guarantee the attainment of the minimum value of the objective function.

2.3.3 Objective Function Selection

The objective function used for the general optimization process also vary for every study. It is because the objective function must be able to allow a given thermodynamic model to predict values are close to the experimental data. According to Bonilla-Petriciolet et al. (2010), there are two main types of objective functions: the least squares criterion and the error-in-variable criterion.

The least squares criterion is one of the most widely used methods for parameter estimation at VLE. This involves the minimization of the sum of the squares of either the error or relative error of the calculated value with respect to the experimental value. Gmehling et al. (2006) frequently use the least squares for the relative errors, which according to Bonilla-Petriciolet et al. (2010), has an advantage of weighing the errors equally regardless



of whether the experimental value is large or small. On the other hand, the error-in-variable method assumes that the experimental errors exist for the thermodynamic properties measured. Thus, the objective function is defined to be the sum of the squares of the errors divided by the variance for a given property. This criterion is also used by several studies.

However, the error-in-variable method causes the objective function to contain several variables, which increases with the number of experimental data. Since there are hundreds of experimental data involved in this study, this method is not recommended as it may slow down the calculations. Thus, the least squares criterion for relative errors used by Gmehling et al. (2006) will be used because of the advantages that were previously mentioned.