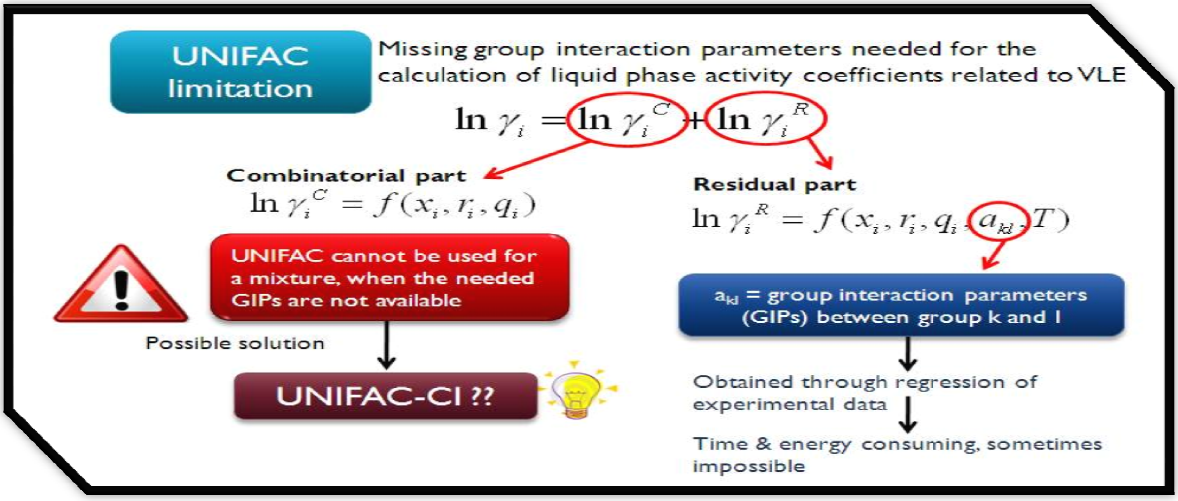
Development and Analysis of Group Contribution Plus Models for Property Prediction of Organic Chemical Systems

Mustaffa, Azizul Azri

Publication date:
2013

Document Version
Publisher's PDF, also known as Version of record

Development and Analysis of Group Contribution Plus Models for Property Prediction of Organic Chemical Systems



Azizul Azri Bin Mustaffa
Ph.D. Thesis
November 2012

DTU Chemical Engineering
Department of Chemical and Biochemical Engineering

Development and Analysis of Group Contribution^{Plus} Models for Property Prediction of Organic Chemical Systems

Ph.D Thesis

Azizul Azri Bin Mustaffa

November 2012

Computer Aided Process-Product Engineering Center
Department of Chemical and Biochemical Engineering
Technical University of Denmark

DTU Chemical Engineering
Department of Chemical and Biochemical Engineering

Address: Computer Aided Process Engineering Center
Department of Chemical and Biochemical Engineering
Technical University of Denmark
Building 229
DK-2800 Kgs. Lyngby
Denmark
Phone: +45 4525 2800
Fax: +45 4588 4588
Web: www.capec.kt.dtu.dk

Print: J&R Frydenberg A/S
København
March 2013

ISBN: 978-87-92481-95-5

PREFACE

This thesis was written as a partial fulfillment of the requirements for the degree of Doctor of Philosophy (Ph.D) in Chemical Engineering at the Technical University of Denmark (DTU). The project has been carried out at the Computer Aided Process-Product Engineering Center (CAPEC) at the Department of Chemical and Biochemical Engineering, from April 2009 until November 2012, under the supervision of Professor Rafiqul Gani (CAPEC) and Professor Georgios Kontogeorgis from the Center for Energy Resources Engineering (CERE). This work has been financed and supported by the Ministry of Higher Education Malaysia (MoHE) and Universiti Teknologi Malaysia (UTM).

First and foremost, I would like to express my gratitude to my main supervisor, Prof. Rafiqul Gani for giving me the opportunity to work with this project and also for the guidance and academic support which have been given to me whenever needed. I am also grateful to my co-supervisor, Prof. Georgios Kontogeorgis for the motivational and academic supports, guidance and also interest in my work. In addition to that, I would like to thank Prof. Jeong Won Kang from Korea University, Seoul for the guidance on the work regarding the consistency tests and uncertainty analysis. Special thanks also to Prof. Zainuddin Abd. Manan from Process Systems Engineering Center (PROSPECT), UTM for the encouragement and support.

Furthermore, I would like to thank previous CAPEC coworkers who include Hugo, Ricardo, Axel, Merlin, Martin, Alicia, Elisa, Philip, Martina, Ravendra, Kamaruddin, Fazli, Oscar, Rasmus, Sascha and also current CAPEC coworkers, Amol, Igor, Sascha, Deenesh, Michele, Alberto, Marina, Peam, Alafiza, Katrine, Hande and Larissa for all the help which have been given when needed and also for the technical and social discussions that we had through my stay here in CAPEC. Special thanks also to Thomas Bisgaard for helping me translating the thesis summary into Danish.

I also would like to express my appreciation for all the support that I got from my friends in Malaysia especially from Silat Cekak Pusaka Hanafi. Special thanks also to Sheikh Dr. Hj. Md. Radzi Hj. Hanafi for the moral support.

Finally and most importantly, I wish to thank my family for their support and understanding throughout the research process and during the writing of this thesis. I am grateful to my parents Mr. Mustaffa and Mrs. Rosnani for their continuous motivational support and prayers. My special thanks also to my wife, Siti Salwa for her unconditional support, patience and love especially as during the period of this Ph.D research, our two lovely sons, Eilman Zulqarnain and Eimran Zubayr were born in Denmark.

AZIZUL AZRI BIN MUSTAFFA
Kgs. Lyngby, November 2012

SUMMARY

Prediction of properties is important in chemical process-product design. Reliable property models are needed for increasingly complex and wider range of chemicals. Group-contribution methods provide useful tool but there is a need to validate them and improve their accuracy when complex chemicals are present in the mixtures. In accordance with that, a combined group-contribution and atom connectivity approach that is able to extend the application range of property models has been developed for mixture properties. This so-called Group-Contribution^{Plus} (GC^{Plus}) approach is a hybrid model which combines group contribution and molecular descriptor theories (such as connectivity indices – CI). Connectivity indices are formalisms defined via graph theoretical concepts intended to describe the topological characteristics of molecular structures. The main idea is the use of connectivity indices to describe the molecular fragmentation that relates properties which is the molecular interactions with the molecular structures. One well known and established group-contribution method is the UNIFAC model, used to predict liquid phase activity coefficients for mixtures. The needed values of the group interaction parameters (GIPs) are obtained by fitting phase equilibrium data. There are, however many gaps in the UNIFAC parameter table due to lack of data. Alternative to performing measurements, which may not be feasible, values of the missing GIPs, can be predicted through the GC^{Plus} approach. The predicted values for the GIPs are then used in the UNIFAC model to calculate activity coefficients. This approach can increase the application range of any “host” UNIFAC model by providing a reliable predictive model towards fast and efficient product development.

This PhD project is focused on the analysis and further development of the GC^{Plus} approach for predicting mixture properties to be called the UNIFAC-CI model. The contributions of this work include an analysis of the developed Original UNIFAC-CI model in order to investigate why the model does not perform as well as the reference UNIFAC model for some systems while performing surprisingly better than the reference model for other systems. In this analysis, it is found that by introducing more structural information to the CHO group through the valence connectivity index (CI), the correlation error involving alkanes-aldehydes system can be reduced. This work is presented in Chapter 3. Furthermore in Chapter 4, as a continuation of the analysis done for systems involving C, H and O atoms, the Original UNIFAC-CI (VLE) model has been further reused and significantly expanded by including nitrogenated, chlorinated and sulfurated systems and the involved atom interaction parameters (AIPs) have been regressed. In addition to that, another set of parameters have been generated for the Original UNIFAC-CI (VLE) model using a quality assessment algorithm, Q_{VLE} (combination of 4 VLE consistency tests) as a weighting factor for each VLE dataset in the objective function for regression of AIPs. The quality factors are useful in identifying anomalous systems which can be problematic in the parameter estimation and can produce parameters which are not accurately representing the systems used for the regression. Moreover, in Chapter 5 the Original UNIFAC-CI (VLE/SLE) model have been developed where the atom interaction parameters (AIPs) are obtained through regression against both VLE and SLE experimental data. The prediction accuracy of SLE systems using the regressed parameters has been slightly increased. Besides that, in Chapter 6, Modified (Dortmund) UNIFAC-CI has been further developed by including chlorinated and sulfurated VLE systems. Finally, in Chapter 7, the developed Original UNIFAC-CI (VLE/SLE) model has been highlighted in selected case studies involving the design of a working solution for hydrogen

peroxide production and solubility investigation of pharmaceutical systems where new group have been created and their interaction parameters are predicted/fine tuned generating a master parameter table specifically for those case studies. Also, the applicability of the Original UNIFAC-CI model is shown for predicting phase equilibria of lipid systems, filling missing GIPs and improving prediction of azeotropic mixture. In Chapter 8, a discussion with concluding remarks and recommendation for future work are presented.

RESUME PÅ DANSK

Forudsigelse af egenskaber er vigtigt i kemisk proces-produktdesign. Pålidelige modeller er påkrævet til komplekse kemikalier og for at dække et bredere udvalg. Gruppebidragsmetoder bidrager hertil med et nyttigt redskab, men der er behov for at validere dem og forbedre deres nøjagtighed i tilfældet af blandinger indeholdende komplekse kemikalier. I denne forbindelse er en kombineret gruppebidrags- og atomstruktursmetode blevet udviklet, som udvider anvendelsen af en række egenskabsmodeller til blandinger. Denne såkaldte Group-Contribution^{Plus}-metode (GC^{Plus}) er en hybrid, som kombinerer gruppebidrags- og molekyledeskriptorteorier (såsom valence connectivity indices – CI). Valence connectivity index er formelt defineret via grafteoretiske begreber, hvis formål er at beskrive de topologiske egenskaber af molekylers strukturer. Ideen er at anvende valence connectivity indices til at beskrive den molekylære fragmentering, der kan relateres til de molekylære interaktioner med de molekylære strukturer. En udbredt og veletableret gruppebidragsmetode er UNIFAC-modellen, der anvendes til at forudsige blandingers aktivitetskoefficienter i væskefasen. De nødvendige gruppeinteraktionsparametre (GIP) opnås ved at tilpasse model til eksperimentelle faseligevægtsdata. Der er imidlertid mange huller i tabellerede UNIFAC-parametre på grund af manglende data. Et alternativ til at udføre målinger, hvilket eventuelt ikke er muligt, er at forudsige manglende GIP-parametre vha. GC^{Plus}-fremgangsmåden. De forudsagte parametre anvendes derefter i UNIFAC-modellen til beregning af aktivitetskoefficienter. Denne tilgang kan øge anvendelsesmuligheden af enhver "værts"-UNIFAC-model ved at give en pålidelig, prædiktiv model til f.eks. hurtig og effektiv produktudvikling.

Fokus i dette ph.d.-projekt er på analyse og videreudvikling af GC^{Plus}-metoden til forudsigelse af blandingers egenskaber. Denne metode kaldes UNIFAC-CI-modellen. Bidragene fra dette arbejde består af en analyse af den udviklede Original UNIFAC-CI-model for at undersøge, hvorfor modellen ikke klarer sig ikke godt som reference-UNIFAC-modellen for visse systemer, men væsentlig bedre end referencemodellen for andre systemer. I denne analyse har det vist sig, ved at indføre mere strukturel information om CHO-gruppen gennem valence connectivity index (CI), at korrelationsfejlen for systemer involverende alkaner-aldehyder kan reduceres. Førnævnte arbejde er præsenteret i kapitel 3. I kapitel 4 er Original UNIFAC-CI (VLE)-modellen, som en forlængelse af analysen udført for systemer med C, H og O-atomer, blevet yderligere genanvendt og markant udvidet ved at inddrage nitrogen-, klor- og svovlholdige systemer, hvor de involverede atominteraktionsparametre (AIPS) er korreleret. Ydermere er et alternativt parametersæt blevet genereret til Original UNIFAC-CI (VLE)-modellen indeholdende en kvalitetsvurderingsalgoritme, Q_{VLE} (kombination af 4 VLE konsistenstests), som en vægtningsfaktor for hver VLE-datasæt i målfunktionen for regressionen af AIPS. Kvalitetsfaktorerne er nyttige til at identificere anormale systemer, som kan være problematiske i parameterestimeringssammenhæng og som producerer unøjagtige parametre. Desuden er Original UNIFAC-CI (VLE / SLE)-modellen blevet udviklet i kapitel 5, hvor atominteraktionsparametre (AIPS) opnås vha. regression af både VLE og SLE eksperimentelle data. Ved hjælp af denne metode er nøjagtigheden af forudsigelserne af SLE-systemer steget. I kapitel 6 er Modified (Dortmund) UNIFAC-CI blevet videreudviklet til at dække bl.a. klor- og svovlholdige VLE-systemer. Endelig afprøves den udviklede Original UNIFAC-CI (VLE/SLE)-model i kapitel 7 for et udvalg af eksempler. Disse inkluderer design af en blanding til fremstilling af hydgrogenperoxid samt en opløselighedsundersøgelse af farmaceutiske systemer. Hertil er nye grupper blevet udviklet, og disses interaktionsparametre er blevet estimeret/finjusteret

resulterende i overordnede parametertabeller, som er specifikke for de pågældende eksempler. Desuden er anvendeligheden af Original UNIFAC-CI-modellen til forudsigelse af faseligevægte af lipidsystemer blevet undersøgt, manglende GIP data er blevet udfyldt og forudsigelser af azeotropiske blandinger er blevet forbedret. I kapitel 8 er en diskussion med afsluttende bemærkninger og anbefalinger for det fremtidige arbejde præsenteret.

TABLE OF CONTENTS

Preface iii
Abstract iv
Resume på Dansk vi
Table of Contents viii
List of Tables xii
List of Figures xvi

Chapter 1: Introduction 1
1.1 Background of the Project 1
1.2 Role and Importance of UNIFAC in Process and Product Design 4
1.3 Why Connectivity Index? Previous Work and Status 5
1.4 Motivation and Objectives 6
1.5 Thesis Organization 7

Chapter 2: Theoretical/Conceptual Background and Analysis Methods/Tools 9
2.1 Review of Available Property Prediction Models for Mixtures Properties 9
2.1.1 Quantum Mechanics / Quantum Chemical Methods 10
2.1.2 Group-Contribution Methods 11
2.1.3 Quantitative Structure-Activity Relationship (QSAR) / Quantitative Structure-Property Relationship (QSPR) Methods 12
2.2 Background of Group Contribution Methods 13
2.2.1 Pure Component Group Contribution Methods 13
2.2.2 Group Contribution Methods for Mixtures 15
2.3 The UNIFAC Group Contribution Method 16
2.3.1 Limitations of UNIFAC 19
2.3.2 UNIFAC Variants 20
2.3.2 UNIFAC Applications 24
2.4 Molecular Description Theory 25
2.4.1 Background of the Connectivity Index 26
2.4.2 Significance of Valence Connectivity Index 26
2.5 Group Contribution^{Plus} Models 28
2.5.1 Background of GC^{Plus} Models 28
2.5.2 UNIFAC-CI Model Formulation 29
2.5.2.1 Original UNIFAC-CI Model 33
2.5.2.2 Modified (Dortmund) UNIFAC-CI Model 36
2.6 Methods and Tools for Analyzing the Models and Their Accuracy 38
2.6.1 Thermodynamic Equations for Phase Equilibrium 39
2.6.2 VLE and SLE Experimental Data and Sources of Data 41
2.6.3 Optimization Scheme and Objective Function for Parameter Regression 43

2.6.4 UNIFAC Terms Analysis 44
2.6.5 VLE Data Quality Criterion, Q_{VLE} 45
2.6.6 Software Tools 45

Chapter 3: A First Analysis of the Original UNIFAC-CI Model 47
3.1 Introduction 47
3.2 Analysis of the Original UNIFAC-CI Model (Hydrocarbons and Oxygenated Systems) 47
3.2.1 Identification of the Problematic Systems 48
3.2.2 Identification of the Problems in the Model Expression and/or Parameters 50
3.2.3 Regression with the Weighted Objective Function 55
3.2.4 Assigning Higher Order CI Parameters 55
3.3 Extrapolation of the Original UNIFAC-CI to SLE Systems (Hydrocarbons and Oxygenated Systems) 59
3.4 Conclusion of Analysis 62

Chapter 4: Further Analysis and Development of Original UNIFAC-CI Models: VLE Systems 63
4.1 Introduction 63
4.2 Development of the Original UNIFAC-CI Model for VLE 63
4.2.1 Background of Parameter Estimation 64
4.2.2 Equilibrium Data 65
4.2.3 Regression Procedure 66
4.2.4 Correlation Results 67
4.2.5 Overall Correlation Results 85
4.3 Predictions of VLE Data using the Regressed Parameters 82
4.3.1 Oxygenated Systems 82
4.3.2 Nitrogenated Systems 83
4.3.3 Chlorinated Systems 85
4.3.4 Sulfurated Systems 87
4.4 Conclusions 87

Chapter 5: Further Analysis and Development of Original UNIFAC-CI Models: Addition of SLE Systems 90
5.1 Introduction 90
5.2 Development of the Original UNIFAC-CI (VLE/SLE) Model 90
5.2.1 Background of Parameter Estimation 90
5.2.2 Equilibrium Data 93
5.2.3 Regression Procedure 94
5.2.4 Correlation Results 94
5.2.4.1 Vapor-Liquid Equilibrium Data 95
5.2.4.2 Solid-Liquid Equilibrium Data 108
5.2.4.3 Overall Correlation Results 117

5.3 Further Analysis of the Original UNIFAC-CI Models 123
5.3.1 Analysis of SLE Predictions using Regressed Parameters 123
5.3.2 Further Analysis of Parameter Regression Work 128
5.3.2.1 Problematic Systems in Parameter Regression 128
5.3.2.2 Discussion on SLE Experimental Data 137
5.4 Conclusions 138

Chapter 6: Development of Modified (Dortmund) UNIFAC-CI Models: VLE Systems 139
6.1 Introduction 139
6.2 Development of the Modified (Dortmund) UNIFAC-CI Model for VLE 139
6.2.1 Background of Parameter Estimation 140
6.2.2 Equilibrium Data 142
6.2.3 Regression Procedure 142
6.2.4 Correlation Results 143
6.2.5 Overall Correlation Results 146
6.3 Predictions of VLE Data using the Regressed Parameters 147
6.3.1 Chlorinated Systems 147
6.3.2 Sulfurated Systems 147
6.4 Conclusions 148

Chapter 7: Application of the Original UNIFAC-CI Model: Case Studies 150
7.1 Introduction 150
7.2 Working Solution Design for Hydrogen Peroxide Production 150
7.2.1 Background of Case Study 153
7.2.2 Phase Equilibria Predictions and Parameter Table 155
7.2.3 Conclusions from the Case Study 162
7.3 Pharmaceutical Systems 162
7.3.1 Background of Case Study 163
7.3.2 Phase Equilibria Predictions and Parameter Table 167
7.3.3 Conclusions for this of Case Study 174
7.4 Lipid Systems 175
7.4.1 Background of Sub-Case Study 1 176
7.4.2 Phase Equilibria Predictions and Parameter Table for Sub-Case Study 1 177
7.4.3 Background of Sub-Case Study 2 179
7.4.4 Phase Equilibria Predictions and Parameter Table for Sub-Case Study 2 180
7.4.3 Conclusions for the Lipid of Case Studies 183
7.5 Systems with Missing GIPs and Azeotropic Systems 183
7.6 Overall Conclusions from All Case Studies 187

Chapter 8: Conclusions and Future Work 189
7.1 Conclusion and Discussion 189

7.2 Contributions of PhD Project 191
7.3 Future Work and Recommendations 192

Nomenclature 194

References 198

Appendices 209

Appendix A: Tables of Atom Stoichiometry and Valence Connectivity Indices for the UNIFAC Groups 209
Appendix B: Atom Interaction Parameter Tables for the UNIFAC-CI Models 213
Appendix C: Prediction of GIPs of the Modified (Dortmund) UNIFAC-CI for Systems Involving Atoms C, H, O, N, Cl and S 219
Appendix D: Calculations of Group Interaction Parameters of CH_2-ACCO and CH_2-ACNH from the CI-Method 222
Appendix E: List of Conference Presentations and Publications 225

LIST OF TABLES

Table 2.1	Original UNIFAC-VLE Subgroup Parameters.	17
Table 2.2	Original UNIFAC-VLE Interaction Parameters, a_{mk}, in Kelvin.	17
Table 2.3	Some of the Most Important UNIFAC Variants.	20
Table 2.4	Some Typical R_k and Q_k Values for Modified (Dortmund) UNIFAC and Original UNIFAC.	22
Table 2.5	Literature of Works using the GC^{Plus} Approach using Valence Connectivity Index.	29
Table 2.6	Atomic, Bond and Path Indices Values for HCOO.	32
Table 2.7	Atom Stoichiometry and CI Values of Group CH_2 and ACOH.	35
Table 2.8	AIPs Needed for the Calculations of GIPs between Groups CH_2 and ACOH.	36
Table 2.9	Types of Phase Equilibria Data Used.	41
Table 2.10	Statistics of Data used in Parameter Regression.	41
Table 3.1	Number of Datasets Used for Each Type of System.	50
Table 3.2	AIPs Used by Several Systems.	51
Table 3.3	Order of Parameter Regression When Introducing a Second Order CI for the CH_3-CHO Interaction.	57
Table 3.4	Correlation Results for Alkane–Aldehyde Systems Compared between Different Representations of the CHO Group.	58
Table 3.5	New Set of AIPs Regressed After the Introduction of a Second Order CI for CHO group.	59
Table 4.1	Statistics of the Data Used in the Parameter Regression.	66
Table 4.2	Correlation Results for C-C Atom Interaction Related Systems.	68
Table 4.3	Correlation Results for C-O, O-C Atom Interaction Related Systems.	69
Table 4.4	Correlation Results for O-O Atom Interaction Related Systems.	73

Table 4.5	Correlation Results for C-N, N-C Atom Interactions Related Systems.	74
Table 4.6	Correlation Results for O-N, N-O Atom Interactions Related Systems.	75
Table 4.7	Correlation Results for N-N Atom Interaction Related Systems.	76
Table 4.8	Correlation Results for C-Cl, Cl-C Atom Interactions Related Systems.	77
Table 4.9	Correlation Results for O-Cl, Cl-O Atom Interactions Related Systems.	79
Table 4.10	Correlation Results for C-S, S-C, O-S, S-O Cl-S, and S-Cl Atom Interaction Related Systems.	80
Table 4.11	Overall correlation results in AARD (%).	81
Table 5.1	Statistics of the Data Used in the Parameter Regression.	93
Table 5.2	Correlation Results for C-C Atom Interactions Related VLE Systems.	95
Table 5.3	Correlation Results for C-O, O-C Atom Interactions Related VLE Systems.	97
Table 5.4	Correlation Results for O-O Atom Interactions Related VLE Systems.	101
Table 5.5	Correlation Results for C-N, N-C Atom Interactions Related VLE Systems.	102
Table 5.6	Correlation Results for O-N, N-O Atom Interactions Related VLE Systems.	103
Table 5.7	Correlation Results for N-N Atom Interactions Related VLE Systems.	104
Table 5.8	Correlation Results for C-Cl, Cl-C Atom Interactions Related VLE Systems.	104
Table 5.9	Correlation Results for O-Cl, Cl-O Atom Interactions Related VLE Systems.	106
Table 5.10	Correlation Results for C-S, S-C, O-S, S-O, Cl-S, and S-Cl Atom Interactions Related VLE Systems.	107
Table 5.11	Correlation Results for C-C Atoms Interactions Related SLE Systems.	108
Table 5.12	Correlation Results for C-O, O-C Atoms Interactions Related SLE Systems.	110
Table 5.13	Correlation Results for O-O Atoms Interactions Related SLE Systems.	114
Table 5.14	Correlation Results for C-N, N-C Atoms Interactions Related SLE Systems.	115
Table 5.15	Correlation Results for O-N, N-O Atoms Interactions Related SLE Systems.	116

Table 5.16 Correlation Results for N-N Atoms Interactions Related SLE System. 116

Table 5.17 Correlation Results for C-Cl, Cl-C Atoms Interactions Related SLE Systems. 116

Table 5.18 Correlation Results for O-Cl, Cl-O Atoms Interactions Related SLE Systems. 117

Table 5.19 Overall Correlation Results in AARD1 for VLE Systems. 118

Table 5.20 Overall Correlation Results in AARD2 for SLE Systems. 119

Table 5.21 Experimental Melting Temperatures and Heats of Fusion Used for the Solutes in Generating the SLE Data. 137

Table 6.1 Statistics of the Data Used in the Parameter Regression. 142

Table 6.2 Correlation Results for C-Cl, Cl-C, O-Cl and Cl-O Atom Interactions Related Systems. 143

Table 6.3 Correlation Results for C-S, S-C Atom Interaction Related Systems. 146

Table 6.4 Overall correlation results in AARD (%).147

Table 7.1 Atomic, Bond and Path Indices Values for ACCO. 154

Table 7.2 Group Interaction Parameters (GIPs) Table (in Kelvin) for Working Solution Design Case Study (GIPs labeled with * are predicted using the CI-method). 156

Table 7.3 Group Interaction Parameters (GIPs) Table (in Kelvin) for Working Solution Design Case Study with Fine Tuned Parameters (GIPs labeled with * are predicted using the CI-method while GIPs labeled with # are fine tuned. The labels are placed at the right side of the GIPs). 161

Table 7.4 Atomic, Bond and Path Indices Values for ACNH. 165

Table 7.5 Atomic, Bond and Path Indices Values for ACCONH. 165

Table 7.6 R_k and Q_k values for group ACNH and ACCONH. 167

Table 7.7 Group Interaction Parameters (GIPs) Table (in Kelvin) for Pharmaceutical Case Study (GIPs labeled with * are predicted using the CI-method while GIPs labeled with ¤ are set to zero). 167

Table 7.8 Group Interaction Parameters (GIPs) Table (in Kelvin) for Pharmaceutical Case Study with Fine Tuned Parameters (GIPs labeled with * are predicted using the CI-method while GIPs labeled with # are fine tuned. The labels are placed at the right side of the GIPs). 173

Table 7.9	Palm Oil Composition with its UNIFAC Group Representation.	176
Table 7.10	Group interaction parameters (in Kelvin) predicted using Original UNIFAC for Deodorization Process of Palm Oil.	177
Table 7.11	Group interaction parameters (in Kelvin) predicted using Original UNIFAC-CI for Deodorization Process of Palm Oil.	178
Table 7.12	Activity Coefficients at 250 ^{0}C and 3.5 mmHg Calculated using UNIFAC-CI and Reference UNIFAC Models.	178
Table 7.13	Group interaction parameters (in Kelvin) for Original UNIFAC for Sub-Case Study 2.	180
Table 7.14	Group interaction parameters (in Kelvin) for Original UNIFAC-CI for Sub-Case Study 2.	180
Table 7.15	Parameter Table (in Kelvin) of the Original UNIFAC Model with the Missing Parameters.	183
Table 7.16	Atom Stoichiometry and CI Values for ACOH-HCOO.	183
Table 7.17	Atom Interaction Parameters (AIPs) Needed for the Calculation of GIPs Related to $a_{CH2\text{-}ACNH}$ and $a_{ACNH\text{-}CH2.}$	184
Table 7.18:	Parameter Table (in Kelvin) of the Original UNIFAC Model with the Predicted Parameters (labeled with *).	185
Table 7.19	Group Interaction Parameters (in Kelvin) of Original UNIFAC-CI Model for Dimethyl Carbonate-Methanol System.	186
Table A.1	Atom Stoichiometry and Connectivity Indices (CI) Values for the Original UNIFAC-CI Groups. (nX refers to the number of atom X and $^0\chi^v$, $^1\chi^v$ and $^2\chi^v$ refers to valence connectivity index of zeroth, first and second order respectively).	209
Table A.2	Stoichiometric and Connectivity Indices (CI) Values for the Modified (Dortmund) UNIFAC-CI Groups. (nX refers to the number of atom X and $^0\chi^v$, $^1\chi^v$ and $^2\chi^v$ refers to valence connectivity index of zeroth, first and second order respectively).	211
Table B.1	Atom Interaction Parameters (AIPs) for the Original UNIFAC-CI (VLE) Model for Groups Involving Atoms C, H, O, N, Cl and S.	213

Table B.2 Atom Interaction Parameters (AIPs) for the Original UNIFAC-CI (VLE) with Q_{VLE} Model for Groups Involving Atoms C, H, O, N, Cl and S. 214

Table B.3 Atom Interaction Parameters (AIPs) for the Original UNIFAC-CI (VLE/SLE) Model for Groups Involving Atoms C, H, O, N, Cl and S. 216

Table B.4 Atom Interaction Parameters (AIPs) for the Modified (Dortmund) UNIFAC-CI (VLE) Model for Groups Involving Atoms C, O, H, Cl and S. 217

Table D.1 Atom Stoichiometry and CI Values for CH_2-ACCO. 222

Table D.2 Atom Interaction Parameters (AIPs) Needed for the Calculation of GIPs Related to $a_{CH2\text{-}ACCO}$ and $a_{ACCO\text{-}CH2}$. 223

Table D.3 Atom Stoichiometry and CI Values for CH_2-ACNH. 223

Table D.4 Atom Interaction Parameters (AIPs) Needed for the Calculation of GIPs Related to $a_{CH2\text{-}ACNH}$ and $a_{ACNH\text{-}CH2}$. 224

LIST OF FIGURES

Figure 1.1	Schematic Diagram of the Significance of the UNIFAC-CI Model.	3
Figure 2.1	Classification of Predictive Models for Mixture Properties.	10
Figure 2.2	Schematical Representation of the Group Contribution Approach.	13
Figure 2.3	Basic Methodology to Generate Missing GIPs.	31
Figure 2.4	Molecular Structure and Hydrogen-Suppressed Graph of HCOO Group.	31
Figure 3.1	Ratio of AARD of Reference UNIFAC and the UNIFAC-CI Models Against Fraction of the Datasets for Systems Containing the C, O, H Atoms. The Crossover Indicates the Percentage of the Data-Sets Which Are Better Correlated by the Reference Model. mg is the Number of Main Groups.	48
Figure 3.2	Average AARD (%) for Each Type of Systems Used in the Regression for Systems Containing Atoms C, O, H Between Reference UNIFAC and UNIFAC-CI.	52
Figure 3.3	Comparison of the Activity Coefficients Calculated Using the Original UNIFAC and UNIFAC-CI Methods with Experimental Data for (a) 1-Butanal-n-Heptane at 318 K, (b) 2-Methylpropanal-n-Heptane at 318 K, (c) 2-Methylpropanal-n-Heptane at 335 K, (d) 1-Pentanal-n-Heptane at 348K and (e) 1-Butanal-n-Heptane at 343 K.	49
Figure 3.4	Values of F with respect to Composition for (a) 1-Pentanal-n-Heptane at 348 K and (b) 1-Butanal-n-Heptane at 343 K.	53
Figure 3.5	Values of F with respect to Composition for (a) 1-Butanal-n-Heptane at 318 K, (b) 2-Methylpropanal-n-Heptane at 318 K and (c) 2-Methylpropanal-n-Heptane at 335 K.	54
Figure 3.6	Ratio of AARD of Reference UNIFAC and UNIFAC-CI Against Fraction of the Datasets for C, O and H Atoms Related Systems with (a) Weight 1, (b) Weight 5, and (c) Weight 10 (in the Objective Function) at Concentrations Lower Than 0.5 for Problematic Systems. The Crossover Indicates the Percentage of the Datasets which are Better Correlated by the Reference Model. mg is the Number of Main Groups.	56

Figure 3.7 UNIFAC-CI VLE Prediction for the Systems (a) 1-Butanal-n-Heptane at 343 K, (b) 2-Methylpropanal-n-Heptane at 335 K, (c) 1-Butanal-n-Heptane at 318 K, (d) 2-Methylpropanal-n-Heptane at 318K and (e) 1-Pentanal-n-Heptane at 348 K. 60

Figure 3.8 UNIFAC-CI SLE Prediction for the Systems: (a) p-Toluic Acid-Acetic-Acid (b) Anthracene-Cyclohexane (c) Epsilon-Caprolactone-1-Propanol (d) Ibuprofen-1-Octanol. 61

Figure 4.1 Overall Regression Procedure for the AIPs Parameter Estimation. 67

Figure 4.2 Correlation Errors in AARD (%) for CI-Models Compared with the Reference UNIFAC Model. 81

Figure 4.3 VLE Diagrams Predicted using Original UNIFAC, UNIFAC-CI (VLE) and UNIFAC-CI (VLE) with Q_{VLE} for Systems of: (a) Trans-2-Butene-1-Propanol at 364.50 K, (b) Acetic Acid Ethylbenzene at 725 mmHg, (c) Toluene-1-Butanol at 363.15 K and (d) 1-Propanol-Di-n-Butyl Ether at 323.15 K. 83

Figure 4.4 VLE Diagrams Predicted using Original UNIFAC, UNIFAC-CI (VLE) and UNIFAC-CI (VLE) with Q_{VLE} for Systems of: (a) n-Hexane-Aniline at 760 mmHg, (b) Acetonitrile-1-Propanol at 66.6 kPa. 84

Figure 4.5 VLE Diagrams Predicted using Original UNIFAC, UNIFAC-CI (VLE) and UNIFAC-CI (VLE) with Q_{VLE} for Systems of: (a) Acetonitrile-2-Methyl-1-Propanol at 760 mmHg and (b) Acetonitrile-Nitromethane at 333.15 K. 85

Figure 4.6 VLE Diagrams Predicted using Original UNIFAC, UNIFAC-CI (VLE) and UNIFAC-CI (VLE) with Q_{VLE} for Systems of: (a) Ethyl Acetate-Monochlorobenzene at 393.15 K, (b) 1-Pentene-Monochlorobenzene at 280 K, (c) Chloroform-1-Butanol at 760 mmHg and (d) Chloroform-n-Hexane at 328.15 K. 86

Figure 4.7 VLE Diagrams Predicted using Original UNIFAC, UNIFAC-CI (VLE) and UNIFAC-CI (VLE) with Q_{VLE} for Systems of: (a) Diethyl Sulfide-n-Heptane at 363.15 K (b) Diethyl Sulfide-n-Hexane at 338.15 K, (c) Dimethyl Sulfoxide-Methyl Ethyl Ketone at 95.3 kPa and (d) Dimethyl Sulfoxide-Methyl Isobutyl Ketone at 95.3 kPa. 88

Figure 5.1 Overall Regression Procedure for the AIPs Regression Work. 94

Figure 5.2 Average Correlation Errors of VLE Systems for the UNIFAC-CI (VLE/SLE) Model Compared with the Reference UNIFAC and UNIFAC-CI (VLE) with Q_{VLE} Models for Different Types of Systems. 118

Figure 5.3 Average Correlation Errors of SLE Systems for the UNIFAC-CI (VLE/SLE) Model Compared with the Reference UNIFAC and UNIFAC-CI (VLE) with Q_{VLE} Models for Different Types of Systems. 120

Figure 5.4 Parity Plots Between Experimental and Calculated Solute Composition using Original UNIFAC and UNIFAC-CI (VLE/SLE) Models for (a) Hydrocarbons, (b) Systems with C-O, O-C Atom Interactions and (c) Systems with O-O Atom Interactions. 121

Figure 5.5 Parity Plots Between Experimental and Calculated Solute Composition using Original UNIFAC and UNIFAC-CI (VLE/SLE) Models for (a) Nitrogenated Systems and (b) Chlorinated Systems. 122

Figure 5.6 SLE diagrams predicted using Original UNIFAC, UNIFAC-CI (VLE) with Q_{VLE} and UNIFAC-CI (VLE/SLE) models for systems of : a) Biphenyl-2-Ethyl-1-Hexanol, b) Phenanthrene-Methyl Ethyl Ketone, c) Naphthalene-2-Propanol, d) Camphor-gamma-Butyrolactone. 124

Figure 5.7 SLE diagrams predicted using Original UNIFAC, UNIFAC-CI (VLE) with Q_{VLE} and UNIFAC-CI (VLE/SLE) models for systems of : a) Naphthalene-Acetic Acid, b) Acetic Acid-Benzene, c) Aspirin-2-Butanone, d) Aspirin-Isopropyl Acetate. 125

Figure 5.8 SLE diagrams predicted using Original UNIFAC, UNIFAC-CI (VLE) with Q_{VLE} and UNIFAC-CI (VLE/SLE) models for systems of : a) Anthracene-Acetonitrile, b) Naphthalene-Nitrobenzene, c) Naphthalene-1,1-Dichloroethane, d) Indane-1,2-Dichloroethane. 127

Figure 5.9 Problematic VLE diagrams predicted using Original UNIFAC, UNIFAC-CI (VLE) with Q_{VLE} and UNIFAC-CI (VLE/SLE) models for systems of: (a) 1,2-Dimethoxyethane-Toluene at 350 K, (b) Phenol-Styrene at 373 K, (c) Acetonitrile-Ethanol at 393 K and (d) n-Heptane-1,4-Dichlorobutane at 298 K. 129

Figure 5.10 Problematic VLE diagrams predicted using Original UNIFAC, UNIFAC-CI (VLE) with Q_{VLE} and UNIFAC-CI (VLE/SLE) models for systems of: (a) Dichloroethane-Toulene at 298 K, (b) Dichloromethane-1-Hexene at 298 K, (c) 1-Pentene-Monochlorobenzene at 320 K and (d) Dichloromethane-Ethyl Acetate at 348 K. 131

Figure 5.11 Problematic SLE diagrams predicted using Original UNIFAC, UNIFAC-CI (VLE) with Q_{VLE} and UNIFAC-CI (VLE/SLE) models for systems of: (a) Naphthalene-1-Butanol, (b) Ibuprofen-n-Heptane, (c) Ibuprofen-Benzene and (d) Pyrene-1,4-Dioxane. 133

Figure 5.12 Problematic SLE diagrams predicted using Original UNIFAC, UNIFAC-CI (VLE) with Q_{VLE} and UNIFAC-CI (VLE/SLE) models for systems of: (a) Epsilon Caprolactone-Toluene, (b) Benzoic Acid-1,4-Dioxane, (c) Benzoic Acid-2-Propanol and (d) Salicylic Acid-1-Octanol. 134

Figure 5.13 Problematic SLE diagrams predicted using Original UNIFAC, UNIFAC-CI (VLE) with Q_{VLE} and UNIFAC-CI (VLE/SLE) models for systems of: (a) Ibuprofen-Ethyl Acetate, (b) Naphthalene-1,2-Dichloroethane, (c) Ibuprofen-Monochlorobenzene and (d) Benzoic Acid-Chloroform. 136

Figure 6.1 Overall Regression Procedure for the AIPs Regression Work. 143

Figure 6.2 Correlation Errors in AARD (%) for Modified (Dortmund) UNIFAC-CI Model Compared with the Reference Modified (Dortmund) UNIFAC Model. 146

Figure 6.3 VLE Diagrams Predicted using Modified (Dortmund) UNIFAC and Modified (Dortmund) UNIFAC-CI (VLE) for Systems of: (a) n-Heptane-Monochlorobenzene at 323 K and (b) n-Hexane-n-Butyl Chloride at 94.4 kPa. 148

Figure 6.4 VLE Diagrams Predicted using Modified (Dortmund) UNIFAC and Modified (Dortmund) UNIFAC-CI (VLE) for Systems of: (a) Benzene-Carbon Disulfide at 298.15 K and (b) Ethyl-Mercaptan-Propylene at 323 K. 149

Figure 7.1 Molecular Structure of Hydrogen Peroxide. 150

Figure 7.2 Hydrogenation Step of the Anthraquinone Process. 151

Figure 7.3 Oxidation Step of the Anthraquinone Process. 152

Figure 7.4 Simplified Process Flow Diagram of a Hydrogen Peroxide Production using the Anthraquinone Process. 152

Figure 7.5 The ACCO group in Anthraquinone. 153

Figure 7.6 Molecular Structure and Hydrogen-Suppressed Graph of ACCO Group. 154

Figure 7.7 SLE Diagram for (a) Anthraquinone 1 and Anthraquinone 2 with Solvent 1, (b) Anthraquinone 1 and Anthraquinone 2 with Solvent 2 and (c) Anthraquinone 1 and Anthraquinone 2 with Solvent 3. 157

Figure 7.8 SLE Diagram for (a) Anthraquinone 1 and Anthraquinone 2 with Solvent 4, (b) Anthraquinone 1 and Anthraquinone 2 with Solvent 5. 158

Figure 7.9 SLE Diagram for (a) Anthrahydroquinone 1 and Anthrahydroquinone 2 with Solvent 1, (b) Anthrahydroquinone 1 and Anthrahydroquinone 2 with Solvent 2 and (c) Anthrahydroquinone 1 and Anthrahydroquinone 2 with Solvent 3. 159

Figure 7.10 SLE Diagram for (a) Anthrahydroquinone 1 and Anthrahydroquinone 2 with Solvent 4, (b) Anthrahydroquinone 1 and Anthrahydroquinone 2 with Solvent 5. 160

Figure 7.11 Overall SLE Diagrams for (a) Anthraquinones (A) and Different Solvents (S) and (b) Anthrahydroquinones (AH) and Different Solvents (S). 161

Figure 7.12 Examples of Pharmaceutical Compounds. 163

Figure 7.13 Molecular Structure and Hydrogen-Suppressed Graph of ACNH Group. 164

Figure 7.14 Molecular Structure and Hydrogen-Suppressed Graph of ACCONH Group. 164

Figure 7.15 SLE Diagrams of APIs with Solvents, a) Heptane, b) Tetrahydrofuran, c) Acetone, d) Methyl Ethyl Ketone, e) Methyl Isobutyl Ketone and f) Acetonitrile Predicted using Original UNIFAC-CI using Fine Tuned GIPs. 168

Figure 7.16 SLE Diagrams of APIs with Solvents, a) Isobutanol, b) Methanol and c) Toluene Predicted using Original UNIFAC-CI using Fine Tuned GIPs. 169

Figure 7.17 SLE Diagrams of APIs with Solvents, a) Heptane, b) Toluene, c) Tetrahydrofuran and d) Acetone Predicted using Original UNIFAC-CI using Fine Tuned GIPs. 171

Figure 7.18 SLE Diagrams of APIs with Solvents, a) Methyl Ethyl Ketone, b) 2-Propanol, c) Butanol and d) Methanol Predicted using Original UNIFAC-CI using Fine Tuned GIPs. 172

Figure 7.19 Overall SLE Diagrams for (a) API 1 and Different Solvents and (b) API 2 and Different Solvents. 174

Figure 7.20 Molecular Structure of Tocopherol and the Subgroup C-O. 179

Figure 7.21 VLE Phase Diagrams of (a) Methyl Laurate-Methanol, (b) Methyl Myristate-Methanol, (c) Methyl Oleate-Methanol, (d) Methyl Laurate-Ethanol, (e) Methyl Myristate-Ethanol, (e) Methyl Oleate-Ethanol. 181

Figure 7.22 VLE Phase Diagrams of (a) Methanol-Glycerol, (b) Ethanol -Glycerol, (c) 1-Propanol-Glycerol, (d) 2-Propanol- Glycerol, (e) 1-Butanol- Glycerol, (e) Water- Glycerol. 182

Figure 7.23 VLE Diagram of Ethyl-Formate and Phenol at (a) 300 K and (b) 320 K. 186

Figure 7.24 VLE Diagram of Dimethyl Carbonate-Methanol at 1 atm. 187

CHAPTER 1

INTRODUCTION

1.1 Background of the Project

Prediction of properties is important in chemical process and product design. According to Gani and O'Connell [1], three different functions of properties have been identified in the solution of Computer Aided Process Engineering (CAPE) problems. First, properties can be used to provide service where the property models generate the needed properties whenever requested by the user. This role is important in process simulation problems. The second function of properties in the solution of CAPE problems is to provide service and also advice where the generation of those relevant properties can also provide feasible problem formulations and to avoid unnecessary calculations [2]. This role is essential in process design and synthesis. For example, if we consider separating a binary mixture, first we can examine the mixture properties through the phase diagram and we can see whether the system has an azeotrope or not. Then, we can decide whether to use a direct distillation (if the system does not have an azeotrope) or extractive distillation (if the system has an azeotrope). Here we can see how the properties are providing the advice for the user. Further examples on the role of properties to provide service and service/advice can be found in [3]. The third function identified is to provide service, advice and also to solve the given CAPE problems. This role is important in process and tools integration such as distillation column design and pinch technology for heat integration where usually the properties can provide the solution strategies. These functions showed the significance of properties in chemical engineering problems and therefore property models which are reliable and have a wide application range need to be available and further developed.

Group-contribution (GC) methods are useful tools but there is a need to validate and improve their accuracy when complex chemicals are present in the mixtures. In GC methods, chemicals or molecules are considered to be consisting of different functional groups which are smaller in size. These fragments of groups are used as building blocks in order to describe the whole molecular structures. In the case of mixtures, they are also considered to be consisting of functional groups instead of molecules. The properties of the chemicals/molecules or mixtures are considered as an additive function of parameters which are related to each of the groups which describe them. Those parameters are considered as the contributions of each of the group fragments to the properties of the compounds or mixtures under study. Normally, the parameters or contributions of each of the functional group are obtained through regression over a set of experimental data of the investigated properties involving a range of chemicals and mixtures. Some examples of GC methods which have been developed for the estimation of properties of pure compounds include those published by Joback and Reid [4], Lydersen [5], Ambrose [6], Constantinou and Gani [7] and Marrero and Gani [8-9]. On the other hand, many GC based property models have also been developed to predict properties of mixtures mainly to predict the non-ideality of the liquid phase using activity coefficients which includes ASOG [10-11], Original UNIFAC [12], Modified UNIFAC (Dortmund) [13] and PSRK [14].

From those lists of GC based models, UNIFAC is one of the most widely used models which is important in process and product design especially in predicting the liquid phase activity coefficients of mixtures. In the production of chemicals, phase equilibria plays an important role especially if it involves separation processes such as distillation which can contribute to the bigger part of the overall operating cost. Therefore, in order to describe the phase equilibria of the processes in the best possible way, property models which are reliable and have a wide application range are desirable. However, the UNIFAC model is known to have several limitations and one of them is the missing group interaction parameters (GIPs) due to lack of experimental data. In the UNIFAC parameters table published by Hansen et al. [15], about 53% of the total possible parameters are missing. The UNIFAC model cannot be used when for a certain mixture the needed GIPs are not available. The best way to fill the missing gaps is by conducting experiments and using those experimental data, the missing GIPs of the UNIFAC model are then estimated. However, performing experiments can be cost, time and energy consuming and sometimes impossible due to safety issues for example. For these reasons experimental data for many systems are not available thus leaving many gaps in the UNIFAC parameter table. Hence, an alternative to experiments approach that could save money, time and also energy would be a great solution in order to increase the application range of the UNIFAC model.

Recently, a combined GC and atom connectivity approach that is able to extend the application range of property models has been developed for mixture properties using UNIFAC as the host GC based model [16-17]. The model is called UNIFAC-CI model. This so-called GC^{Plus} approach is a hybrid model which combines GC based models and valence connectivity indices (CI) which are molecular descriptors. The main idea is the use of CI to describe molecular fragmentations that relates the molecular interactions with the molecular structures. The needed values of the group interaction parameters (GIPs) are normally obtained by fitting phase equilibrium data. As an alternative to performing measurements, values of the missing GIPs can be predicted through the GC^{Plus} approach. The predicted values for the GIPs are then used in the UNIFAC model to calculate activity coefficients. In the development of this method, a relationship have been established between the GIPs and the i) atom stoichiometry of the UNIFAC groups, ii) the valence connectivity indices which can be calculated from each UNIFAC groups and iii) atom interaction parameters (AIPs) which are regressed against only available data that are currently in the literature, instead of GIPs [18]. In this way, just like many compounds and mixtures that can be described by different functional groups, UNIFAC group interactions can be described by atom interactions which are complimented by the information provided by the connectivity indices. The experimental data used in the parameter regression are consisting of different compounds. In an atomistic scale, each compound is formed by different atoms. Therefore, missing GIPs which are formed by the same set of atoms can be possibly represented and predicted. A schematic diagram of this idea is illustrated in Figure 1.1. In this PhD project, the initially developed UNIFAC-CI [16-18] model which has been developed through the GC^{Plus} approach is analyzed in order to further improve the performance and reliability of the model. In addition to that, the analyzed model will be further developed in order to improve predictions of phase equilibria not only for vapor-liquid equilibrium (VLE) but also for solid-liquid equilibrium (SLE) where the AIPs will be regressed simultaneously against VLE and SLE data. Besides increasing the application range of Original UNIFAC, other UNIFAC variants can also be used as a host model and in this work the GC^{Plus} approach have been implemented to the Modified UNIFAC (Dortmund) model, continuing the work which has been initiated by González [18]. Finally, the application of the analyzed and developed

UNIFAC-CI will be highlighted in innovative process and product design problems and also in cases where no parameters and/or experimental data exist thereby providing a true and reliable predictive power to properties estimation in process-product design and synthesis.

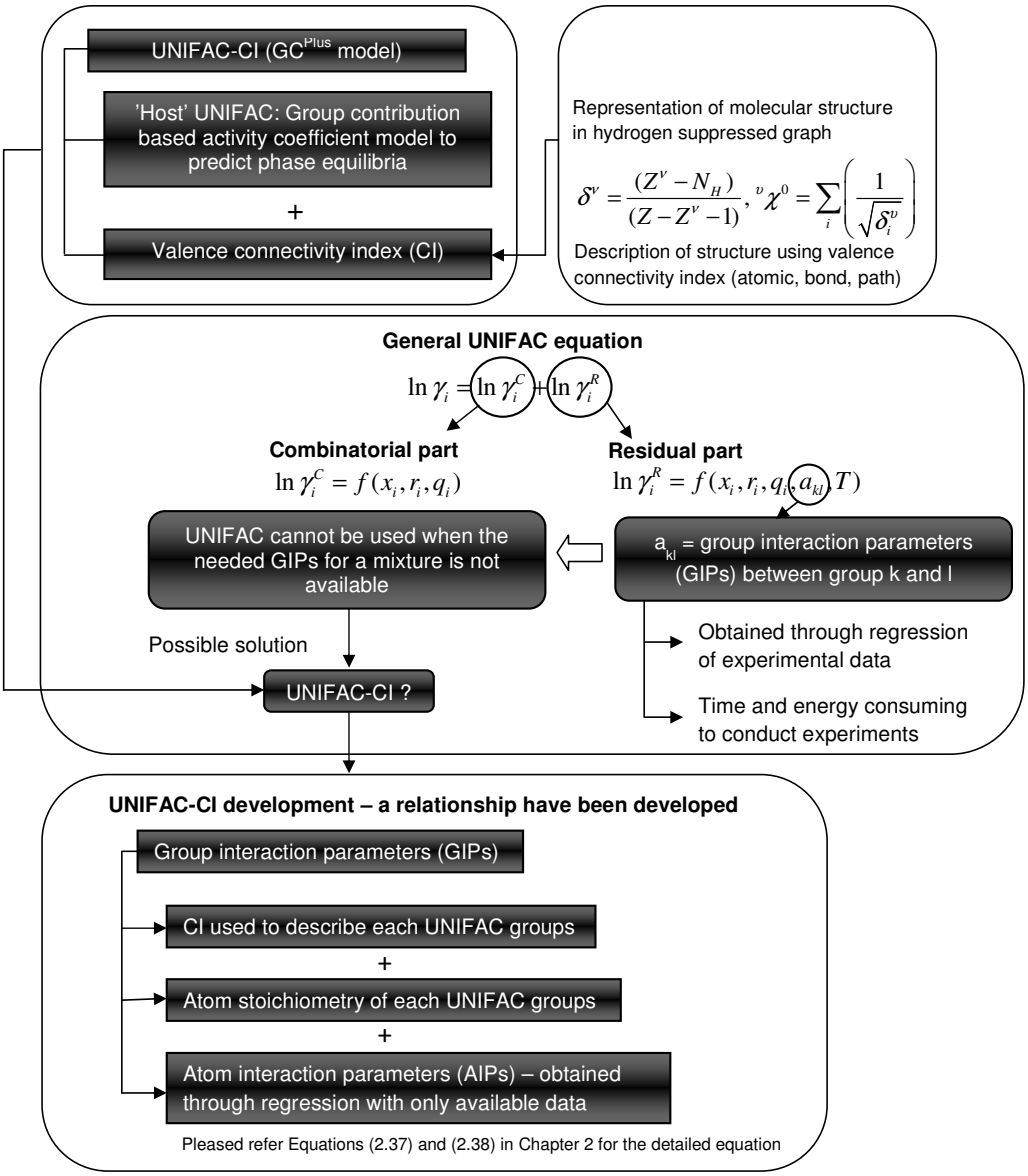


Figure 1.1: Schematic Diagram of the Significance of the UNIFAC-CI Model.

1.2 Role and Importance of UNIFAC in Process and Product Design

In the production of chemical products, typically many separation processes are involved such as distillation, absorption and extraction which require thermodynamic models to describe the phase equilibria of the process. Usually the phase equilibria are measured at a later design stages. However, for preliminary design purposes in order to screen the options and alternatives that are available such as the alternatives for the best separation techniques, an approximate and predictive model which can be used in a broad range is desirable. Such predictive thermodynamic models can ideally provide the estimation of phase equilibria in a fast and easy way for a wide range of compounds and mixtures.

The importance of this kind of models is not only of paramount importance in process design where the chemical system is known but the property values need to be predicted or generated. It is also essential in product design where the target properties are known but the chemical system is unknown. The problems which can be encountered in product design can be related to the design of molecules, mixtures, blends and formulations such as solvents, fuels and emulsions. There can be thousands of compounds and millions of possible combinations of compounds that are available and it can be difficult and complex to find which chemical system satisfies the set of properties which have been targeted. In this situation, a property model which can predict the property/phase equilibria value in a fast and efficient way is vital so that the best possible molecules or mixtures which satisfy the target properties can be narrowed down and thus reducing the number of alternatives and removing those compounds of mixtures that are not feasible [2].

The requirement of a property model which can provide predictions of properties in a fast and efficient way, for a broad range of compounds and conditions and thus can be used to solve problems in process and product design have been the driving force in the development of the UNIQUAC Functional-Group Activity Coefficients (simply known as UNIFAC). It is a group contribution (GC) based model to predict liquid phase activity coefficients involving for example, vapor-liquid equilibrium (VLE), solid-liquid equilibrium (SLE) and liquid-liquid equilibrium (LLE) calculations. Besides predicting phase behavior, GC methods such as UNIFAC have been also used in solving problems related to estimation of solvent effects on chemical reaction rates, calculation of critical micelle concentrations for surfactants solutions, calculations of flash points of flammable liquid mixtures and the calculations of viscosities of liquid mixtures, estimation of excess enthalpies and many more [2].

Since the first article published in 1975 by Fredenslund et al. [12] introducing the UNIFAC method, more than 2200 articles [19] have been published until July 2012 involving reviews, direct application and parameter updates and also for parameter use in regressions, modeling, simulations, comparison of performance with other models and many more. UNIFAC has been proven to be a reliable predictive model, with parameters and groups constantly updated and improved as new experimental data appear. Besides that, UNIFAC is computationally efficient, easy to program and also fast in calculations and because of this it is widely used by commercial simulators such as ASPEN, CHEMCAD, HYSIS, PROII, ProSim and many more for the simulation of chemical processes.

In 2010, the Working Party on Thermodynamic and Transport Properties of the European Federation of Chemical Engineering (EFCE) has carried out an investigation on the industrial requirements for thermodynamic and transport properties, where 28 companies have participated [20]. One of the main results of the investigation is that eventhough new methods such as SAFT and COSMO-RS have been increasingly used and further developed, the use of traditional method such as cubic equations of state and UNIFAC group contribution approach cannot be replaced and is still very much needed.

1.3 Why Valence Connectivity Index? Previous Work and Status

Molecular descriptors such as valence connectivity indices (CI) are topological indices used to translate molecular structure through mathematical characterization into numerical index to encode information about size, branching, cyclization, unsaturization and hetero-atomic content of molecules. Besides this important information which is embedded in CI, they are also convenient due to their simplicity in calculations, two-dimensional representation of molecular structure and their benefit when used in routine engineering calculations.

As we have discussed in the previous section, one of the limitations of the GC based property models is the missing group contributions and interaction parameters which are ideally obtained by regressing them against experimental data. CI can be used as complimentary information to predict those missing parameters. This idea was first developed by Gani et al. [21] through the establishment of a methodology to predict, with the aid of valence connectivity index (CI), the missing group contribution of the Marrero and Gani [8] pure component property prediction model. This hybrid model is also known as the GC^{Plus} approach. This article [21], published on the introduction of this method and approach has been cited 26 times until July 2012 [19].

The same approach, was further extended for the prediction of pure component properties of polymers involving properties of the glassy and rubbery amorphous volume, amorphous volume, crystalline volume, glass transition temperature, solubility parameter and refractive index at 298.15 K [22]. The developed GC^{Plus} model for polymers have also been used in Computer Aided Molecular Design (CAMD) problems for the design of polymer products where the polymer repeat unit structures properties can be predicted according to specified constraints [23-24]. Furthermore, the hybrid model combining GC and CI has been used for the calculation of solid solubility for solvent selection where CI-based parameters are generated in order to be able to predict the three Hansen solubility parameters which are available for organic chemicals with C, H, O, N, Cl, S, F, Br, I and P atoms [25]. The same approach was also further developed for the estimation of surface tension and viscosity [26]. Recently, the GC^{Plus} approach which has been implemented for the Marrero and Gani GC model in [21] has been updated and improved and includes calculations of the uncertainties of the estimated property values [9].

Besides pure component GC based models, the GC^{Plus} approach has been also implemented to GC models for the prediction of mixture properties. Valence connectivity indices have been used to predict the missing group interaction parameters (GIPs) of the UNIFAC model. A first investigation has been done by González et al. [16-18] for developing this hybrid model (called UNIFAC-CI). In this work, two UNIFAC models have been used in the development, the Original UNIFAC model

developed by Fredenlund et al. [12] and the Modified (Dortmund) UNIFAC model by Weidlich et al. [13]. A relationship has been established between the GIPs which are considered as missing with the atom stoichiometry of the UNIFAC group and CI which can be easily calculated for each of the group and also the atom interaction parameters (AIPs) which are obtained through regression with experimental data. For the Original UNIFAC-CI model, the atom interaction parameters (AIPs) have been regressed against VLE data for systems involving atom C, H, O, N, Cl and S while for Modified (Dortmund) UNIFAC AIPs for systems containing C, H, O and N atoms are available and the overall performance is published in [18]. The current Ph.D project continue on the base which have been developed by González et al. in order to analyze, further develop and highlight the application of UNIFAC-CI in important process-product design problems.

1.4 Motivation, Scope and Objectives

This Ph.D project is devoted to the analysis and further development of UNIFAC-CI models and also to highlight the application of the analyzed and developed models in innovative process-product design problems. As discussed above, property models such as group contribution models are useful for preliminary calculations of compound properties and phase behavior of mixtures. UNIFAC, which has been widely used by the chemical engineering community, is still relevant and advanced models which maybe more accurate such as SAFT still are not able to replace the UNIFAC model in tackling process and product design problems at the early stage of the problem solving step. However, UNIFAC has some limitations and one of them is the missing group interaction parameters (GIPs) which without them, it cannot be used to predict activity coefficients of mixtures and thermodynamic properties in general.

Previous efforts have been reported for extending the application range of UNIFAC. One of them [18] was by introducing valence connectivity indices which can be used to predict the missing GIPs through an established relationship between the GIPs and the CI. Here we use the number of different atoms and atom interaction parameters (AIPs) which are obtained by matching it with only the available data. If this approach is proved successfully, we can minimize the experiments that must be conducted. In order to further validate this idea, it is important to further analyze the initially developed UNIFAC-CI model so that its performance can be improved. Its capabilities/limitations are investigated on different applications.

The scope of the project is limited to organic systems which can be found in many industrial biochemical and chemical processes involving distillation, crystallization, extraction and pervaporation. Electrolytes, non-condensable gases and polymers are not considered in this work. Besides that, the project will focus on two UNIFAC activity coefficient models involving two main UNIFAC models (Original UNIFAC, Modified (Dortmund) UNIFAC) for developing hybrid models based on the GC^{Plus} concept initially developed by Gani et al. [21] and further extended to mixtures by González et al. [16-18]. Moreover, this project focuses on developing models capable of predicting different types of phase equlibria system involving both VLE and SLE and to illustrate applications for process and product design problems involving those types of phase behavior.

The objectives of this current work is, first to revisit the UNIFAC-CI model developed by González et al. [16-18] and to further analyze the model. The purpose of this analyses is to improve the

performance by investigating and understanding why the prediction or correlation is good for certain systems but less satisfactory for others. Once the problematic systems identified, possible solutions are implemented which hopefully will improve the performance of the models for all systems. In addition, the second objective of this project is to further develop UNIFAC-CI. For the case of Modified (Dortmund) UNIFACI-CI, developed [18] model parameters are available for systems related to atoms C, H, O and N. This model is further developed to also include parameters with respect to atoms Cl and S which are also included in Original UNIFAC-CI. Furthermore, both UNIFAC-CI models (Original and Modified (Dortmund)) are further developed by regressing AIPs simultaneously against VLE and SLE data. This is done in order to see whether the predictions of the solid solubility can be improved while maintaining the accuracy for the VLE calculations. Finally, the third objectives of this project is to illustrate the application of the developed CI-models to relevant process and product design problems involving pharmaceutical and lipid systems and also design of a working solution in the production of hydrogen peroxide. The final result from this project will include improved UNIFAC parameter table for each host models and also an analysis on how the parameters can be estimated based on combinations of VLE and SLE data.

1.5 Thesis Organization

This Ph.D thesis is divided into seven chapters including this current introduction chapter. Here, the background of the project is discussed. In addition, group contribution (GC) based property models are briefly described. We highlighted the role and importance of UNIFAC in process and product design studies. The limitation of the UNIFAC model with the missing group interaction parameters (GIPs) is discussed where molecular descriptors such as connectivity indices (CI) can be used to predict the missing parameters without need of additional experimental data. The previous work done using the same concept is presented. The previous work starts from the development of a hybrid model which combines GC model and CI used for pure component properties estimation and some applications to mixtures properties involving UNIFAC as the host model. Finally the motivation and objectives of the current project are discussed.

Chapter 2 will present the theoretical and conceptual background of different aspects of the project which are considered building blocks of the overall work. In the first part of Chapter 2, a review of available property prediction models for mixtures properties is highlighted ranging from quantum chemical to group contribution methods. GC methods are discussed next especially UNIFAC and its many variants illustrating their purpose of development, differences, accuracy and applications. Furthermore, applications of UNIFAC from literatures on process and product design will be presented. Then, the background of molecular description theory especially the connectivity index is discussed together with its importance in engineering. In the next section, the development of the GC^{Plus} approach which combines GC based host models and valence connectivity indices is presented. Chapter 2 concludes with a description of the methods and tools which have been used in analyzing the developed models.

In Chapter 3, a first analysis of the Original UNIFAC-CI model is presented where the developed UNIFAC-CI model [16-18] is revisited and its performance is analyzed in order to see why the

model works well in certain systems while not performing well in other systems. The analysis and how the predictions of those problematic systems are improved are highlighted in this chapter.

The UNIFAC-CI models are further developed by regressing the atom interaction parameters simultaneously against VLE and SLE data (for systems related to C, H, O, N, Cl and S atoms). This is done in order to see whether the prediction of SLE systems can be improved (instead of the usual way of just extrapolating the predictions for SLE systems using VLE generated parameters) and still maintaining the accuracy of predictions of VLE systems. These further developments are reported in Chapter 4 for the Original UNIFAC-CI model and in Chapter 5 for the Modified (Dortmund) UNIFAC-CI model. Additionally in the latter chapter, the initial developments of the model with only VLE generated parameters are also discussed.

In Chapter 6, the developed CI-models are applied in innovative process-product design problems involving a working solution design, pharmaceutical and lipid systems highlighting the importance of GC^{Plus} based property models for mixtures for solving problems which could not be handled before. Finally, the Ph.D thesis is concluded with a summary of the main contributions of this work and a short presentation of future work and recommendations for improvements and future developments.

CHAPTER 2

THEORETICAL/CONCEPTUAL BACKGROUND AND ANALYSIS METHODS/TOOLS

2.1 Review of Available Property Prediction Models for Mixture Properties

The prediction of properties is very important since normally in process design engineering, not all the needed data at all different conditions, can be measured. In order to predict those properties, property models need to be developed. These models can be divided into different classes, depending on the level of empiricism, as highlighted in Figure 2.1. For the theoretical computational chemistry methods, calculations are made directly from first principles and ideally no data are needed. The methods involve quantum and statistical mechanical calculations. Between the theoretical and empirical approaches are the semi-empirical methods which use various forms of equations or models. Here some adjustable parameters are needed which are determined from experimental data. Examples of this kind of models include the SAFT equations of state, group contribution methods and corresponding states formulations which may be predictive within the range of data regressed.

Property models can also be categorized into pure component property prediction and mixture property prediction methods. For pure compounds, properties such as boiling point, melting point, critical properties, vapor pressure, heat of formation, heat of vaporization and others have been modeled. An extensive compilation and review of these prediction methods is given by Poling, Prausnitz and O' Connell [27]. On the other hand, for mixtures, phase equilibria properties are important in chemical engineering calculations involving vapor liquid equilibrium (VLE), liquid-liquid equilibrium (LLE) and solid-liquid equilibrium (SLE). In process plants, separation processes involving these phase equilibria condition contribute to the majority of the total operating cost and therefore they need to be properly predicted or approximated [28].

Predictive models are very useful. They can be used to calculate properties of a certain system under study without having any previous knowledge of that system. This system can be for example a mixture at certain temperature and/or pressure with two or more compounds. In the next sub sections, it is intended to discuss different types of property prediction model which can be used to predict mixture properties. These include (i) quantum mechanics/quantum chemical methods, (ii) group contribution methods and (iii) quantitative structure-activity relationship (QSAR)/quantitative structure-property relationship (QSPR) methods. These three type of models are different in terms of the level of empiricism appearing in the model construction, their assumptions and, their background theory as well as the advantages and limitations in their application.

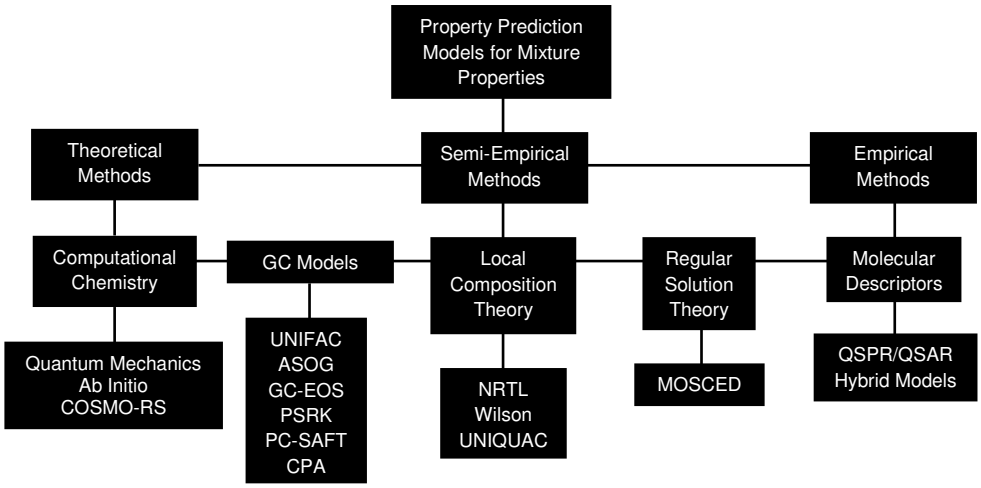


Figure 2.1: Classification of Predictive Models for Mixture Properties.

2.1.1 Quantum Mechanics / Quantum Chemical Methods

Initially the quantum mechanical methods have been primarily used for the calculation of heats of formation, heat capacities, heats of reaction, like molecular conformations, reaction pathways and transition states and most of them are restricted to a single molecule or an ideal gas [28]. However, it is now able to predict thermophysical properties and phase behavior of fluids and mixtures that are not ideal gases. These types of methods are based on calculations using first principles which are supposed to require less empirical information. The starting point for the quantum mechanical methods is the Schrödinger equation which cannot be solved exactly for multi-electron systems and approximations need to be made. In order to do this, there are a lot of software packages available which include ab initio and density functional calculation methods such as Gaussian, Turbomole and Gamess [29]. Furthermore the level of calculations needs to be decided and this is not an easy task to do. Approximations can be made according to the level of theory used and the accuracy with which the electron density is represented. In order to obtain the best results, high level of theory and accurate representation of the electron densities need to be used. However, using the coupled cluster method which is a high level of theory, the calculations will be very computationally intensive. Extrapolation methods with some empirical corrections can be used but the chemical accuracy is not sufficient for phase behavior calculations. With this kind of accuracy, it is not possible yet to predict phase equilibria directly from first principles using quantum mechanical methods [29]

However, the quantum mechanical method can be used somewhat indirectly for performing phase equilibria calculation. In this part, three approaches of using the quantum mechanical method are highlighted [30]. In the first approach, computational quantum mechanics is used to generate information on the multidimensional potential energy surface for the interaction between a pair of molecules (as a function of intermolecular separation and relative orientation). Using the generated potential, second virial coefficient can be calculated and using molecular simulation, the

thermodynamic properties and phase equilibria of a mixture can be predicted. This approach is very computationally intensive and also limited to small molecules. A detailed discussion on this approach can be found in [29] and [30]. In a less computationally intensive approach, the quantum mechanical method can be used to improve group contribution approaches. This can be done by providing corrections based on the charge and dipole moment of each functional group that is unique to the molecule in which it appears. In the group contribution approach, the behavior of a functional group is assumed to be the same no matter for which compound/mixture it is used. This is essentially incorrect whenever two strong polar groups are close to each other on the same molecule which is also referred to as the proximity effect. In addition to this advantage of the quantum mechanical method, it can be also used to determine parameters in existing thermodynamic models. An example of this is seen in the work of Wolbach and Sandler [31] where they used molecular orbital quantum mechanical calculations to determine the association parameter values for the Statistical Association Fluid Theory (SAFT) equation of state for mixtures containing water, methanol and other compounds.

The third approach is to use the quantum mechanics-based continuum solvation (polarizable) models to predict excess Gibbs energies, activity coefficient and phase behavior. This is known as the Conductor-Like Screening Model (COSMO-RS) developed by Klamt and co-workers [32-33]. The fundamental idea of this approach is that a molecule is divided into very small surface elements and the charge density of each surface element is obtained by using a quantum electrostatic calculation. The special characteristic of each molecule is its sigma profile that is a representation of charge density versus likelihood of occurrences. Using this sigma profile and a statistical-mechanical analysis, excess Gibbs energy at any composition can be computed. This model has several advantages which includes ability to handle the proximity effect and distinguish between isomers. Moreover, there are only a few adjustable parameters which respect to the diameter of each type of atom, parameters related to hydrogen bonding and the area of a surface element which are needed for the calculation of the sigma profiles and it does not require a large database like the group contribution method. This model has been the most useful and successful quantum mechanics-based model for thermodynamic property calculations.

2.1.2 Group-Contribution Methods

In group contribution methods, the molecular structure is decomposed into building blocks and the property of that molecule is estimated by the summation of the contributions of these building blocks. The building blocks refer to functional groups. The assumption is that a property value of any group has the same contribution in all compounds where it appears. Moreover, the property value of the compound is a function of the contributions of all the groups needed for a unique representation of the molecular structure of the compound.

The advantage of using group contribution methods is that compared to quantum mechanical methods, they are simple and easy to use. For example, to calculate the standard heat of vaporization of a compound, only a simple linear summation of the contributions of the structural groups representing the molecular structure of that compound is needed. They are computationally simple to use, and in a predictive way because the same functional group can be used to represent molecular structure of more than one compound. They are thus widely used and are suitable for

process engineering design and simulation. Group contribution methods are particularly useful when qualitatively correct predictions of the properties are needed and the demand for quantitative accuracy is not very high. This is due to the limitations that the group contribution methods have as they are unable to distinguish between isomers and thay are also unable to capture the proximity effects when many strong polar groups are present in a molecule. Moreover, group contribution methods are limited to compounds for which the needed groups have been defined and parameterized through fitting to experimental data Thus the properties of some compounds cannot be predicted if they cannot be represented by the existing groups. In order to overcome those limitations with proximity effects, higher group contributions have been introduced in the work of Constantinou and Gani [34], Marrero and Gani [8] and also Kang et al [35].

The details on the background and the development of these group contribution approaches are further discussed in section 2.2 since it is the background of the models specifically used in this PhD project.

2.1.3 Quantitative Structure-Activity Relationship (QSAR) / Quantitative Structure-Property Relationship (QSPR) Methods

The QSAR/QSPR methods have been developed by correlating chemical structures with a defined property of activity. QSPR modeling has the potential to provide reliable property estimates based on detailed information regarding chemical structures. However it does not completely eliminate the need for chemical synthesis or experimental validation, but only a smaller amount of chemicals need to be tested. The advantage of this method to model the behavior or organic chemicals has brought the attention of researchers working in the field of computer-aided molecular design to further investigate the relationship between the chemical structure and thermo-physical behavior.

The main argument of this QSPR method is that the three-dimensional structure of a molecule encodes all the information which determines the properties of that molecule. Once the structural information of that molecule can be decoded, then its properties could be determined from the knowledge of the structure. Quantum mechanics is used to construct the molecule in terms of a series of molecular descriptors and then using those descriptors, correlations for the property are developed. The basic steps for the development of QSAR/QSPR models involve (i) structure generation, (ii) structure optimization, (iii) descriptor generation, (iv) descriptor reduction, (v) QSPR model development and, (vi) QSPR model validation [36].

QSPR approaches have been used mainly for the prediction of pure compound physical properties and some thermodynamic properties such as vapor pressures, boiling point, melting point etc [36]. Only limited work has been done for mixture properties such as the estimation of the activity coefficients at infinite dilution of aqueous systems, partition coefficients, solubilities and Henry's law constants. Since QSPR can only estimate properties at a single temperature, Ravindranath et al. [36] have proposed a hybrid approach integrating excess Gibbs energy models such as NRTL and UNIQUAC with the QSPR method by developing structure-based model parameters using the QSPR approach for those activity coefficient models. This hybrid model is capable of giving a priori prediction of VLE phase behavior. Furthermore, Bunz et al. [37] have implemented the QSPR method to model carbon dioxide-hydrocarbon phase behavior using an equation of state.

2.2 Background of Group Contribution Methods

The basic concept of the Group Contribution (GC) methods is that the compounds or mixtures under study are considered as ensembles of groups which are the building blocks and can be either atoms [38], bonds [39-40] or functional groups [4, 39-41]. The idea is that the contributions of these groups to a specific chemical property are the same no matter what compounds or mixtures they representing. An example of a schematic representation of the GC approach is shown in Figure 2.2 for n-hexane and 1-butanol. The molecular structure of n-Hexane can be represented by 2 CH_3, 4 CH_2 groups while 1-Butanol can be divided into 1 CH_3, 3 CH_2, and 1 OH. Both structures use the same type of groups to define them. In this GC approach, the contributions of the CH_3 and CH_2 to the properties of n-Hexane and 1-Butanol are considered to be the same.

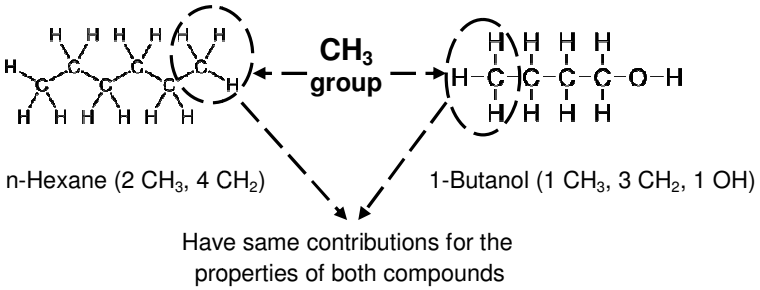


Figure 2.2: Schematical Representation of the Group Contribution Approach.

In order to predict the properties P, the additivity rule is implemented defined as,

$$f(P) = \sum_i N_i C_i \tag{2.1}$$

In Equation (2.1) above, C_i represents the contributions of atoms, bonds and first order functional group while N_i is the number of occurrences of each of those groups in defining the compounds or mixtures.

2.2.1 Pure Component Group Contribution Methods

For the estimation of pure component properties, many group contribution methods have been developed e.g. by Lydersen [5], Ambrose [6] and Klincewicz and Reid [42]. The first two models use GC obtained by analyzing incremental changes in physical properties within homologous series while the third model uses a least-squares regression analysis using data from [6]. On the other hand, Benson et al. [43] used an approach based on the extrapolation of the additivity rules used for solute species at infinite dilution to properties such as enthalpy of formation, heat capacity and entropy. Franklin [44] and Verma and Doraisamy [45] developed methods to predict enthalpy of formation. The model of Franklin [44] was established for branched paraffins which calculate the

group contribution values from the heat content and free energy functions of molecules containing the groups investigated. The model of Verma and Doraisamy [45] is intended for a wider range of organic compound where the temperature dependence of any group was expressed as a linear function and two equations are required to cover temperature ranges between 300 and 1500^{0}C.

Moreover, the Joback and Reid [4] GC model can be used to predict eleven important and commonly used pure component thermodynamic properties. Nine of them are temperature-independent properties estimated by a simple sum of GC add and two are temperature dependent properties (ideal gas heat capacity, dynamic viscosity of liquids). The heat capacity polynomial uses four parameters (calculated by GC) while the viscosity equation uses only two. The Joback and Reid model is well known due to the fact it can be used to predict a wide range of properties for single compounds and also it is easy to use in terms of the mathematics and description of compounds using groups. As it has been stated by the authors in [4], high accuracy is not claimed for this method but it is often as accurate as methods at the time it was developed. There are some limitations of this model which include aromatic compounds which are not differentiated from normal ring containing compounds, the database for obtaining the parameters is small and covers only a limited number of compounds and there were also problems formula used for the prediction of the normal boiling point.

In order to overcome some of the limitations of previous methods, for example the inability to distinguish between isomers, higher orders GC methods have been developed by Constantinou and Gani [34] and Marrero and Gani [8]. In [34], two levels of contributions are considered where in the first level, compounds are described by first order groups which are considered as building blocks. Using the contributions from those groups, a first order estimation is obtained. In the second level, the compounds are described by second order groups which provide more structural information to the compound and thus the ability to distinguish between isomers. Using the contributions of these second order groups, better predictions are obtained. The properties that can predicted from the Constantinou and Gani model are normal boiling point, normal melting point, critical pressure, critical temperature, critical volume, standard enthalpy of vaporization at 298 K, standard Gibbs energy and standard enthalpy of formation at 298 K. Even with the second order groups, the range of applicability of GC methods is still limited since the properties of large, complex and multifunctional compounds which appear in biochemical and environmental studies cannot be accurately estimated. In order to increase the application range of GC methods, Marrero and Gani [8] have proposed a model where the estimation is performed in three levels i.e. the compounds can be described up to a third order. Using the first order groups, a wide variety of compounds can be described mainly simple and mono-functional compounds. However, these first order groups are unable to describe fully the proximity effect and also unable to distinguish between isomers and therefore second order groups and the second level of estimation were introduced. The second level of prediction permits the description of compounds which are poly-functional, polar or nonpolar of medium size, C = 3-6, and aromatic or cycloaliphatic compounds with only one ring. The first and second order groups were unable to describe fully the molecular fragments of compounds and therefore third order groups were added to enhance this structural information. The third order groups make it possible to represent complex heterocyclic and large (C = 7-60) poly-functional acyclic compounds. This multilevel property model can be used with higher accuracy and reliability compared to the previous property models. Recently [9], the Marrero and Gani model has been further revised in terms of the parameters and the uncertainties of the

predicted properties have been evaluated for determining the reliability of the predictions. The Marrero and Gani model is presented as

$$f(P) = \sum_i N_i C_i + w \sum_j M_j D_j + z \sum_k O_k E_k \quad (2.2)$$

In the equation above, C_i, D_j and E_k represents the contributions of first, second and third order groups of i, j and k while N_i, M_j and O_k are the number of occurrences of each of those groups in defining the compounds or mixtures. In the first level of estimation, the constants w and z are set to zero. When second level of estimation is considered, the constants w and z are set to unity and zero respectively. Both constants w and z are set to unity in the third level.

2.2.2 Group Contribution Methods for Mixtures

The extension of the GC approach to mixtures is of interest since many functional groups can already be described for many compounds. It is of interest with the same number groups to represent the overwhelming number of mixtures. The estimation of thermodynamic properties of liquid mixtures using the GC approach was first proposed by Langmuir [46]. He found that the addition of each successive CH_2 group to a hydrocarbon chain has about the same effect on volume, boiling point and solubility and therefore assumed that the field of force about any particular group or radical in a large organic molecule is characteristic of that group and as a first approximation is independent of the nature of the rest of the molecule. Therefore if the contributions of each group within a molecule to the free energy of a transfer process are independent, they are also additive.

This suggestion however received little attention until Derr et al. [47] and Derr and Papadopoulos [48] used GC to correlate heat of mixing of non-electrolytes for binary solutions of hydrocarbon based on the short-range character of intermolecular forces. In [47], the authors assumed that the energy of interaction of two molecules is the sum of the terms contributed by the contacts between parts of the two molecules. The contacts between different classes of groups have more weights compared to groups which are from the same class. However, it was assumed that the contributions of each contact depend only on the two groups and not the other parts or the concentration of the two molecules involved or on any other molecules present. Based on this, only the interactions of pairs are accounted. Furthermore Wilson and Deal [49] have developed the solution-of-groups method for activity coefficients.

The development of these ideas have encouraged Derr and Deal [10] to develop a model called the Analytical Solution of Groups (ASOG) model for correlating and predicting activity coefficients of liquid mixtures. Parameters are obtained from pairs of structural groups which are regressed from experimental data. The model is considered as a sum of two terms as described in Equation (2.3):

$$\ln \gamma_i = \ln \gamma_i^{FH} + \ln \gamma_i^{R} \quad (2.3)$$

In Equation (2.3) above, the first part is the combinatorial or size contribution which is calculated from the Flory-Huggins (FH) theory and the second part is related to the energetic interaction between the structural groups (residual part). The size contribution includes a linear functionality that expresses the ratio of solute groups to the total number of groups in the average liquid molecule. This method was further extended by Ronc and Ratcliff [11] for the prediction of the excess free energies of liquid mixtures. Kojima and Tochigi [50] further increased the application range of the ASOG model by adding more parameters which makes it possible to describe a wider range of compounds and mixtures. The use of GC approach in the development of ASOG model has inspired Fredenslund et al. [12] to develop the UNIFAC model which combines the solution-of-groups concept with the UNIQUAC model to predict activity coefficients. The UNIFAC model is presented next.

2.3 The UNIFAC Group Contribution Method

The UNIFAC model was initially proposed by Fredenslund et al. [12] in 1975. They combined the solution-of-functional-groups concept which was implemented in development of the ASOG [10-11] models with UNIQUAC [51] which is a model for activity coefficients based on the quasi chemical theory for liquid mixtures. The basic idea in the implementation of the solution-of-functional-groups concept is to use phase equilibria experimental data, available in literature e.g. in [52], to estimate the phase behavior of systems for which no experimental data are available. The assumption is that the mixtures are not only consisting of molecules or compounds but are combinations of functional groups. These groups describe each molecule in the mixture and the activity coefficients of liquid mixtures are related to the interactions between those smaller structural groups. Therefore, through a regression of activity coefficients or phase equilibria data, the parameters describing the interactions between pairs of functional groups for non-electrolyte systems are obtained. Using these generated parameters, activity coefficients of systems which have not been studied experimentally before but having the same functional groups (as for systems used in parameter fitting) can be predicted.

The structural groups which define the mixtures are called subgroups and some of them are listed in column 2 of Table 2.1 (defined for the Original UNIFAC model). The relative volume and relative surface area of those subgroups are defined as R_k and Q_k respectively where the notation k represents each of the subgroups. In the Table 2.1, examples of the representation of molecules by each subgroup are given. For a certain molecule, there may several ways to represent them using different number and type of subgroups. It is believed that the combination which contains the least number of groups is the correct one. Besides the pure properties of the subgroups, the activity coefficients also depend on the interaction between subgroups. However, similar subgroups are assigned to the same main group as shown in the first two columns of Table 2.1. This is because in terms of group energy interactions, all subgroups which belong to the same main group are considered as identical. Therefore, the parameters which characterize the group interactions are between the main group pairs. For example if we consider a binary mixture of n-Butane (2 CH_3, 2 CH_2) with Ethanol (1 CH_3, 1 CH_2, 1 OH), there are 3 different types of subgroups but for the calculations of activity coefficients, the values of interaction parameters needed are only between two main groups which are CH_2 and OH (interactions between the same main group is zero). Example values of these interactions parameters can be found in Table 2.2.

Table 2.1: Original UNIFAC-VLE [15] Subgroup Parameters.

Main group	Subgroup	*k*	R_k	Q_k	Examples of molecules and their constituent groups	
1 "CH_2"	CH_3	1	0.9011	0.848	n-Butane:	2 CH_3, 2 CH_2
	CH_2	2	0.6744	0.540	Isobutane	3 CH_3, 1 CH
	CH	3	0.4469	0.228	2,2-Dimethyl propane	4 CH_3, 1 C
	C	4	0.2195	0.000		
3 "ACH"	ACH	10	0.5313	0.400	Benzene	6 ACH
4 "$ACCH_2$"	$ACCH_3$	12	1.2663	0.968	Toluene	5 ACH, 1 $ACCH_3$
	$ACCH_2$	13	1.0396	0.660	Ethylbenzene	1 CH_3, 5 ACH, 1 $ACCH_2$
5 "OH"	OH	15	1.0000	1.200	Ethanol	1 CH_3, 1 CH_2, 1 OH
9 "CH_2CO"	CH_3CO	19	1.6724	1.488	Acetone	1 CH_3CO, 1 CH_3
	CH_2CO	20	1.4457	1.180	3-Pentanone	2 CH_3, 1 CH_2CO, 1 CH_2

Table 2.2: Original UNIFAC-VLE [15] Interaction Parameters, a_{mk}, in Kelvin.

	1 CH_2	**3 ACH**	**4 $ACCH_2$**	**5 OH**	**9 CH_2CO**
1 CH_2	0.00	61.13	76.50	986.50	476.40
3 ACH	-11.12	0.00	167.00	636.10	25.77
4 $ACCH_2$	-69.70	-146.80	0.00	803.20	-52.10
5 OH	156.40	89.60	25.82	0.00	84.00
9 CH_2CO	26.76	140.10	365.80	164.50	0.00

Until now, almost 40 years after the first publication of UNIFAC, there have been many versions of the UNIFAC model. The main common feature of all UNIFAC versions is that they are based on the group contribution concept where the liquid phase activity coefficient of a component in a mixture is obtained by summing all contributions of each functional group represented in the mixture. Most UNIFAC models have two contributions, combinatorial and residual one. The combinatorial term accounts for the differences in the molecular size and shape and it is a function of the liquid composition and the group volume and surface areas The residual term accounts for the energetic interactions between molecules and groups and it is a function of the liquid composition, the group volume and surface areas, temperature as well as the interactions between the functional groups which are obtained through the group interaction parameters (GIPs) that are fitted to experimental data.

The general UNIFAC equation is as follows with the superscripts *C* and *R* indicating the combinatorial and residual contributions.

$$\ln \gamma_i = \ln \gamma_i^C + \ln \gamma_i^R \tag{2.4}$$

The entropic effect due to differences in the molecular size and shape are accounted for in the combinatorial term ($\ln \gamma_i^C$) while the residual contribution ($\ln \gamma_i^R$) accounts for the enthalpic effect

due to the molecular energetic interactions. The UNIFAC equations in the rest of this section are based on Original UNIFAC by Fredenslund et al. [12]. The other versions of the UNIFAC model differ in how the combinatorial part is being formulated and how the temperature dependence is defined for the group interaction parameters in the residual part.

The combinatorial term is written as [53]

$$\ln \gamma_i^C = 1 - J_i + \ln J_i - 5q_i\left(1 - \frac{J_i}{L_i} + \ln \frac{J_i}{L_i}\right) \tag{2.5}$$

In this part, only pure component properties enter into this equation. The values of J_i and L_i are calculated using Equations (2.6) and (2.7) as follows

$$J_i = \frac{r_i}{\sum_j r_j x_j} \tag{2.6}$$

$$L_i = \frac{q_i}{\sum_j q_j x_j} \tag{2.7}$$

Subscript *i* represent the species or compounds while *j* indicates summation over all compounds. The parameters r_i and q_i are calculated from the summation of the group volume and surface area R_k and Q_k respectively according to the equations below

$$r_i = \sum_k \upsilon_k^{(i)} R_k \tag{2.8}$$

$$q_i = \sum_k \upsilon_k^{(i)} Q_k \tag{2.9}$$

where $\upsilon_k^{(i)}$ is the number of groups of type *k* in compound *i*. The group parameters R_k and Q_k are obtained from the van der Waals group volume and surface areas V_{wk} and A_{wk} given by Bondi [54]

$$R_k = \frac{V_{wk}}{15.17} \tag{2.10}$$

$$Q_k = \frac{A_{wk}}{2.5 \times 10^9} \tag{2.11}$$

The normalization factors 15.17 and 2.5 x 10^9 are those given by Abrams and Prausnitz [51]. They are based on a standard segment defined as a sphere such that for a linear polymethylene. The residual term is written as follows

$$\ln \gamma_i^R = q_i \left[1 - \sum_k \left(\theta_k \frac{\beta_{ik}}{s_k} - e_{ki} \ln \frac{\beta_{ik}}{s_k} \right) \right] \tag{2.12}$$

The values of e_{ki}, θ_{ki}, ϑ_{ki} and s_k terms are described as follows

$$e_{ki} = \frac{\upsilon_k^{(i)} Q_k}{q_i} \tag{2.13}$$

$$\beta_{ik} = \sum_m e_{mi} \tau_{mk} \tag{2.14}$$

$$\theta_k = \frac{\sum_i x_i q_i e_{ki}}{\sum_j x_j q_j} \tag{2.15}$$

$$s_k = \sum_m \theta_m \tau_{mk} \tag{2.16}$$

$$\tau_{mk} = \exp \frac{-a_{mk}}{T} \tag{2.17}$$

Subscripts *k* identifies subgroups and *m* is a dummy index running over all subgroups. The notation a_{mk} represents the group interactions parameters (which are in the unit of Kelvin) which are obtained from phase equilibria data and for the Original UNIFAC model are reported by Hansen et al. in [15]. It should be pointed out that a_{mk} is not equal to a_{km} and it is assumed to be temperature independent and the only temperature dependancy is that of the Boltzmann factors as shown in Equation (2.17).

2.3.1 Limitations of UNIFAC

UNIFAC [15] has several limitations [2, 55-57]. The first one is that it cannot differentiate between isomers due to the assumptions of the solution-of-functional-groups concept using only first order description of molecules. In addition, it has been found that using the interaction parameters characterized by vapor liquid equilibrium (VLE) data only, the prediction of liquid-liquid equilibrium (LLE) is not satisfactory. This is an important deficiency of the basic assumption according to which the parameters fitted do not know specifically what they are being used for. Also, due to proximity effects (occurring when polar groups are close to each other), the representation of dilute systems and the prediction of complex systems containing water and multifunctional chemicals can be very poor.

Besides that, the application ranges of UNIFAC are limited to low pressures (between 10-15 atm) and a temperature range of 275-425 K which also depends on the range of temperature of the

phase equilibria data used to regress the interaction parameters. An extrapolation of predictions outside these ranges is not advisable at least for the Original UNIFAC model. Furthermore, another serious limitation of Original UNIFAC [15] is the weak temperature dependency of the interaction parameters which leads to poor predictions for properties such as heat of mixing (H^E) and infinite dilution activity coefficients (γ^{∞}). Moreover, UNIFAC is not applicable for non-condensable gases, electrolytes and polymers.

2.3.2 UNIFAC Variants

More than 2000 articles [19] have been published until July 2012 on UNIFAC. Among those articles, a series of revisions and extensions the UNIFAC parameter table has been carried out [58-63] due to the increase of new experimental data which in the versions after new groups and new parameters are introduced. The parameter table reported by Hansen et al., [15] was the last one published in the open literature for Original UNIFAC. Since then, the further revision and extension of the UNIFAC parameter table has been done by the UNIFAC Consortium [64] and most of the parameters only available for the members or sponsors of the consortium.

Besides the work on the further revision and extension of the UNIFAC model, work on the development of other versions of UNIFAC has also been carried out in order to overcome the limitations of Original UNIFAC and to further extend its application range. Some of the most important UNIFAC variants are listed in Table 2.3. In order to overcome the limitation of the UNIFAC-VLE model [15] to predict LLE systems, Magnussen et al. [65] have developed a separate parameter table specifically suitable for LLE systems but only at 25 ^{0}C. The UNIFAC-LLE model is identical to the VLE version, only the interaction parameters are different.

Table 2.3: Some of the Most Important UNIFAC Variants [55].

UNIFAC Variant	Temperature Dependency of Interaction Parameters	Data used in Parameter Estimation	Reference
Original VLE	Independent of temperature $a_{mn} = a_{mn,0}$	VLE	Fredenslund et al. [12] Hansen et al. [15]
Original LLE	Independent of temperature	LLE	Magnussen et al. [65]
Linear UNIFAC	Linearly dependent on temperature: $a_{mn} = a_{mn,0} + a_{mn,1}(T - T_0)$	VLE	Hansen et al. [66]
Modified (Lyngby) UNIFAC	Logarithmic dependency $a_{mn} = a_{mn,0} + a_{mn,1}(T - T_0) + a_{mn,2}\left(T \ln \frac{T_0}{T} + T - T_0\right)$	VLE, H^E	Larsen et al. [67]
Modified (Dortmund) UNIFAC	Quadratic dependency	VLE, H^E, γ^{∞}	Weidlich et al. [13]

	$a_{mn} = a_{mn,0} + a_{mn,1}T + a_{mn,2}T^2$		
KT-UNIFAC	Linearly dependent on temperature (first-order): $a_{mn} = a_{mn,0} + a_{mn,1}(T - T_0)$	VLE, H^E, γ^∞	Kang et al. [35]

In the Original UNIFAC model, the interaction parameters are considered to be independent of temperature. Therefore, quantitative predictions of excess enthalpies, H^E could not be obtained. In order to improve this and other things, the Modified (Lyngby) UNIFAC by Larsen et al. [67] and Modified (Dortmund) UNIFAC by Weidlich et al. [13] were developed. For both models, modifications have been done in the combinatorial and the residual part of UNIFAC. For the combinatorial part, an exponent type term were used in both modified UNIFAC models where Equations (2.5)-(2.7) of Original UNIFAC are replaced by Equations (2.18) and (2.19) below for the Lyngby version

$$\ln \gamma_i^C = \ln \frac{\Phi_i}{x_i} + 1 - \frac{\Phi_i}{x_i} \tag{2.18}$$

$$\Phi_i = \frac{x_i r_i^{2/3}}{\sum_j x_j r_j^{2/3}} \tag{2.19}$$

whereby for the Dortmund version, Equations (2.3)-(2.7) are replaced by Equations (2.20)-(2.21) as described below.

$$\ln \gamma_i^C = \ln \frac{\varphi_i'}{x_i} + 1 - \frac{\varphi_i'}{x_i} - 5q_i \left(\ln \frac{\Phi_i}{\vartheta_i} + 1 - \frac{\Phi_i}{\vartheta_i} \right) \tag{2.20}$$

$$\varphi_i' = \frac{x_i r_i^{3/4}}{\sum_j x_j r_j^{3/4}} \tag{2.21}$$

$$\Phi_i = \frac{x_i r_i}{\sum_j x_j r_j} \tag{2.22}$$

In addition to that, in the residual part, temperature dependent interaction parameters were used where they have a logarithmic and quadratic dependency towards temperature for the Lyngby and Dortmund versions respectively (as shown in Table 2.3). Due to this temperature dependency, the predictions of VLE, H^E and γ^∞ have improved especially for the Dortmund version since it is based on more experimental data. These modified UNIFAC models can also extrapolate reliably the predictions of VLE at higher temperatures compared to Original UNIFAC. It should be pointed out

that for the Modified (Dortmund) UNIFAC model, the R_k and Q_k values are treated as adjustable parameters (not directly obtained from Bondi [54]) and are generated together with the interaction parameters through regression of experimental data. Some typical values of the regressed R_k and Q_k values compared with the ones calculated by Bondi [54] are presented in Table 2.4. Since the first development of Modified (Dortmund) UNIFAC which was published in [13], numerous revisions and extensions of the model have been presented in order to improve its reliability [68-71]. In the scientific papers published in [67] and [68], LLE and SLE data of eutectic systems were included in the parameter estimation. Just like the Original UNIFAC, since 1996, the further revision and extension of this model was done by the UNIFAC Consortium in University of Oldenburg, Germany and most of the parameters are only available for consortium members and sponsors. Currently, this is the UNIFAC version that has the most extensive parameter table.

Table 2.4: Some Typical R_k and Q_k Values for Modified (Dortmund) UNIFAC and Original UNIFAC.

Groups	Mod. (Dort.) UNIFAC		Original UNIFAC	
	R_k	Q_k	R_k	Q_k
CH_2	0.6325	0.7081	0.9011	0.8480
ACH	0.3763	0.4321	0.5313	0.4000
$ACCH_2$	0.9100	0.7962	1.0396	0.6600
ACOH	1.0800	0.9750	0.8952	0.6800

In the case of the Linear UNIFAC model [66], the combinatorial term used is the same as in Original UNIFAC with linearly temperature dependent interaction parameters used in the residual part. Eventhough it seems to be a step backwards considering the development of the modified UNIFAC models, the exponent-type combinatorials of the modified models extrapolate badly to athermal polymer solutions [55]. The Flory-Huggins term in the combinatorial part of the Original UNIFAC extrapolates better to polymers solution. Since the focus was on VLE data alone up to high temperatures, the linear temperature dependent parameters are used in Linear UNIFAC. The parameter table for this model was not published in open literature but it is available as a technical report which is available from the Technical University of Denmark [66].

The Original UNIFAC model cannot distinguish between isomers and handle systems with proximity effects. In order to overcome this limitation, Kang et al. [35] have proposed a model called the KT-UNIFAC model where the estimations are obtained in two levels. Mixtures are modeled by first-order group contributions in the first level which forms the basic structure and second-order group contributions in the second level depending on the complexity of the compounds. The introduction of this second-order contributions to some extent can overcome the proximity effects and differentiate between isomers. Initial investigations of the addition of this second order terms to UNIFAC have been carried out by Abildskov et al. [72-73]. Using this so called second order UNIFAC model, activity coefficients can be calculated from the summation of the first order combinatorial part, first order residual part and second order residual part which are described by Equation (2.23). The second order residual part (R2) is used to account for second-order effects on molecular interactions. A detail derivation of the second-order residual part of the model can be found in [35].

$$\ln \gamma_i = \ln_i^C + \ln \gamma_i^R + \omega_{R2} \ln \gamma_i^{R2} \tag{2.23}$$

When the second order residual part is not considered, the constant $_{wR2}$ is set zero. The combinatorial part is the same as the Original UNIFAC model while the first order residual part is the same as the Linear UNIFAC model. However the group descriptions of this model have been enhanced with a larger set of first-order groups in order to be able to describe mixtures in more detail. The first-order groups were intended to describe a large range or organic compounds while the second-order groups were used to provide more structural information such as the molecular fragments of compounds which cannot be provided by the first-order groups. All the group descriptions for this model follow those descriptions made by Marrero and Gani [8] to allow consistency between pure component and mixtures property models.

Besides the UNIFAC variants listed in Table 2.3, there are other special versions of UNIFAC e.g. the Water-UNIFAC models developed by Hooper et al. [74] and Chen et al. [75]. The former model is used to predict LLE of water and hydrocarbon systems while the later model is used for octanol-water partition coefficient calculations. In Hooper's Water-UNIFAC model, the combinatorial term used is that of Larsen et al. [67] while the interaction parameters are determined independently between interaction parameter, a_{mn} (with m = water and n = organic compound) and the interaction parameter, a_{nm} with the quadratic and linear temperature dependency respectively in temperature range 20-250 ^{0}C. Interaction parameters involving water are considered to be temperature dependent while parameters which involve only organic compounds are temperature independent. The data used for parameter fitting in Hooper's Water-UNIFAC are VLE and LLE data involving water-organic compound systems.

In order to predict the phase equilibria of polymer solutions, efforts have also been done to develop suitable UNIFAC models by implementing the free volume (FV) concept [76-80]. This concept was used by Oishi and Prausnitz [76] in the so-called UNIFAC-FV model equation shown in the equation below:

$$\ln \gamma_i = \ln \gamma_i^C + \ln \gamma_i^R + \ln \gamma_i^{FV} \tag{2.24}$$

The UNIFAC model is based on a two-liquid lattice theory of mixtures which does not account for changes in free volume caused by mixing. But in polymer solutions, free volume (FV) effects are important. In polymer solutions, the molecules are tightly packed and therefore the effect of FV is not negligible. The effect of FV is compensated by the term $\ln \gamma_i^{FV}$ in Equation (2.24) where the equation of state theory proposed by Flory [81] is used. Elbro et al. [77] used a slightly different approach where instead of adding an additional term to UNIFAC, the combinatorial and FV effect are combined. This new term is derived from the generalized van der Waals partition function and quantifies the entropy by means of FV of each compound defined as Vf=V-Vw. The main input parameters for this model are the values of the specific volume. Elbro et al. model is one of the most successful models for polymer solutions. Research on the extension of the application range and reliability of these models have been carried out by Kontogeorgis et al. [78], Kouskoumvekaki et al. [79] and Liu et al. [80].

In addition to the models described above several other versions of UNIFAC have been developed for specific purposes for example to predict phase equilibria involving electrolytes by Kikic et al. [82] and Aznar et al. [83] using the Modified (Dortmund) UNIFAC model. Furthermore there are UNIFAC versions suitable for VLE of associating mixtures that include an extra term on association proposed by Fu et al. [84], namely UNIFAC-AG and UNIFAC-AM. They employ functional-group-based association and molecular-based association respectively. In 2009, Lei et al. [85] extended the Original UNIFAC parameter table by regressing parameters with respect to activity coefficient at infinite dilution data of ionic liquids and recently in 2011, Diedrichs et al. [86] developed the Pharma Modified UNIFAC model to predict solubility of active pharmaceuticals ingredients in alkanes, alcohols and water.

There have been even more versions of the UNIFAC model developed for extending the application range of the Original UNIFAC model. For different purposes, UNIFAC has been one preferred choice of models by researchers to be further extended either by adding new terms or regressing new parameters, which shows the usefulness and versatility of the model. However, it would be more versatile if these many UNIFAC variants could be combined in one version where only one parameter table is used to predict phase equilibria of many systems. This is of course a very challenging task. The Modified (Dortmund) UNIFAC approaches this target.

2.3.3 UNIFAC Applications

The UNIFAC models have been widely applied to many chemical engineering problems especially in process and product design involving phase equilibria calculations of binary and multicomponent system where experimental data are not available. Some of the specific applications of UNIFAC models include calculations of:

- VLE [12], LLE [65] and SLE systems [87-88]
- Activities in polymer solutions [77-80]
- Vapor pressures of pure components [89]
- Solvent effects on chemical reaction rates [90]
- Critical micelle concentrations for surfactants solutions [91]
- Flash points of flammable liquid mixtures and solvent mixtures [92-93]
- Solubilities of gases [94]
- Solvent selection for extraction [95]
- Excess enthalpies [96]
- Viscosities of liquid mixtures [97]
- Octanol-water partition coefficients [98-100]
- Solubilities of active pharmaceutical ingredients in solvents [86]
- Solubility of antibiotics in mixed solvents [101]
- Flavor sorption in packaging polymers [102]
- Reid vapor pressure of gasoline [103]
- Infinite dilution activity coefficients [104]
- Henry's law constant [105-106]
- Design of a distillation column including azeotrope and extractive distillation [107]

These applications are either by applying the Original UNIFAC model or by using one of its variants which have been modified to solve these specific problems.

2.4 Molecular Description Theory

Throughout the years, molecular descriptors have been an important tool in the fields of chemistry, pharmaceutical sciences, environmental protection policy, health research and also for chemical engineering. Molecular descriptors are defined through some mathematical treatment of chemical information encoded inside a molecule which is considered as the real object. According to Todeschini and Consonni [108], molecular descriptors are obtained through a numerical procedure which translates the chemical information contained in molecular formula and structure into numbers and it can also be a result from some standardized experiment.

Generally, molecular descriptors can be categorized into two approaches. The first approach is through experimental measurements of properties such as partition coefficient (*log P*), molar refractivity, dipole moment and polarizability, while the second approach is the theoretical molecular descriptors which are derived from a symbolic representation of the molecule. They are further classified to different molecular representation such as 0-, 1-, 2-, 3- and 4-dimensional descriptors. The basic requirements for optimal descriptors are that:

- they should have structural interpretation,
- good correlation with at least one property,
- should be able to discriminate among isomers,
- can be applied to local structure and be generalized to higher descriptors,
- simple and should not be trivially related to other descriptors,
- constructed efficiently and should also change gradually with gradual change in structures.

Molecular descriptors can have physico-chemical, graph theoretical (topological) or quantum mechanical origin depending on the way they are described [109]. In traditional QSAR studies which involve correlation analysis, other than using physico-chemical descriptors such as Hammett σ, a few physico-chemical properties (log P, molar refraction) are used to describe the correlation. In this way, the relations between the descriptors and properties act as a descriptor to the considered biological property or activity. This provides a mixture of property-property and structure-property relationship. On the other hand, graph theoretical descriptors are defined from mathematical descriptions of a molecular structure where usually hydrogen atoms are excluded. Examples of this type of descriptors are the Wiener number, *W* [110] and the topological index, *Z* proposed by Hosoya [111]. The former is obtained from the total length of all the distances between each pair of atoms in a molecule while the latter is defined from the total sets of non-adjacent bonds in a structure. Other descriptors in this group are the connectivity index, *X*, Hyper-Wiener and Path Eigenvalue. In addition, quantum chemical descriptors are derived from models which are based on quantum chemical calculations. The ones that are normally used include computed atomic charges and highest occupied molecular orbital-lowest occupied molecular orbital (HOMO-LOMO) energies based on molecular orbital calculations.

2.4.1 Background of the Connectivity Index

Among those molecular descriptors described in section 2.4, the connectivity indices (CI) are well known for their ability in correlating physico-chemical properties of compounds such as boiling point, partition coefficient, molecular refraction, heat of atomization, solubilities, toxicities and many more. The idea was first introduced by Kier et al. [112] based on the use of mathematical characterization of compounds. It is further discussed by Randíc [113] with the construction of connectivity index (^1X), a bond additive mathematical invariant of molecules, which was originally designed to parallel relative magnitudes of the boiling point in smaller alkanes. This approach was further extended by Kier et al. in [114] with the introduction of higher order connectivity indices (^mX) followed by another important development by Kier and Hall [115] in order to consider hetero-atomic content of molecules (other than only C, O and H atoms) with the introduction of the valence connectivity indices ($^mX^v$). This type of molecular descriptors is highly interdisciplinary which has been used in many areas including chemical engineering due to their simplicity when used in routine engineering calculations.

2.4.2 Significance of Valence Connectivity Index

As discussed in section 2.4.1, there are several types of connectivity indices and among those, the valence connectivity indices are of great interest since they are available at different order. This can makes it possible to describe larger molecular fragments when necessary and also they can handle compounds with different atomic content.

The important starting points in describing the valence connectivity index are the atoms which define the molecules. In the valence state, the atoms can be represented by two cardinal numbers which is δ, the number of bonded neighbors and δ^v, the number of valence electron, both excluding hydrogen atoms. With the description of the δ^v value of atoms appearing in a certain molecular structure, the zeroth-, first- and also second-order valence connectivity indices can be calculated using Equations (2.25)-(2.27)

$$^0\chi^v = \sum_i (\delta_i^v)^{-1/2} \tag{2.25}$$

$$^1\chi^v = \sum_k (\delta_i^v \delta_j^v)_k^{-1/2} \tag{2.26}$$

$$^2\chi^v = \sum_m (\delta_i^v \delta_j^v \delta_l^v)_m^{-1/2} \tag{2.27}$$

In the equations above, the indices *i*, *j* and *l* represent different atoms, *k* represents the number of bonds and *m* is the number of two-edge path appearing in the molecule under study.

The relationship between the values δ and δ^v can be shown from the Equation (2.28), where δ_i is the count of sigma bond (excluding hydrogen) electrons contributed by atom *i* while δ^v is the count

of all valence electron including those sigma electrons, the number of π orbital electrons (p_i) and the number of lone pair electrons (n_i) on atom i excluding bonding hydrogen (h_i).

$$\delta_i^v = \delta_i + p_i + n_i = \sigma_i + p_i + n_i - h_i \tag{2.28}$$

$$\delta^v = Z^v - h_i \tag{2.29}$$

These units which occupy the atom under study and its connection to other bonded atoms provide information on volume and electronic character. The valence connectivity index (CI) has been also correlated to experimental or calculated volume using different methodologies and a special feature is that the volume will increase as the value of the valence connectivity index increases. Besides volume, important information about electronic character that can be extracted from CI is electronegativity and from electronegativity equalization, properties which include ionicity, bond dipole, partial atomic charge and bond strength can be obtained [116-118]. In Equation (2.29), Z^v represents the number of valence electrons and that δ^v is number of valence electrons not involve in bond to hydrogen since hydrogen atoms are suppressed. The hydrogen atoms in reality have been implicitly taken into account when determining those δ^v and δ values which have been modified reflecting also the number of hydrogen atoms on the atom under study.

As discussed, by using the δ^v values, valence connectivity indices can be calculated by Equations (2.25)-(2.27). But one might ask what the significance of those expressions is. We now know that information on volume and electronegativity can be obtained but the question is how it is represented by those equations. Kier and Hall [119] have provided several ways to describe the volume and electronegativity information using the bond index in Equation (2.26) rewritten as in Equation (2.30).

$${}^1\chi^v = \sum (\delta_A^v \delta_B^v)^{-1/2} \tag{2.30}$$

The statements describing volume and electronic contributions from each atom are described:

- "The properties of electronegativity and volume encoded in δ^v_A for an atom (A) are considered to be contributed equally from all valence electrons or their orbitals. Thus, one electron from atom A forming one σ bond will possess the fraction $1/\delta^v_A$ of the properties encoded in δ^v_A." The same goes for atom B.

- "If the property under consideration is the electronegativity, then orbitals from A and B contributing the fractions $1/\delta^v_A$ and $1/\delta^v_B$, respectively, will become adjusted toward an equal intermediate electronegativity, which may be ascribed to the bond." This is what has been described by Sanderson [116-117] as the electronegativity equalization who initially proposed that the geometric mean of the atom electronegativities leads to the best description of the bond electronegativity with the algorithm, $[(1/\delta^v_A)(1/\delta^v_B)]^{1/2}$

- "The volume contributed by atom A to bond A-B is the fraction $1/\delta^v_A$ of the total. The fraction $1/\delta^v_B$ is the volume contributed from atom B. The bond has a 'volume' due to orbital

overlap that can be approximated by $2[(1/\delta^{v}_{A})\ (1/\delta^{v}_{B})]^{1/2}$. With a constant of 2, $[(1/\delta^{v}_{A})\ (1/\delta^{v}_{B})]^{1/2}$ reflects the relative volume of the bond A-B."

2.5 Group Contribution^{Plus} Models

In this section, initially the background of the Group Contribution^{Plus} or GC^{Plus} method will be introduced where previous and current work using this approach will be reviewed. Furthermore the extension of this approach from pure components to mixtures properties will be highlighted with the development of the UNIFAC-CI models, a GC^{Plus} approach using the UNIFAC GC-based activity coefficient model as the host model and with the aid of the valence connectivity indices (CI).

2.5.1 Background of the GC^{Plus} Models

The Group Contribution^{Plus} or GC^{Plus} models are hybrid models which combine a host GC-based property model with molecular descriptors. The latter can be used as complimentary information to overcome the limitation of the property model due to missing contributions or parameters of certain groups or pair of structural groups as a result of lack of experimental data. In GC-based models, experimental data are needed in order to fit the parameters of groups needed for predicting the targeted properties.

The GC^{Plus} idea was proposed by Gani et al. [21], which enhanced the application range of the Marrero and Gani [8] model by implementing valence connectivity indices (CI) with an established relationship to predict missing group contributions. Using this approach missing group contributions can be predicted without conducting experiments With promising results obtained from the work, the implementation of this approach was further extended for GC-based property models for polymers which were developed by Satyanarayana et al. [22-24] using the same molecular descriptors. Furthermore, the hybrid model was implemented for models used to predict Hansen solubility parameters important for solvent selection process [25] and also for the estimation of surface tension and viscosity [26].

Being quite successful in increasing the application range of selected pure component property estimation models, the GC^{Plus} approach was extended for mixture properties. UNIFAC is a successful GC-based model to predict phase equilibria of liquid mixtures. There are still many parameters that are missing in the UNIFAC-VLE parameter table published by Hansen [15] and therefore with the aid of CI, the missing interaction parameters could be predicted in principle. Therefore, González et al. [16-18] have taken the initiative to develop GC^{Plus} models for UNIFAC using valence connectivity indices. Three host UNIFAC models have been selected for the development which are Original UNIFAC [12] and Modified (Dortmund) UNIFAC by Weidlich et al. [13] for VLE calculations and also Original UNIFAC LLE [65] for LLE calculations. A summary of the work done using the GC^{Plus} approach is listed in Table 2.5

Table 2.5: Literature of Works using the GC^{Plus} Approach using Valence Connectivity Index.

Property Model	Molecular Descriptor	Targeted Properties/Application	Reference
Marrero & Gani Pure Component Prediction Model	Valence Connectivity Index	Normal boiling point, Normal melting point, Critical temperature, Critical pressure, Critical volume, Standard heat of formation, Standard Gibbs energy, Standard heat of fusion, Standard heat of vaporization at 298 K, Octanol-water partition coefficient	Gani et al. [21]
Marrero & Gani Pure Component Predictions Model for Polymer		Glassy amorphous volume, Rubbery amorphous volume at 298 K, Amorphous Volume at 298 K, Crystalline Volume at 298 K, Glass transition temperature, Solubility parameter at 298 K and refractive Index at 298 K, Computer aided polymer design	Satyarayana et al. [22-23]
Hansen Solubility Parameter GC Model based on Marrero & Gani Model		Hansen solubility parameters, useful for solubility calculations for solvent selection	Modaressi et al. [25]
Surface tension and Viscosity GC Model based on Marrero & Gani Model		Surface tension at 298 K, Viscosity at 300 K	Conte et al. [26]
Original UNIFAC-VLE Model		Activity coefficient of liquid mixture, for prediction of VLE & SLE systems	González et al. [16-17]
Modified (Dortmund) UNIFAC Model		Activity coefficient of liquid mixtures, for prediction of VLE systems	González [18]
Original UNIFAC-LLE Model		Activity coefficient of liquid mixtures, for prediction of LLE systems	González [18]

2.5.2 UNIFAC-CI Model Formulation

The GC^{Plus} approach for mixtures using UNIFAC as the host GC-based model and valence connectivity indices (CI) as the molecular descriptor is called the UNIFAC-CI model. The main purpose of this hybrid model is to predict the missing group interaction parameters (GIPs) of the UNIFAC model without the need to conduct new experimental data which are normally used to fit those parameters. The UNIFAC model cannot be used when the needed GIPs are not available. Therefore, the UNIFAC groups and their interactions can be considered as the most important variables for the UNIFAC model. These UNIFAC groups are functional and molecular groups that consist of different atoms which are bonded together and the way they are connected to each other gives a special chemical character to each groups. Valence connectivity indices (CI) are molecular descriptors that can transform this chemical character into numerical index. This numerical index can store information about size, branching, unsaturization and hetero-content of molecules for each of the UNIFAC group. By using the information on the atoms and their connections, just as

groups that are used to represent molecules or mixtures that are not included in the regression step, atom connectivities are used to represent groups that are also not included in the parameter regression step provided that the set of data used for regression contain molecules or mixtures that are consist of those interested atoms.

In order to be able to predict the missing GIPs, a relationship has been derived between the GIPs and three main items which are (i) the CIs which are used to describe all the UNIFAC groups and are considered at three different orders, (ii) the number of atoms occurring in each group excluding those hydrogen atoms and, (iii) atom interaction parameters (AIPs) which are obtained by matching them with available data in the literature. The CIs used to characterize the UNIFAC groups are in three orders, zeroth (atom), first (bond) and second (path) orders depending on how complex is the structure of the group. Higher orders can be used, but the size of the groups can already adequately be defined until second order CI also in order to make sure the calculations are at a manageable level. The CIs and the atom constituents of the UNIFAC group can be easily calculated. Once these values have been defined, the next step is to conduct a parameter regression in order to regress the AIPs. Before that, suitable amount of available data (VLE, LLE, SLE or any phase equilibria data) need to be compiled. Using those data, the AIPs are generated through parameter regression. Once the AIPs are available and using the derived relationship, it is possible to (i) generate values for missing GIPs on the UNIFAC parameter table, (ii) re-estimate one or more GIPs and (iii) create a new group and estimate its GIPs. The general methodology in order to generate the missing GIPs through the GC^{Plus} is summarized in Figure 2.3.

In the second step in Figure 2.3, all the UNIFAC groups of the host model need to be defined in terms of the number of different types of atoms that they have and also the values of the valence connectivity indices (CI) of zeroth, first and second order whenever applicable. It is fairly simple to determine the atom stoichiometry directly by observing the structure or the formula of the functional group. For the calculation of the CIs, there are several steps that need to be followed. The first step is to represent the structure of the functional groups in terms of a hydrogen-suppressed graph. For example, the difference between the molecular structure and their corresponding hydrogen-suppressed graph for group HCOO is shown in Figure 2.4. By the way, for group HCOO, there are 1 carbon atom and 2 oxygen atoms. Since it is a “hydrogen-suppressed graph”, the hydrogen atom branch is excluded (on the right side). The non-hydrogen atoms become vertices *1*, *2* and *3* while the bonds connecting them are labeled as *a* and *b*. The omission of the hydrogen and the double bond is compensated by the manner in which the atomic (valence) index δ^v for each vertex is defined.

The atomic index, δ^v for each atom/vertex is defined as follow in Equation (2.31).

$$\delta^v = \frac{(Z^v - N_H)}{(Z - Z^v - 1)} \tag{2.31}$$

In the equation above, Z^v is the number of valence electron in the atom, N_H is the number of hydrogen atom attached to the atom and Z is the atomic number. According to Figure 2.3, the values of each of the item defined are summarized in Table 2.5. The atomic index calculated does

not only contain information about the nature of the atom associated to the vertex but also about the way it is bonded to its surrounding atoms.

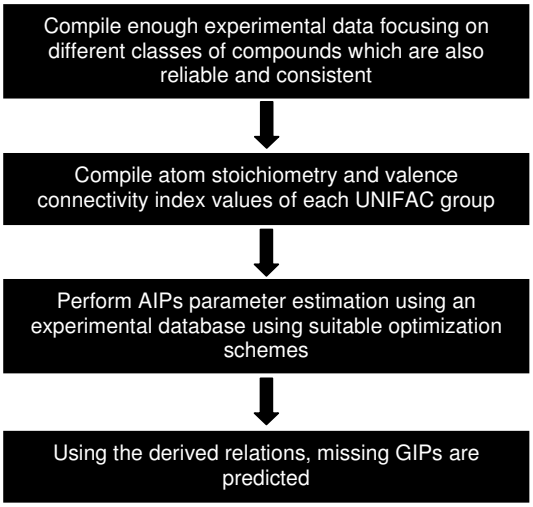


Figure 2.3: Basic Methodology to Generate Missing GIPs.

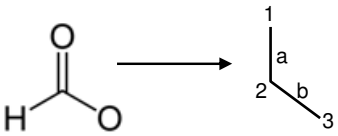


Figure 2.4: Molecular Structure and Hydrogen-Suppressed Graph of HCOO Group.

The zeroth-order (atomic) connectivity index ${}^{v}X^{0}$ is defined as a summation over the vertices of the hydrogen suppressed graph and is calculated as follows.

$$ {}^{v}\chi^{0} = \sum_{i} \left(\frac{1}{\sqrt{\delta_i^{v}}} \right) \qquad i = 1, L \tag{2.32} $$

In the equation above, L is the number of vertices/atoms in the graph and δ^{v} is the valence atomic indices which have been calculated for each atom for HCOO in Table 2.6. In addition, the first-order valence bond indices β^{v} can be defined for each bond by using the δ^{v} corresponding to the pair of bonding atoms using Equation (2.33).

Table 2.6: Atomic, Bond and Path Indices Values for HCOO.

Atom	1	2	3
Z^v	6	4	6
N_H	0	1	0
Z	8	6	8
δ^v	6	3	6
β^v (1-2)	18		-
β^v (2-3)	-	18	
ε^v (1-2-3)	108		

$$\beta^v = \delta_i^v \bullet \delta_j^v \tag{2.33}$$

The calculated values of β^v for HCOO are given in Table 2.6. Using these values, the first-order valence connectivity index ${}^v\chi^1$ which is defined as the summation over the edges of the hydrogen suppressed graph are calculated as follows.

$${}^v\chi^1 = \sum_i \left(\frac{1}{\sqrt{\beta_i^v}} \right) \quad \text{i = 1, M} \tag{2.34}$$

In the equation above, *M* is the number of path of length one-edge of the hydrogen-suppressed graph and β^v is calculated using Equation (2.33). Furthermore for path of length two-edges, second order bond indices ε^v can be calculated as follows.

$$\varepsilon^v = \delta_i^v \bullet \delta_j^v \bullet \delta_k^v \tag{2.35}$$

The value of ε^v calculated for the HCOO group is given in Table 2.6 where there is only 1 path of two-edges identified. Using this value, the second order valence connectivity index ${}^v\chi^2$ can be calculated which is the summation over path of length two-edges of the hydrogen-suppressed graph defined as follows.

$${}^v\chi^2 = \sum_i \left(\frac{1}{\sqrt{\varepsilon_i^v}} \right) \quad \text{i = 1, N} \tag{2.36}$$

In the equation above, *N* is the number of path of two-edges in the graph while ε^v is calculated using Equation (2.35). Using the values obtained in Table 2.6 for HCOO and Equations (2.32), (2.34) and (2.36), the values of the zeroth, first and second CIs are 1.3938, 0.4714 and 0.0962 respectively. A compilation of these CI values for each of the UNIFAC group are listed in the Appendix A.

Furthermore, to look into more detail of the equations and the relationship between GIPs and the compiled atom stoichiometry and CIs values and how it is implemented specifically for different host UNIFAC models, two models Original UNIFAC-VLE and Modified (Dortmund) UNIFAC have been selected to highlight the development of the UNIFAC-CI models.

2.5.2.1 Original UNIFAC-CI Model

For the development of the Original UNIFAC-CI model, González et al [16] has chosen the Original UNIFAC parameter table published by Hansen et al. [15] as basis. The version published in [15] is the result of a series of revision and extension of the originally published article by Fredenslund et al. [12]. There are 50 main groups in the parameter table matrix and it is chosen since this version is the last one where its GIPs were published in open literature and more than 50 % of the parameters are missing. This Original UNIFAC model is intended for VLE calculations since the GIPs are obtained through regression with only VLE data. However, it has been demonstrated that these VLE-generated GIPs can also be used to predict SLE systems [88-89] often reasonably well.

The Original UNIFAC-CI model has been developed to deal with systems containing only atoms C, O, N, Cl and S since a suitable amount of VLE systems are available which consist of these five atoms. Therefore, this UNIFAC-CI model can only be used in the case where missing GIPs are between the pair of groups formed by either one to five of these atoms. In order to predict the missing GIPs, a generic relation has been derived between GIPs, a_{mn} (between main groups m and n where m is less than n) and (i) the number of C-atoms, (ii) the number of O-atoms, (iii) the number of N-atoms, (iv) the number of Cl-atoms, (v) the number of S-atoms in each UNIFAC group, (vi) the CIs described for each group and the (vii) atom interaction parameters (AIPs) defined in Equations (2.37) and (2.38).

$$
\begin{aligned}
a_{mn} = {} & b_{C-C}(A_{mn}^{C-C})_0 + b_{C-O}(A_{mn}^{C-O})_0 + b_{C-N}(A_{mn}^{C-N})_0 + b_{C-Cl}(A_{mn}^{C-Cl})_0 + b_{C-S}(A_{mn}^{C-S})_0 + b_{O-C}(A_{mn}^{O-C})_0 + b_{O-O}(A_{mn}^{O-O})_0 + b_{O-N}(A_{mn}^{O-N})_0 \\
& + b_{O-Cl}(A_{mn}^{O-Cl})_0 + b_{O-S}(A_{mn}^{O-S})_0 + b_{N-C}(A_{mn}^{N-C})_0 + b_{N-O}(A_{mn}^{N-O})_0 + b_{N-N}(A_{mn}^{N-N})_0 + b_{N-Cl}(A_{mn}^{N-Cl})_0 + b_{N-S}(A_{mn}^{N-S})_0 + b_{Cl-C}(A_{mn}^{Cl-C})_0 \\
& \underbrace{+ b_{Cl-O}(A_{mn}^{Cl-O})_0 + b_{Cl-N}(A_{mn}^{Cl-N})_0 + b_{Cl-Cl}(A_{mn}^{Cl-Cl})_0 + b_{Cl-S}(A_{mn}^{Cl-S})_0 + b_{S-C}(A_{mn}^{S-C})_0 + b_{S-O}(A_{mn}^{S-O})_0 + b_{S-N}(A_{mn}^{S-N})_0 + b_{S-Cl}(A_{mn}^{S-Cl})_0 + b_{S-S}(A_{mn}^{S-S})_0}_{0th-order\,interactions} \\
& + c_{C-C}(A_{mn}^{C-C})_1 + c_{C-O}(A_{mn}^{C-O})_1 + c_{C-N}(A_{mn}^{C-N})_1 + c_{C-Cl}(A_{mn}^{C-Cl})_1 + c_{C-S}(A_{mn}^{C-S})_1 + c_{O-C}(A_{mn}^{O-C})_1 + c_{O-O}(A_{mn}^{O-O})_1 + c_{O-N}(A_{mn}^{O-N})_1 \\
& + c_{O-Cl}(A_{mn}^{O-Cl})_1 + c_{O-S}(A_{mn}^{O-S})_1 + c_{N-C}(A_{mn}^{N-C})_1 + c_{N-O}(A_{mn}^{N-O})_1 + c_{N-N}(A_{mn}^{N-N})_1 + c_{N-Cl}(A_{mn}^{N-Cl})_1 + c_{N-S}(A_{mn}^{N-S})_1 + c_{Cl-C}(A_{mn}^{Cl-C})_1 \\
& \underbrace{+ c_{Cl-O}(A_{mn}^{Cl-O})_1 + c_{Cl-N}(A_{mn}^{Cl-N})_1 + c_{Cl-Cl}(A_{mn}^{Cl-Cl})_1 + c_{Cl-S}(A_{mn}^{Cl-S})_1 + c_{S-C}(A_{mn}^{S-C})_1 + c_{S-O}(A_{mn}^{S-O})_1 + c_{S-N}(A_{mn}^{S-N})_1 + c_{S-Cl}(A_{mn}^{S-Cl})_1 + c_{S-S}(A_{mn}^{S-S})_1}_{1st-order\,interactions} \\
& + d_{C-C}(A_{mn}^{C-C})_2 + d_{C-O}(A_{mn}^{C-O})_2 + d_{C-N}(A_{mn}^{C-N})_2 + d_{C-Cl}(A_{mn}^{C-Cl})_2 + d_{C-S}(A_{mn}^{C-S})_2 + d_{O-C}(A_{mn}^{O-C})_2 + d_{O-O}(A_{mn}^{O-O})_2 + d_{O-N}(A_{mn}^{O-N})_2 \\
& + d_{O-Cl}(A_{mn}^{O-Cl})_2 + d_{O-S}(A_{mn}^{O-S})_2 + d_{N-C}(A_{mn}^{N-C})_2 + d_{N-O}(A_{mn}^{N-O})_2 + d_{N-N}(A_{mn}^{N-N})_2 + d_{N-Cl}(A_{mn}^{N-Cl})_2 + d_{N-S}(A_{mn}^{N-S})_2 + d_{Cl-C}(A_{mn}^{Cl-C})_2 \\
& \underbrace{+ d_{Cl-O}(A_{mn}^{Cl-O})_2 + d_{Cl-N}(A_{mn}^{Cl-N})_2 + d_{Cl-Cl}(A_{mn}^{Cl-Cl})_2 + d_{Cl-S}(A_{mn}^{Cl-S})_2 + d_{S-C}(A_{mn}^{S-C})_2 + d_{S-O}(A_{mn}^{S-O})_2 + d_{S-N}(A_{mn}^{S-N})_2 + d_{S-Cl}(A_{mn}^{S-Cl})_2 + d_{S-S}(A_{mn}^{S-S})_2}_{2nd-order\,interactions} \\
& + e_{C-C}(A_{mn}^{C-C})_3 + e_{C-O}(A_{mn}^{C-O})_3 + e_{C-N}(A_{mn}^{C-N})_3 + e_{C-Cl}(A_{mn}^{C-Cl})_3 + e_{C-S}(A_{mn}^{C-S})_3 + e_{O-C}(A_{mn}^{O-C})_3 + e_{O-O}(A_{mn}^{O-O})_3 + e_{O-N}(A_{mn}^{O-N})_3 \\
& + e_{O-Cl}(A_{mn}^{O-Cl})_3 + e_{O-S}(A_{mn}^{O-S})_3 + e_{N-C}(A_{mn}^{N-C})_3 + e_{N-O}(A_{mn}^{N-O})_3 + e_{N-N}(A_{mn}^{N-N})_3 + e_{N-Cl}(A_{mn}^{N-Cl})_3 + e_{N-S}(A_{mn}^{N-S})_3 + e_{Cl-C}(A_{mn}^{Cl-C})_3 \\
& \underbrace{+ e_{Cl-O}(A_{mn}^{Cl-O})_3 + e_{Cl-N}(A_{mn}^{Cl-N})_3 + e_{Cl-Cl}(A_{mn}^{Cl-Cl})_3 + e_{Cl-S}(A_{mn}^{Cl-S})_3 + e_{S-C}(A_{mn}^{S-C})_3 + e_{S-O}(A_{mn}^{S-O})_3 + e_{S-N}(A_{mn}^{S-N})_3 + e_{S-Cl}(A_{mn}^{S-Cl})_3 + e_{S-S}(A_{mn}^{S-S})_3}_{3rd-order\,interactions}
\end{aligned}
\tag{2.37}
$$

While for the parameters a_{nm} (where m is less than n)

$$
\begin{aligned}
a_{nm} = {} & \overline{b_{C-C}}(A_{nm}^{C-C})_0 + \overline{b_{C-O}}(A_{nm}^{C-O})_0 + \overline{b_{C-N}}(A_{nm}^{C-N})_0 + \overline{b_{C-Cl}}(A_{nm}^{C-Cl})_0 + \overline{b_{C-S}}(A_{nm}^{C-S})_0 + \overline{b_{O-C}}(A_{nm}^{O-C})_0 + \overline{b_{O-O}}(A_{nm}^{O-O})_0 + \overline{b_{O-N}}(A_{nm}^{O-N})_0 \\
& + \overline{b_{O-Cl}}(A_{nm}^{O-Cl})_0 + \overline{b_{O-S}}(A_{nm}^{O-S})_0 + \overline{b_{N-C}}(A_{nm}^{N-C})_0 + \overline{b_{N-O}}(A_{nm}^{N-O})_0 + \overline{b_{N-N}}(A_{nm}^{N-N})_0 + \overline{b_{N-Cl}}(A_{nm}^{N-Cl})_0 + \overline{b_{N-S}}(A_{nm}^{N-S})_0 + \overline{b_{Cl-C}}(A_{nm}^{Cl-C})_0 \\
& \underbrace{+\overline{b_{Cl-O}}(A_{nm}^{Cl-O})_0 + \overline{b_{Cl-N}}(A_{nm}^{Cl-N})_0 + \overline{b_{Cl-Cl}}(A_{nm}^{Cl-Cl})_0 + \overline{b_{Cl-S}}(A_{nm}^{Cl-S})_0 + \overline{b_{S-C}}(A_{nm}^{S-C})_0 + \overline{b_{S-O}}(A_{nm}^{S-O})_0 + \overline{b_{S-N}}(A_{nm}^{S-N})_0 + \overline{b_{S-Cl}}(A_{nm}^{S-Cl})_0 + \overline{b_{S-S}}(A_{nm}^{S-S})_0}_{0th-order\ interactions} \\
& + \overline{c_{C-C}}(A_{nm}^{C-C})_1 + \overline{c_{C-O}}(A_{nm}^{C-O})_1 + \overline{c_{C-N}}(A_{nm}^{C-N})_1 + \overline{c_{C-Cl}}(A_{nm}^{C-Cl})_1 + \overline{c_{C-S}}(A_{nm}^{C-S})_1 + \overline{c_{O-C}}(A_{nm}^{O-C})_1 + \overline{c_{O-O}}(A_{nm}^{O-O})_1 + \overline{c_{O-N}}(A_{nm}^{O-N})_1 \\
& + \overline{c_{O-Cl}}(A_{nm}^{O-Cl})_1 + \overline{c_{O-S}}(A_{nm}^{O-S})_1 + \overline{c_{N-C}}(A_{nm}^{N-C})_1 + \overline{c_{N-O}}(A_{nm}^{N-O})_1 + \overline{c_{N-N}}(A_{nm}^{N-N})_1 + \overline{c_{N-Cl}}(A_{nm}^{N-Cl})_1 + \overline{c_{N-S}}(A_{nm}^{N-S})_1 + \overline{c_{Cl-C}}(A_{nm}^{Cl-C})_1 \\
& \underbrace{+\overline{c_{Cl-O}}(A_{nm}^{Cl-O})_1 + \overline{c_{Cl-N}}(A_{nm}^{Cl-N})_1 + \overline{c_{Cl-Cl}}(A_{nm}^{Cl-Cl})_1 + \overline{c_{Cl-S}}(A_{nm}^{Cl-S})_1 + \overline{c_{S-C}}(A_{nm}^{S-C})_1 + \overline{c_{S-O}}(A_{nm}^{S-O})_1 + \overline{c_{S-N}}(A_{nm}^{S-N})_1 + \overline{c_{S-Cl}}(A_{nm}^{S-Cl})_1 + \overline{c_{S-S}}(A_{nm}^{S-S})_1}_{1st-order\ interactions} \\
& + \overline{d_{C-C}}(A_{nm}^{C-C})_2 + \overline{d_{C-O}}(A_{nm}^{C-O})_2 + \overline{d_{C-N}}(A_{nm}^{C-N})_2 + \overline{d_{C-Cl}}(A_{nm}^{C-Cl})_2 + \overline{d_{C-S}}(A_{nm}^{C-S})_2 + \overline{d_{O-C}}(A_{nm}^{O-C})_2 + \overline{d_{O-O}}(A_{nm}^{O-O})_2 + \overline{d_{O-N}}(A_{nm}^{O-N})_2 \\
& + \overline{d_{O-Cl}}(A_{nm}^{O-Cl})_2 + \overline{d_{O-S}}(A_{nm}^{O-S})_2 + \overline{d_{N-C}}(A_{nm}^{N-C})_2 + \overline{d_{N-O}}(A_{nm}^{N-O})_2 + \overline{d_{N-N}}(A_{nm}^{N-N})_2 + \overline{d_{N-Cl}}(A_{nm}^{N-Cl})_2 + \overline{d_{N-S}}(A_{nm}^{N-S})_2 + \overline{d_{Cl-C}}(A_{nm}^{Cl-C})_2 \\
& \underbrace{+\overline{d_{Cl-O}}(A_{nm}^{Cl-O})_2 + \overline{d_{Cl-N}}(A_{nm}^{Cl-N})_2 + \overline{d_{Cl-Cl}}(A_{nm}^{Cl-Cl})_2 + \overline{d_{Cl-S}}(A_{nm}^{Cl-S})_2 + \overline{d_{S-C}}(A_{nm}^{S-C})_2 + \overline{d_{S-O}}(A_{nm}^{S-O})_2 + \overline{d_{S-N}}(A_{nm}^{S-N})_2 + \overline{d_{S-Cl}}(A_{nm}^{S-Cl})_2 + \overline{d_{S-S}}(A_{nm}^{S-S})_2}_{2nd-order\ interactions} \\
& + \overline{e_{C-C}}(A_{nm}^{C-C})_3 + \overline{e_{C-O}}(A_{nm}^{C-O})_3 + \overline{e_{C-N}}(A_{nm}^{C-N})_3 + \overline{e_{C-Cl}}(A_{nm}^{C-Cl})_3 + \overline{e_{C-S}}(A_{nm}^{C-S})_3 + \overline{e_{O-C}}(A_{nm}^{O-C})_3 + \overline{e_{O-O}}(A_{nm}^{O-O})_3 + \overline{e_{O-N}}(A_{nm}^{O-N})_3 \\
& + \overline{e_{O-Cl}}(A_{nm}^{O-Cl})_3 + \overline{e_{O-S}}(A_{nm}^{O-S})_3 + \overline{e_{N-C}}(A_{nm}^{N-C})_3 + \overline{e_{N-O}}(A_{nm}^{N-O})_3 + \overline{e_{N-N}}(A_{nm}^{N-N})_3 + \overline{e_{N-Cl}}(A_{nm}^{N-Cl})_3 + \overline{e_{N-S}}(A_{nm}^{N-S})_3 + \overline{e_{Cl-C}}(A_{nm}^{Cl-C})_3 \\
& \underbrace{+\overline{e_{Cl-O}}(A_{nm}^{Cl-O})_3 + \overline{e_{Cl-N}}(A_{nm}^{Cl-N})_3 + \overline{e_{Cl-Cl}}(A_{nm}^{Cl-Cl})_3 + \overline{e_{Cl-S}}(A_{nm}^{Cl-S})_3 + \overline{e_{S-C}}(A_{nm}^{S-C})_3 + \overline{e_{S-O}}(A_{nm}^{S-O})_3 + \overline{e_{S-N}}(A_{nm}^{S-N})_3 + \overline{e_{S-Cl}}(A_{nm}^{S-Cl})_3 + \overline{e_{S-S}}(A_{nm}^{S-S})_3}_{3rd-order\ interactions}
\end{aligned}
\tag{2.38}
$$

In Equations (2.37) and (2.38), $b_{XY}, c_{XY}, d_{XY}, e_{XY}, \overline{b_{XY}}, \overline{c_{XY}}, \overline{d_{XY}}, \overline{e_{XY}}$, are the atom interaction parameters (AIPs) between atoms *X* and *Y* which are obtained through regression with available experimental data while $(A_{mn}^{XY})_i$ are the coefficients that are calculated using Equations (2.39)-(2.42). It is used in corresponding to the AIPs in terms of atom interaction between *X* and *Y* at different orders, *i*.

$$(A_{mn}^{XY})_0 = \frac{n_X^{(m)}\,{}^{v}\chi_{(n)}^{0} - n_Y^{(n)}\,{}^{v}\chi_{(m)}^{0}}{{}^{v}\chi_{(n)}^{0}\;{}^{v}\chi_{(m)}^{0}} \tag{2.39}$$

$$(A_{mn}^{XY})_1 = \frac{n_X^{(m)}\,{}^{v}\chi_{(n)}^{1} - n_Y^{(n)}\,{}^{v}\chi_{(m)}^{0}}{{}^{v}\chi_{(n)}^{1}\;{}^{v}\chi_{(m)}^{0}} \tag{2.40}$$

$$(A_{mn}^{XY})_2 = \frac{n_X^{(m)}\,{}^{v}\chi_{(n)}^{1} - n_Y^{(n)}\,{}^{v}\chi_{(m)}^{1}}{{}^{v}\chi_{(n)}^{1}\;{}^{v}\chi_{(m)}^{1}} \tag{2.41}$$

$$(A_{mn}^{XY})_3 = \frac{n_X^{(m)}\,{}^{v}\chi_{(n)}^{2} - n_Y^{(n)}\,{}^{v}\chi_{(m)}^{0}}{{}^{v}\chi_{(n)}^{2}\;{}^{v}\chi_{(m)}^{0}} \tag{2.42}$$

In Equations (2.39)-(2.42), $n_X^{(m)}$ is the number of atom type X in main group m while ${}^{v}\chi_{(m)}^{i}$ is the *i*-th order valence connectivity index for main group m. From Equations (2.37) and (2.38), the atom interaction parameters are considered at four different levels,

(i) level 1 – interaction between 0^{th} and 0^{th} order CI,
(ii) level 2 - interaction between 0^{th} and 1^{st} order CI,
(iii) level 3 - interaction between 1^{st} and 1^{st} order CI and
(iv) level 4 - interaction between 0^{th} and 2^{nd} order CI.

The interaction between 1^{st} and 2^{nd} and 2^{nd} and 2^{nd} order CI were neglected since the size and structure of the UNIFAC groups are relatively small compared to the molecules that they represent. Please note the maximum AIPs that need to be regressed is 200 which are far smaller than the GIPs with 1722 for groups containing atoms C, O, N, Cl and S. However a parallel parameter table for molecular groups (CH_3OH, H_2O, PYR, CCl_4, DOH and DMSO) needs to be provided and the missing GIPs with respect to these molecular groups need to be generated using their corresponding AIPs table. This small amount of adjustable parameters which seems to be an advantage still cannot in general be expected to give high precision in prediction of phase equilibria. However, this UNIFAC-CI model can be considered useful to provide the missing GIPs to be used together with published Original UNIFAC GIPs and thus increasing the application range of the reference model.

In order to predict a particular GIP, not all terms in Equations (2.37) and (2.38) need to be used. It depends on what types of atom the main groups have and what order of connectivity indices can be defined. To demonstrate this, the following example shows how we can predict the GIPs between group CH_2 and ACOH, $a_{CH2\text{-}ACOH}$. Let us say that this GIPs is missing and we need to use the Original UNIFAC-CI method to predict it. The value of CI and the atom constituent of main group CH_2 and ACOH are presented in Table 2.7.

Table 2.7: Atom Stoichiometry and CI Values of Group CH_2 and ACOH.

	n_C	n_O	$^0\chi^v$	$^1\chi^v$	$^2\chi^v$
CH_2	1	0	0.7071	0.0000	0.0000
ACOH	1	1	0.9472	0.2236	0.0000

Using this information, Equation (2.37) and (2.38) can be simplified as below.

$$a_{CH_2-ACOH} = b_{C-C}(A^{C-C}_{CH_2-ACOH})_0 + b_{C-O}(A^{C-O}_{CH_2-ACOH})_0 + c_{C-C}(A^{C-C}_{CH_2-ACOH})_1 + c_{C-O}(A^{C-O}_{CH_2-ACOH})_1 \quad (2.43)$$

$$a_{ACOH-CH_2} = \overline{b_{C-C}}(A^{C-C}_{ACOH-CH_2})_0 + \overline{b_{O-C}}(A^{O-C}_{ACOH-CH_2})_0$$

The 200 terms in the generic equation have been tremendously reduced to only 6 terms for this interaction which show the simplicity of the calculations which are of great interest for users. By further expanding the coefficients in the equation above, the following equation is obtained.

$$a_{CH_2-ACOH} = b_{C-C}\frac{n_C^{CH_2}\,{}^v\chi^0_{ACOH} - n_C^{ACOH}\,{}^v\chi^0_{CH_2}}{{}^v\chi^0_{ACOH}\,{}^v\chi^0_{CH_2}} + b_{C-O}\frac{n_C^{CH_2}\,{}^v\chi^0_{ACOH} - n_O^{ACOH}\,{}^v\chi^0_{CH_2}}{{}^v\chi^0_{ACOH}\,{}^v\chi^0_{CH_2}}$$

$$+c_{C-C}\frac{n_C^{CH_2}\,{}^{v}\chi^{1}_{ACOH}-n_C^{ACOH}\,{}^{v}\chi^{0}_{CH_2}}{{}^{v}\chi^{1}_{ACOH}\,{}^{v}\chi^{0}_{CH_2}}+c_{C-O}\frac{n_C^{CH_2}\,{}^{v}\chi^{1}_{ACOH}-n_O^{ACOH}\,{}^{v}\chi^{0}_{CH_2}}{{}^{v}\chi^{1}_{ACOH}\,{}^{v}\chi^{0}_{CH_2}}$$

$$a_{ACOH-CH_2}=\overline{b_{C-C}}\frac{n_C^{ACOH}\,{}^{v}\chi^{0}_{CH_2}-n_C^{CH_2}\,{}^{v}\chi^{0}_{ACOH}}{{}^{v}\chi^{0}_{CH_2}\,{}^{v}\chi^{0}_{ACOH}}+\overline{b_{O-C}}\frac{n_O^{ACOH}\,{}^{v}\chi^{0}_{CH_2}-n_C^{CH_2}\,{}^{v}\chi^{0}_{ACOH}}{{}^{v}\chi^{0}_{CH_2}\,{}^{v}\chi^{0}_{ACOH}} \tag{2.44}$$

The AIPs published by Gonzalez [18] for the above example are presented in Table 2.7.

Table 2.8: AIPs Needed for the Calculations of GIPs between Groups CH_2 and ACOH.

AIPs	Values	AIPs	Values
b_{C-C}	977.79	c_{C-O}	-17.44
b_{C-O}	-1134.82	$\overline{b_{C-C}}$	-145.10
c_{C-C}	-108.11	$\overline{b_{O-C}}$	63.36

Using the information obtained in Table 2.7-2.8 and Equation (2.44), the GIPs between main groups CH_2 and ACOH can be calculated as follows.

$$\begin{aligned}a_{CH_2-ACOH} &= (977.79)\left[\frac{(1)(0.9472)-(1)(0.7071)}{(0.9472)(0.7071)}\right]+(-1134.82)\left[\frac{(1)(0.9472)-(1)(0.7071)}{(0.9472)(0.7071)}\right]\\ &\quad+(-108.11)\left[\frac{(1)(0.2236)-(1)(0.7071)}{(0.2236)(0.7071)}\right]+(-17.44)\left[\frac{(1)(0.2236)-(1)(0.7071)}{(0.2236)(0.7071)}\right]\\ &= 327.65\\ a_{ACOH-CH_2} &= (-145.10)\left[\frac{(1)(0.7071)-(1)(0.9472)}{(0.7071)(0.9472)}\right]+(63.36)\left[\frac{(1)(0.7071)-(1)(0.9472)}{(0.7071)(0.9472)}\right]\\ &= 29.30\end{aligned}$$

2.5.2.2 Modified (Dortmund) UNIFAC-CI Model

González [18] has chosen the Modified (Dortmund) UNIFAC version published by Gmehling et al. [68] for the development of the Modified (Dortmund) UNIFAC-CI model since it is the most accessible publication providing detail information about the explanation of the model, availability of the GIPs and since the recent revisions do not give the parameter values for all newly determined GIPs and the subgroup representations of the UNIFAC main groups. The parameter table matrix for this version has 45 main groups and they are slightly different from the version published by Hansen et al. [15]. There are also still many missing parameters in the parameter table. As it was discussed in section 2.3.2, compared to the Original UNIFAC model, this modified

model has three GIPs which overall have quadratic dependency towards temperature. This is shown in Equation (2.45).

$$a_{mn,overall} = a_{mn} + b_{mn}T + c_{mn}T^2 \tag{2.45}$$

In order to relate these GIPs with the atom stoichiometry, CIs used to defined each main groups and the AIPs, the relationship in Equation (2.46) and (2.47) has been derived to predict the missing GIPs of the Modified (Dortmund) UNIFAC model.

$$\begin{aligned}
a_{mn} = & \underbrace{b^1_{C-C}(A^{C-C}_{mn})_0 + b^1_{C-O}(A^{C-O}_{mn})_0 + b^1_{C-N}(A^{C-N}_{mn})_0 + b^1_{O-C}(A^{O-C}_{mn})_0 + b^1_{O-O}(A^{O-O}_{mn})_0 + b^1_{O-N}(A^{O-N}_{mn})_0 + b^1_{N-C}(A^{N-C}_{mn})_0 + b^1_{N-O}(A^{N-O}_{mn})_0 + b^1_{N-N}(A^{N-N}_{mn})_0}_{0th-order\,interaction} \\
& + \underbrace{c^1_{C-C}(A^{C-C}_{mn})_1 + c^1_{C-O}(A^{C-O}_{mn})_1 + c^1_{C-N}(A^{C-N}_{mn})_1 + c^1_{O-C}(A^{O-C}_{mn})_1 + c^1_{O-O}(A^{O-O}_{mn})_1 + c^1_{O-N}(A^{O-N}_{mn})_1 + c^1_{N-C}(A^{N-C}_{mn})_1 + c^1_{N-O}(A^{N-O}_{mn})_1 + c^1_{N-N}(A^{N-N}_{mn})_1}_{1st-order\,interaction} \\
& + \underbrace{d^1_{C-C}(A^{C-C}_{mn})_2 + d^1_{C-O}(A^{C-O}_{mn})_2 + d^1_{C-N}(A^{C-N}_{mn})_2 + d^1_{O-C}(A^{O-C}_{mn})_2 + d^1_{O-O}(A^{O-O}_{mn})_2 + d^1_{O-N}(A^{O-N}_{mn})_2 + d^1_{N-C}(A^{N-C}_{mn})_2 + d^1_{N-O}(A^{N-O}_{mn})_2 + d^1_{N-N}(A^{N-N}_{mn})_2}_{2nd-order\,interaction} \\
& + \underbrace{e^1_{C-C}(A^{C-C}_{mn})_3 + e^1_{C-O}(A^{C-O}_{mn})_3 + e^1_{C-N}(A^{C-N}_{mn})_3 + e^1_{O-C}(A^{O-C}_{mn})_3 + e^1_{O-O}(A^{O-O}_{mn})_3 + e^1_{O-N}(A^{O-N}_{mn})_3 + e^1_{N-C}(A^{N-C}_{mn})_3 + e^1_{N-O}(A^{N-O}_{mn})_3 + e^1_{N-N}(A^{N-N}_{mn})_3}_{3rd-order\,interaction}
\end{aligned}$$

$$\begin{aligned}
b_{mn} = & \underbrace{b^2_{C-C}(A^{C-C}_{mn})_0 + b^2_{C-O}(A^{C-O}_{mn})_0 + b^2_{C-N}(A^{C-N}_{mn})_0 + b^2_{O-C}(A^{O-C}_{mn})_0 + b^2_{O-O}(A^{O-O}_{mn})_0 + b^2_{O-N}(A^{O-N}_{mn})_0 + b^2_{N-C}(A^{N-C}_{mn})_0 + b^2_{N-O}(A^{N-O}_{mn})_0 + b^2_{N-N}(A^{N-N}_{mn})_0}_{0th-order\,interaction} \\
& + \underbrace{c^2_{C-C}(A^{C-C}_{mn})_1 + c^2_{C-O}(A^{C-O}_{mn})_1 + c^2_{C-N}(A^{C-N}_{mn})_1 + c^2_{O-C}(A^{O-C}_{mn})_1 + c^2_{O-O}(A^{O-O}_{mn})_1 + c^2_{O-N}(A^{O-N}_{mn})_1 + c^2_{N-C}(A^{N-C}_{mn})_1 + c^2_{N-O}(A^{N-O}_{mn})_1 + c^2_{N-N}(A^{N-N}_{mn})_1}_{1st-order\,interaction} \\
& + \underbrace{d^2_{C-C}(A^{C-C}_{mn})_2 + d^2_{C-O}(A^{C-O}_{mn})_2 + d^2_{C-N}(A^{C-N}_{mn})_2 + d^2_{O-C}(A^{O-C}_{mn})_2 + d^2_{O-O}(A^{O-O}_{mn})_2 + d^2_{O-N}(A^{O-N}_{mn})_2 + d^2_{N-C}(A^{N-C}_{mn})_2 + d^2_{N-O}(A^{N-O}_{mn})_2 + d^2_{N-N}(A^{N-N}_{mn})_2}_{2nd-order\,interaction} \\
& + \underbrace{e^2_{C-C}(A^{C-C}_{mn})_3 + e^2_{C-O}(A^{C-O}_{mn})_3 + e^2_{C-N}(A^{C-N}_{mn})_3 + e^2_{O-C}(A^{O-C}_{mn})_3 + e^2_{O-O}(A^{O-O}_{mn})_3 + e^2_{O-N}(A^{O-N}_{mn})_3 + e^2_{N-C}(A^{N-C}_{mn})_3 + e^2_{N-O}(A^{N-O}_{mn})_3 + e^2_{N-N}(A^{N-N}_{mn})_3}_{3rd-order\,interaction}
\end{aligned}$$

$$\begin{aligned}
c_{mn} = & \underbrace{b^3_{C-C}(A^{C-C}_{mn})_0 + b^3_{C-O}(A^{C-O}_{mn})_0 + b^3_{C-N}(A^{C-N}_{mn})_0 + b^3_{O-C}(A^{O-C}_{mn})_0 + b^3_{O-O}(A^{O-O}_{mn})_0 + b^3_{O-N}(A^{O-N}_{mn})_0 + b^3_{N-C}(A^{N-C}_{mn})_0 + b^3_{N-O}(A^{N-O}_{mn})_0 + b^3_{N-N}(A^{N-N}_{mn})_0}_{0th-order\,interaction} \\
& + \underbrace{c^3_{C-C}(A^{C-C}_{mn})_1 + c^3_{C-O}(A^{C-O}_{mn})_1 + c^3_{C-N}(A^{C-N}_{mn})_1 + c^3_{O-C}(A^{O-C}_{mn})_1 + c^3_{O-O}(A^{O-O}_{mn})_1 + c^3_{O-N}(A^{O-N}_{mn})_1 + c^3_{N-C}(A^{N-C}_{mn})_1 + c^3_{N-O}(A^{N-O}_{mn})_1 + c^3_{N-N}(A^{N-N}_{mn})_1}_{1st-order\,interaction} \\
& + \underbrace{d^3_{C-C}(A^{C-C}_{mn})_2 + d^3_{C-O}(A^{C-O}_{mn})_2 + d^3_{C-N}(A^{C-N}_{mn})_2 + d^3_{O-C}(A^{O-C}_{mn})_2 + d^3_{O-O}(A^{O-O}_{mn})_2 + d^3_{O-N}(A^{O-N}_{mn})_2 + d^3_{N-C}(A^{N-C}_{mn})_2 + d^3_{N-O}(A^{N-O}_{mn})_2 + d^3_{N-N}(A^{N-N}_{mn})_2}_{2nd-order\,interaction} \\
& + \underbrace{e^3_{C-C}(A^{C-C}_{mn})_3 + e^3_{C-O}(A^{C-O}_{mn})_3 + e^3_{C-N}(A^{C-N}_{mn})_3 + e^3_{O-C}(A^{O-C}_{mn})_3 + e^3_{O-O}(A^{O-O}_{mn})_3 + e^3_{O-N}(A^{O-N}_{mn})_3 + e^3_{N-C}(A^{N-C}_{mn})_3 + e^3_{N-O}(A^{N-O}_{mn})_3 + e^3_{N-N}(A^{N-N}_{mn})_3}_{3rd-order\,interaction}
\end{aligned} \tag{2.46}$$

$$\begin{aligned}
a_{nm} = & \underbrace{\overline{b^1_{C-C}}(A^{C-C}_{nm})_0 + \overline{b^1_{C-O}}(A^{C-O}_{nm})_0 + \overline{b^1_{C-N}}(A^{C-N}_{nm})_0 + \overline{b^1_{O-C}}(A^{O-C}_{nm})_0 + \overline{b^1_{O-O}}(A^{O-O}_{nm})_0 + \overline{b^1_{O-N}}(A^{O-N}_{nm})_0 + \overline{b^1_{N-C}}(A^{N-C}_{nm})_0 + \overline{b^1_{N-O}}(A^{N-O}_{nm})_0 + \overline{b^1_{N-N}}(A^{N-N}_{nm})_0}_{0th-order\,interaction} \\
& + \underbrace{\overline{c^1_{C-C}}(A^{C-C}_{nm})_1 + \overline{c^1_{C-O}}(A^{C-O}_{nm})_1 + \overline{c^1_{C-N}}(A^{C-N}_{nm})_1 + \overline{c^1_{O-C}}(A^{O-C}_{nm})_1 + \overline{c^1_{O-O}}(A^{O-O}_{nm})_1 + \overline{c^1_{O-N}}(A^{O-N}_{nm})_1 + \overline{c^1_{N-C}}(A^{N-C}_{nm})_1 + \overline{c^1_{N-O}}(A^{N-O}_{nm})_1 + \overline{c^1_{N-N}}(A^{N-N}_{nm})_1}_{1st-order\,interaction} \\
& + \underbrace{\overline{d^1_{C-C}}(A^{C-C}_{nm})_2 + \overline{d^1_{C-O}}(A^{C-O}_{nm})_2 + \overline{d^1_{C-N}}(A^{C-N}_{nm})_2 + \overline{d^1_{O-C}}(A^{O-C}_{nm})_2 + \overline{d^1_{O-O}}(A^{O-O}_{nm})_2 + \overline{d^1_{O-N}}(A^{O-N}_{nm})_2 + \overline{d^1_{N-C}}(A^{N-C}_{nm})_2 + \overline{d^1_{N-O}}(A^{N-O}_{nm})_2 + \overline{d^1_{N-N}}(A^{N-N}_{nm})_2}_{2nd-order\,interaction} \\
& + \underbrace{\overline{e^1_{C-C}}(A^{C-C}_{nm})_3 + \overline{e^1_{C-O}}(A^{C-O}_{nm})_3 + \overline{e^1_{C-N}}(A^{C-N}_{nm})_3 + \overline{e^1_{O-C}}(A^{O-C}_{nm})_3 + \overline{e^1_{O-O}}(A^{O-O}_{nm})_3 + \overline{e^1_{O-N}}(A^{O-N}_{nm})_3 + \overline{e^1_{N-C}}(A^{N-C}_{nm})_3 + \overline{e^1_{N-O}}(A^{N-O}_{nm})_3 + \overline{e^1_{N-N}}(A^{N-N}_{nm})_3}_{3rd-order\,interaction}
\end{aligned}$$

$$\begin{aligned}
b_{nm} = & \underbrace{\overline{b^2_{C-C}}(A^{C-C}_{nm})_0 + \overline{b^2_{C-O}}(A^{C-O}_{nm})_0 + \overline{b^2_{C-N}}(A^{C-N}_{nm})_0 + \overline{b^2_{O-C}}(A^{O-C}_{nm})_0 + \overline{b^2_{O-O}}(A^{O-O}_{nm})_0 + \overline{b^2_{O-N}}(A^{O-N}_{nm})_0 + \overline{b^2_{N-C}}(A^{N-C}_{nm})_0 + \overline{b^2_{N-O}}(A^{N-O}_{nm})_0 + \overline{b^2_{N-N}}(A^{N-N}_{nm})_0}_{0th-order\,interaction} \\
& + \underbrace{\overline{c^2_{C-C}}(A^{C-C}_{nm})_1 + \overline{c^2_{C-O}}(A^{C-O}_{nm})_1 + \overline{c^2_{C-N}}(A^{C-N}_{nm})_1 + \overline{c^2_{O-C}}(A^{O-C}_{nm})_1 + \overline{c^2_{O-O}}(A^{O-O}_{nm})_1 + \overline{c^2_{O-N}}(A^{O-N}_{nm})_1 + \overline{c^2_{N-C}}(A^{N-C}_{nm})_1 + \overline{c^2_{N-O}}(A^{N-O}_{nm})_1 + \overline{c^2_{N-N}}(A^{N-N}_{nm})_1}_{1st-order\,interaction}
\end{aligned}$$

$$+\underbrace{\overline{d^2_{C-C}}(A^{C-C}_{nm})_2+\overline{d^2_{C-O}}(A^{C-O}_{nm})_2+\overline{d^2_{C-N}}(A^{C-N}_{nm})_2+\overline{d^2_{O-C}}(A^{O-C}_{nm})_2+\overline{d^2_{O-O}}(A^{O-O}_{nm})_2+\overline{d^2_{O-N}}(A^{O-N}_{nm})_2+\overline{d^2_{N-C}}(A^{N-C}_{nm})_2+\overline{d^2_{N-O}}(A^{N-O}_{nm})_2+\overline{d^2_{N-N}}(A^{N-N}_{nm})_2}_{2nd-order\,\text{int}eraction}$$

$$+\underbrace{\overline{e^2_{C-C}}(A^{C-C}_{nm})_3+\overline{e^2_{C-O}}(A^{C-O}_{nm})_3+\overline{e^2_{C-N}}(A^{C-N}_{nm})_3+\overline{e^2_{O-C}}(A^{O-C}_{nm})_3+\overline{e^2_{O-O}}(A^{O-O}_{nm})_3+\overline{e^2_{O-N}}(A^{O-N}_{nm})_3+\overline{e^2_{N-C}}(A^{N-C}_{nm})_3+\overline{e^2_{N-O}}(A^{N-O}_{nm})_3+\overline{e^2_{N-N}}(A^{N-N}_{nm})_3}_{3rd-order\,\text{int}eraction}$$

$$c_{nm}=\underbrace{\overline{b^3_{C-C}}(A^{C-C}_{nm})_0+\overline{b^3_{C-O}}(A^{C-O}_{nm})_0+\overline{b^3_{C-N}}(A^{C-N}_{nm})_0+\overline{b^3_{O-C}}(A^{O-C}_{nm})_0+\overline{b^3_{O-O}}(A^{O-O}_{nm})_0+\overline{b^3_{O-N}}(A^{O-N}_{nm})_0+\overline{b^3_{N-C}}(A^{N-C}_{nm})_0+\overline{b^3_{N-O}}(A^{N-O}_{nm})_0+\overline{b^3_{N-N}}(A^{N-N}_{nm})_0}_{0th-order\,\text{int}eraction}$$

$$+\underbrace{\overline{c^3_{C-C}}(A^{C-C}_{nm})_1+\overline{c^3_{C-O}}(A^{C-O}_{nm})_1+\overline{c^3_{C-N}}(A^{C-N}_{nm})_1+\overline{c^3_{O-C}}(A^{O-C}_{nm})_1+\overline{c^3_{O-O}}(A^{O-O}_{nm})_1+\overline{c^3_{O-N}}(A^{O-N}_{nm})_1+\overline{c^3_{N-C}}(A^{N-C}_{nm})_1+\overline{c^3_{N-O}}(A^{N-O}_{nm})_1+\overline{c^3_{N-N}}(A^{N-N}_{nm})_1}_{1st-order\,\text{int}eraction}$$

$$+\underbrace{\overline{d^3_{C-C}}(A^{C-C}_{nm})_2+\overline{d^3_{C-O}}(A^{C-O}_{nm})_2+\overline{d^3_{C-N}}(A^{C-N}_{nm})_2+\overline{d^3_{O-C}}(A^{O-C}_{nm})_2+\overline{d^3_{O-O}}(A^{O-O}_{nm})_2+\overline{d^3_{O-N}}(A^{O-N}_{nm})_2+\overline{d^3_{N-C}}(A^{N-C}_{nm})_2+\overline{d^3_{N-O}}(A^{N-O}_{nm})_2+\overline{d^3_{N-N}}(A^{N-N}_{nm})_2}_{2nd-order\,\text{int}eraction}$$

$$+\underbrace{\overline{e^3_{C-C}}(A^{C-C}_{nm})_3+\overline{e^3_{C-O}}(A^{C-O}_{nm})_3+\overline{e^3_{C-N}}(A^{C-N}_{nm})_3+\overline{e^3_{O-C}}(A^{O-C}_{nm})_3+\overline{e^3_{O-O}}(A^{O-O}_{nm})_3+\overline{e^3_{O-N}}(A^{O-N}_{nm})_3+\overline{e^3_{N-C}}(A^{N-C}_{nm})_3+\overline{e^3_{N-O}}(A^{N-O}_{nm})_3+\overline{e^3_{N-N}}(A^{N-N}_{nm})_3}_{3rd-order\,\text{int}eraction} \quad (2.47)$$

The Modified (Dortmund) UNIFAC-CI model developed by González [18] can handle systems which are consist of atoms C, O and N which are also represented by Equations (2.46) and (2.47). In the above equations, the coeffcients $(A^{XY}_{mn})_i$ can also be calculated using Equations (2.39)-(2.42) just like for the Original UNIFAC-CI model. Since there are three set of GIPs, therefore there are also three set of AIPs which are labeled by number 1, 2, 3 that correspond to the a_{mn}, b_{mn} and c_{mn} GIPs respectively. However in the work of González [18], the parameter c_{mn} has been set to zero because there were not enough experimental data for taking into account extreme sensitivity to temperature effects. Only VLE data were used to regress the remaining adjustable parameters. The calculations of the missing GIPs are more or less the same as for the Original UNIFAC-CI model, but with one more additional set with respect to b_{mn}. Without the use of experimental data such as the activity coefficient at infinite dilution and excess enthalpy to regress the parameters just like the host modified UNIFAC model, it can found that the prediction of those properties using CI-generated parameters are not reliable. However, as a future work which will be conducted here in the Computer Aided Process-Product Engineering Center (CAPEC), these data will be included to regress all interaction parameters and the performance can be compared with the host model in a fair manner.

2.6 Methods and Tools for Analyzing the Models and Their Accuracy

In this section, all the methods and tools which have been used in this PhD work are decribed in order to analyze the UNIFAC-CI models developed by González [16-18] and to further develop the model. These tools are the thermodynamic equations of phase equilibria including VLE and SLE, the set of experimental data used in parameter regression and their sources, the objective function and optimization scheme used in the parameter regression step, the UNIFAC terms used to analyze the model and also the quality assessment algorithm used to check the consistency of the VLE data used in the parameter estimation.

2.6.1 Thermodynamic Equations for Phase Equilibrium

In this work, two types of phase equilibria calculations are considered the vapor-liquid equilibrium (VLE) and solid-liquid equilibrium (SLE). For the VLE calculations, the gamma-phi (*γ-φ*) approach [52] is considered where fugacity coefficients need to be calculated in the vapor phase while the liquid phase is described by activity coefficients. All equilibrium conditions start with an isofugacity criterion, which means that in order for the vapor and liquid phase to be in equilibrium, the vapor fugacity need to be the same as the liquid fugacity:

$$\hat{f}_i^V = \hat{f}_i^L \tag{2.48}$$

In Equation (2.48), for any component *i*, *V* and *L* represent the vapor and liquid phase respectively while the hat indicates the fugacities in the mixture. In the vapor phase, the fugacity is described by the following expression,

$$\hat{f}_i^V = y_i \hat{\phi}_i P \tag{2.49}$$

while in the liquid phase the expression used is given below,

$$\hat{f}_i^L = x_i \gamma_i f_i \tag{2.50}$$

In these equations, y_i and x_i are the vapor and liquid composition respectively, P is the total pressure, f_i is the pure liquid fugacity, γ_i is the liquid phase activity coefficients and $\hat{\phi}_i$ is the vapor phase fugacity coefficient. Combining Equations (2.48)-(2.50), we get,

$$y_i \hat{\phi}_i P = x_i \gamma_i f_i \tag{2.51}$$

which can be further modified as,

$$y_i \Phi_i P = x_i \gamma_i P_i^{sat} \tag{2.52}$$

with

$$\Phi_i = \frac{\hat{\phi}_i}{\phi_i^{sat}} \exp\left[-\frac{V_i^L (P - P_i^{sat})}{RT} \right] \tag{2.53}$$

Equation (2.52) is the general γ-φ formulation of VLE. The exponential term (Poynting factor) in Equation (2.53) can be neglected at low to moderate pressures and its omission introduces often negligible error. Most VLE systems considered in this PhD project, they are at low to moderate pressure and therefore the vapor phase is assumed to be ideal. Using this assumption, Φ_i is set to unity and Equation (2.5.2) reduces to the Modified Raoult's Law equation as in Equation (2.54):

$$y_i P = x_i \gamma_i P_i^{sat} \tag{2.54}$$

In Equation (2.54), the vapor pressure P_i^{sat} can be calculated using established equations such as the Antoine equation, the Wagner equation and also DIPPR 101 equation [27]. The activity coefficients γ_i, are in this project calculated using UNIFAC [12]. There are also VLE systems used in this project which contains carboxylic acids for which the vapor phase is not ideal. To account for this, Equation (2.52) is used and the virial equation is used to calculate the fugacity coefficients using virial coefficients based on the method of Hayden and O' Connell which can be found in [120].

Just like VLE, the basis for the description of SLE is the isofugacity criterion expressed below for the liquid phase, *L* and solid phase S. Subscript 2 refers to solute and 1 to the solvent.

$$\hat{f}_2^L = \hat{f}_2^S \tag{2.55}$$

Equation (2.55) can be written further, assuming that there is no solubility of solvent 1:

$$\gamma_2 x_2 f_2^L = f_2^S \tag{2.55}$$

In the equation above, γ_2 is the activity coefficient, x_2 is the solute's composition and f_2^L and f_2^S are the pure liquid and solid fugacities respectively. Just like for VLE, the activity coefficients are calculated using UNIFAC. Equation (2.55) is rearranged as:

$$\gamma_2 x_2 = \frac{f_2^S}{f_2^L} \tag{2.56}$$

with
$$\ln\left(\frac{f_2^S}{f_2^L}\right) = \left[\frac{\Delta H_{fus,2}}{RT_{m,2}}\left(1-\frac{T_{m,2}}{T}\right)+\frac{\Delta C_{p,2}}{T_{m,2}}\left(\frac{T_{m,2}}{T}-1-\ln\frac{T_{m,2}}{T}\right)\right] \tag{2.57}$$

In Equation (2.57), the second part on the right hand side normally is often neglected as its contribution to the solubility calculations is small. Therefore combining Equation (2.56) and a simplified version of Equation (2.57), the SLE equation reduces to the expression below, which is used throughout the project.

$$\gamma_2 x_2 = \exp\left[\frac{\Delta H_{fus,2}}{RT_{m,2}}\left(1-\frac{T_{m,2}}{T}\right)\right] \tag{2.58}$$

In the above equation, solubility (x_2) of solute 2 in solvent 1 is calculated by determining the activity coefficient γ_2 at system temperature T, where $T_{m,2}$ and $\Delta H_{fus,2}$ *are* the solute's melting temperature and heat of fusion.

2.6.2 VLE and SLE Experimental Data and Sources of Data

In this project, two types of data were used for parameter regression and also for verification purposes; vapor-liquid equilibrium VLE) and solid liquid equilibrium (SLE) data. A classification of the data used for the correlation of experimental data is shown in Table 2.9.

Table 2.9: Types of Phase Equilibria Data Used.

Phase Equilibria Information	Type of Data
VLE	P, T, x_i, y_i
	P, T, x_i
SLE	T, x_i, γ_i

Most of the VLE data have been extracted from the CAPEC database [121] and from open literature. The time span of these data covers from 1930 to 2011 and the experimental set-ups in these data include vapor recirculation still, ebulliometer, equilibrium cell, modified static apparatus and others. For the regression of atom interaction parameters (AIPs) in this work, only a moderate amount of VLE data was used with a total of about 400 datasets This is since due one of the goal of the project is to allow predictions without the need for a large amounts of experimental data or using new experimental data. These datasets are heterogeneous in terms of different types of compounds and systems such as alkanes, aromatics, ketones, aldehydes, esters, ethers, amines, acids, amides, nitro alkane, chloro-alkane and many more. These systems are classified in terms of hydrocarbon, oxygenated, nitrogenated, chlorinated and sulfurated system depending on the atoms that they are consisting of. Another thing to point out is that the consistencies of all VLE data have been tested using a quality assessment algorithm which will be further discussed in the upcoming section. The statistics of VLE and SLE data used for parameter regression is displayed in Table 2.10 showing the number of systems and data points used and the UNIFAC main groups that they cover.

Table 2.10: Statistics of Data used in Parameter Regression.

Phase Equilibria Information	Types of system	Type of data	No. of systems	Data Points	Main Groups Involved
VLE	Hydrocarbons	T, x_i, y_i, P	49	742	CH_2, C=C, ACH, $ACCH_2$, OH,
	Oxygenated	T, x_i, y_i, P	164	2782	ACOH, CH_2CO, CHO, CCOO, HCOO, CH_2O, COOH
	Nitrogenated	T, x_i, y_i, P	28	785	CH_2, ACH, $ACCH_2$, OH, CH_2CO, CCOO, CNH_2, CNH, $(C)_3N$,
		T, x_i, P	30		

					$ACNH_2$, PYR, CCN, CNO_2
	Chlorinated	*T, x_i, y_i, P*	50	1240	CH_2, C=C, ACH, $ACCH_2$, OH, CH_2CO, CCOO, CH_2O, CCl, CCl_2, CCl_3, CCl_4, ACCl
		T, x_i, P	36		
	Sulfurated	*T, x_i, y_i, P*	14	158	CH_2, C=C, ACH, OH, CH_2CO, CCOO, CH_2O, CCl_3, CCl_4, DMSO, CH_2S
	Total		**371**	**5707**	
SLE	Hydrocarbons	T, x_i, γ_i	53	92	CH_2, C=C, ACH, $ACCH_2$, OH, CH_3OH, CH_2CO, CCOO, CH_2O, COOH
	Oxygenated	T, x_i, γ_i	176	352	
	Nitrogenated	T, x_i, γ_i	19	77	CH_2, ACH, $ACCH_2$, OH, CHO, COOH, $(C)_2NH$, $ACNH_2$, PYR, CCN, $ACNO_2$, DMF
	Chlorinated	T, x_i, γ_i	17	46	CH_2, ACH, $ACCH_2$, ACOH, CCOO, COOH, CCl, CCl_2, CCl_3, CCl_4, ACCl
	Sulfurated	T, x_i, γ_i	8	41	CH_2, ACH, $ACCH_2$, COOH, DMSO, THIOPHENE
	Total		**273**	**608**	

For SLE, most of the data are obtained from the DECHEMA Chemistry Data Series for Solubility and Related Properties of Large Complex Chemicals which is a database in a form of book [122-123] and only the data with experimental values of the solute melting temperatures and heats of fusion are chosen. In that database, for solutes for which melting temperatures and/or heats of fusion are not available, they are predicted using the Marrero and Gani prediction method [8]. Since the datasets are going to be used for parameter estimation work, those datasets with predicted pure components will not be used in the parameter estimation to avoid introducing more errors. In addition, only those systems where the compounds can be described by the UNIFAC groups are chosen. Systems with compounds that cannot be fully described by UNIFAC groups will be used to test the prediction power of the UNIFAC-CI model.

2.6.3 Optimization Scheme and Objective Function for Parameter Regression

In order to fit thermodynamic model parameters with respect to measured data for non-ideal models such as UNIFAC and UNIFAC-CI, a suitable optimization scheme needs to be chosen. In this work the method of sum of weighted squares has been chosen since it has been demonstrated [118] that this method is an efficient technique for estimation of thermodynamic model parameters because of the optimization scenario inherent to the development of the UNIFAC-CI models is very complex.

The basic idea of this method is the minimization of the sum of squared deviations between the measured data and the calculated values. In this optimization scheme, three important criteria are required (i) a numerical rule for successively updating iterations, (ii) a method for deciding when to stop the process and, (iii) starting values to get the iterative process under way, since the closer the initial guess to the least estimates that minimize the objective function, the faster the convergence and the more reliable the iterative algorithm. In this work, a Newton's method is chosen to find the minimum using a modified Levenberg-Marquardt algorithm to force the Newton method in the direction of steepest descent. The algorithm is implemented in Harwell FORTRAN subroutine VA07AD, modified to subroutine XVA07A for AIPs regression, and used as the minimization algorithm. The details of this algorithm can be found in [125].

Besides the optimization scheme, suitable objective functions also need to be carefully selected. In this work, it is intended to match parameters with VLE and SLE experimental data. In order to get the best results, for VLE data, the following objective function (OF) has been chosen,

$$OF = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{P_{i,\exp} - P_{i,calc}}{P_{i,\exp}}\right)^2 \tag{2.59}$$

while for SLE data, the following OF have been selected.

$$OF = \frac{1}{N}\sum_{i=1}^{N}\left(x_{i,\exp} - x_{i,calc}\right)^2 \tag{2.60}$$

In those objective functions, N is the number of experimental data points, i is an index running over all species in the mixtures, $P_{i,exp}$ and $P_{i,calc}$ are the experimental and calculated pressure respectively, and $x_{i,exp}$ and $x_{i,calc}$ are both experimental and calculated liquid compositions. In addition to the above objective functions, in Equation (2.61) a parameter regularization term is added (whenever needed) to ensure that the regressed parameters did not move too far away from the previous step which have been presented with comparable accuracy.

$$OF = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{P_{i,\exp} - P_{i,calc}}{P_{i,\exp}}\right)^2 + \frac{1}{N}\sum_{i=1}^{N}\left(x_{i,\exp} - x_{i,calc}\right)^2 + w_{reg}\sum_{j}(AIP_j - AIP_j^{IG})^2 \tag{2.61}$$

When simultaneous regressions are done for both VLE and SLE, both OF in Equations (2.59) and (2.60) are used together. In Equation (2.61), AIP_j is the current value of interaction parameter *j*, AIP_j^{IG} is the corresponding initial guess and w_{reg} is a weighting factor used to control the influence of regularization in the optimization. In Equations (2.59), the pressure is calculated based on the assumptions whether the vapor phase is ideal or not. If it is ideal, then the total pressure is calculated from the Modified Raoult's Law relationship (Equation (2.54)) but for systems with carboxylic acids, the fugacity coefficients need to be calculated through the method of Hayden and O' Connell [120] using Equation (2.52). For the liquid composition in Equation (2.60), it is obtained through an iterative step using the SLE Equation (2.58) and the UNIFAC equation [12] which is used to calculate the activity coefficient where the value of the liquid composition is satisfying both equations.

2.6.4 UNIFAC Terms Analysis

One of the objectives of this project is to conduct some analysis in order to improve the performance of the initially developed UNIFAC-CI model by González [18]. The work on this analysis has been published in [126] and will be presented in detail in Chapter 3. Since the UNIFAC model is a group contribution based model, the basis for the analysis is to find which group contribution gives the worse results or errors. In order to do that, the terms in the UNIFAC model which represents the group contributions need to be extracted. The UNIFAC model expressions presented in [53] by Smith et al. and also in Equations (2.4)-(2.17) are used as the basis for analysis.

The UNIFAC model can be divided into the combinatorial and the residual terms. Let us say we a have system (used for correlation) which can be very well predicted using the Original UNIFAC model (all needed parameters are available). However for the Original UNIFAC-CI model, the prediction (using CI-generated GIPs, obtained from correlation) is not good. By comparing the prediction of activity coefficients (γ) made by the Original UNIFAC model and the Original UNIFAC-CI model for that same system the contribution to γ by the combinatorial term is the same for both models. So, the next step is to check in detail the residual part. By observing the residual part in Equation (2.12), a new term which represent the contributions of subgroup *k* has been identified for each components *i* which is:

$$F_i = \sum_k \left(\theta_k \frac{\beta_{ik}}{s_k} - e_{ki} \ln \frac{\beta_{ik}}{s_k} \right) \tag{2.62}$$

Using Equation (2.62), the group contributions to the predictions of activity coefficients by both reference and CI-based UNIFAC models can be determined. Any deviations of the group contributions between both UNIFAC-CI and UNIFAC models will determine the source of the problem and further possible solution can be used to improve the performance of the CI-based model.

2.6.5 VLE Data Quality Criterion, Q_{VLE}

The consistencies of the VLE data used for the parameter regression step have been tested using a quality assessment algorithm proposed by Kang et al. [127]. It is based on four widely used consistency tests for VLE namely Herington, Van Ness, Point (Differential) and Infinite Dilution test which are based on the Gibbs-Duhem requirements. In addition to that, the check on the consistencies between the VLE data and the corresponding pure component vapor pressures are also included. For each of the four tests, quality factors $F_{test,i}$ can be evaluated between 0.025 to 0.25 resulting in the sum of factors with values between 0.1 to 1. The sum of the four quality factors is 1 when all tests are passed as shown in Equation (2.63).

$$F_{test1,\max} + F_{test2,\max} + F_{test3,\max} + F_{test4,\max} = 1 \tag{2.63}$$

For the check of the consistency between the 'end point' of the VLE curve, a quality factor F_{pure} is generated which have a value of 1. Simultaneous use of all five tests provides the opportunity to establish an overall VLE data quality factor, Q_{VLE}. The overall quality factors obtained for each dataset indicate the quality of each datasets and they can be used as weighting factors in the regression of VLE data. Using all the quality factors, F, the overall VLE data quality factor Q_{VLE} is calculated as follows,

$$Q_{VLE} = F_{pure}(F_{test1} + F_{test2} + F_{test3} + F_{test4}), Q_{VLE} \leq 1 \tag{2.64}$$

The consistency tests based on the Gibbs-Duhem equation can be performed only for systems with *P*, *T*, *x* and *y* data. For *P*, *T*, *x* data, only pure component consistency test can be performed. Q_{VLE} and the *F* factors were formulated so that if a particular test cannot be performed, the F factor is set to $0.5F_{test,i\ max}$. $F_{test,i\ max}$ is 0.25 for tests 1 to 4 and 1 for the pure component consistency test. Therefore for *P*, *T*, *x* data, the maximum Q_{VLE} value is 0.5 since all four tests cannot be performed. In a hypothetical case where none of the tests can be applied, Q_{VLE} is 0.25. When a dataset passed all the consistency tests, the Q_{VLE} obtained is 1.0, which is the maximum and the best value. The overall quality factors obtained indicate the quality of the datasets. The idea is that when the datasets fail certain tests (while passing others), they are given a lower quality factor (lower than 1). A dataset is considered problematic when the value of Q_{VLE} is less or equal to 0.05 and therefore the dataset should be corrected by removing outliers and completely removed from correlation. The detail descriptions on how to calculate each of the *F* factor are described by Kang et al [127]. This algorithm has been implemented in software and it is available from the authors.

2.6.6 Software Tools

The parameter regression work has been conducted using FORTRAN programming including generation of phase diagrams. In addition to that, the ICAS (Integrated Computer Aided System) software [128] developed and revised annually in CAPEC, DTU has been also extensively used. ICAS was used to generate phase diagrams (VLE and SLE) predictions using the UNIFAC models using its utility tools, for parameter tuning of UNIFAC interaction parameters in testing the applicability of the UNIFAC-CI model to predict certain system, using its TML (Thermodynamic

Model Parameter Estimation) toolbox and also for predicting pure component properties from a given molecular structure or SMILES from the ProPred (Component Property Prediction) toolbox. Moreover, the consistency test which generates the overall quality factor, Q_{VLE} for a tested dataset can be performed using the TDE-Equilibria software developed by Kang et al. [127].

CHAPTER 3

A FIRST ANALYSIS OF THE ORIGINAL UNIFAC-CI MODEL

3.1 Introduction

In the developed UNIFAC-CI method [16-18], an expression was established for relating the GIPs to the number of atoms involved in the UNIFAC groups, the connectivity indices and a set of atom interaction parameters (AIPs). The atom stoichiometry and the values of the CIs can be obtained directly from the group definition while the AIPs were regressed using available experimental data. An important issue to note is that the same set of experimental data used to regress the GIPs of the reference UNIFAC model, is also used to obtain the AIPs, thereby providing the predictive nature of the UNIFAC-CI method. González et al. [16-17] published the AIPs and the corresponding GIPs for the Original UNIFAC model [15] for groups formed by C, H, O, N, Cl and S atoms and for the Modified (Dortmund) UNIFAC [64] model for groups formed by C, H, O, and N atoms. Just like the reference-UNIFAC model, the UNIFAC-CI model has also been found not to work for some systems [129-130] and therefore further analysis of the UNIFAC-CI model in terms of their performance and accuracy is done to increase its reliability and flexibility [126].

The objective of the work presented in this chapter is to revisit the UNIFAC-CI model and the method for the prediction of the GIPs in order to identify the reasons the method does not perform as well as the reference UNIFAC model for some systems while performing surprisingly better than the reference model for other systems. The UNIFAC-CI model should not be expected to perform better than the reference with a smaller set of parameters. Nevertheless, the aim here is to identify through a systematic investigation, where the reference and the UNIFAC-CI models are performing less satisfactorily for some of the systems. Then, having identified some of the causes, perform a new regression of the AIPs, evaluate the performance of the UNIFAC-CI model and the reference UNIFAC model with its filled GIPs. This chapter reports the results from the analysis, the regression of the AIPs and the performance of the reference models with their new GIPs.

3.2 Analysis of the Original UNIFAC-CI Model (Hydrocarbons and Oxygenated Systems)

In this section, a detailed analysis of the performance and trends of predictions of vapour–liquid phase equilibrium with the Original UNIFAC-CI model are presented. The cases where the model using the predicted GIPs perform well and cases where the performance is unreliable are investigated. The causes for the unreliable performance of the UNIFAC-CI model are explained and results from one of the possible solutions that gave very good results are presented.

3.2.1 Identification of the Problematic Systems

We have restricted our analysis in this work to the systems that include the atoms C, H and O since nearly 58% of the VLE data used for the reference UNIFAC GIPs concern these three atoms. Therefore, analyzing the causes of problems first through these systems and implanting the final solution on the others is a good starting point. The AIPs, which involve the groups formed by C, H and O atoms, were regressed using 228 datasets [16]. In Figure 3.1, the ratio of the average absolute relative deviation (AARD) values of reference UNIFAC and the UNIFAC-CI models obtained for the 228 datasets are plotted against the fraction of datasets. It can be observed that for the systems with two main groups (mg), UNIFAC-CI performs better for 53% of the datasets while 54% of the datasets with three main groups were better predicted by the reference UNIFAC model. Meanwhile, for the systems with four main groups, UNIFAC-CI performs better compared to the reference model for 55% of the datasets.

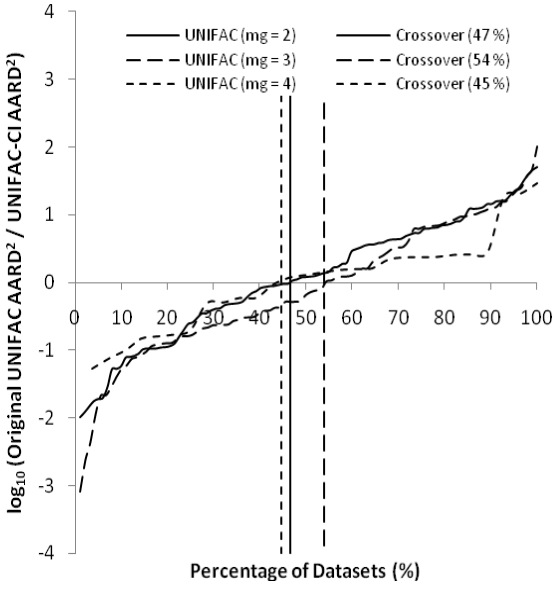


Figure 3.1: Ratio of AARD of Reference UNIFAC and the UNIFAC-CI Models against Fraction of the Datasets for Systems Containing the C, O, H atoms. The Crossover Indicates the Percentage of the Datasets which are Better Correlated by the Reference Model. mg is the Number of Main Groups.

The error is expressed as the average absolute relative deviation (AARD) in pressure for both reference UNIFAC and UNIFAC-CI models,

$$AARD(\%) = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{P_{\exp} - P_{calc}}{P_{\exp}}\right| \times 100 \tag{3.1}$$

where N is the number of data points in the whole set of data, P_{exp} is the experimental pressure and P_{calc} is the calculated pressure. In this work, we used constants of a vapour pressure equation to calculate the pure component vapour pressures (from DIPPR 101). Then, the pure component vapour pressures calculated were checked and they matched with the experimental data for all systems.

The values of AARD for each dataset are calculated in terms of molecular types and are shown in Figure 3.2. This analysis clearly identifies the types of systems that are well correlated and those that are not. This figure also shows that all systems which involve the C–C interactions, such as the alkane/cyclic alkane–cyclic alkane, alkane–alkene, alkane/cyclic alkane–aromatic, and alkene–aromatic are well correlated. Some systems are in fact better correlated than the reference UNIFAC model, such as alkane/cyclic alkane–ethoxy, alcohol–carboxylic acid, ether–ethoxy alcohol and ethanediol–ether. However, the correlation results of several systems for the UNIFAC-CI model are not good, such as, the alkane–aldehyde, hydroxyl aromatic–aromatic and ester–aromatic systems. Note that only one dataset was used in the regression for the last two systems. This is the current state of the performance of the UNIFAC-CI model developed by González [18] for systems involving C, O, and H atoms. The statistics of the number of datasets used for each type of systems are presented in Table 3.1. Please note that, for the systems which involve carboxylic acids, the association constants based on the method of Hayden and O'Connell [120] were employed in the calculations of the vapour phase fugacities. Furthermore, a glycol group instead of two alcohol groups was used to represent 1, 2-ethanediol.

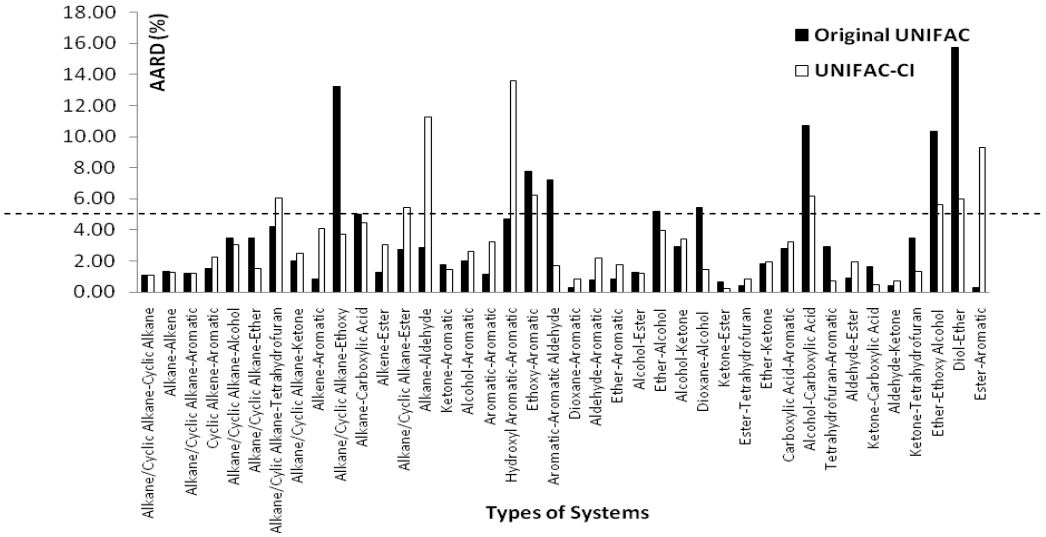


Figure 3.2: Average AARD (%) for each Type of Systems used in the Regression for Systems Containing Atoms C, O, H between Reference UNIFAC and UNIFAC-CI.

Table 3.1: Number of Datasets used for Each Type of System.

Types of systems	Number of datasets
Alkane/Cyclic Alkane-Aromatic	35
Alkane/Cyclic Alkane-Alcohol	23
Ketone-Aromatic	19
Alkane/Cyclic Alkane-Ether	16
Alkane/Cyclic Alkane-Ketone	12
Alcohol-Aromatic	10
Alkane/Cyclic Alkane-Cyclic Alkane	8
Alkane-Alkene, Aldehyde-Aromatic	7
Ether-Aromatic	6
Alkane/Cyclic Alkane-Tetrahydrofuran, Alkane/Cyclic Alkane-Ester, Alkane-Aldehyde, Alcohol-Ester, Ether-Alcohol, Aldehyde-Ester, Ether-Ketone	5
Alkane/Cyclic Alkane-Ethoxy, Alkane-Carboxylic Acid, Alkene-Ester, Tetrahydrofuran-Aromatic	4
Cyclic Alkene-Aromatic, Alkene-Aromatic, Alcohol-Ketone, Carboxylic Acid-Aromatic	3
Ethoxy-Aromatic, Aromatic-Aromatic Aldehyde, Dioxane-Aromatic, Dioxane-Alcohol, Ketone-Ester, Ester-Tetrahydrofuran, Alcohol-Carboxylic Acid	2
Ketone-Carboxylic Acid, Aldehyde-Ketone, Ketone-Tetrahydrofuran, Ether-Ethoxy Alcohol, Diol-Ether, Ester-Aromatic, Aromatic-Aromatic, Hydroxyl Aromatic-Aromatic	1

For further analysis, a limit of 5% AARD was assumed to be an acceptable value and the datasets which have higher AARD values with respect to UNIFAC-CI method were further extracted and are highlighted by the dashed line in Figure 3.2. However, the systems showing AARD values greater than 10% were then selected for further analysis. From Figure 3.2, the alkane–aldehyde systems (1-butanal-n-heptane at 343 K, 1-butanal-n-heptane at 318 K, 1-pentanal-n-heptane at 348 K, 2-methylpropanal-n-heptane at 335K and 2-methylpropanal-nheptane at 318 K) and the phenol–styrene system were found to have AARD greater than 10% (with the largest AARD being 13%).

3.2.2 Identification of the Problems in the Model Expression and/or Parameters

The AIPs employed by the five alkane–aldehyde systems are listed in Table 3.2. For each of the binary systems, activity coefficients were then calculated and analyzed for the reference and the UNIFAC-CI models and compared with experimental data in Figure 3.3. This figure shows that the predicted activity coefficients by the reference UNIFAC model agree well with the experimental values but not so for the UNIFAC-CI model. The problem appears to be at both infinite dilution regions for the binary system. It also looks like the activity coefficient values are under-predicted in all cases.

Table 3.2: AIPs used by several systems.

Systems	AIPs used
Alkane-Aldehyde	b_{C-C}, b_{C-O}, c_{C-C}, c_{C-O}, $\overline{b_{C-C}}$, $\overline{b_{O-C}}$
Alkane-Alcohol	b_{C-C}, b_{C-O}, $\overline{b_{C-C}}$, $\overline{b_{O-C}}$
Alkane-Ether	b_{C-C}, b_{C-O}, c_{C-C}, c_{C-O}, $\overline{b_{C-C}}$, $\overline{b_{O-C}}$

At the next step, it was decided to check the relationship between the different terms of the residual activity coefficient contribution and the corresponding GIP values. A new term, F, was defined for this purpose which represent the contributions of subgroup *k* for each components *i*,

$$F_i = \sum_k \left(\theta_k \frac{\beta_{ik}}{s_k} - e_{ki} \ln \frac{\beta_{ik}}{s_k} \right) \tag{3.2}$$

F is obtained from the residual contribution of the activity coefficient of the reference UNIFAC model,

$$\ln \gamma_i^R = q_i \left[1 - \sum_k \left(\theta_k \frac{\beta_{ik}}{s_k} - e_{ki} \ln \frac{\beta_{ik}}{s_k} \right) \right] \tag{3.3}$$

By analyzing Equations (3.2) and (3.3), we can see that the higher the value of F for a certain group (subscript *k*), the less is the contribution from that specific group. Also, the larger the value of F, the smaller is the residual contribution to the activity coefficient. The F-values of Equation (3.2) are shown in Figures 3.4 and 3.5 for all five alkane–aldehyde systems as a function of composition. Figures 3.4 and 3.5 show that the contributions of the CHO group are higher compared to the CH_3 group for both reference UNIFAC and UNIFAC-CI models. Contributions from group CH_3 are almost the same for the reference UNIFAC and UNIFAC-CI models but different for group CHO (at the infinite dilution range of both compound of the binary system) suggesting that the difference between the two models is caused by the AIPs related to the CHO group. The figure also reveals problems of the UNIFAC-CI model at low concentrations, which is in the near infinite dilution region both compounds of the binary systems.

From this analysis, several alternatives were identified in order to improve the performance of the UNIFAC-CI model:

- Use more weights on the objective function (for parameter regression) for those problematic systems at lower concentrations.
- Introduce a higher order CI for group CHO or introduce a new unique CI parameter only for the alkane–aldehyde systems for the interaction which involves C–O atom interaction.

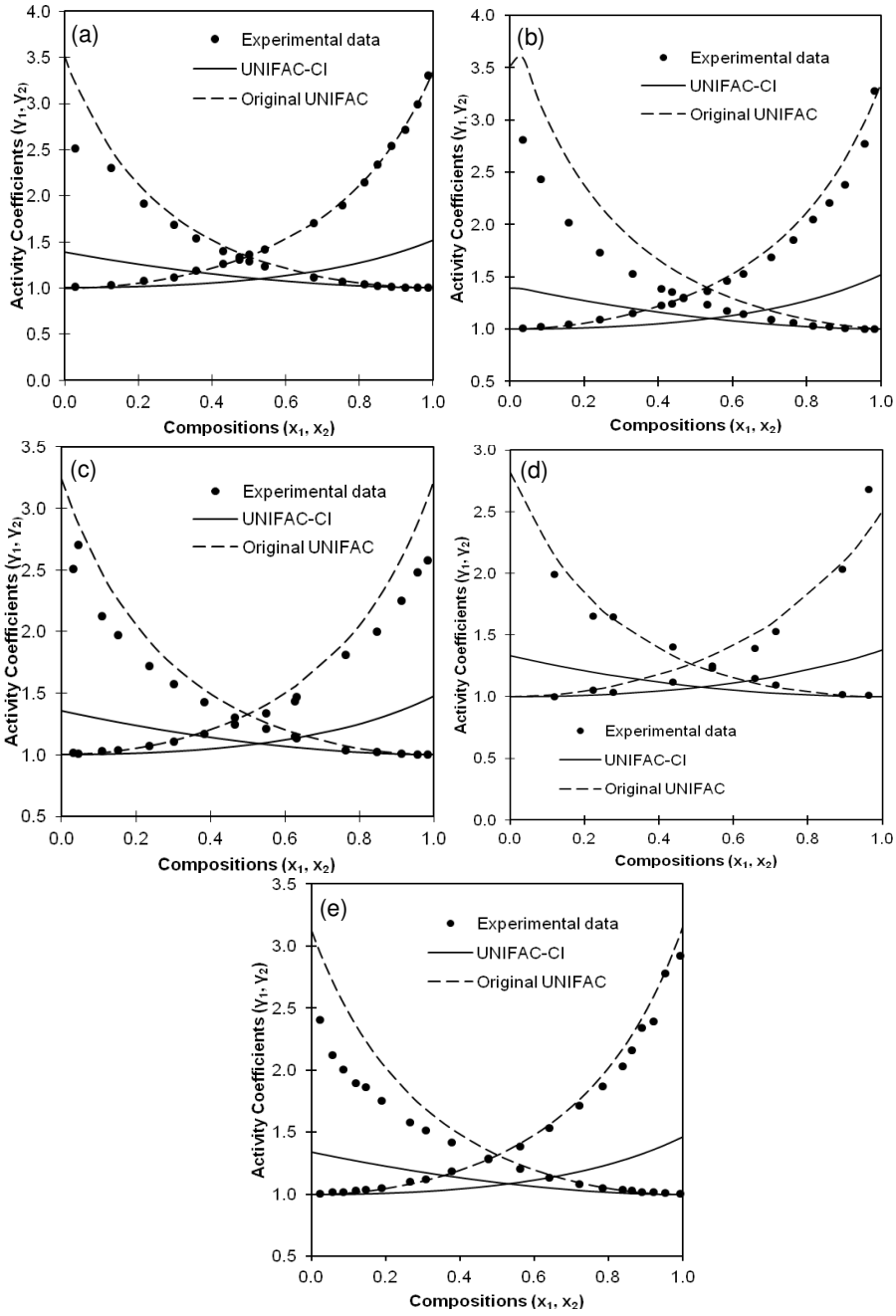


Figure 3.3: Comparison of the Activity Coefficients Calculated using the Original UNIFAC and UNIFAC-CI Methods with Experimental Data for (a) 1-Butanal-n-Heptane at 318 K, (b) 2-Methylpropanal-n-Heptane at 318 K, (c) 2-Methylpropanal-n-Heptane at 335 K, (d) 1-Pentanal-n-Heptane at 348 K and (e) 1-Butanal-n-Heptane at 343 K.

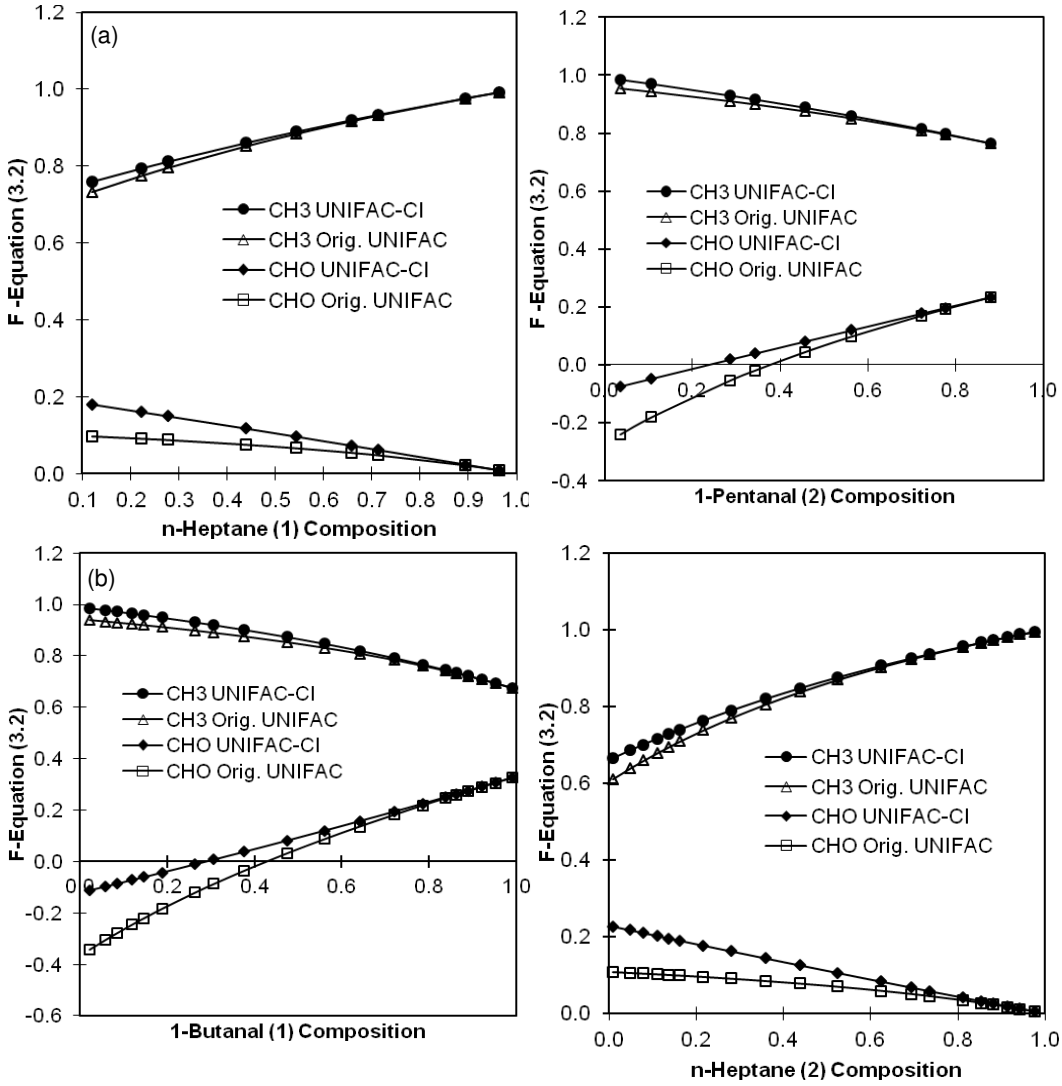


Figure 3.4: Values of F with respect to Composition for (a) 1-Pentanal-n-Heptane at 348K and (b) 1-Butanal-n-Heptane at 343 K.

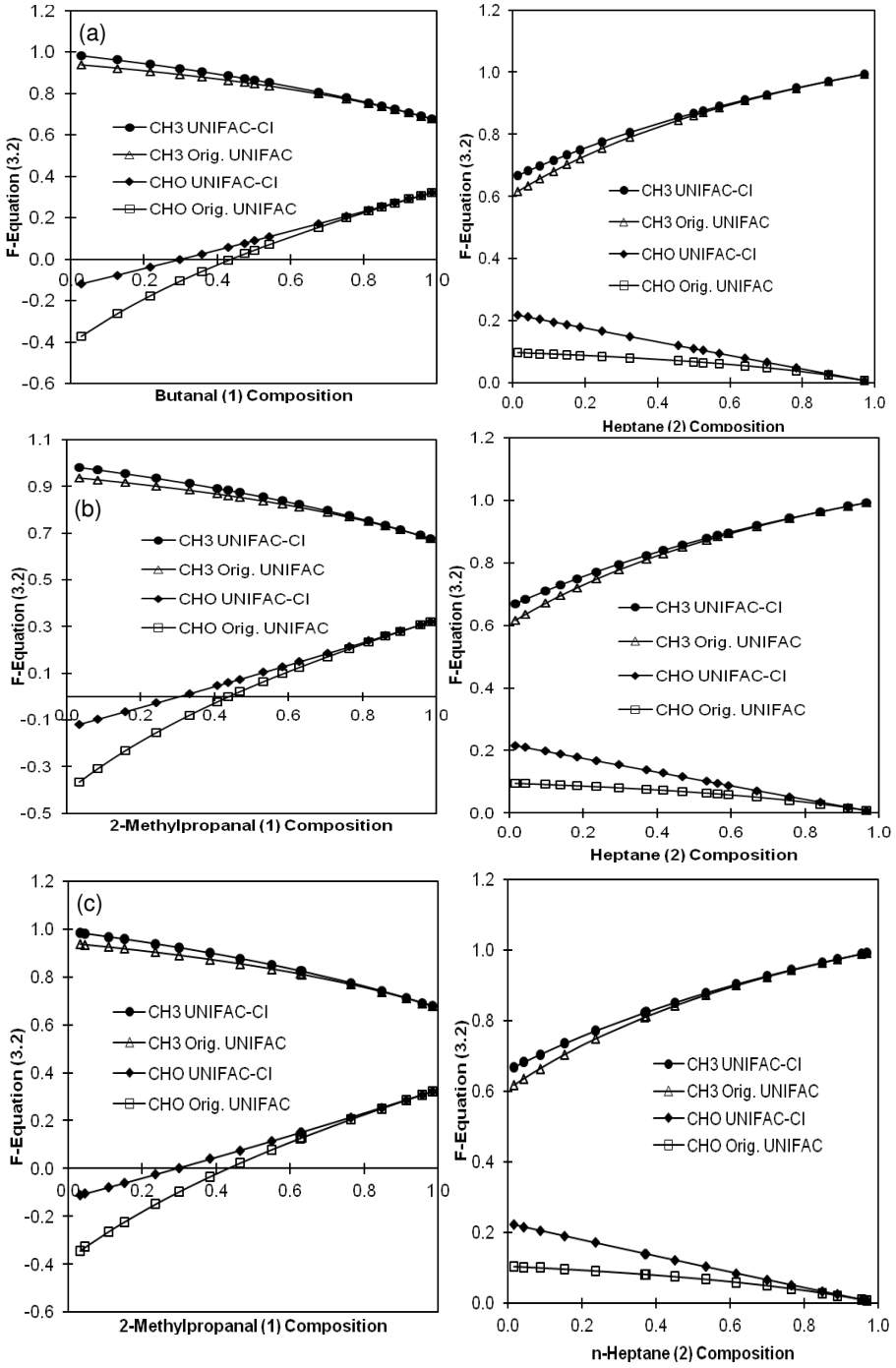


Figure 3.5: Values of F with respect to Composition for (a) 1-Butanal-n-Heptane at 318 K, (b) 2-Methylpropanal-n-Heptane at 318 K and (c) 2-Methylpropanal-n-Heptane at 335 K.

- Add infinite dilution activity coefficient data for regression of the interaction parameters. However, this alternative will not be applied as yet in order to have a fair comparison with the reference UNIFAC model where their GIPs were not regressed from infinite dilution activity coefficient data.

The performance of the UNIFAC-CI model involving hydrocarbons (alkane-cyclic alkane, alkane-alkene, cyclic alkane/alkane-aromatic, cyclic alkene/alkene-aromatic) which only involves C-C interactions is quite good and does not need further improvement. Therefore, for further analysis only systems which involve C, O, and H atoms were investigated. This means that only C-O, O-C and O-O related AIPs are refitted (involving 171 datasets).

3.2.3 Regression with the Weighted Objective Function

Due to the problem that has been identified for the UNIFACCI model, which is especially pronounced at lower concentrations, higher weights were added (in the objective function) for problematic systems (also for systems other than alkane–aldehyde that have AARD> 5%) at concentrations lower than 0.5 in order to improve their correlation. Weights of 5 and 10 have been added only for problematic systems while other systems remained at weight 1. The results of correlation through the ratio of AARD values between the reference UNIFAC and UNIFAC-CI models against the fraction of datasets with the high weights at lower concentrations are presented in Figure 3.6. All 171 systems were regressed in this step.

Figure 3.6 shows that for the systems with two main groups, the reference UNIFAC model performs better for 50% of the datasets at weight 1 and changed to 47% and 49% when adding weight 5 and 10, respectively. On the other hand, for the systems with three main groups UNIFAC-CI performs better for 51% of the datasets at weight 1, 50% and 53% when adding weight 5 and 10, respectively. Furthermore, 43% of the datasets with four main groups were better predicted by the reference model at weight 1 and the deviation increases to 54% and 69% when adding weights 5 and 10.

However, with the higher weights the model performance improved slightly for most of the problematic systems, but some other systems became worse. This indicated the need for a complex maximum likelihood based simultaneous regression of all parameters. For this, a good starting point is necessary and many trials are necessary. This option is time consuming and therefore the second option (introduce new higher order parameters) was considered next.

3.2.4 Assigning Higher Order CI Parameters

A second order CI ($^{v}X^{2}$ = 0.1179) was introduced for the CHO group specifically for the alkane–aldehyde systems which are represented by group interaction CH_3–CHO. Therefore, a new AIP which is e_{C-O} is introduced to the original set of AIPs as presented in Table 3.2. All 171 datasets were regressed sequentially according to the AIPs that they employ. The scheme for this sequential regression is as follows:,

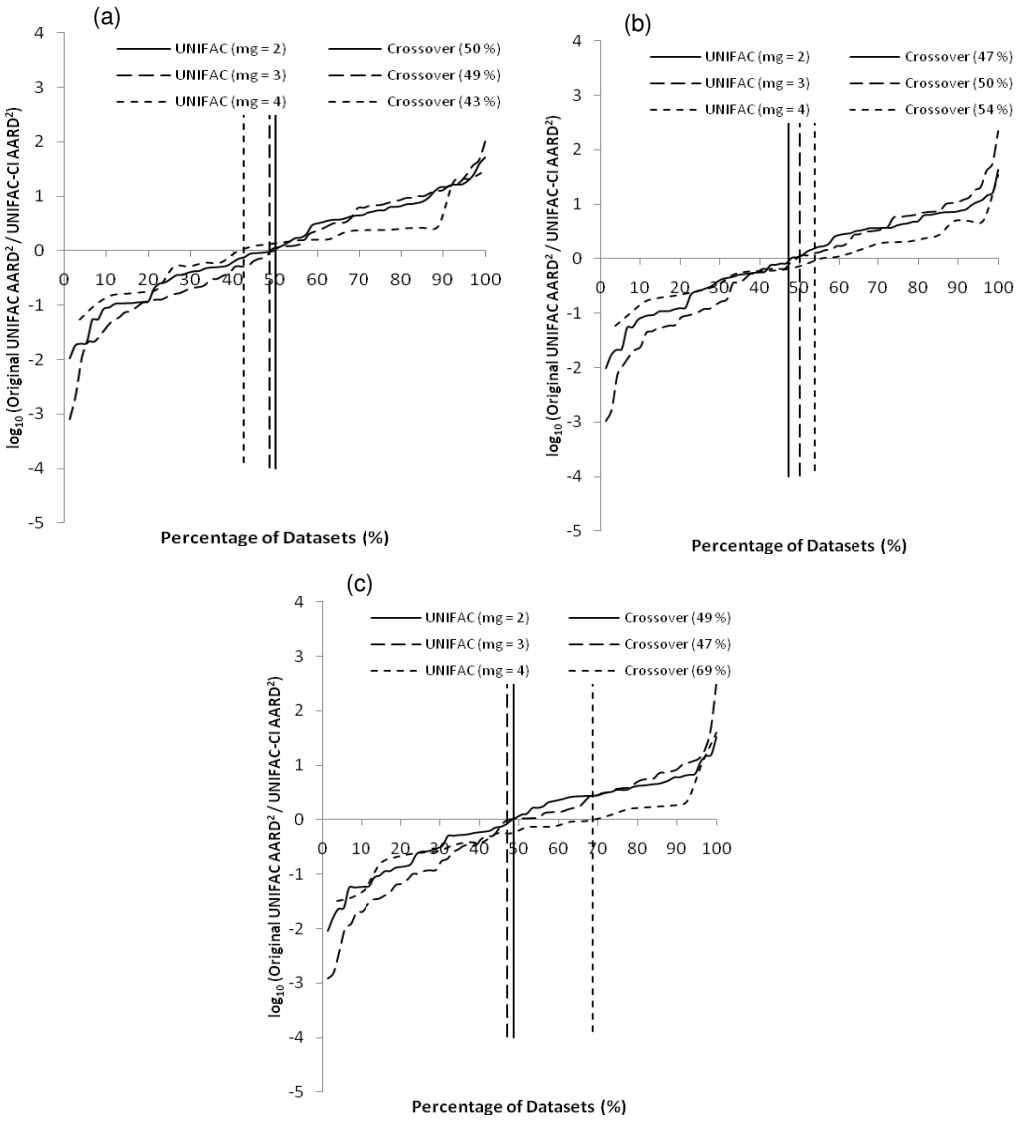


Figure 3.6: Ratio of AARD of Reference UNIFAC and UNIFAC-CI against Fraction of the Datasets for C, O and H Atoms Related Systems with (a) Weight 1, (b) Weight 5, and (c) Weight 10 (in the Objective Function) at Concentrations Lower than 0.5 for Problematic Systems. The Crossover Indicates the Percentage of the Datasets which are better Correlated by the Reference Model. mg is the Number of Main Groups.

Step 1: Regress the parameters for the systems alkane/cyclic alkane–alcohol
Step 2: Keeping all parameters regressed in Step 1 fixed, regress only the new parameters introduced by the systems alkane/cyclic alkane–ethoxy, alkane/cyclic alkane–ether, alkane/cyclic alkane–tetrahydrofuran

Step 3: Keeping all parameters regressed in Step 2, regress only the new parameters introduced by the systems alkane–aldehyde, alkane/cyclic alkane–ester, alkane–carboxylic acid, alkane/cyclic alkane–ketone
Step 4: Keeping all parameters regressed in Step 3, regress only new parameters introduced by the systems alcohol-aromatic
Step 5: Keeping all parameters regressed in Step 4, regress only new parameters introduced by the systems aromatic hydroxyl-aromatic, ethoxy-aromatic, ester-aromatic, alkene-ester, aldehyde-aromatic, ketone-aromatic, ether-aromatic, carboxylic acid-aromatic, aromatic-aromatic aldehyde, tetrahydrofuran-aromatic, dioxane-aromatic
Step 6: Keeping all parameters regressed in Step 5, regress only new parameters introduced by the systems ethoxy alcohol-ether, alcohol-ether, dioxane-alcohol
Step 7: Keeping all parameters regressed in Step 6, regress only new parameters introduced by the systems ketone-alcohol, alcohol-carboxylic acid, ester-alcohol
Step 8: Keeping all parameters fixed at Step 7, regress only the parameters introduced by the system aldehyde–ester, ether–ketone, tetrahydrofuran–ester, tetrahydrofuran–ketone, aldehyde–ketone, carboxylic acid–ketone, ketone–ester, diol–ether

Only in step-3 where the alkane–aldehyde system appears, the new higher order CI and its parameter is introduced. The complete regression-order list and their corresponding binary systems and the involved AIPs are shown in Table 3.3.

Table 3.3: Order of Parameter Regression When Introducing a Second Order CI for the CH_3-CHO Interaction.

Types of systems	AIPs	Order of Regression
Alkane/Cyclic Alkane-Alcohol	$\boldsymbol{b_{C\text{-}O}, bh_{O\text{-}C}}$	1
Alkane/Cyclic Alkane-Ethoxy, Alkane/Cyclic Alkane-Ether, Alkane/Cyclic Alkane-Tetrahydrofuran	$b_{C\text{-}O}, \boldsymbol{c_{C\text{-}O}}, bh_{O\text{-}C}$	2
Alkane-Aldehyde, Alkane/Cyclic Alkane-Ester, Alkane-Carboxylic Acid, Alkane/Cyclic, Alkane-Ketone	$b_{C\text{-}O}, c_{C\text{-}O}, \boldsymbol{e_{C\text{-}O}}, bh_{O\text{-}C}$	3
Alcohol-Aromatic	$b_{C\text{-}O}, bh_{O\text{-}C}, \boldsymbol{ch_{O\text{-}C}, eh_{O\text{-}C}}$	4
Aromatic Hydroxyl-Aromatic, Ethoxy-Aromatic, Ester-Aromatic, Alkene-Ester, Aldehyde-Aromatic, Ketone-Aromatic, Ether-Aromatic, Carboxylic Acid-Aromatic, Aromatic-Aromatic Aldehyde, Tetrahydrofuran-Aromatic, Dioxane-Aromatic	$b_{C\text{-}O}, c_{C\text{-}O}, \boldsymbol{d_{C\text{-}O}}, e_{C\text{-}O}, bh_{O\text{-}C}, ch_{O\text{-}C}, \boldsymbol{dh_{O\text{-}C}}, eh_{O\text{-}C}$	5
Ethoxy Alcohol-Ether, Alcohol-Ether, Dioxane-Alcohol	$b_{C\text{-}O}, \boldsymbol{b_{O\text{-}C}}, \boldsymbol{b_{O\text{-}O}}, c_{C\text{-}O}, \boldsymbol{c_{O\text{-}C}}, \boldsymbol{c_{O\text{-}O}}, \boldsymbol{bh_{C\text{-}O}}, bh_{O\text{-}C}, \boldsymbol{bh_{O\text{-}O}}$	6
Ketone-Alcohol, Alcohol-Carboxylic Acid, Ester-Alcohol	$b_{C\text{-}O}, b_{O\text{-}C}, b_{O\text{-}O}, c_{C\text{-}O}, c_{O\text{-}C}, c_{O\text{-}O}, e_{C\text{-}O}, \boldsymbol{e_{O\text{-}C}}, \boldsymbol{e_{O\text{-}O}}, bh_{C\text{-}O}, bh_{O\text{-}C}, bh_{O\text{-}O},$	7
Aldehyde-Ester, Ether-Ketone, Tetrahydrofuran-	$b_{C\text{-}O}, b_{O\text{-}C}, b_{O\text{-}O}, c_{C\text{-}O}, c_{O\text{-}C}, c_{O\text{-}$	8

Ester, Tetrahydrofuran-Ketone, Aldehyde-Ketone, Carboxylic Acid-Ketone, Ketone-Ester, Diol-Ether	$_{O}$, $d_{C\text{-}O}$, $\mathbf{d_{O\text{-}C}}$, $\mathbf{d_{O\text{-}O}}$, $e_{C\text{-}O}$, $e_{O\text{-}C}$, $e_{O\text{-}O}$, $bh_{C\text{-}O}$, $bh_{O\text{-}C}$, $bh_{O\text{-}O}$, $\mathbf{ch_{C\text{-}O}}$, $ch_{O\text{-}C}$, $\mathbf{ch_{O\text{-}O}}$, $\mathbf{dh_{C\text{-}O}}$, $dh_{O\text{-}C}$, $\mathbf{dh_{O\text{-}O}}$, $\mathbf{eh_{C\text{-}O}}$, $eh_{O\text{-}C}$, $\mathbf{eh_{O\text{-}O}}$,

Note that:

- $bh_{X-Y} = \overline{b_{X-Y}}$ with X, Y = C or O atoms
- Bolded AIPs are the new parameters introduced in each regression step and are regressed. Unbolded AIPs are fixed in each step.

It was found that when the second order CI is introduced for the CHO group for interactions with all other groups, the correlation results of all other systems which contain aldehyde, such as, aldehyde–aromatic, aromatic–aromatic aldehyde, aldehyde–ester and aldehyde–ketone became worse. Therefore, the second order CI interaction was introduced only in the step-3 to account specifically for the CH_3–CHO group interaction.

Very good correlation results were obtained for the alkane–aldehyde systems although some systems involving alkane–ketone and cyclic alkane–ester show a slight (but negligible) increase in the AARD. The correlation results for the alkane–aldehyde systems are given in Table 3.4 and compared against the AARD corresponding to only using the zeroth and first order CI for the CHO group. Figure 3.7 shows the vapour–liquid equilibrium (VLE) phase diagram for the alkane–aldehyde systems for UNIFAC-CI when the zeroth and first order CI were assigned for the CHO group and when second order CI was added.

Table 3.4: Correlation Results for Alkane–Aldehyde Systems Compared between Different Representations of the CHO Group.

Systems	AARD (%)	
	CHO group represented by 0^{th} and 1^{st} order CI	CHO group represented by 0^{th}, 1^{st} & 2nd order CI
1-Butanal-n-Heptane (318 K)	12.18	1.27
2-Methylpropanal-n-Heptane (318 K)	11.51	1.06
2-Methylpropanal-n-Heptane (335 K)	11.43	1.01
1-Butanal-n-Heptane (343 K)	10.35	2.02
1-Pentanal-n-Heptane (348 K)	10.75	0.45

Figure 3.7 shows that when the second order CI was introduced for the CHO group, the prediction by the UNIFAC-CI model agrees well with the experimental data compared to when only zeroth and first order CI were used for describing the CHO group. This shows that by revising the group representation through CI might improve the correlation for certain problematic systems. That is,

providing more structural information to the model. The new sets of AIPs for systems that have been refitted are presented in Table 3.5.

Generally, whenever higher order CI parameters are added and new AIPs are regressed, these new AIPs can be used to predict the corresponding GIPs and thus the phase behavior. These new sets of regressed AIPs can always be used to predict unknown/missing interactions. Currently the parameters (AIPs) are available for atoms C, H, O, N, Cl and S. Therefore, any GIPs which consist of these atoms can possibly be predicted. Furthermore, the same experimental dataset is used to regress the GIPs as well as AIPs. However, the AIPs are then used to predict missing GIPs. Therefore, even with additional AIPs it is pure prediction of the GIPs. Note that it is not proposed that the UNIFAC-CI model should replace the reference UNIFAC model. What is proposed is that the UNIFAC-CI model could be used to fill in the gaps of the reference UNIFAC model group interaction parameters, especially when experimental data is not available.

Table 3.5: New Set of AIPs Regressed after the Introduction of a Second Order CI for CHO Group.

AIPs	Value	AIPs	Value
b_{C-C}	977.7980	c_{C-O}	-65.5433
c_{C-C}	-108.1096	e_{C-O}	67.5523
e_{C-C}	-109.4275	$\overline{b_{C-C}}$	-145.1011
b_{C-O}	-1149.5886	$\overline{b_{O-C}}$	71.5612

3.3 Extrapolation of the Original UNIFAC-CI to SLE Systems (Hydrocarbons and Oxygenated Systems)

In the work of Gmehling et al. [87] and Jakob et al. [88], the Original UNIFAC model with GIPs regressed from VLE data alone showed moderately good results when they were used to predict solid–liquid equilibrium (SLE). Here, the applicability of using the UNIFAC GIPs generated through CIs is illustrated. Extrapolation results for four binary systems are highlighted: p-toluic acid–acetic acid, anthracene–cyclohexane, epsilon–caprolactone–1-propanol and ibuprofen–1-octanol. The GIPs used in this SLE prediction are from the AIPs published in [18].

Figure 3.8 (a) shows that the predictions for the system of *p*-toluic acid–acetic acid for UNIFAC-CI model conforms closely to the experimental data compared to the original UNIFAC and ideal models. In addition to that, Figure 3.8 (b) shows that the SLE predictions for anthracene, a solid polycyclic aromatic hydrocarbon consisting of three fused benzene rings with cyclohexane with UNIFAC and UNIFAC-CI are also close to each other and follow closely the trend of the experimental data (note that this is a purely hydrocarbon system and VLE analysis was also favorable).

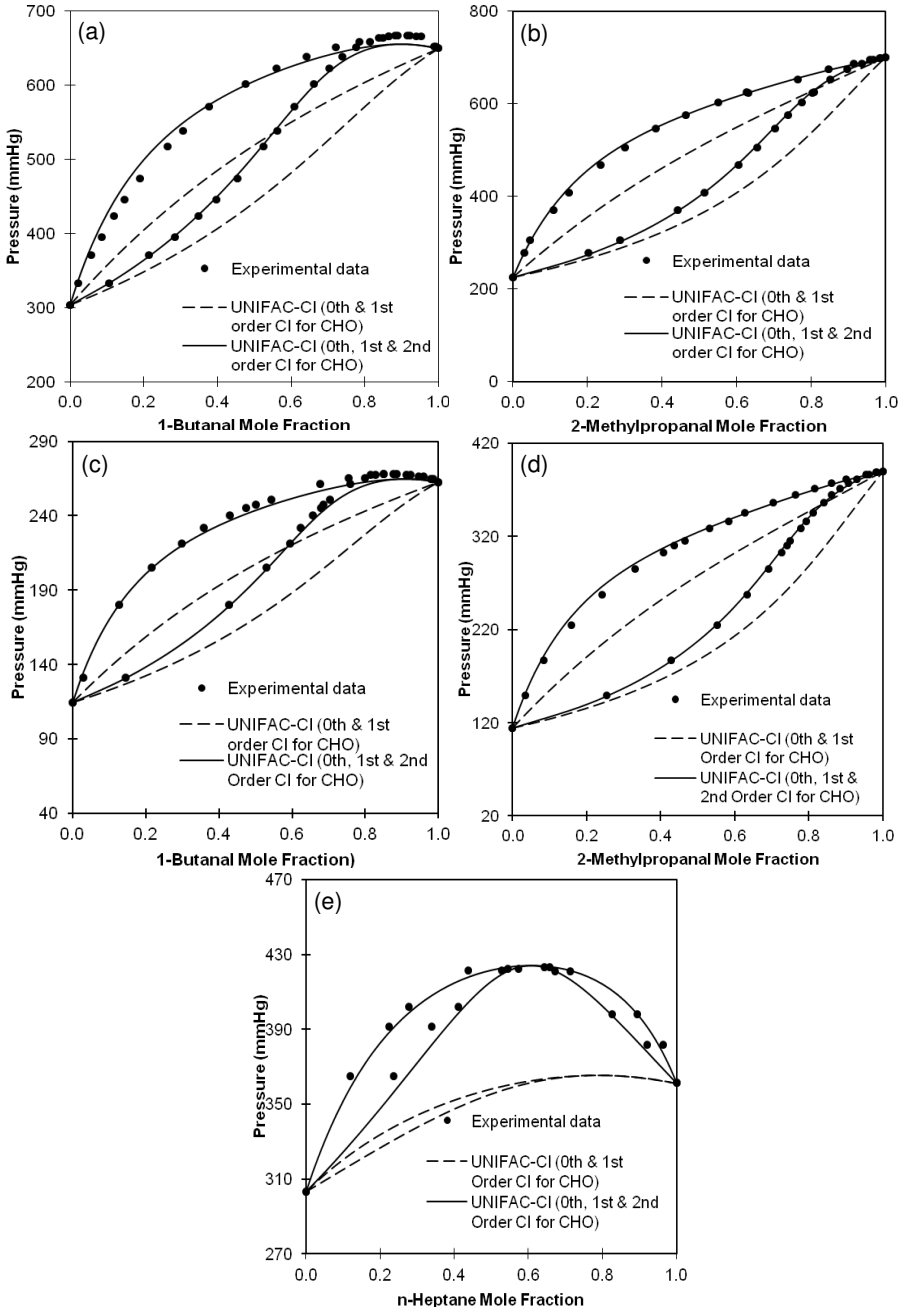


Figure 3.7: UNIFAC-CI VLE Prediction for the Systems (a) 1-Butanal-n-Heptane at 343 K, (b) 2-Methylpropanal-n-Heptane at 335 K, (c) 1-Butanal-n-Heptane at 318 K, (d) 2-Methylpropanal-n-Heptane at 318K and (e) 1-Pentanal-n-Heptane at 348 K.

Furthermore, as shown in Figure 3.8 (c) the predictions for both reference UNIFAC and UNIFAC-CI models are quite good compared to the experimental data for the system of epsilon–caprolactone–1-propanol. Finally, the extrapolation of UNIFAC-CI method is demonstrated for system of ibuprofen ($C_{13}H_{18}O_2$) and 1-octanol. Although only two experimental data points are available, both reference UNIFAC and UNIFAC-CI models provide satisfactory prediction with a small difference at the eutectic point of both methods. The values of melting temperatures and heats of fusion of the compounds are predicted using the Marrero and Gani [8] group contribution method with values reported in [122].

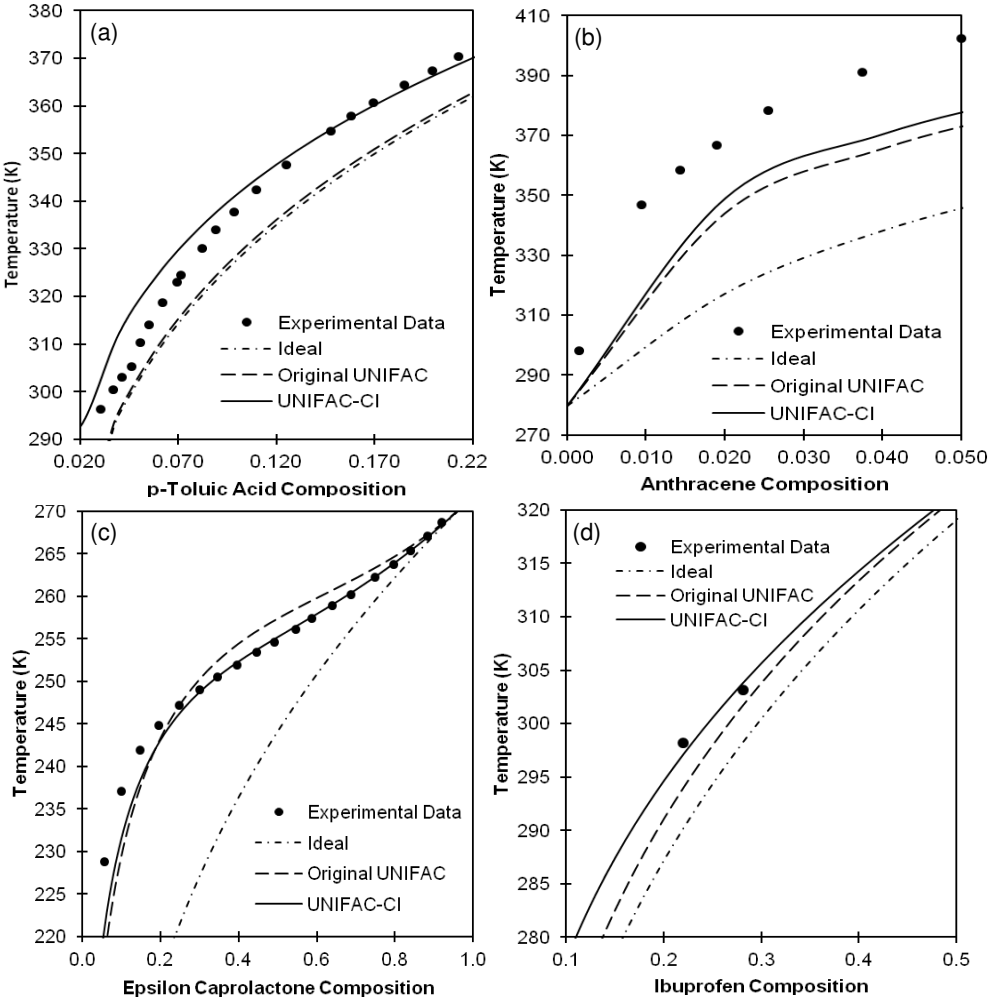


Figure 3.8: UNIFAC-CI SLE Prediction for the Systems: (a) p-Toluic Acid-Acetic-Acid, (b) Anthracene-Cyclohexane, (c) Epsilon-Caprolactone-1-Propanol and (d) Ibuprofen-1-Octanol. Experimental Data are from [122].

The above SLE results confirm that the GIPs generated from CIs which were regressed from VLE data alone can be extrapolated for predictions of some SLE systems with satisfactory results. This is evaluated in more detail in Chapter 5.

3.4 Conclusion of the Analysis

The UNIFAC-CI models developed by González et al. [16-18] are powerful predictive tools which can be used in cases where reference UNIFAC model parameters are missing. However, UNIFAC-CI model does not always perform well and deviations are seen also against the reference model, for certain systems. An analysis was therefore initiated, at first in order to study molecules containing C, H and O with special focus at low concentrations. Furthermore, by revising the CI description of group CHO, which appeared to be the cause for unreliable predictions, the performance of UNIFAC-CI model improved, especially for alkane–aldehyde systems. This group revision can be continued to include other systems with N, S, and other atoms. Next, the extrapolation capabilities with respect to applications in SLE have been highlighted through predictions of SLE curves for four binary systems. It should be noted that these results are extrapolations both in terms of temperature and equilibrium data, as the GIPs were predicted from AIPs regressed from VLE data.

CHAPTER 4

FURTHER ANALYSIS AND DEVELOPMENT OF THE ORIGINAL UNIFAC-CI MODELS: VAPOR-LIQUID EQUILIBRIUM SYSTEMS

4.1 Introduction

In this chapter, we present the further development and analysis of the Original UNIFAC-CI model for vapor-liquid equilibrium (VLE) systems. Previously in the early development of the Original UNIFAC-CI model, Gonzalez et al. [18], have regressed the atom interaction parameters (AIPs) against vapor-liquid equilibrium (VLE) experimental data involving systems with C, O, N Cl and S atoms. The objective of that work was to increase the application range of UNIFAC and to regress and obtain the best AIPs for the CI-model. The work has been continued by Mustaffa et al. [120], with a detailed analysis of the performance and trends of predictions of VLE systems involving only C, H, O related systems using the CI-model. Mustaffa et al. investigated also why the UNIFAC-CI model is working well for certain systems but less satisfactory for other systems. The work of Mustaffa et al. has been presented in Chapter 3.

In the work done by Mustaffa et al. [126], as a result of the detailed analysis, new set of AIPs have been revised for systems with C, H and O atoms. Therefore, the AIPs for the remaining systems with N, Cl and S atoms also need to be revised. The result of the correlations in terms of deviations errors and predictions are presented in this chapter. In addition to that, another set of AIPs for VLE systems (involving systems with C, O, N, Cl and S atoms) have been regressed using the Q_{VLE} quality factors obtained from a quality assessment algorithm (consistency tests) as weighting factors in the objective functions. The results of the correlations and predictions of these 2 sets of AIPs are compared and discussed.

4.2 Development of the Original UNIFAC-CI Model for VLE

In this section, we present the parameter regression work for the further development of the Original UNIFAC-CI model. The atom interaction parameters (AIPs) of Original UNIFAC-CI are regressed against VLE data. This work is divided into 2 parts. In the first part, as a continuation of the analysis and regression made by Mustaffa et al. [126] for C, H, O atoms related system, the regression of the remaining AIPs for systems containing N, Cl and S atoms is performed. All *P, T,* x_i, y_i VLE data used in the first part have passed the Van Ness [131] VLE consistency check and no weights have been assigned to the datasets in the objective function for parameter regression. In the second part of the regression work, the same set of VLE data (including the C, O, H atoms related systems) have been checked using a quality assessment algorithm developed by Kang et al. [127] described in section 2.6.5 which uses four VLE consistency tests (Herington, Van Ness,

Infinite Dilution and Differential). This will generate a VLE quality factor, Q_{VLE} (with a maximum value of 1) for each of the dataset and these factors are used as weighting factors for each dataset in the objective function. The correlation results of these 2 sets of parameters are compared together with the reference model (i.e. Original UNIFAC).

4.2.1 Background of Parameter Estimation

The minimization technique used to regress the parameters in this work is the least squares technique using the Modified Levenberg-Marquardt approach [119] with the algorithm described in section 2.6.3 which has been slightly modified to regress AIPs for the Original UNIFAC-CI models. The same technique of parameter regression was used in the previous work [18].

In this work, the parameter regression is based on the *P, T, x_i, y_i* VLE data and only isothermal data were used for the regression. Since the unit and magnitude of the pressures reported in each of the VLE data are different (mmHg, torr, kPa, Pa), the objective function below has been chosen to regress the parameters which is expressed as the average relative pressure quadratic deviation.

The overall objective function (OF) used is as follows:

$$OF = \sum_{set} Q_{VLE} \left[\frac{1}{N} \sum_{i=1}^{N} \left(\frac{P_{i-\exp} - P_{i-calc}}{P_{i-\exp}} \right)^2 \right] + w_{reg} \sum_{i=1}^{N} \left(AIP_i - AIP_i^{IG} \right)^2 \tag{4.1}$$

where P_{exp} is the experimental pressure, N is the number of experimental data points used for the estimation, AIP_j is the current value of the CI-interaction parameter j, AIP^{IG}_j its corresponding initial guess and w_{reg} (values between $1x10^2$ and $1x10^{10}$) is a weighing value used to increase and decrease the influence of regularization in the optimization. Q_{VLE} is the VLE quality factor described by Kang et al. [127] generated for each dataset.

The equilibrium pressure, P_{i-calc} was calculated in two different ways depending on whether the systems need an association term (such system involving carboxylic acids) or not. For systems without association term, the total pressure is calculated as follows:

$$P_{i-calc} = \sum_i x_i \gamma_i P_i^{sat} (POY_i) \tag{4.2}$$

where *i* is an index running over all species in the mixture and POY_i is the Poynting factor. However, for the systems needing the association term, the equilibrium pressure is calculated using fugacity coefficient and association based on the method of Hayden and O'Connell [120] described briefly as follows:

$$P_{i-calc} = \sum_i \frac{x_i \gamma_i P_i^{sat} (POY_i)}{\Phi_i} \tag{4.3}$$

$$\ln \Phi_i = \left[2\sum_{j=1}^{N} y_j B_{ij} - B \right] \frac{P}{RT} \tag{4.4}$$

where Φ_i is the fugacity coefficient, y_j is the vapor mole fraction, $B_{ij}(T)$ is the second virial coefficient characterizing pair interactions between *i* and *j* molecule and B is the second virial coefficient. The cross second virial coefficient, B_{ij} can be calculated directly from PVT data, from statistical mechanical formulas or from empirical and semitheoretical correlations [120]. On the other hand, the second virial coefficient, B are calculated using the equations below which are contributed by the different types of intermolecular forces which are described as bound, metastably bound, free pairs and B_{chem} which include the association contribution.

$$B = B_{free} + B_{metastable} + B_{bound} + B_{chem} \tag{4.5}$$

The bound and metastably bound contribution of the second virial coefficient are calculated in Equation (4.6) while the free pairs contribution is calculated in Equation (4.7):

$$B_{metastable} + B_{bound} = b_0 A \exp\left[\Delta H / kT / \varepsilon\right] \tag{4.6}$$

$$\text{with } b_0 = \frac{2\pi}{3} N_0 \sigma^3,\ A = -0.3 - 0.05\mu^*,\ \Delta H = 1.99 + 0.2\mu^{*2},\ \mu^* = \mu^2 / \varepsilon\sigma^3$$

$$B_{free} = b_0 \left[(0.94 - 1.47/T^{*'} - 0.85/T^{*'2} + 1.015/T^{*'3}) - \mu^{*'}(0.75 - 3/T^{*'} + 2.1/T^{*'2} + 2.1/T^{*'3}) \right] \tag{4.7}$$

where ε and σ are the effective nonpolar potential parameters while μ is the moleculat dipole moment. Moreover, by taking into account the contribution of association, B_{chem} is calculated as follows:

$$B_{chem} = b_0 \exp\left\{\eta\left[650/(\varepsilon/k + 300) - 4.27\right]\right\} \times \left\{1 - \exp\left[1500\eta/T\right]\right\} \tag{4.8}$$

where η is the association parameter.

Moreover, the regularization term in the objective function is used in order to ensure that the regressed parameters do not go too far from the initial values and only used whenever needed. This would be useful as the previously investigated systems can still be represented with comparable accuracy. The higher the value of the weighing value, w_{reg} the higher the influence of the regularization.

4.2.2 Equilibrium Data

The input for the parameter estimation are VLE experimental data (involving C, H, O, N Cl and S atoms) and the statistics of data used in the regression work are summarized in Table 4.1. All VLE

data have been initally tested for consistency using a quality assesment algorithm which combines four widely used consistency tests (Herington, Van Ness, Differential and Infinite Dilution tests) and a check between the consistency of the binary and the pure component vapor pressure [127].

For the parameter regression work, a total of 363 VLE datasets with 5627 data points consisting of *P, T,* x_i*,* y_i and *P, T,* x_i data have been used. Only a moderate amount of experimental data were used because the purpose of the development of the Original UNIFAC-CI models is to be able to predict phase equilibria with a limited amount of experimental data and without using new experimental data [18].

Table 4.1: Statistics of the Data Used in the Parameter Regression.

Phase Equilibria Information	Types of system	Type of data	No. of systems	Data Points	Main Groups Involved
VLE	Hydrocarbons	*P, T,* x_i*,* y_i	49	742	CH_2, C=C, ACH, $ACCH_2$, OH, ACOH, CH_2CO, CHO, CCOO, HCOO, CH_2O, COOH
	Oxygenated	*P, T,* x_i*,* y_i	161	2752	
	Nitrogenated	*P, T,* x_i*,* y_i	28	785	CH_2, ACH, $ACCH_2$, OH, CH_2CO, CCOO, CNH_2, CNH, $(C)_3N$, $ACNH_2$, PYR, CCN, CNO_2
		P, T, x_i	30		
	Chlorinated	*P, T,* x_i*,* y_i	46	1200	CH_2, C=C, ACH, $ACCH_2$, OH, CH_2CO, CCOO, CH_2O, CCl, CCl_2, CCl_3, CCl_4, ACCl
		P, T, x_i	36		
	Sulfurated	*P, T,* x_i*,* y_i	13	148	CH_2, C=C, ACH, OH, CH_2CO, CCOO, CH_2O, CCl_3, CCl_4, DMSO, CH_2S
	Total		**363**	**5627**	

4.2.3 Regression Procedure

In this work, the atom interactions parameters (AIPs) are regressed in series. Initially, only systems with C-C atom interactions are used for the parameter regression. For this step, a total of 49 data sets of VLE systems were used. The parameters involved in the regression of AIPs related to C-C interaction are $b_{C\text{-}C}$, $c_{C\text{-}C}$, $d_{C\text{-}C}$, $e_{C\text{-}C}$, $bh_{C\text{-}C}$, $ch_{C\text{-}C}$, $dh_{C\text{-}C}$, $eh_{C\text{-}C}$.

Next, systems related to the C-O and O-C interactions are used for parameter regression where the AIPs, $b_{C\text{-}O}$, $c_{C\text{-}O}$, $d_{C\text{-}O}$, $e_{C\text{-}O}$, $b_{O\text{-}C}$, $c_{O\text{-}C}$, $d_{O\text{-}C}$, $e_{O\text{-}C}$, $bh_{C\text{-}O}$, $ch_{C\text{-}O}$, $dh_{C\text{-}O}$, $eh_{C\text{-}O}$, $bh_{O\text{-}C}$, $ch_{O\text{-}C}$, $dh_{O\text{-}C}$, $eh_{O\text{-}C}$ are regressed while fixing the AIPs related to the C-C interactions which were regressed earlier. A total of 129 data sets of VLE systems were used in this step. Following in the next sequence are the systems related to the O-O interactions, where the AIPs, $b_{O\text{-}O}$, $c_{O\text{-}O}$, $d_{O\text{-}O}$, $e_{O\text{-}O}$, $bh_{O\text{-}O}$, $ch_{O\text{-}O}$, $dh_{O\text{-}O}$, $eh_{O\text{-}O}$ are regressed while fixing the AIPs related to the C-C, C-O and O-C interactions which were regressed earlier. A total of 32 data sets of VLE systems were used in this step.

The regression continued in the same procedure where the parameters related to interactions C-N, O-N, N-N, C-Cl, O-Cl, C-S, O-S and Cl-S were added sequentially. In each step, the previously fitted parameters were fixed. The overall regression procedure is illustrated in Figure 4.1.

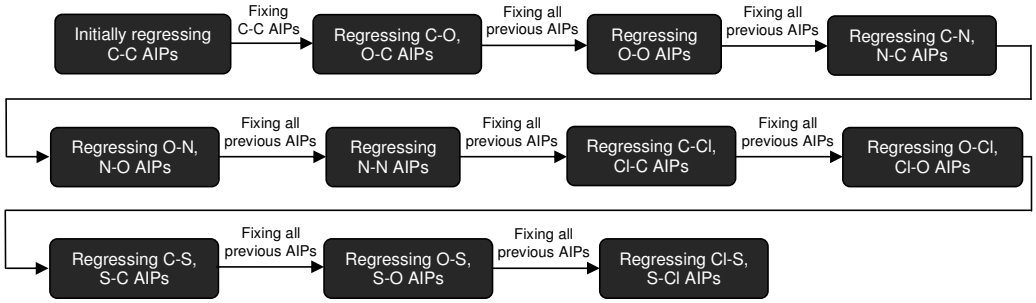


Figure 4.1: Overall Regression Procedure for the AIPs Parameter Estimation.

4.2.4 Correlation Results

The correlation error or deviation between the experimental data and the regressed values are defined in terms of the average absolute relative deviation (AARD) shown in Equation (4.9):

$$AARD(\%) = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{P_{i-\exp} - P_{i-calc}}{P_{i-\exp}}\right| \times 100 \tag{4.9}$$

In this section, detailed correlation results of each of the datasets involved in the parameter regression work are presented and described. There are 2 sets of parameters which are regressed against the VLE systems. For the first set, the quality factor for each of the dataset is assigned as 1 and labeled by UNIFAC-CI (VLE) in Tables 4.2-4.11. For the second set, all the datasets were assigned a weighting factor, Q_{VLE} (obtained from the quality assessment algorithm described in [121]) which is used in the objective function and the values are also reported in Tables 4.2-4.11. This result is labeled as UNIFAC-CI (VLE) with Q_{VLE}. At the end of this section, the overall correlation results are presented and discussed.

The correlation results for systems related to only C-C interactions which are hydrocarbons (involving aromatics-alkane and alkane-alkene systems) are presented in Table 4.2. The average AARD obtained for the Original UNIFAC model is 1.1 % while for the UNIFAC-CI (VLE) and UNIFAC-CI (VLE) with Q_{VLE} are 1.3 and 1.2 % respectively. For the CI-models, we can say that for hydrocarbons, the performance is as good as the reference model. However, we can see that there is a very small decrease of error when Q_{VLE} factors are used in the parameter estimation.

Table 4.2: Correlation Results for C-C Atom Interaction Related Systems.

No.	Systems	Q_{VLE}	AARD (%)		
			Original UNIFAC	UNIFAC-CI (VLE)	UNIFAC-CI (VLE) with Q_{VLE}
1	Cyclohexane-Toluene at 323 K	1.00	0.94	1.69	1.53
2	m-Xylene-Decane at 394 K	0.32	1.76	3.64	3.47
3	Methylcyclohexane-Benzene at 348 K	1.00	0.29	1.69	1.65
4	Methylcyclohexane-Toluene at 348 K	1.00	0.37	0.57	0.41
5	1-Hexene-Hexane at 333 K	0.83	0.50	0.45	0.12
6	1-Hexene-Hexane at 333 K	1.00	0.56	0.50	0.13
7	Ethylbenzene-Nonane at 333 K	0.95	0.24	0.69	0.54
8	Propylbenzene-Octane at 343 K	0.96	0.38	0.30	0.26
9	Propylbenzene-Nonane at 363 K	1.00	0.90	0.35	0.47
10	Tetradecane-Benzene at 303 K	0.35	5.41	1.50	1.59
11	Benzene-n-Heptane at 298 K	1.00	1.49	0.42	0.42
12	1-Hexene-Hexane at 313 K	0.80	0.59	0.51	0.07
13	Ethylbenzene-Cyclooctane at 343 K	0.60	0.38	0.73	0.60
14	Ethylbenzene-Cyclooctane at 353 K	0.59	0.38	0.30	0.22
15	Ethylbenzene-Cyclooctane at 373 K	0.81	0.69	0.21	0.26
16	Benzene-2,2,4-Trimethylpentane at 313 K	1.00	2.25	0.64	0.67
17	Toluene-2,2,4-Trimethylpentane at 313 K	1.00	0.75	1.55	1.36
18	o-Xylene-2,2,4-Trimethylpentane at 313 K	1.00	1.98	1.95	1.82
19	Propylbenzene-2,2,4-Trimethylpentane at 313 K	1.00	1.79	2.45	2.25
20	Benzene-Cyclohexene at 293 K	0.76	0.98	2.00	1.74
21	2,2,4-Trimethylpentane-Toluene at 373 K	1.00	1.24	0.26	0.41
22	Benzene-n-Heptane at 333 K	0.67	0.49	2.00	1.97
23	Benzene-Cyclohexene at 298 K	0.97	1.10	1.70	1.38
24	Benzene-Cyclohexene at 348 K	1.00	1.93	2.18	2.54
25	Benzene-Cyclohexane at 283 K	1.00	2.19	1.11	1.14
26	Benzene-Cyclohexane at 333 K	1.00	0.25	0.88	0.85
27	1-Heptene-n-Heptane at 328 K	0.94	0.48	0.52	0.87
28	1-Heptene-n-Octane at 328 K	1.00	3.14	3.19	3.58
29	Benzene-Cyclohexane at 298 K	1.00	0.79	0.38	0.35
30	Benzene-n-Heptane at 328 K	0.64	1.07	0.98	0.97
31	n-Hexane-Benzene at 333 K	0.68	0.14	1.23	1.20
32	Methylcyclopentane at 333 K	0.76	0.43	1.07	1.05
33	n-Hexane-Benzene at 298 K	1.00	1.56	0.28	0.28
34	Benzene-n-Heptane at 298 K	1.00	1.55	0.33	0.32
35	Benzene-Cyclohexane at 343 K	1.00	0.17	1.39	1.37

36	1-Hexene-Benzene at 298 K	1.00	0.25	3.82	3.33
37	1-Hexene-n-Hexane at 328 K	1.00	0.37	0.32	0.05
38	1-Hexene-n-Octane at 328 K	0.82	2.27	2.18	1.64
39	Benzene-1-Heptene at 328 K	1.00	0.35	1.21	1.05
40	Benzene-n-Heptane at 328 K	1.00	0.31	1.09	1.06
41	1-Heptene-Toluene at 328 K	1.00	1.54	5.62	6.27
42	n-Heptane-Toluene at 328 K	1.00	0.41	1.03	0.88
43	Benzene-Cyclohexane at 293 K	1.00	1.26	0.44	0.46
44	Benzene-Cyclohexane at 313 K	1.00	1.04	0.41	0.37
45	Methylcyclopentane-Benzene at 298 K	1.00	1.48	0.22	0.22
46	Methylcyclopentane-Benzene at 313 K	1.00	0.79	0.57	0.54
47	n-Heptane-Ethylbenzene at 298 K	0.64	1.03	1.68	1.46
48	Benzene-Toluene at 334 K	0.55	1.00	2.80	2.37
49	Cyclohexane-Toluene at 318 K	1.00	0.79	1.49	1.33
	Average		**1.06**	**1.28**	**1.20**

In Table 4.3, the correlation results with respect to C-O and O-C atom interactions are shown. There are systems such as aromatics-ethers, alkanes-ethers, esters-alkenes, esters-alkanes, alkanes-aldehydes, aromatics-aldehydes, alcohols-alkane, alcohols-aromatics, ketones-aromatics, ketones-alkanes, carboxylic acids-aromatics and carboxylic acids-alkanes are presented. The average AARD obtained for the reference model is 2.6 %. However, the average correlation error of the CI-model with Q_{VLE} equal to 1 is a bit higher with 3.7 %. When the generated Q_{VLE} factors are used, the average correlation error of the CI-model improved to 2.6 % which is as good as the reference model.

Table 4.3: Correlation Results for C-O, O-C Atom Interaction Related Systems.

No.	**Systems**	**Q_{VLE}**	**AARD (%)**		
			Original UNIFAC	**UNIFAC-CI (VLE)**	**UNIFAC-CI (VLE) with Q_{VLE}**
1	1,2-Dimethoxyethane-Toluene at 350 K	0.80	9.18	15.87	9.40
2	1,2-Dimethoxyethane-Methylcyclohexane at 350 K	1.00	16.76	17.22	6.38
3	Tetrahydrofuran-Hexane at 313 K	0.98	3.99	4.66	5.51
4	Tetrahydrofuran-Hexane at 333 K	0.98	3.51	4.13	4.97
5	Cyclohexane-tert-Butanol at 328 K	1.00	4.58	3.04	2.06
6	Butyraldehyde-n-Heptane at 318 K	0.78	1.52	1.27	11.07
7	2-Methylpropanal-n-Heptane at 318 K	0.87	3.48	1.06	10.55
8	1-Butanol-n-Heptane at 363 K	1.00	0.98	2.76	3.61
9	Methyl-Acetate-1-Hexene at 323 K	1.00	1.02	2.27	2.15

10	1-Hexene-Ethyl-Acetate at 333 K	0.88	0.28	1.64	1.03
11	Butyl-Acetate-1-Decene at 363 K	0.48	2.29	7.55	4.79
12	2-Butanone-Benzene at 328 K	0.79	1.61	2.62	1.79
13	2-Butanol-Hexane at 333 K	1.00	2.53	2.06	2.36
14	Butyl-Acetate-n-Heptane at 347 K	0.24	1.04	11.29	6.10
15	Ethyl-Butyrate-n-Heptane at 373 K	0.77	2.10	4.56	1.76
16	2-Methyl-1-Propanol-Ethylbenzene at 313 K	1.00	2.86	3.98	3.88
17	2-Methyl-1-Propanol-p-Xylene at 313 K	1.00	5.82	3.88	3.50
18	tert-Butanol-Benzene at 313 K	1.00	1.26	4.12	3.10
19	tert-Butanol-Toluene at 313 K	1.00	0.60	0.61	0.60
20	tert-Butanol-Ethylbenzene at 313 K	1.00	2.07	1.53	1.24
21	Methyl-Butyl-Ether-Benzene at 343 K	0.79	0.13	2.78	2.18
22	Diisopropyl-Ether-Benzene at 343 K	0.94	1.02	0.76	1.37
23	2-Propanol-Hexane at 328 K	0.97	1.83	3.45	4.49
24	2-Propanol-Octane at 353 K	0.92	5.30	2.57	1.96
25	2-Methyl-1-Propanol-n-Heptane at 333 K	1.00	0.73	3.02	3.82
26	Phenol-Styrene at 373 K	0.93	4.04	13.00	11.69
27	1-Propanol_Octane at 363 K	1.00	2.28	1.60	2.70
28	1-Butanol-n-Heptane at 333 K	1.00	1.40	4.80	5.55
29	TAME-Heptane at 313 K	0.51	3.03	3.12	0.84
30	TAME-Heptane at 313 K	1.00	3.71	3.80	1.28
31	TAME-Cyclohexane at 313 K	1.00	2.69	2.74	0.64
32	TAME-Benzene at 313 K	1.00	0.93	4.03	3.42
33	1-Butanol-Octane at 373 K	0.95	1.06	3.36	4.20
34	2-Butanol-Octane at 358 K	0.74	5.55	2.77	1.53
35	tert-Butanol-Octane at 343 K	0.98	7.12	4.63	3.28
36	Acetone-Cyclohexane at 303 K	0.89	2.41	4.66	2.79
37	2-Propanol-Octane at 348 K	1.00	4.02	1.72	1.98
38	Hexane-Methyl-tert-Butyl-Ether at 313 K	1.00	3.38	3.44	0.74
39	Dipentyl-Ether_Undecane at 403 K	0.27	0.85	4.71	0.37
40	Diisopropyl-Ether-n-Heptane at 323 K	0.49	0.28	1.02	0.25
41	Acetone-Cyclohexane at 313 K	0.94	1.47	0.31	2.12
42	Diethoxymethane-n-Heptane at 323 K	0.30	6.07	6.30	0.97
43	Diethoxymethane-n-Heptane at 343 K	0.28	5.38	5.59	1.08
44	1-Butanol-Decane at 373 K	1.00	2.42	1.22	0.88
45	1,2-Dimethoxyethane-2,4-Dimethylpentane at 343 K	1.00	16.04	15.40	4.47
46	n-Heptane-3-Pentanol at 368 K	0.62	5.63	4.01	2.89
47	2-Methyl-1-Propanol-Octane at 373 K	1.00	2.47	1.17	1.49
48	Octane-Methyl-Tert-Amyl-Ether at 323 K	0.83	4.26	4.55	1.79
49	1-Propanol-Octane at 358 K	0.92	1.97	3.57	4.73
50	TAME-2,2,4-Trimethylpentane at 311 K	0.55	3.82	4.13	1.49

51	TAME-Methylcyclohexane at 293 K	0.64	4.41	4.68	2.07
52	TAME-Methylcyclohexane at 311 K	0.61	4.00	4.39	1.99
53	TAME-Amyl-Ether-Methylcyclohexane at 333 K	0.68	3.96	4.30	2.08
54	TAME-Toluene at 293 K	0.49	0.98	0.45	0.33
55	Toluene-tert-Butyl-Ethyl-Ether at 333 K	0.77	0.84	0.38	1.00
56	Methyl-Ethyl-Ketone-n-Heptane at 318 K	1.00	2.75	1.63	2.47
57	n-Heptane-Methyl-Isobutyl-Ketone at 343 K	0.75	3.06	2.90	3.14
58	Benzene-Acetic-Acid at 298 K	0.46	1.29	4.20	3.28
59	n-Heptane-3-Pentanone at 353 K	0.96	2.08	2.74	1.99
60	2-Methyl-1-Propanol-Toluene at 353 K	0.70	0.37	1.07	0.83
61	1,4-Dioxane-Toluene at 353 K	0.81	0.27	2.45	0.38
62	1,4-Dioxane-Toluene at 373 K	0.61	0.18	2.87	0.71
63	Acetone-n-Hexane at 328 K	0.98	1.28	2.27	1.01
64	Benzene-2-Butanol at 318 K	1.00	0.91	2.67	1.65
65	Benzene-2-Methyl-2-Propanol at 318 K	1.00	1.45	4.63	3.36
66	Acetone-n-Hexane at 293 K	0.58	2.22	6.40	3.44
67	Acetone-Benzene at 298 K	1.00	0.94	4.81	0.40
68	Ethyl-Acetate-Cyclohexane at 293 K	0.92	2.91	8.01	2.66
69	2-Pentanone-Toluene at 323 K	0.96	0.98	2.88	1.25
70	Di-n-Propyl-Ether-n-Heptane at 343 K	0.80	1.33	1.55	0.32
71	Benzene-Di-n-Propyl-Ether at 343 K	0.80	0.49	0.72	0.49
72	Benzene-1,2-Dimethoxyethane at 343 K	0.82	3.73	0.88	1.85
73	Benzene-2-Pentanone at 323 K	0.83	0.46	2.62	1.14
74	n-Heptane-3-Pentanone at 327 K	1.00	0.61	1.15	0.55
75	Tetrahydrofuran-Cyclohexane at 313 K	0.85	4.24	4.82	6.10
76	Tetrahydrofuran-Cyclohexane at 323 K	0.97	3.44	4.08	5.37
77	Tetrahydrofuran-Cyclohexane at 333 K	0.79	3.57	4.19	5.24
78	Tetrahydrofuran-Benzene at 303 K	1.00	3.46	1.15	0.91
79	Tetrahydrofuran-Benzene at 313 K	0.87	2.39	0.77	0.68
80	Tetrahydrofuran-Benzene at 323 K	0.88	2.27	0.62	0.56
81	Tetrahydrofuran-Benzene at 333 K	0.94	2.17	0.57	0.49
82	n-Heptane-2-Butanol at 338 K	0.97	3.30	1.00	1.17
83	Acetone-Benzene at 298 K	0.97	1.15	5.08	0.97
84	Methyl-Ethyl-Ketone-Benzene at 298 K	0.83	1.00	3.29	1.02
85	1-Propanol-2,2,4-Trimethylpentane at 328 K	1.00	1.81	2.20	3.50
86	Ethanol-Toluene at 333 K	0.92	0.76	1.16	1.17
87	1-Hexene-Ethyl-Acetate at 313 K	0.26	1.06	2.01	1.92
88	n-Hexane-Isopropanol at 331 K	0.16	6.15	3.75	2.73
89	n-Heptane-Propionic-Acid at 323 K	0.11	7.35	4.07	4.03
90	Acetone-n-Hexane at 313 K	1.00	0.54	3.70	1.27
91	_Acetone-n-Heptane at 313 K	1.00	0.64	4.37	1.49

92 Acetone-n-Octane at 313 K	0.97	1.37	5.09	2.43
93 Acetone-Benzene at 313 K	1.00	1.11	5.24	0.61
94 Acetone-Toluene at 313 K	1.00	2.67	9.96	2.67
95 Acetone-Ethylbenzene at 313 K	0.84	3.58	9.51	2.32
96 n-Heptane-2-Butanol at 348 K	1.00	4.07	1.45	1.09
97 Methyl-Ethyl-Ketone-Benzene at 328 K	0.67	0.66	4.23	0.51
98 n-Heptane-1-Butanol at 363 K	1.00	0.98	2.76	3.60
99 n-Heptane-2-Methyl-1-Propanol at 333 K	1.00	0.73	3.02	3.81
100 Methyl-Ethyl-Ketone-Benzene at 313 K	0.65	1.48	2.03	1.66
101 2-Methylpropanal-n-Heptane at 335 K	0.96	3.32	1.01	10.34
102 Ethyl-Acetate-n-Heptane at 323 K	1.00	0.87	12.57	6.52
103 Cyclohexane-2-Methyl-2-Propanol at 328 K	1.00	4.58	3.04	2.06
104 1-Butanal-Toluene at 313 K	0.59	0.23	4.92	0.91
105 1-Butanal-Benzene at 313 K	1.00	0.54	0.92	1.94
106 1-Butanal-Toluene at 327 K	0.48	1.69	8.43	3.75
107 1-Propanal-Benzene at 313 K	0.71	0.81	0.57	3.19
108 1-Propanal-Toluene at 313 K	0.59	0.61	3.49	0.60
109 Benzene-1-Pentanal at 313 K	0.76	0.23	1.08	1.61
110 1-Pentanal-Toluene at 313 K	0.77	0.53	4.01	1.37
111 Ethylbenzene-Benzaldehyde at 348 K	0.36	8.13	0.67	1.71
112 Styrene-Benzaldehyde at 363 K	0.63	4.42	2.74	1.46
113 Methyl-Tert-Butyl-Ether-Hexane at 313 K	0.73	3.34	4.51	0.73
114 n-Propyl-Formate-Benzene at 303 K	1.00	0.22	7.62	4.89
115 n-Heptane-1-Pentanal at 348 K	0.54	1.25	0.45	8.94
116 Acetone-Cumene at 293 K	0.83	4.60	10.14	2.39
117 1-Butanal-n-Heptane at 343 K	0.59	3.14	2.02	9.58
118 n-Octane-Methyl-Tert-Pentyl-Ether at 323 K	0.83	4.26	4.55	1.79
119 Toluene-3-Methyl-1-Butanol at 368 K	0.73	1.50	1.10	1.07
120 Cyclohexane-Methyl-Ethyl-Ketone at 323 K	0.84	3.27	2.55	3.10
121 Methyl-Propionate-n-Heptane at 323 K	0.21	5.11	2.41	0.98
122 Methyl-Ethyl-Ketone-Toluene at 330 K	1.00	0.47	3.66	0.60
123 Diisopropyl-Ether-Dimethylpentane at 343 K	0.78	0.65	0.79	1.06
124 Toluene-Methyl-Ethyl-Ketone at 323 K	0.83	0.87	3.66	1.23
125 Toluene-2-Pentanone at 323 K	0.96	0.99	2.88	1.25
126 Toluene-Methyl-Isobutyl-Ketone at 323 K	1.00	2.89	4.06	0.25
127 Acetone-Benzene at 323 K	0.87	0.93	5.36	0.41
128 Methyl-Ethyl-Ketone-Benzene at 323 K	0.84	1.05	2.64	1.24
129 Methyl-Isobutyl-Ketone-Benzene at 323 K	0.88	1.65	3.67	0.42
Average		**2.59**	**3.73**	**2.59**

Table 4.4 shows the correlation results for systems related to O-O atom interactions and they involve ethers-esters, ethers-ketones, alcohols-ketones, esters-aldehydes, alcohols-esters, alcohols-ethers, carboxylic acids-ketones, esters-ketones and aldehydes-ketones. The average correlation error for the UNIFAC-CI (VLE) model is 2.6 % while for the UNIFAC-CI (VLE) model with Q_{VLE}, the average AARD increases significantly to 4.3 %. For the UNIFAC-CI (VLE) model with Q_{VLE}, the problematic systems (> 10 %) are contributed by the tetrahydrofuran-ester (2 systems) and alcohol-ketone (2 systems) systems. For the reference UNIFAC model, the average prediction error obtained is 2.2 % which is a bit better than the CI-model without the Q_{VLE} factor.

All the results for the C-C, C-O, O-C and O-O atom interactions obtained for the UNIFAC-CI (VLE) model reported in Tables 4.2-4.4 are the ones that were obtained as a result of the analysis done by Mustaffa et al. [126] which has not been published yet. For the purpose of comparison with the UNIFAC-CI (VLE) model with Q_{VLE}, the results are presented here.

Table 4.4: Correlation Results for O-O Atom Interaction Related Systems.

No.	Systems	Q_{VLE}	AARD (%)		
			Original UNIFAC	UNIFAC-CI (VLE)	UNIFAC-CI (VLE) with Q_{VLE}
1	Tetrahydrofuran-Ethyl-Acetate at 313 K	0.92	0.27	4.29	11.34
2	Tetrahydrofuran-Ethyl-Acetate at 333 K	0.99	0.44	5.73	12.24
3	Dibutyl-Ether-2-Heptanone at 393 K	0.97	0.30	1.38	0.45
4	Dibutyl-Ether-3-Heptanone at 363 K	0.06	1.43	2.56	1.70
5	Dibutyl-Ether-4-Heptanone at 363 K	0.46	1.26	2.45	1.55
6	Ethanol-Acetone at 397 K	0.53	3.62	3.36	12.16
7	Ethyl-Acetate-2-Methylpropanal at 313 K	0.66	1.78	0.16	1.09
8	Ethanol-Methyl-Butyl-Ether at 338 K	1.00	2.56	7.10	8.18
9	Ethanol-Dipropyl-Ether at 308 K	0.97	2.45	8.60	9.47
10	Ethanol-Methyl-Proponate at 346 K	0.59	1.06	0.45	4.19
11	Ethyl-Acetate-2-Propanol at 333 K	0.77	2.42	4.51	2.88
12	Methyl-tert-Butyl-Ether-tert-Butanol at 339 K	0.14	5.86	1.53	1.67
13	Ethanol-Diisopropyl-Ether at 353 K	0.75	4.76	1.26	0.38
14	Acetone-Diisopropyl-Ether at 343 K	0.80	4.20	0.97	3.31
15	Acetone-Ethanol at 353 K	1.00	3.11	2.01	10.08
16	Diisopropyl-Ether-2-Methoxy-Ethanol at 341 K	0.72	8.97	5.34	7.64
17	1,4-Dioxane-2-Methyl-1-Propanol at 353 K	0.65	3.68	0.41	3.54
18	1,4-Dioxane-2-Methyl-1-Propanol at 373 K	0.65	4.74	1.73	1.77
19	Acetone-Acetic-Acid at 303 K	0.12	1.02	1.51	1.35
20	Ethanol-Methyl-Ethyl-Ketone at 298 K	0.95	0.57	4.29	3.98
21	Ethanol-Di-n-Butyl-Ether at 323 K	0.83	6.17	2.74	7.05
22	Acetone-Tetrahydrofuran at 302 K	1.00	3.21	0.98	1.01

23	Acetone-Methyl-Acetate at 328 K	0.84	0.45	1.93	0.28
24	Methyl-Propionate-1-Propanol at 328 K	0.72	0.49	0.72	4.94
25	Methyl-Acetate-1-Butanal at 313 K	0.78	0.34	4.87	3.05
26	1-Propanal-Ethyl-Acetate at 303 K	0.75	1.23	1.24	2.43
27	1-Propanal-Methyl-Acetate at 303 K	0.46	0.38	2.54	0.53
28	1-Butanal-n-Propyl-Acetate at 333 K	0.52	0.20	0.61	0.31
29	1-Propanal-Methyl-Ethyl-Ketone at 318 K	0.35	0.35	1.84	9.29
30	1-Propanol-Methyl-Propionate at 328 K	0.81	0.43	0.72	4.93
31	1-Propanol-Methyl-N-Butyrate at 333 K	1.00	1.20	0.38	3.42
32	Acetone-Diethyl-Ether at 303 K	0.76	0.32	4.92	2.33
	Average		**2.16**	**2.60**	**4.33**

Now we move to the nitrogenated systems which involve C, H, O and N atoms. In Table 4.5, the correlation errors of the systems related to C-N, N-C and the previous atoms interactions are presented which involve alkyl-pyridines-aromatics, alkyl-pyridines-alkanes, nitriles-aromatics and amines-alkanes systems. The results show that the average error for the UNIFAC-CI (VLE) model is 2.3 % which is more or less the same as the UNIFAC-CI (VLE) model with Q_{VLE} with 2.4 %. The average AARD obtained for the CI-models are better compared to the reference model with 3.2 %.

For the tri-n-butylamine-n-hexane system, since the Q_{VLE} factor obtained for this dataset in less than 0.05, for the UNIFAC-CI (VLE) model with Q_{VLE}, it is removed from the list of experimental for parameter regression.

Table 4.5: Correlation Results for C-N, N-C Atom Interactions Related Systems.

No.	Systems	Q_{VLE}	AARD (%)		
			Original UNIFAC	UNIFAC-CI (VLE)	UNIFAC-CI (VLE) with Q_{VLE}
1	4-Methylpyridine-Benzene at 313 K	0.97	1.16	1.09	1.05
2	4-Methylpyridine-Toluene at 313 K	0.81	1.11	0.58	0.57
3	3-Methylpyridine-Benzene at 313 K	0.73	0.69	1.20	1.15
4	3-Methylpyridine-Toluene at 313 K	0.94	1.10	0.37	0.35
5	2-Methylpyridine-Benzene at 313 K	0.80	0.33	1.40	1.36
6	2-Methylpyridine-Toluene at 313 K	0.90	0.70	0.81	0.77
7	2-Methylpiridine-n-Octane at 313 K	1.00	4.22	3.02	2.97
8	2-Methylpyridine-n-Nonane at 313 K	1.00	4.34	3.58	3.46
9	2-Methylpyridine-n-Heptane at 313 K	1.00	4.39	4.10	4.02
10	Toluene-Acetonitrile at 293k	0.85	0.54	0.61	0.79
11	Toluene-Acetonitrile at 343k	1.00	1.82	0.52	0.75
12	Toluene-Acetonitrile at 393k	0.79	4.35	0.53	0.70

13	Ethylbenzene-Propionitrile at 313k	0.59	1.47	1.28	1.45
14	Ethylbenzene-Propionitrile at 353k	0.88	1.07	1.00	1.17
15	Ethylbenzene-Propionitrile at 393k	0.71	0.73	0.92	1.10
16	n-Hexane-Triethylamine at 298 K	0.50	5.21	1.83	2.20
17	Triethylamine-n-Octane at 298 K	0.26	4.32	1.17	1.50
18	Tri-n-Butylamine-n-Hexane at 298 K	0.03	2.39	1.26	N/A
19	Dimethylamine-n-Hexane at 258 K	0.09	11.94	6.32	8.52
20	Dimethylamine-n -Hexane at 268 K	0.14	9.44	4.41	6.16
21	Dimethylamine-n -Hexane at 278 K	0.43	7.07	2.53	4.13
22	Dimethylamine-n -Hexane at 288 K	0.28	4.69	1.83	2.21
23	Dimethylamine-n -Hexane at 298 K	0.50	2.92	3.00	1.85
24	Dimethylamine-n -Hexane at 308 K	0.50	1.64	5.29	3.27
25	Dimethylamine-n -Hexane at 318 K	0.50	2.59	7.65	5.68
	Average		**3.21**	**2.25**	**2.38**

Next, the regression errors for O-N, N-O related systems (alcohols-nitriles, ketones-anilines, esters-anilines, amines-ketones, amines-esters) are presented in Table 4.6 with the average correlations errors of 5.3, 2.3 and 2.4 % respectively for the Original UNIFAC, UNIFAC-CI (VLE) and UNIFAC-CI (VLE) with Q_{VLE} respectively. Again, the performance of the CI-models are similar and better compared to the reference model. For the UNIFAC-CI model with Q_{VLE}, the acetone-ethanol system at 277 K was removed due to its low Q_{VLE} value. For the reference model, high correlation errors (> 10%) were found for the nitrile-alcohol systems.

Table 4.6: Correlation Results for O-N, N-O Atom Interactions Related Systems.

No.	Systems	Q_{VLE}	AARD (%)		
			Original UNIFAC	UNIFAC-CI (VLE)	UNIFAC-CI (VLE) with Q_{VLE}
1	n-Butyronitrile-2-Butanol at 278 K	0.24	17.48	6.30	7.38
2	n-Butyronitrile-2-Butanol at 288 K	0.40	14.28	3.27	4.37
3	n-Butyronitrile-2-Butanol at 293 K	0.79	12.59	1.65	2.76
4	n-Butyronitrile-2-Butanol at 298 K	0.81	11.51	0.80	1.83
5	n-Butyronitrile-2-Butanol at 303 K	1.00	10.53	0.53	0.99
6	n-Butyronitrile-2-Butanol at 308 K	1.00	9.79	0.89	0.54
7	Acetonitrile-Ethanol at 293 K	0.50	9.11	2.73	2.20
8	Acetonitrile-Ethanol at 343 K	0.50	4.34	6.58	5.88
9	Acetonitrile-Ethanol at 393 K	0.50	0.65	9.93	9.12
10	Acetone-Aniline at 277 K	0.04	2.33	2.66	N/A
11	Acetone-Aniline at 313 K	0.35	3.60	2.56	2.39

12	Acetone-Aniline at 386 K	0.50	4.00	2.44	2.65
13	Ethyl-Acetate-Aniline at 297 K	0.24	1.11	0.80	2.91
14	Aniline-Ethyl-Acetate at 348 K	0.12	1.62	0.26	0.74
15	Aniline-Ethyl-Acetate at 397 K	0.50	3.48	0.68	2.41
16	Diethylamine-Acetone at 298 K	0.36	0.74	1.42	0.26
17	Diethylamine-Acetone at 347 K	0.50	1.46	1.11	0.24
18	Diethylamine-Acetone at 398 K	0.07	1.59	1.10	0.20
19	Diethylamine-Ehtyl-Acetate at 297 K	0.33	0.26	0.99	1.08
20	Diethylamine-Ethyl-Acetate at 348 K	0.50	0.63	0.22	0.08
21	Diethylamine-Ethyl-Acetate at 398 K	0.29	1.10	0.86	0.65
Average			**5.34**	**2.27**	**2.43**

In addition, for the regression of N-N AIPs related datasets involving 12 systems of nitro-alkanes-nitriles, pyridines-amines, nitriles-anilines, amines-nitriles and amines-amines systems, the correlation results are listed in Table 4.7. The average errors show that the performance of the CI-models is as good as the reference UNIFAC model. When Q_{VLE} factors are added, the average errors for the CI-model slightly increased from 1.7 to 1.0 %. However, the diethylamine-pyridine system was removed due to the fact that its Q_{VLE} value is 0.03 which is less that the allowable limit of 0.05. A dataset which have a Q_{VLE} value of less than 0.05 is considered as problematic.

Table 4.7: Correlation Results for N-N Atom Interaction Related Systems.

No.	Systems	Q_{VLE}	AARD (%)		
			Original UNIFAC	**UNIFAC-CI (VLE)**	**UNIFAC-CI (VLE) with Q_{VLE}**
1	Nitromethane-Acetonitrile at 298 K	0.80	0.20	1.58	1.28
2	Nitromethane-Acetonitrile at 348 K	0.81	0.19	0.34	0.42
3	Nitromethane-Acetonitrile at 398 K	0.57	0.27	1.24	1.22
4	Diethylamine-Pyridine at 363.15 K	0.03	1.76	4.26	N/A
5	Acetonitrile-Aniline at 293 K	0.28	1.37	0.82	1.04
6	Acetonitrile-Aniline at 343 K	0.50	3.36	1.34	0.85
7	Acetonitrile-Aniline at 393 K	0.38	4.95	2.12	1.88
8	Diethylamine-Acetonitrile at 298 K	0.33	0.40	2.39	1.22
9	Diethylamine-Acetonitrile at 347 K	0.50	1.33	1.95	0.50
10	Diethylamine-Acetonitrile at 398 K	0.28	2.18	2.37	1.34
11	n-Propylamine-Diethylamine at 297 K	0.32	0.40	1.62	0.73
12	n-Propylamine-Diethylamine at 347 K	0.50	0.38	0.48	0.89
Average			**1.40**	**1.71**	**1.03**

The chlorinated systems are shown in Table 4.8 for systems with C-Cl, Cl-C interactions (alkyl-chloride-aromatics, alkyl-chlorides-alkanes, chloro-alkanes-alkanes, chloro-aromatics-aromatics, carbon tetrachlorides-alkanes, chloroalkanes-aromatics and carbon tetrachlorides-aromatics). The average AARD value for Original UNIFAC is 2.7 % while the correlation errors increases to 3.5 % for both Cl models. The problematic systems for the Cl-models include the alkane-chloroalkane (2 systems) systems which are also the same for the reference model and alkene-chloroalkane (1 system) system.

Table 4.8: Correlation Results for C-Cl, Cl-C Atom Interactions Related Systems.

No.	Systems	Q_{VLE}	AARD (%)		
			Original UNIFAC	UNIFAC-Cl (VLE)	UNIFAC-Cl (VLE) with Q_{VLE}
1	sec-Butyl-Chloride-Toluene at 323 K	0.51	0.71	1.89	1.76
2	sec-Butyl-Chloride-n-Heptane at 323 K	0.82	4.50	3.30	3.67
3	sec-Butyl-Chloride-n-Heptane at 333 K	0.81	4.32	3.17	3.53
4	1,2-Dichloroethane-n-Heptane at 343 K	0.95	1.03	4.56	2.03
5	n-Butyl-Chloride-Toluene at 323 K	0.79	0.33	2.61	2.35
6	n-Butyl-Chloride-n-Heptane at 323 K	1.00	2.94	1.20	1.76
7	tert-Butyl-Chloride-n-Heptane at 323 K	0.70	4.57	3.81	4.04
8	n-Butyl-Chloride-n-Heptane at 323 K	0.82	2.94	1.20	1.76
9	n-Butyl-Chloride-Toluene at 323 K	0.57	0.33	2.61	2.35
10	sec-Butyl-Chloride-Toluene at 323 K	0.51	0.71	1.89	1.76
11	n-Butyl-Chloride-Toluene at 298 K	0.87	1.06	1.84	1.65
12	n-Butyl-Chloride-Toluene at 348 K	0.68	0.47	2.78	2.48
13	n-Butyl-Chloride-Toluene at 398 K	0.55	0.17	3.90	3.55
14	Benzene-Monochlorobenzene at 298 K	0.82	0.54	3.92	3.92
15	Benzene-Monochlorobenzene at 348 K	0.59	0.88	3.15	3.14
16	Benzene-Monochlorobenzene at 398 K	0.44	1.38	2.35	2.34
17	n-Pentane-Dichloromethane at 298 K	1.00	8.83	6.01	6.26
18	n-Pentane-Dichloromethane at 348 K	1.00	5.66	3.17	3.38
19	n-Pentane-Dichloromethane at 398 K	0.93	3.47	1.24	1.43
20	1,2-Dichloroethane-n-Heptane at 340 K	0.93	1.03	4.56	2.03
21	n-Heptane-Dichloromethane at 298 K	0.77	8.19	4.55	4.86
22	n-Heptane-1,2-Dichloroethane at 298 K	0.55	2.73	5.04	1.85
23	n-Heptane-1,3-Dichloropropane at 298 K	0.17	15.96	7.57	10.83
24	n-Heptane-1,4-Dichlorobutane at 298 K	0.13	15.97	9.47	11.75
25	Carbon-Tetrachloride-n-Hexane at 298 K	1.00	2.35	2.76	2.09
26	Carbon-Tetrachloride-n-Heptane at 298 K	0.95	2.76	3.63	2.83

27	Chloroform-n-Hexane at 298 K	1.00	1.03	2.32	2.96
28	Chloroform-n-Heptane at 298 K	1.00	0.70	1.88	2.60
29	n-Heptane-n-Butyl-Chloride at 298 K	0.50	3.79	1.97	2.54
30	n-Pentane-n-Butyl-Chloride at 318 K	0.50	3.10	1.60	2.31
31	n-Pentane-n-Butyl-Chloride at 338 K	0.50	2.84	1.49	2.18
32	Dichloromethane-Benzene at 298 K	0.50	2.91	3.27	1.52
33	Dichloromethane-Benzene at 348 K	0.50	1.46	5.82	4.06
34	Dichloromethane-Toluene at 298 K	0.25	1.83	7.33	8.02
35	Dichloromethane-Toluene at 347 K	0.50	0.51	2.05	2.21
36	Tetrachloroethane-N-Hexane at 298 K	0.50	3.01	3.34	2.82
37	Tetrachloroethane-N-Hexane at 308 K	0.50	1.74	4.45	3.91
38	1,2-Dichloroethane-N-Hexane at 298 K	0.50	5.34	1.45	1.60
39	p-Xylene-Carbon-Tetrachloride at 313 K	0.50	4.29	3.01	2.31
40	p-Xylene-Carbon-Tetrachloride at 323 K	0.45	3.30	1.99	1.75
41	Dichloromethane-1-Hexene at 298 K	0.50	4.28	13.56	13.25
42	Monochlorobenzene-Ethylbenzene at 293 K	0.50	0.21	1.01	1.12
43	1-Pentene-Monochlorobenzene at 320 K	0.50	2.14	8.81	8.11
44	1-Pentene-Monochlorobenzene at 360 K	0.50	4.22	3.05	2.42
45	o-Xylene-Carbon-Tetrachloride at 303 K	0.32	0.39	2.19	2.83
46	p-Xylene-Carbon_Tetrachloride at 303 K	0.20	0.72	2.46	2.84
47	m-Xylene-Carbon_Tetrachloride at 303 K	0.37	0.90	2.63	2.71
48	Cumene-Carbon-Tetrachloride at 303 K	0.28	1.33	6.20	6.33
49	n-Heptane- Carbon-Tetrachloride at 323 K	0.48	1.30	3.56	2.94
50	n-Butyl-Chloride-n-Heptane at 323 K	0.29	2.97	1.16	1.73
51	Carbon-Tetrachloride-Benzene at 260 K	0.08	1.01	3.41	4.38
52	Carbon-Tetrachloride-Benzene at 265 K	0.33	0.88	2.92	3.89
53	Carbon-Tetrachloride-Benzene at 266 K	0.33	0.84	2.76	3.73
54	Carbon-Tetrachloride-Benzene at 271 K	0.33	0.76	2.35	3.31
55	Carbon-Tetrachloride-Benzene at 273 K	0.33	0.71	2.17	3.08
56	Carbon-Tetrachloride-Toluene at 273 K	0.50	1.30	4.67	4.20
	Average		**2.74**	**3.52**	**3.52**

In addition, the correlation results for systems related to O-Cl, Cl-O interactions are presented in Table 4.9. The systems involved in this part are related to chloro-alkanes-ethers, esters-chloroform, ethers-chloroalkanes, alcohols-chloroforms, ketones-chloroforms and ketones-chloro-aromatics. The average correlation results for the CI-models are better than the reference model with 3.0 and 2.4 % respectively for UNIFAC-CI (VLE) and UNIFAC-CI (VLE) with Q_{VLE}, compared to 5.9 % for Original UNIFAC. For the CI-models, no dataset obtained AARD more than 10 %. However, for the reference UNIFAC model, problematic systems can be found for the chloroalkane-ether (6 systems) systems which are well correlated by the CI-models.

Table 4.9: Correlation Results for O-Cl, Cl-O Atom Interactions Related Systems.

No.	Systems	Q_{VLE}	AARD (%)		
			Original UNIFAC	UNIFAC-Cl (VLE)	UNIFAC-Cl (VLE) with Q_{VLE}
1	Dichloroethane-Di-n-Butyl-Ether at 330 K	0.81	21.22	2.15	3.47
2	Dichloroethane-Di-n-Butyl-Ether at 350 K	0.80	10.90	1.58	1.97
3	Dichloroethane-Di-n-Butyl-Ether at 370 K	1.00	16.19	2.88	3.57
4	Trichloroethane-Di-n-Butyl-Ether at 323 K	0.77	7.77	1.08	2.06
5	Ethanol-Monochlorobenzene at 298 K	0.82	2.74	0.94	1.39
6	Ethanol-Monochlorobenzene at 398 K	0.64	3.02	6.34	4.26
7	Dichloromethane-Ethyl-Acetate at 298 K	1.00	1.81	2.36	2.61
8	Dichloromethane-Ethyl-Acetate at 348 K	0.74	0.71	0.65	1.86
9	Dichloromethane-Ethyl-Acetate at 398 K	0.55	0.87	2.36	1.88
10	1-Chloropentane-Di-n-Butyl-Ether at 313 K	0.63	4.38	5.06	3.83
11	1-Chloropentane-Di-n-Butyl-Ether at 323 K	0.86	3.54	4.21	3.15
12	1,2-Dichloroethane-Di-n-Butyl-Ether at 330 K	0.67	21.22	2.15	3.47
13	Dichloroethane-Di-n-Butyl-Ether at 350 K	0.80	10.90	1.58	1.97
14	Dichloroethane-Di-n-Butyl-Ether at 370 K	1.00	16.19	2.88	3.57
15	Trichloroethane-Di-n-Butyl-Ether at 323 K	1.00	7.77	1.08	2.06
16	Trichloroethane-Di-n-Butyl-Ether at 343 K	0.66	6.90	1.59	2.10
17	Ethanol-Chloroform at 303 K	0.98	2.49	1.62	1.65
18	Ethanol-Chloroform at 313 K	1.00	2.04	1.20	1.20
19	Ethanol-Chloroform at 323 K	0.86	1.88	1.21	1.13
20	Acetone-Monochlorobenzene at 313 K	0.49	3.76	3.55	1.07
21	Acetone-Monochlorobenzene at 353 K	0.49	2.52	5.33	0.43
22	Acetone-Monochlorobenzene at 386 K	0.49	2.32	6.56	0.52
23	Chloroform-Acetone at 313 K	0.82	0.60	5.56	3.99
24	Chloroform-Acetone at 323 K	0.82	0.39	4.50	3.11
25	Chloroform-Acetone at 303 K	0.88	0.65	5.62	3.99
26	1-Butanol-Carbon-Tetrachloride at 323 K	0.21	1.16	4.92	2.82
	Average		**5.92**	**3.04**	**2.43**

For the sulfurated systems, the correlation results are presented in Tables 4.10 for C-S and S-C, O-S and S-O and Cl-S and S-Cl interactions related systems. In total, 13 datasets were used for regression and involve systems with diethyl sulfides-cycloalkanes, diethyl sulfides-ethoxy-alkanes, dimethyl-sulfoxides-ketones, dimethyl-sulfoxides-esters and others. The average AARD for the Cl-models, UNIFAC-Cl (VLE) and UNIFAC-Cl (VLE) with Q_{VLE} are slightly better than Original UNIFAC with 0.6 and 0.7 % compared to 1.0 % which is very good.

Table 4.10: Correlation Results for C-S, S-C, O-S, S-O Cl-S, S-Cl Atom Interaction Related Systems.

No.	Systems	Q_{VLE}	AARD (%)		
			Original UNIFAC	UNIFAC-CI (VLE)	UNIFAC-CI (VLE) with Q_{VLE}
1	Diethyl-Sulfide-Cyclohexane at 343 K	1.00	0.55	0.24	0.12
2	Diethyl-Sulfide-Cyclohexane at 353 K	1.00	0.24	0.18	0.27
3	Dimethyl-Sulfoxide-Acetone at 298 K	0.37	0.79	0.48	0.44
4	Dimethyl-Sulfoxide-Acetone at 308 K	0.33	0.16	0.18	0.18
5	Dimethyl-Sulfoxide-Acetone at 318 K	0.52	0.70	0.41	0.41
6	Dimethyl-Sulfoxide-Ethyl-Acetate at 298 K	0.39	0.74	0.48	0.79
7	Dimethyl-Sulfoxide-Ethyl-Acetate at 308 K	0.43	0.75	0.50	0.68
8	Dimethyl-Sulfoxide-Ethyl-Acetate at 318 K	0.50	0.80	0.54	0.77
9	Diethyl-Sulfide-2-Ethoxy-2-Methylpropane at 333 K	1.00	1.55	0.07	0.10
10	Diethyl-Sulfide-2-Ethoxy-2-Methylpropane at 343 K	0.97	1.62	0.07	0.07
11	Chloroform-Dimethyl-Sulfide at 298 K	0.89	1.58	2.42	2.54
12	Chloroform-Diethyl-Sulfide at 298 K	1.00	3.65	2.27	2.13
13	Carbon-Tetrachloride-Diethyl-Sulfide at 298 K	0.88	0.33	0.08	0.10
	Average		**1.03**	**0.61**	**0.66**

4.2.5 Overall Correlation Results

The correlation results reported in Tables 4.2-4.10 are summarized in Table 4.11 and Figure 4.2 according to the types of systems. For all the hydrocarbons and oxygenated systems the performance of the CI-models is more or less as good as the reference model which obtained an AARD of 2.2 %. When the Q_{VLE} factors were used, the average AARD value for the CI-model reduced slightly from 2.9 to 2.5 %. In addition, for the nitrogenated system, the correlation errors obtained for the CI-models are more or less the same with 2.1 % which is better than the reference model (3.6 %). When the Q_{VLE} factors are added, the average correlation error for the CI-model is slightly reduced from 3.4 to 3.2 % for the chlorinated system. Overall we can see that the performance of the CI-models (with 2.85 and 2.55 % AARDs) are as good as the reference UNIFAC model (with 2.7 % AARD). When we add the Q_{VLE} quality factors for the CI-model, only a slight decrease of the average correlation error can be seen (from 2.9 to 2.6 % suggesting that the inclusion of the Q_{VLE} factor in the regression of UNIFAC parameters can provide only a small contribution in reducing the correlation errors. Furthermore, by generating these quality factors, anomalous or problematic experimental data can be identified and corrected, or removed.

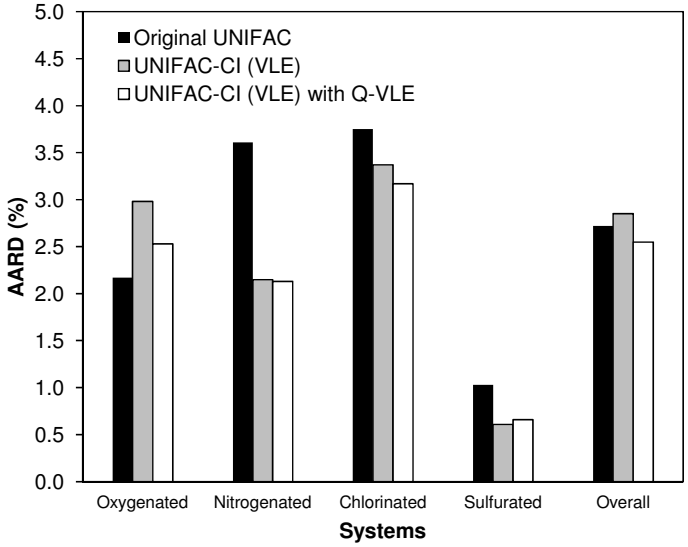


Figure 4.2: Correlation Errors in AARD (%) for CI-Models Compared with the Reference UNIFAC Model.

Table 4.11: Overall Correlation Results in AARD (%).

Types of systems	AARD (%)		
	Original UNIFAC	**UNIFAC-CI (VLE)**	**UNIFAC-CI (VLE) with Q_{VLE}**
C-C related systems	1.06	1.28	1.20
C-O, O-C related systems	2.59	3.73	2.59
O-O related systems	2.16	2.60	4.33
Overall hydrocarbon & oxygenated systems	**2.17**	**2.98**	**2.53**
C-N, N-C related systems	3.21	2.25	2.38
O-N, N-O related systems	5.34	2.27	2.43
N-N related systems	1.40	1.71	1.03
Overall nitrogenated systems	**3.61**	**2.15**	**2.13**
C-Cl, Cl-C related systems	2.74	3.52	3.52
O-Cl, Cl-O related systems	5.92	3.04	2.43
Overall chlorinated systems	**3.75**	**3.37**	**3.17**
C-S, S-C, O-S, S-O, Cl-S, S-Cl related systems	1.03	0.61	0.66
Overall sulfurated systems	**1.03**	**0.61**	**0.66**
All systems	**2.72**	**2.85**	**2.55**

4.3 Predictions of VLE Data using the Regressed Parameters

In this section, the predictions of VLE systems using the regressed interaction parameters of the CI-models are presented and compared with the predictions made by the reference UNIFAC model. The predictions will demonstrate the performance of the regressed parameters and their reliability in predicting systems which are not used in the parameter regression but with the same types of systems.

4.3.1 Oxygenated Systems

The predictions of VLE systems with C, O and H atoms are presented in Figure 4.3 with 3 isothermal and 1 isobaric systems. The group interaction parameters (GIPs) involved in the prediction of these systems include for the interactions between groups CH_2-C=C, C=C-OH, CH_2-ACH, CH_2-COOH, ACH-COOH, ACH-$ACCH_2$, CH_2-$ACCH_2$, OH-$ACCH_2$, CH_2-CH_2O, OH-CH_2O which are also used to represent some of the systems used in the parameter regression. For the isothermal system of trans-2-butene-1-propanol at 364.50 K as displayed in Figure 4.3 (a), the predictions by Original UNIFAC are in good agreement with the experimental data especially at the pressure-liquid composition (*P-x*) line. However, the data are slightly underpredicted by both CI-models which have the same prediction accuracy. In Figure 4.3 (b) for the isobaric system of acetic-acid-ethylbenzene at 725 mmHg, the azeotrope is best captured by UNIFAC-CI (VLE) with Q_{VLE}, followed by Original UNIFAC and UNIFAC-CI (VLE). For the remaining experimental data, the temperature-liquid composition (*T-x*) data are in good agreement with the predictions by Original UNIFAC and UNIFAC-CI (VLE) at lower concentrations but deviate at the higher concentration region which is better predicted by UNIFAC-CI (VLE) with Q_{VLE}. On the other hand, the predictions by Original UNIFAC are close to the temperature-vapor composition (*T-y*) data, followed by the CI-models.

Furthermore, for the isothermal system of toluene-1-butanol at 363.15 K in Figure 4.3 (c), the trends of the predictions by the UNIFAC models are the same as the experimental data but they are not in exact agreement with the experimental data. The azeotrope is closely captured by Original UNIFAC followed by the CI-models which have the same prediction accuracy. The predictions made by Original UNIFAC and the CI-models are almost the same at lower concentrations but started to deviate when moving towards the azeotrope. In Figure 4.3 (d), it is clearly seen that the system is best predicted by Original UNIFAC. The predictions made by the CI-models are not very good which are expected due to the fact that the correlation errors for this type of system (alcohol-ether) in the parameter regression step is quite high.

Between the two CI-models, the predictions made by UNIFAC-CI (VLE) are better at higher concentrations but when approaching the azeotrope at lower concentration, UNIFAC-CI (VLE) with Q_{VLE} performs better.

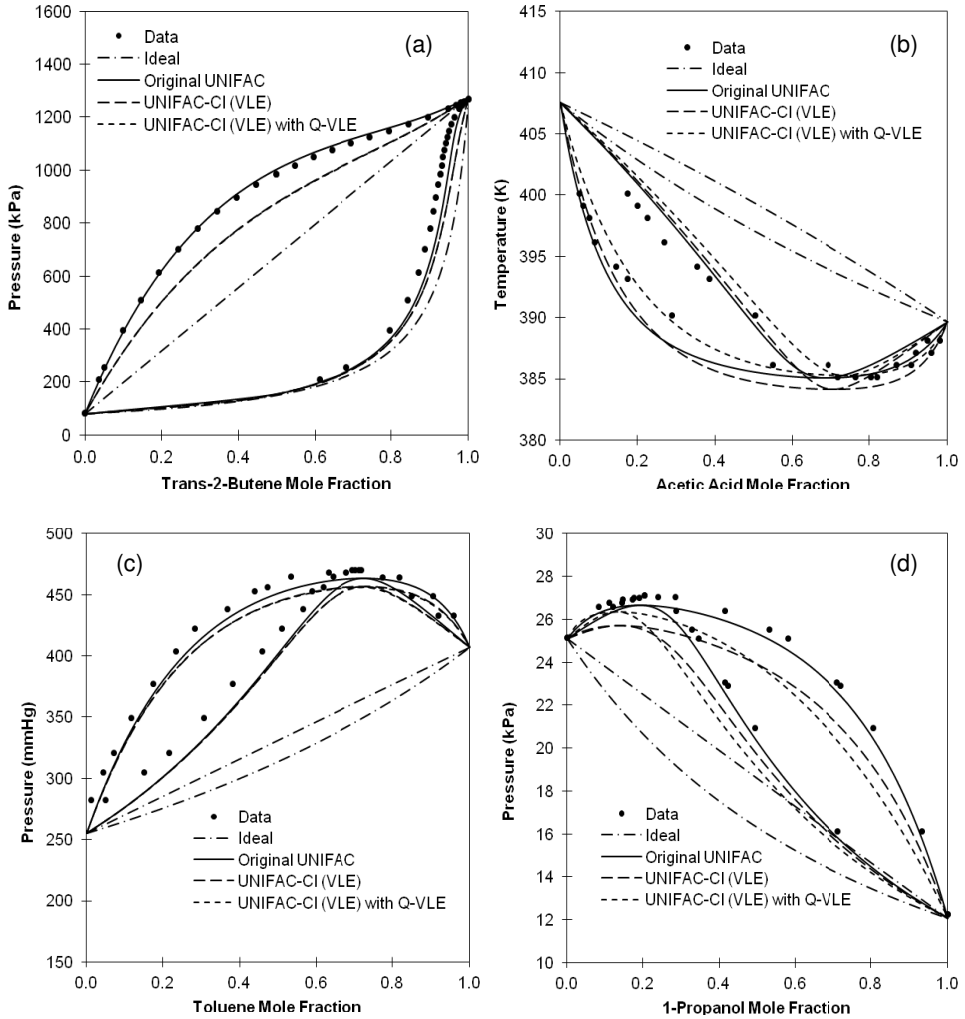


Figure 4.3: VLE Diagrams Predicted using Original UNIFAC, UNIFAC-CI (VLE) and UNIFAC-CI (VLE) with Q_{VLE} for Systems of: (a) Trans-2-Butene-1-Propanol at 364.50 K, (b) Acetic Acid Ethylbenzene at 725 mmHg, (c) Toluene-1-Butanol at 363.15 K and (d) 1-Propanol-Di-n-Butyl Ether at 323.15 K.

4.3.2 Nitrogenated Systems

Figure 4.4 and 4.5 illustrate four VLE phase diagrams involving 3 isobaric and 1 isothermal systems for nitrogenated system (systems containing C, H, O and N atoms). The GIPs involved are for the interactions between groups CH_2-$ACNH_2$, CH_2-CCN, OH-CCN and CCN-CNO_2. Figure 4.4 (a) shows the phase diagram for the system of n-hexane-aniline at constant pressure of 760 mmHg. From the figure, it can be seen that for Original UNIFAC, the predictions are in good

agreement with the experimental data except for some data points. The accuracy of the predictions of the CI-models for the temperature-vapor composition (*T-y*) line is almost the same as Original UNIFAC. However, the temperature-liquid composition (*T-x*) data are slightly underpredicted by the CI-models and can be clearly seen between n-hexane mole fraction of 0.1 to 0.8. Overall, the system can be adequately be predicted using the CI-models. In addition, Figure 4.4 (b) present the isobaric VLE phase diagram of acetonitrile-1-propanol at 66.6 kPa. The predictions by Original UNIFAC for this system is not very good. However, for the CI-models the predictions are closer to the experimental data especially at higher concentrations.

The same type of system (alcohol-nitrile) is also presented in Figure 4.5 (a) involving the system of acetonitrile-2-methyl-1-propanol. However for this system, the experimental data are best predicted by Original UNIFAC compared to the CI-models eventhough the predictions are not in good agreement with the data. For both CI-models, the predictions underpredict the data expecially at the temperature-liquid composition (*T-x*) line. In Table 4.6, the correlation results for most of the nitrile-alcohol systems are found to be quite high for Original UNIFAC and the CI-models and therefore, it is expected that the predictions involving this type of system are not very good as demonstrated by Original UNIFAC in Figure 4.4 (b) and the CI-models in Figure 4.5 (a). Moreover, in Figure 4.5 (b) the VLE diagram of acetonitrile-nitromethane at constant temperature of 333.15 K is presented. The predictions of Original UNIFAC and both the CI-models are in good agreement with the experimental data. However, if the diagram is closely examined, the predictions by the CI-models are more accurate compared to the Original UNIFAC model. The correlation errors presented in Table 4.7 involving this type of system (nitrile-nitroalkane) are also very low.

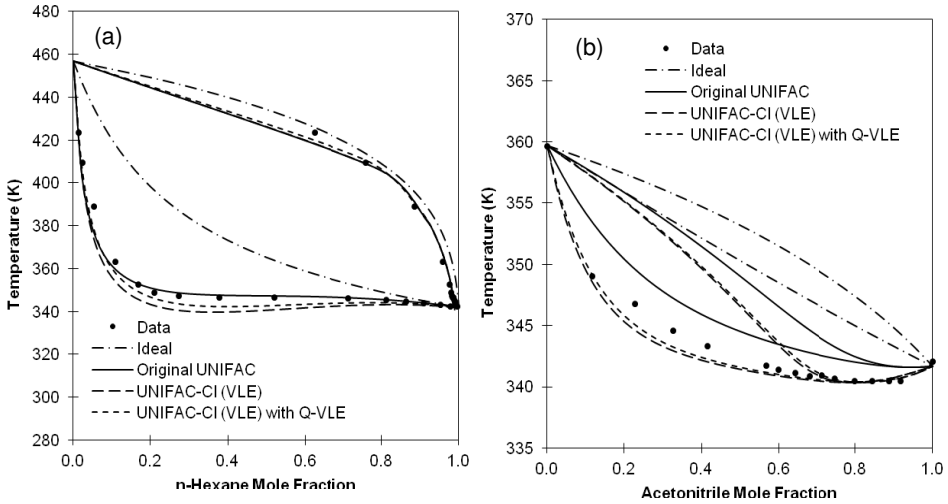


Figure 4.4: VLE Diagrams Predicted using Original UNIFAC, UNIFAC-CI (VLE) and UNIFAC-CI (VLE) with Q_{VLE} for Systems of: (a) n-Hexane-Aniline at 760 mmHg, (b) Acetonitrile-1-Propanol at 66.6 kPa.

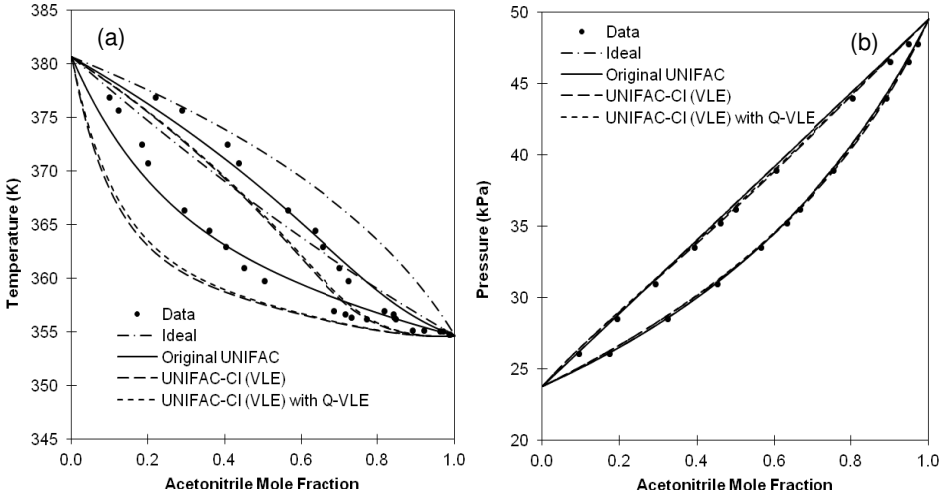


Figure 4.5: VLE Diagrams Predicted using Original UNIFAC, UNIFAC-CI (VLE) and UNIFAC-CI (VLE) with Q_{VLE} for Systems of: (a) Acetonitrile-2-Methyl-1-Propanol at 760 mmHg and (b) Acetonitrile-Nitromethane at 333.15 K.

4.3.3 Chlorinated Systems

The VLE phase diagrams for chlorinated systems (systems containing C, H, O, N and Cl atoms) involving 3 isothermal and 1 isobaric systems are illustrated in Figure 4.6. The group interactions involved include ACH-ACCl, CH_2COO-ACCl, C=C-ACCl, CH_2-CCl_3, OH-CCl_3. In Figure 4.6 (a), the predictions made by Original UNIFAC and UNIFAC-CI (VLE) with Q_{VLE} for the systems of ethyl-acetate-monochlorobenzene at 393.15 K are in close agreement with the experimental data. However, the UNIFAC-CI (VLE) model overpredicted the data. This type of system (ester-chloroaromatic) is not used in the parameter regression step but can still be predicted by Original UNIFAC and UNIFAC-CI (VLE) with Q_{VLE} with acceptable accuracy. However, systems with esters and chloroaromatic compounds have been used in the parameter estimation but with different combinations (ester-chloroalkane, alcohol-chloroaromatic, ketone-chloroaromatic). Furthermore, Figure 4.6 (b) shows the VLE diagram for 1-pentene-monochlorobenzene at 280 K. In the parameter regression step, the same type of system has been used but at temperatures 320 and 360 K. From the figure, it can be seen that the system is well predicted by Original UNIFAC.

However, it is underpredicted by the CI-models. At the pressure-vapor composition line, the accuracy of predictions for Original UNIFAC and the CI-models is the same. In Table 4.8, the correlation errors for 1-pentene-monochlorobenene at 320 K for the CI-models are also quite high (> 8 %).

Figure 4.6 (c) shows the VLE phase diagrams for the system of chloroform-1-butanol at atmospheric pressure (760 mmHg). The temperature-vapor composition (*T-y*) data are well predicted by Original UNIFAC and both the CI-models. However all the UNIFAC models were unable to capture the temperature-liquid composition (*T-x*) data especially at lower concentrations.

Furthermore, the VLE phase diagram for the isothermal system of chloroform-n-hexane at 328.15 K is presented in Figure 4.6 (d). The experimental data are in a close agreement with the prediction made by Original UNIFAC. Both the UNIFAC-CI (VLE) and UNIFAC-CI (VLE) with Q_{VLE} models however, slightly underpredict the system but with still acceptable accuracy. The correlation errors for this type of system (chloroform-alkane) reported in Table 4.8 are also quite low.

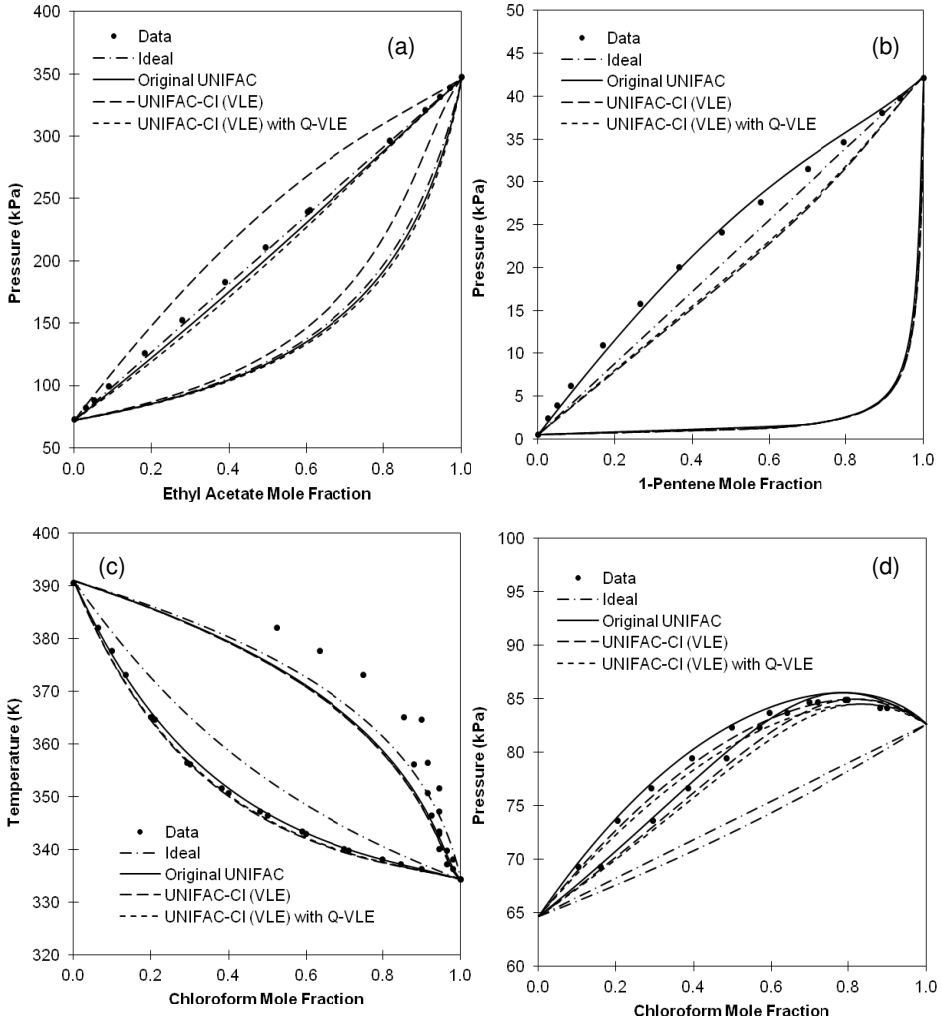


Figure 4.6: VLE Diagrams Predicted using Original UNIFAC, UNIFAC-CI (VLE) and UNIFAC-CI (VLE) with Q_{VLE} for Systems of: (a) Ethyl Acetate-Monochlorobenzene at 393.15 K, (b) 1-Pentene-Monochlorobenzene at 280 K, (c) Chloroform-1-Butanol at 760 mmHg and (d) Chloroform-n-Hexane at 328.15 K.

4.3.4 Sulfurated Systems

The VLE predictions of sulfurated systems (systems containing C, H, O and S atoms) using Original UNIFAC, UNIFAC-CI (VLE) and UNIFAC-CI (VLE) with Q_{VLE} for 2 isothermal and 2 isobaric systems are presented in Figure 4.7. The GIPs involved are for the interactions between groups CH_2-CH_2S, CH_2-DMSO and CH_2CO-DMSO. In Figure 4.7 (a) for the system of diethyl sulfide-n-heptane at 363.15 K, the predictions are not in a good agreement with the experimental data eventhough they have the same trends. However, the predictions made by the CI-models are better compared to the Original UNIFAC models as they are closer to the experimental data especially at higher concentrations. For the systems of diethyl sulfide-n-hexane at 338.15 K displayed in Figure 4.7 (b), the prediction accuracies of Original UNIFAC and both CI-models are almost the same and they are also very close (slightly unpredict) to the experimental data. The correlation results (AARD %) of diethyl sulfide-cycloalkane/alkane type of system in Table 4.10 are also quite low (< 1 %).

Figures 4.7 (c) and (d) illustrate the systems of dimethyl sulfoxide-methyl ethyl ketone and dimethyl sulfoxide-methyl isobutyl ketone at constant pressure of 95.3 kPa. According to González [18], for systems involving dimethyl sulfoxide, a separate parameter table is used for the prediction. However, since systems with dimethyl sulfoxide are also used for the prediction of parameters involving sulfur-contained functional groups (see Table 4.10), the prediction capability of those parameters to predict systems with dimethyl sulfoxide are also displayed here. In Figure 4.6 (c), only the temperature-liquid composition (*T-x*) data are available. All models underpredict the experimental with the UNIFAC-CI (VLE) with Q_{VLE} being the closest model to the data followed by UNIFAC-CI (VLE) and Original UNIFAC. On the other hand, in Figure 4.6 (d) at lower concentrations, the system is quite well predicted by Original UNIFAC but deviate at higher concentrations. The best predictions are made by UNIFAC-CI (VLE) which follows the data closely throughout the entire concentration range. The worst predictions are made by UNIFAC-CI (VLE) with Q_{VLE} which over predict the experimental data. However, it is found that in Table 4.10 that the correlation results for all the UNIFAC models for dimethyl sulfoxide-ketone systems are very good.

4.4 Conclusions

In this chapter, the atom interaction parameters (AIPs) of the UNIFAC-CI model involving VLE systems have been regressed and the results of the correlations were compared with the prediction accuracy of the reference Original UNIFAC model. In the regression of AIPs, two approaches have been used. In the first approach, a data quality factor (Q_{VLE}) which has been obtained for each VLE system is used as weighting factor in the objective function for regressing the parameters. However, in the second approach these quality factors are set to unity. The result of the correlation and also the predictions using the regressed parameters between these two sets of AIPs of the CI-models have been compared. From the overall correlation result displayed in Table 4.11, the addition of the Q_{VLE} factor in the regression of the AIPs has slightly reduced the correlation errors. Furthermore, it is also useful in identifying problematic systems which can produce parameters (from parameter estimation) which are not accurately representing the systems used for the regression.

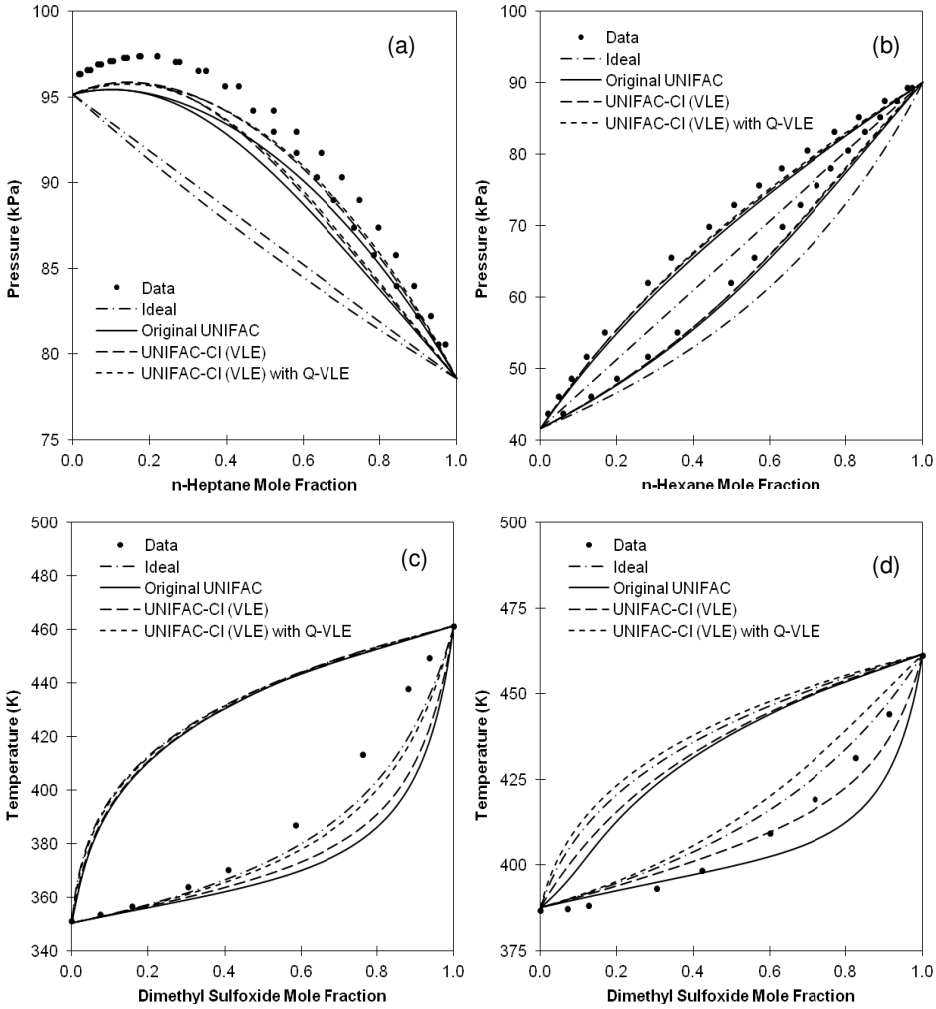


Figure 4.7: VLE Diagrams Predicted using Original UNIFAC, UNIFAC-CI (VLE) and UNIFAC-CI (VLE) with Q_{VLE} for Systems of: (a) Diethyl Sulfide-n-Heptane at 363.15 K (b) Diethyl Sulfide-n-Hexane at 338.15 K, (c) Dimethyl Sulfoxide-Methyl Ethyl Ketone at 95.3 kPa and (d) Dimethyl Sulfoxide-Methyl Isobutyl Ketone at 95.3 kPa.

Moreover, in section 4.3 the predictions of VLE systems using the regressed parameters have been analyzed. The systems used for predictions are of the same type as the systems used in the parameter estimation. It is found that the prediction accuracy of the CI-models and Original UNIFAC are more or less the same as the correlation errors obtained for the same types of system. In order to use the regressed parameters for systems or UNIFAC main groups which have not been used in the parameter regression step, good predictions are not guaranteed. However, it is the purpose of this work to only use a limited amount of data in the parameter estimation. Then, using the regressed parameters for the prediction of other systems which have not been used in

parameter regression but at least having the same UNIFAC main groups. In theory, possibly any UNIFAC group-interactions involving C, H, O, N, Cl and S atoms can be created and predicted. The AIPs obtained from this work are presented in Appendix B in Tables B.1 and B.2.

CHAPTER 5

FURTHER ANALYSIS AND DEVELOPMENT OF THE ORIGINAL UNIFAC-CI MODELS: ADDITION OF SOLID-LIQUID EQUILIBRIUM SYSTEMS

5.1 Introduction

In this chapter, we present the further development and analysis of the Original UNIFAC-CI model with the addition of solid-liquid equilibrium (SLE) experimental data to vapor-liquid equilibrium data (VLE) data for the regression of the atom interaction parameters (AIPs). In Chapter 4, the AIPs are regressed against only VLE data. In this chapter, a comparison has been made between the parameters regressed from VLE experimental data alone and the parameters regressed from VLE and SLE experimental data simultaneously. Both correlation results and predictions will be shown using the obtained parameters, in order to see the significance of adding SLE data in the parameter estimation and the effect to VLE. Besides predicting VLE, it is known that the UNIFAC model have also been used to predict SLE (and in principle other types of phase behavior). Therefore, it is important to investigate to what extent the Original UNIFAC-CI model can succesfully predict SLE especially when the needed parameters are missing.

5.2 Development of the Original UNIFAC-CI (VLE/SLE) Model

In this section, we present the parameter regression work for the development of the Original UNIFAC-CI (VLE/SLE) model. The AIPs of the CI-model are regressed simultaneously against VLE and SLE data. The overall correlation results of all systems used in the regression are presented and analyzed. The correlation results are compared with the prediction errors of the same systems when using the AIPs obtained from regression against only VLE data. In this work the comparison are made with the Original UNIFAC-CI (VLE) with the Q_{VLE} model developed in Chapter 4. In addition, several phase diagram generated with UNIFAC using the regressed parameters will also be presented and compared with the predictions made by the reference model (ie. Original UNIFAC).

5.2.1 Background of Parameter Estimation

As described in Section 2.6.3, the optimization technique used to regress the parameters in this work is the least squares technique using the Modified Levenberg-Marquardt approach with the algorithm described in [125] which has been slightly modified to regress AIPs for the Original UNIFAC-CI models. The same technique of parameter regression was used in the previous work [16-18].

In this work, the parameter regression is based on the *P, T, x_i, y_i* and *P, T, x_i*, VLE data and *T, x_i, $T_{m,i}$, $\Delta H_{fus,i}$* SLE data. Only isothermal data VLE data are used for the regression. Since the unit and magnitude of the pressures reported in each of the VLE data are different (mmHg, torr, kPa, Pa), the objective function below (Equation (5.1)) has been chosen to regress the parameters expressed as the average relative pressure quadratic deviation as representated by the first term in Equation (5.1). For the regression with SLE data, the second term in Equation (5.1) was chosen as the absolute deviation between the experimental and calculated liquid compositions.

The overall objective function (OF) used is as follows:

$$\mathrm{OF} = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{P_{i-\exp} - P_{i-calc}}{P_{i-\exp}}\right)^2 + \frac{1}{N}\sum_{i=1}^{N}\left(x_{i-\exp} - x_{i-calc}\right)^2 + \mathrm{w}_{reg}\sum_{i=1}^{N}(\mathrm{AIP}_i - \mathrm{AIP}_i^{IG})^2 \qquad (5.1)$$

where $x_{i\text{-}exp}$ is the experimental composition of the solute, $P_{i\text{-}exp}$ is the experimental pressure, N is the number of experimental data points used for the estimation, AIP_j is the current value of the CI-interaction parameter j, AIP^{IG}_j its corresponding initial guess and w_{reg} (values between $1\text{x}10^2$ and $1\text{x}10^{10}$) is a weighing value used to increase and decrease the influence of regularization in the optimization.

The equilibrium pressure, $P_{i\text{-}calc}$ was calculated in two different ways depending on whether the VLE systems need an association term (such system involving carboxylic acids) or not. For systems without association term, the pressure is calculated as follows:

$$\mathrm{P}_{i-calc} = \sum_i \mathrm{x}_i \gamma_i \mathrm{P}_i^{sat}(\mathrm{POY}_i) \qquad (5.2)$$

where *i* is an index running over all species in the mixture and POY_i is the Poynting factor

However, for the VLE systems needing the association term, the equilibrium pressure is calculated using the fugacity coefficient in the vapor phase estimated based on the method of Hayden and O'Connell [120] described briefly as follows:

$$P_{i-calc} = \sum_i \frac{x_i \gamma_i P_i^{sat}(POY_i)}{\Phi_i} \qquad (5.3)$$

$$\ln \Phi_i = \left[2\sum_{j=1}^{N} y_j B_{ij} - B\right]\frac{P}{RT} \qquad (5.4)$$

where Φ_i is the fugacity coefficient, y_j is the vapor mole fraction, $B_{ij}(T)$ is the second virial coefficient characterizing pair interactions between *i* and *j* molecule and B is the second virial coefficient. The cross second virial coefficient, B_{ij} can be calculated directly from PVT data, from statistical mechanical formulas or from empirical and semitheoretical correlations [120]. On the other hand, the second virial coefficient, B are calculated using the equations below which are

contributed by the different types of intermolecular forces which are described as bound, metastably bound, free pairs and B_{chem} which include the association contribution.

$$B = B_{free} + B_{metastable} + B_{bound} + B_{chem} \tag{5.5}$$

The bound and metastably bound contribution of the second virial coefficient are calculated in Equation (4.6) while the free pairs contribution is calculated in Equation (4.7):

$$B_{metastable} + B_{bound} = b_0 A \exp\left[\Delta H / kT / \varepsilon\right] \tag{5.6}$$

$$\text{with } b_0 = \frac{2\pi}{3} N_0 \sigma^3,\ A = -0.3 - 0.05\mu^*,\ \Delta H = 1.99 + 0.2\mu^{*2},\ \mu^* = \mu^2 / \varepsilon\sigma^3$$

$$B_{free} = b_0 \left[(0.94 - 1.47 / T^{*'} - 0.85 / T^{*'2} + 1.015 / T^{*'3}) - \mu^{*'} (0.75 - 3 / T^{*'} + 2.1 / T^{*'2} + 2.1 / T^{*'3} \right] \tag{5.7}$$

where ε and σ are the effective nonpolar potential parameters while μ is the moleculat dipole moment. Moreover, by taking into account the contribution of association, B_{chem} is calculated as follows:

$$B_{chem} = b_0 \exp\left\{\eta\left[650 / \left(\varepsilon / k + 300\right) - 4.27\right]\right\} \times \left\{1 - \exp\left[1500\eta / T\right]\right\} \tag{5.8}$$

where η is the association parameter.

On the other hand, the liquid composition, $x_{i\text{-}calc}$ are calculated using the SLE equation expressed in Equation (5.9) as follows,

$$x_{i-calc} = \frac{1}{\gamma_i} \left\{ \exp\left[\frac{\Delta H_{fus,i}}{R T_{m,i}} \left(1 - \frac{T_{m,i}}{T} \right) \right] \right\} \tag{5.9}$$

where $T_{m,i}$ and $\Delta H_{fus,i}$ are the solute's melting temperature and heat of fusion respectively, R is the gas constant, T is the temperature and γ_i is the activity coefficient. $x_{i\text{-}calc}$ is obtained through an iteration procedure by matching the temperature and composition used to calculated the activity coefficient. With respect to the objective function in Equation (5.1), when VLE data are used, only the first term is used while when SLE data are used, the second term was used in regressing the model parameters. In addition, the regularization term (third term in Equation (5.1)) is used in order to ensure that the regressed parameters does not go to far from the initial values and only used whenever needed. This would be useful as the previously investigated systems can still be represented with comparable accuracy. The higher the value of the weighing value, w_{reg} the higher the influence of the regularization.

5.2.2 Equilibrium Data

The input for the parameter estimation are VLE and SLE experimental data (involving C, H, O, N Cl and S atoms) and the statistics of data used in the regression work are summarized in Table 5.1. For the SLE data, only systems with experimental melting temperature and heats of fusion are used which have been compiled in the CAPEC database [121]. All the SLE data have been collected from the DECHEMA Chemistry Data Series database [122-123]. On the other hand, all VLE data have been initally tested for consistency using a quality assesment algorithm which combines four widely used consistency tests (Herington, Van Ness, Differential and Infinite Dilution tests) and a check between the consistency of the binary and the pure component vapor pressure [127].

For the parameter regression work, a total of 371 VLE datasets with 5707 data points consisting of P, T, x_i, y_i and P, T, x_i data and 273 (T, x_i) SLE datasets with 608 data points have been used. Only a moderate amount of experimental data were used because the purpose of the development of the Original UNIFAC-CI models is to be able to predict phase equilibria with a limited amount of experimental data and without using new experimental data [16].

Table 5.1: Statistics of the Data Used in the Parameter Regression.

Phase Equilibria Information	Types of system	Type of data	No. of systems	Data Points	Main Groups Involved
VLE	Hydrocarbons	P, T, x_i, y_i	49	791	CH_2, C=C, ACH, $ACCH_2$, OH, ACOH, CH_2CO, CHO, CCOO, HCOO, CH_2O, COOH
	Oxygenated	P, T, x_i, y_i	161	2932	
	Nitrogenated	P, T, x_i, y_i	25	808	CH_2, ACH, $ACCH_2$, OH, CH_2CO, CCOO, CNH_2, CNH, $(C)_3N$, $ACNH_2$, PYR, CCN, CNO_2
		P, T, x_i	30		
	Chlorinated	P, T, x_i, y_i	47	1258	CH_2, C=C, ACH, $ACCH_2$, OH, CH_2CO, CCOO, CH_2O, CCl, CCl_2, CCl_3, CCl_4, ACCl
		P, T, x_i	34		
	Sulfurated	P, T, x_i, y_i	14	172	CH_2, C=C, ACH, OH, CH_2CO, CCOO, CH_2O, CCl_3, CCl_4, DMSO, CH_2S
	Total		**360**	**5961**	
SLE	Hydrocarbons	T, x_i	53	82	CH_2, C=C, ACH, $ACCH_2$, OH, CH_3OH, CH_2CO, CCOO, CH_2O, COOH
	Oxygenated	T, x_i	173	323	
	Nitrogenated	T, x_i	15	60	CH_2, ACH, $ACCH_2$, OH, CHO, COOH, $(C)_2NH$, $ACNH_2$, PYR, CCN, $ACNO_2$, DMF
	Chlorinated	T, x_i	17	37	CH_2, ACH, $ACCH_2$, ACOH, CCOO, COOH, CCl, CCl_2, CCl_3, CCl_4, ACCl
	Total		**258**	**502**	

5.2.3 Regression Procedure

Similar to the work in Chapter 4, in this work, the atom interactions parameters (AIPs) are regressed in series. Initially, only systems with C-C atom interactions are used for the parameter regression. For this step, 49 VLE and 53 SLE experimental data were used. The parameters involved in the regression of AIPs related to C-C interaction are $b_{C\text{-}C}$, $c_{C\text{-}C}$, $d_{C\text{-}C}$, $e_{C\text{-}C}$, $bh_{C\text{-}C}$, $ch_{C\text{-}C}$, $dh_{C\text{-}C}$, $eh_{C\text{-}C}$.

For the next step, systems related to the C-O and O-C interactions are added in the parameter regression where the AIPs, $b_{C\text{-}O}$, $c_{C\text{-}O}$, $d_{C\text{-}O}$, $e_{C\text{-}O}$, $b_{O\text{-}C}$, $c_{O\text{-}C}$, $d_{O\text{-}C}$, $e_{O\text{-}C}$, $bh_{C\text{-}O}$, $ch_{C\text{-}O}$, $dh_{C\text{-}O}$, $eh_{C\text{-}O}$, $bh_{O\text{-}C}$, $ch_{O\text{-}C}$, $dh_{O\text{-}C}$, $eh_{O\text{-}C}$ are regressed while the AIPs which were regressed earlier are fixed. A total of 129 VLE and 134 SLE systems were used. Following in the next sequence are the systems related to the O-O interactions, where the AIPs, $b_{O\text{-}O}$, $c_{O\text{-}O}$, $d_{O\text{-}O}$, $e_{O\text{-}O}$, $bh_{O\text{-}O}$, $ch_{O\text{-}O}$, $dh_{O\text{-}O}$, $eh_{O\text{-}O}$ are regressed while fixing the AIPs related to the C-C, C-O and O-C interactions which were regressed earlier. A total of 32 data sets of VLE systems and 39 data sets of SLE systems were used in this step.

The regression continued in the same procedure where the parameters related to interactions C-N, O-N, N-N, C-Cl, O-Cl, C-S, O-S and Cl-S were added sequentially. In each step, the previously fitted parameters were fixed. The overall regression procedure is illustrated in Figure 5.1. In the last 3 steps, only regression with VLE data alone are involved as there are no SLE data related to that interaction was used in this work.

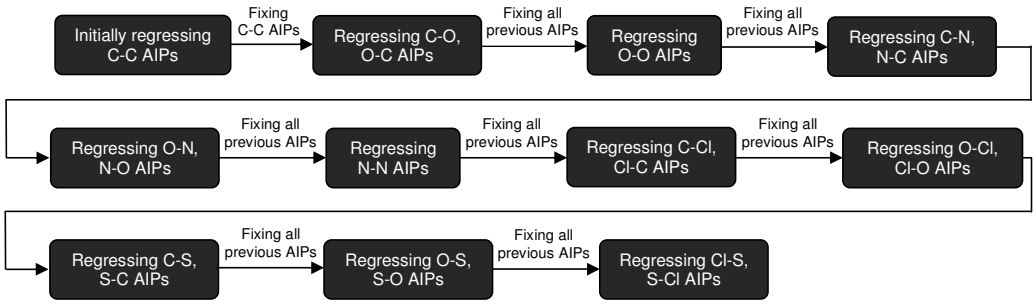


Figure 5.1: Overall Regression Procedure for the AIPs Regression Work.

5.2.4 Correlation Results

The correlation error or deviation between the experimental data and the regressed values are defined in terms of the average absolute relative deviation (AARD) shown in Equations (5.10) and (5.11). For VLE data, the correlation error is defined in the following way:

$$\text{AARD1}(\%) = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{\text{P}_{i-\text{exp}} - \text{P}_{i-calc}}{\text{P}_{i-\text{exp}}}\right| \times 100 \tag{5.10}$$

For SLE data, the measure of correlation error was chosen to be:

$$\text{AARD2}(\%) = \frac{1}{N}\sum_{i=1}^{N}\left|x_{i-\text{exp}} - x_{i-calc}\right| \times 100 \tag{5.11}$$

The correlation error terms equations chosen in order to get a fair comparison between the datasets of the same type and also because of the way the datasets have been set up in the earlier step (in the objective function) to regress the AIPs.

In this section, detailed correlation results of each of the datasets involved in the parameter regression work are presented and described. In Section 5.2.4.1, the correlation results for VLE systems are presented with the parameters that were regressed against VLE data alone (labeled by UNIFAC-CI (VLE) with Q_{VLE}) and simultaneously against VLE and SLE data (labeled by UNIFAC-CI (VLE/SLE)). In Section 5.2.4.2 the correlation errors for the SLE systems are presented with the parameters which were regressed against both type of systems and also compared with the predictions accuracy using the parameters regressed against VLE data alone. At the end of this chapter, the overall correlation results are presented and discussed.

5.2.4.1 Vapor-Liquid Equilibrium Data

The correlation results for systems related to only C-C interactions which are hydrocarbons are presented in Table 5.2. The average AARD1 obtained for the Original UNIFAC model is 1.1 % while for the UNIFAC-CI (VLE/SLE) and UNIFAC-CI (VLE) with Q_{VLE} are 1.5 and 1.2 % respectively. For the CI-models, we can say that for hydrocarbons, the performance is as good as for the reference model. However, we can see that there is a slight increase of error when SLE data are added to the parameter estimation.

Table 5.2: Correlation Results for C-C Atom Interactions Related VLE Systems.

No.	Systems	AARD1 (%)		
		Original UNIFAC	UNIFAC-CI (VLE/SLE)	UNIFAC-CI (VLE) with Q_{VLE}
1	Cyclohexane-Toluene at 323 K	0.94	1.67	1.53
2	m-Xylene-Decane at 394 K	1.76	4.57	3.47
3	Methylcyclohexane-Benzene at 348 K	0.29	2.45	1.65
4	Methylcyclohexane-Toluene at 348 K	0.37	0.51	0.41
5	1-Hexene-Hexane at 333 K	0.50	0.90	0.12

6	1-Hexene-Hexane at 333 K	0.56	1.00	0.13
7	Ethylbenzene-Nonane at 333 K	0.24	0.77	0.54
8	Propylbenzene-Octane at 343 K	0.38	0.38	0.26
9	Propylbenzene-Nonane at 363 K	0.90	0.27	0.47
10	Tetradecane-Benzene at 303 K	5.41	0.68	1.59
11	Benzene-n-Heptane at 298 K	1.49	0.94	0.42
12	1-Hexene-Hexane at 313 K	0.59	1.08	0.07
13	Ethylbenzene-Cyclooctane at 343 K	0.38	0.81	0.60
14	Ethylbenzene-Cyclooctane at 353 K	0.38	0.36	0.22
15	Ethylbenzene-Cyclooctane at 373 K	0.69	0.20	0.26
16	Benzene-2,2,4-Trimethylpentane at 313 K	2.25	0.27	0.67
17	Toluene-2,2,4-Trimethylpentane at 313 K	0.75	1.61	1.36
18	o-Xylene-2,2,4-Trimethylpentane at 313 K	1.98	2.89	1.82
19	Propylbenzene-2,2,4-Trimethylpentane at 313 K	1.79	2.67	2.25
20	Benzene-Cyclohexene at 293 K	0.98	1.86	1.74
21	2,2,4-Trimethylpentane-Toluene at 373 K	1.24	0.30	0.41
22	Benzene-n-Heptane at 333 K	0.49	2.77	1.97
23	Benzene-Cyclohexene at 298 K	1.10	1.53	1.38
24	Benzene-Cyclohexene at 348 K	1.93	2.34	2.54
25	Benzene-Cyclohexane at 283 K	2.19	0.50	1.14
26	Benzene-Cyclohexane at 333 K	0.25	1.47	0.85
27	1-Heptene-n-Heptane at 328 K	0.48	0.10	0.87
28	1-Heptene-n-Octane at 328 K	3.14	2.67	3.58
29	Benzene-Cyclohexane at 298 K	0.79	1.01	0.35
30	Benzene-n-Heptane at 328 K	1.07	1.39	0.97
31	n-Hexane-Benzene at 333 K	0.14	1.90	1.20
32	Methylcyclopentane at 333 K	0.43	1.68	1.05
33	n-Hexane-Benzene at 298 K	1.56	0.66	0.28
34	Benzene-n-Heptane at 298 K	1.55	0.95	0.32
35	Benzene-Cyclohexane at 343 K	0.17	2.04	1.37
36	1-Hexene-Benzene at 298 K	0.25	3.51	3.33
37	1-Hexene-n-Hexane at 328 K	0.37	0.79	0.05
38	1-Hexene-n-Octane at 328 K	2.27	2.90	1.64
39	Benzene-1-Heptene at 328 K	0.35	1.18	1.05
40	Benzene-n-Heptane at 328 K	0.31	1.79	1.06
41	1-Heptene-Toluene at 328 K	1.54	5.22	6.27
42	n-Heptane-Toluene at 328 K	0.41	1.04	0.88
43	Benzene-Cyclohexane at 293 K	1.26	0.21	0.46
44	Benzene-Cyclohexane at 313 K	1.04	1.16	0.37
45	Methylcyclopentane-Benzene at 298 K	1.48	0.68	0.22
46	Methylcyclopentane-Benzene at 313 K	0.79	1.28	0.54

47	n-Heptane-Ethylbenzene at 298 K	1.03	1.91	1.46
48	Benzene-Toluene at 334 K	1.00	4.14	2.37
49	Cyclohexane-Toluene at 318 K	0.79	1.53	1.33
	Average	**1.06**	**1.52**	**1.20**

In Table 5.3, the correlation results with respect to C-O and O-C atom interactions related systems such as aromatics-ethers, alkanes-ethers, esters-alkenes, esters-alkanes, alkanes-aldehydes, aromatics-aldehydes, alcohols-alkane, alcohols-aromatics, ketones-aromatics, ketones-alkanes, carboxylic acids-aromatics and carboxylic acids-alkanes are presented where the average AARD1 obtained for all 3 models (reference model, UNIFAC-CI (VLE/SLE) and UNIFAC-CI (VLE) with Q_{VLE} models) are on average the same with 2.6 %. Individually, when looking at problematic systems for both CI-models, the systems of alkanes-aldehydes and aromatic alcohol-aromatic obtained AARD1 values of more than 10 %. Addition of SLE data to the regression does not seem to affect the correlation error of VLE systems.

Table 5.3: Correlation Results for C-O, O-C Atom Interactions Related VLE Systems.

No.	Systems	AARD1 (%)		
		Original UNIFAC	**UNIFAC-CI (VLE/SLE)**	**UNIFAC-CI (VLE) with Q_{VLE}**
1	1,2-Dimethoxyethane-Toluene at 350 K	9.18	9.76	9.40
2	1,2-Dimethoxyethane-Methylcyclohexane at 350 K	16.76	6.69	6.38
3	Tetrahydrofuran-Hexane at 313 K	3.99	5.51	5.51
4	Tetrahydrofuran-Hexane at 333 K	3.51	4.96	4.97
5	Cyclohexane-tert-Butanol at 328 K	4.58	2.42	2.06
6	Butyraldehyde-n-Heptane at 318 K	1.52	10.85	11.07
7	2-Methylpropanal-n-Heptane at 318 K	3.48	10.31	10.55
8	1-Butanol-n-Heptane at 363 K	0.98	2.89	3.61
9	Methyl-Acetate-1-Hexene at 323 K	1.02	1.89	2.15
10	1-Hexene-Ethyl-Acetate at 333 K	0.28	1.00	1.03
11	Butyl-Acetate-1-Decene at 363 K	2.29	4.83	4.79
12	2-Butanone-Benzene at 328 K	1.61	2.89	1.79
13	2-Butanol-Hexane at 333 K	2.53	2.07	2.36
14	Butyl-Acetate-n-Heptane at 347 K	1.04	6.35	6.10
15	Ethyl-Butyrate-n-Heptane at 373 K	2.10	1.81	1.76
16	2-Methyl-1-Propanol-Ethylbenzene at 313 K	2.86	5.56	3.88
17	2-Methyl-1-Propanol-p-Xylene at 313 K	5.82	1.18	3.50
18	tert-Butanol-Benzene at 313 K	1.26	0.77	3.10
19	tert-Butanol-Toluene at 313 K	0.60	1.17	0.60
20	tert-Butanol-Ethylbenzene at 313 K	2.07	1.12	1.24

21	Methyl-Butyl-Ether-Benzene at 343 K	0.13	1.68	2.18
22	Diisopropyl-Ether-Benzene at 343 K	1.02	2.06	1.37
23	2-Propanol-Hexane at 328 K	1.83	3.87	4.49
24	2-Propanol-Octane at 353 K	5.30	2.31	1.96
25	2-Methyl-1-Propanol-n-Heptane at 333 K	0.73	3.07	3.82
26	Phenol-Styrene at 373 K	4.04	11.51	11.69
27	1-Propanol_Octane at 363 K	2.28	1.98	2.70
28	1-Butanol-n-Heptane at 333 K	1.40	4.79	5.55
29	TAME-Heptane at 313 K	3.03	0.83	0.84
30	TAME-Heptane at 313 K	3.71	1.27	1.28
31	TAME-Cyclohexane at 313 K	2.69	0.64	0.64
32	TAME-Benzene at 313 K	0.93	3.00	3.42
33	1-Butanol-Octane at 373 K	1.06	3.43	4.20
34	2-Butanol-Octane at 358 K	5.55	2.12	1.53
35	tert-Butanol-Octane at 343 K	7.12	3.90	3.28
36	Acetone-Cyclohexane at 303 K	2.41	2.14	2.79
37	2-Propanol-Octane at 348 K	4.02	1.62	1.98
38	Hexane-Methyl-tert-Butyl-Ether at 313 K	3.38	0.74	0.74
39	Dipentyl-Ether_Undecane at 403 K	0.85	0.37	0.37
40	Diisopropyl-Ether-n-Heptane at 323 K	0.28	0.26	0.25
41	Acetone-Cyclohexane at 313 K	1.47	1.33	2.12
42	Diethoxymethane-n-Heptane at 323 K	6.07	1.02	0.97
43	Diethoxymethane-n-Heptane at 343 K	5.38	1.13	1.08
44	1-Butanol-Decane at 373 K	2.42	1.09	0.88
45	1,2-Dimethoxyethane-2,4-Dimethylpentane at 343 K	16.04	4.81	4.47
46	n-Heptane-3-Pentanol at 368 K	5.63	3.50	2.89
47	2-Methyl-1-Propanol-Octane at 373 K	2.47	1.34	1.49
48	Octane-Methyl-Tert-Amyl-Ether at 323 K	4.26	1.79	1.79
49	1-Propanol-Octane at 358 K	1.97	3.95	4.73
50	TAME-2,2,4-Trimethylpentane at 311 K	3.82	1.48	1.49
51	TAME-Methylcyclohexane at 293 K	4.41	2.05	2.07
52	TAME-Methylcyclohexane at 311 K	4.00	1.99	1.99
53	TAME-Amyl-Ether-Methylcyclohexane at 333 K	3.96	2.08	2.08
54	TAME-Toluene at 293 K	0.98	0.67	0.33
55	Toluene-tert-Butyl-Ethyl-Ether at 333 K	0.84	0.75	1.00
56	Methyl-Ethyl-Ketone-n-Heptane at 318 K	2.75	3.01	2.47
57	n-Heptane-Methyl-Isobutyl-Ketone at 343 K	3.06	3.43	3.14
58	Benzene-Acetic-Acid at 298 K	1.29	4.13	3.28
59	n-Heptane-3-Pentanone at 353 K	2.08	1.78	1.99
60	2-Methyl-1-Propanol-Toluene at 353 K	0.37	1.57	0.83
61	1,4-Dioxane-Toluene at 353 K	0.27	0.37	0.38

62	1,4-Dioxane-Toluene at 373 K	0.18	0.91	0.71
63	Acetone-n-Hexane at 328 K	1.28	1.23	1.01
64	Benzene-2-Butanol at 318 K	0.91	1.24	1.65
65	Benzene-2-Methyl-2-Propanol at 318 K	1.45	0.90	3.36
66	Acetone-n-Hexane at 293 K	2.22	2.71	3.44
67	Acetone-Benzene at 298 K	0.94	0.57	0.40
68	Ethyl-Acetate-Cyclohexane at 293 K	2.91	3.05	2.66
69	2-Pentanone-Toluene at 323 K	0.98	0.08	1.25
70	Di-n-Propyl-Ether-n-Heptane at 343 K	1.33	0.31	0.32
71	Benzene-Di-n-Propyl-Ether at 343 K	0.49	0.75	0.49
72	Benzene-1,2-Dimethoxyethane at 343 K	3.73	1.26	1.85
73	Benzene-2-Pentanone at 323 K	0.46	2.21	1.14
74	n-Heptane-3-Pentanone at 327 K	0.61	0.49	0.55
75	Tetrahydrofuran-Cyclohexane at 313 K	4.24	6.10	6.10
76	Tetrahydrofuran-Cyclohexane at 323 K	3.44	5.36	5.37
77	Tetrahydrofuran-Cyclohexane at 333 K	3.57	5.23	5.24
78	Tetrahydrofuran-Benzene at 303 K	3.46	1.05	0.91
79	Tetrahydrofuran-Benzene at 313 K	2.39	0.77	0.68
80	Tetrahydrofuran-Benzene at 323 K	2.27	0.64	0.56
81	Tetrahydrofuran-Benzene at 333 K	2.17	0.73	0.49
82	n-Heptane-2-Butanol at 338 K	3.30	1.04	1.17
83	Acetone-Benzene at 298 K	1.15	1.45	0.97
84	Methyl-Ethyl-Ketone-Benzene at 298 K	1.00	1.55	1.02
85	1-Propanol-2,2,4-Trimethylpentane at 328 K	1.81	2.75	3.50
86	Ethanol-Toluene at 333 K	0.76	2.20	1.17
87	1-Hexene-Ethyl-Acetate at 313 K	1.06	1.65	1.92
88	n-Hexane-Isopropanol at 331 K	6.15	3.02	2.73
89	n-Heptane-Propionic-Acid at 323 K	7.35	4.47	4.03
90	Acetone-n-Hexane at 313 K	0.54	0.70	1.27
91	_Acetone-n-Heptane at 313 K	0.64	0.94	1.49
92	Acetone-n-Octane at 313 K	1.37	1.67	2.43
93	Acetone-Benzene at 313 K	1.11	0.73	0.61
94	Acetone-Toluene at 313 K	2.67	4.68	2.67
95	Acetone-Ethylbenzene at 313 K	3.58	3.67	2.32
96	n-Heptane-2-Butanol at 348 K	4.07	1.16	1.09
97	Methyl-Ethyl-Ketone-Benzene at 328 K	0.66	0.44	0.51
98	n-Heptane-1-Butanol at 363 K	0.98	2.89	3.60
99	n-Heptane-2-Methyl-1-Propanol at 333 K	0.73	3.07	3.81
100	Methyl-Ethyl-Ketone-Benzene at 313 K	1.48	2.63	1.66
101	2-Methylpropanal-n-Heptane at 335 K	3.32	10.12	10.34
102	Ethyl-Acetate-n-Heptane at 323 K	0.87	6.95	6.52

103	Cyclohexane-2-Methyl-2-Propanol at 328 K	4.58	2.42	2.06
104	1-Butanal-Toluene at 313 K	0.23	0.43	0.91
105	1-Butanal-Benzene at 313 K	0.54	1.94	1.94
106	1-Butanal-Toluene at 327 K	1.69	3.23	3.75
107	1-Propanal-Benzene at 313 K	0.81	3.21	3.19
108	1-Propanal-Toluene at 313 K	0.61	0.30	0.60
109	Benzene-1-Pentanal at 313 K	0.23	1.59	1.61
110	1-Pentanal-Toluene at 313 K	0.53	0.86	1.37
111	Ethylbenzene-Benzaldehyde at 348 K	8.13	0.95	1.71
112	Styrene-Benzaldehyde at 363 K	4.42	1.70	1.46
113	Methyl-Tert-Butyl-Ether-Hexane at 313 K	3.34	0.74	0.73
114	n-Propyl-Formate-Benzene at 303 K	0.22	6.65	4.89
115	n-Heptane-1-Pentanal at 348 K	1.25	8.78	8.94
116	Acetone-Cumene at 293 K	4.60	2.67	2.39
117	1-Butanal-n-Heptane at 343 K	3.14	9.37	9.58
118	n-Octane-Methyl-Tert-Pentyl-Ether at 323 K	4.26	1.79	1.79
119	Toluene-3-Methyl-1-Butanol at 368 K	1.50	0.93	1.07
120	Cyclohexane-Methyl-Ethyl-Ketone at 323 K	3.27	3.47	3.10
121	Methyl-Propionate-n-Heptane at 323 K	5.11	0.82	0.98
122	Methyl-Ethyl-Ketone-Toluene at 330 K	0.47	0.84	0.60
123	Diisopropyl-Ether-Dimethylpentane at 343 K	0.65	1.06	1.06
124	Toluene-Methyl-Ethyl-Ketone at 323 K	0.87	0.52	1.23
125	Toluene-2-Pentanone at 323 K	0.99	0.09	1.25
126	Toluene-Methyl-Isobutyl-Ketone at 323 K	2.89	1.46	0.25
127	Acetone-Benzene at 323 K	3.44	3.41	4.60
128	Methyl-Ethyl-Ketone-Benzene at 323 K	1.05	2.34	1.24
129	Methyl-Isobutyl-Ketone-Benzene at 323 K	1.65	0.70	0.42
	Average	**2.61**	**2.59**	**2.62**

Table 5.4 shows the correlation results for systems related to O-O atom interactions and these involve ethers-esters, ethers-ketones, alcohols-ketones, esters-aldehydes, alcohols-esters, alcohols-ethers, carboxylic acids-ketones, esters-ketones and aldehydes-ketones. The prediction error obtained for the reference model is 2.2 %. Suprisingly, when SLE systems are added in the parameter regression, the correlation error increases from 4.3 % for the UNIFAC-CI (VLE) with Q_{VLE} model to 2.0 % for the UNIFAC-CI (VLE/SLE) model which is similar to the reference UNIFAC model.

The problematic systems (AARD1 > 10 %) identified for the UNIFAC-CI (VLE) with Q_{VLE} model are related to systems involving tetrahydrofuran-ester (2 systems), alcohol-ketone (2 systems). However, the inclusion of SLE systems for regression have reduced the correlation errors of these systems between 0.4 to 2.8 %.

Table 5.4: Correlation Results for O-O Atom Interactions Related VLE Systems.

No.	Systems	AARD1 (%)		
		Original UNIFAC	UNIFAC-CI (VLE/SLE)	UNIFAC-CI (VLE) with Q_{VLE}
1	Tetrahydrofuran-Ethyl-Acetate at 313 K	0.27	0.38	11.34
2	Tetrahydrofuran-Ethyl-Acetate at 333 K	0.44	0.91	12.24
3	Dibutyl-Ether-2-Heptanone at 393 K	0.30	0.17	0.45
4	Dibutyl-Ether-3-Heptanone at 363 K	1.43	1.50	1.70
5	Dibutyl-Ether-4-Heptanone at 363 K	1.26	1.35	1.55
6	Ethanol-Acetone at 397 K	3.62	2.81	12.16
7	Ethyl-Acetate-2-Methylpropanal at 313 K	1.78	1.28	1.09
8	Ethanol-Methyl-Butyl-Ether at 338 K	2.56	5.82	8.18
9	Ethanol-Dipropyl-Ether at 308 K	2.45	6.97	9.47
10	Ethanol-Methyl-Proponate at 346 K	1.06	0.35	4.19
11	Ethyl-Acetate-2-Propanol at 333 K	2.42	2.94	2.88
12	Methyl-tert-Butyl-Ether-tert-Butanol at 339 K	5.86	1.27	1.67
13	Ethanol-Diisopropyl-Ether at 353 K	4.76	1.05	0.38
14	Acetone-Diisopropyl-Ether at 343 K	4.20	3.95	3.31
15	Acetone-Ethanol at 353 K	3.11	1.60	10.08
16	Diisopropyl-Ether-2-Methoxy-Ethanol at 341 K	8.97	5.25	7.64
17	1,4-Dioxane-2-Methyl-1-Propanol at 353 K	3.68	1.23	3.54
18	1,4-Dioxane-2-Methyl-1-Propanol at 373 K	4.74	0.27	1.77
19	Acetone-Acetic-Acid at 303 K	1.02	0.70	1.35
20	Ethanol-Methyl-Ethyl-Ketone at 298 K	0.57	3.76	3.98
21	Ethanol-Di-n-Butyl-Ether at 323 K	6.17	2.35	7.05
22	Acetone-Tetrahydrofuran at 302 K	3.21	1.18	1.01
23	Acetone-Methyl-Acetate at 328 K	0.45	0.11	0.28
24	Methyl-Propionate-1-Propanol at 328 K	0.49	0.71	4.94
25	Methyl-Acetate-1-Butanal at 313 K	0.34	3.18	3.05
26	1-Propanal-Ethyl-Acetate at 303 K	1.23	2.29	2.43
27	1-Propanal-Methyl-Acetate at 303 K	0.38	0.92	0.53
28	1-Butanal-n-Propyl-Acetate at 333 K	0.20	1.00	0.31
29	1-Propanal-Methyl-Ethyl-Ketone at 318 K	0.35	0.80	9.29
30	1-Propanol-Methyl-Propionate at 328 K	0.43	0.76	4.93
31	1-Propanol-Methyl-N-Butyrate at 333 K	1.20	0.52	3.42
32	Acetone-Diethyl-Ether at 303 K	0.32	2.01	2.33
	Average	**2.16**	**1.86**	**4.33**

Now we move to nitrogenated systems which involve C, H, O and N atoms. In Table 5.5, the correlation errors of the systems related to C-N, N-C and the previous atoms interactions are presented which involve alkyl-pyridines-aromatics, alkyl-pyridines-alkanes, nitriles-aromatics and amines-alkanes systems. The results show that the average error for the UNIFAC-CI (VLE/SLE) model is 2.3% which is close to the UNIFAC-CI (VLE) with Q_{VLE} model with 2.4 %. However, the correlation errors of these CI-models are better than the reference UNIFAC model with 3.2 %. For this part, we can also see that the addition of SLE systems does not really affect the AARD1 values obtained for VLE systems.

Table 5.5: Correlation Results for C-N, N-C Atom Interactions Related VLE Systems.

No.	Systems	AARD1 (%)		
		Original UNIFAC	**UNIFAC-CI (VLE/SLE)**	**UNIFAC-CI (VLE) with Q_{VLE}**
1	4-Methylpyridine-Benzene at 313 K	1.16	1.03	1.05
2	4-Methylpyridine-Toluene at 313 K	1.11	0.72	0.57
3	3-Methylpyridine-Benzene at 313 K	0.69	1.03	1.15
4	3-Methylpyridine-Toluene at 313 K	1.10	0.38	0.35
5	2-Methylpyridine-Benzene at 313 K	0.33	1.21	1.36
6	2-Methylpyridine-Toluene at 313 K	0.70	0.62	0.77
7	2-Methylpiridine-n-Octane at 313 K	4.22	2.90	2.97
8	2-Methylpyridine-n-Nonane at 313 K	4.34	3.50	3.46
9	2-Methylpyridine-n-Heptane at 313 K	4.39	3.93	4.02
10	Toluene-Acetonitrile at 293k	0.54	0.59	0.79
11	Toluene-Acetonitrile at 343k	1.82	0.49	0.75
12	Toluene-Acetonitrile at 393k	4.35	0.55	0.70
13	Ethylbenzene-Propionitrile at 313k	1.47	1.08	1.45
14	Ethylbenzene-Propionitrile at 353k	1.07	0.93	1.17
15	Ethylbenzene-Propionitrile at 393k	0.73	0.90	1.10
16	n-Hexane-Triethylamine at 298 K	5.21	1.60	2.20
17	Triethylamine-n-Octane at 298 K	4.32	1.15	1.50
18	Dimethylamine-n-Hexane at 258 K	11.94	6.61	8.52
19	Dimethylamine-n -Hexane at 268 K	9.44	4.80	6.16
20	Dimethylamine-n -Hexane at 278 K	7.07	2.96	4.13
21	Dimethylamine-n -Hexane at 288 K	4.69	2.24	2.21
22	Dimethylamine-n -Hexane at 298 K	2.92	3.11	1.85
23	Dimethylamine-n -Hexane at 308 K	1.64	5.41	3.27
24	Dimethylamine-n -Hexane at 318 K	2.59	7.79	5.68
	Average	**3.24**	**2.31**	**2.38**

Next, the regression errors for O-N, N-O related systems (alcohols-nitriles, ketones-anilines, esters-anilines, amines-ketones, amines-esters) are presented in Table 5.6 with the average correlations errors of 5.5, 2.4 and 2.5 % respectively for the Original UNIFAC, UNIFAC-CI (VLE) with Q_{VLE} and UNIFAC-CI (VLE/SLE) models respectively. The same trend was found in this part where the inclusion of SLE system still does not significantly affect the correlation error of the VLE systems. However, the performance of these CI-models are better than the Original UNIFAC model for most of the VLE systems listed in Table 5.6. High prediction errors (> 10%) which were found for the n-butyronitrile-2-butanol systems for the Original UNIFAC have been reduced significantly when using the CI-models.

Table 5.6: Correlation Results for O-N, N-O Atom Interactions Related VLE Systems.

No.	Systems	AARD1 (%)		
		Original UNIFAC	**UNIFAC-CI (VLE/SLE)**	**UNIFAC-CI (VLE) with Q_{VLE}**
1	n-Butyronitrile-2-Butanol at 278 K	17.48	6.68	7.38
2	n-Butyronitrile-2-Butanol at 288 K	14.28	3.61	4.37
3	n-Butyronitrile-2-Butanol at 293 K	12.59	1.97	2.76
4	n-Butyronitrile-2-Butanol at 298 K	11.51	1.02	1.83
5	n-Butyronitrile-2-Butanol at 303 K	10.53	0.44	0.99
6	n-Butyronitrile-2-Butanol at 308 K	9.79	0.66	0.54
7	Acetonitrile-Ethanol at 293 K	9.11	3.58	2.20
8	Acetonitrile-Ethanol at 343 K	4.34	7.52	5.88
9	Acetonitrile-Ethanol at 393 K	0.65	10.89	9.12
10	Acetone-Aniline at 313 K	3.60	2.60	2.39
11	Acetone-Aniline at 386 K	4.00	3.29	2.65
12	Ethyl-Acetate-Aniline at 297 K	1.11	1.33	2.91
13	Aniline-Ethyl-Acetate at 348 K	1.62	0.61	0.74
14	Aniline-Ethyl-Acetate at 397 K	3.48	1.74	2.41
15	Diethylamine-Acetone at 298 K	0.74	0.56	0.26
16	Diethylamine-Acetone at 347 K	1.46	0.66	0.24
17	Diethylamine-Acetone at 398 K	1.59	0.60	0.20
18	Diethylamine-Ehtyl-Acetate at 297 K	0.26	1.19	1.08
19	Diethylamine-Ethyl-Acetate at 348 K	0.63	0.10	0.08
20	Diethylamine-Ethyl-Acetate at 398 K	1.10	0.51	0.65
	Average	**5.49**	**2.48**	**2.43**

In addition, for the regression of N-N AIPs related datasets involving 11 systems of nitro-alkanes-nitriles, pyridines-amines, nitriles-anilines, amines-nitriles and amines-amines systems, the

correlation results are listed in Table 5.7. The average errors show that the performance of the CI-models are similar about 1.0 % compared to the reference UNIFAC models with 1.4 %.

Table 5.7: Correlation Results for N-N Atom Interactions Related VLE Systems.

No.	Systems	AARD1 (%)		
		Original UNIFAC	UNIFAC-CI (VLE/SLE)	UNIFAC-CI (VLE) with Q_{VLE}
1	Nitromethane-Acetonitrile at 298 K	0.20	0.57	1.28
2	Nitromethane-Acetonitrile at 348 K	0.19	0.21	0.42
3	Nitromethane-Acetonitrile at 398 K	0.27	0.51	1.22
4	Acetonitrile-Aniline at 293 K	1.37	1.08	1.04
5	Acetonitrile-Aniline at 343 K	3.36	0.76	0.85
6	Acetonitrile-Aniline at 393 K	4.95	1.71	1.88
7	Diethylamine-Acetonitrile at 298 K	0.40	1.68	1.22
8	Diethylamine-Acetonitrile at 347 K	1.33	0.50	0.50
9	Diethylamine-Acetonitrile at 398 K	2.18	1.77	1.34
10	n-Propylamine-Diethylamine at 297 K	0.40	0.69	0.73
11	n-Propylamine-Diethylamine at 347 K	0.38	0.84	0.89
	Average	**1.37**	**0.94**	**1.03**

From nitrogenated to chlorinated systems, Table 5.8 highlights the correlation results with respect to C-Cl, Cl-C interactions related systems such as alkyl-chloride-aromatics, alkyl-chlorides-alkanes, chloro-alkanes-alkanes, chloro-aromatics-aromatics, carbon tetrachlorides-alkanes, chloroalkanes-aromatics and carbon tetrachlorides-aromatics. The average AARD1 value for Original UNIFAC is 2.7 % while for the CI-models the values are 3.6 % for both UNIFAC-CI (VLE) and UNIFAC-CI (VLE/SLE) models. Eventough the AARD1 values for the CI-model are higher than the reference UNIFAC model, addition of SLE data still does not significantly affect the average error of the VLE systems. However, 2 alkane-chloroalkane systems which obtained AARD1 > 10 % for the UNIFAC-CI (VLE) with Q_{VLE} have been reduced to < 10 % for the UNIFAC-CI (VLE/SLE) model.

Table 5.8: Correlation Results for C-Cl, Cl-C Atom Interactions Related VLE Systems.

No.	Systems	AARD1 (%)		
		Original UNIFAC	UNIFAC-CI (VLE/SLE)	UNIFAC-CI (VLE) with Q_{VLE}
1	sec-Butyl-Chloride-Toluene at 323 K	0.71	1.06	1.76
2	sec-Butyl-Chloride-n-Heptane at 323 K	4.50	3.35	3.67

3	sec-Butyl-Chloride-n-Heptane at 333 K	4.32	3.22	3.53
4	1,2-Dichloroethane-n-Heptane at 343 K	1.03	4.56	2.03
5	n-Butyl-Chloride-Toluene at 323 K	0.33	1.39	2.35
6	n-Butyl-Chloride-n-Heptane at 323 K	2.94	1.28	1.76
7	tert-Butyl-Chloride-n-Heptane at 323 K	4.57	3.85	4.04
8	n-Butyl-Chloride-n-Heptane at 323 K	2.94	1.28	1.76
9	n-Butyl-Chloride-Toluene at 323 K	0.33	1.39	2.35
10	sec-Butyl-Chloride-Toluene at 323 K	0.71	1.06	1.76
11	n-Butyl-Chloride-Toluene at 298 K	1.06	1.85	1.65
12	n-Butyl-Chloride-Toluene at 348 K	0.47	1.39	2.48
13	n-Butyl-Chloride-Toluene at 398 K	0.17	2.39	3.55
14	Benzene-Monochlorobenzene at 298 K	0.54	3.94	3.92
15	Benzene-Monochlorobenzene at 348 K	0.88	3.17	3.14
16	Benzene-Monochlorobenzene at 398 K	1.38	2.38	2.34
17	n-Pentane-Dichloromethane at 298 K	8.83	6.10	6.26
18	n-Pentane-Dichloromethane at 348 K	5.66	3.24	3.38
19	n-Pentane-Dichloromethane at 398 K	3.47	1.29	1.43
20	1,2-Dichloroethane-n-Heptane at 340 K	1.03	4.56	2.03
21	n-Heptane-Dichloromethane at 298 K	8.19	4.65	4.86
22	n-Heptane-1,2-Dichloroethane at 298 K	2.73	5.00	1.85
23	n-Heptane-1,3-Dichloropropane at 298 K	15.96	7.74	10.83
24	n-Hexane-1,4-Dichlorobutane at 298 K	15.97	9.68	11.75
25	Carbon-Tetrachloride-n-Hexane at 298 K	2.35	3.06	2.09
26	Carbon-Tetrachloride-n-Heptane at 298 K	2.76	4.04	2.83
27	Chloroform-n-Hexane at 298 K	1.03	2.14	2.96
28	Chloroform-n-Heptane at 298 K	0.70	1.75	2.60
29	n-Heptane-n-Butyl-Chloride at 298 K	3.79	2.05	2.54
30	n-Pentane-n-Butyl-Chloride at 318 K	3.10	1.66	2.31
31	n-Pentane-n-Butyl-Chloride at 338 K	2.84	1.54	2.18
32	Dichloromethane-Benzene at 298 K	2.91	2.13	1.52
33	Dichloromethane-Benzene at 348 K	1.46	4.57	4.06
34	Dichloromethane-Toluene at 298 K	1.83	8.31	8.02
35	Dichloromethane-Toluene at 347 K	0.51	2.54	2.21
36	Tetrachloroethane-N-Hexane at 298 K	3.01	3.26	2.82
37	Tetrachloroethane-N-Hexane at 308 K	1.74	4.33	3.91
38	1,2-Dichloroethane-N-Hexane at 298 K	5.34	1.43	1.60
39	p-Xylene-Carbon-Tetrachloride at 313 K	4.29	2.31	2.31
40	p-Xylene-Carbon-Tetrachloride at 323 K	3.30	1.66	1.75
41	Dichloromethane-1-Hexene at 298 K	4.28	11.35	13.25
42	Monochlorobenzene-Ethylbenzene at 293 K	0.21	0.93	1.12
43	1-Pentene-Monochlorobenzene at 320 K	2.14	9.77	8.11

44	1-Pentene-Monochlorobenzene at 360 K	4.22	4.02	2.42
45	o-Xylene-Carbon-Tetrachloride at 303 K	0.39	2.76	2.83
46	p-Xylene-Carbon_Tetrachloride at 303 K	0.72	2.76	2.84
47	m-Xylene-Carbon_Tetrachloride at 303 K	0.90	2.61	2.71
48	Cumene-Carbon-Tetrachloride at 303 K	1.33	6.96	6.33
49	n-Heptane- Carbon-Tetrachloride at 323 K	1.30	3.92	2.94
50	n-Butyl-Chloride-n-Heptane at 323 K	1.01	4.44	4.38
51	Carbon-Tetrachloride-Benzene at 260 K	0.88	3.98	3.89
52	Carbon-Tetrachloride-Benzene at 265 K	0.84	3.83	3.73
53	Carbon-Tetrachloride-Benzene at 266 K	0.76	3.43	3.31
54	Carbon-Tetrachloride-Benzene at 271 K	0.71	3.21	3.08
55	Carbon-Tetrachloride-Benzene at 273 K	1.30	5.60	4.20
	Average	**2.74**	**3.57**	**3.55**

In addition, the correlation results for systems related to O-Cl, Cl-O interactions are presented in Table 5.9. The systems involved in this part are related to chloro-alkanes-ethers, esters-chloroform, ethers-chloroalkanes, alcohols-chloroforms, ketones-chloroforms and ketones-chloro-aromatics. The average correlation results for the CI-models are better than the reference model with 2.9 and 2.4 % respectively for UNIFAC-CI (VLE/SLE) and UNIFAC-CI (VLE) with Q_{VLE} compared to 5.2 % for the Original UNIFAC. This time we can see that the addition of SLE systems for regression have slightly increase the average correlation error for the CI-model.

Table 5.9: Correlation Results for O-Cl, Cl-O Atom Interactions Related VLE Systems.

No.	Systems	AARD1 (%)		
		Original UNIFAC	**UNIFAC-CI (VLE/SLE)**	**UNIFAC-CI (VLE) with Q_{VLE}**
1	Dichloroethane-Di-n-Butyl-Ether at 330 K	2.74	5.25	1.39
2	Dichloroethane-Di-n-Butyl-Ether at 350 K	3.02	2.47	4.26
3	Dichloroethane-Di-n-Butyl-Ether at 370 K	1.81	5.53	2.61
4	Trichloroethane-Di-n-Butyl-Ether at 323 K	0.71	3.84	1.86
5	Ethanol-Monochlorobenzene at 298 K	0.87	4.38	1.88
6	Ethanol-Monochlorobenzene at 398 K	4.38	4.79	3.83
7	Dichloromethane-Ethyl-Acetate at 298 K	3.54	3.85	3.15
8	Dichloromethane-Ethyl-Acetate at 348 K	1.41	10.58	3.47
9	Dichloromethane-Ethyl-Acetate at 398 K	10.90	2.84	1.97
10	1-Chloropentane-Di-n-Butyl-Ether at 313 K	16.19	5.11	3.57
11	1-Chloropentane-Di-n-Butyl-Ether at 323 K	7.77	1.46	2.06
12	1,2-Dichloroethane-Di-n-Butyl-Ether at 330 K	6.90	1.28	2.10

13	Dichloroethane-Di-n-Butyl-Ether at 350 K	2.49	1.81	1.65
14	Dichloroethane-Di-n-Butyl-Ether at 370 K	2.04	1.35	1.20
15	Trichloroethane-Di-n-Butyl-Ether at 323 K	1.88	1.31	1.13
16	Trichloroethane-Di-n-Butyl-Ether at 343 K	3.76	0.80	1.07
17	Ethanol-Chloroform at 303 K	2.52	0.22	0.43
18	Ethanol-Chloroform at 313 K	2.32	1.01	0.52
19	Ethanol-Chloroform at 323 K	0.60	0.37	3.99
20	Acetone-Monochlorobenzene at 313 K	0.39	0.84	3.11
21	Acetone-Monochlorobenzene at 353 K	0.65	0.73	3.99
22	Acetone-Monochlorobenzene at 386 K	1.16	1.06	2.82
23	Chloroform-Acetone at 313 K	21.22	4.83	3.47
24	Chloroform-Acetone at 323 K	10.90	2.84	1.97
25	Chloroform-Acetone at 303 K	16.19	5.11	3.57
26	1-Butanol-Carbon-Tetrachloride at 323 K	7.77	1.46	2.06
	Average	**5.16**	**2.89**	**2.43**

For the sulfurated systems, the correlation results are presented in Tables 5.10 for C-S and S-C, O-S and S-O and Cl-S and S-Cl interactions related systems. In total, 14 datasets were used for regression and involve systems with diethyl sulfides-cycloalkanes, diethyl sulfides-ethoxy-alkanes, dimethyl-sulfoxides-ketones, dimethyl-sulfoxides-esters and others. For this part only VLE data are used for the regression of parameters. Again very small effect is seen by adding SLE data in the regression.

Table 5.10: Correlation Results for C-S, S-C, O-S, S-O, Cl-S, S-Cl Atom Interactions Related VLE Systems.

No.	Systems	AARD1 (%)		
		Original UNIFAC	**UNIFAC-Cl (VLE/SLE)**	**UNIFAC-Cl (VLE) with Q_{VLE}**
1	Diethyl-Sulfide-Cyclohexane at 343 K	0.55	0.54	0.24
2	Diethyl-Sulfide-Cyclohexane at 353 K	0.24	0.91	0.18
3	Dimethyl-Sulfoxide-Acetone at 298 K	0.79	0.47	0.48
4	Dimethyl-Sulfoxide-Acetone at 308 K	0.16	0.16	0.18
5	Dimethyl-Sulfoxide-Acetone at 318 K	0.70	0.45	0.41
6	Dimethyl-Sulfoxide-Ethyl-Acetate at 298 K	0.74	0.61	0.48
7	Dimethyl-Sulfoxide-Ethyl-Acetate at 308 K	0.75	0.62	0.50
8	Dimethyl-Sulfoxide-Ethyl-Acetate at 318 K	0.80	0.65	0.54
9	Diethyl-Sulfide-2-Ethoxy-2-Methylpropane at 333 K	1.45	0.07	0.53
10	Diethyl-Sulfide-2-Ethoxy-2-Methylpropane at 343 K	1.50	0.07	0.65

11	Carbon-Tetrachloride-Thiophene at 343 K	2.46	0.29	
12	Chloroform-Dimethyl-Sulfide at 298 K	1.58	2.47	2.42
13	Chloroform-Diethyl-Sulfide at 298 K	3.65	2.47	2.27
14	Carbon-Tetrachloride-Diethyl-Sulfide at 298 K	0.33	0.07	0.08
	Average	**1.12**	**0.70**	**0.69**

5.2.4.2 Solid-Liquid Equilibrium Data

In this section, the correlation results for SLE systems are presented which are described by AARD2. For hydrocarbon, only C-C interactions are regressed and Table 5.11 show the average correlation error of 2.8, 2.2 and 2.3 % respectively for Original UNIFAC, UNIFAC-CI (VLE/SLE) and UNIFAC-CI (VLE) with Q_{VLE} models which are considered tobe very similar.

Table 5.11: Correlation Results for C-C Atoms Interactions Related SLE Systems.

No.	SLE Systems	No. of Data Points	AARD2 (%)		
			Original UNIFAC	UNIFAC-CI (VLE/SLE)	UNIFAC-CI (VLE) with Q_{VLE}
1	Biphenyl-Hexane	1	7.63	4.18	5.43
2	Biphenyl-Cyclohexane	1	8.13	5.62	6.66
3	Biphenyl-Octane	1	3.42	0.11	1.1
4	Biphenyl-Nonane	1	2.47	1.09	0.11
5	Biphenyl-Decane	1	1.58	2.01	0.81
6	Biphenyl-Heptane	1	5.00	1.49	2.72
7	Biphenyl-Cyclooctane	1	1.02	2.46	1.17
8	Biphenyl-Isooctane	1	6.02	2.58	3.74
9	Biphenyl-Hexadecane	1	1.96	5.89	4.62
10	Acenaphthene-Methylcyclohexane	1	2.48	3.00	2.74
11	Acenaphthene-Hexane	1	2.40	2.82	2.57
12	Acenaphthene-Cyclohexane	1	4.50	5.02	4.80
13	Acenaphthene-Octane	1	0.69	1.11	0.85
14	Acenaphthene-Nonane	1	0.42	0.85	0.57
15	Acenaphthene-Heptane	1	1.41	1.83	1.58
16	Acenaphthene-Hexadecane	1	1.42	0.91	1.25
17	Acenaphthene-Cyclooctane	1	0.26	0.77	0.49
18	Phenanthrene-Methylcyclohexane	1	7.54	5.02	5.87
19	Phenanthrene-Hexane	1	5.07	3.10	3.72

20	Phenanthrene-Cyclohexane	1	10.25	7.66	8.57
21	Phenanthrene-Octane	1	3.31	1.45	2.03
22	Phenanthrene-Nonane	1	2.95	1.10	1.67
23	Phenanthrene-Decane	1	2.27	0.40	0.98
24	Phenanthrene-Heptane	1	4.01	2.11	2.70
25	Phenanthrene-Cyclooctane	1	5.02	2.60	3.39
26	Phenanthrene-Hexadecane	1	0.90	1.19	0.56
27	Anthracene-Methylcyclohexane	1	0.17	0.09	0.11
28	Anthracene-Hexane	1	0.10	0.05	0.06
29	Anthracene-Cyclohexane	7	2.28	1.29	1.61
30	Anthracene-Octane	1	0.05	0.01	0.01
31	Anthracene-Heptane	1	0.07	0.02	0.03
32	Anthracene-Cyclooctane	1	0.09	0.01	0.04
33	Anthracene-Isooctane	1	0.10	0.05	0.07
34	Anthracene-Hexadecane	1	0.08	0.16	0.13
35	Pyrene-p-Xylene	1	4.28	4.64	3.37
36	Pyrene-m-Xylene	1	4.06	4.86	3.59
37	Pyrene-o-Xylene	1	3.50	5.42	4.15
38	Pyrene-Toluene	1	5.11	0.68	2.05
39	Naphthalene-Ethylbenzene	1	0.85	4.43	3.13
40	Naphthalene-Toluene	7	1.65	4.43	2.64
41	Naphthalene-Hexane	1	2.97	0.63	1.42
42	Naphthalene-Cyclohexane	1	4.32	1.98	2.82
43	Naphthalene-Hexadecane	1	2.39	5.36	4.40
44	Phenanthrene-Cyclohexane	9	13.63	11.38	12.3
45	Anthracene-Toluene	4	0.51	0.14	0.08
46	Anthracene-Heptane	7	0.13	0.05	0.07
47	trans-Stilbene-Methylcyclohexane	1	1.56	0.89	1.25
48	trans-Stilbene-Hexane	1	1.34	0.78	1.07
49	trans-Stilbene-Cyclohexane	1	1.93	1.25	1.65
50	trans-Stilbene-Octane	1	0.95	0.40	0.67
51	trans-Stilbene-Heptane	1	1.14	0.58	0.86
52	trans-Stilbene-2,2,4-Trimethylpentane	1	1.23	0.71	0.96
53	trans-Stilbene-Cyclooctane	1	0.69	0.04	0.38
	Average		**2.78**	**2.20**	**2.26**

In addition, 134 SLE systems have been added to regress parameters with respect to the UNIFAC-CI (VLE/SLE) models. As shown in Table 5.12, when SLE systems are included in the parameter estimation, the deviations for the UNIFAC-CI (VLE/SLE) has slightly reduced from 2.4 to 2.2 %.

The correlation errors obtained are similar to the reference model (1.9 %). Among the SLE systems used for parameter regression, the problematic systems (AARD2 > 10 %) identified for the CI-models involve naphthalene-alcohols, ibuprofen-alkane, epsilon caprolactone-aromatics, ibuprofen-aromatic and pyrene-1,4-dioxane. For these systems, the correlation errors have been slightly reduced when they are used together with the VLE systems for regression. For the first three systems described, the same trend is also observed for the Original UNIFAC model.

Table 5.12: Correlation Results for C-O, O-C Atoms Interactions Related SLE Systems.

No.	SLE Systems	No. of Data Points	AARD2 (%)		
			Original UNIFAC	UNIFAC-CI (VLE/SLE)	UNIFAC-CI (VLE) with Q_{VLE}
1	Benzoic Acid-Benzene	1	0.66	5.28	4.29
2	Naphthalene-Ethanol	9	9.84	8.95	11.08
3	Naphthalene-2-Propanol	12	9.47	8.61	11.63
4	Naphthalene-Acetone	12	2.13	2.92	2.04
5	Naphthalene-1-Propanol	11	11.33	10.42	13.51
6	Naphthalene-1-Butanol	19	14.03	12.84	16.56
7	Naphthalene-1-Pentanol	1	1.62	1.73	2.86
8	Camphor-Hexane	1	7.69	7.90	7.74
9	Biphenyl-2-Ethyl-1-Hexanol	1	1.19	0.40	0.99
10	Biphenyl-2-Methyl-1-Pentanol	1	1.00	0.65	0.85
11	Biphenyl-4-Methyl-2-Pentanol	1	2.14	1.78	0.27
12	Biphenyl-1-Hexanol	1	0.41	0.76	2.26
13	Biphenyl-1-Heptanol	1	0.56	1.14	2.59
14	Biphenyl-1-Octanol-2	1	0.33	1.12	2.50
15	Biphenyl-3-Methyl-1-Butanol	1	1.22	1.09	0.42
16	Biphenyl-DibutylEther	1	1.74	1.34	2.17
17	Biphenyl-2-Pentanol	1	0.36	0.23	1.28
18	Biphenyl-2-Propanol	1	0.47	0.75	0.59
19	Biphenyl-1-Propanol	1	0.65	0.37	1.70
20	Biphenyl-1-Butanol	1	0.34	0.26	1.72
21	Biphenyl-1-Pentanol	1	0.72	0.86	2.36
22	Biphenyl-t-Butanol	1	0.58	0.61	0.61
23	Biphenyl-2-Methyl-2-Butanol	1	1.04	1.19	2.49
24	Biphenyl-2-Methyl-1-Propanol	1	1.57	1.65	0.19
25	Biphenyl-2-Butanol	1	0.47	0.56	0.91
26	Acenaphthene-2-Ethyl-1-Hexanol	1	0.02	0.18	0.13
27	Acenaphthene-2-Methyl-1-Pentanol	1	0.39	0.50	0.24
28	Acenaphthene-4-Methyl-2-Pentanol	1	0.75	0.87	0.61

29	Acenaphthene-Tetrahydrofuran	1	0.68	0.10	0.27
30	Acenaphthene-1-Hexanol	1	0.65	0.53	0.79
31	Acenaphthene-1-Heptanol	1	0.79	0.63	0.92
32	Acenaphthene-1-Octanol	1	0.73	0.52	0.83
33	Acenaphthene-3-Methyl-1-Butanol	1	0.37	0.45	0.22
34	Acenaphthene-ButylAcetate	1	0.77	5.22	4.73
35	Acenaphthene-1,4-Dioxane	1	0.96	1.21	0.94
36	Acenaphthene-Ethyl Acetate	1	0.98	4.83	4.11
37	Acenaphthene-Di-n-Butyl Ether	1	1.13	0.13	0.67
38	Acenaphthene-2-Pentanol	1	0.28	0.36	0.13
39	Acenaphthene-Ethanol	1	0.12	0.10	0.20
40	Acenaphthene-2-Propanol	1	0.21	0.24	0.09
41	Acenaphthene-1-Propanol	1	0.16	0.13	0.27
42	Acenaphthene-1-Butanol	1	0.25	0.20	0.39
43	Acenaphthene-1-Pentanol	1	0.47	0.39	0.62
44	Acenaphthene-tert-Butanol	1	0.09	0.13	0.02
45	Acenaphthene-2-Methyl-2-Butanol	1	0.50	0.44	0.63
46	Acenaphthene-2-Methyl-1-Propanol	1	0.44	0.49	0.30
47	Acenaphthene-2-Butanol	1	0.26	0.31	0.12
48	Acenaphthene-Methyl Ethyl Ketone	1	2.07	1.42	2.89
49	Fluorene-1-Octanol	1	0.15	0.33	0.71
50	Ibuprofen-Cyclohexane	1	1.47	8.63	7.73
51	Ibuprofen-Heptane	1	3.16	10.89	9.77
52	Ibuprofen-Benzene	1	21.75	19.28	20.31
53	Phenanthrene-2-Ethy-1-Hexanol	1	1.30	0.94	0.38
54	Phenanthrene-2-Methy-1-Pentanol	1	1.29	1.12	0.55
55	Phenanthrene-4-Methy-2-Pentanol	1	1.35	1.18	0.61
56	Phenanthrene-Cyclohexanone	1	2.59	5.05	3.81
57	Phenanthrene-Tetrahydrofuran	1	0.91	0.22	0.02
58	Phenanthrene-1-Hexanol	1	0.04	0.12	0.69
59	Phenanthrene-1-Heptanol	1	0.31	0.58	1.15
60	Phenanthrene-1-Octanol	1	1.26	1.62	2.18
61	Phenanthrene-3-Methyl-1-Butanol	1	0.91	0.84	0.29
62	Phenanthrene-Butyl Acetate	1	0.31	4.09	1.25
63	Phenanthrene-1,4-Dioxane	1	3.19	8.48	8.58
64	Phenanthrene-EthylAcetate	1	3.47	2.93	0.68
65	Phenanthrene-Di-n-Butyl Ether	1	4.36	5.26	5.84
66	Phenanthrene-2-Pentanol	1	0.76	0.68	0.13
67	Phenanthrene-Ethanol	1	0.31	0.19	0.53
68	Phenanthrene-2-Propanol	1	0.39	0.47	0.02
69	Phenanthrene-1-Propanol	1	0.00	0.08	0.37

70	Phenanthrene-1-Butanol	1	0.16	0.17	0.34
71	Phenanthrene-1-Pentanol	1	0.01	0.06	0.61
72	Phenanthrene-2-Methyl-2-Butanol	1	0.22	0.15	0.31
73	Phenanthrene-2-Methyl-1-Propanol	1	0.92	0.94	0.42
74	Phenanthrene-2-Butanol	1	0.76	0.78	0.26
75	Phenanthrene-Methyl Ethyl Ketone	1	3.91	5.63	4.18
76	Anthracene-2-Ethyl-1-Hexanol	1	0.00	0.02	0.03
77	Anthracene-2-Methyl-1-Pentanol	1	0.00	0.00	0.02
78	Anthracene-4-Methyl-2-Pentanol	1	0.02	0.02	0.00
79	Anthracene-2-Methoxyethanol	1	0.17	0.10	0.17
80	Anthracene-2-Ethoxyethanol	1	0.22	0.13	0.21
81	Anthracene-2-Buthoxyethanol	1	0.25	0.04	0.21
82	Anthracene-1-Octanol	1	0.08	0.09	0.11
83	Anthracene-3-Methyl-Butanol	1	0.01	0.01	0.01
84	Anthracene-1,4-Dioxane-6	6	0.59	2.56	2.63
85	Anthracene-Di-n-Butyl Ether	1	0.06	0.19	0.21
86	Anthracene-Methyl-t-Butyl Ether	1	0.15	0.58	0.58
87	Anthracene-2-Pentanol	1	0.00	0.00	0.02
88	Anthracene-2-Propanol	13	0.01	0.01	0.01
89	Anthracene-Di-n-PentylEther	1	0.09	0.17	0.19
90	Anthracene-1-Propanol	1	0.01	0.01	0.03
91	Anthracene-1-Butanol	1	0.02	0.02	0.03
92	Anthracene-1-Pentanol	1	0.03	0.03	0.05
93	Anthracene-2-Methyl-1-Propanol	1	0.02	0.02	0.00
94	Anthracene-2-Butanol	1	0.01	0.01	0.01
95	Pyrene-4-Methyl-2-Pentanol	1	0.16	0.12	0.03
96	Pyrene-3-Methyl-1-Butanol	1	0.53	0.43	0.29
97	Pyrene-ButylAcetate	1	0.53	2.17	1.15
98	Pyrene-1,4-Dioxane	1	5.91	11.90	12.07
99	Pyrene-2-Pentanol	1	0.02	0.04	0.17
100	Pyrene-Ethanol	1	0.14	0.11	0.19
101	Pyrene-2-Propanol	1	0.03	0.05	0.06
102	Pyrene-Acetone	1	0.33	0.05	0.48
103	Pyrene-1-Propanol	1	0.11	0.09	0.20
104	Pyrene-1-Butanol	1	0.16	0.16	0.28
105	Pyrene-1-Pentanol	1	0.31	0.33	0.46
106	Pyrene-2-Methyl-2-Butanol	1	0.10	0.12	0.23
107	Pyrene-2-Methyl-1-Propanol	1	0.15	0.15	0.03
108	Pyrene-2-Butanol	1	0.03	0.04	0.09
109	Epsilon Caprolactone-Benzene	3	29.61	30.45	29.74
110	Epsilon Caprolactone-Toluene	12	23.58	29.04	22.06

111	p-Toluic Acid-Cyclohexane	1	0.20	0.92	0.79
112	p-Toluic Acid-Benzene	1	1.27	2.45	2.51
113	Naphthalene-tert-Butanol	9	11.79	11.56	14.04
114	Naphthalene-2-Methyl-1-Propanol	9	9.94	9.25	12.86
115	Naphthalene-2-Butanol	8	10.87	10.23	13.82
116	Ibuprofen-Toluene	5	5.18	7.01	6.50
117	Anthracene-Methyl Ethyl Ketone	5	0.15	0.22	0.16
118	trans-Stilbene-Ethanol	1	0.11	0.08	0.11
119	trans-Stilbene-2-Propanol	1	0.09	0.12	0.07
120	trans-Stilbene-1-Propanol	1	0.03	0.01	0.05
121	trans-Stilbene-1-Butanol	1	0.00	0.01	0.03
122	trans-Stilbene-1-Pentanol	1	0.01	0.01	0.03
123	trans-Stilbene-2-Methyl-1-Propanol	1	0.21	0.22	0.18
124	trans-Stilbene-2-Butanol	1	0.16	0.17	0.12
125	trans-Stilbene-2-Ethyl-1-Hexanol	1	0.32	0.27	0.27
126	trans-Stilbene-2-Methyl-1-Pentanol	1	0.26	0.25	0.22
127	trans-Stilbene-4-Methyl-2-Pentanol	1	0.36	0.35	0.32
128	trans-Stilbene-1-Hexanol	1	0.02	0.00	0.02
129	trans-Stilbene-1-Heptanol	1	0.07	0.10	0.12
130	trans-Stilbene-1-Octanol-2	1	0.07	0.12	0.12
131	trans-Stilbene-3-Methyl-1-Butanol	1	0.20	0.20	0.16
132	trans-Stilbene-Di-n-Butyl Ether	1	1.17	1.55	1.93
133	trans-Stilbene-Methyl-t-Butyl Ether	1	1.72	3.43	3.74
134	trans-Stilbene-2-Pentanol	1	0.17	0.17	0.13
	Average		**1.85**	**2.25**	**2.36**

Table 5.13 shows the correlation error with respect to O-O interaction related systems with the average errors obtained for the Original UNIFAC, UNIFAC-CI (VLE/SLE) and UNIFAC-CI (VLE) with Q_{VLE} of 9, 12 and 13 % respectively. We can see that when SLE systems were added to regress the AIPs the correlation error reduced a bit from 13 to 12 though still somewhat higher than the reference model. Several of the systems listed below are problematic in the sense that the AARD2 is more than 10 %. There are also many problematic systems for the Original UNIFAC model. Most of the SLE used in this part are involve acids especially benzoic acid, salicylic acid and ibuprofen which are considered to be difficult to correlate.

Table 5.13: Correlation Results for O-O Atoms Interactions Related SLE Systems.

No.	SLE Systems	No. of Data Points	AARD2 (%)		
			Original UNIFAC	UNIFAC-CI (VLE/SLE)	UNIFAC-CI (VLE) with Q_{VLE}
1	Benzoic Acid-2-Ethoxyethanol	1	12.55	20.16	24.57
2	Benzoic Acid-1-Hexanol	1	9.49	10.94	13.61
3	Benzoic Acid-1-Octanol	1	3.29	3.63	7.09
4	Benzoic Acid-Dioxane	1	7.89	16.04	20.24
5	Benzoic Acid-Ethanol	1	8.98	14.20	13.82
6	Benzoic Acid-Acetic Acid	1	1.14	5.64	4.42
7	Benzoic Acid-2-Propanol	1	10.03	14.09	14.77
8	Benzoic Acid-Acetone	1	1.86	5.67	3.32
9	Benzoic Acid-1-Propanol	1	8.61	12.65	13.33
10	Benzoic Acid-1-Butanol	1	10.71	13.72	15.22
11	Benzoic Acid-1-Pentanol	1	8.87	11.03	13.17
12	Benzoic Acid-2-Methyl-1-Propanol	1	5.76	8.78	10.28
13	Benzoic Acid-PropionicAcid	1	4.40	7.01	6.31
14	Benzoic Acid-Gamma Butyrolactone	1	1.28	3.80	15.33
15	Benzoic Acid-Acetophenone	1	6.46	2.79	0.05
16	Salicyclic Acid-Benzyl Alcohol	1	9.90	2.34	7.04
17	Salicyclic Acid-2-Ethoxyethanol	1	1.70	7.89	19.92
18	Salicyclic Acid-1-Octanol	1	7.74	15.47	17.99
19	Salicyclic Acid-Ethanol	1	6.53	8.60	9.55
20	Salicyclic Acid-2-Propanol	1	0.10	13.09	14.18
21	Salicyclic Acid-1-Pentanol	1	0.35	10.63	12.50
22	Salicyclic Acid-tert-Butanol	1	5.94	17.24	18.58
23	Salicyclic Acid-2-Methyl-1-Propanol	1	2.42	9.08	10.62
24	Salicyclic Acid-2-Butanol	1	1.97	13.47	15.01
25	Camphor-Tetrahydrofuran	1	14.77	13.50	13.51
26	Camphor-2-Ethoxyethanol	1	20.43	10.08	21.74
27	Camphor-Ethyl Acetate	1	13.01	15.18	13.28
28	Camphor-1-Propanol	1	20.19	14.40	16.77
29	Camphor-Gamma Butyrolactone	1	12.90	0.96	2.70
30	Ibuprofen-Ethyl Acetate	1	1.72	28.60	28.96
31	Ibuprofen-Acetic Acid	1	13.93	12.23	12.79
32	Ibuprofen-Acetone	1	2.50	0.53	1.83
33	Ibuprofen-Propionic Acid	1	5.57	3.31	3.70
34	Epsilon Caplolactone-1-Propanol	10	51.99	54.24	31.56

35	Epsilon Caplolactone-2-Pentanone	11	26.36	9.63	11.13
36	Benzoic Acid-1-Heptanol	1	7.40	8.26	11.35
37	Ibuprofen-Acetone	5	8.87	12.22	9.63
38	Ibuprofen-4-Methyl-2-Pentanone	5	8.57	13.26	10.49
39	Ibuprofen-EthylAcetate	5	9.66	16.65	16.97
	Average		**9.12**	**11.82**	**13.01**

For the nitrogenated mixtures, a total of 15 SLE systems were added to regress the AIPs. In Table 5.14 for C-N, N-C related atom interaction the correlation errors for both CI-models are the same (4.4 %) suggesting that the use of SLE system in the parameter regression only give a very small effect on the correlation errors of the VLE systems. The average error for the reference UNIFAC model is lower compared to the CI-models (2.1 %). Furthermore, for the systems involving O-N, N-O atom interactions (Table 5.15), the inclusion of SLE data has significantly reduced the correlation errors of the CI-models from 19 to 10 % which are also lower than the Original UNIFAC model with 15 %. Moreover, for the N-N interactions related system (Table 5.16), only 1 SLE dataset was used and the inclusion of this dataset does not seem to reduce the correlation error. However, the average errors of the CI-models are better compared to the reference model.

Table 5.14: Correlation Results for C-N, N-C Atoms Interactions Related SLE Systems.

No.	SLE Systems	No. of Data Points	AARD2 (%)		
			Original UNIFAC	UNIFAC-CI (VLE/SLE)	UNIFAC-CI (VLE) with Q_{VLE}
1	Pyrene-Pyridine	5	4.67	8.61	8.45
2	Phenanthrene-Acetonitrile	1	2.49	2.55	2.65
3	Anthracene-Pyridine	6	0.87	3.74	3.74
4	Naphthalene-Pyridine	5	4.26	4.41	4.36
5	Acenaphthene-Pyridine	5	1.26	4.21	4.34
6	Biphenyl-Pyridine	6	0.34	3.35	3.30
7	Fluorene-Pyridine	5	2.28	7.64	7.62
8	Phenanthrene-Pyridine	10	2.36	5.08	5.03
9	Anthracene-Acetonitrile	5	0.06	0.12	0.12
	Average		**2.07**	**4.41**	**4.40**

Table 5.15: Correlation Results for O-N, N-O Atoms Interactions Related SLE Systems.

No.	SLE Systems	No. of Data Points	AARD2 (%)		
			Original UNIFAC	UNIFAC-CI (VLE/SLE)	UNIFAC-CI (VLE) with Q_{VLE}
1	Benzoic Acid-N-Methylformamide	1	18.31	7.17	31.65
2	Benzoic Acid-N,N-Dimethylacetamide	1	36.48	25.34	49.82
3	4-Nitroaniline-Acetone	1	3.64	12.30	10.22
4	Naphthalene-Nitrobenzene	7	6.32	2.86	3.55
5	BenzoicAcid-Nitrobenzene	1	9.28	0.06	1.12
Average			**14.81**	**9.55**	**19.27**

Table 5.16: Correlation Results for N-N Atoms Interactions Related SLE System.

No.	Systems	No. of Data Points	AARD2 (%)		
			Original UNIFAC	UNIFAC-CI (VLE/SLE)	UNIFAC-CI (VLE) with Q_{VLE}
1	p-Nitroaniline-Acetonitrile	1	11.09	5.25	5.22

For the chlorinated systems, in Table 5.17 the correlation errors for the C-Cl and Cl-C related systems are presented. The average deviations for the Original UNIFAC is 2.3 % while for the VLE-regressed parameters, the average deviation is higher with 6.7 %. However, when SLE systems are used together with the VLE data to regress the AIPs, the average error is still more or less the same. Problematic systems which obtained AARD2 > 10 % are phenanthrene-carbon tetrachloride and naphthalene-1,2-dichloroethane.

Table 5.17: Correlation Results for C-Cl, Cl-C Atoms Interactions Related SLE Systems.

No.	Systems	No. of Data Points	AARD2 (%)		
			Original UNIFAC	UNIFAC-CI (VLE/SLE)	UNIFAC-CI (VLE) with Q_{VLE}
1	Phenanthrene-CarbonTetrachloride	1	9.02	10.63	10.30
2	Pyrene-1,2-Dichloroethane	1	1.14	6.83	6.90
3	Pyrene-1-Chlorobutane	1	0.62	2.69	2.81
4	Pyrene-CarbonTetrachloride	1	2.69	4.02	3.91
5	Naphthalene-1,1-Dichloroethane	7	0.72	2.20	1.84
6	Naphthalene-1,2-Dichloroethane	7	1.56	24.23	24.39
7	Naphthalene-Monochlorobenzene	7	2.20	0.95	0.94

8	Naphthalene-CarbonTetrachloride	10	0.33	2.20	2.13
	Average		**2.28**	**6.72**	**6.65**

For the O-Cl and Cl-O AIPs related systems, Table 5.18 shows that the average error for the reference Original UNIFAC is 9.3 % while for the CI-model using the VLE-regressed parameters, the average deviation obtained was 10 %. However, when SLE data were added the average correlation error was slightly reduced from 10 to 9.5 %.

Table 5.18: Correlation Results for O-Cl, Cl-O Atoms Interactions Related SLE Systems.

No.	Systems	No. of Data Points	AARD2 (%)		
			Original UNIFAC	UNIFAC-CI (VLE/SLE)	UNIFAC-CI (VLE) with Q_{VLE}
1	Benzoic Acid-1,2-Dichloroethane	1	1.41	6.63	6.66
2	Benzoic Acid-Monochlorobenzene	1	2.80	0.50	2.46
3	Benzoic Acid-Chloroform	1	0.08	12.44	11.98
4	Salicyclic Acid-Chloroform	1	10.11	0.43	0.60
5	Ibuprofen-1,2-Dichloroethane	1	18.40	7.09	7.01
6	Ibuprofen-Monochlorobenzene	1	26.09	27.67	29.60
7	Ibuprofen-Chloroform	1	8.92	6.47	9.62
8	Acetylsalicylic Acid-Chloroform	1	12.35	20.59	20.47
9	Ibuprofen-Chloroform	3	3.31	3.37	2.32
	Average		**9.27**	**9.47**	**10.08**

5.2.4.3 Overall Correlation Results

The correlation results reported in Tables 5.2-5.18 are summarized in Table 5.19 and Figure 5.2 for VLE systems and in Table 5.20 and Figure 5.3 for SLE systems according the the types of systems. According to Table 5.19, for the hydrocarbon and oxygenated VLE systems, the addition of SLE systems for the UNIFAC-CI (VLE/SLE) model seems to be reducing the average correlation error from 2.5 % by the UNIFAC-CI (VLE) with Q_{VLE} model to 2.2 %. The same trend can be found for the nitrogenated system but with a very small reduction of the correlation error. However the performance of the CI-models for the nitrogenated system is better than the Original UNIFAC model. However, for the chlorinated systems, addition of the SLE system has slightly increased the average error from 3.1 to 3.4 % with both CI-models performing better than the Original UNIFAC model which has 3.8 % prediction error. Overall we can see that, the inclusion of SLE systems have slightly reduced the correlation error from 2.6 to 2.4 %. When observing Figure 5.2, for most types of system the performance of the CI-models is also better compared to the Original UNIFAC model. Overall however, the differences are small between the three models which is considered a positive result in the development of CI-models (which is based on few data only)

Table 5.19: Overall Correlation Results in AARD1 for VLE Systems.

Types of systems	AARD1 (%)		
	Original UNIFAC	UNIFAC-CI (VLE/SLE)	UNIFAC-CI (VLE) with Q_{VLE}
C-C related systems	1.06	1.52	1.20
C-O, O-C related systems	2.59	2.59	2.59
O-O related systems	2.16	1.86	4.33
Overall hydrocarbon & oxygenated systems	**2.17**	**2.23**	**2.53**
C-N, N-C related systems	3.21	2.31	2.38
O-N, N-O related systems	5.34	2.48	2.43
N-N related systems	1.40	0.94	1.03
Overall nitrogenated systems	**3.61**	**2.10**	**2.13**
C-Cl, Cl-C related systems	2.74	3.57	3.52
O-Cl, Cl-O related systems	5.92	2.89	2.43
Overall chlorinated systems	**3.75**	**3.35**	**3.17**
C-S, S-C, O-S, S-O, Cl-S, S-Cl related systems	1.03	0.70	0.66
Overall sulfurated systems	**1.03**	**0.70**	**0.66**
All systems	**2.72**	**2.40**	**2.55**

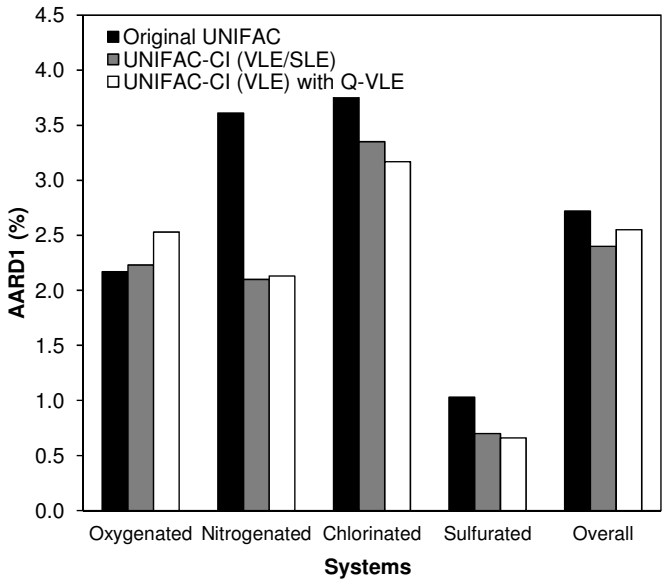


Figure 5.2: Average Correlation Errors of VLE Systems for the UNIFAC-CI (VLE/SLE) Model Compared with the Reference UNIFAC and UNIFAC-CI (VLE) with Q_{VLE} Models for Different Types of Systems.

In Table 5.20, for the hydrocarbon and oxygenated SLE systems, the prediction error using the VLE-regressed CI-model obtained is 4.2 %. However, when the data were added together with the VLE systems in the parameter estimation, the average error reduced to 3.9 %. For the nitrogenated systems, the same trend can be found where the inclusion of SLE systems in the parameter estimation has reduced significantly the average error from 9.4 % when using the UNIFAC-CI (VLE) with Q_{VLE} model to 6.2 %. Furthermore, for the chlorinated systems, the inclusion of those systems with VLE data has also slightly reduced the average error from 8.5 to 8.2 %. Overall, we can see that the addition of SLE data in order to regress the AIPs has reduced to 4.3 % from 4.8 when using the VLE-based CI-model.

From Figure 5.3, we can clearly see that for all types of systems, when the SLE systems are added to regress the AIPs, there is a slight increase of performance compared to the prediction using VLE-regressed only parameters of the UNIFAC-CI models. Although the performance of the Original UNIFAC is better than the UNIFAC-CI (VLE/SLE) models, the performance of the CI-model using the parameters regressed simultaneously against VLE and SLE data can be considered as acceptable. In addition to that, it is not the objective of the CI-models to be better than the reference model but to increase the application range of UNIFAC to predict SLE systems whenever there are missing parameters. However, individually analysis of the system need to be done to see the limitations and problems of the regressed parameters, which will be highlighted in the next sections.

Table 5.20: Overall Correlation Results in AARD2 for SLE Systems.

Types of systems	AARD2 (%)		
	Original UNIFAC	**UNIFAC-CI (VLE/SLE)**	**UNIFAC-CI (VLE) with Q_{VLE}**
C-C related systems	2.78	2.20	2.26
C-O, O-C related systems	1.85	2.25	2.36
O-O related systems	9.12	11.82	13.01
Overall hydrocarbon & oxygenated systems	**3.32**	**3.89**	**4.17**
C-N, N-C related systems	2.07	4.41	4.40
O-N, N-O related systems	14.81	9.55	19.27
N-N related systems	11.09	5.25	5.22
Overall nitrogenated systems	**6.91**	**6.18**	**9.41**
C-Cl, Cl-C related systems	2.28	6.72	6.65
O-Cl, Cl-O related systems	9.27	9.47	10.08
Overall chlorinated systems	**5.99**	**8.17**	**8.47**
All systems	**3.71**	**4.31**	**4.76**

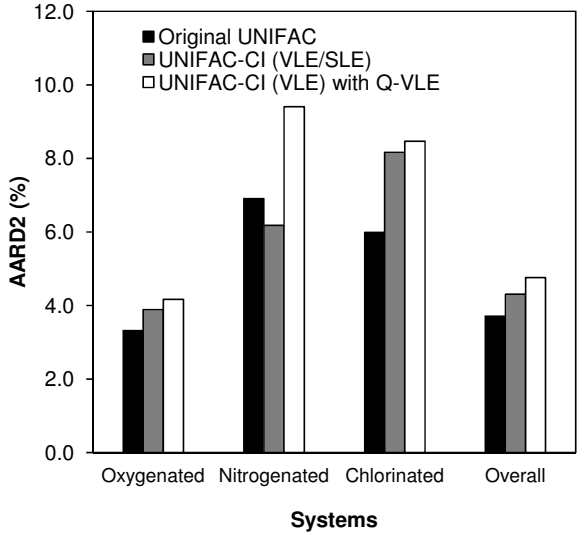


Figure 5.3: Average Correlation Errors of SLE Systems for the UNIFAC-CI (VLE/SLE) Model Compared with the Reference UNIFAC and UNIFAC-CI (VLE) with Q_{VLE} Models for Different Types of Systems.

Furthermore, Figure 5.4 shows the parity plots between each of the experimental data points and the corresponding calculated solute composition which are predicted using the UNIFAC-CI (VLE/SLE) model compared with the reference Original UNIFAC model for hydrocarbons and oxygenated systems. In Figure 5.4 (a) for the hydrocarbons we can see that most of the points are at the diagonal line except for several points for the UNIFAC-CI (VLE/SLE) model and the trends are more or less the same for the Original UNIFAC model. Moreover, for the C-O, O-C related systems displayed in Figure 5.4 (b), most of the data points are concentrated at the lower solute compositions which are also near to the diagonal line for both UNIFAC models. However, there are also data points which are scattered and quite far from the diagonal line.

For the O-O related systems displayed in Figure 5.4 (c), we can see that for the Original UNIFAC model, there are poor predictions for some data points which are far from the diagonal line. However, when the UNIFAC-CI (VLE/SLE) model is used to predict the SLE systems, the problematic points that can be seen for the reference UNIFAC model have been reduced suggesting an improve of performance of the CI-model compared to the Original UNIFAC model. The systems which are considered as problematic are analyzed further in Section 5.3.2. The problematic systems include Napthalene-1-Butanol, Ibuprofen-n-Heptane, Ibuprofen-Benzene, Epsilon Caprolactone-Toluene, Benzoic Acid-2-Propanol and Salicylic Acid-1-Octanol.

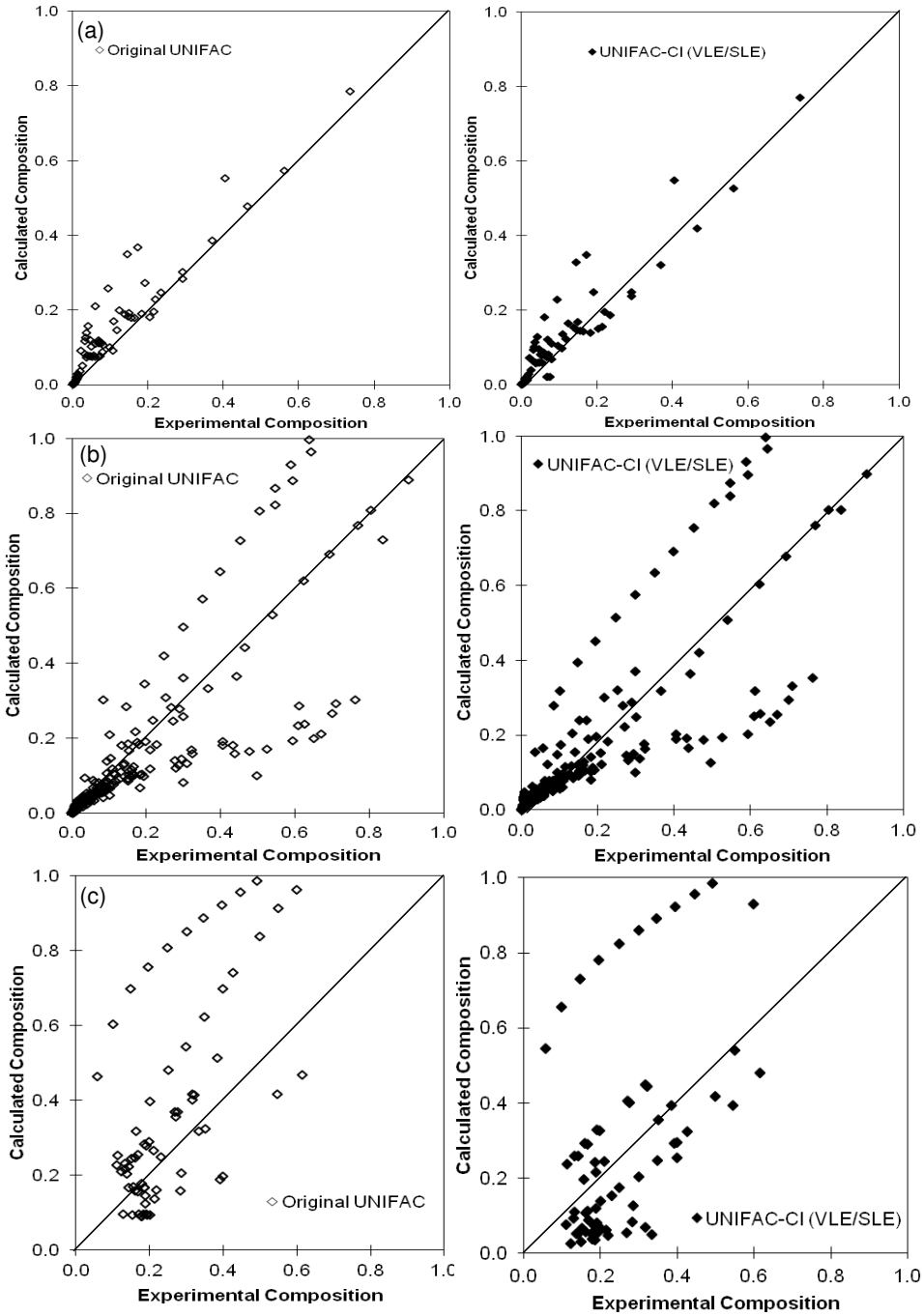


Figure 5.4: Parity Plots Between Experimental and Calculated Solute Composition using Original UNIFAC and UNIFAC-CI (VLE/SLE) Models for (a) Hydrocarbons, (b) Systems with C-O, O-C Atom Interactions and (c) Systems with O-O Atom Interactions.

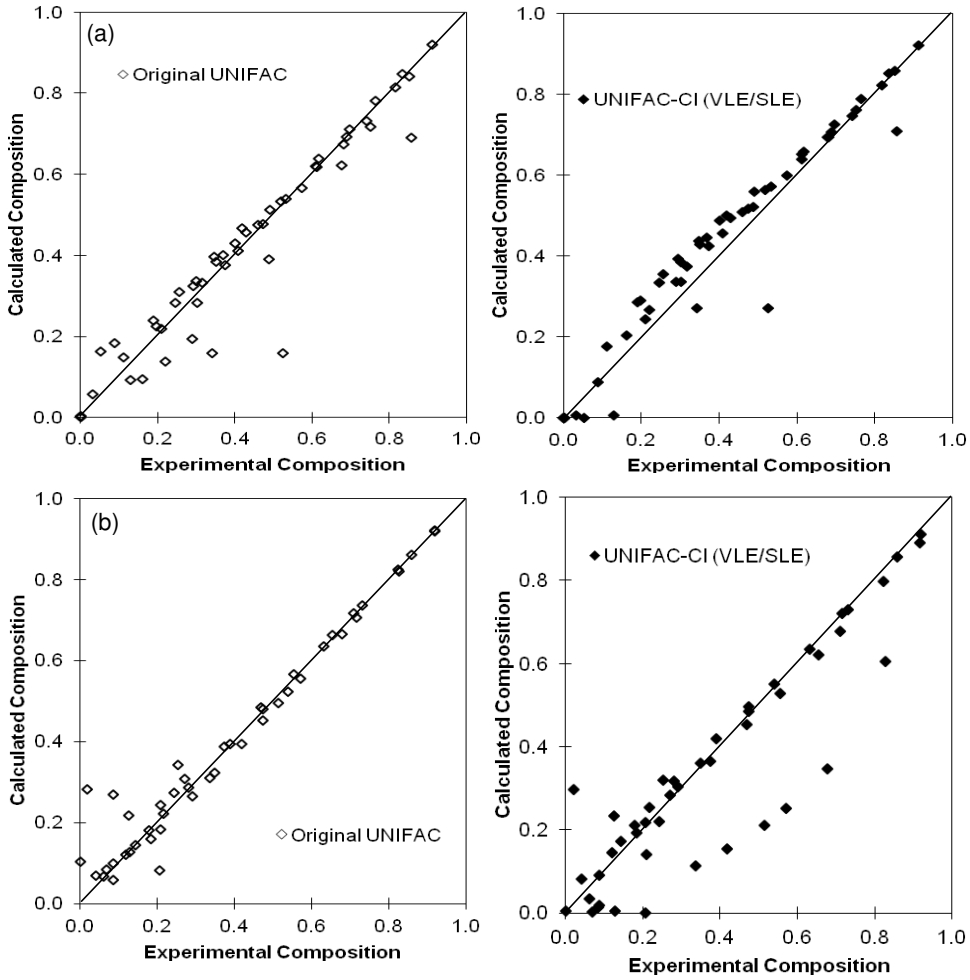


Figure 5.5: Parity Plots Between Experimental and Calculated Solute Composition using Original UNIFAC and UNIFAC-CI (VLE/SLE) Models for (a) Nitrogenated Systems and (b) Chlorinated Systems.

Furthermore, Figure 5.5 shows the same parity plots for nitrogenated and chlorinated systems. From Figure 5.5 (a), we can see that for both Original UNIFAC and UNIFAC-CI (VLE/SLE) model, most of the data points are near to the diagonal line except for some points. Similar trends can also be found for the chlorinated systems by observing the parity plot displayed in Figure 5.5 (b). The problematic systems which deviate from the diagonal lines in the parity plots in 5.5 are further investigated in Section 5.3.2. The problematic systems include Naphthalene-1,2-Dichloroethane, Ibuprofen-Monochlorobenzene and Benzoic Acid-Chloroform.

5.3 Further Analysis of the Original UNIFAC-CI Models

In this section, the work in Section 5.2 will be further analyzed in terms of the individual datasets and the performance of the regressed parameters especially in predicting SLE. Problematic systems for example will be analyzed to identify the possible origin of problems and how they can be resolved. Besides that, other issues will also be discussed which contributed to the correlation results obtained with the CI-models.

5.3.1 Analysis of SLE Predictions using Regressed Parameters

In this section, the parameters regressed in the work described in section 5.2 will be tested by generating SLE diagrams for selected systems. In other words, the diagrams will be generated with the UNIFAC-CI model using the parameters regressed with VLE data alone and will be compared with UNIFAC-CI having parameters regressed simultaneously against VLE and SLE data. Their predictions will also be compared with the reference Original UNIFAC model and assumption of ideal solution.

Figure 5.6 shows the SLE diagrams for 4 binary systems (Biphenyl-2-Ethyl-1-Hexanol, Phenanthrene-Methyl Ethyl Ketone, Naphthalene-2-Propanol and Camphor-gamma-Butyrolactone) which were included in the training set to regress the AIPs parameters. From Figure 5.6 (a) for the Biphenyl-2-Ethyl-Hexanol system, we can see that all 3 models provide a very good prediction of the system as all 3 curves follow closely the single experimental data point. However, if we look closer, the equilibrium curve predicted by the UNIFAC-CI (VLE/SLE) model is the best as it goes exactly through the single data point. For system Phenanthrene-Methyl Ethyl Ketone in Figure 5.6 (b), the SLE curve generated by the Original UNIFAC and UNIFAC-CI (VLE) with Q_{VLE} models are closer to the experimental point compared with the UNIFAC-CI (VLE/SLE) model but it is still acceptable.

For the Naphthalene-2-Propanol system in Figure 5.6 (c), the predictions made by the UNIFAC-CI (VLE/SLE) model follow closely the experimental data points especially at lower concentrations and are better than the other 2 models. Moreover, in Figure 5.6 (d) for the Camphor-gamma-Butyrolactone system, the SLE diagrams generated by all 3 models follow closely the experimental data. However, when investigating the curve at constant temperature (of the experimental data point), in terms of the liquid composition, the predictions made by the UNIFAC-CI (VLE/SLE) model is the closest to the data point.

In addition to the four systems in Figure 5.6, Figure 5.7 shows the SLE diagrams of four systems with two of them having eutectic points. These four systems were not included in the parameter estimation.

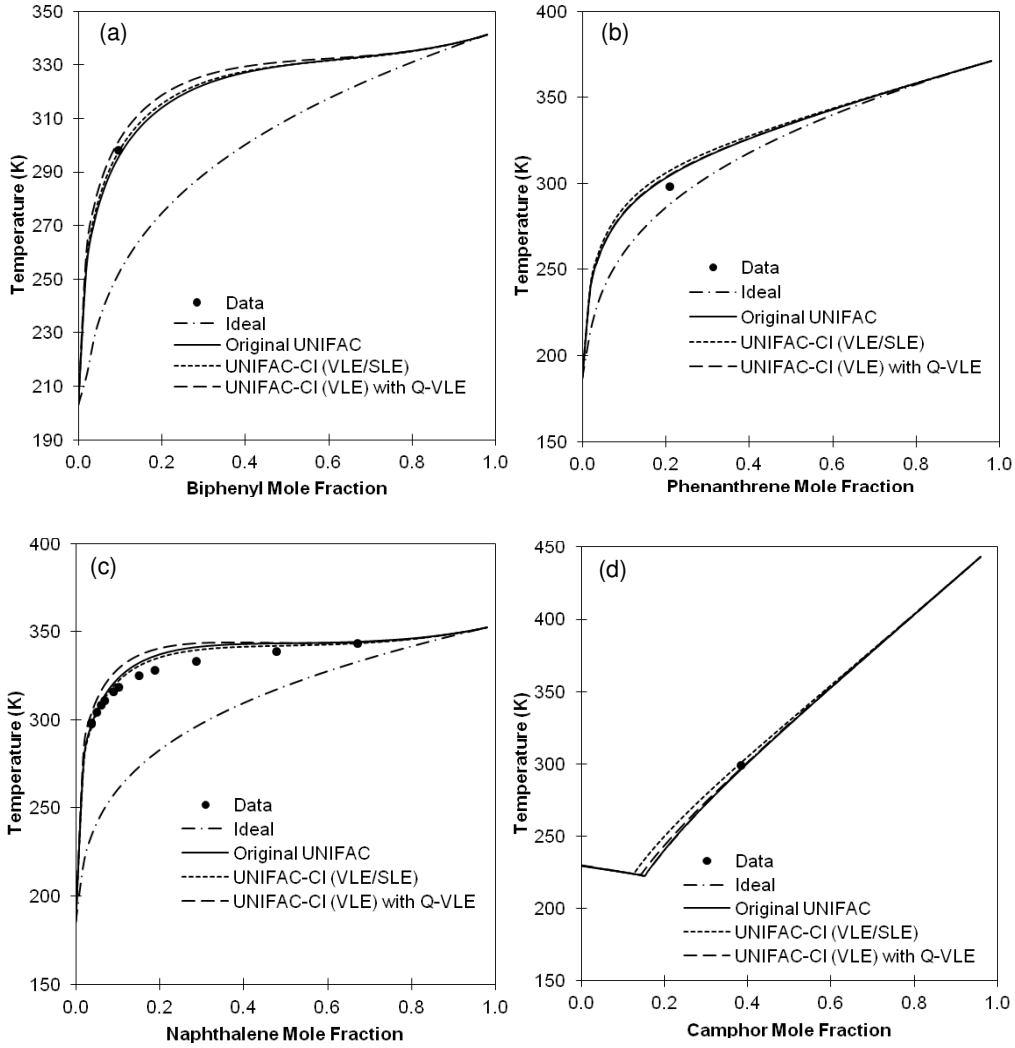


Figure 5.6: SLE Diagrams Predicted using Original UNIFAC, UNIFAC-CI (VLE) with Q_{VLE} and UNIFAC-CI (VLE/SLE) Models for Systems of : a) Biphenyl-2-Ethyl-1-Hexanol, b) Phenanthrene-Methyl Ethyl Ketone, c) Naphthalene-2-Propanol, d) Camphor-gamma-Butyrolactone.

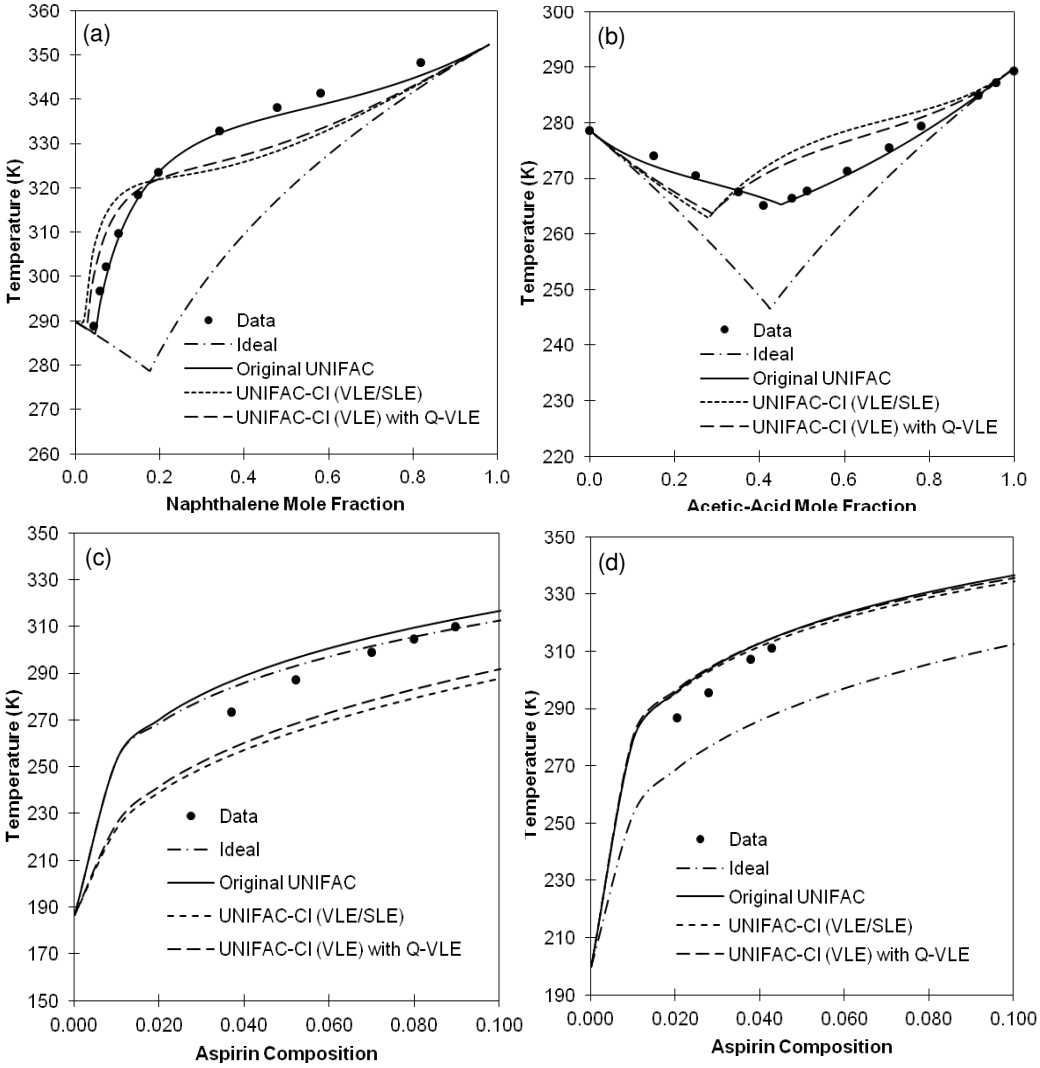


Figure 5.7: SLE Diagrams Predicted using Original UNIFAC, UNIFAC-CI (VLE) with Q_{VLE} and UNIFAC-CI (VLE/SLE) Models for Systems of : a) Naphthalene-Acetic Acid, b) Acetic Acid-Benzene, c) Aspirin-2-Butanone, d) Aspirin-Isopropyl Acetate.

In Figure 5.7 (a) (Naphthalene-Acetic Acid), the Original UNIFAC model is in very good agreement with the experimental data. However, for the CI-models, it can be observed that there are disagreements with the data points especially at concentrations above 0.2 with the UNIFAC-CI (VLE) - Q_{VLE} model being slightly better than UNIFAC-CI (VLE/SLE). On the other hand, Figure 5.7 (b) illustrates the SLE curves for Acetic Acid-Benzene. Here Original UNIFAC provides again better predictions compared to the CI-models. The trends of prediction by the CI-models are the

same with the UNIFAC-CI (VLE) - Q_{VLE} providing a closer prediction of the SLE curve to the experimental data compared to UNIFAC-CI (VLE/SLE). However, both CI-models were unable to capture the eutectic point as well as the reference UNIFAC model.

In addition to that, the prediction of the Aspirin-2-Butanone SLE system in Figure 5.7 (c) shows that the best prediction is made by Original UNIFAC especially at higher concentrations. Nevertheless, the SLE diagrams generated by CI-models show the same trends as the reference UNIFAC model but at a lower temperature range. By comparing the predictions made by both CI-models, the VLE-based model seems to be slightly closer to the experimental data points. For the next system, Aspirin-Isopropyl Acetate illustrated in Figure 5.7 (d), all models (excluding the ideal curve) generated SLE curves which are very close to each other and most importantly also close to the experimental data. When, observing closely we can see that the predictions made by the UNIFAC-CI (VLE/SLE) model is the closest one to the data points.

In addition to the previous 8 SLE systems, more SLE phase diagrams are presented in Figure 5.8 involving nitrogenated and chlorinated systems are tested and analyzed for the regressed atom parameters and compared with the reference model. The first system involved is Anthracene-Acetonitrile displayed in Figure 5.8 (a). We can see that the experimental data are available only at very low concentrations. However, when we zoomed-in more closer (mole fraction range of 0-0.40) we can see that the prediction made by CI-models were closer to the experimental data points compared to Original UNIFAC. The SLE curves generated by the CI-models are also found to be very close to each other. Next, in Figure 5.8 (b) the SLE phase diagram of the system Naphthalene-Nitrobenzene is displayed where it is shown that the system is an ideal system. Predictions by all 3 models showed that the Original UNIFAC model is in close agreement with the experimental data. For the CI-models, we can see that the predictions made by the UNIFAC-CI (VLE/SLE) is closer to the data points followed by the VLE-regressed CI-model. However the eutectic point was not well captured by the CI-models for this system.

On the other hand, Figures 5.8 (c) and (d) illustrate two chlorinated systems involving Naphthalene-1,1-Dichloroethane and Indane-1,2-Dichloroethane respectively. For the first system, we can see that the system is also close to ideal with the best predictions made by the Original UNIFAC model. The predictions made by the CI-models are close to each other and slightly overpredict the experimental data. For the Indane-1,2-Dichloroethane system, the agreement with the experimental data is best for Original UNIFAC especially at higher solute concentration. For the CI-models, the predictions are also quite close to the provided experimental data

From all 12 examples of SLE systems discussed in this section, we can say that adding SLE systems when regressing parameters related to the UNIFAC-CI models improves (a bit) the prediction performance of UNIFAC-CI to estimate phase equilibria related to SLE for some of the systems. For some other systems, the performance of the VLE-regressed CI-model is better or as good as the UNIFAC-CI (VLE/SLE) model.

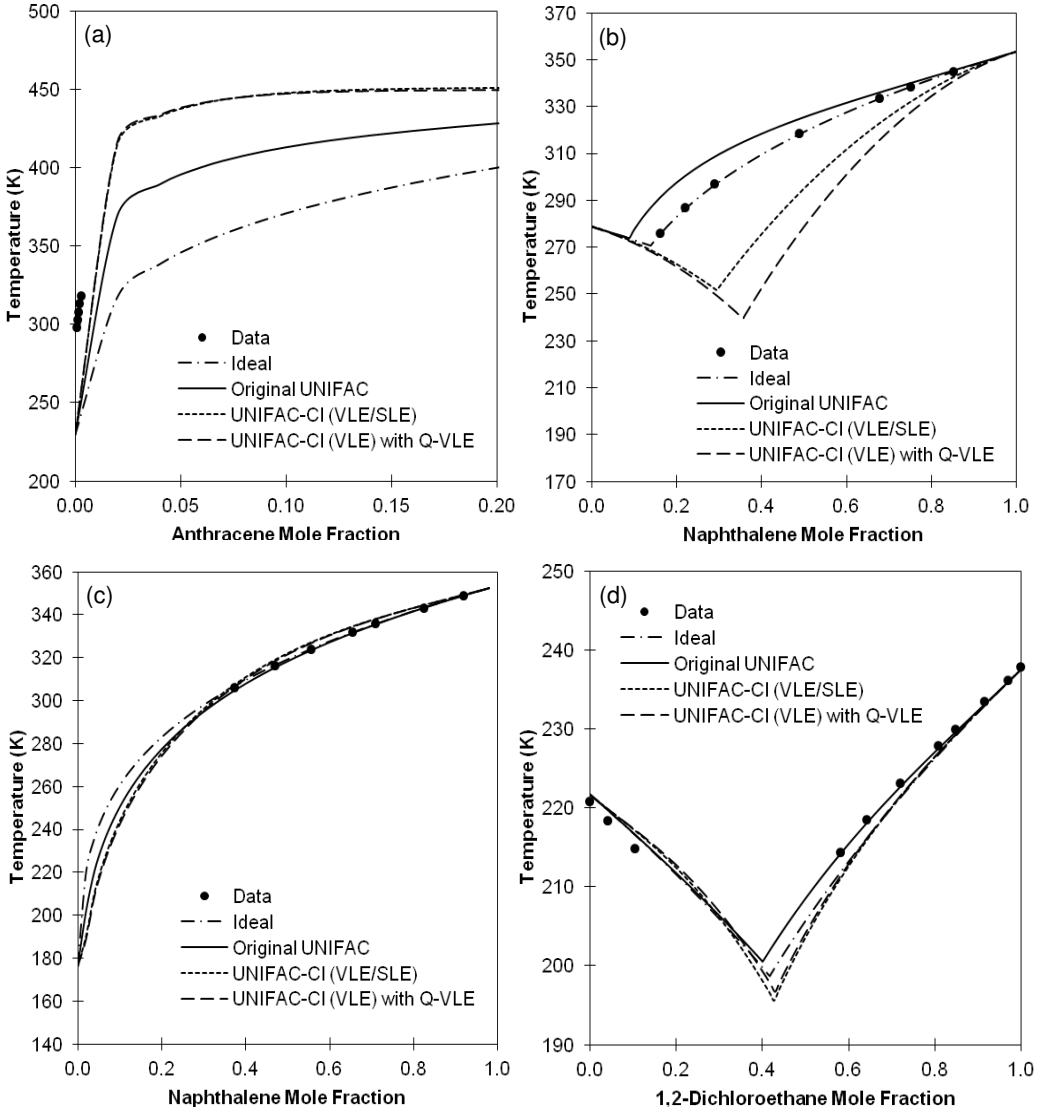


Figure 5.8: SLE Diagrams Predicted using Original UNIFAC, UNIFAC-CI (VLE) with Q_{VLE} and UNIFAC-CI (VLE/SLE) Models for Systems of : a) Anthracene-Acetonitrile, b) Naphthalene-Nitrobenzene, c) Naphthalene-1,1-Dichloroethane, d) Indane-1,2-Dichloroethane.

5.3.2 Further Analysis of Parameter Regression Work

In this section, further analysis of the parameter estimation described in Section 5.2 is carried out in order to discuss some of the problems that arise during parameter regression and also related to some problematic systems that contribute to the high correlation error values obtained.

5.3.2.1 Problematic Systems in Parameter Regression

From the correlation results presented in Tables 5.2-5.18, VLE systems with AARD1 values more than 8 % and SLE systems with AARD2 values more than 10 % are considered to be problematic. The problematic VLE systems involve alkanes-aldehydes, aromatic alcohols-aromatics, ethers-aromatics for oxygenated systems, while for nitrogenated systems the problematic systems are nitriles-alcohols. In addition to that, systems such as alkanes-chloroalkanes, chloroalkanes-aromatics, chloroalkanes-alkenes, chloroaromatics-alkenes, chloroalkanes-esters show high correlation errors. On the other hand, for the SLE systems, aromatics-alcohols, aromatics-carboxylic acids, aromatics-dioxanes, cyclic esters-aromatics, cyclic esters-alcohols, aromatic carboxylic acids-alcohols, camphors-alcohols, camphors-esters, aromatic carboxylic acids-amides anilines-ketones, aromatics-carbon tetrachlorides, aromatics-chloroalkanes, aromatic carboxylic acids-chloroforms and carboxylic acids-chloroaromatics Most systems related carboxylic acids and alcohols are problematic.

For further analysis and discussion, only selected problematic systems with the UNIFAC-CI (VLE/SLE) having AARD1 > 8 % and AARD2 > 10 % are investigated further. Figure 5.9 shows four problematic isothermal VLE systems involving 1,2-Dimethoxyethane-Toluene at 350, Phenol-Styrene at 373 K, Acetonitrile-Ethanol at 393 K and n-Heptane-1,4-Dichlorobutane at 298 K. From the observation of the data points and also the pure component vapor pressures, no clear outliers can be found. However, for the system in Figure 5.9 (a), the predictions with all 3 models are not in good agreement with the experimental data. The AARD1 values obtained for the CI-models are 9.8 and 9.4 % respectively for the UNIFAC-CI (VLE/SLE) and UNIFAC-CI (VLE) with Q_{VLE} models which are close to the reference model with 9.2 %. Using the quality assessment algorithm (described in Section 2.6.5) to check the consistency of the data set, the quality factor, Q_{VLE} obtained is 0.80 which is quite good. Besides that, for the Phenol-Styrene system in Figure 5.9 (b) the predictions made by the CI-models underpredicted the experimental data with AARD1 values of 11.5 and 11.7 %. The prediction made by the Original UNIFAC model is better and close to the experimental data points obtaining prediction error of 4.0 %. The Q_{VLE} value obtained for this data set is 0.93 which is also good. In the parameter regression step, only 1 system (Phenol-Styrene) involving group ACOH is used. Therefore, any use of the parameters related to this group should be used with caution as good quality of prediction is not guaranteed.

Furthermore, for the nitrogenated system of Acetonitrile-Ethanol shown in Figure 5.9 (c), the prediction made by the reference UNIFAC model is in a good agreement with the experimental data except at higher concentration where the azeotrope is slightly underpredicted. However for the CI-models, the data points are overpredicted by both UNIFAC-CI (VLE/SLE) and UNIFAC-CI (VLE) with Q_{VLE} models with AARD1 of 11 and 9 % respectively. The azeotrope are overpredicted

significantly by both models by approximately 50 kPa. The quality factor obtained for this data set is 0.50 since it is a *P,T,x* data which is also good.

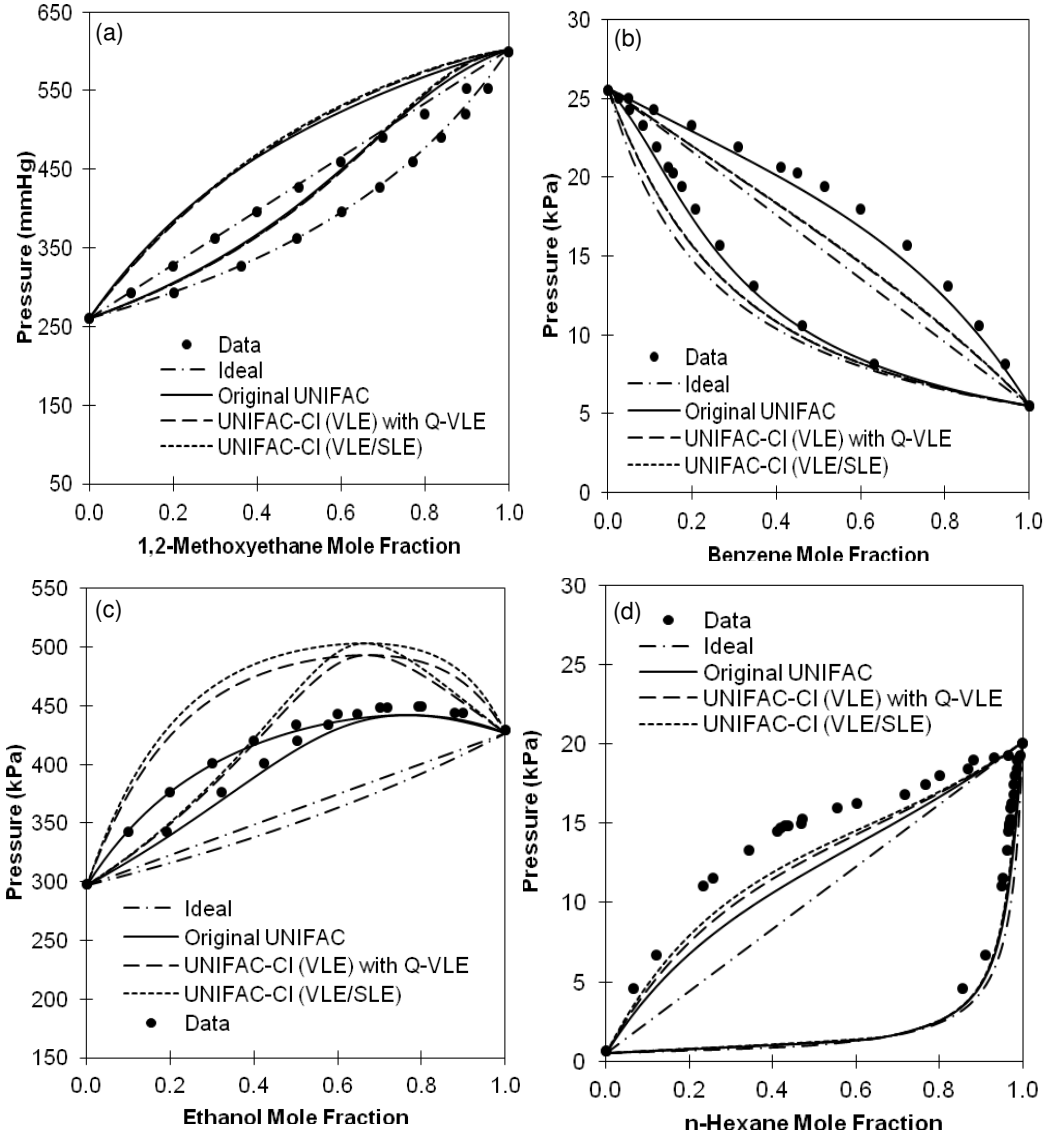


Figure 5.9: Problematic VLE diagrams predicted using Original UNIFAC, UNIFAC-CI (VLE) with Q_{VLE} and UNIFAC-CI (VLE/SLE) models for systems of: (a) 1,2-Dimethoxyethane-Toluene at 350 K, (b) Phenol-Styrene at 373 K, (c) Acetonitrile-Ethanol at 393 K and (d) n-Heptane-1,4-Dichlorobutane at 298 K.

There are 9 systems involving nitriles-alcohols that have been used in the parameter estimation step and only this data set obtained AARD1 > 8 %. In addition to that, for the system of n-Heptane-1,4-Dichlorobutane shown in Figure 5.9 (d), the CI-models underpredcited the data points, especially for the *P-x* line. The AARD1 values obtained for both UNIFAC-CI (VLE/SLE) and UNIFAC-CI (VLE) with Q_{VLE} models are 10 and 12 % respectively. However, the predictions made by the CI-modes are better than for the reference UNIFAC model. The Q_{VLE} value generated for this data set is 0.13 which is quite bad. This is due to some inconsistency of the data points which can be seen by zooming-in at the *P-x* data at the ranges on composition between 0.4-0.5 and 0.8-0.9.

Moreover, Figure 5.10 highlights four more problematic VLE systems involving Dichloroethane-Toulene at 298 K, Dichloromethane-1-Hexene at 298 K, 1-Pentene-Monochlorobenzene at 320 K and Dichloromethane-Ethyl Acetate at 348 K. The predictions by the CI-models in the first system in Figure 5.10 (a) showed more or less same behavior with negative deviations against the experimental data with an average deviation AARD1 around 8.0 % for both UNIFAC-CI (VLE/SLE) and UNIFAC-CI (VLE) - Q_{VLE}. However, the predictions made by the Original model is very good. The Q_{VLE} value obtained for this *P, T, x* data set is 0.25 which is quite low. Furthermore, for the system displayed in Figure 5.10 (b) which obtained a Q_{VLE} of 0.5, the Original UNIFAC model slightly underpredicted the system with AARD1 of 4.3 %. However, the predictions made by the CI-models were unable to capture the data points closely with AARD1 values of 11-13 %. In the regression step, only one system involving alkene-chloroalkane was used. Therefore, predictions using the interaction parameters between groups C=C and CCl_2 should be used with caution.

Moreover, for the chlorinated system of 1-Pentene-Monochlorobenzene in Figure 5.10 (c), a slightly negative deviation can be found for both the CI-models with average pressure deviations of 8-10 %. The prediction made by the Original UNIFAC is very good with AARD1 equals to 2 %. The quality factor generated for this *P, T, x* system is 0.50 which good. Next, for the Dichloromethane-Ethyl Acetate system illustrated in Figure 5.10 (d), the best prediction is obtained by the Original UNIFAC followed by the UNIFAC-CI (VLE) with Q_{VLE} and UNIFAC-CI (VLE/SLE) models with AARD1 of 1.4, 4 and 11 % respectively. The quality factor for this data set is 0.74 which is not bad.

By analyzing the correlation results in Tables 5.2-5.18, there are only several systems with correlation errors above 8 % with respect to the UNIFAC-CI (VLE/SLE) model, and for some of them the results can be considered as acceptable. However, the high deviations of these systems might be because of the quantity of the same type of systems or group-interactions that are used in the parameter estimation are not many. For example, there is only one system of phenol and styrene involving group interactions between ACOH-ACH and ACOH-C=C that was used in the parameter estimation. Therefore, these interaction parameters need to be used with caution when trying to predict other systems since the correlation errors obtained for this system are already quite high. On the other hand, there are also cases where for the same type of systems using the same interaction parameter, for example the nitriles-alcohols systems, some data sets obtained low deviation error while one or two systems obtained high errors. This is maybe due to some errors of inconsistencies in the data sets themselves.

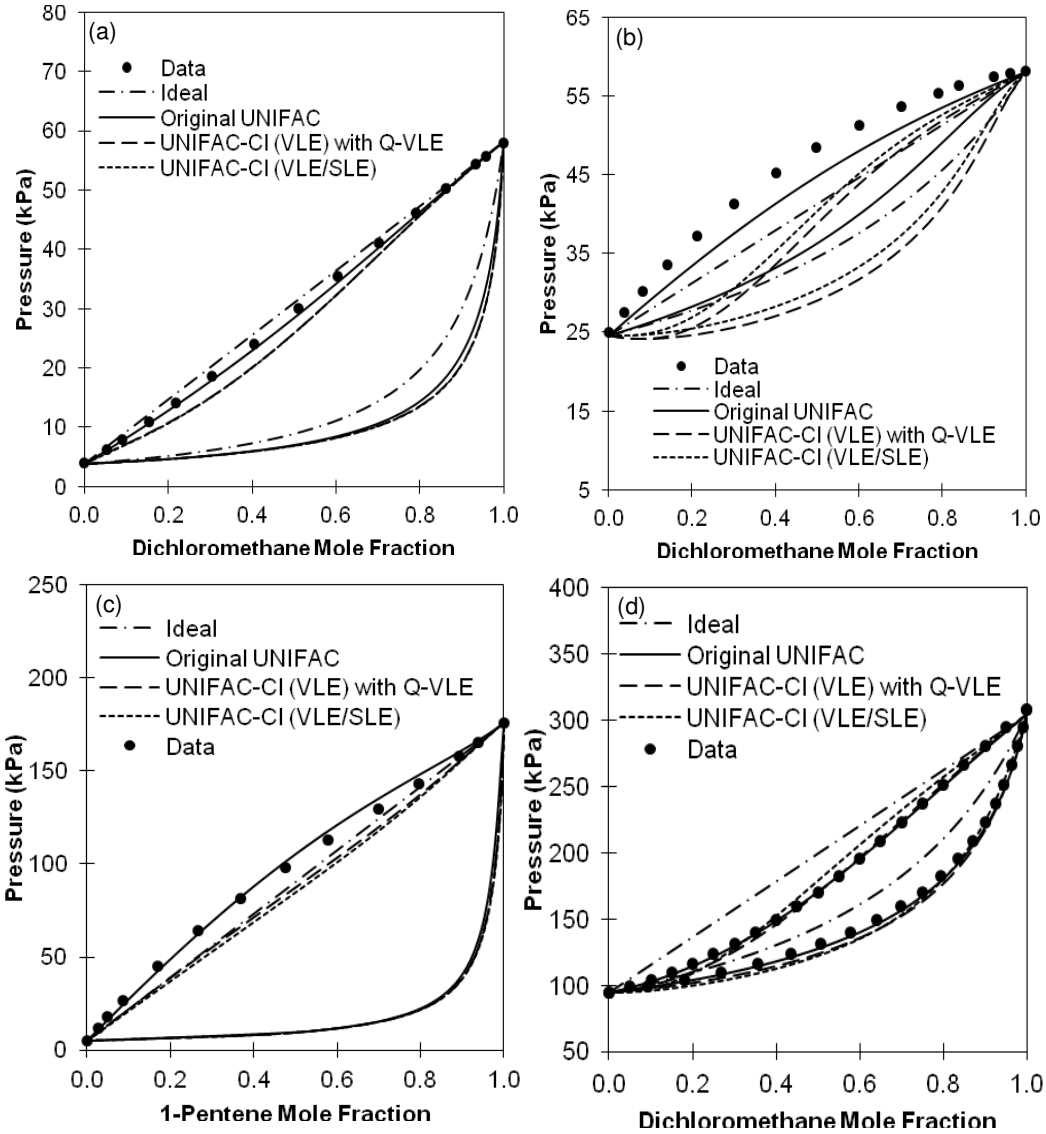


Figure 5.10: Problematic VLE diagrams Predicted using Original UNIFAC, UNIFAC-CI (VLE) with Q_{VLE} and UNIFAC-CI (VLE/SLE) models for systems of: (a) Dichloroethane-Toulene at 298 K, (b) Dichloromethane-1-Hexene at 298 K, (c) 1-Pentene-Monochlorobenzene at 320 K and (d) Dichloromethane-Ethyl Acetate at 348 K.

Next, we discuss the problematic SLE systems with AARD2 > 10 % as shown in Figures 5.11, 5.12 and 5.13. Figure 5.11 shows the SLE diagrams for the systems of Naphthalene-1-Butanol, Ibuprofen-n-Heptane, Ibuprofen-Benzene and Pyrene-1,4-Dioxane. In Figure 5.11 (a), the

predictions by all three UNIFAC models are close with each other obtaining AARD2 values of 14, 13 and 17 % respectively for the Original UNIFAC, UNIFAC-CI (VLE/SLE) and UNIFAC-CI (VLE) - Q_{VLE} models. The predictions are in good agreement with the experimental data except at the range between 0.1-0.5 of the solute concentration. However, the predictions are considered as acceptable and the prediction error has been slightly reduced when this dataset is included in the parameter regression step. There are several other systems with naphthalene and alcohols and their correlation errors are also quite high but still acceptable. Next, in Figure 5.11 (b), the SLE diagrams shown are with respect to Ibuprofen-n-Heptane system. Only 1 data point is available for this system. The system is best predicted by the Original UNIFAC model with AARD2 value of 3 %. However, the correlation errors obtained by the two CI-models are a bit high with 11 and 10 %. We can see from the figure that the CI-models underpredict this single data point.

In addition to that, for the system of Ibuprofen-Benzene in Figure 5.11 (c), the correlation errors obtained are quite high for all three UNIFAC models. We can see that each of the SLE curve generated by all the models have an eutectic point and also quite close to each other. However, the predicted diagrams are quite far from the experimental data. The AARD2 obtained for the UNIFAC models are between 19 to 22 %. In Figure 5.11 (d), there is also one data point available and the SLE curves generated by all these UNIFAC models also have eutectic points. The predictions made by Original UNIFAC is the closest to the data point with AARD2 of 6 %. The predictions made by the CI-models are more or less the same with AARD2 around 12 %.

Furthermore, four more problematic SLE systems that were used in the regression step are displayed in Figure 5.12 (Epsilon Caprolactone-Toluene, Benzoic Acid-1,4-Dioxane, Benzoic Acid-2-Propanol and Salicylic Acid-1-Octanol). Focusing on Figure 5.12 (a) for the system of Epsilon Caprolactone and Toluene, all SLE diagrams showed underestimations of the experimental data points. However, the predictions made by the CI-models are closer to the data points compared to the reference Original UNIFAC model. Besides this system, other system with Epsilon Caprolactone with Benzene was included in the regression step and also showed high deviations with data. Therefore, when predicting this kind of system using the regressed parameters, good predictions are not guaranteed.

Moreover, for the Benzoic Acid-1,4-Dioxane system illustrated in Figure 5.12 (b), only one data point is available and the best predictions is obtained by Original UNIFAC model AARD2 of 8 %. The correlation errors generated for the CI-models are higher (16-20 %). The SLE diagrams generated by the CI-models were unable to capture the eutectic point of the system suggesting why they overpredicted the data point. However, we can see that when this system is included in the parameter estimation, the prediction performance is slightly improved.

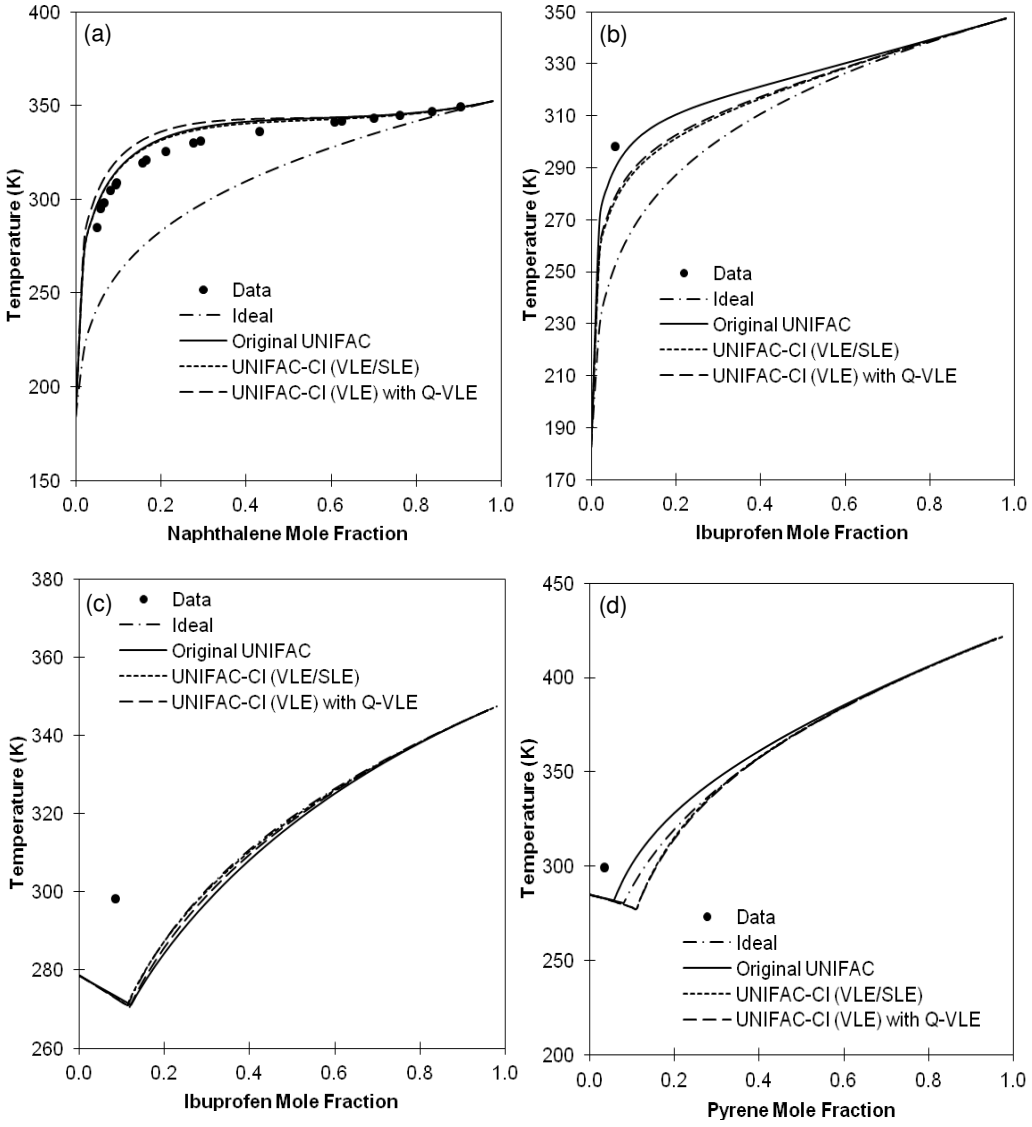


Figure 5.11: Problematic SLE Diagrams Predicted using Original UNIFAC, UNIFAC-CI (VLE) with Q_{VLE} and UNIFAC-CI (VLE/SLE) Models for Systems of: (a) Naphthalene-1-Butanol, (b) Ibuprofen-n-Heptane, (c) Ibuprofen-Benzene and (d) Pyrene-1,4-Dioxane.

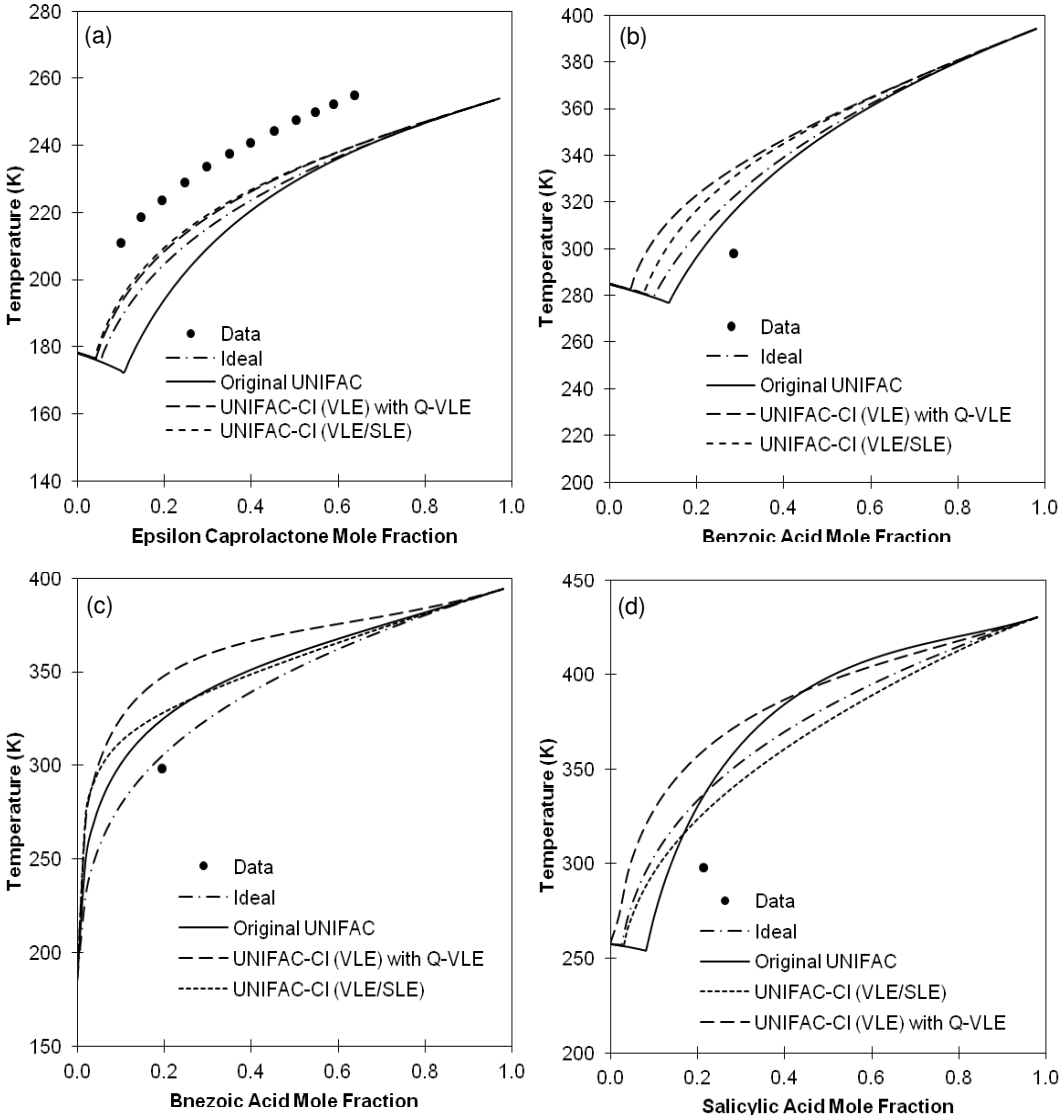


Figure 5.12: Problematic SLE Diagrams Predicted using Original UNIFAC, UNIFAC-CI (VLE) with Q_{VLE} and UNIFAC-CI (VLE/SLE) Models for Systems of: (a) Epsilon Caprolactone-Toluene, (b) Benzoic Acid-1,4-Dioxane, (c) Benzoic Acid-2-Propanol and (d) Salicylic Acid-1-Octanol.

Next in Figure 5.12 (c), with only one data point available, the system of Benzoic Acid-2-Propanol shows a near ideal behavior. All three UNIFAC models were unable to capture the experimental data accurately. The predictions made by Original UNIFAC is the closest to the data with deviation

10.0 %, while the two CI-models have deviations 14-15 %. However, we can observe that the SLE diagram generated by the UNIFAC-CI (VLE/SLE) model is closer to the curve generated by the Original UNIFAC model and also closer to the data point compared to the VLE-regressed CI-model. We can assume that when this data is used together with VLE systems to regress parameters, the performance is slightly improved. Furthermore, for the system of Salicylic Acid-1-Octanol in Figure 5.12 (d), the best performance is again obtained by the Original UNIFAC model generating SLE curve which is closest to the experimental data with deviation error of 8 %. On the other hand, the CI-models obtained correlation errors of 16-18 % respectively for the two CI-models which are quite far from the experimental data. However, we can see that the performance of the UNIFAC-CI (VLE/SLE) model is slightly better than the CI-model using only VLE-regresssed parameters. There are also other systems with Salicyclic Acid-Alcohols that were used in the parameter estimation and most of them also showed high deviation errors. Looking at these trends, we suggest that the regressed parameters to predict this kind of system should be used with caution.

Another four problematic SLE systems are shown in Figure 5.13 (Ibuprofen-Ethyl Acetate, Naphthalene-1,2-Dichloroethane, Ibuprofen-Monochlorobenzene and Acetylsalicylic Acid-Chloroform). In Figure 5.13 (a), there are two sets of experimental for the system Ibuprofen-Ethyl Acetate from different sources. Data 1 has only one data point while Data 2 have five data points and these data sets are not in agreement with each other suggesting that there might be errors or uncertainties when the experiments were done. Nevertheless, for Data 1, the predictions made by the Original UNIFAC model is in good agreement with the single data point with a deviation error of 2 %. The performance of the CI-models is quite poor (29 %). On the other hand, for Data 2, the Original UNIFAC is still the best with an accuracy of 10 % which is higher compared to Data 1. For the CI-models, the deviation errors are still high (17 %). Next, according to Figure 5.13 (b), the system of Naphthalene-1,2-Dichloroethane shows a near ideal behavior and it is well predicted by Original UNIFAC with an accuracy of 2 %. Both CI-models show a high overestimation of the experimental data with correlation errors around 24 %. The SLE curves generated by the CI-models do not produce eutectic points similar to the diagram generated by Original UNIFAC model.

For the problematic SLE systems displayed in Figures 5.11-5.13, the regressed interaction parameters need to be used with caution when trying to predict other SLE systems which are similar and good prediction is not guaranteed. Nevertheless, even for the problematic systems we can see that adding SLE data to regress parameters can slightly improve the predictions of SLE systems.

Moreover, in Figure 5.13 (c) for system Ibuprofen-Monochlorobenzene, all three UNIFAC models were unable to capture the single experimental data point, The deviation errors obtained by the Original UNIFAC, UNIFAC-CI (VLE/SLE) and the UNIFAC-CI (VLE) with Q_{VLE} models are 26, 27 and 21 respectively which are quite high. Furthermore, for the system of Benzoic Acid-Chloroform shown in Figure 5.13 (d), the SLE curve predicted by the Original UNIFAC is in a very good agreement with the single experimental data point. However, the CI-models overpredicted the data point (12 %).

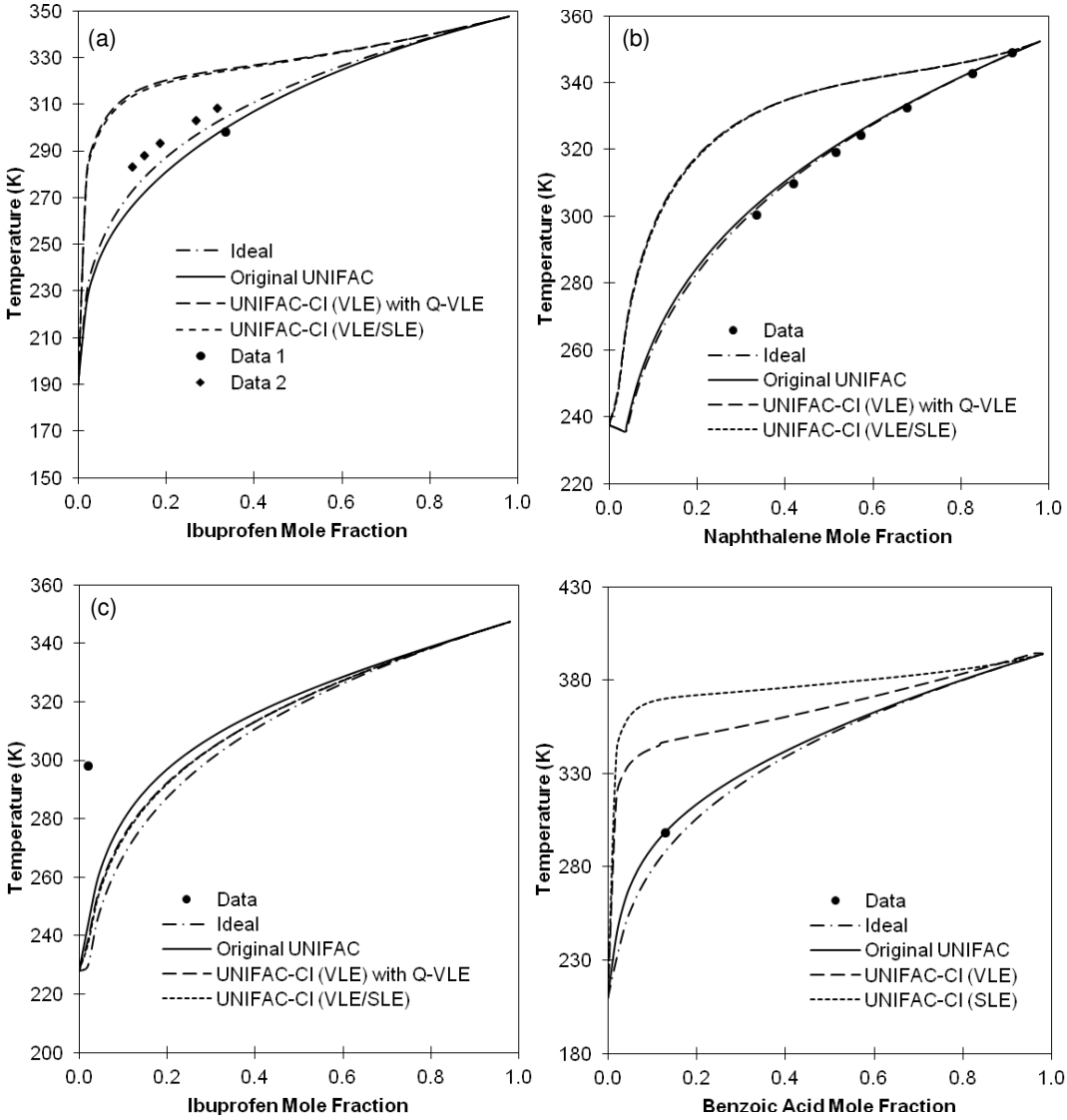


Figure 5.13: Problematic SLE Diagrams Predicted using Original UNIFAC, UNIFAC-CI (VLE) with Q_{VLE} and UNIFAC-CI (VLE/SLE) Models for Systems of: (a) Ibuprofen-Ethyl Acetate, (b) Naphthalene-1,2-Dichloroethane, (c) Ibuprofen-Monochlorobenzene and (d) Benzoic Acid-Chloroform.

5.3.2.2 Discussion on SLE Experimental Data

In this section, we would like to discuss the SLE experimental data that were used in the parameter regression work. Among the 258 SLE datasets that were used, 85 % of the systems have only one data point. The problem with only one data points is that we cannot really see the trends of the experimental data over the whole concentration range and therefore it might also be a problem for the property prediction model to totally capture the behavior of the system and to produce the right set of parameters. However, according to the source from where these data were taken, there are no consistency type tests that were used, but the experimental dat have carefully checked the data and only data which clearly erroneous are excluded.

Besides that, for the SLE experimental data, the only systems used in the parameter estimation are those for which experimental pure component properties (melting temperatures and heats of fusion) are available. The experimental melting points and heats of fusion used for the solutes are presented in Table 5.21.

Table 5.21: Experimental Melting Temperatures and Heats of Fusion used for the Solutes in Generating the SLE Data.

No.	Compound /Solute	Melting Point (K)	Heats of Fusion (kJ/kmol)
1	Benzoic Acid	395.55	18070
2	Salicyclic Acid	431.15	19590
3	4-Hydrobenzoic Acid	487.65	30860
4	m-Toluic Acid	381.85	15690
5	Naphthalene	353.35	18980
6	Biphenyl	342.15	18580
7	Camphor	453.15	5839
8	Acenaphthene	366.55	21462
9	Fluorene	387.95	19578
10	Ibuprofen	349.15	21900
11	Anthraquinone	599.15	32552
12	Phenanthrene	372.35	16463
13	Anthracene	488.15	29370
14	Pyrene	424.35	17360
15	Epsilon Caprolactone	255.15	13800
16	p-Toluic Acid	452.75	22720
17	Trans-Stilbene	396.15	27690

5.4 Conclusions

In this chapter, the atom interaction parameters (AIPs) of the UNIFAC-CI model have been regressed simultaneously against VLE and SLE data. The results of the correlations were

compared with those checked with Original UNIFAC model and the UNIFAC-CI with Q_{VLE} models with parameters regressed against only VLE data. From the overall correlation results displayed in Table 5.20 and Figure 5.3, the addition of the SLE data together with VLE data in the regression of the AIPs has slightly increased the accuracy of predictions related to SLE systems while still maintaining the reliability of predictions of the VLE systems.

The performance of the UNIFAC-CI (VLE/SLE) model to predict SLE systems has been investigated in more detail in Section 5.3.1 and we can see that for some systems, the prediction using the VLE-regressed CI-model is better. However, for most of the systems investigated, there is a slight improvement of performance of the CI-model with the inclusion of SLE system in the parameter estimation step. Moreover, some problematic VLE and SLE systems which obtained high deviation errors have been analyzed in Section 5.3.2. One problem identified is that most of the system have only one data point which makes it hard for UNIFAC to capture the behavior of the system throughout the whole concentration range. Especially during the parameter regression step in order to obtain the best AIPs. In addition to that, the amount of the VLE data and data points are larger compared to the SLE systems. Therefore, we can say that the AIPs are more favourable to the VLE systems suggesting why there is only a slight increase of performance when SLE systems are included. Nevertheless, the results obtained for this work are satisfactory and acceptable.

It is well known that Original UNIFAC can also be used to predict SLE system with high accuracy and this can also be seen in Figures 5.11-5.13 since their group interaction parameters have been regressed extensively against a large database. However, it is the purpose of this work to only use a limited amount of data in the parameter regression. Then, using the regressed parameters for the prediction of other systems which have not been used in parameter regression but at least having the same UNIFAC main groups. The AIPs obtained from this work are presented in Table B.3 in Appendix B.

CHAPTER 6

DEVELOPMENT OF THE MODIFIED (DORTMUND) UNIFAC-CI MODELS: VAPOR-LIQUID EQUILIBRIUM SYSTEMS

6.1 Introduction

In this chapter, we present the further development and analysis of the Modified (Dortmund) UNIFAC-CI model for the prediction of vapor-liquid equilibrium (VLE) systems. Previously in an earlier development of the Modified (Dortmund) UNIFAC-CI model, Gonzalez et al. [18], have regressed the atom interaction parameters (AIPs) against vapor-liquid equilibrium (VLE) experimental data involving systems with C, O and N atoms. In this work, systems with respect to atoms Cl and S are added and the related AIPs are regressed. The objective of this work is to further increase the application range of the developed CI-model in order to be able to predict systems with C, O, N, Cl and S atoms. The result of the correlations in terms of deviations errors of the CI-model are compared with the predictions errors of reference Modified (Dortmund) UNIFAC-CI model. The VLE predictions made by the CI-model for selected systems using the regressed parameters are presented and discussed.

6.2 Development of the Modifed (Dortmund) UNIFAC-CI Model for VLE

The Modified (Dortmund) UNIFAC version published by Gmehling et al. [64] has been chosen by González [18] for the development of the Modified (Dortmund) UNIFAC-CI model. This is the most accessible publication presenting complete information about the Modified UNIFAC model and availability of the GIPs. Recent revisions do exist but they do not give the parameter values for all newly determined GIPs and the subgroup representations of the UNIFAC main groups. The parameter table matrix for this version has 45 main groups and they are slightly different from the version published by Hansen et al. [15]. There are also still many missing parameters in the parameter table. As it was discussed in section 2.3.2, compared to the Original UNIFAC model, this modified model has three GIPs which overall have quadratic dependency towards temperature. This is shown in Equation (6.1).

$$a_{mn,overall} = a_{mn} + b_{mn}T + c_{mn}T^2 \tag{6.1}$$

In order to relate these GIPs with the atom stoichiometry, CIs used to define each main groups and the AIPs, the relationships are defined as in Equations (2.46)-(2.47) discussed in Section 2.5.2.2 for Modified (Dortmund) UNIFAC-CI model for systems with atoms C, H, O and N. However, in this chapter, since systems with Cl and S atoms are added, Equations (2.46)-(2.47) have been

expanded to take into account the CIs and AIPs with respect these two atoms. The derived relationship for the prediction of GIPs of the Modified (Dortmund) UNIFAC-CI for systems involving atoms C, H, O, N, Cl and S are presented in Appendix C.

In the development of Original UNIFAC, only the GIPs are regressed while the group volume (R_k) and surface are (Q_k) are obtained from Bondi [54]. However, in the Modified (Dortmund) UNIFAC model developed by Gmehling et al. [68], besides the GIPs, the group volume (R_k) and surface are (Q_k) are also treated as adjustable parameters and regressed against the experimental data. However, one of the objective of this work is to fill any missing GIPs in the Modified (Dortmund) UNIFAC parameter table. Therefore, in this work, in order to use the same the R_k and Q_k values obtained by Gmehling et al. [68], those values are set to constant and only the AIPs are regressed against the experimental data.

6.2.1 Background of Parameter Estimation

The minimization technique used to regress the parameters in this work is the least squares technique using the Modified Levenberg-Marquardt approach which algorithm are described in section 2.6.3 and [119] for the regression AIPs for the Modified (Dortmund) UNIFAC-CI models. The same technique of parameter regression was used in the previous work [18].

In this work, the parameter regression is based on the *P, T, x_i, y_i* and *P, T, x_i,* VLE data involving sulfurated and chlorinated systems and only isothermal data were used for the regression. The unit and magnitude of the pressures reported in each of the VLE data are different (mmHg, torr, kPa, Pa). The objective function below has been chosen to regress the parameters together with the activity coefficient terms whenever *P, T, x_i, y_i* data are used.

$$OF = \frac{1}{N}\sum_{i=1}^{N}\left[\left(\frac{P_{i-\exp}-P_{i-calc}}{P_{i-\exp}}\right)^2+\left(\frac{\gamma_{i-\exp}-\gamma_{i-calc}}{P_{i-\exp}}\right)^2\right]+w_{reg}\sum_{i=1}^{N}\left(AIP_i-AIP_i^{IG}\right)^2 \qquad (6.1)$$

where P_{exp} is the experimental pressure, $\gamma_{i,\exp}$ is the experimental activity coefficient and N is the number of experimental data points used for the estimation, AIP_j is the current value of the CI-interaction parameter j, AIP^{IG}_j its corresponding initial guess and w_{reg} (values between $1x10^2$ and $1x10^{10}$) is a weighting value used to increase and decrease the influence of regularization in the optimization.

Just like for the Original UNIFAC-CI model, the equilibrium pressure, $P_{i\text{-}calc}$ was calculated in two different ways depending on whether the systems need an association term (systems involving carboxylic acids) or not. For systems without an association term, the pressure is calculated as follows:

$$P_{i-calc} = \sum_i x_i \gamma_i P_i^{sat}(POY_i) \qquad (6.2)$$

where i is an index running over all species in the mixture and POY_i is the Poynting factor. However, for the systems needing the association term, the equilibrium pressure is calculated using fugacity coefficient and association based on the method of Hayden and O'Connell [120] described briefly as follows:

$$P_{i-calc} = \sum_i \frac{x_i \gamma_i P_i^{sat}(POY_i)}{\Phi_i} \tag{6.3}$$

$$\ln \Phi_i = \left[2\sum_{j=1}^{N} y_j B_{ij} - B \right] \frac{P}{RT} \tag{6.4}$$

where Φ_i is the fugacity coefficient, y_j is the vapor mole fraction, $B_{ij}(T)$ is the second virial coefficient characterizing pair interactions between *i* and *j* molecule and B is the second virial coefficient. The cross second virial coefficient, B_{ij} can be calculated directly from PVT data, from statistical mechanical formulas or from empirical and semitheoretical correlations [114]. On the other hand, the second virial coefficient, B are calculated using the equations below which are contributed by the different types of intermolecular forces which are described as bound, metastably bound, free pairs and B_{chem} which include the association contribution.

$$B = B_{free} + B_{metastable} + B_{bound} + B_{chem} \tag{6.5}$$

The bound and metastably bound contribution of the second virial coefficient are calculated in Equation (4.6) while the free pairs contribution is calculated in Equation (4.7):

$$B_{metastable} + B_{bound} = b_0 A \exp\left[\Delta H / kT / \varepsilon\right] \tag{6.6}$$

$$\text{with } b_0 = \frac{2\pi}{3} N_0 \sigma^3,\ A = -0.3 - 0.05\mu^*,\ \Delta H = 1.99 + 0.2\mu^{*2},\ \mu^* = \mu^2 / \varepsilon\sigma^3$$

$$B_{free} = b_0 \left[(0.94 - 1.47/T^{*'} - 0.85/T^{*'2} + 1.015/T^{*'3}) - \mu^{*'}(0.75 - 3/T^{*'} + 2.1/T^{*'2} + 2.1/T^{*'3} \right] \tag{6.7}$$

where ε and σ are the effective nonpolar potential parameters while μ is the moleculat dipole moment. Moreover, by taking into account the contribution of association, B_{chem} is calculated as follows:

$$B_{chem} = b_0 \exp\left\{\eta\left[650/(\varepsilon/k + 300) - 4.27\right]\right\} \times \left\{1 - \exp\left[1500\eta/T\right]\right\} \tag{6.8}$$

where η is the association parameter.

In addition, the experimental activity coefficients, $\gamma_{i,\exp}$ were calculated as follows:

$$\gamma_{i-\exp} = \frac{y_i P}{x_i P_i^{sat}} \tag{6.9}$$

where x_i and y_i are the experimental liquid and vapor composition respectively, P is the pressure and P_i^{sat} is the vapor pressure. Moreover, the regularization term is used in order to ensure that the regressed parameters do not deviate significantly from the initial values and only used whenever needed. This would be useful as the previously investigated systems can still be represented with comparable accuracy. The higher the value of the weighting value, w_{reg} the higher the influence of the regularization.

6.2.2 Equilibrium Data

The input for the parameter estimation are VLE experimental data (involving C, H, O, N Cl and S atoms) and the statistics of data used in the regression work are summarized in Table 6.1. All VLE data have been initally tested for consistency using a quality assesment algorithm which combines four widely used consistency tests (Herington, Van Ness, Differential and Infinite Dilution tests) and a check between the consistency of the binary and the pure component vapor pressure.

For the parameter regression work, a total of 363 VLE datasets involving sulfurated and chlorinated systems with 5627 data points consisting of *P, T, x_i, y_i* and *P, T, x_i* data have been used. Only a moderate amount of experimental data were used because the purpose of the development of the Modified (Dortmund) UNIFAC-CI models is to be able to predict phase equilibria with a limited amount of experimental data [18].

Table 6.1: Statistics of the Data Used in the Parameter Regression.

Phase Equilibria Information	Types of system	Type of data	No. of systems	Data Points	Main Groups Involved
VLE	Chlorinated	*P, T, x_i, y_i*	38	1098	CH_2, C=C, ACH, $ACCH_2$, OH,CH_2O, CCl, CCl_2, CCl_3, CCl_4, ACCl
		P, T, x_i	34		
	Sulfurated	*P, T, x_i, y_i*	4	45	CH_2, C=C, ACH, CS_2, CH_3SH
	Total		**76**	**1143**	

6.2.3 Regression Procedure

In this work, the atom interactions parameters (AIPs) are regressed in series. The AIPs with respect to atom interactions C-C, C-O, O-O, C-N, O-N and N-N have been regressed previously in [18]. Therefore the next step is to regress AIPs with respect to atom Cl. Initially, only systems with C-Cl, Cl-C, O-Cl and Cl-O atom interactions are used for the parameter regression. For this step, a total of 72 data sets of VLE systems were used.

Next, sulfurated systems related to the C-S, S-C, O-S, S-O interactions are used for parameter regression where all previously regressed parameters are fixed. A total of 76 data sets of VLE systems were used in this step. The overall regression procedure is illustrated in Figure 6.1.

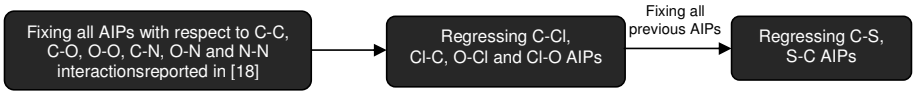


Figure 6.1: Overall Regression Procedure for the AIPs Regression Work.

6.2.4 Correlation Results

The correlation error or deviation between the experimental data and the regressed values are defined in terms of the average absolute relative deviation (AARD) shown in Equation (6.10):

$$AARD(\%) = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{P_{i-\exp} - P_{i-calc}}{P_{i-\exp}}\right| \times 100 \qquad (6.10)$$

In this section, detailed correlation results of each of the datasets involved in the parameter regression work are presented and described. At the end of this section, the overall correlation results are presented and discussed.

The correlation results for systems related to C-Cl and O-Cl interactions involving chloroalkanes-aromatics, chloroalkane-alkanes, aromatics-chloroaromatics, chloroalkanes-ethers and alcohols-chloroaromatics are presented in Table 6.2. The average AARD obtained for the reference Modified (Dortmund) UNIFAC is 3 % while for the Modified (Dortmund) UNIFAC-CI model is 4 %. The correlation error obtained for the CI-model is about 0.8 % compared to the reference model which is an acceptable deviation. There are systems which are problematic for the CI-model (with AARD > 10%). These are systems of Chloroform-Hexane at 298 K, Chloroform-Heptane at 298 K and Ethanol-Monochlorobenzene at 398 K There are also systems which obtained AARD > 10 % for the reference model which are the n-Heptane-1,3-Dichloropropane and 1,4-Dichloropropane systems at 298 K. Overall for the other systems, the correlation results for the CI-model are satisfactory.

Table 6.2: Correlation Results for C-Cl, Cl-C, O-Cl and Cl-O Atom Interactions Related Systems.

No.	Systems	AARD (%)	
		Modified (Dortmund) UNIFAC	**Modified (Dortmund) UNIFAC-CI**
1	sec-Butyl-Chloride-Toluene at 323 K	1.47	1.48

2	sec-Butyl-Chloride-n-Heptane at 323 K	5.03	2.91
3	sec-Butyl-Chloride-n-Heptane at 333 K	5.01	3.15
4	1,2-Dichloroethane-n-Heptane at 343 K	4.64	1.77
5	n-Butyl-Chloride-Toluene at 323 K	0.66	2.09
6	n-Butyl-Chloride-n-Heptane at 323 K	2.98	0.09
7	tert-Butyl-Chloride-n-Heptane at 323 K	5.97	4.83
8	n-Butyl-Chloride-n-Heptane at 323 K	2.98	0.09
9	n-Butyl-Chloride-Toluene at 323 K	0.66	2.09
10	sec-Butyl-Chloride-Toluene at 323 K	1.47	1.48
11	n-Butyl-Chloride-Toluene at 298 K	0.69	1.51
12	n-Butyl-Chloride-Toluene at 348 K	0.62	5.28
13	n-Butyl-Chloride-Toluene at 398 K	1.51	9.10
14	Benzene-Monochlorobenzene at 298 K	0.37	1.55
15	Benzene-Monochlorobenzene at 348 K	0.96	1.91
16	Benzene-Monochlorobenzene at 398 K	1.83	2.48
17	n-Pentane-Dichloromethane at 298 K	6.67	7.66
18	n-Pentane-Dichloromethane at 348 K	6.12	3.61
19	n-Pentane-Dichloromethane at 398 K	6.53	3.32
20	1,2-Dichloroethane-n-Heptane at 340 K	4.64	1.77
21	n-Heptane-Dichloromethane at 298 K	5.22	6.39
22	n-Heptane-1,2-Dichloroethane at 298 K	1.66	6.68
23	n-Heptane-1,3-Dichloropropane at 298 K	13.96	5.22
24	n-Heptane-1,4-Dichlorobutane at 298 K	13.58	3.66
25	Carbon-Tetrachloride-n-Hexane at 298 K	0.27	5.38
26	Carbon-Tetrachloride-n-Heptane at 298 K	0.57	5.66
27	Chloroform-n-Hexane at 298 K	0.84	12.63
28	Chloroform-n-Heptane at 298 K	1.26	14.81
29	n-Heptane-n-Butyl-Chloride at 298 K	3.23	0.43
30	Dichloromethane-Benzene at 298 K	1.52	1.00
31	Dichloromethane-Benzene at 348 K	1.78	0.87
32	Dichloromethane-Toluene at 298 K	2.65	3.00
33	Dichloromethane-Toluene at 347 K	3.40	6.47
34	Tetrachloroethane-N-Hexane at 298 K	6.28	4.19
35	Tetrachloroethane-N-Hexane at 308 K	6.40	7.06
36	1,2-Dichloroethane-N-Hexane at 298 K	4.51	4.14
37	p-Xylene-Carbon-Tetrachloride at 313 K	4.36	4.15
38	p-Xylene-Carbon-Tetrachloride at 323 K	3.41	4.56
39	Dichloromethane-1-Hexene at 298 K	3.89	8.01
40	Monochlorobenzene-Ethylbenzene at 293 K	0.58	2.59
41	1-Pentene-Monochlorobenzene at 320 K	2.05	6.30
42	1-Pentene-Monochlorobenzene at 360 K	1.17	2.08

43	o-Xylene-Carbon-Tetrachloride at 303 K	1.11	2.18
44	p-Xylene-Carbon_Tetrachloride at 303 K	1.06	2.04
45	m-Xylene-Carbon_Tetrachloride at 303 K	0.58	1.29
46	Cumene-Carbon-Tetrachloride at 303 K	2.83	2.48
47	n-Heptane- Carbon-Tetrachloride at 323 K	0.27	0.53
48	n-Butyl-Chloride-n-Heptane at 323 K	2.92	0.59
49	Carbon-Tetrachloride-Benzene at 260 K	0.48	1.47
50	Carbon-Tetrachloride-Benzene at 265 K	0.41	0.77
51	Carbon-Tetrachloride-Benzene at 266 K	0.39	0.75
52	Carbon-Tetrachloride-Benzene at 271 K	0.34	0.79
53	Carbon-Tetrachloride-Benzene at 273 K	0.31	0.65
54	Carbon-Tetrachloride-Toluene at 273 K	1.95	2.21
55	Dichloroethane-Di-n-Butyl-Ether at 330 K	2.41	3.59
56	Dichloroethane-Di-n-Butyl-Ether at 350 K	0.84	1.56
57	Dichloroethane-Di-n-Butyl-Ether at 370 K	2.83	2.47
58	Trichloroethane-Di-n-Butyl-Ether at 323 K	2.53	3.53
59	1-Chloropentane-Di-n-Butyl-Ether at 313 K	1.12	4.42
60	1-Chloropentane-Di-n-Butyl-Ether at 323 K	1.23	4.08
61	1,2-Dichloroethane-Di-n-Butyl-Ether at 330 K	2.41	3.59
62	Dichloroethane-Di-n-Butyl-Ether at 350 K	0.84	1.56
63	Dichloroethane-Di-n-Butyl-Ether at 370 K	2.83	2.47
64	Trichloroethane-Di-n-Butyl-Ether at 323 K	2.53	3.53
65	Trichloroethane-Di-n-Butyl-Ether at 343 K	2.81	7.38
66	Ethanol-Monochlorobenzene at 298 K	4.83	1.29
67	Ethanol-Monochlorobenzene at 348 K	4.28	6.08
68	Ethanol-Monochlorobenzene at 398 K	1.44	20.63
69	Ethanol-Chloroform at 303 K	3.82	2.00
70	Ethanol-Chloroform at 313 K	3.74	1.61
71	Ethanol-Chloroform at 323 K	4.35	1.70
72	1-Butanol-Carbon-Tetrachloride at 323 K	1.72	1.01
	Average	**2.81**	**3.63**

For the sulfurated systems, the correlation results are presented in Table 6.3 for C-S and S-C interactions related systems. In total, 4 datasets were used for the regression involving systems of benzene-carbon disulfide and ethyl mercaptan-propylene. We can see from Table 6.3 that the average correlation error obtained for the Modified (Dortmund) UNIFAC-CI model is 0.6 % which is better than the reference model with an average prediction error of 3 %.

Table 6.3: Correlation Results for C-S, S-C Atom Interaction Related Systems.

No.	Systems	AARD (%)	
		Modified (Dortmund) UNIFAC	Modified (Dortmund) UNIFAC-CI
1	Benzene-Carbon-Disulfide at 293.15 K	1.81	0.73
2	Benzene-Carbon-Disulfide at 303.15 K	1.21	0.35
3	Ethyl-Mercaptan-Propylene at 253 K	7.09	0.14
4	Ethyl-Mercaptan-Propylene at 323 K	2.80	1.19
Average		**3.23**	**0.60**

6.2.5 Overall Correlation Results

The correlation results reported in Tables 6.2 and 6.3 are summarized in Table 6.4 and Figure 6.2 according to the types of systems. Overall for the chlorinated and sulfurated systems used in the parameter regression step, the average deviation error obtained for the reference Modified (Dortmund) UNIFAC-CI model is 3 %. On the other hand, for the CI-model, the AARD obtained is slightly higher with 4 %. The correlation error obtained for the CI-model is slightly higher due to some problematic systems (Chloroform-Hexane, Chloroform-Heptane and Ethanol-Monochlorobenzene).

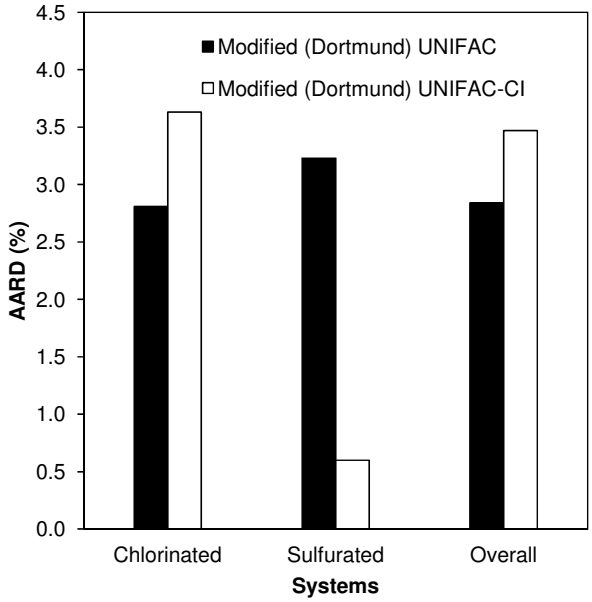


Figure 6.2: Correlation Errors in AARD (%) for Modified (Dortmund) UNIFAC-CI Model Compared with the Reference Modified (Dortmund) UNIFAC Model.

Table 6.4: Overall Correlation Results in AARD (%).

Types of systems	AARD (%)	
	Modified (Dortmund) UNIFAC	**Modified (Dortmund) UNIFAC-CI**
All chlorinated systems	2.81	3.63
All sulfurated systems	3.23	0.60
All systems	**2.84**	**3.47**

6.3 Predictions of VLE Data using the Regressed Parameters

In this section, the predictions of VLE systems using the regressed interaction parameters of the CI-models are presented and compared with the predictions obtained by the reference Modified (Dortmund) UNIFAC model. The predictions will demonstrate the performance of the regressed parameters and the reliability of the model in predicting systems which are not used in the parameter regression but for which parameters are available.

6.3.1 Chlorinated Systems

The VLE phase diagrams for chlorinated systems (systems containing C, H and Cl atoms) involving 1 isothermal and 1 isobaric system are illustrated in Figure 6.3. The group interactions involved include CH_2-ACH, CH_2-ACCl, ACH-ACCl and CH_2-CCl. In Figure 6.3 (a), the predictions made by Modified (Dortmund) UNIFAC for the system of n-heptane-monochlorobenzene at 323 K are in very close agreement with the experimental data. However, the predictions made by the CI-model slightly overpredict the experimental data especially for the pressure-liquid composition (P-*x*) data. On the other hand, for the systems of n-hexane-n-butyl chloride at 94.4 kPa displayed in Figure 6.3 (b), the predictions made by the Modified (Dortmund) UNIFAC-CI are better and closer to the experimental data compared to the reference model. Eventhough the predictions made by the CI-model do not represent the systems very accurately for both systems, the predictions are acceptable as they are close to the experimental data.

6.3.2 Sulfurated Systems

The VLE predictions of sulfurated systems (systems containing C, H and S atoms) using Modified (Dortmund) UNIFAC and Modified (Dortmund) UNIFAC-CI for two isothermal systems are presented in Figure 6.4. The GIPs involved are for the interactions between groups ACH-CS_2, CH_2-CH_2SH and C=C-CH_2SH. In Figure 6.4 (a), for the system of benzene-carbon disulfide at 298.15 K the predictions made by the reference Modified (Dormund) UNIFAC and the CI-model are in very good agreement with the experimental data. This maybe due to the fact that there are 2 other systems of benzene-carbon disulfide at different temperatures that were used in the parameter regression step and obtained low deviation errors. Furthermore, Figure 6.4 (b) shows the VLE

diagram for the system of ethyl mercaptan-propylene at 323 K. This system was used in the parameter regression step obtaining AARD of 3 % for the reference Modified (Dortmund) UNIFAC model and 1 % for the CI-model. From Figure 6.4 (b), we can see that the predictions made by both UNIFAC models are in very close agreement with the experimental data. It is expected that the predictions of the same type of system with the same UNIFAC group interactions will generate phase diagrams with the same accuracy.

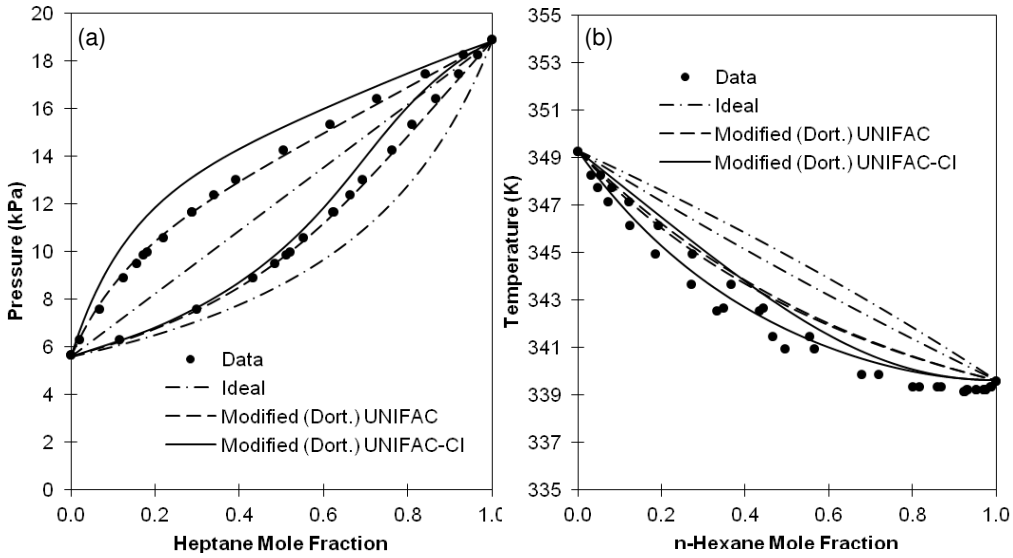


Figure 6.3: VLE Diagrams Predicted using Modified (Dortmund) UNIFAC and Modified (Dortmund) UNIFAC-CI (VLE) for Systems of: (a) n-Heptane-Monochlorobenzene at 323 K and (b) n-Hexane-n-Butyl Chloride at 94.4 kPa.

6.4 Conclusions

In this chapter, the atom interaction parameters (AIPs) of the Modified (Dortmund) UNIFAC-CI model involving chlorinated and sulfurated VLE systems have been regressed and the results of the correlations were compared with the prediction accuracy of the reference Modified (Dortmund) UNIFAC model using GIPs reported in [68]. The correlation errors obtained for the CI-model are as good as the reference model except for a few problematic system which showed deviation errors more than 10 %.

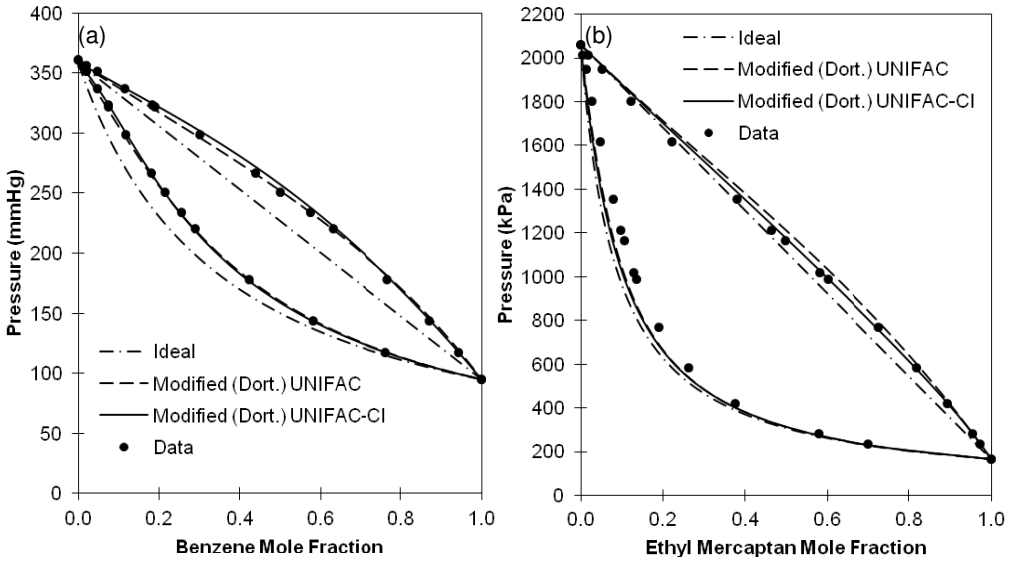


Figure 6.4: VLE Diagrams Predicted using Modified (Dortmund) UNIFAC and Modified (Dortmund) UNIFAC-CI (VLE) for Systems of: (a) Benzene-Carbon Disulfide at 298.15 K and (b) Ethyl-Mercaptan-Propylene at 323 K.

Furthermore, in section 6.3 the predictions of several VLE systems using the regressed parameters have been presented. The systems used for predictions are of the same type as the systems used in the parameter regression. It is found that the prediction accuracy of the CI-model are acceptable and close to the experimental data. However, as for the Original UNIFAC-CI model, when we in order to use the regressed parameters for systems or UNIFAC main groups which have not been used in the parameter regression step, good predictions are not guaranteed. It is the purpose of this work to only use a limited amount of data for parameter regression. Then, to use the regressed parameters for the prediction of other systems which have not been used in parameter regression but at least having the same UNIFAC main groups.

Currently the Modified (Dortmund) UNIFAC-CI model has been regressed against only VLE data compared to the reference Modified (Dortmund) UNIFAC-CI which has been regressed against other data as well such as activity coefficients at infinite dilution and excess enthalphies. Therefore, this work will be a starting point for the further development of the Modified (Dortmund) UNIFAC-CI model which will be comparable to the reference UNIFAC model.

CHAPTER 7

APPLICATION OF THE ORIGINAL UNIFAC-CI MODEL: CASE STUDIES

7.1 Introduction

In this chapter, the application of the developed and analyzed Original UNIFAC-CI model will be highlighted in four case studies. These are a working solution design for the production of hydrogen peroxide, an investigation of the solubilities of active pharmaceutical ingredients in different solvents, the performance of the CI-model in the predictions of systems involving lipids and also the application of the CI-model in specific systems where the group interaction parameters of the Original UNIFAC are missing including azeotropic systems. The atom interactions parameters (AIPs) that were used in these case studies are those regressed simultaneously against vapor-liquid equilibria (VLE) and solid-liquid equilibria (SLE) systems presented in Chapter 5. The case studies that are presented in Sections 7.2-7.5 will highlight the advantages and some limitations or the application range of this Original UNIFAC-CI model.

7.2 Working Solution Design for Hydrogen Peroxide Production

Hydrogen peroxide (H_2O_2) is the simplest peroxide of the peroxide family and it consists of two hydrogen atoms and two oxygen atoms that are single bonded with each other. The O-O group is called the peroxide group. The molecular structure of hydrogen peroxide is displayed in Figure 7.1. This chemical is a weakly acidic (pure H_2O_2 has a pH of 6.2) colorless liquid and slightly more viscous than water. It is also a strong oxidizer and therefore it is often used as a bleaching or cleaning agent. In New Zealand, the peroxide is mainly used to oxidize many chemical compounds including lignins, cyanides, sulphides and phenols. Hydrogen peroxides can be reacted to produce HO or HOO^- or other species depending on the selected conditions [132].

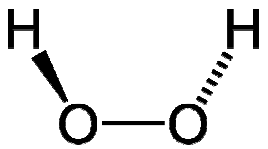


Figure 7.1: Molecular Structure of Hydrogen Peroxide.

Traditionally, hydrogen peroxide was produced by the electrolysis of ammonium bisulfate (NH_4HSO_4) or an aqueous solution of sulfuric acid, followed by hydrolysis of the peroxodisulfate $S_2O_8^{2-}$ that is formed. However, today the widely used process to produce hydrogen peroxide is the

anthraquinone process involving the autoxidation of a 2-alkyl anthrahydroquinone to the corresponding 2-alkyl anthraquinone. This process is a cyclic operation where the alkyl anthraquinone is reused. In this cycle, several processes are involved which are in sequence starting from hydrogenation followed by filtration, oxidation and extraction stages. In the hydrogenation stage, an alkyl anthraquinone is dissolved in solvents and the combination is described as the working solution. The working solution containing the dissolved anthraquinone is hydrogenated using hydrogen gas in a slurry-type hydrogenator using alumina loaded with a small amount of palladium catalyst. In this step, the alkyl anthraquinone and tetrahydroalkyl anthraquinone are converted to alkyl anthrahydroquinone and tetrahydroalkyl anthrahydroquinone. The tetra-form of the quinone is preferable since it is easier to be hydrogenised. The hydrogenation process is illustrated in the chemical equation shown in Figure 7.2.

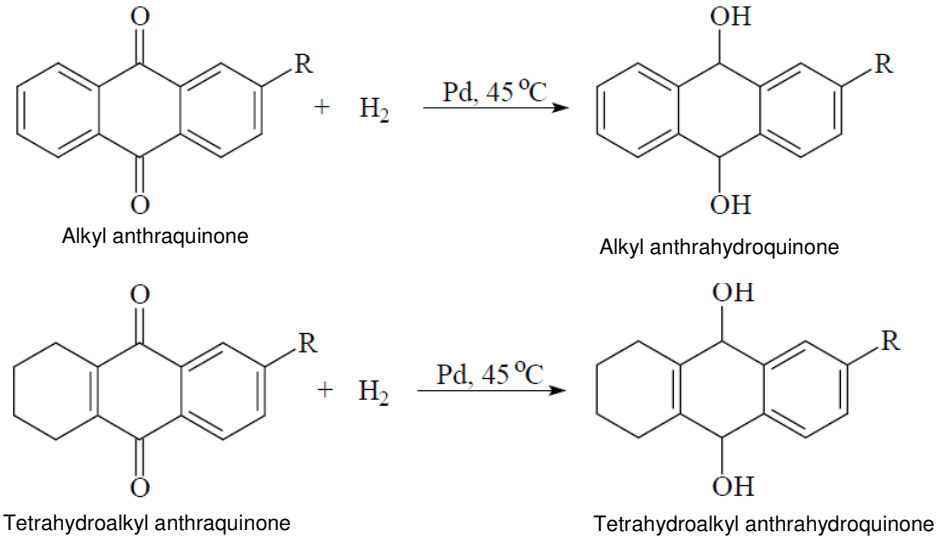


Figure 7.2: Hydrogenation Step of the Anthraquinone Process.

The next step is the filtration of the hydrogenated anthraquinone to remove any trace of catalyst that can decompose the hydrogen peroxide in the next stages which can reduce yields and create potential hazards and unwanted by-products. After that, another important step is the oxidization of the working solution (containing the hydrogenated quinone) where air is blown through it forming hydrogen peroxide in an organic phase. This step is also called as the auto-oxidation since no catalyst is used. The chemical equation for this process is shown in Figure 7.3. Furthermore, in order to extract the hydrogen peroxide from the working solution and since it is present in an organic phase, demineralized water is added from the top of a liquid-liquid extraction column where the working solution and the hydrogen peroxide mixture have been fed to. The column is designed to ensure that a maximum contact between the water and the working solution is obtained.

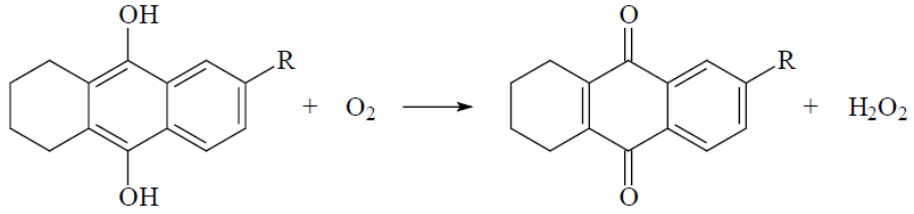


Figure 7.3: Oxidation Step of the Anthraquinone Process.

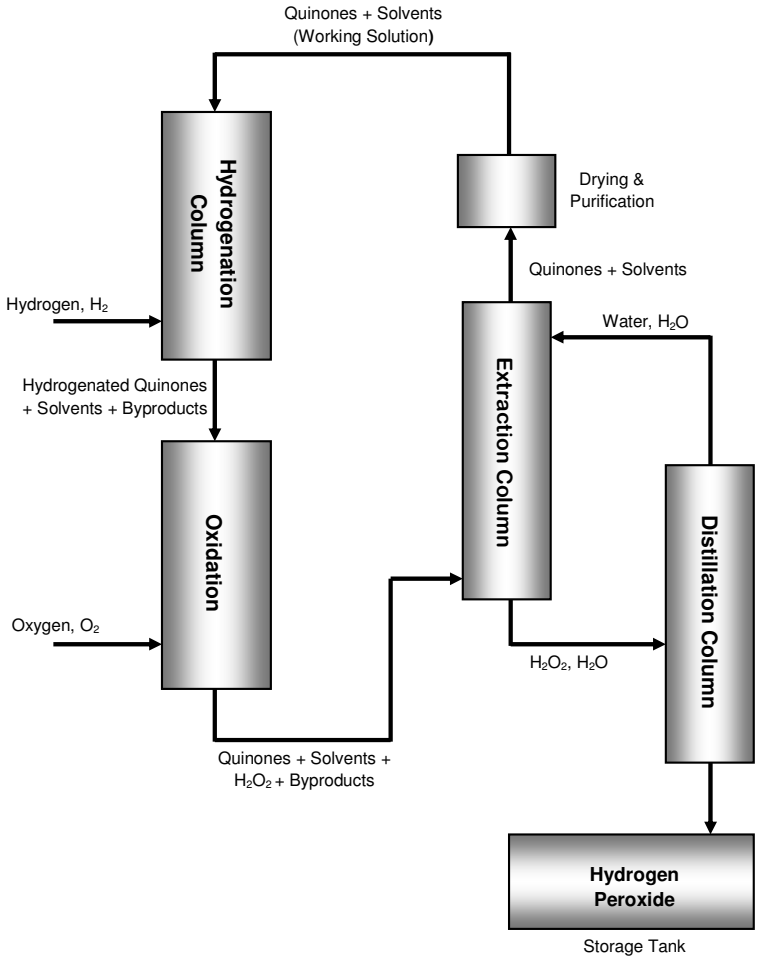


Figure 7.4: Simplified Process Flow Diagram of a Hydrogen Peroxide Production using the Anthraquinone Process.

The water will extract hydrogen peroxide at the bottom of the column while the working solution leaves the top of the column without any hydrogen peroxide. The working solution now containing alkyl anthraquinone and tetrahydroalkyl anthraquinone is pumped back into the hydrogenator for the next cycle. The crude hydrogen peroxide is purified in a distillation column to remove water and stored in a storage tank. The simplified process flow diagram of the production of hydrogen peroxide is illustrated in Figure 7.4.

7.2.1 Background of Case Study

In this case study, we are interested to study the solubility of the quinones in different solvents in order to maximize the solubility of the quinones in the working solution and to increase the efficiently in producing hydrogen peroxide. This case study is in collaboration with a company based in the United States of America which produces hydrogen peroxide. However, all the data provided by the company are confidential and therefore the identity of the quinones and interested solvents is not reported. The quinones interested for this study is labeled as Anthraquinone 1, an alkyl anthraquinone and it corresponding tetrahydroalkyl anthraquinone labeled as Anthraquinone 2. These quinones are hydrogenated in the peroxide process to the corresponding Anthrahydroquinone 1 and Anthrahydroquinone 2. The solvents which are of interest are labeled as Solvent 1, Solvent 2, Solvent 3, Solvent 4 and Solvent 5

We are interested in using the Original UNIFAC-CI model which has been developed in Chapter 5 for this case study since the anthraquinone could not be fully represented by the Original UNIFAC model. Therefore a new group described as ACCO for the anthraquinone has been created and its related interaction parameters are predicted using the regressed atom interaction parameters (AIPs) and Equations (2.37)-(2.42). The new ACCO group which has been created is shown in Figure 7.5 as an example for anthraquinone.

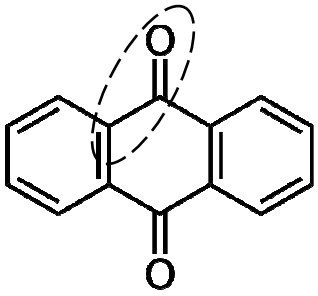


Figure 7.5: The ACCO Group in Anthraquinone.

In order to calculate the related group interaction parameters (GIPs), valence connectivity indices (CI) for this new group need to be calculated using Equations (2.31)-(2.36). First, the molecular structure of ACCO is transformed to a hydrogen-suppressed graph as shown in Figure 7.6. Each non-hydrogen atoms are labeled as 1, 2 and 3 while the bonds became edges *a* and *b*. The

omission of hydrogen and double bonds is compensated by the manner in which Equation (2.31) has been defined. The values needed for the calculation of the delta values, δ^v such as the number of atoms *Z*, number of valence electron Z^v and the number of hydrogen N_H atom attached are listed in Table 7.1. Using Equation (2.31), the delta values for each vertex *1*, *2* and *3* are calculated and reported in row number 4 of Table 7.1.

Calculations of the Delta Values:

1) Equation (2.31) is as follows:

$$\delta^v = \frac{(Z^v - N_H)}{(Z - Z^v - 1)}$$

2) Using the values listed in Table 7.1, δ^v for atoms *1*, *2*, *3* are calculated as follows:

Atom 1: $\delta^v = \frac{(Z^v - N_H)}{(Z - Z^v - 1)} = \frac{(6-0)}{(8-6-1)} = 6$

Atom 2 and 3: $\delta^v = \frac{(Z^v - N_H)}{(Z - Z^v - 1)} = \frac{(4-0)}{(6-4-1)} = 4$

Furthermore, using Equations (2.33) and (2.35), the first order and second order valence bond indices β^v and ε^v are calculated and they are also reported in Table 7.1. Using all those values reported in this table, the zeroth, first and second order valence connectivity indices are calculated.

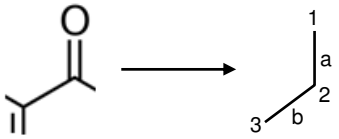


Figure 7.6: Molecular Structure and Hydrogen-Suppressed Graph of ACCO Group.

Table 7.1: Atomic, Bond and Path Indices Values for ACCO.

Atom	1	2	3
Z^v	6	4	4
N_H	0	0	0
Z	8	6	6
δ^v	6	4	4
β^v (1-2)	24		-
β^v (2-3)	-	16	
ε^v (1-2-3)	96		

Calculations of the Valence Connectivity Indices:

1) Zeroth order CI: $^{v}\chi^{0} = \sum_{i}\left(\frac{1}{\sqrt{\delta_i^{v}}}\right) = \frac{1}{\sqrt{6}} + \frac{1}{\sqrt{4}} + \frac{1}{\sqrt{4}} = 1.4083$

2) First order CI: $^{v}\chi^{1} = \sum_{i}\left(\frac{1}{\sqrt{\beta_i^{v}}}\right) = \frac{1}{\sqrt{24}} + \frac{1}{\sqrt{16}} = 0.4541$

3) Second order CI: $^{v}\chi^{2} = \sum_{i}\left(\frac{1}{\sqrt{\varepsilon_i^{v}}}\right) = \frac{1}{\sqrt{96}} = 0.1021$

After calculating the valence connectivity indices (CI) for ACCO, the next step is to calculate the group interaction parameters (GIPs) between ACCO and other groups which are describing the solvents and also other groups in the anthraquinone. The calculations of the GIPs between group CH_2 and ACCO are shown in Appendix D.

When all the needed GIPs for ACCO have been generated, they are used together with the UNIFAC model to generate solubility diagrams for the working solution and to investigate which is the best which can maximize the solubility of the quinones. In order to use UNIFAC, important parameters which are needed are the group volume, R_k and surface area, Q_k values. These values are obtained from Bondi [54]. The R_k and Q_k values for group ACCO are 1.1365 and 0.724 respectively.

7.2.2 Phase Equilibria Predictions and Parameter Table

In the anthraquinone process for the production of hydrogen peroxide, the alkyl anthraquinone and tetrahydroalkyl anthraquinone are dissolved in a mixture of different solvents to form the working solution. However, for this case study, the solubilities of the pure quinones in single solvent are investigated first. This investigation can be a good starting point in order to decide which combination of solvents can be used in order to maximize the solubility of the quinones. For this work, solubility data for Anthraquinone 1 and Anthraquinone 2 (the tetra hydro form) are available for three solvents (Solvent 1, Solvent 2 and Solvent 3).

From the newly created group ACCO, its corresponding group interaction parameters (GIPs) have been predicted using the CI-method and the parameter table obtained for this case study is displayed in Table 7.2. In this case study, we are using all the parameters which are already available for the Original UNIFAC model. Group ACCO has been created to be able to fully describe the anthraquinone. For group ACCO, the GIPs are predicted using the CI-method. In Table 7.2, there is also group PO_4 which is not available in the Original UNIFAC parameter table but the group is needed to describe one the solvents. Since the AIPs available are only related to atoms C, H, O, N, Cl and S, the GIPs related to PO_4 are set to zero.

Table 7.2: Group Interaction Parameters (GIPs) Table (in Kelvin) for Working Solution Design Case Study (GIPs labeled with * are predicted using the CI-method).

	CH_2	ACH	$ACCH_2$	OH	ACOH	CCOO	CH_2N	CON	ACCO	PO_4
CH_2	0	61	77	987	1333	232	207	391	974*	0
ACH	-11	0	167	636	1329	6	91	-1985*	560*	0
$ACCH_2$	-70	-147	0	803	885	5688	24	-1496*	1201*	0
OH	156	90	26	0	-260	101	-323	-383	-2356*	0
ACOH	276	25	244	-452	0	-449	1526*	12629*	541*	0
CCOO	115	86	-170	245	-37	0	-197	10283*	780*	0
CH_2N	-84	-224	110	59	-3515*	2889	0	-1176*	1048*	0
CON	28	-196*	-945*	395	-3621*	1596*	-190*	0	828*	0
ACCO	-24*	272*	265*	149*	1009*	2386*	860*	2953*	0	0
PO_4	0	0	0	0	0	0	0	0	0	0

Initially it was intended to use all the parameters which have been obtained in Table 7.2 to predict the solubility of the anthraquinones in all the solvents. However, the predictions obtained were not very good and there were also convergence problems in the SLE calculations. This problem is due to the high negative values (usually > -1000) of the GIPs that were predicted using the CI-method. Therefore, we have decided to fine tune one or more pairs of the GIPs predicted using the CI-method against the experimental data which have been given by the company. Based on sensitivity analysis, the most sensitive pair of GIPs was selected to be fine tuned against the given experimental data. The GIPs were fined tuned using the Thermodynamic Model Parameter Estimation (TML) sub-program in the ICAS software [133].

For the system of Anthraquinone 1 and Solvent 1, 3 pairs of GIPs involving ACH-CON, $ACCH_2$-CON and CH_2N-CON have been fined tuned due to the high negative values that were obtained when the original GIPs were predicted using the CI-method. The same set of parameters can also be used to predict the system of Anthraquinone 2 and Solvent 1. The SLE diagrams for this system are illustrated in Figure 7.7 (a). For systems with Solvent 2, all the GIPs for the Original UNIFAC model plus the ACCO related CI-generated GIPs have been used without fine tuning the parameters since they have been also used to predict the previous systems. The SLE diagrams generated are displayed in Figure 7.7 (b). For Solvent 3 involving group PO_4, one pair of GIPs related to this group has been fine tuned related to the ACCO-PO_4 interaction. The SLE diagrams generated using this fine tuned GIPs are presented in Figure 7.7 (c).

For Solvent 4 and 5, no experimental data are available. Therefore the SLE systems related to those solvents are generated using the available parameters which consist of the original GIPs and the CI-generated GIPs. The SLE diagrams for Anthraquinone 1 and 2 and those two solvents are illustrated in Figure 7.8.

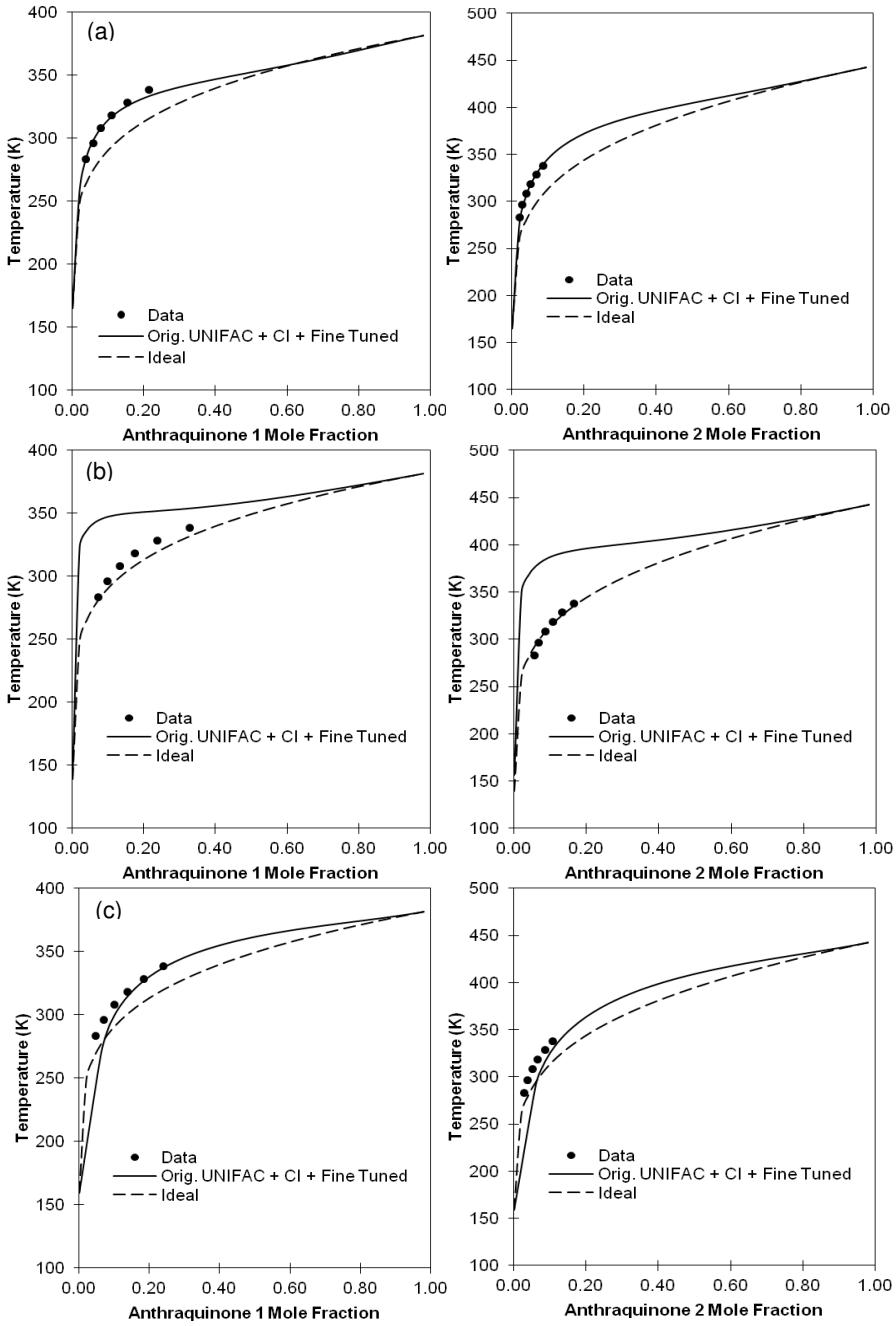


Figure 7.7: SLE Diagram for (a) Anthraquinone 1 and Anthraquinone 2 with Solvent 1, (b) Anthraquinone 1 and Anthraquinone 2 with Solvent 2 and (c) Anthraquinone 1 and Anthraquinone 2 with Solvent 3. (--- represents calculations using the ideal solubility assumption).

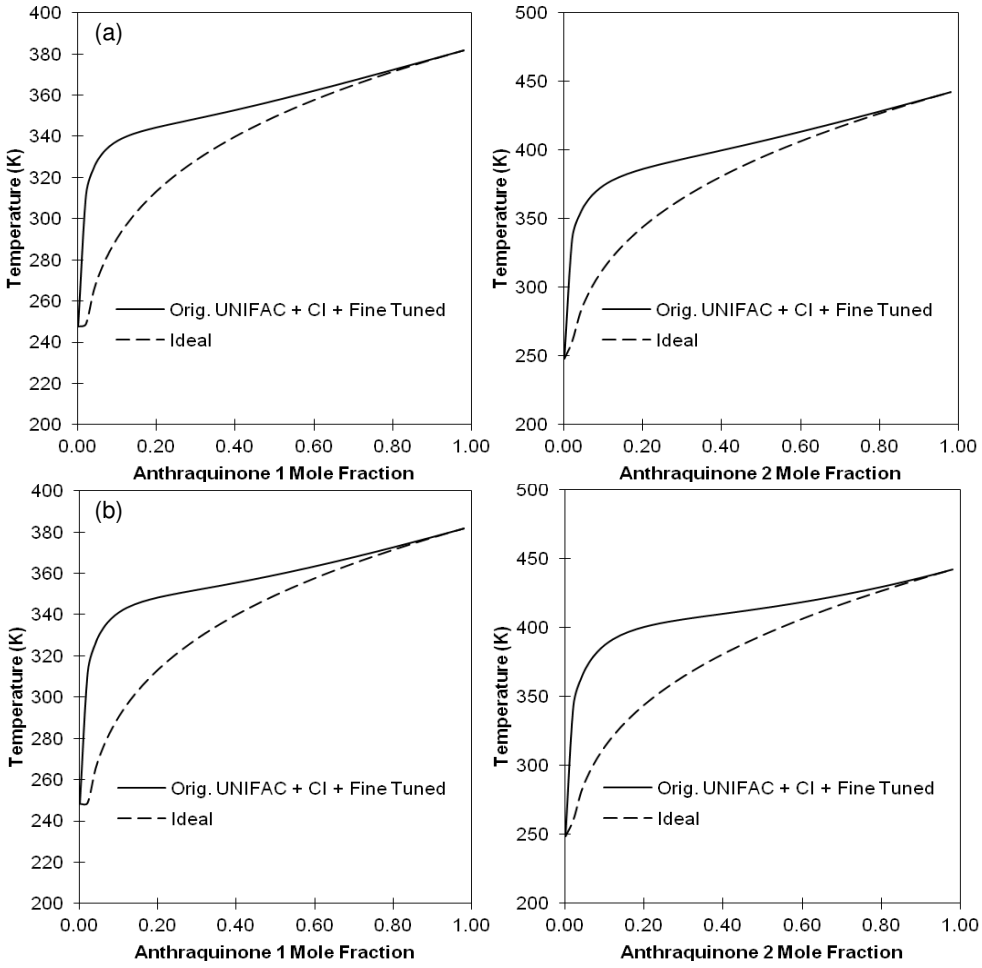


Figure 7.8: SLE Diagram for (a) Anthraquinone 1 and Anthraquinone 2 with Solvent 4, (b) Anthraquinone 1 and Anthraquinone 2 with Solvent 5. (--- represents calculations using the ideal solubility assumption).

Besides the solubilities of the anthraquinones, the SLE diagrams related to the hydrogenated anthraquinones have also been investigated. However, for these systems no experimental data were given. Therefore, initially the SLE curves were generated using the parameters which are available and predicted using the CI-method. However for the systems of Anthrahydroquinone 1 and Anthrahydroquinone 1 with Solvent 1, there are still CI-generated parameters which are highly negative which are related to ACOH-CH_2N and ACOH-CON and we expect those problems with the convergence of the SLE calculations. Therefore, imaginary points have been generated and those GIPs which are highly negative are fine tuned.

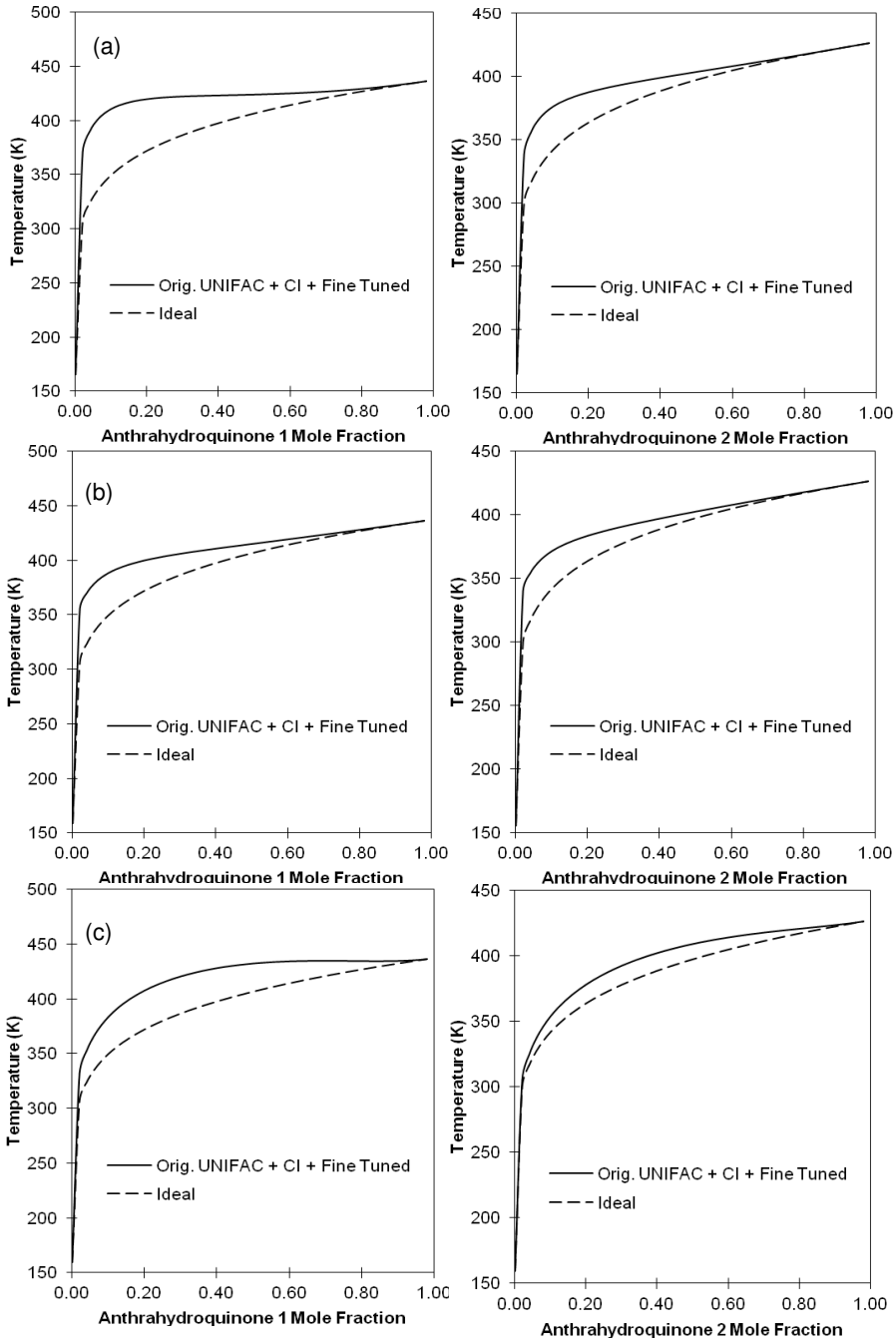


Figure 7.9: SLE Diagram for (a) Anthrahydroquinone 1 and Anthrahydroquinone 2 with Solvent 1, (b) Anthrahydroquinone 1 and Anthrahydroquinone 2 with Solvent 2 and (c) Anthrahydroquinone 1 and Anthrahydroquinone 2 with Solvent 3. (--- represents calculations using the ideal solubility assumption).

The results of the SLE curves are presented in Figure 7.9 (a). On the other hand, for the anthrahydroquinones with Solvent 2, no GIPs were fined tuned and the related SLE curves are displayed in Figure 7.9 (b).

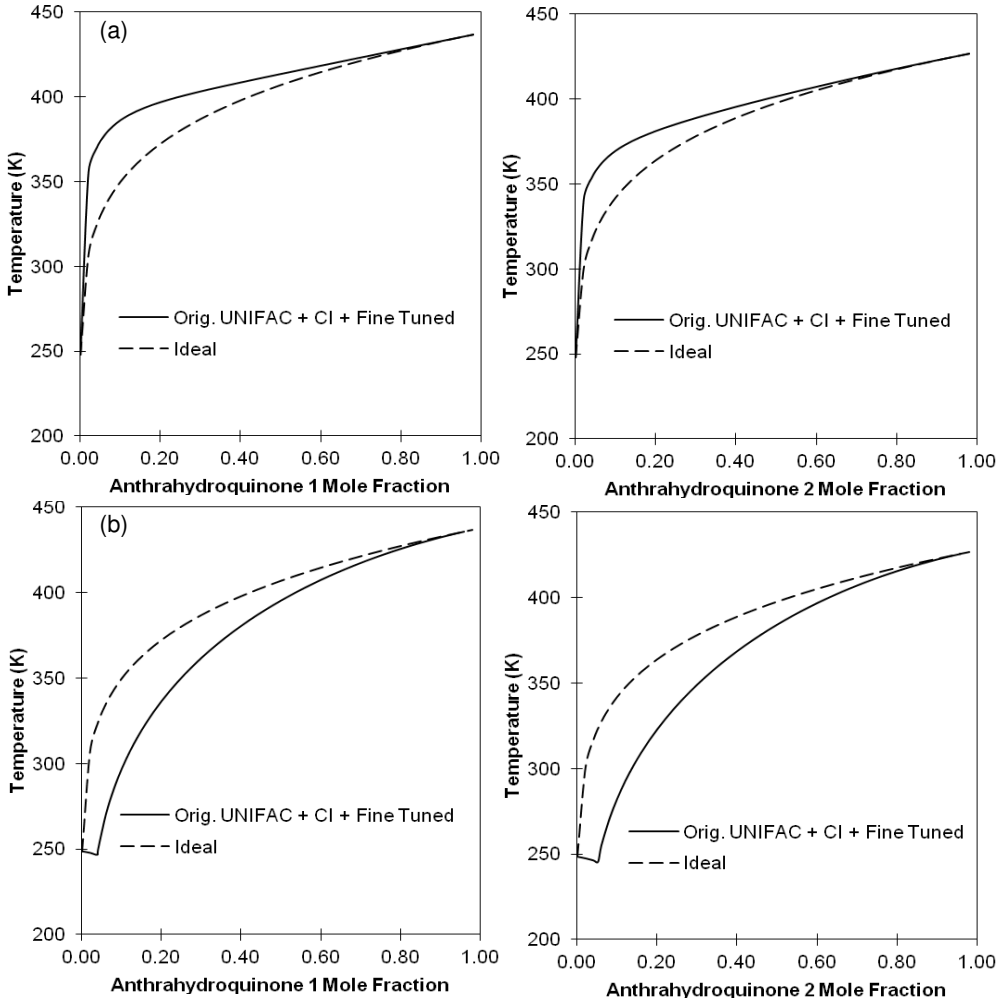


Figure 7.10: SLE Diagram for (a) Anthrahydroquinone 1 and Anthrahydroquinone 2 with Solvent 4, (b) Anthrahydroquinone 1 and Anthrahydroquinone 2 with Solvent 5. (--- represents calculations using the ideal solubility assumption).

Furthermore, for the systems between the hydrogenated quinones and Solvent 3, imaginary points were generated and 1 pair of parameters related to ACOH-PO_4 has been fine tuned. The SLE diagrams generated using the fine tuned parameters are illustrated in Figure 7.9 (c). In addition, in

Figure 7.10, the SLE phase diagrams of the anthrahydroquinones with Solvent 4 and 5 have been generated without fine tuning any of the remaining interaction parameters since the diagrams obtained are satisfactory. The final parameter table obtained for the prediction of SLE systems related to this working solution case study is presented in Table 7.3.

Table 7.3: Group Interaction Parameters (GIPs) Table (in Kelvin) for Working Solution Design Case Study with Fine Tuned Parameters (GIPs labeled with * are predicted using the CI-method while GIPs labeled with # are fine tuned. The labels are placed at the right side of the GIPs)

	CH_2	ACH	$ACCH_2$	OH	ACOH	CCOO	CH_2N	CON	ACCO	PO_4
CH_2	0	61	77	987	1333	232	207	391	974*	0
ACH	-11	0	167	636	1329	6	91	435 #	560*	0
$ACCH_2$	-70	-147	0	803	885	5688	24	-577 #	1201*	0
OH	156	90	26	0	-260	101	-323	-383	-2356*	0
ACOH	276	25	244	-452	0	-449	-859 #	8203 #	541*	680 #
CCOO	115	86	-170	245	-37	0	-197	10283*	780*	0
CH_2N	-84	-224	110	59	8258 #	2889	0	575 #	1048*	0
CON	28	185 #	-657 #	395	9227 #	1596*	-264 #	0	828*	0
ACCO	-24*	272*	265*	149*	1009*	2386*	860*	2953*	0	-952 #
PO_4	0	0	0	0	-389 #	0	0	0	876 #	0

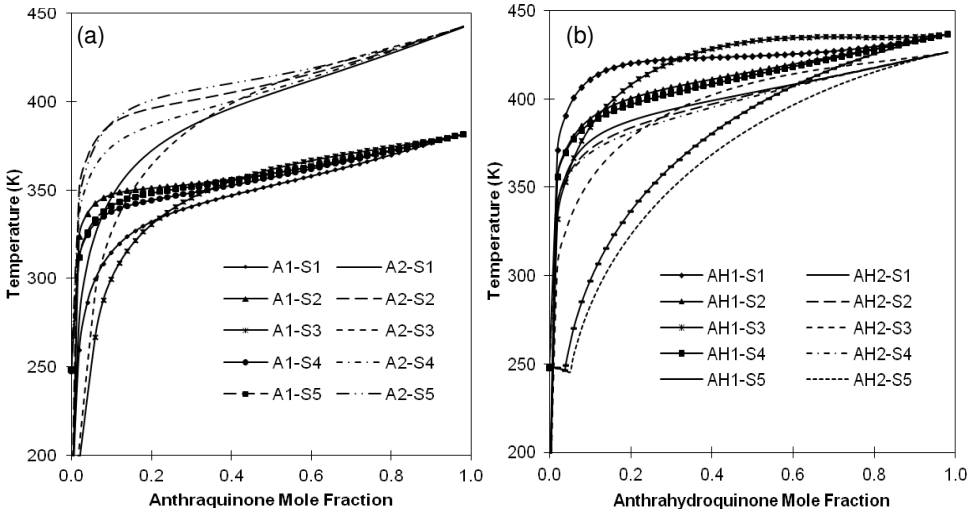


Figure 7.11: Overall SLE Diagrams for (a) Anthraquinones (A) and Different Solvents (S) and (b) Anthrahydroquinones (AH) and Different Solvents (S).

By using all the GIPs reported in Table 7.3, the SLE diagrams shown in Figures 7.7-7.10 have been generated. By combining all these individual SLE diagrams, a single diagram for anthraquinone and anthrahydroquinone have been generated in Figure 7.11 to identify which combinations of quinone and solvent provide the highest solubility of quinones. In Figure 7.11 (a), the SLE diagrams for Anthraquinone 1 and 2 together with Solvents 1-5 have been generated. According to the figure, the systems of Anthraquinone 1-Solvent 3 have the highest solubility of Anthraquinone 1 at lower temperatures until 334 K. Above that temperature, the solubility of Anthraquinone 1 is higher in Anthraquinone 1-Solvent 1. For Anthraquinone 2, the highest solubility can be found for system Anthraquinone 2-Solvent 3 until 388 K. However, the solubility of Anthraquinone 2 is higher above that temperature for Anthraquinone 2-Solvent 1. The same trends can be found for both Anthraquinone 1 and 2. Therefore, the best combinations of solute-solvent systems are Anthraquinone 1 and 2 with Solvent 3 at lower temperatures and Anthraquinone 1 and 2 with Solvent 1 at higher temperatures. However, for the anthrahydroquinones, the best system with the highest solubility is the Anthrahydroquinone 2 with Solvent 5.

7.2.3 Conclusions from the Case Study

In this case study, the Original UNIFAC-CI model has been used to predict the SLE behavior of anthraquinone with several interested solvents in order to design the best working solution for the production of hydrogen peroxide. The work presented here can be considered as the initial stage in order to design the final working solution which consists of a mixture of different anthraquinones and solvents. The systems involving anthraquinone cannot be fully represented by the Original UNIFAC model [15]. Therefore, using the AIPs which have been regressed for the Original UNIFAC-CI model, a new group ACCO has been created and its related group interaction parameters (GIPs) have been predicted using the CI-method. Using the GIPs which are already available in the UNIFAC parameter table [15] together with the GIPs which have been estimated, the needed SLE diagrams can be generated.

However, one limitation of this Original UNIFAC-CI model is that there is a probability that the GIPs predicted using regressed AIPs for groups which are not considered in the parameter regression step are highly negative. When trying to predict SLE system with these highly negative GIPs (usually > -1000), convergence of the SLE calculation is a problem. Therefore, some of those GIPs need to be fine tuned against the experimental data. The result is a master parameter table containing the original parameters, parameters predicted from the CI-method and the fine tuned parameters which can be used specifically for this case study. From a total of 14 pairs of GIPs which were initially predicted using the CI-method, only 7 pairs were fined tuned. Using the final parameter table, SLE diagrams have been generated and the best system which has the highest solubility of anthraquinone can be determined.

7.3 Pharmaceutical Systems

In the pharmaceutical industry, active pharmaceutical ingredients (API) are very important substances since they can provide specific reactions in organism even with a small quantity. These APIs include drugs which are used for therapeutic purposes on humans and animals and also

cosmetics and agrochemicals. The number of different functional groups that formed these APIs is quite large. Each API molecule is normally consisting of 10-50 non-hydrogen atoms and most of them contain hetero-atoms such as oxygen, nitrogen, sulfur and halogens. Some examples of those APIs are illustrated in Figure 7.12. In order to extract and purify the APIs, the screening of solvents and the investigation of the solubility of the APIs in solvents are very important.

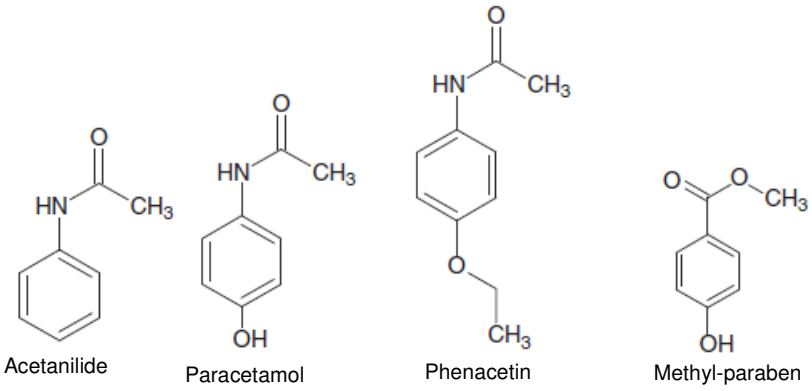


Figure 7.12: Examples of Pharmaceutical Compounds.

Usually, the APIs will decompose before reaching their melting point. Therefore, the APIs have been extracted from natural products with the help of a solvent. In addition, solvents are also used in the production of synthetic APIs where they are used as reaction media. Both types of APIs need to be obtained at high purity so that they can be used in producing medicine for human. In order to obtain these pure APIs, solution crystallization is used for the purification step, producing API in the solid form. Therefore the knowledge of the API solubility is very important since many factors in terms of yield, selectivity and also the cost of the processes will be affected.

In order to conduct experiments for solvents screening, cost can be the main issue since the APIs will not always be available for the measurements. Also, the required solubility data increase rapidly as a function of the number of solvents, different solvent mixtures, composition and temperatures. However, from a limited amount of experimental data that can be conducted for certain combinations of APIs-solvents, solubility models can be developed either using empirical methods involving a direct correlation between solubility and individual compounds or using predictive thermodynamic models.

7.3.1 Background of Case Study

In this case study, the solubility of APIs in selected solvents has been investigated using the Original UNIFAC-CI model that we have developed. Since the APIs are consisting of many different functional groups, modeling the solubility of these systems could be a difficult task. This case study is in collaboration with a pharmaceutical company which provided confidential experimental data. Therefore, the identity of those APIs cannot be revealed.

There are two APIs involved labeled as API 1 and API 2 and 12 different solvents. Both API 1 and API 2 consist of 4 different groups. In this study, only the solubility of API in pure single solvents will be investigated. The company has provided only one experimental data for each of the 9 API 1-solvents and also for 9 API 2-solvents systems. In order to model the solubility of those systems using the Original UNIFAC-CI model, new groups need to be created since API 1 and 2 cannot be fully represented using the current UNIFAC groups. Those new groups are ACCO, ACNH and ACCONH that can be found in many pharmaceutical compounds. For each of those new groups, valence connectivity indices (CI) need to be calculated using Equations (2.31)-(2.36). First, the ACNH and ACCONH groups are represented in terms of the hydrogen-suppressed graph displayed in Figure 7.13 and 7.14 (for the calculation of the CI involving ACCO, please refer to previous case study).

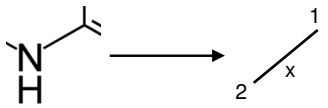


Figure 7.13: Molecular Structure and Hydrogen-Suppressed Graph of ACNH Group.

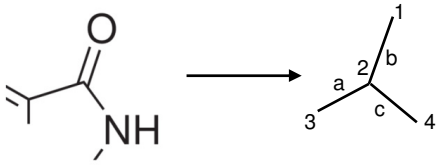


Figure 7.14: Molecular Structure and Hydrogen-Suppressed Graph of ACCONH Group.

For ACNH, each non-hydrogen atoms are labeled as 1 and 2 while the bond became edge x while for ACCONH the non-hydrogen atoms are labeled as 1, 2, 3 and 4 while the bonds become edges *a*, *b* and *c*. The omission of hydrogen and double bonds is compensated by the manner in which Equation (2.31) have been defined. The values needed for the calculation of the delta values, δ^v such as the number of atoms *Z*, number of valence electron Z^v and the number of hydrogen N_H atom attached are listed in Table 7.4 and 7.5. Using Equation (2.31), the δ^v values for each vertex are calculated and reported in row number 4 of Table 7.4 and 7.5.

Calculations of the Delta Values:

1) Equation (2.31) is as follows:

$$\delta^v = \frac{(Z^v - N_H)}{(Z - Z^v - 1)}$$

2) Using the values listed in Table 7.4, δ^v for atoms 1 and 2 are calculated as follows:

Atom 1: $\delta^v = \frac{(Z^v - N_H)}{(Z - Z^v - 1)} = \frac{(4-0)}{(6-4-1)} = 4$

Atom 2: $\delta^v = \frac{(Z^v - N_H)}{(Z - Z^v - 1)} = \frac{(5-1)}{(7-5-1)} = 4$

3) Using the values listed in Table 7.5, δ^v for atoms 1-4 are calculated as follows:

Atom 1: $\delta^v = \frac{(Z^v - N_H)}{(Z - Z^v - 1)} = \frac{(6-0)}{(8-6-1)} = 6$

Atom 2 and 3: $\delta^v = \frac{(Z^v - N_H)}{(Z - Z^v - 1)} = \frac{(4-0)}{(6-4-1)} = 4$

Atom 4: $\delta^v = \frac{(Z^v - N_H)}{(Z - Z^v - 1)} = \frac{(5-1)}{(7-5-1)} = 4$

Table 7.4: Atomic, Bond and Path Indices Values for ACNH.

Atom	**1**	**2**
Z^v	4	5
N_H	0	1
Z	6	7
δ^v	4	4
β^v (1-2)	16	

Table 7.5: Atomic, Bond and Path Indices Values for ACCONH.

Atom	**1**	**2**	**3**	**4**
Z^v	6	4	4	5
N_H	0	0	0	1
Z	8	6	6	7
δ^v	6	4	4	4
β^v (1-2)	24			
β^v (2-3)	16			
β^v (2-4)	16			
ε^v (1-2-3)	96			
ε^v (1-2-4)	96			
ε^v (3-2-4)	64			

Furthermore, using Equations (2.33) and (2.35), the first order and second order valence bond indices β^v and ε^v are calculated and they are reported in Table 7.4 and 7.5. Using all those values

reported in the tables, the zeroth, first and second order valence connectivity indices are calculated as shown below:

Calculations of the Valence Connectivity Indices:

For ACNH:

1) Zeroth order CI: $^{v}\chi^{0} = \sum_{i}\left(\frac{1}{\sqrt{\delta_i^v}}\right) = \frac{1}{\sqrt{4}} + \frac{1}{\sqrt{4}} = 1.0000$

2) First order CI: $^{v}\chi^{1} = \sum_{i}\left(\frac{1}{\sqrt{\beta_i^v}}\right) = \frac{1}{\sqrt{16}} = 0.500$

For ACCONH:

1) Zeroth order CI: $^{v}\chi^{0} = \sum_{i}\left(\frac{1}{\sqrt{\delta_i^v}}\right) = \frac{1}{\sqrt{6}} + \frac{1}{\sqrt{4}} + \frac{1}{\sqrt{4}} + \frac{1}{\sqrt{4}} = 1.9083$

2) First order CI: $^{v}\chi^{1} = \sum_{i}\left(\frac{1}{\sqrt{\beta_i^v}}\right) = \frac{1}{\sqrt{24}} + \frac{1}{\sqrt{16}} + \frac{1}{\sqrt{16}} = 0.7041$

3) Second order CI: $^{v}\chi^{2} = \sum_{i}\left(\frac{1}{\sqrt{\varepsilon_i^v}}\right) = \frac{1}{\sqrt{96}} + \frac{1}{\sqrt{96}} + \frac{1}{\sqrt{64}} = 0.3291$

After calculating the valence connectivity indices (CI) for ACNH and ACCONH, the next step is to calculate the group interaction parameters (GIPs) between those groups and other groups which represent the solvents and also other groups in the anthraquinone. The calculations of the GIPs between group CH_2 and ACNH are presented in Appendix D.

When all the needed GIPs for ACNH and ACCONH have been generated, they are used together with the UNIFAC model to generate solubility diagrams and to investigate which solvent is the best for both APIs. Again the group volume, R_k and surface area, Q_k values are needed and they are obtained from Bondi [54]. The R_k and Q_k values for group ACNH and ACCONH are presented in Table 7.6.

Table 7.6: R_k and Q_k Values for Group ACNH and ACCONH.

Group	R_k	Q_k
ACNH	0.8978	0.5160
ACCONH	1.6691	1.1200

7.3.2 Phase Equilibria Predictions and Parameter Table

From the newly created groups, ACCO, ACNH and ACCONH, their corresponding group interaction parameters (GIPs) have been predicted using the CI-method and the parameter table obtained for this case study is displayed in Table 7.7. In this case study, we are using all the parameters which are already available for the Original UNIFAC model. Groups ACCO, ACNH and ACCONH have been created to be able to fully describe the APIs. For groups ACCO, ACNH and ACCONH the GIPs are predicted using the CI-method.

Table 7.7: Group Interaction Parameters (GIPs) Table (in Kelvin) for Pharmaceutical Case Study. (GIPs labeled with * are predicted using the CI-method while GIPs labeled with ¤ are set to zero).

	CH_2	ACH	$ACCH_2$	OH	CH_3OH	CH_2CO	CCOO	CH_2O	CCN	ACCL	$ACNO_2$	ACNH	ACCO	ACCONH
CH_2	0	61	77	987	697	476	232	252	597	11	543	1885*	974*	-1304*
ACH	-11	0	167	636	637	26	6	32	213	187	195	1519*	560*	-2099*
$ACCH_2$	-70	-147	0	803	603	-52	5688	213	6096	-211	4448	1681*	1201*	-1610*
OH	156	90	26	0	-137	84	101	28	7	124	157	-4540*	-2356*	-2404*
CH_3OH	17	-50	-45	249	0	23	-11	-129	53	-28	0 ¤	0 ¤	-11*	0 ¤
CH_2CO	27	140	366	165	109	0	-214	-104	482	-120	549	628*	882*	2833*
CCOO	115	86	-170	245	250	372	0	-236	495	442	1927*	380*	780*	3442*
CH_2O	83	52	66	238	238	191	461	0	-19	135	3564*	1653*	656*	5079*
CCN	25	-23	-138	185	163	-288	-267	39	0	-5	340*	9382*	1089*	-345*
ACCL	107	-97	403	326	613	518	-171	-25	364	0	2213	2630*	1007*	242*
$ACNO_2$	5541	1824	-128	562	0 ¤	-102	1643*	-2648*	243*	-123	0	8958*	625*	6136*
ACNH	51*	-310*	-311*	-70*	0 ¤	-673*	-597*	-3193*	-796*	27*	210*	0	1196*	257*
ACCO	-24*	272*	265*	149*	-459*	745*	2386*	1302*	-3337*	949*	3701*	529*	0	2834*
ACCONH	-4*	16*	-848*	821*	0 ¤	2182*	2452*	-2068*	196*	188*	4701*	-90*	3209*	0

Initially it was intended to use all the parameters which have been obtained in Table 7.7 to predict the solubility of the APIs in all the solvents. However, the predictions obtained were not very good and there were also convergence problems with the calculations. This problem is due to the high negative values (usually > -1000) of the GIPs that were predicted using the CI-method. Therefore, we have decided to fine tune one or more pairs of the GIPs predicted using the CI-method against the experimental data which have been provided by the company.

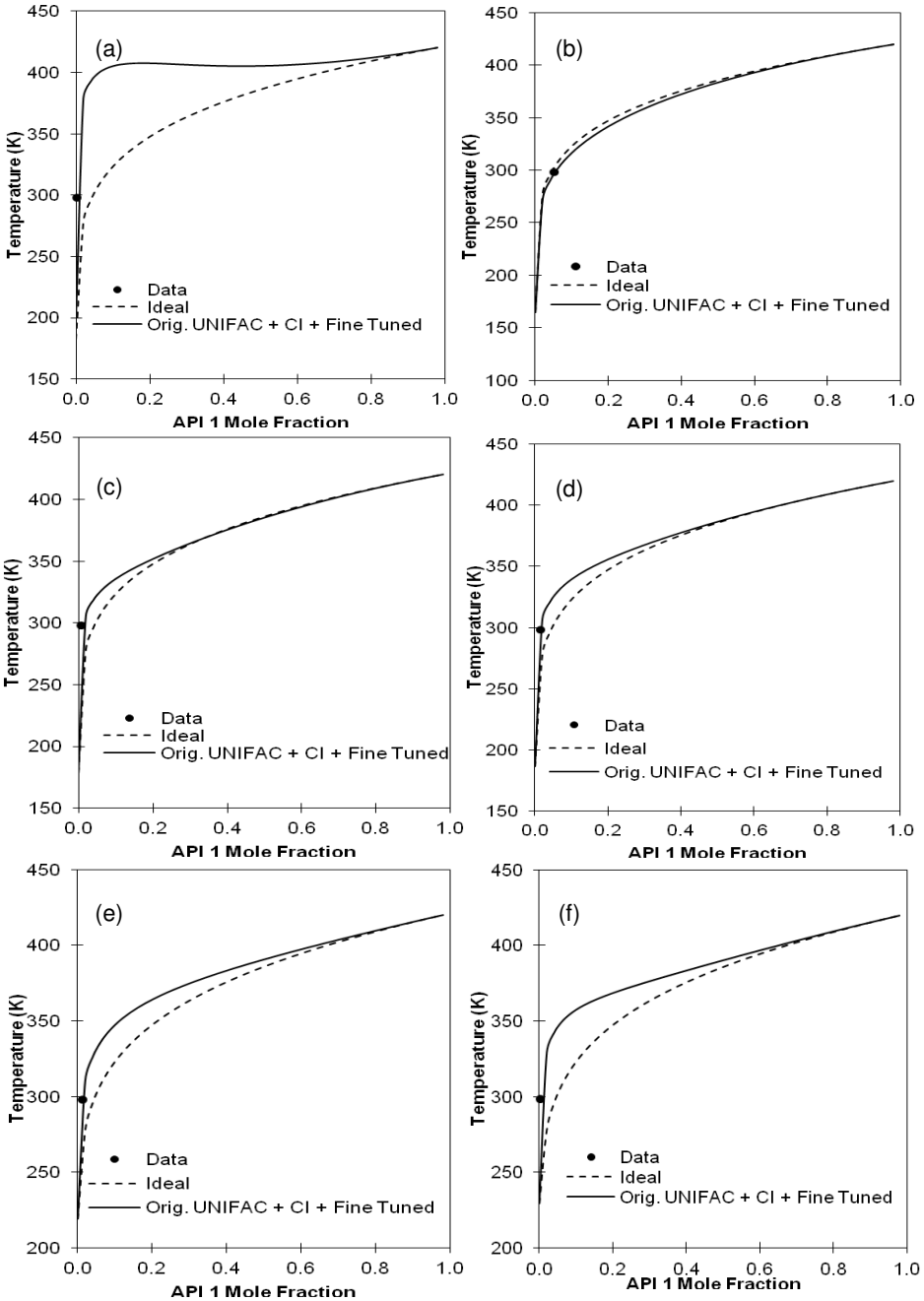


Figure 7.15: SLE Diagrams of APIs with Solvents, a) Heptane, b) Tetrahydrofuran, c) Acetone, d) Methyl Ethyl Ketone, e) Methyl Isobutyl Ketone and f) Acetonitrile Predicted using Original UNIFAC-CI using Fine Tuned GIPs. (--- represents calculations using the ideal solubility assumption).

Based on the sensitivity analysis, the most sensitive pair of GIPs was selected to be fine tuned against the given experimental data. The GIPs were fined tuned using the Thermodynamic Model Parameter Estimation (TML) sub-program in the ICAS software [133].

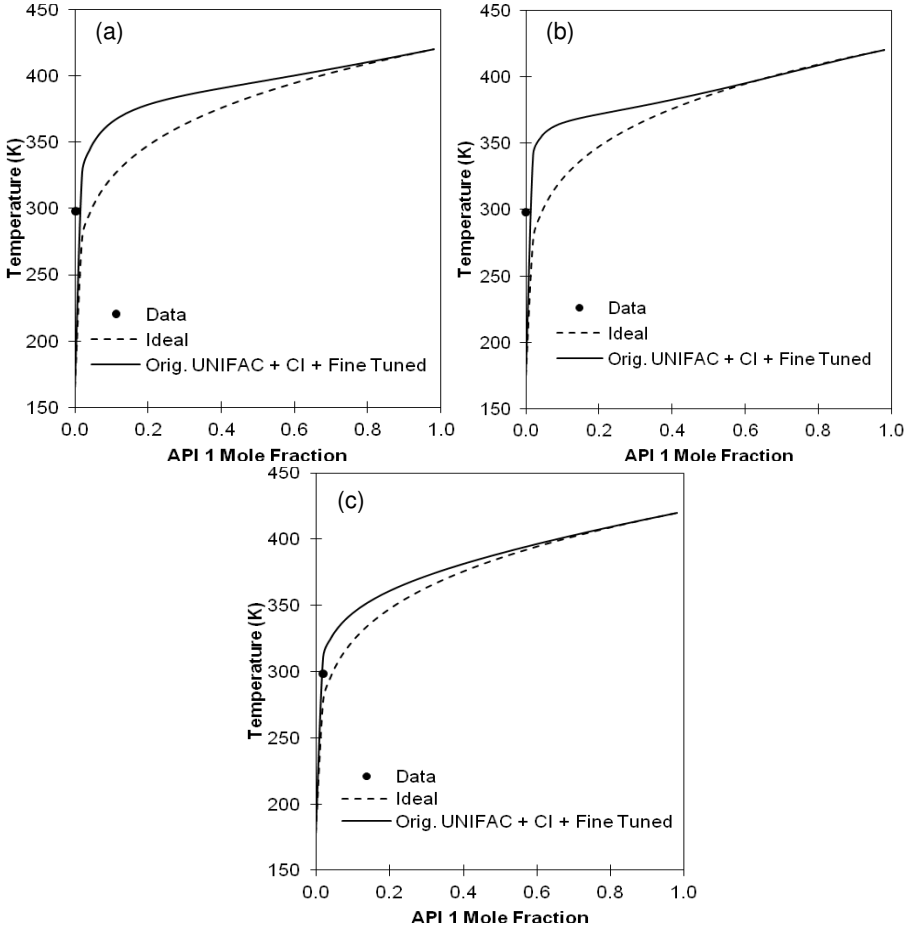


Figure 7.16: SLE Diagrams of APIs with Solvents, a) Isobutanol, b) Methanol and c) Toluene Predicted using Original UNIFAC-CI using Fine Tuned GIPs. (--- represents calculations using the ideal solubility assumption).

For the system of API 1-Heptane, 1 pair of GIPs related to ACH-ACNH interactions has been fined tuned against the single experimental data and the SLE diagram generated is shown in Figure 7.15 (a). Furthermore, for API 1-Tetrahydrofuran 2 pairs of GIPs related to CH_2O-$ACNO_2$ and CH_2O-ACNH have been fined tuned. Using the fine tuned parameters together with other GIPs, the SLE curve generated using the Original UNIFAC-CI model is illustrated in Figure 7.15 (b). Next, for the

system of API 1-Acetone, another pair of GIPs which was initially generated using the CI-model has been fine tuned with respect to the CH_2CO-ACNH interaction. Using these fine tune parameters, the SLE diagram for this system has been generated and displayed in Figure 7.15 (c). Since for systems API 1-Methyl Ethyl Ketone and API 1-Methyl Isobutyl Ketone, the GIPs that are needed are the same as for API 1 with Acetone, no further GIPs are fined tuned. The SLE predictions of those systems are presented in Figure 7.15 (d) and (e). Moreover, when investigating the solubility of API 1 in Acetonitrile, another pair of GIPs with respect to CCN-ACNH has been fine tuned against the single data point. The generated phase diagram is shown in Figure 7.15 (f).

Figure 7.16 (a) shows the SLE diagram for the system of API 1-Isobutanol where 1 pair of GIPs from the CI-generated parameters has been fine tuned. The fine tuned parameters are with respect to the OH-ACNH group interaction. Furthermore, for the system of API 1-Methanol, 1 pair of GIPs related to CH_3OH-$ACNO_2$ which has been initially set to zero has been fine tuned to improve the prediction of the system using the Original UNIFAC-CI model. The predictions are shown in Figure 7.16 (b). However when using Toluene as the solvent, no additional GIPs need to be fine tuned as the predictions using the available parameters are satisfactory. The SLE curve for this system is displayed in Figure 7.16 (c).

Figure 7.17 shows the SLE diagrams for API 2 in four solvents involving Heptane, Toluene, Tetrahydrofuran and Acetone. For the systems of API 2-Heptane and API 2-Toluene, the SLE predictions were obtained without fine tuning any additional GIPs. In fact, there are no fine tuned GIPs that were used to predict these systems. The predictions made are in a good agreement with the single data point for each system. However, for system involving API 2 and Tetrahydrofuran, a pair of parameters involving CH_2O-ACCONH has been fine tuned to improve the predictions so that they are closer to the experimental data point. Furthermore, another pair of parameters related to the CH_2CO-ACCO interaction has been fine tuned for system involving API 2 and Acetone. Using the same set of GIPs, the system of API 2 and Methyl Ethyl Ketone is predicted using the Original UNIFAC model and the SLE curve obtained are shown in Figure 7.18 (a).

For the system between API 2 and the alcohols (2-Propanol and Butanol), the interaction parameters involving the OH-ACCO interactions were fine tuned to improve the SLE prediction and the phase diagrams generated are illustrated in Figures 7.18 (b) and (c). Moreover, when using the methanol as the solvent for API 2, the predictions made with the initially available GIPs are not satisfactory. Therefore, the GIPs for the CH_3OH-ACCONH (which has been set to zero initially) has been fine tuned.

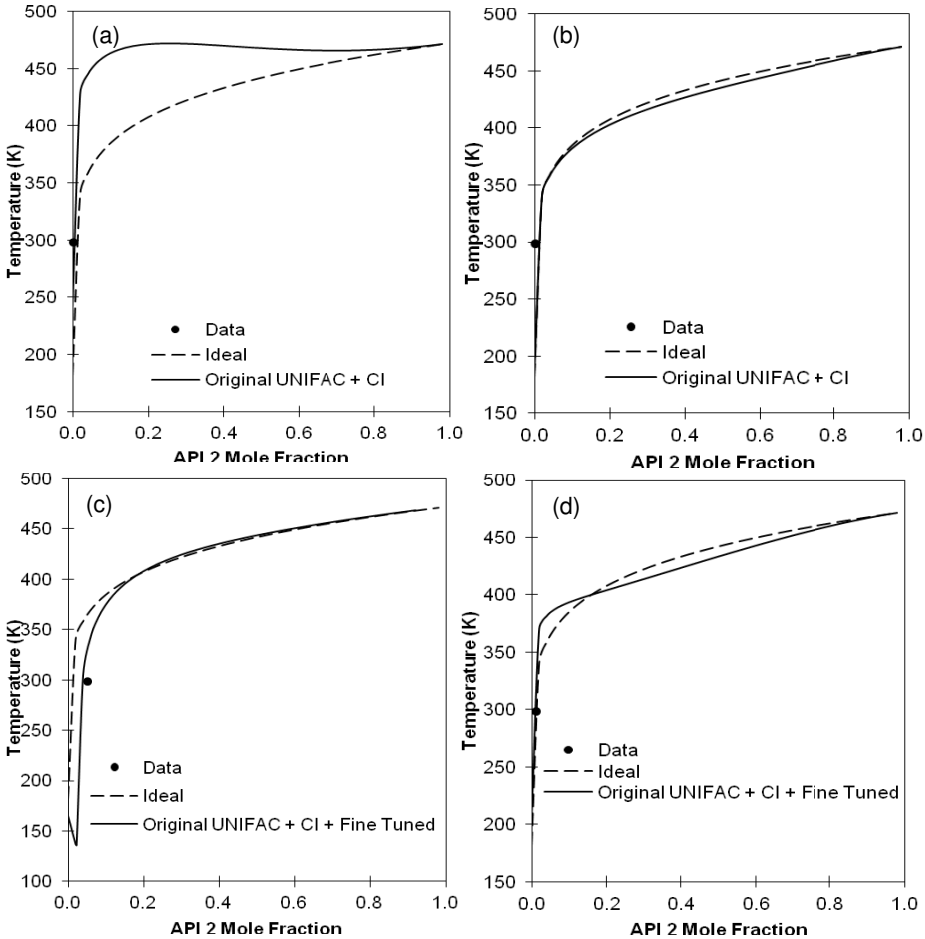


Figure 7.17: SLE Diagrams of APIs with Solvents, a) Heptane, b) Toluene, c) Tetrahydrofuran and d) Acetone Predicted using Original UNIFAC-CI using Fine Tuned GIPs. (--- represents calculations using the ideal solubility assumption).

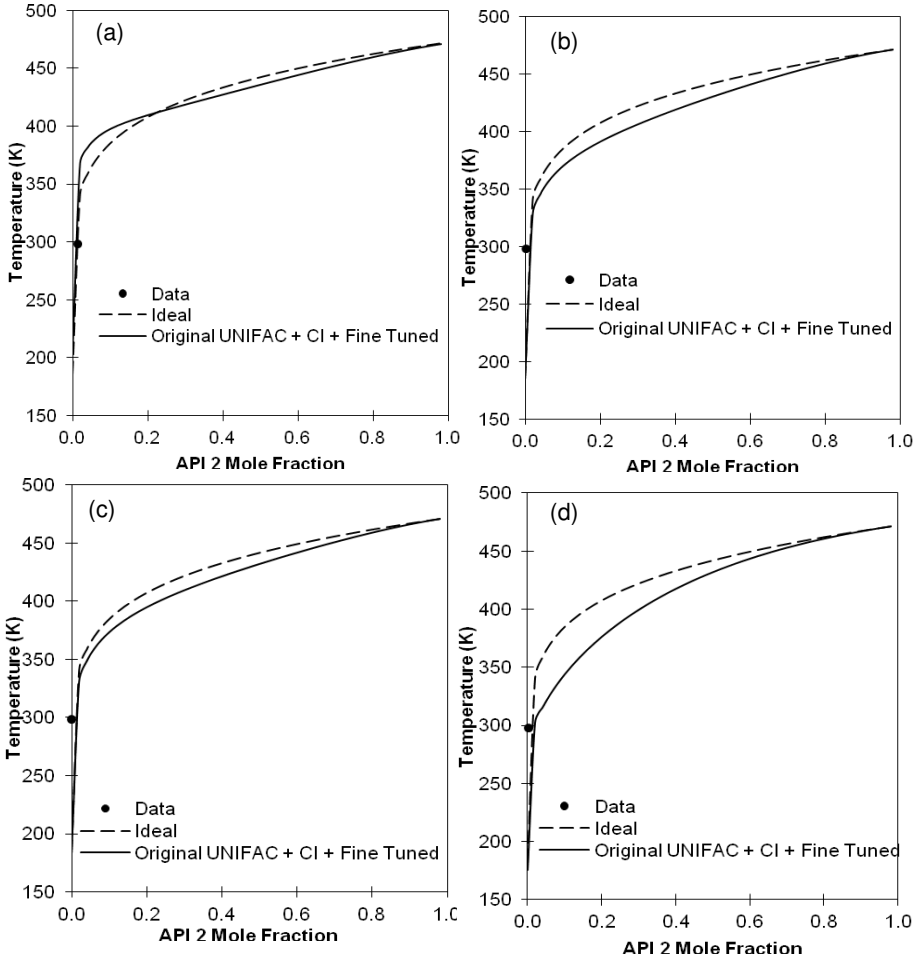


Figure 7.18: SLE Diagrams of APIs with Solvents, a) Methyl Ethyl Ketone, b) 2-Propanol, c) Butanol and d) Methanol Predicted using Original UNIFAC-CI using Fine Tuned GIPs. (--- represents calculations using the ideal solubility assumption).

Table 7.8: Group Interaction Parameters (GIPs) Table (in Kelvin) for Pharmaceutical Case Study with Fine Tuned Parameters. (GIPs labeled with * are predicted using the CI-method while GIPs labeled with # are fine tuned. The labels are placed at the right side of the GIPs).

	CH_2	ACH	$ACCH_2$	OH	CH_3OH	CH_2CO	CCOO	CH_2O	CCN	ACCL	$ACNO_2$	ACNH	ACCO	ACCONH
CH_2	0	61	77	987	697	476	232	252	597	11	543	1885*	974*	-1304*
ACH	-11	0	167	636	637	26	6	32	213	187	195	7862 #	560*	-2099*
$ACCH_2$	-70	-147	0	803	603	-52	5688	213	6096	-211	4448	1681*	1201*	-1610*
OH	156	90	26	0	-137	84	101	28	7	124	157	-4540*	734 #	-2404*
CH_3OH	17	-50	-45	249	0	23	-11	-129	53	-28	498 #	0 ¤	-11*	182 #
CH_2CO	27	140	366	165	109	0	-214	-104	482	-120	549	782 #	-371 #	2833*
CCOO	115	86	-170	245	250	372	0	-236	495	442	1927*	380*	-271 #	3442*
CH_2O	83	52	66	238	238	191	461	0	-19	135	4740 #	2405 #	656*	5909 #
CCN	25	-23	-138	185	163	-288	-267	39	0	-5	340*	1071 #	1089*	-345*
ACCL	107	-97	403	326	613	518	-171	-25	364	0	2213	2630*	1007*	242*
$ACNO_2$	5541	1824	-128	562	1810 #	-102	1643*	140 #	243*	-123	0	8958*	625*	6136*
ACNH	51*	8913 #	-311*	-70*	0 ¤	1818 #	-597*	-620 #	2364 #	27*	210*	0	1196*	257*
ACCO	-24*	272*	265*	-386 #	-459*	-78 #	-700 #	1302*	-3337*	949*	3701*	529*	0	2834*
ACCONH	-4*	16*	-848*	821*	-372 #	2182*	2452*	-864 #	196*	188*	4701*	-90*	3209*	0

There results of all the GIPs that have been fined tuned for this pharmaceutical case study are highlighted in Table 7.8. By comparing the GIPs in Table 7.7 and 7.8, we can see that from the total of 37 pairs of GIPs that were predicted using the CI-method, only 9 pairs of GIPs or 24 % were fine tuned.

By using all the GIPs reported in Table 7.8, the SLE diagrams shown in Figures 7.15-7.18 have been generated. Those individual SLE diagrams have been combined into a single diagram for API 1 and API 2 displayed in Figure 7.19 to identify which solvent provide the highest solubility of the APIs. In Figure 7.11 (a), the SLE diagrams for API 1 together with all the interested solvents have been generated. According to the figure, the highest solubility of API 1 can be found when using Tetrahydrofuran as the solvent. On the other hand, for API 2, the highest solubility can be found for system API 2-Tetrahydrofuran until 315 K. However, the solubility of API 2 is higher above 315 K when using Methanol as the solvent. Therefore, the best solvent that can be used for API 1 and API is Tetrahydrofuran. However at temperature higher than 315 K, Methanol can be used as the solvent for API 2.

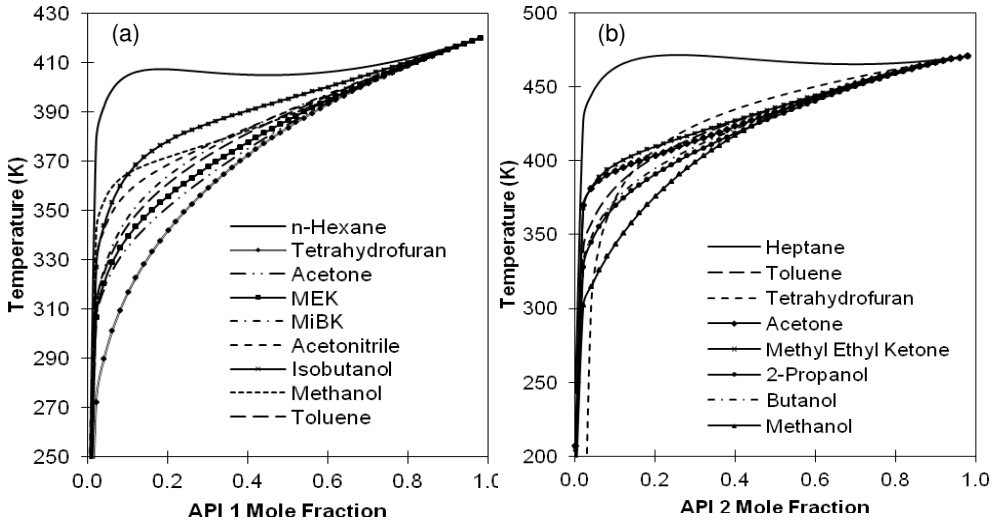


Figure 7.19: Overall SLE Diagrams for (a) API 1 and Different Solvents and (b) API 2 and Different Solvents.

7.3.3 Conclusions for this of Case Study

In this case study, the Original UNIFAC-CI model has been used to predict the solubility of APIs in several interested solvents. The purpose is to find which solvent can produce the highest solubility of the APIs under study. The systems involving the APIs cannot be fully represented by the Original UNIFAC model [15]. Therefore, using the AIPs which have been regressed for the Original UNIFAC-CI model, new groups involving ACCO, ACNH and ACCONH have been created and the related group interaction parameters (GIPs) have been predicted using the CI-method. Using the GIPs which are already available in the UNIFAC parameter table [15] together with the GIPs which have been estimated, the needed SLE diagrams can be generated.

However, one limitation of this Original UNIFAC-CI model is that there is a probability that the GIPs predicted using regressed AIPs for groups which are not considered in the parameter regression step, high negative GIPs will be obtained. When trying to predict SLE system with these highly negative GIPs (usually > -1000), convergence of the SLE calculation may be a problem. Therefore, some of those GIPs need to be fine tuned against the experimental data. The result is a master parameter table containing the original parameters, parameters predicted from the CI-method and the fine tuned parameters which can be used specifically for this case study. From a total of 37 pairs of GIPs which were initial predicted using the CI-method, only 9 pairs were fined tuned. Using the final parameter table, SLE diagrams have been generated and the best system which has the highest solubility of APIs can be determined.

7.4 Lipid Systems

Lipids which include vegetable oils and fats are very important for human nutrition and also in the chemical industry because they are a source of energy and fat-soluble vitamins. Over the past few decades, the world's fats and edible oils production has been growing rapidly, far beyond the need for human nutrition [134]. This overproduction combined with the growing consumer preferences for healthier food products and the interest in bio-fuels, has led the oleo chemical industry to face major challenges in terms of design and development of better products and more sustainable processes. However, although the oleo chemical industry is mature and based on well established processes, the complex systems that lipid compounds form, the lack of accurate predictive models for their physical properties and unit operation models for their processing have limited the use of computer-aided methods and tools for process synthesis, modeling and simulation within this industry.

Group-contribution models for the prediction of pure lipid compounds and a database of relevant properties and model parameters have been developed by Diaz-Tovar et al. [134]. The development of the property prediction models is achieved by a) identifying the most significant and widely produced edible oils/fats, as well as their corresponding representative families of chemical species, (b) molecular description of the identified chemical species in terms of the property model, (c) creating a list of the physical chemical properties needed for model-based design and analysis of edible oil and biodiesel processes, (d) collecting the available experimental data from different sources for the identified lipid compounds and their corresponding properties and (e) selecting and adopting the appropriate models to predict the necessary properties, to fill-out the gaps in the lipid-database and to make it suitable for applications with other computer-aided tools.

With the developed models and database, other than pure lipid compounds, mixtures involving lipid compounds can now be modeled and predicted using suitable models. In [135] the development of a computer aided multilevel modeling network for the systematic design and analysis of processes employing lipid technologies have been presented. This is achieved by decomposing the problem into four levels of modeling i) pure component property modeling and a lipid database of collected experimental data from industry and generated data from validated predictive property models, ii) modeling of phase behavior of relevant lipid mixtures using the UNIFAC-CI model, development of a master parameter table and calculations of the activity coefficients, iii) development of a model library consisting of new and adopted process models of unit operations involved in lipid processing technologies, validation of the developed models using operating data collected from existing process plants, and application of validated models in design and analysis of unit operations; iv) use of information and models developed as building blocks in the development of methods and tools for computer-aided synthesis and design of process flowsheets (CAFD).

In this section, the aim is to highlight the second level of the multilevel modeling network involving phase equilibria prediction of lipid systems using the Original UNIFAC-CI model that we have developed. There are two sub-case studies that will be presented in the next section and they involve phase equilibria calculations of a multicomponent system of a deodorization process and predictions of several VLE systems involving fatty acid esters and glycerols with alcohols.

7.4.1 Background of Sub-Case Study 1

Processing of edible oils/lipids involves separation processes such as fatty acids distillation, fatty alcohols fractionation, physical refining and deodorization of edible oils. In all these processes, phase equilibria play an important role and therefore they need to be predicted reliably. The information obtained from the equilibrium calculations is important for the design and operations of equipments with simulation results to evaluate the separation processes [136]. In this first case study, we will highlight the use of the Original UNIFAC-CI model to predict the activity coefficients of a multicomponent system involved in the deodorization process of palm oil. The deodorization process is necessary to remove free fatty acids (FFA) and odoriferous compounds which naturally present in the oil or created during processing by vaporizing the compounds through the injection of steam at high temperatures and low pressures and thus separating them from the triacylglycerol (TAG) mixtures [137-138]. Since the deodorization process occurs at high temperatures and low pressures, the vapor-liquid equilibria (VLE) can be simplified by setting the vapor phase as ideal and assuming that only the liquid phase affects the non-ideality of the system.

Table 7.9: Palm Oil Composition with its UNIFAC Group Representation.

Compound*	Type	Mass (%)	UNIFAC Group Representation
PPP		5.50	CH_3 (3), CH_2 (41), CH (1), CH_2COO (3)
POP		36.32	CH_3 (3), CH_2 (41), CH (1), CH=CH (1), CH_2COO (3)
POS	TAG	6.09	CH_3 (3), CH_2 (43), CH (1), CH=CH (1), CH_2COO (3)
PLiP		9.90	CH_3 (3), CH_2 (39), CH (1), CH=CH (2), CH_2COO (3)
POO		20.80	CH_3 (3), CH_2 (41), CH (1), CH=CH (2), CH_2COO (3)
POLi		9.53	CH_3 (3), CH_2 (39), CH (1), CH=CH (3), CH_2COO (3)
PP-OH		2.25	CH_3 (2), CH_2 (28), CH (1), OH (1), CH_2COO (2)
PO-OH	DAG	4.46	CH_3 (2), CH_2 (28), CH (1), CH=CH (1), OH (1), CH_2COO (2)
PLi-OH		1.04	CH_3 (2), CH_2 (26), CH (1), CH=CH (2), OH (1), CH_2COO (2)
P-OH-OH		0.30	CH_3 (1), CH_2 (15), CH (1), OH (2), CH_2COO (1)
O-OH-OH	MAG	0.17	CH_3 (1), CH_2 (15), CH (1), CH=CH (1), OH (2), CH_2COO (1)
Li-OH-OH		0.04	CH_3 (1), CH_2 (13), CH (1), CH=CH (2), OH (2), CH_2COO (1)
Palmitic Acid		1.75	CH_3 (1), CH_2 (15), COOH (1)
Oleic Acid	FFA	1.41	CH_3 (1), CH_2 (14), CH=CH (1), COOH (1)
Linoleic Acid		0.34	CH_3 (1), CH_2 (12), CH=CH (2), COOH (1)
Tocopherol		0.10	CH_3 (5), CH_2 (11), CH (3), AC (2), $ACCH_3$ (3), ACOH (1), C-O (1)

*Abbreviations: P is palmitic acid, S is stearic acid, O is oleic acid and Li is linoleic acid
*Example: PPP is tripalmitin which is a combination of three palmitic acid (P) derivatives

The first step is to identify the compounds involved in the system to be studied. According to Ceriani et al. [137], palm oil can be represented by six triacylglycerols (TAG), three diacylglycerols (DAG), three monoacylglycerols (MAG), three free fatty acids (FFAs) and tocopherol. Next, the identified compounds need to be decomposed into several UNIFAC functional groups. Table 7.12 shows the palm oil composition and its group representation where the number in the bracket after the UNIFAC group indicates the number of occurrences of that group in each compound. The TAGs can be represented by the groups CH_3, CH_2, CH (UNIFAC main group 1), CH_2COO

(UNIFAC main group 11) and also CH=CH (UNIFAC main group 2) depending on whether the unsaturated fatty acid chain appears in the compound or not. DAGs and MAGs can be represented by groups CH_3, CH_2, CH, CH_2COO, OH (UNIFAC main group 5) and CH=CH while FFAs are represented by groups CH_3, $_{CH2}$, COOH (UNIFAC main group 20) and also CH=CH.

7.4.2 Phase Equilibria Predictions and Parameter Table for Sub-Case Study 1

The group interaction parameters (GIPs) needed for the calculations of the activity coefficients of the compounds are presented in Table 7.11 which are reestimated using the UNIFAC-CI method. For the reference UNIFAC model [15], the GIPs are regressed directly against experimental data and can be used to represent compounds that are not included in the regression step. However, for mixtures or compounds for which GIPs are missing, the normal procedure will be to collect the necessary experimental data and regress the missing parameters with the new data which can be time consuming, costly and even infeasible to conduct the experiments. To overcome this problem, the Original UNIFAC-CI [16-18] model was developed by taking into account the interactions between atoms (instead of groups) through connectivity indices. Therefore, atom connectivities can be used to represent functional groups that are not present in the regression step. In this way, when a specific GIP is missing, the connectivity indices for that specific group are calculated and using the corresponding regressed atom interaction parameters (AIPs), the missing GIPs are predicted through an established relation (see Equations (2.37)-(2.38)). Currently the model parameters are available for systems containing C, H, O, N Cl and S atoms. The advantage of this UNIFAC-CI model is that phase equilibria for any lipid systems which contain these five atoms can be predicted even when the GIPs for the reference model are not available.

For this deodorization problem, all GIPs for the reference UNIFAC parameter table are available and therefore the parameters are re-estimated using the UNIFAC-CI method to examine the applicability of this model. The GIPs of the Original UNIFAC model are presented in Table 7.10.

Table 7.10: Group Interaction Parameters (in Kelvin) Predicted using Original UNIFAC for Deodorization Process of Palm Oil.

	CH_2	**C=C**	**ACH**	**$ACCH_2$**	**OH**	**ACOH**	**CCOO**	**CH_2O**	**COOH**
CH_2	0.0	86.0	61.1	76.5	986.5	1333.0	232.1	251.5	663.5
C=C	-35.4	0.0	38.8	74.2	524.1	526.1	37.9	214.5	318.9
ACH	-11.1	3.4	0.0	167.0	636.1	1329.0	6.0	32.1	537.4
$ACCH_2$	-69.7	-113.6	-146.8	0.0	803.2	884.9	5688.0	213.1	872.3
OH	156.4	457.0	89.6	25.8	0.0	-259.7	101.1	28.1	199.0
ACOH	275.8	217.5	25.3	244.2	-451.6	0.0	-449.4	-162.9	408.9
CCOO	114.8	132.1	85.8	-170.0	245.4	-36.7	0.0	-235.7	660.2
CH_2O	83.4	26.5	52.1	65.7	237.7	-178.5	461.3	0.0	664.6
COOH	315.3	1264.0	62.3	89.9	-151.0	-11.0	-256.3	-338.5	0.0

Table 7.11: Group Interaction Parameters (in Kelvin) Predicted using Original UNIFAC-CI for Deodorization Process of Palm Oil.

	CH_2	C=C	ACH	$ACCH_2$	OH	ACOH	CCOO	CH_2O	COOH
CH_2	0.0	172.1	124.1	241.4	956.5	320.1	382.2	160.6	529.3
C=C	-94.5	0.0	1136.0	950.7	274.7	651.9	900.7	641.1	1013.1
ACH	-51.3	-274.6	0.0	145.8	586.6	-220.9	39.3	-231.7	151.7
$ACCH_2$	-39.1	1000.9	320.5	0.0	674.1	416.7	680.1	406.0	792.55
OH	26.7	481.2	6.71	299.9	0.0	-373.5	192.1	16.3	1730.6
ACOH	46.2	547.4	904.0	995.9	226.58	0.0	-200.7	-591.71	98.3
CCOO	40.3	-265.4	101.9	183.1	217.9	409.9	0.0	-421.04	266.1
CH_2O	66.7	65.9	384.9	514.34	257.1	962.7	2134.3	0.0	164.56
COOH	111.1	73.3	342.3	521.78	-92.6	480.6	1516.4	824.13	0.0

Table 7.12: Activity Coefficients at 250 ^{0}C and 3.5 mmHg Calculated using UNIFAC-CI and Reference UNIFAC Models.

Compound	Type	Activity Coefficients	
		UNIFAC-CI	Reference UNIFAC
PPP	TAG	1.01	1.01
POP		1.00	1.00
POS		1.00	1.00
PliP		1.00	1.00
POO		1.00	1.00
POLi		1.00	1.00
PP-OH	DAG	1.10	1.15
PO-OH		1.13	1.15
PLi-OH		1.18	1.18
P-OH-OH	MAG	2.15	2.45
O-OH-OH		2.19	2.57
Li-OH-OH		2.39	0.77
Palmitic Acid	FFA	1.05	0.77
Oleic Acid		1.01	0.75
Linoleic Acid		0.93	0.72
Tocopherol		1.73	2.58

Table 7.12 shows the activity coefficients (related to VLE) for each compound of the multicomponent system (their compositions are reported in Table 7.9) at 250 ^{0}C and 3.5 mmHg, often used operation conditions of the deodorization process. The calculations are based on the Original UNIFAC-CI model and are compared with the reference UNIFAC using original GIPs. The results show that the activity coefficients calculated by both models for TAGs are almost equal to unity. For the DAGs, the activity coefficients are almost the same with both models. Furthermore for the MAGs, more or less the same activity coefficient values are obtained by both models except for P-OH-OH. For this MAG, the activity coefficient predicted by the reference UNIFAC model is 0.77 while the Original UNIFAC-CI model produced a value of 2.39. In addition to that, the FFAs the average activity coefficient values calculated by the reference UNIFAC and UNIFAC-CI models are 0.75 and 0.99 respectively which are a bit different.

For tocopherol, initially the molecular structure cannot be fully represented by the UNIFAC group. The molecular structure of tocopherol is displayed in Figure 7.20. However, a new subgroup, C-O has been added under UNIFAC group 13, CH_2O in order to be able to present the molecular structure. The volume, R_k and surface area, Q_k parameters for this new subgroup are obtained from Bondi and the group interaction parameters involving main group CH_2O are used for predictions. The activity coefficients obtained for tocopherol using both reference UNIFAC and Original UNIFAC-CI models are 2.58 and 1.73 respectively. Overall the activity coefficients calculated by Original UNIFAC-CI are in good agreement with the values predicted by the reference model except for some compounds where slight deviations exist.

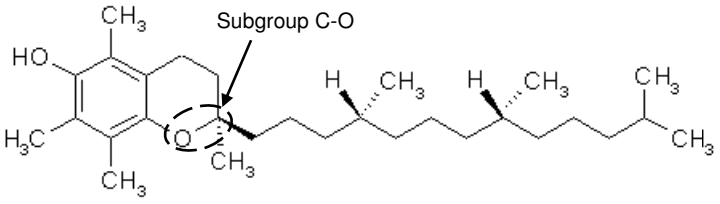


Figure 7.20: Molecular Structure of Tocopherol and the Subgroup C-O.

7.4.3 Background of Sub-Case Study 2

In Europe, the most used biofuel for road vehicles is biodiesel. In the production of biodiesel, there are different separation and purification processes that are needed especially for the unit operations that contain a lot of glycerol for example for the recovery of unreacted alcohol and the removal of water. In order to design the recovery and purification systems in a successful way, the phase equilibria knowledge for the compounds involving water-glycerol and alcohol-glycerol are needed.

In addition to that, it is also of great interest to have the capability to describe the phase equilibria of binary systems formed by a fatty acid ester and an alcohol. These mixtures can be found after the transesterification unit where they must be separated in order to purify the biodiesel stream and to recover the unreacted alcohol.

VLE experimental data for systems involving fatty acid ester-alcohol, glycerol-water and glycerol-alcohol are available from [139] and [140]. Using the interaction parameters which have been obtained for the Original UNIFAC-CI model, the predictions using these models are compared with the experimental data and also the predictions made by the reference UNIFAC model since all the lipid compound can be fully represented by the available UNIFAC groups.

7.4.4 Phase Equilibria Predictions and Parameter Table for Sub-Case Study 2

The group interaction parameters needed for the prediction of the systems described in Section 7.4.3 for both Original UNIFAC and Original UNIFAC-CI models are shown in Tables 7.13 and 7.14. Using those GIPs the prediction of fatty acid esters and alcohol are presented in Figure 7.21. In this figure, the Original UNIFAC-CI model is labeled as UNIFAC-CI (VLE/SLE).

Table 7.13: Group Interaction Parameters (in Kelvin) for Original UNIFAC for Sub-Case Study 2.

	CH_2	C=C	OH	CH_3OH	H_2O	CCOO
CH_2	0.0	86.0	986.5	697.2	1318.0	232.1
C=C	-35.4	0.0	524.1	787.6	270.6	37.9
OH	156.4	457.0	0.0	-137.1	353.5	101.1
CH_3OH	16.5	-12.5	249.1	0.0	-181.0	-10.72
H_2O	300.0	496.1	-229.1	324.5	0.0	78.9
CCOO	114.8	132.1	245.4	249.6	200.8	0.0

Table 7.14: Group Interaction Parameters (in Kelvin) for Original UNIFAC-CI for Sub-Case Study 2.

	CH_2	C=C	OH	CH_3OH	H_2O	CCOO
CH_2	0.0	172.1	956.5	631.3	495.1	382.2
C=C	-94.5	0.0	274.7	219.9	0.0	900.7
OH	26.7	481.2	0.0	97.6	-47.3	192.1
CH_3OH	-72.6	3956.9	-65.4	0.0	-421.6	-53.7
H_2O	72.4	6429.6	181.4	159.2	0.0	
CCOO	40.3	-265.4	217.9	301.9		0.0

From the VLE phase diagrams displayed in Figure 7.21, we can see that the predictions made by the UNIFAC-CI (VLE/SLE) model are as good as the reference Original UNIFAC model and also in a very good agreement with experimental data. In fact for several systems, for example for the systems of Methyl Laurate-Ethanol, Methyl Myristate-Ethanol, and Methyl Oleate-Ethanol, the predictions made by the CI-model are slightly closer to the experimental data than the reference model.

Furthermore, the phase diagrams involving alcohol-glycerol and water-glycerol are presented in Figure 7.22. The predictions made by the UNIFAC-CI (VLE/SLE) model are also as successful as the reference UNIFAC model. However, the predictions made by both models were unable to accurately capture all the experimental data. Nevertheless, the VLE diagrams generated from the UNIFAC models overall are satisfactory and acceptable.

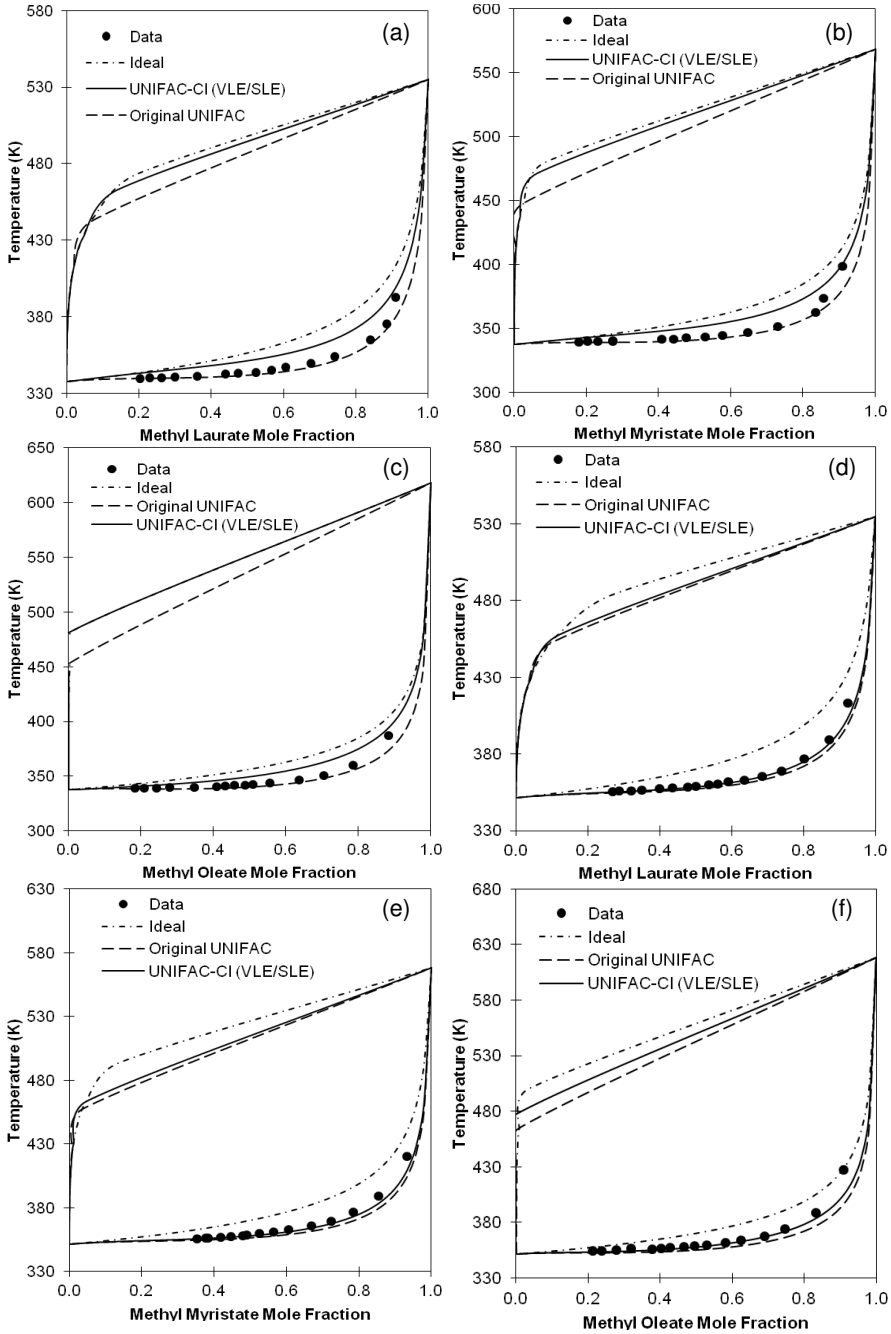


Figure 7.21: VLE Phase Diagrams of (a) Methyl Laurate-Methanol, (b) Methyl Myristate-Methanol, (c) Methyl Oleate-Methanol, (d) Methyl Laurate-Ethanol, (e) Methyl Myristate-Ethanol, (f) Methyl Oleate-Ethanol.

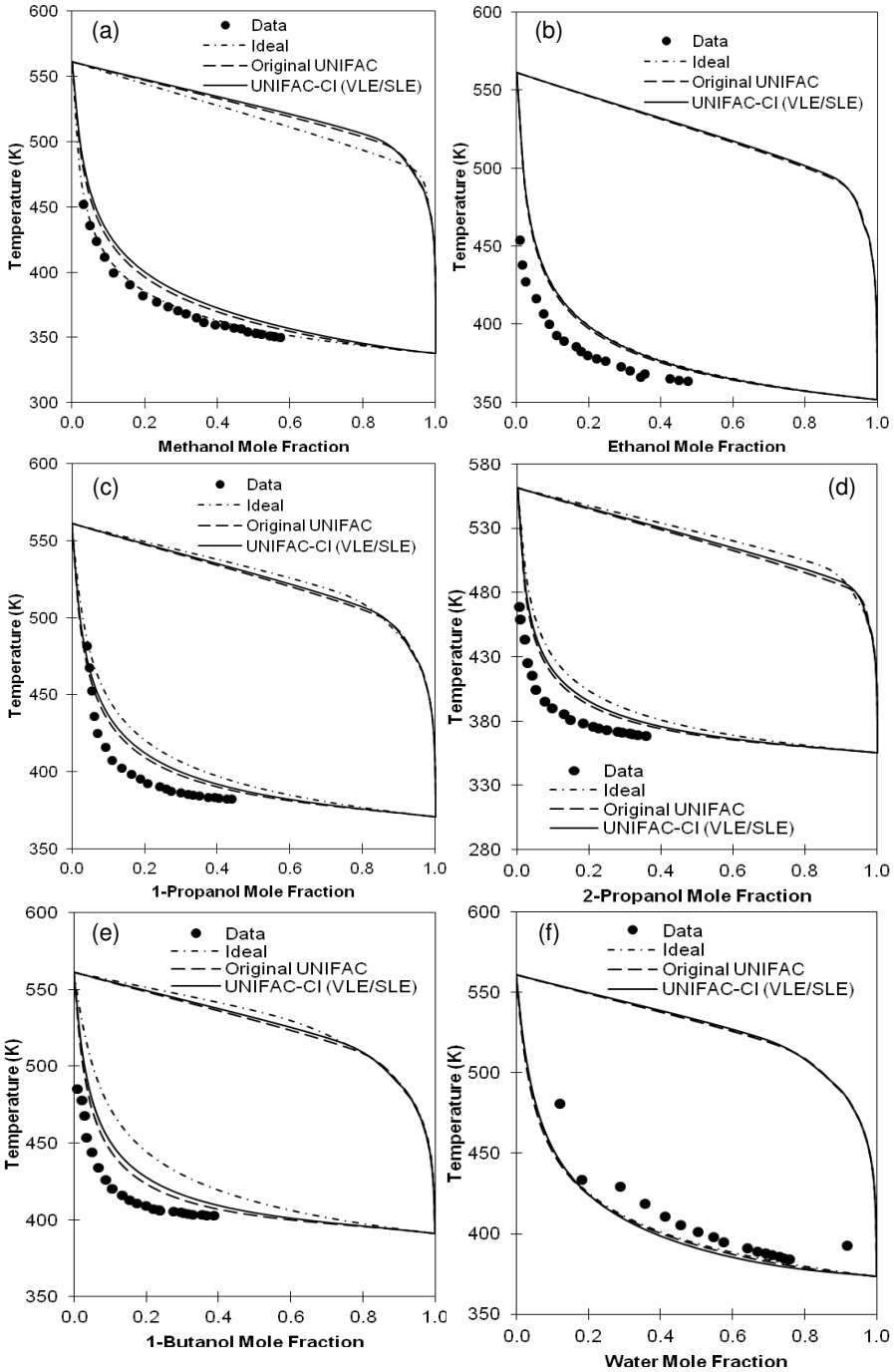


Figure 7.22: VLE Phase Diagrams of (a) Methanol-Glycerol, (b) Ethanol -Glycerol, (c) 1-Propanol-Glycerol, (d) 2-Propanol- Glycerol, (e) 1-Butanol- Glycerol, (f) Water- Glycerol.

7.4.3 Conclusions for the Lipid of Case Studies

For lipid systems case study, 2 sub-case studies have been presented involving the phase equlibria investigation of a multicomponent system of a palm oil deodorization process and the predictions of VLE systems involving fatty acid ester-alcohols and alcohols-glycerols. Based on sub-case study 1 and 2, we can conclude that the Original UNIFAC-CI model can be used to predict lipid systems reliably and with high accuracy when comparing the predictions with the experimental data.

7.5 Systems with Missing GIPs and Azeotropic Systems

In this section, the VLE phase equilibria of a system, Ethyl Formate-Phenol at constant temperatures of 300 and 320 K will be predicted using the UNIFAC model. However, there are missing parameters in the Original UNIFAC [15] parameter table with respect to ACOH-HCOO group interactions. Therefore, these missing parameters are predicted using the CI-method.

Table 7.15: Parameter Table (in Kelvin) of the Original UNIFAC Model with the Missing Parameters.

	CH_2	**ACH**	**ACOH**	**HCOO**
CH_2	0.00	61.13	1333.00	507.00
ACH	-11.12	0.00	1329.00	287.10
ACOH	275.80	25.34	0.00	N/A
HCOO	329.30	18.12	N/A	0.00

The calculations of the GIPs with respect to ACOH-HCOO interactions are shown below:

Calculations of the GIPs between ACOH and HCOO

1) The stoichiometry and the CI values needed are given in Table 7.16

Table 7.16: Atom Stoichiometry and CI Values for ACOH-HCOO.

Group	n_C	n_O	$^{v}\chi^{0}$	$^{v}\chi^{1}$	$^{v}\chi^{2}$
ACOH	1	1	0.9472	0.2236	0.0000
HCOO	1	2	1.3938	0.4714	0.0962

2) Using Equations (2.37)-(2.42) and the atom interaction parameters (AIPs) presented in Table 7.20, the group interaction parameters $a_{ACOH\text{-}HCOO}$ and $a_{HCOO\text{-}ACOH}$ are calculated as follows:

Table 7.17: Atom Interaction Parameters (AIPs) Needed for the Calculation of GIPs Related to $a_{CH2\text{-}ACNH}$ and $a_{ACNH\text{-}CH2}$.

AIPs	Values	AIPs	Values
b_{C-C}	969.1222	d_{O-O}	452.3623
c_{C-C}	-112.3279	e_{O-O}	-384.7823
d_{C-C}	88.5224	$\overline{b_{C-C}}$	-161.3158
e_{C-C}	-111.7174	$\overline{c_{C-C}}$	-259.6039
b_{C-O}	-1163.8140	$\overline{d_{C-C}}$	299.2265
c_{C-O}	-15.1636	$\overline{b_{C-O}}$	323.4421
d_{C-O}	59.0360	$\overline{c_{C-O}}$	-120.2872
e_{C-O}	69.1201	$\overline{d_{C-O}}$	236.8295
b_{O-C}	-317.6756	$\overline{b_{O-C}}$	32.5118
c_{O-C}	-789.8060	$\overline{c_{O-C}}$	-82.1519
d_{O-C}	-757.6192	$\overline{d_{O-C}}$	252.7937
e_{O-C}	363.2116	$\overline{b_{O-O}}$	-437.0453
b_{O-O}	643.3884	$\overline{c_{O-O}}$	28.5467
c_{O-O}	1126.7728	$\overline{d_{O-O}}$	-26.1669

$$
\begin{aligned}
a_{ACOH-HCOO} = {} & b_{C-C}\frac{n_C^{(ACOH)}\chi_0^{(HCOO)} - n_C^{(HCOO)}\chi_0^{(ACOH)}}{\chi_0^{(HCOO)}\chi_0^{(ACOH)}} + c_{C-C}\frac{n_C^{(ACOH)}\chi_1^{(HCOO)} - n_C^{(HCOO)}\chi_0^{(ACOH)}}{\chi_1^{(HCOO)}\chi_0^{(ACOH)}} + \\
& d_{C-C}\frac{n_C^{(ACOH)}\chi_1^{(HCOO)} - n_C^{(HCOO)}\chi_1^{(ACOH)}}{\chi_1^{(HCOO)}\chi_1^{(ACOH)}} + e_{C-C}\frac{n_C^{(ACOH)}\chi_2^{(HCOO)} - n_C^{(HCOO)}\chi_0^{(ACOH)}}{\chi_2^{(HCOO)}\chi_0^{(ACOH)}} + \\
& b_{C-O}\frac{n_C^{(ACOH)}\chi_0^{(HCOO)} - n_O^{(HCOO)}\chi_0^{(ACOH)}}{\chi_0^{(HCOO)}\chi_0^{(ACOH)}} + c_{C-O}\frac{n_C^{(ACOH)}\chi_1^{(HCOO)} - n_O^{(HCOO)}\chi_0^{(ACOH)}}{\chi_1^{(HCOO)}\chi_0^{(ACOH)}} + \\
& d_{C-O}\frac{n_C^{(ACOH)}\chi_1^{(HCOO)} - n_O^{(HCOO)}\chi_1^{(ACOH)}}{\chi_1^{(HCOO)}\chi_1^{(ACOH)}} + e_{C-O}\frac{n_C^{(ACOH)}\chi_2^{(HCOO)} - n_O^{(HCOO)}\chi_0^{(ACOH)}}{\chi_2^{(HCOO)}\chi_0^{(ACOH)}} + \\
& b_{O-O}\frac{n_O^{(ACOH)}\chi_0^{(HCOO)} - n_O^{(HCOO)}\chi_0^{(ACOH)}}{\chi_0^{(HCOO)}\chi_0^{(ACOH)}} + c_{O-O}\frac{n_O^{(ACOH)}\chi_1^{(HCOO)} - n_O^{(HCOO)}\chi_0^{(ACOH)}}{\chi_1^{(HCOO)}\chi_0^{(ACOH)}} + \\
& d_{O-O}\frac{n_O^{(ACOH)}\chi_1^{(HCOO)} - n_O^{(HCOO)}\chi_1^{(ACOH)}}{\chi_1^{(HCOO)}\chi_1^{(ACOH)}} + e_{O-O}\frac{n_O^{(ACOH)}\chi_2^{(HCOO)} - n_O^{(HCOO)}\chi_0^{(ACOH)}}{\chi_2^{(HCOO)}\chi_0^{(ACOH)}} +
\end{aligned}
$$

$$a_{HCOO-ACOH} = \overline{b_{C-C}} \frac{n_C^{(HCOO)}\chi_0^{(ACOH)} - n_C^{(ACOH)}\chi_0^{(HCOO)}}{\chi_0^{(ACOH)}\chi_0^{(HCOO)}} + \overline{c_{C-C}} \frac{n_C^{(HCOO)}\chi_1^{(ACOH)} - n_C^{(ACOH)}\chi_0^{(HCOO)}}{\chi_1^{(ACOH)}\chi_0^{(HCOO)}} +$$
$$\overline{d_{C-C}} \frac{n_C^{(HCOO)}\chi_1^{(ACOH)} - n_C^{(ACOH)}\chi_1^{(HCOO)}}{\chi_1^{(ACOH)}\chi_1^{(HCOO)}} + \overline{b_{C-O}} \frac{n_C^{(HCOO)}\chi_0^{(ACOH)} - n_O^{(ACOH)}\chi_0^{(HCOO)}}{\chi_0^{(ACOH)}\chi_0^{(HCOO)}} +$$
$$\overline{c_{C-O}} \frac{n_C^{(HCOO)}\chi_1^{(ACOH)} - n_O^{(ACOH)}\chi_0^{(HCOO)}}{\chi_1^{(ACOH)}\chi_0^{(HCOO)}} + \overline{d_{C-O}} \frac{n_C^{(HCOO)}\chi_1^{(ACOH)} - n_O^{(ACOH)}\chi_1^{(HCOO)}}{\chi_1^{(ACOH)}\chi_1^{(HCOO)}} +$$
$$\overline{b_{O-C}} \frac{n_O^{(HCOO)}\chi_0^{(ACOH)} - n_C^{(ACOH)}\chi_0^{(HCOO)}}{\chi_0^{(ACOH)}\chi_0^{(HCOO)}} + \overline{c_{O-C}} \frac{n_O^{(HCOO)}\chi_1^{(ACOH)} - n_C^{(ACOH)}\chi_0^{(HCOO)}}{\chi_1^{(ACOH)}\chi_0^{(HCOO)}} +$$
$$\overline{d_{O-C}} \frac{n_O^{(HCOO)}\chi_1^{(ACOH)} - n_C^{(ACOH)}\chi_1^{(HCOO)}}{\chi_1^{(ACOH)}\chi_1^{(HCOO)}} + \overline{b_{O-O}} \frac{n_O^{(HCOO)}\chi_0^{(ACOH)} - n_O^{(ACOH)}\chi_0^{(HCOO)}}{\chi_0^{(ACOH)}\chi_0^{(HCOO)}} +$$
$$\overline{c_{O-O}} \frac{n_O^{(HCOO)}\chi_1^{(ACOH)} - n_O^{(ACOH)}\chi_0^{(HCOO)}}{\chi_1^{(ACOH)}\chi_0^{(HCOO)}} + \overline{d_{O-O}} \frac{n_O^{(HCOO)}\chi_1^{(ACOH)} - n_O^{(ACOH)}\chi_1^{(HCOO)}}{\chi_1^{(ACOH)}\chi_1^{(HCOO)}} +$$

3) By substituting all the AIPs, the atom stoichiometry and the CI values, the GIPs obtained are 260.41 and 68.79 respectively for $a_{ACOH\text{-}HCOO}$ and $a_{HCOO\text{-}ACOH}$.

Table 7.18: Parameter Table (in Kelvin) of the Original UNIFAC Model with the Predicted Parameters. (labeled with *).

	CH_2	**ACH**	**ACOH**	**HCOO**
CH_2	0.00	61.13	1333.00	507.00
ACH	-11.12	0.00	1329.00	287.10
ACOH	275.80	25.34	0.00	260.41*
HCOO	329.30	18.12	68.79*	0.00

The complete parameter table together with the GIPs which were predicted using the CI-method is presented in Table 7.18. Using those parameters, the VLE phase diagrams of Ethyl Formate-Phenol are predicted using the Original UNIFAC model (with the predicted GIPs) and displayed in Figure 7.23. From the generated phase diagrams, we can see that the prediction made by the UNIFAC model is in very good agreement with the pressure-vapor composition (P-y) data at both temperatures. However, the model was unable to capture accurately the pressure-liquid composition (P-x) data. Nevertheless the predictions are overall in satisfactory agreement to the experimental data.

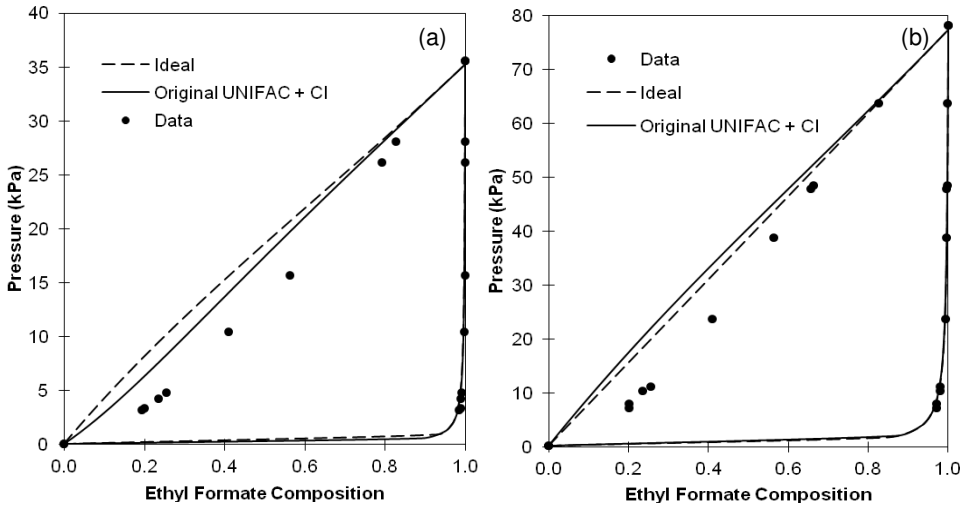


Figure 7.23: VLE Diagram of Ethyl-Formate and Phenol at (a) 300 K and (b) 320 K.

Another system that that we would like to present is the system of Dimethyl Carbonate-Methanol. All the Original UNIFAC parameters are already available for this system. However, this system has an azeotrope which is not well predicted by the Original UNIFAC model. Therefore, the original GIPs are reestimated using the CI-method. All the GIPs that have been reestimated are presented in Table 7.19.

Table 7.19: Group Interaction Parameters (in Kelvin) of Original UNIFAC-CI Model for Dimethyl Carbonate-Methanol System.

	CH_2	CH_3OH	CH_2O	COO
CH_2	0.00	631.25	160.56	430.61
CH_3OH	-72.62	0.00	-58.05	279.48
CH_2O	66.67	193.99	0.00	-163.43
COO	119.00	1088.80	142.18	0.00

Using all the GIPs that are reported in Table 7.19, the VLE phase diagram involving the Dimethyl Carbonate-Methanol system at 1 atm is presented in Figure 7.24, compared with the predictions made by the Original UNIFAC and the experimental data. The predictions made by the Original UNIFAC do not generate any azeotrope. On the other hand, when the Original UNIFAC-CI is used to predict the system, a VLE diagram with an azeotrope is produced with the same trends as the experimental data which is better than the reference UNIFAC model. Unfortunately, the VLE diagram generated was unable to capture quantitatively accurate the experimental data and also the azeotrope. Nevertheless, the predictions made by the CI-model are better than the Original UNIFAC and close to the experimental data.

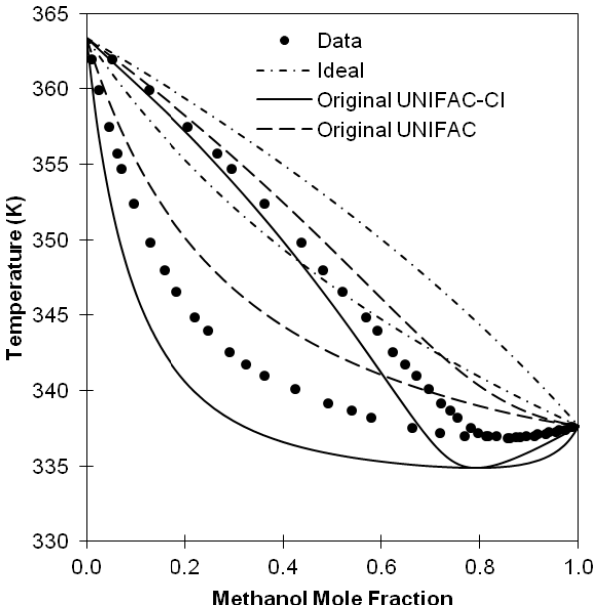


Figure 7.24: VLE Diagram of Dimethyl Carbonate-Methanol at 1 atm.

7.6 Overall Conclusions from All Case Studies

In this chapter, the application of the developed Original UNIFAC-CI model has been highlighted in different case studies involving the design of a working solution for the hydrogen peroxide production, solubility investigation involving pharmaceutical systems, phase equilibria predictions of lipids system and examples of systems with missing parameters and also prediction of azeotropic mixture.

In the first two case studies, new functional groups have been created in order to be able to fully represent the solute's molecule under study. For the new groups, the valence connectivity indices and the interactions (GIPs) with other groups have been calculated. However, one of the problems is that some of the GIPs obtained have high negative values (> -1000). These high negative values usually will cause convergence problem for the SLE calculations. Therefore, one of the solutions is to fine tune one or more of the GIPs which have been predicted using the CI-method against experimental data. Good results have been obtained using the fine tuned GIPs. Using the GIPs which were obtained from Hansen et al. [15], predicted from the CI-method and fine tuned against experimental, a master parameter table specifically for the case study have been obtained. This parameter table can be used for phase equilibria prediction of other systems involving the same solute with other molecules/solvent having the same groups listed in the table.

In addition, the applicability of the Original UNIFAC-CI also has been highlighted in lipid systems for two sub-case studies involving the phase equlibria investigation of a palm oil deodorization process and the predictions of VLE systems involving fatty acid ester-alcohols and alcohols-glycerols. This case study shows that the performance of the CI-model using all the CI-generated

GIPs is as good as the reference Original UNIFAC model using GIPs reported in [15] when predicting the acitivity coefficients of the multicomponent systems in the deodorization process and also the VLE phase diagrams of the fatty acid ester-alcohols and alcohols-glycerols systems. Furthermore, the use of CI-generated GIPs to predict the missing interaction parameters involving ethyl formate-phenol has been highlighted showing good agreement with the experimental data. In the final case study involving phase equilibria of an azeotropic system (methyl carbonate-methanol), the predictions obtained by the Original UNIFAC-CI model using CI-generated GIPs are better compared to the reference UNIFAC model.

In these case studies, the limitations of the model with its predicted GIPs have been highlighted and we discuss methods of solving the problems encountered. In addition to that, the importance of the CI-model to increase the application range of the Original UNIFAC to investigate phase equilibria of different kind of systems has also been highlighted.

CHAPTER 8

CONCLUSIONS AND FUTURE WORK

6.1 Conclusion and Discussion

In this work, analysis and further development of the GC^{Plus} models for mixtures of organic chemical systems developed previously by Gonzalez et al [16-18] have been carried out. In the initial stage of the PhD project, the Original UNIFAC-CI model developed in [18] has been further analyzed in order to investigate why the correlation errors are high for some systems involving C, H and O atoms. Therefore, after the problematic systems have been identified, the model expression, the related parameters and also the contribution of each UNIFAC group to the systems have been further checked. Most of the problematic cases are with respect to the alkanes-aldehydes systems. From the investigation and comparison of the UNIFAC group contributions of the alkanes-aldehydes systems predicted by the CI-model with the reference Original UNIFAC model, which is close to the experimental data, high deviation are especially observed for aldehyde-containing systems particularly at lower concentrations. Therefore, several possible solutions have been listed and tested which include: 1) using higher weights for data points at lower concentrations in the objective function for the problematic systems, 2) assigning higher order CI parameter for group CHO and 3) including activity coefficients at infinite dilution data when regressing the AIPs. The first solution has been implemented but was unsuccesfull in improving the correlation. Furthermore, the second solution has been implemented where a second order CI were introduced for the CHO group, specifically only for group interactions involving CH_2-CHO (which describes the interaction between the alkanes and aldehydes). After the higher order CI has been introduced, the regression for systems involving C, H and O atom was done again. Using this solution, the correlation errors involving the alkanes-aldehydes systems have been reduced. The details of this work were reported in Chapter 3.

Since the regression involving systems with C, H and O atoms have been carried out again, the Original UNIFAC-CI model was subsequestly further developed by regressing nitrogenated, chlorinated and sulfurated systems involving C, H, O, N, Cl and S atoms. In addition to this version, another set of atom interaction parameters (AIPs) have been regressed against VLE data using a quality assessment algorithm, Q_{VLE} which combines four widely used VLE consistency tests (the Herington, Van Ness, Infinite Dilution and Differential tests). This quality factors are used as a weighting factor for each system in the objective function for the regression of AIPs. The idea is that, for some VLE systems that failed certain tests while passing others, they are given lower weights instead of being removed from the regression step. The details of the regression work were presented in Chapter 4. When comparing both version of AIPs, the addition of the Q_{VLE} factor in the regression of the AIPs has slightly reduced the correlation errors. Furthermore, it is also useful in identifying anomalous systems which can be problematic in the parameter estimation and can results to parameters which are not accurately representing the systems used for the regression.

Furthermore, in Chapter 5 another set of AIPs has been generated by regressing the parameters simultaneously against both VLE and SLE data. The correlation errors obtained are compared with the regression made only against VLE data and the prediction errors of the Original UNIFAC model. The purpose of this work is to investigate whether by including SLE systems in the parameter estimation, we can improve SLE calculations while preserving the prediction accuracy for VLE systems. A total of 258 SLE systems involving 502 data points have been included in the parameter estimation while the same amount of VLE systems as Chapter 4 are used (360 systems with a total of 5961 data points). From the overall correlation results obtained, it is found that the addition of the SLE data together with VLE data in the regression of the AIPs slightly increase the accuracy of predictions related to SLE systems while still maintaining the accuracy of predictions of the VLE systems. There are no substantial improvements in the correlation/prediction of SLE data maybe because the amount of VLE data is much bigger compared to the SLE data. Moreover, most of the SLE systems have only one data point. It is difficult for UNIFAC to capture the behavior of such system throughout the entire concentration range, especially during the parameter regression step in order to obtain the best AIPs. Therefore, we can say that the AIPs are more favourable to the VLE systems suggesting that there is only a slight increase of performance when SLE systems are included. Nevertheless, the results obtained for this work are satisfactory and acceptable.

Besides using the Original UNIFAC model [15] as the host model for the development of GC^{Plus} models for mixtures, the Modified (Dortmund) UNIFAC model developed by Gmehling et al. [68] has also been chosen as the host UNIFAC model. The Modified (Dortmund) UNIFAC-CI model was initially developed by Gonzalez [18] for VLE systems involving the C, H, O and N atoms. As presented in Chapter 6, the developed CI-model has been further expanded for VLE systems involving chlorinated and sulfurated systems. The development of this version of the Modified (Dortmund) UNIFAC-CI can be a good starting point for its further development which can make it fairly comparable with the host UNIFAC model which has GIPs regressed not only on VLE but also on activity coefficients at infinite dilution and excess enthalpy data.

Finally, the application of the CI-models using the regressed AIPs are highlighted in Chapter 7 for selected case studies involving the design of a working solution, solubility investigation involving pharmaceuticals systems, phase equilibria of lipid systems and some other special systems including azeotropic ones. For these case studies it is decided only to use the analyzed and developed Original UNIFAC-CI model. From all the work done previously for the Original UNIFAC-CI model as reported in Chapters 3-5, the set of AIPs which were regressed simultaneously against both VLE and SLE have been chosen for these case studies in Chapter 7. In the case studies involving the design of a working solution for hydrogen peroxide production and the solubility investigation of pharmaceutical systems, both involve SLE. The solutes for both case studies cannot be fully described by the UNIFAC groups. Therefore, new groups have been created and their valence connectivity indices (CI) and related interaction parameters have been generated. However, one problem that arises was the highly negative values that were generated for some GIPs which is problematic for the convergence of the SLE calculations. Therefore, some CI generated GIPs have been fine tuned against the experimental data. Finally a master parameter table containing GIPs of the Original UNIFAC [15], CI-generated parameters and the fine tuned parameters have been obtained for each of the case studies. For the lipid systems, the phase equlibria investigation of a multicomponent system of a palm oil deodorization process and the

predictions of VLE systems involving fatty acid ester-alcohols and alcohols-glycerols systems have been studied. The predictions made by the CI-model are compared with the reference Original UNIFAC model and the results obtained illustrate that the Original UNIFAC-CI model can be used to predict lipid systems reliably and with high accuracy. Furthermore, the applicability of the CI-model in filling missing GIPs in the UNIFAC parameter table [15] has been also highlighted for system of phenol-ethyl formate. Besides that, the improvement of the predictions of the CI-model compared with the Original UNIFAC model involving an azeotropic system was also presented.

6.2 Contributions of the PhD Project

From the work presented in Chapters 3-7, the achievements and contributions of this PhD project can be summarized as follows:

1) Analysis of the developed Original UNIFAC-CI model, where it is found that by introducing more structural information to the CHO group through the valence connectivity index (CI), the correlation error involving alkanes-aldehydes system can be reduced.

2) As a continuation of the analysis done for systems involving C, H and O atoms, the Original UNIFAC-CI model has been further reused and significantly expanded by including nitrogenated, chlorinated and sulfurated systems and the involved atom interaction parameters (AIPs) have been regressed.

3) Development of the Original UNIFAC-CI (VLE) model using a quality assessment algorithm, Q_{VLE} (combination of 4 VLE consistency tests) as a weighting factor for each VLE dataset in the objective function for regression of AIPs. The quality factors are useful in identifying anomalous systems which can be problematic in the parameter estimation and can produce parameters which are not accurately representing the systems used for the regression.

4) Development of the Original UNIFAC-CI (VLE/SLE) model for which AIPs are obtained through regression against both VLE and SLE experimental data. The prediction accuracy of SLE systems using the regressed parameters have been slightly increased.

5) Further development of the Modified (Dortmund) UNIFAC-CI involving chlorinated and sulfurated VLE systems.

6) Application of the developed Original UNIFAC-CI (VLE/SLE) model in selected case studies involving the design of a working solution for hydrogen peroxide production and solubility investigation of pharmaceutical systems where new group have been created and their interaction parameters are predicted/fine tuned generating a master parameter table specifically for those case studies. Also, the applicability of the Original UNIFAC-CI model is shown for predicting phase equilibria of lipid systems, filling missing GIPs and improving prediction of azeotropic mixture.

The list of conference presentations and publications with respect to this PhD project are presented in Appendix E.

6.3 Future Work and Recommendations

The GC^{Plus} models for mixtures have been analyzed and further developed successfully in this work. However, the analyzed and developed GC^{Plus} models are not perfect and of course there is room for improvements. Therefore, we would like to give some recommendations and suggestions for future work. The recommendations are as follows:

1) One of the purpose of this work is to use a limited or available amount of data in order to regress the atom interaction parameters (AIPs). However, the reliability and accuracy of the predictions using the regressed parameters will increase if more data involving different types of systems described by different UNIFAC groups are used in the parameter regression step. Currently, using the predicted GIPs from the regressed AIPs for phase equilibria predicitons, reliable predictions are not guaranteed if the UNIFAC groups representing the investigated systems are not included in the list of systems used for the parameter regression. The recommendation is therefore to use systems which can be described by most of the UNIFAC groups that can be found in the UNIFAC parameter table, for the regression of AIPs.

2) For the development of the Original UNIFAC-CI (VLE/SLE) model, more SLE systems need to be included in order to regress the AIPs. This wil possibly improve the performance and accuracy of the model for SLE predictions. Also, it would be useful if a consistency test for SLE systems can be developed in order to make sure that the SLE systems used are consistent and appropriate weight can be used for each system during parameter regression. Such SLE consistency test do not exist and this is a major challenge but it will be highly useful.

3) For the development of the Modified (Dortmund) UNIFAC-CI model, new types of experimental data can be used together with the VLE data in order to have a fair comparison with the reference model developed by Gmehling et al. [64]. Those experimental data include activity coefficients at infinite dilution and excess enthalpy data which can improve the model predictions at lower concentrations and also over a wider temperature range.

4) Use of a method for the evaluation of UNIFAC interaction parameters as presented by Kang et al. [127] and Anderson et al. [141] and modified for the regression of AIPs. In this method, the uncertainties in the experimental data are taken into account through the use of Q_{VLE}, a quality factor and the objective function is minimized using a maximum likelihood method which incorporates the uncertainties of the variables. In addition, a variance-covariance matrix of the regression analysis can be generated which gives an estimate of the uncertainties in the regressed parameters. Using these uncertainties in the model parameters, the uncertainties in the predicted properties can also be estimated. The inclusion of uncertainties in the predicted properties can provide more reliable property predictions.

5) When generating the optimal values for the AIPs, sometimes very high positive and negative values are generated for the GIPs. This issue is not a big problem during

correlation since the regression of the parameters can be controlled using the regularization term included in the objective function eventhough sometimes it cannot be avoided since acceptable correlation errors have been obtained even with high positive/negative values. However, when generating new GIPs for predictions, such high values especially the negative ones can lead to poor predictions. For SLE systems, GIPs with more than -1000 can cause convergence problem for the SLE calculations. Therefore, it should be possible to bound the generated GIPs to rational limits when using AIPs in order to avoid difficulties in predictions.

6) As suggested by Dr. Gerard Krooshof, a Senior Expert Thermodynamics from DSM, the CI-models can also be improved by including the closure relation as reported in [142] which has been applied to UNIQUAC and NRTL. The latter requires that the non-randomness parameter has to be fixed to a constant value for all binaries. The article in [142] shows that the closure relation does not imply that the overall fit is better, but it becomes more robust. The article also shows that without the closure relation the root mean square becomes smaller. However, for an atom based prediction method, it is prefered to have internal consistency (closure relation) even more than accuracy, because outliers are problematic for prediction purposes.

NOMENCLATURE

Symbols

a_{kl}	group interaction parameter between group *k* and *l*
$(A_{mn}^{X\text{-}Y})_n$	coefficient involving CIs and stochiometry information (Equations (2.39)-(2.42))
A_{wk}	van der Waals group surface area
B	virial coefficient
B_{bound}	bound contribution of the virial coefficient
B_{chem}	chemical theory contribution of the virial coefficient
B_{free}	free pairs contribution of the virial coefficient
$B_{metastable}$	metastable bound contribution of the virial coefficient
b_{mn}	UNIFAC group interaction parameter between group *m* and *n*
b_{X-Y}	atom interaction parameter of level 1 for atoms X and Y (for a_{mn})
$\overline{b_{X-Y}}$	atom interaction parameter of level 1 for atoms X and Y (for a_{nm})
b^1_{X-Y}	atom interaction parameter of level 1 for atoms X and Y (for b_{mn})
$\overline{b^1_{X-Y}}$	atom interaction parameter of level 1 for atoms X and Y (for b_{nm})
b^2_{X-Y}	atom interaction parameter of level 1 for atoms X and Y (for c_{mn})
$\overline{b^2_{X-Y}}$	atom interaction parameter of level 1 for atoms X and Y (for c_{nm})
C_i	contributions first order functional group *i* in compound
c_{mn}	UNIFAC group interaction parameter between group *m* and *n*
c_{X-Y}	atom interaction parameter of level 2 for atoms X and Y (for a_{mn})
$\overline{c_{X-Y}}$	atom interaction parameter of level 2 for atoms X and Y (for a_{nm})
c^1_{X-Y}	atom interaction parameter of level 2 for atoms X and Y (for b_{mn})
$\overline{c^1_{X-Y}}$	atom interaction parameter of level 2 for atoms X and Y (for b_{nm})
c^2_{X-Y}	atom interaction parameter of level 2 for atoms X and Y (for c_{mn})
$\overline{c^2_{X-Y}}$	atom interaction parameter of level 2 for atoms X and Y (for c_{nm})
D_j	contributions second order functional group *j* in compound
d_{X-Y}	atom interaction parameter of level 3 for atoms X and Y (for a_{mn})
$\overline{d_{X-Y}}$	atom interaction parameter of level 3 for atoms X and Y (for a_{nm})
d^1_{X-Y}	atom interaction parameter of level 3 for atoms X and Y (for b_{mn})
$\overline{d^1_{X-Y}}$	atom interaction parameter of level 3 for atoms X and Y (for b_{nm})
d^2_{X-Y}	atom interaction parameter of level 3 for atoms X and Y (for c_{mn})
$\overline{d^2_{X-Y}}$	atom interaction parameter of level 3 for atoms X and Y (for c_{nm})
E_k	contributions third order functional group *i* in compound
e_{X-Y}	atom interaction parameter of level 4 for atoms X and Y (for a_{mn})
$\overline{e_{X-Y}}$	atom interaction parameter of level 4 for atoms X and Y (for a_{nm})

e^{1}_{X-Y}	atom interaction parameter of level 4 for atoms X and Y (for b_{mn})
$\overline{e^{1}_{X-Y}}$	atom interaction parameter of level 4 for atoms X and Y (for b_{nm})
e^{2}_{X-Y}	atom interaction parameter of level 4 for atoms X and Y (for c_{mn})
$\overline{e^{2}_{X-Y}}$	atom interaction parameter of level 4 for atoms X and Y (for c_{nm})
f_i^V	vapor fugacity of component *i*
f_i^L	liquid fugacity of component *i*
H^E	entalphy heat of mixing
h_i	number of bonded hydrogen on atom *i*
K	temperature unit of Kelvin
log P	partition coefficient
M_j	number of occurences of second order groups *j* in compound
mg	main groups
N	number of data points
N_H	number of hydrogen atoms attached
n_i	lone-pair electrons on atom *i*
N_i	number of occurences of first order groups *i* in compound
O_k	number of occurences of third order groups *k* in compound
P	properties (in Equation (2.1))
P	pressure
p_i	pi orbital electrons on atom *i*
P_i^{sat}	vapor pressure of component *i*
P_{exp}	experimental pressure
P_{calc}	calculated pressure
q_i	molecular surface area parameter for component *i*
Q_k	surface area parameter for subgroup *k*
Q_{VLE}	VLE data quality factor
R	gas constant
r_i	molecular volume parameter for component *i*
R_k	volume parameter for subgroup *k*
T	temperature
T_m	melting temperature
V_{wk}	van der Waals group volume area
w_{reg}	regularization factor
x_i	liquid mole fraction of component *i*
$x_{i\text{-}exp}$	experimental liquid mole fraction of component *i*
$x_{i\text{-}calc}$	calculated liquid mole fraction of component *i*
y_i	vapor mole fraction of component *i*
Z^v	number of valence electron in the atom
Z	atomic number
ΔH_{fus}	heat of fusion

Greek Symbols

β^{v}	first order valence bond index
γ	activity coefficient
γ^{C}	combinatorial contribution of the activity coefficient
γ^{R}	residual contribution of the activity coefficient
γ^{R2}	second order residual contribution of the activity coefficient
γ^{∞}	activity coefficient at infinite dilution
$\gamma_{i-\exp}$	experimental activity coefficient
γ_{i-calc}	calculated activity coefficient
δ_i^{v}	number of valence electron of atom vertex i
ε	effective nonpolar potential parameter
ε^{v}	second order valence bond index
η	association parameter
μ	molecular dipole moment
σ_i	sigma orbital electrons on atom *i*
ϕ_i	fugacity coefficient of component *i*
$\upsilon_k^{(i)}$	number of groups *k* in molecule *i*
${}^{v}\chi^{0}$	zeroth order valence connetivity index
${}^{v}\chi^{1}$	first order valence connetivity index
${}^{v}\chi^{2}$	second order valence connetivity index

Acronyms

AARD	Average Absolute Relative Deviation
AARD1	Average Absolute Relative Deviation (for pressure)
AARD2	Average Absolute Relative Deviation (for liquid mole fraction)
AIPs	Atom Interaction Parameters
API	Active Pharmaceutical Ingredients
ASOG	Analytical Solution of Groups Model
CAMD	Computer Aided Molecular Design
CAPE	Computer Aided Process Engineering
CAFD	Computer Aided Flowsheet Design
CI	Valence Connectivity Index
COSMO-RS	Conductor-Like Screening Model for Real Solvents
CPA	Cubic Plus Association
DAG	Dialcylglyceride
EFCE	European Federation of Chemical Engineering
FFA	Free Fatty Acid

FH	Flory-Huggins
FV	Free Volume
GC	Group-Contribution
GC^{Plus}	Group-Contribution^{Plus}
GIPs	Group Interaction Parameters
LLE	Liquid-liquid Equilibrium
OF	Objective Function
POY	Poynting factor
PSRK	Predictive Soave-Redlich-Kwong
QSAR	Quantitative Structure-Activity Relationship
QSPR	Quantitative Structure-Property Relationship
MAG	Monoalcylgliceride
MOSCED	Modified Separation of Cohesive Energy Density Model
NRTL	Non-Random Two Liquid Model
SAFT	Statistical Associating Fluid Theory
SLE	Solid Liquid Equilibrium
TAG	Trialcylglyceride
UNIFAC	Universal Functional Group Activity Coefficient Model
UNIQUAC	Universal Quasi-Chemical Model
VLE	Vapor Liquid Equilibrium

REFERENCES

[1] Gani R., and O'Connell J. P., "Properties and CAPE: From Present Uses to Future Challenges", *Comput. Chem. Eng.*, **25**, 3-14 (2001).

[2] Kontogeorgis G. M., and Gani R., "Computer Aided Property Estimation for Process and Product Design", *Computer-Aided Chemical Engineering*, Amsterdam, Vol. 19 (2004).

[3] Gani R., and Pistikopoulos E. N., "Property Modelling and Simulation for Product and Process Design", *Fluid Phase Equilibr.* **194-197**, 43-59 (2002).

[4] Joback K.G., and Reid R.C., "Estimation of Pure-Component Properties from Group-Contributions", *Chem. Eng. Commun.* **57**, 233 - 243 (1987).

[5] Lydersen A. L., "Estimation of Critical Properties of Organic Compounds", College Engineering University Wisconsin, Engineering Experimental Station Report 3, Madison, WI, April 1955

[6] Ambrose D., Correlation and Estimation of Vapor-Liquid Critical Properties. I. Critical Temperatures of Organic Compounds, National Physical Laboratory, Teddington, UK, NPL Report Chem., 92, September 1978

[7] Gani R., and Constantinou L., "Molecular Structure Based Estimation of Properties for Process Design", *Fluid Phase Equilibr.* **116**, 75 – 86 (1996).

[8] Marrero J., and Gani R., "Group-Contribution Based Estimation of Pure Component Properties", *Fluid Phase Equilibr.,* **183-184**, 183-208 (2001).

[9] Hukkerikar A. S., Sarup B., Kate A. T., Abildskov J., Sin G., and Gani R., "Group-Contribution^{+} (GC^{+}) Based Estimation of Properties of Pure Components: Improved Property Estimation and Uncertainty Analysis", *Fluid Phase Equilibr.,* **321**, 25-43 (2012).

[10] Derr E. L., and Deal C. H., "Analytical Solutions of Groups: Correlation of Activity Coefficients Through Structural Group Parameters," *1. Chem. E. Symp. Ser. No. 32* (Instn. Chem. Engrs., London) **3**, 40 (1969).

[11] Ronc M., and Ratcliff G. A., "Prediction of Excess Free Energies of Liquid Mixtures by an Analytic Group Solution Model," *Can.* J. *Chem. Eng.,* **49**, 875 (1971).

[12] Fredenslund Aa., Jones R. L., and Prausnitz J. M., "Group Contribution Estimation of Activity Coefficients in Nonideal Liquid Mixtures", *AIChE Journal*, **21**, 1086 (1975)

[13] Weidlich U., and Gmehling J., "A Modified UNIFAC Model. 1- Prediction of VLE, h^{E}, and γ^{∞}", *Ind. Eng. Chem. Res.*, **26**, 1372 (1987).

[14] Holderbaum T., and Gmehling J., "PSRK: A Group Contribution Equation of State Based on UNIFAC", *Fluid Phase Equilibria*, **70**, 251 (1991).

[15] Hansen K. H., Rasmussen P., Fredenslund Aa., Schiller M., and Gmehling J., "Vapor-Liquid Equilibria by UNIFAC Group Contribution. 5. Revision and Extension", *Ind. Eng. Chem. Res.,* **30**, 2355 (1991).

[16] González H E, Abildskov J, Gani R, Rousseaux P, and Le Bert B. "A Method for Prediction of UNIFAC Group Interaction Parameters*", AIChE J.* **53**, 1620 (2007).

[17] González H. E., Abildskov J., and Gani R., "Computer-Aided Framework for Pure Component Properties and Phase Equilibria Prediction for Organic Systems", *Fluid Phase Equilibria*, **261**, 199 (2007).

[18] González H. E., "Develoment of Group-Contribution^{Plus} Models for Properties of Organic Chemical Systems", Ph.D Dissertation, Technical University of Denmark, (2009).

[19] SciVerse Scopus (http://www.scopus.com)

[20] Hendriks E., Kontogeorgis G. M., Dohrn R., de Hemptinne J-C., Economou I. G., Zilnik L. F., and Vesovic V., "Industrial Requirements for Thermodynamics and Transport Properties", *Ind. Eng. Chem. Res.,* **49**, 11131-11141 (2010).

[21] Gani R., Harper P., and Hostrup M., "Automatic Creation of Missing Groups through Connectivity Indices for Pure-Compound Property Prediction", *Ind. Eng. Chem. Res*., **44**, 7262 (2005).

[22] Satyanarayana K. C., Gani R., and Abildskov J., "Polymer Property Modeling using Grid Technology for Design of Structured Products", *Fluid Phase Equilibria*, **261**, 58 (2007).

[23] Satyanarayana K. C., Gani R., and Abildskov J., "Computer-Aided Polymer Design using Group Contribution Plus Property Models", *Comput. Chem. Eng*., **33**, 1004-1013 (2009).

[24] Satyanarayana K. C., Abildskov J., Gani R., Tsolou G., and Mavrantzas V. G., "Computer Aided Polymer Design using Multi-Scale Modelling", *Braz. J. Chem. Eng.,* **27**, 369-380 (2010).

[25] Modaressi H., Conte E., Abildskov J., Gani R., and Crafts P., "Model-Based Calculation of Solid Solubility for Solvent Selections - A Review", *Ind. Eng. Chem. Res,. **47,*** 5234–5242 (2008).

[26] Conte E., Martinho A., Matos H., and Gani R., "Combined Group Contribution and Atom Connectivity Index based Methods for Estimation of Surface Tension and Viscosity", *Ind. Eng. Chem. Res*., **47**, 5751 (2008).

[27] Poling B. E., Prausnitz J. M., and O´Connell J. P., *The Properties of Gases and Liquids*, McGraw-Hill, New York (2001).

[28] Prausnitz J. M., and Tavares F. W., "Thermodynamics of Fluid-Phase Equilibria for Standard Chemical Engineering Operations", *AIChE J*., **50 (4)**, 739-761 (2004).

[29] Sandler S. I., and Castier M., “Computational Quantum Mechanics: An Underutilized Tool in Thermodynamics”, *Pure Appl. Chem.,* **79**, 1345-1356 (2007).

[30] Sandler S. I., “Quantum Mechanics: A New Tool for Engineering Thermodynamics”, *Fluid Phase Equilibr.,* **210**, 147-160 (2003).

[31] Wolbach J. P., and Sandler S. I., “Using Molecular Orbital Calculations to Describe the Phase Behavior of Cross-Associating Mixtures”, *Ind. Eng. Chem. Res.,* **37***,* 2917-2928 (1998).

[32] Klamt A., “Conductor-Like Screening Model for Real Solvents: A New Approach to the Quantitative Calculation of Solvation Phenomena” *J. Phys. Chem.,* **99***,* 2224-2235 (1995).

[33] Klamt A., and Eckert F., “COSMO-RS: A Novel and Efficient Method for the A *Priori* Prediction of Thermophysical Data of Liquids”, *Fluid Phase Equilibr.*, **172**, 43 (2000).

[34] Constantinou L., and Gani R., “New Group Contribution Method for Estimating Properties of Pure Compunds”, *AIChE J.,* **40**, 1697-1710 (1994).

[35] Kang J. W., Abildskov J., and Gani R., ”Estimation of Mixture Properties from First- and Second-Order Group Contributions with the UNIFAC Model”, *Ind. Eng. Chem. Res.,* **41**, 3260-3273 (2002).

[36] Ravindranath D., Neely B. J., Robinson Jr. R. L., and Gasem K. A. M., “QSPR Generalization of Activity Coefficient Models for Predicting Vapor–Liquid Equilibrium Behavior”, *Fluid Phase Equilibr.*, **257**, 53 (2007).

[37] Bunz A. P., Braun B., and Janowsky R., “Quantitative Structure–Property Relationships and Neural Networks: Correlation and Prediction of Physical Properties of Pure Components and Mixtures from Molecular Structure”, *Fluid Phase Equilibr.*, **367**, 158-160 (1999).

[38] van Krevelen D. W., “Properties of Polymers. Correlation with Chemical Structure”, *Elsevier*, Amsterdam (1990).

[39] Horvath A. L., “Molecular Design”, *Elsevier*, Amsterdam (1992).

[40] Reid R. C., Prausnitz J. M., and Poling B. E., “The Properties of Gases and Liquids”, 4th Edition, *McGraw-Hill Book Co.,* (1987).

[41] Constantinou L., and Gani R., “New Group Contribution Method for Estimating Properties of Pure Compunds”, *AIChE J.,* **40**, 1697-1710 (1994).

[42] Klincewicz K. M., and Reid R. C., “Estimation of Critical Properties with Group Contribution Methods”, *AIChE J.,* **30**, 137 (1984).

[43] Benson S. W., and Buss J., “Additivity Rules for the Estimation of Molecular Properties. Thermodynamic Properties”, *J. Chem. Phys.*, **9**, 3 (1958).

[44] Franklin J. L., “Prediction of Heat and Free Energies of Organic Compounds”,*Ind. Eng. Chem.*, **41**, 5 (1949).

[45] Verma K. K., and Doraiswamy, L. K., “Estimation of Heats of Formation of Organic Compounds”, *Ind. Eng. Chem. Fundam.* **4**, 389 (1965).

[46] Langmuir I., *The Distribution and Orientation of Molecules*, Third Colloid Symposium Monograph, The Chemical Catalog Company Inc., New York (1925).

[47] Redlich O., Derr E. L., and Pierotti G., “Group Interaction I. A Model for Interactions in Solution”, *J. Am. Chem. Soc.*, **81**, 2283 (1959).

[48] Derr E. L., and Papadopoulos M., “Group Interaction II. A test of the Group Model on Binary Solutions of Hydrocarbons”, *J. Am. Chem. Soc.*, **81**, 2285 (1959).

[49] Wilson G. M., and Deal C. H., “Activity Coefficients and Molecular Structure”, *Ind. Chem. Fund.,* **1**, 20 (1962).

[50] Kojima K., and Tochigi K., “Prediction of Vapor-Liquid Equilibria by the ASOG Method”, *Kodansha-Elsevier*, Tokyo (1979).

[51] Abrams D. S., and Prausnitz J. M., “Statistical Thermodynamics of Liquid Mixtures. A New Expression for the Excess Gibbs Energy of Partly and Completely Miscible Systems”, *AIChE Journal*, **21**, 116 (1975).

[52] Gmehling J., Onken U., and Arlt W., “Vapor-Liquid Equilibrium Data Collection”, *Chemistry Data Series*, Vol. I, Parts 1-8, DECHEMA, Frankfurt Main (1974-1990).

[53] Smith J.M., Van Ness H. C., and Abbott M. M., “Introduction to Chemical Engineering Thermodymanic”, 6th Edition, *McGraw-Hill International Edition* (2001).

[54] Bondi A., *Physical Properties of Molecular Crystals, Liquids and Glasses*, John Wiley & Sons, New York (1968).

[55] Kontogeorgios, G. M., and Folas, G., “Thermodynamic Models for Industrial Applications: From Classical and Advanced Mixing Rules to Association Theories”, *John Wiley and Sons Ltd.* (2010).

[56] Fredenslund Aa., “UNIFAC and Related Group-Contribution Models for Phase Equilibria”, *Fluid Phase Equilibr.,* **52**, 135-150 (1989)

[57] Fredenslund Aa., and Rasmussen P., "From UNIFAC to SUPERFAC – and Back?", *Fluid Phase Equilibr.,* **24**, 115-150 (1985).

[58] Skjold-Jørgensen S., Kolbe B., Gmehling, J., and Rasmussen P., "Vapor-Liquid Equilibria by UNIFAC Group Contribution. Revision and Extension", *Ind. Eng. Chem. Process Des. Dev.,* **18**, 714-722 (1979).

[59] Gmehling J., Rasmussen P., and Fredenslund Aa., "Vapor-Liquid Equilibria by UNIFAC Group Contribution. Revision and Extension. 2", *Ind. Eng. Chem. Process Des. Dev.,* **21**, 118-127 (1982).

[60] Macedo E. A., Weidlich U., Gmehling J., and Rasmussen P., "Vapor-Liquid Equilibria by UNIFAC Group Contribution. Revision and Extension. 3", *Ind. Eng. Chem. Process Des. Dev.,* **22**, 676-678 (1983).

[61] Tiegs D., Gmehling J., Rasmussen P., and Fredenslund Aa., "Vapor-Liquid Equilibria by UNIFAC Group Contribution. Revision and Extension. 4", *Ind. Eng. Chem. Res.,* **26**, 159-161 (1987).

[62] Hansen H. K., Rasmussen P., Fredenslund Aa., Schiller M. and Gmehling J., "Vapor-Liquid Equilibria by UNIFAC Group Contribution. Revision and Extension. 5", *Ind. Eng. Chem. Res.,* **30**, 2355-2358 (1991).

[63] Wittig R., Lohmann J., and Gmehling J., "Vapor-Liquid Equilibria by UNIFAC Group Contribution. Revision and Extension. 6", *Ind. Eng. Chem. Res.,* **42**, 183-188 (2003).

[64] UNIFAC Consortium, http://unifac.ddbst.de/

[65] Magnussen T., Rasmussen P., and Fredenslund Aa., "UNIFAC Parameter Table for Prediction of Liquid-Liquid Equilibria", *Ind. Eng. Chem. Process Des. Dev.*, **20**, 331-339 (1981).

[66] Hansen H. K., Coto B., and Kulhmann B., "UNIFAC with Linearly Temperature-Dependent Group-Interaction Parameters", IVC-SEP Internal Report SEP 9212, Institut for Kemiteknik, Technical University of Denmark (1992).

[67] Larsen B. L., Rasmussen P., and Fredenslund Aa., "A Modified UNIFAC Group-Contribution Model for Prediction of Phase Equilibria and Heats of Mixing", *Ind. Eng. Chem. Res.*, **26**, 2274-2286 (1987).

[68] Gmehling J., Li J., and Schiller M., "A Modified UNIFAC Model. 2. Present Parameter Matrix and Results for Different Thermodymanic Properties", *Ind. Eng. Chem. Res.,* **32**, 178-193 (1993).

[69] Gmehling J., Lohmann J., Jakob A., Li, J., and Joh R., "A Modified UNIFAC (Dortmund) Model. 3. Revision and Extension", *Ind. Eng. Chem. Res, **37**,* 4876-4882 (1998).

[70] Gmehling J., Wittig R., Lohmann J., and Joh R., "A Modified UNIFAC (Dortmund) Model. 4. Revision and Extension", *Ind. Eng. Chem. Res, **41**,* 1678-1688 (2002).

[71] Jakob A., Grensemann H., Lohmann J., and Gmehling J., "Further Development of Modified UNIFAC (Dortmund): Revision and Extension 5", *Ind. Eng. Chem. Res.,* **45**, 7924-7933 (2006).

[72] Abildskov J., Gani R., Rasmussen P., O'Connel J. P., "Beyond Basic UNIFAC", *Fluid Phase Equilibr.,* **158-160**, 349-356 (1999).

[73] Abildskov J., Constantinou L., Gani R., "Towards the Development of a Second-Order Approximation in Activity Coefficient Models Based on Group Contributions", *Fluid Phase Equilibr.,* **118**, 1 (1996).

[74] Hooper H. H., Michel S., and Prausnitz J. M., "Correlation of Liquid-Liquid Equilibria for Some Water-Organic Liquid Systems in the Region 20-250^{0}C", *Ind. Eng. Chem. Res.,* **27**, 2182-2187 (1988).

[75] Chen F., Holten-Andersen J., and Tyle H., "New Development of the UNIFAC Model for Enviromental Applications", *Chemosphere*, **26**, 1325-1354 (1993).

[76] Oishi T., and Praunitz J. M., "Estimation of Solvent Activities in Polymer Solutions Using a Group-Contribution Method, *Ind. Eng. Chem. Process Des. Dev*., **17**, 333-339 (1978).

[77] Elbro H. S., Fredenslund Aa., and Rasmussen P.,"A New Simple Equation for the Prediction of Solvent Activities in Polymer Solutions", *Macromolecules*, **23**, 4707-4714 (1990).

[78] Kontogeorgis G. M., Fredenslund Aa., and Tassios D. P., "Simple Activity Coefficient Model for the Prediction of Solvent Activities in Polymer Solutions", *Ind. Eng. Chem. Res.,* **32**, 362-372 (1993).

[79] Kouskoumvekaki I. A., Michelsen M. L., Kontogeorgis G. M., "An Improved Entropic Expression for Polymer Solutions" *Fluid Phase Equilibr.,* **202**, 325–335 (2002).

[80] Liu Q. L., and Cheng Z. F., "A Modified UNIFAC Model for the Prediction of Phase Equilibrium for Polymer Solutions", *J. Polym. Sci.,* **43**, 2541-2547 (2005).

[81] Flory P. J., "Fifteenth Spiers Memorial Lecture: Thermodynamics of Polymer Solutions" *Discuss. Farday Soc.,* **49**, 7 (1970).

[82] Kikic I., Fermeglia M., and Rasmussen P., "UNIFAC Prediction of Vapor-Liquid Equilibrium in Mixed Solvent-Salt Systems", Chem. Eng. Sci., **46**, 2775-2780 (1991).

[83] Aznar M., and Telles A. S., "Prediction of Electrolyte Vapor Liquid Equilibrium by UNIFAC Dortmund", *Braz. J. Chem. Eng,.* **18**, 127-137 (2001).

[84] Fu Y. H., Orbey H., and Sandler S. I., "Prediction of Vapor-Liquid Equilibria of Associating Mixtures with UNIFAC Models That Include Association", *Ind. Eng. Chem. Res.,* **35**, 4656-4666 (1996).

[85] Lei Z., Zhang J., Li Q., and Chen B., "UNIFAC Model for Ionic Liquids", *Ind. Eng. Chem. Res.,* **48**, 2697-2704 (2009).

[86] Diedrichs A., and Gmehling J., "Solubility Calculation of Active Pharmceutical Ingredients in Alkanes, Alcohols, Water and their Mixtures using Various Activity Coefficient Models", *Ind. Eng. Chem. Res.,* **50**, 1757-1769 (2011).

[87] Gmehling J., Anderson T. F., and Prausnitz J. M., "Solid-Liquid Equilibria using UNIFAC", *Ind. Eng. Chem. Fund.,* **17**, 269-273 (1978).

[88] Jakob A., Joh R., Rose C., and Gmehling J., "Solid-Liquid Equilibria in Binary Mixtures of Organic Compounds", *Fluid Phase Equilibr.,* **113**, 117-126 (1995).

[89] Jensen T., Fredenslund Aa., Rasmussen P., "Pure Component Vapor Pressures Using UNIFAC Group Contribution", *Ind. Eng. Chem. Fund.,* **20,** 239–246 (1981).

[90] Lo H. S., and Paulatis M. E., "Estimation of Solvent Effects on Chemical Reaction Rates using UNIFAC Group Contribution", *AIChE J.,* **27**, 842 (1981).

[91] Cheng H. Y., Kontogeorgis G. M., and Stenby E. H., "Prediction of Micelle Formation for Aqueous Polyoxyethylene Alcohol Solutions with the UNIFAC Model", *Ind. Eng. Chem. Res.,* **41***,* 892-898 (2002).

[92] Tegtmeier U., and Misselhorn K., "Use of the UNIFAC Method for the Calculation of Extraction Parameters", *Chem. Ing. Techn.,* **53**, 542 (1981).

[93] Gmehling J., Rasmussen P., "Flash Points of Flammable Liquids using UNIFAC", *Ind. Eng Chem. Fund.,* **21,**186–188 (1982).

[94] Sander B., Skjold-Jørgensen S., Rasmussen P., "Gas Solubility Calculations. I. UNIFAC", *Fluid Phase Equilibr.,* **11**, 105–126 (1983).

[95] Brignole E. A., Bottini S., and Gani R., "A Strategy for the Design and Selection of Solvents for Separation Processes", *Fluid Phase Equilibr.,* **29**, 125-132 (1986).

[96] Dang D., Tasslos D. P., "Prediction of Enthalpies of Mixing with a UNIFAC Model", *Ind. Eng. Chem. Proc. Des. Develop.,* **25**, 22–31 (1986).

[97] Cao W., Knudsen K., Fredenslund Aa., and Rasmussen P., "Group-Contribution Viscosity Predictions of Liquid Mixtures using UNIFAC-VLE Parameters, *Ind. Eng. Chem. Res.,* **32**, 2088-2092 (1993).

[98] Li A., Doucette W. J., and Andren A. W., "Estimation of Aqueous Solubility, Octanol/Water Partition Coefficient, and Henry's Law Constant for Polychlorinated Biphenyls using UNIFAC", *Chemosphere,* **29**, 657-669 (1994).

[99] Wienke G., anf Gmehling J., "Prediction of Octanol-Water Partition Coefficients, Henry Coeffcients and Water Solubilities using UNIFAC", *Toxicol. Environ. Chem.,* **65**, 57-86 (1998).

[100] Derawi S. O., Kontogeorgis G. M., and Stenby E. H., "Application of Group Contribution Models to the Calculation of the Octanol-Water Partition Coefficient", *Ind. Eng. Chem. Res.,* **40***,* 434-443 (2001).

[101] Gupta R. B., and Heidemann R. H., "Solubility Models for Amino Acids and Antibiotics", *AIChE J.,* **37**, 333-341 (1990).

[102] Li S., and Paik J. S., "Flavor Sorption Estimation by UNIFAC Group Contribution Model", *Trans. ASAE,* **39**, 1013-1017 (1996).

[103] Hatzioannidis I., Voutsas E. C., Lois E., and Tassios D. P., "Measurement and Prediction of Reid Vapor Pressure of Gasoline in the Presence of Additives", *J. Chem. Eng. Data*, **43***,* 386-392 (1998).

[104] Paksoy H. O., Ornektekin S., Balcl B., Demirel Y., "The Performance of UNIFAC and Related Group Contribution Models Part I. Prediction of Infinite Dilution Activity Coefficients", *Thermochim. Acta,* **287**, 235-249 (1996).

[105] Ornektekin S., Paksoy H. O., Demirel Y., "The Performance of UNIFAC and Related Group Contribution Models Part II. Prediction of Henry's Law Constants", *Thermochim. Acta,* **287**, 251-259 (1996)

[106] Voutsas E. C., and Tassios D. P., "Prediction of Infinite-Dilution Activity Coefficients in Binary Mixtures with UNIFAC. A Critical Evaluation" *Ind. Eng. Chem. Res.,* **35***,* 1438-1445 (1996).

[107] Fredenslund Aa., Gmehling J., Michelsen M. L., Rasmussen P., and Prausnitz P. M., "Computerized Design of Multicomponent Distillation Columns using the UNIFAC Group Contribution Method for Calculation of Activity Coefficients", *Ind. Eng. Chem., Process Des. Dev.,* **16**, (1977).

[108] Todeschini R., and Consonni V., *Molecular Descriptors for Chemoinformatics*, Wiley-VCH (2009).

[109] Randíc M., "The Connectivity Index 25 Years After", *J. Mol. Graph. Model.,* **20**, 19-35 (2001).

[110] Wiener H., "Structural Determination of Paraffin Boiling Points", *J. Am. Chem. Soc*., **69**, 17-20 (1947).

[111] Hosoya H., "Topological Index: A Newly Proposed Quantity Characterizing the Topological Nature of Structural Isomers of Saturated Hydrocarbons", *B. Chem. Soc. Jpn*., **44** 2332-2339 (1971).

[112] Kier L. B., Murray W. J., Randíc M., and Hall L. H., "Molecular Connectivity I. Relationship to Nonspecific Local Anesthesia", *J. Pharm. Sci*., **64**, 1971-1974 (1975).

[113] Randíc M., "On the Characterization of Molecular Branching, *J. Am. Chem. Soc*., **97**, 6609-6615 (1975).

[114] Kier L. B., Murray W. J., Randíc M., and Hall L. H., "Molecular Connectivity V. Connectivity Series Applied to Density, *J. Pharm. Sci.,* **65**, 1226-1230 (1975).

[115] Kier L. B., and Hall L. H., "Molecular Connectivity VII. Specific treatment of Heteroatoms, *J. Pharm. Sci*., **65**, 1806-1809 (1975).

[116] Sanderson R. T., "Electronegativities in Inorganic Chemistry II", *J. Chem. Educ*., **31**, (1954).

[117] Sanderson R. T., *Chemical Bonds and Bond Energy*, Academic, New York, N.Y. (1976).

[118] Sanderson R. T., "Principles of Electronegativity. II. Applications", *J. Chem. Educ*., **65**, 227 (1988).

[119] Kier L. B., and Hall L. H., "Derivation and Significance of Valence Molecular Connectivity", *J. Pharm. Sci*., **70**, 583 (1981).

[120] Hayden J. G., and O' Connell J. P., "A Generalized Method for Predicting Second Virial Coefficients", *Ind. Eng. Chem. Process Des. Dev.,* **14**, 209-216 (1975).

[121] Nielsen T.L., Abildskov J., Harper P.M., Papaeconomou I., Gani R., "The CAPEC Database". *J. Chem. Eng. Data*, **46**, 1041 (2001).

[122] Marrero J., and Abildskov J., "Solubility and Related Properties of Large Comple Chemicals", Vol. XV, Part 1: Organic Soultes Ranging from C4 to C40, DECHEMA Chemistry Data Series (2003).

[123] Abildskov J., "Solubility and Related Properties of Large Comple Chemicals", Vol. XV, Part 2: Organic Soultes Ranging from C2 to C41, DECHEMA Chemistry Data Series (2005).

[124] Kemeny S., Manczinger J., Skjold-Jørgensen S., "Reductions of Thermodynamic Data by Means of the Multiresponse Maximum Likelihood Principle*". AIChE J*., **28**, 20 (1982).

[125] Fletcher R., "A Modified Marquardt Subroutine for Non-Linear Least Squares", Harwell Report, AERE R.6799, (1971).

[126] Mustaffa A. A., Kontogeorgis G. M., and Gani R., "Analysis and Application of GC^{Plus} Models for Property Prediction of Organic Chemical Systems", *Fluid Phase Equilibr.,* **302**, 274-283 (2011).

[127] Kang J. W., Diky V., Chirico R. D., Magee J. W., Muzny C. D., Abdulgatov I., Kazakov A. F., and Frenkel M., "Quality Assessment Algorithm for Vapor-Liquid Equilibrium Data", *J. Chem. Eng. Data.,* **55**, 3631-3640 (2010).

[128] http://www.capec.kt.dtu.dk/Software/ICAS-and-its-Tools/

[129] Mohs A., Jakob A., and Gmehling J., "Analysis of a Concept for Predicting Missing Group Interaction Parameters of the UNIFAC Model using Connectivity Indices", *AIChE J.,* **55**, 1614-1625 (2009).

[130] Gani R., and González H. E., "Letter to the Editor", *AIChE J.,* **55**, 1626-1627 (2009).

[131] Van Ness H. C., "Thermodynamics in the Treatment of (Vapor + Liquid) Equilibria", *J. Chem. Thermodynamics*, **27**, 113 (1995).

[132] http://nzic.org.nz/ChemProcesses/production/1E.pdf

[133] http://www.capec.kt.dtu.dk/Software/ICAS-and-its-Tools/

[134] Diaz-Tovar C. A., Gani R., and Sarup B., "Lipid Technology: Property Prediction and Process Design/Analysis in the Edible Oil and Biodiesel Industries", *Fluid Phase Equilibr.*, **302**, 284-293 (2011).

[135] Mustaffa A. A., Diaz-Tovar C. A., Hukkerikar A., Quaglia A., Sin G., Kontogeorgis G., Sarup B., and Gani R., "Building a Multilevel Modeling Network for Lipid Processing Systems" *4th International Conference on Modeling, Simulation and Applied Optimization (ICMSAO),* 19-21 April 2011, Kuala Lumpur, Malaysia

[136] Ceriani, R., Meirelles, A. J. A., "Predicting Vapor-Liquid Equilibria of Fatty Systems", *Fluid Phase Equilibr.*, **215**, 227-236 (2004).

[137] Ceriani, R., Meirelles, A.J., Gani, R., "Simulation of Thin-Film Deodorizers in Palm Oil Refining", *J. Food Process Eng.*, **33**, 208-225 (2010).

[138] Gavin, A. M., "Edible Oil Deodorization", *J. Am. Oil Chem. Soc.*, **55**, 783-791 (1978).

[139] Oliveira M. B., Miguel S. I., Queimada A. J., and Coutinho J. A. P., “Phase Equilibria of Ester + Alcohol Systems and Their Description with the Cubic-Plus-Association Equation of State, *Ind. Eng. Chem. Res.,* **49***,* 3452–3458 (2010).

[140] Oliveira M. B., Teles A. R. R., Queimada A. J., Coutinho J. A. P., “Phase Equilibria of Glycerol Containing Systems and their Description with the Cubic-Plus-Association (CPA) Equation of State” *Fluid Phase Equilibr.,* **280**, 22–29 (2009).

[141] Anderson T. F., Abrams D. S., and Gren E. A., “Evaluation of Parameters for Nonlinear Thermodynamic Models”, *AIChE J.*, **24**, 20-29 (1978).

[142] Sahoo R. K., Banerjee T., Ahmad S. A., and Khanna A., “Improved Binary Parameters using GA for Multi-Component Aromatic Extraction: NRTL Model Without and With Closure Equations”, *Fluid Phase Equilibr.,* **239**, 107–119 (2006).

APPENDIX A

TABLES OF ATOM STOICHIOMETRY AND VALENCE CONNECTIVITY INDICES FOR THE UNIFAC GROUPS

This appendix lists the atom stoichiometry and the valence connectivity index (CI) calculated for each UNIFAC group for the Original UNIFAC-CI and Modified (Dortmund) UNIFAC-CI models respectively in Tables A.1 and A.2. This information can be used in Equations (2.39) to (2.42) for Original UNIFAC-CI and Equations (C.1)-(C.2) for Modified (Dortmund) UNIFAC-CI together with the regressed atom interaction parameters (AIPs) presented in Appendix B.

Table A.1: Atom Stoichiometry and Connectivity Indices (CI) Values for the Original UNIFAC-CI Groups (nX refers to the number of atom X and $^{0}\chi^{v}$, $^{1}\chi^{v}$ and $^{2}\chi^{v}$ refers to valence connectivity index of zeroth, first and second order respectively)

Index	Group	nC	nO	nN	nCl	nS	$^{0}\chi^{v}$	$^{1}\chi^{v}$	$^{2}\chi^{v}$
1	CH_2	1	0	0	0	0	0.707107	0.000000	0.000000
2	C=C	2	0	0	0	0	1.000000	0.250000	0.000000
3	ACH	1	0	0	0	0	0.577350	0.666667	0.192450
4	$ACCH_2$	2	0	0	0	0	1.207107	0.353553	0.000000
5	OH	0	1	0	0	0	0.447214	0.000000	0.000000
6	CH_3OH	1	1	0	0	0	1.447214	0.447214	0.000000
7	H_2O	0	1	0	0	0	0.500000	0.000000	0.000000
8	ACOH	1	1	0	0	0	0.947214	0.223607	0.000000
9	CH_2CO	2	1	0	0	0	1.615355	0.557678	0.144338
10	CHO	1	1	0	0	0	0.985599	0.235702	0.000000
11	CCOO	2	2	0	0	0	1.816497	0.658248	0.287457
12	HCOO	1	2	0	0	0	1.393847	0.471405	0.096225
13	CH_2O	1	1	0	0	0	1.115355	0.288675	0.000000
14	CNH_2	1	0	1	0	0	1.577350	0.577350	0.000000
15	CNH	1	0	1	0	0	1.500000	0.500000	0.000000
16	$(C)_3N$	1	0	1	0	0	1.447214	0.447214	0.000000
17	$ACNH_2$	1	0	1	0	0	1.077350	0.288675	0.000000
18	PYRIDINE	5	0	1	0	0	1.982894	0.655263	0.218421

19	CCN	2	0	1	0	0	1.447214	0.473607	0.111803
20	COOH	1	2	0	0	0	1.355462	0.427731	0.091287
21	CCl	1	0	0	1	0	2.133893	1.133893	0.000000
22	CCl_2	1	0	0	2	0	2.974894	2.087498	0.909137
23	CCl_3	1	0	0	3	0	3.979031	1.963961	2.226922
24	CCl_4	1	0	0	4	0	5.035574	2.267787	3.857143
25	ACCl	1	0	0	1	0	1.633893	0.566947	0.000000
26	CNO_2	1	2	1	0	0	2.224745	0.741582	0.401375
27	$ACNO_2$	1	2	1	0	0	1.724745	0.537457	0.234708
28	CS_2	1	0	0	0	2	1.316497	0.408248	0.083333
29	CH_3SH	1	0	0	0	1	1.447214	0.447214	0.000000
30	FURFURAL	5	2	0	0	0	3.467649	1.749057	0.943474
31	DOH	2	2	0	0	0	2.308641	1.132456	0.447214
32	N/A	0	0	0	0	0	0.000000	0.000000	0.000000
33	N/A	0	0	0	0	0	0.000000	0.000000	0.000000
34	C=-C	2	0	0	0	0	1.000000	0.250000	0.000000
35	DMSO	2	1	0	0	1	2.816497	0.983163	0.741582
36	ACRY	3	0	1	0	0	2.231671	0.920530	0.333224
37	N/A	0	0	0	0	0	0.000000	0.000000	0.000000
38	N/A	0	0	0	0	0	0.000000	0.000000	0.000000
39	DMF	3	1	1	0	0	3.432812	1.388328	1.069021
40	N/A	0	0	0	0	0	0.000000	0.000000	0.000000
41	COO	1	2	0	0	0	1.524564	0.546874	0.129099
42	N/A	0	0	0	0	0	0.000000	0.000000	0.000000
43	N/A	0	0	0	0	0	0.000000	0.000000	0.000000
44	NMP	5	1	1	0	0	2.799734	0.754668	0.310897
45	N/A	0	0	0	0	0	0.000000	0.000000	0.000000
46	CON	1	1	1	0	0	1.432812	0.493901	0.105409
47	OCCOH	2	2	0	0	1	2.269675	1.104903	0.427731
48	CH_2S	1	0	0	0	1	1.408248	0.408248	0.000000
49	MORPHOLINE	4	1	1	0	0	1.682051	0.471132	0.131838
50	THIOPHENE	4	0	0	0	1	1.517649	0.457225	0.136880

Note: 1) For group CHO, a second order CI with a value of 0.1179 were added for the calculation of CH_2-CHO group interaction for the UNIFAC-CI (VLE) model. For details, please refer Chapter 3.

2) The group denoted as N/A refers to UNIFAC groups which contain atoms other than C, H, O, N, Cl and S and are not considered in this work.

Table A.2: Stoichiometric and Connectivity Indices (CI) Values for the Modified (Dortmund) UNIFAC-CI Groups (nX refers to the number of atom X and $^{0}\chi^{v}$, $^{1}\chi^{v}$ and $^{2}\chi^{v}$ refers to valence connectivity index of zeroth, first and second order respectively)

Index	Group	nC	nO	nN	nCl	nS	$^{0}\chi^{v}$	$^{1}\chi^{v}$	$^{2}\chi^{v}$
1	CH_2	1	0	0	0	0	0.707107	0.000000	0.000000
2	C=C	2	0	0	0	0	1.000000	0.250000	0.000000
3	ACH	1	0	0	0	0	0.577350	0.666667	0.192450
4	$ACCH_2$	2	0	0	0	0	1.207107	0.353553	0.000000
5	OH	0	1	0	0	0	0.447214	0.000000	0.000000
6	CH_3OH	1	1	0	0	0	1.447214	0.447214	0.000000
7	H_2O	0	1	0	0	0	0.500000	0.000000	0.000000
8	ACOH	1	1	0	0	0	0.714475	0.223607	0.000000
9	CH_2CO	2	1	0	0	0	1.615355	0.557678	0.144338
10	CHO	1	1	0	0	0	0.985599	0.235702	0.000000
11	CCOO	2	2	0	0	0	1.816497	0.658248	0.287457
12	HCOO	1	2	0	0	0	1.393847	0.471405	0.096225
13	CH_2O	1	1	0	0	0	1.115355	0.288675	0.000000
14	CNH_2	1	0	1	0	0	1.577350	0.577350	0.000000
15	CNH	1	0	1	0	0	1.500000	0.500000	0.000000
16	$(C)_3N$	1	0	1	0	0	1.447214	0.447214	0.000000
17	$ACNH_2$	1	0	1	0	0	0.844612	0.154303	0.000000
18	PYRIDINE	5	0	1	0	0	1.654012	0.455942	0.125680
19	CCN	2	0	1	0	0	1.447214	0.473607	0.111803
20	COOH	1	2	0	0	0	1.355462	0.427731	0.091287
21	CCl	1	0	0	1	0	2.133893	1.133893	0.000000
22	CCl_2	1	0	0	2	0	2.974894	2.087498	0.909137
23	CCl_3	1	0	0	3	0	3.979031	1.963961	2.226922
24	CCl_4	1	0	0	4	0	5.035574	2.267787	3.857143
25	ACCl	1	0	0	1	0	1.633893	0.566947	0.000000
26	CNO_2	1	2	1	0	0	2.224745	0.741582	0.401375
27	$ACNO_2$	1	2	1	0	0	1.492006	0.442442	0.157128
28	CS_2	1	0	0	0	2	1.316497	0.408248	0.083333
29	CH_3SH	1	0	0	0	1	1.447214	0.447214	0.000000
30	FURFURAL	5	2	0	0	0	2.277456	0.723711	0.227611
31	DOH	2	2	0	0	0	2.308641	1.132456	0.447214
32	N/A	0	0	0	0	0	0.000000	0.000000	0.000000
33	N/A	0	0	0	0	0	0.000000	0.000000	0.000000

34	C=-C	2	0	0	0	0	1.000000	0.250000	0.000000
35	DMSO	2	1	0	0	1	2.816497	0.983163	0.741582
36	ACRY	3	0	1	0	0	2.231671	0.920530	0.333224
37	N/A	0	0	0	0	0	0.000000	0.000000	0.000000
38	ACF	1	0	0	0	0	0.645226	0.243872	0.073085
39	DMF	3	1	1	0	0	3.432812	1.388328	1.069021
40	CF_2	1	0	0	0	0	1.255929	0.377964	0.071429
41	COO	1	2	0	0	0	1.524564	0.546874	0.129099
42	c-CH_2	1	0	0	0	0	0.707107	1.000000	0.353553
43	c-CH_2O	2	1	0	0	0	1.115355	0.577350	0.204124
44	HCOOH	1	2	0	0	0	1.432812	0.493901	0.105409
45	$CHCl_3$	1	0	0	3	0	3.401680	1.963961	2.226922

APPENDIX B

ATOM INTERACTION PARAMETER TABLES FOR THE UNIFAC-CI MODELS

This appendix lists the atom interaction parameters (AIPs) which have been regressed for the Original UNIFAC-CI (VLE), Original UNIFAC-CI (VLE) with Q_{VLE}, Original UNIFAC-CI (VLE/SLE) and Modified (Dortmund) UNIFAC-CI (VLE) models respectively in Tables B.1, B.2, B.3 and B.4. This information can be used in Equations (2.39) to (2.42) together with the atom stoichiometry and CIs values presented in Tables A.1 and A.2.

Table B.1: Atom Interaction Parameters (AIPs) for the Original UNIFAC-CI (VLE) Model for Groups Involving Atoms C, H, O, N, Cl and S

AIPs	Value	AIPs	Value	AIPs	Value	AIPs	Value
bC-C	977.798010	dN-N	573.940034	bhC-C	-145.101080	dhN-N	-2572.089882
cC-C	-108.109618	eN-N	497.295700	chC-C	-281.567688	ehN-N	-498.708977
dC-C	104.616161	bN-Cl	0.000000	dhC-C	321.430980	bhN-Cl	0.000000
eC-C	-109.427513	cN-Cl	0.000000	ehC-C	261.303370	chN-Cl	0.000000
bC-O	-1149.588635	dN-Cl	0.000000	bhC-O	398.873452	dhN-Cl	0.000000
cC-O	-65.433629	eN-Cl	0.000000	chC-O	-4.318437	ehN-Cl	0.000000
dC-O	52.793952	bN-S	0.000000	dhC-O	245.592055	bhN-S	0.000000
eC-O	67.552287	cN-S	0.000000	ehC-O	-145.789583	chN-S	0.000000
bC-N	-1064.223191	dN-S	0.000000	bhC-N	0.000000	dhN-S	0.000000
cC-N	-514.845402	eN-S	0.000000	chC-N	0.000000	ehN-S	0.000000
dC-N	-85.055436	bCl-C	0.000000	dhC-N	0.000000	bhCl-C	124.354615
eC-N	139.293043	cCl-C	0.000000	ehC-N	0.000000	chCl-C	-11.845058
bC-Cl	-1326.720289	dCl-C	0.000000	bhC-Cl	0.000000	dhCl-C	43.954029
cC-Cl	1001.851262	eCl-C	0.000000	chC-Cl	0.000000	ehCl-C	-214.189508
dC-Cl	10.665755	bCl-O	0.000000	dhC-Cl	0.000000	bhCl-O	-1078.077642
eC-Cl	439.457629	cCl-O	0.000000	ehC-Cl	0.000000	chCl-O	-374.131842
bC-S	157.014181	dCl-O	0.000000	bhC-S	0.000000	dhCl-O	-38.593238
cC-S	731.330274	eCl-O	0.000000	chC-S	0.000000	ehCl-O	53.130759
dC-S	-570.394496	bCl-N	0.000000	dhC-S	0.000000	bhCl-N	0.000000
eC-S	1302.784061	cCl-N	0.000000	ehC-S	0.000000	chCl-N	0.000000
bO-C	-317.675623	dCl-N	0.000000	bhO-C	71.561153	dhCl-N	0.000000
cO-C	-789.806009	eCl-N	0.000000	chO-C	-35.732062	ehCl-N	0.000000
dO-C	-757.619173	bCl-Cl	0.000000	dhO-C	196.484939	bhCl-Cl	0.000000

eO-C	363.211586	cCl-Cl	0.000000	ehO-C	-29.140294	chCl-Cl	0.000000
bO-O	637.944192	dCl-Cl	0.000000	bhO-O	-515.401867	dhCl-Cl	0.000000
cO-O	1125.880929	eCl-Cl	0.000000	chO-O	-46.692178	ehCl-Cl	0.000000
dO-O	501.581648	bCl-S	21.942784	dhO-O	111.317117	bhCl-S	0.000000
eO-O	-384.001962	cCl-S	-685.618830	ehO-O	-188.160430	chCl-S	0.000000
bO-N	-4084.903877	dCl-S	-351.030653	bhO-N	0.000000	dhCl-S	0.000000
cO-N	732.738450	eCl-S	0.000000	chO-N	0.000000	ehCl-S	0.000000
dO-N	2002.380097	bS-C	0.000000	dhO-N	0.000000	bhS-C	-59.498556
eO-N	-1614.735117	cS-C	0.000000	ehO-N	0.000000	chS-C	-129.914433
bO-Cl	2438.066568	dS-C	0.000000	bhO-Cl	0.000000	dhS-C	-111.146487
cO-Cl	-2170.907517	eS-C	0.000000	chO-Cl	0.000000	ehS-C	-46.097451
dO-Cl	212.443154	bS-O	0.000000	dhO-Cl	0.000000	bhS-O	1132.651416
eO-Cl	-961.611159	cS-O	0.000000	ehO-Cl	0.000000	chS-O	42.186971
bO-S	809.255822	dS-O	0.000000	bhO-S	0.000000	dhS-O	285.668646
cO-S	-465.125263	eS-O	0.000000	chO-S	0.000000	ehS-O	9.862765
dO-S	50.649440	bS-N	0.000000	dhO-S	0.000000	bhS-N	0.000000
eO-S	0.000138	cS-N	0.000000	ehO-S	0.000000	chS-N	0.000000
bN-C	0.000000	dS-N	0.000000	bhN-C	61.100263	dhS-N	0.000000
cN-C	0.000000	eS-N	0.000000	chN-C	320.430632	ehS-N	0.000000
dN-C	0.000000	bS-Cl	0.000000	dhN-C	-224.701949	bhS-Cl	2.025435
eN-C	0.000000	cS-Cl	0.000000	ehN-C	-129.521998	chS-Cl	179.936421
bN-O	0.000000	dS-Cl	0.000000	bhN-O	-323.315611	dhS-Cl	-341.309545
cN-O	0.000000	eS-Cl	0.000000	chN-O	1142.620037	ehS-Cl	276.236177
dN-O	0.000000	bS-S	0.000000	dhN-O	-724.832431	bhS-S	0.000000
eN-O	0.000000	cS-S	0.000000	ehN-O	-365.593645	chS-S	0.000000
bN-N	-1852.384110	dS-S	0.000000	bhN-N	14063.227926	dhS-S	0.000000
cN-N	-2127.656326	eS-S	0.000000	chN-N	68.840361	ehS-S	0.000000

Table B.2: Atom Interaction Parameters (AIPs) for the Original UNIFAC-CI (VLE) with Q_{VLE} Model for Groups Involving Atoms C, H, O, N, Cl and S

AIPs	Value	AIPs	Value	AIPs	Value	AIPs	Value
bC-C	977.221160	dN-N	-0.101016	bhC-C	-143.790305	dhN-N	-346.987918
cC-C	-105.355284	eN-N	407.170479	chC-C	-281.478471	ehN-N	-319.370130
dC-C	108.436597	bN-Cl	0.000000	dhC-C	322.029593	bhN-Cl	0.000000
eC-C	-109.244669	cN-Cl	0.000000	ehC-C	260.603848	chN-Cl	0.000000
bC-O	-1126.870946	dN-Cl	0.000000	bhC-O	0.000000	dhN-Cl	0.000000
cC-O	-13.957176	eN-Cl	0.000000	chC-O	0.000000	ehN-Cl	0.000000
dC-O	40.542377	bN-S	0.000000	dhC-O	0.000000	bhN-S	0.000000
eC-O	66.753710	cN-S	0.000000	ehC-O	0.000000	chN-S	0.000000

bC-N	-1035.684806	dN-S	0.000000	bhC-N	0.000000	dhN-S	0.000000
cC-N	-453.514351	eN-S	0.000000	chC-N	0.000000	ehN-S	0.000000
dC-N	-87.848470	bCl-C	0.000000	dhC-N	0.000000	bhCl-C	120.473226
eC-N	141.045288	cCl-C	0.000000	ehC-N	0.000000	chCl-C	-18.549072
bC-Cl	-1337.833788	dCl-C	0.000000	bhC-Cl	0.000000	dhCl-C	50.761090
cC-Cl	982.869465	eCl-C	0.000000	chC-Cl	0.000000	ehCl-C	-211.415453
dC-Cl	3.347233	bCl-O	0.000000	dhC-Cl	0.000000	bhCl-O	-534.493831
eC-Cl	424.578644	cCl-O	0.000000	ehC-Cl	0.000000	chCl-O	-355.100521
bC-S	95.998356	dCl-O	0.000000	bhC-S	0.000000	dhCl-O	601.427235
cC-S	596.500305	eCl-O	0.000000	chC-S	0.000000	ehCl-O	-73.672079
dC-S	-114.740939	bCl-N	0.000000	dhC-S	0.000000	bhCl-N	0.000000
eC-S	35.921502	cCl-N	0.000000	ehC-S	0.000000	chCl-N	0.000000
bO-C	0.000000	dCl-N	0.000000	bhO-C	37.130338	dhCl-N	0.000000
cO-C	0.000000	eCl-N	0.000000	chO-C	-39.436304	ehCl-N	0.000000
dO-C	0.000000	bCl-Cl	0.000000	dhO-C	203.142668	bhCl-Cl	0.000000
eO-C	0.000000	cCl-Cl	0.000000	ehO-C	-27.786015	chCl-Cl	0.000000
bO-O	-200.543673	dCl-Cl	0.000000	bhO-O	-768.403907	dhCl-Cl	0.000000
cO-O	105.960612	eCl-Cl	0.000000	chO-O	156.814079	ehCl-Cl	0.000000
dO-O	162.075344	bCl-S	3044.988384	dhO-O	114.540372	bhCl-S	0.000000
eO-O	-15.380554	cCl-S	-611.508120	ehO-O	-452.368469	chCl-S	0.000000
bO-N	-454.814719	dCl-S	-1148.007777	bhO-N	0.000000	dhCl-S	0.000000
cO-N	307.183562	eCl-S	223.675291	chO-N	0.000000	ehCl-S	0.000000
dO-N	-108.981382	bS-C	0.000000	dhO-N	0.000000	bhS-C	152.267246
eO-N	-129.528705	cS-C	0.000000	ehO-N	0.000000	chS-C	-60.074867
bO-Cl	-136.972449	dS-C	0.000000	bhO-Cl	0.000000	dhS-C	2.980572
cO-Cl	-201.858021	eS-C	0.000000	chO-Cl	0.000000	ehS-C	-205.227812
dO-Cl	223.471160	bS-O	0.000000	dhO-Cl	0.000000	bhS-O	64.931509
eO-Cl	239.262595	cS-O	0.000000	ehO-Cl	0.000000	chS-O	79.892022
bO-S	47.599396	dS-O	0.000000	bhO-S	0.000000	dhS-O	142.947404
cO-S	-177.728626	eS-O	0.000000	chO-S	0.000000	ehS-O	335.562695
dO-S	201.955253	bS-N	0.000000	dhO-S	0.000000	bhS-N	0.000000
eO-S	107.732415	cS-N	0.000000	ehO-S	0.000000	chS-N	0.000000
bN-C	0.000000	dS-N	0.000000	bhN-C	59.237359	dhS-N	0.000000
cN-C	0.000000	eS-N	0.000000	chN-C	309.568156	ehS-N	0.000000
dN-C	0.000000	bS-Cl	0.000000	dhN-C	-207.770838	bhS-Cl	2905.062558
eN-C	0.000000	cS-Cl	0.000000	ehN-C	-128.904857	chS-Cl	1162.893873
bN-O	0.000000	dS-Cl	0.000000	bhN-O	-103.640548	dhS-Cl	2670.488955
cN-O	0.000000	eS-Cl	0.000000	chN-O	1430.564713	ehS-Cl	3709.522051
dN-O	0.000000	bS-S	0.000000	dhN-O	-617.780393	bhS-S	0.000000
eN-O	0.000000	cS-S	0.000000	ehN-O	-598.463202	chS-S	0.000000
bN-N	430.803627	dS-S	0.000000	bhN-N	-38.937322	dhS-S	0.000000

cN-N	-1792.219574	eS-S	0.000000	chN-N	186.725418	ehS-S	0.000000

Table B.3: Atom Interaction Parameters (AIPs) for the Original UNIFAC-CI (VLE/SLE) Model for Groups Involving Atoms C, H, O, N, Cl and S

AIPs	Value	AIPs	Value	AIPs	Value	AIPs	Value
bC-C	969.122164	dN-N	108.503380	bhC-C	-161.315844	dhN-N	-333.543668
cC-C	-112.327909	eN-N	561.058762	chC-C	-259.603896	ehN-N	-278.211298
dC-C	88.522386	bN-Cl	0.000000	dhC-C	299.226508	bhN-Cl	0.000000
eC-C	-111.717389	cN-Cl	0.000000	ehC-C	252.792688	chN-Cl	0.000000
bC-O	-1163.814039	dN-Cl	0.000000	bhC-O	323.442096	dhN-Cl	0.000000
cC-O	-15.163621	eN-Cl	0.000000	chC-O	-120.287237	ehN-Cl	0.000000
dC-O	59.035990	bN-S	0.000000	dhC-O	236.829459	bhN-S	0.000000
eC-O	69.120115	cN-S	0.000000	ehC-O	-145.789583	chN-S	0.000000
bC-N	-1191.536346	dN-S	0.000000	bhC-N	0.000000	dhN-S	0.000000
cC-N	-652.422035	eN-S	0.000000	chC-N	0.000000	ehN-S	0.000000
dC-N	-67.576772	bCl-C	0.000000	dhC-N	0.000000	bhCl-C	134.655374
eC-N	137.660406	cCl-C	0.000000	ehC-N	0.000000	chCl-C	-15.621784
bC-Cl	-1317.073177	dCl-C	0.000000	bhC-Cl	0.000000	dhCl-C	41.158380
cC-Cl	991.616627	eCl-C	0.000000	chC-Cl	0.000000	ehCl-C	-202.908827
dC-Cl	8.391793	bCl-O	0.000000	dhC-Cl	0.000000	bhCl-O	-1019.511670
eC-Cl	437.221808	cCl-O	0.000000	ehC-Cl	0.000000	chCl-O	-1054.455401
bC-S	239.326978	dCl-O	0.000000	bhC-S	0.000000	dhCl-O	1371.282721
cC-S	646.149175	eCl-O	0.000000	chC-S	0.000000	ehCl-O	65.093387
dC-S	-639.814667	bCl-N	0.000000	dhC-S	0.000000	bhCl-N	0.000000
eC-S	134.556937	cCl-N	0.000000	ehC-S	0.000000	chCl-N	0.000000
bO-C	-317.675623	dCl-N	0.000000	bhO-C	32.511818	dhCl-N	0.000000
cO-C	-789.806009	eCl-N	0.000000	chO-C	-82.151883	ehCl-N	0.000000
dO-C	-757.619173	bCl-Cl	0.000000	dhO-C	252.793700	bhCl-Cl	0.000000
eO-C	363.211586	cCl-Cl	0.000000	ehO-C	-17.157956	chCl-Cl	0.000000
bO-O	643.388444	dCl-Cl	0.000000	bhO-O	-437.045326	dhCl-Cl	0.000000
cO-O	1126.772823	eCl-Cl	0.000000	chO-O	28.546660	ehCl-Cl	0.000000
dO-O	452.362264	bCl-S	2488.266317	dhO-O	-26.166928	bhCl-S	0.000000
eO-O	-384.782304	cCl-S	-393.480184	ehO-O	-220.794151	chCl-S	0.000000
bO-N	-3089.620789	dCl-S	10.722425	bhO-N	0.000000	dhCl-S	0.000000
cO-N	975.681455	eCl-S	146.606266	chO-N	0.000000	ehCl-S	0.000000
dO-N	1705.784968	bS-C	0.000000	dhO-N	0.000000	bhS-C	201.969758
eO-N	-1378.916361	cS-C	0.000000	ehO-N	0.000000	chS-C	-99.511017
bO-Cl	832.321242	dS-C	0.000000	bhO-Cl	0.000000	dhS-C	-290.031240
cO-Cl	-416.512632	eS-C	0.000000	chO-Cl	0.000000	ehS-C	-28.655348

dO-Cl	2790.745664	bS-O	0.000000	dhO-Cl	0.000000	bhS-O	1105.941314
eO-Cl	282.840502	cS-O	0.000000	ehO-Cl	0.000000	chS-O	65.857232
bO-S	20.835842	dS-O	0.000000	bhO-S	0.000000	dhS-O	508.635927
cO-S	-324.894353	eS-O	0.000000	chO-S	0.000000	ehS-O	-18.411163
dO-S	245.800167	bS-N	0.000000	dhO-S	0.000000	bhS-N	0.000000
eO-S	-179.186033	cS-N	0.000000	ehO-S	0.000000	chS-N	0.000000
bN-C	0.000000	dS-N	0.000000	bhN-C	38.453296	dhS-N	0.000000
cN-C	0.000000	eS-N	0.000000	chN-C	324.061387	ehS-N	0.000000
dN-C	0.000000	bS-Cl	0.000000	dhN-C	-243.811414	bhS-Cl	5739.539386
eN-C	0.000000	cS-Cl	0.000000	ehN-C	-132.258378	chS-Cl	365.822284
bN-O	0.000000	dS-Cl	0.000000	bhN-O	-266.981833	dhS-Cl	2563.500748
cN-O	0.000000	eS-Cl	0.000000	chN-O	1252.721576	ehS-Cl	3238.875030
dN-O	0.000000	bS-S	0.000000	dhN-O	-733.607364	bhS-S	0.000000
eN-O	0.000000	cS-S	0.000000	ehN-O	-354.200368	chS-S	0.000000
bN-N	593.998428	dS-S	0.000000	bhN-N	-22.216573	dhS-S	0.000000
cN-N	-2371.763379	eS-S	0.000000	chN-N	144.261071	ehS-S	0.000000

Note: bX-Y, cX-Y, dX-Y, eX-Y, bhX-Y, chX-Y, dhX-Y and ehX-Y (between atom X and Y) in Table B.1, B.2 and B.3 refers to b_{X-Y}, c_{X-Y}, d_{X-Y}, e_{X-Y}, $\overline{b_{X-Y}}$, $\overline{c_{X-Y}}$, $\overline{d_{X-Y}}$, $\overline{e_{X-Y}}$ respectively in Equations (2.37) and (2.38).

Table B.4: Atom Interaction Parameters (AIPs) for the Modified (Dortmund) UNIFAC-CI (VLE) Model for Groups Involving Atoms C, O, H, Cl and S

AIPs	Value	AIPs	Value	AIPs	Value	AIPs	Value
bCl-C	571.859179	b1C-Cl	1.140529	bhC-Cl	0.000000	bh1C-Cl	0.000000
cCl-C	165.318592	c1C-Cl	-0.717766	chC-Cl	0.000000	ch1C-Cl	0.000000
dCl-C	-216.065623	d1C-Cl	0.622262	dhC-Cl	0.000000	dh1C-Cl	0.000000
eCl-C	325.079009	e1C-Cl	-1.029595	ehC-Cl	0.000000	eh1C-Cl	0.000000
bCl-O	-1107.308261	b1C-S	4.379278	bhC-S	0.000000	bh1C-S	0.000000
cCl-O	389.026923	c1C-S	-1.705668	chC-S	0.000000	ch1C-S	0.000000
dCl-O	107.042728	d1C-S	-0.324135	dhC-S	0.000000	dh1C-S	0.000000
eCl-O	57.984770	e1C-S	-0.243158	ehC-S	0.000000	eh1C-S	0.000000
bS-C	-1308.803742	b1O-Cl	16.549055	bhO-Cl	0.000000	bh1O-Cl	0.000000
cS-C	-684.856080	c1O-Cl	-5.415553	chO-Cl	0.000000	ch1O-Cl	0.000000
dS-C	0.000000	d1O-Cl	0.000000	dhO-Cl	0.000000	dh1O-Cl	0.000000
eS-C	-859.189823	e1O-Cl	-4.732760	ehO-Cl	0.000000	eh1O-Cl	0.000000
bCl-C	0.000000	b1Cl-C	0.000000	bhCl-C	-441.834370	bh1Cl-C	2.723888
cCl-C	0.000000	c1Cl-C	0.000000	chCl-C	-34.240608	ch1Cl-C	0.155094
dCl-C	0.000000	d1Cl-C	0.000000	dhCl-C	-24.971517	dh1Cl-C	0.150155

eCl-C	0.000000	e1Cl-C	0.000000	ehCl-C	113.749878	eh1Cl-C	-0.237555
bCl-O	0.000000	b1Cl-O	0.000000	bhCl-O	1758.307511	bh1Cl-O	-7.275750
cCl-O	0.000000	c1Cl-O	0.000000	chCl-O	-5115.861196	ch1Cl-O	20.520730
dCl-O	0.000000	d1Cl-O	0.000000	dhCl-O	-1124.402463	dh1Cl-O	3.031514
eCl-O	0.000000	e1Cl-O	0.000000	ehCl-O	0.000000	eh1Cl-O	0.000000
bS-C	0.000000	b1S-C	0.000000	bhS-C	-2295.483140	bh1S-C	7.465592
cS-C	0.000000	c1S-C	0.000000	chS-C	29.586806	ch1S-C	-0.238330
dS-C	0.000000	d1S-C	0.000000	dhS-C	170.181503	dh1S-C	-0.643862
eS-C	0.000000	e1S-C	0.000000	ehS-C	-289.006414	eh1S-C	1.283210

Note: bX-Y, cX-Y, dX-Y, eX-Y, b1X-Y, c1X-Y, d1X-Y, e1X-Y, bhX-Y, chX-Y, dhX-Y, ehX-Y, bh1X-Y, ch1X-Y, dh1X-Y and eh1X-Y (between atom X and Y) in Table B.4 refers to b_{X-Y}, c_{X-Y}, d_{X-Y}, e_{X-Y}, b^1_{X-Y}, c^1_{X-Y}, d^1_{X-Y}, e^1_{X-Y}, $\overline{b_{X-Y}}$, $\overline{c_{X-Y}}$, $\overline{d_{X-Y}}$, $\overline{e_{X-Y}}$, $\overline{b^1_{X-Y}}$, $\overline{c^1_{X-Y}}$, $\overline{d^1_{X-Y}}$, and $\overline{e^1_{X-Y}}$ respectively in Equations (C.1) and (C.2).

APPENDIX C

PREDICTION OF GIPS OF THE MODIFIED (DORTMUND) UNIFAC-CI FOR SYSTEMS INVOLVING ATOMS C, H, O, N, CL AND S

As it was discussed in Section 2.3.2, compared to the Original UNIFAC model, this modified model has three GIPs which overall have quadratic dependency towards temperature. This is shown in Equation (C.1).

$$a_{mn,overall} = a_{mn} + b_{mn}T + c_{mn}T^2 \tag{C.1}$$

In order to relate these GIPs with the atom stoichiometry, CIs used to defined each main groups and the AIPs, the relationship in Equation (C.1) and (C.2) have been derived to predict the missing GIPs of the Modified (Dortmund) UNIFAC model.

$$\begin{aligned}
a_{mn} = {} & b^1_{C-C}(A^{C-C}_{mn})_0 + b^1_{C-O}(A^{C-O}_{mn})_0 + b^1_{C-N}(A^{C-N}_{mn})_0 + b^1_{C-Cl}(A^{C-Cl}_{mn})_0 + b^1_{C-S}(A^{C-S}_{mn})_0 + b^1_{O-C}(A^{O-C}_{mn})_0 + b^1_{O-O}(A^{O-O}_{mn})_0 + b^1_{O-N}(A^{O-N}_{mn})_0 \\
& + b^1_{O-Cl}(A^{O-Cl}_{mn})_0 + b^1_{O-S}(A^{O-S}_{mn})_0 + b^1_{N-C}(A^{N-C}_{mn})_0 + b^1_{N-O}(A^{N-O}_{mn})_0 + b^1_{N-N}(A^{N-N}_{mn})_0 + b^1_{N-Cl}(A^{N-Cl}_{mn})_0 + b^1_{N-S}(A^{N-S}_{mn})_0 + b^1_{Cl-C}(A^{Cl-C}_{mn})_0 \\
& + \underbrace{b^1_{Cl-O}(A^{Cl-O}_{mn})_0 + b^1_{Cl-N}(A^{Cl-N}_{mn})_0 + b^1_{Cl-S}(A^{Cl-S}_{mn})_0 + b^1_{S-C}(A^{S-C}_{mn})_0 + b^1_{S-O}(A^{S-O}_{mn})_0 + b^1_{S-N}(A^{S-N}_{mn})_0 + b^1_{S-Cl}(A^{S-Cl}_{mn})_0 + b^1_{S-S}(A^{S-S}_{mn})_0}_{0th-order\,interaction} \\
& + c^1_{C-C}(A^{C-C}_{mn})_1 + c^1_{C-O}(A^{C-O}_{mn})_1 + c^1_{C-N}(A^{C-N}_{mn})_1 + c^1_{C-Cl}(A^{C-Cl}_{mn})_1 + c^1_{C-S}(A^{C-S}_{mn})_1 + c^1_{O-C}(A^{O-C}_{mn})_1 + c^1_{O-O}(A^{O-O}_{mn})_1 + c^1_{O-N}(A^{O-N}_{mn})_1 \\
& + c^1_{O-Cl}(A^{O-Cl}_{mn})_1 + c^1_{O-S}(A^{O-S}_{mn})_1 + c^1_{N-C}(A^{N-C}_{mn})_1 + c^1_{N-O}(A^{N-O}_{mn})_1 + c^1_{N-N}(A^{N-N}_{mn})_1 + c^1_{N-Cl}(A^{N-Cl}_{mn})_1 + c^1_{N-S}(A^{N-S}_{mn})_1 + c^1_{Cl-C}(A^{Cl-C}_{mn})_1 \\
& + \underbrace{c^1_{Cl-O}(A^{Cl-O}_{mn})_1 + c^1_{Cl-N}(A^{Cl-N}_{mn})_1 + c^1_{Cl-S}(A^{Cl-S}_{mn})_1 + c^1_{S-C}(A^{S-C}_{mn})_1 + c^1_{S-O}(A^{S-O}_{mn})_1 + c^1_{S-N}(A^{S-N}_{mn})_1 + c^1_{S-Cl}(A^{S-Cl}_{mn})_1 + c^1_{S-S}(A^{S-S}_{mn})_1}_{1st-order\,interaction} \\
& + d^1_{C-C}(A^{C-C}_{mn})_2 + d^1_{C-O}(A^{C-O}_{mn})_2 + d^1_{C-N}(A^{C-N}_{mn})_2 + d^1_{C-Cl}(A^{C-Cl}_{mn})_2 + d^1_{C-S}(A^{C-S}_{mn})_2 + d^1_{O-C}(A^{O-C}_{mn})_2 + d^1_{O-O}(A^{O-O}_{mn})_2 + d^1_{O-N}(A^{O-N}_{mn})_2 \\
& + d^1_{O-Cl}(A^{O-Cl}_{mn})_2 + d^1_{O-S}(A^{O-S}_{mn})_2 + d^1_{N-C}(A^{N-C}_{mn})_2 + d^1_{N-O}(A^{N-O}_{mn})_2 + d^1_{N-N}(A^{N-N}_{mn})_2 + d^1_{N-Cl}(A^{N-Cl}_{mn})_2 + d^1_{N-S}(A^{N-S}_{mn})_2 + d^1_{Cl-C}(A^{Cl-C}_{mn})_2 \\
& + \underbrace{d^1_{Cl-O}(A^{Cl-O}_{mn})_2 + d^1_{Cl-N}(A^{Cl-N}_{mn})_2 + d^1_{Cl-S}(A^{Cl-S}_{mn})_2 + d^1_{S-C}(A^{S-C}_{mn})_2 + d^1_{S-O}(A^{S-O}_{mn})_2 + d^1_{S-N}(A^{S-N}_{mn})_2 + d^1_{S-Cl}(A^{S-Cl}_{mn})_2 + d^1_{S-S}(A^{S-S}_{mn})_2}_{2nd-order\,interaction} \\
& + e^1_{C-C}(A^{C-C}_{mn})_3 + e^1_{C-O}(A^{C-O}_{mn})_3 + e^1_{C-N}(A^{C-N}_{mn})_3 + e^1_{C-Cl}(A^{C-Cl}_{mn})_3 + e^1_{C-S}(A^{C-S}_{mn})_3 + e^1_{O-C}(A^{O-C}_{mn})_3 + e^1_{O-O}(A^{O-O}_{mn})_3 + e^1_{O-N}(A^{O-N}_{mn})_3 \\
& + e^1_{O-Cl}(A^{O-Cl}_{mn})_3 + e^1_{O-S}(A^{O-S}_{mn})_3 + e^1_{N-C}(A^{N-C}_{mn})_3 + e^1_{N-O}(A^{N-O}_{mn})_3 + e^1_{N-N}(A^{N-N}_{mn})_3 + e^1_{N-Cl}(A^{N-Cl}_{mn})_3 + e^1_{N-S}(A^{N-S}_{mn})_3 + e^1_{Cl-C}(A^{Cl-C}_{mn})_3 \\
& + \underbrace{e^1_{Cl-O}(A^{Cl-O}_{mn})_3 + e^1_{Cl-N}(A^{Cl-N}_{mn})_3 + e^1_{Cl-S}(A^{Cl-S}_{mn})_3 + e^1_{S-C}(A^{S-C}_{mn})_3 + e^1_{S-O}(A^{S-O}_{mn})_3 + e^1_{S-N}(A^{S-N}_{mn})_3 + e^1_{S-Cl}(A^{S-Cl}_{mn})_3 + e^1_{S-S}(A^{S-S}_{mn})_3}_{3rd-order\,interaction}
\end{aligned}$$

$$\begin{aligned}
b_{mn} = {} & b^2_{C-C}(A^{C-C}_{mn})_0 + b^2_{C-O}(A^{C-O}_{mn})_0 + b^2_{C-N}(A^{C-N}_{mn})_0 + b^2_{C-Cl}(A^{C-Cl}_{mn})_0 + b^2_{C-S}(A^{C-S}_{mn})_0 + b^2_{O-C}(A^{O-C}_{mn})_0 + b^2_{O-O}(A^{O-O}_{mn})_0 + b^2_{O-N}(A^{O-N}_{mn})_0 \\
& + b^2_{O-Cl}(A^{O-Cl}_{mn})_0 + b^2_{O-S}(A^{O-S}_{mn})_0 + b^2_{N-C}(A^{N-C}_{mn})_0 + b^2_{N-O}(A^{N-O}_{mn})_0 + b^2_{N-N}(A^{N-N}_{mn})_0 + b^2_{N-Cl}(A^{N-Cl}_{mn})_0 + b^2_{N-S}(A^{N-S}_{mn})_0 + b^2_{Cl-C}(A^{Cl-C}_{mn})_0 \\
& + \underbrace{b^2_{Cl-O}(A^{Cl-O}_{mn})_0 + b^2_{Cl-N}(A^{Cl-N}_{mn})_0 + b^2_{Cl-S}(A^{Cl-S}_{mn})_0 + b^2_{S-C}(A^{S-C}_{mn})_0 + b^2_{S-O}(A^{S-O}_{mn})_0 + b^2_{S-N}(A^{S-N}_{mn})_0 + b^2_{S-Cl}(A^{S-Cl}_{mn})_0 + b^2_{S-S}(A^{S-S}_{mn})_0}_{0th-order\,interaction} \\
& + c^2_{C-C}(A^{C-C}_{mn})_1 + c^2_{C-O}(A^{C-O}_{mn})_1 + c^2_{C-N}(A^{C-N}_{mn})_1 + c^2_{C-Cl}(A^{C-Cl}_{mn})_1 + c^2_{C-S}(A^{C-S}_{mn})_1 + c^2_{O-C}(A^{O-C}_{mn})_1 + c^2_{O-O}(A^{O-O}_{mn})_1 + c^2_{O-N}(A^{O-N}_{mn})_1 \\
& + c^2_{O-Cl}(A^{O-Cl}_{mn})_1 + c^2_{O-S}(A^{O-S}_{mn})_1 + c^2_{N-C}(A^{N-C}_{mn})_1 + c^2_{N-O}(A^{N-O}_{mn})_1 + c^2_{N-N}(A^{N-N}_{mn})_1 + c^2_{N-Cl}(A^{N-Cl}_{mn})_1 + c^2_{N-S}(A^{N-S}_{mn})_1 + c^2_{Cl-C}(A^{Cl-C}_{mn})_1 \\
& + \underbrace{c^2_{Cl-O}(A^{Cl-O}_{mn})_1 + c^2_{Cl-N}(A^{Cl-N}_{mn})_1 + c^2_{Cl-S}(A^{Cl-S}_{mn})_1 + c^2_{S-C}(A^{S-C}_{mn})_1 + c^2_{S-O}(A^{S-O}_{mn})_1 + c^2_{S-N}(A^{S-N}_{mn})_1 + c^2_{S-Cl}(A^{S-Cl}_{mn})_1 + c^2_{S-S}(A^{S-S}_{mn})_1}_{1st-order\,interaction}
\end{aligned}$$

$$
\begin{aligned}
&+d^2_{C-C}(A^{C-C}_{mn})_2+d^2_{C-O}(A^{C-O}_{mn})_2+d^2_{C-N}(A^{C-N}_{mn})_2+d^2_{C-Cl}(A^{C-Cl}_{mn})_2+d^2_{C-S}(A^{C-S}_{mn})_2+d^2_{O-C}(A^{O-C}_{mn})_2+d^2_{O-O}(A^{O-O}_{mn})_2+d^2_{O-N}(A^{O-N}_{mn})_2\\
&+d^2_{O-Cl}(A^{O-Cl}_{mn})_2+d^2_{O-S}(A^{O-S}_{mn})_2+d^2_{N-C}(A^{N-C}_{mn})_2+d^2_{N-O}(A^{N-O}_{mn})_2+d^2_{N-N}(A^{N-N}_{mn})_2+d^2_{N-Cl}(A^{N-Cl}_{mn})_2+d^2_{N-S}(A^{N-S}_{mn})_2+d^2_{Cl-C}(A^{Cl-C}_{mn})_2\\
&+\underbrace{d^2_{Cl-O}(A^{Cl-O}_{mn})_2+d^2_{Cl-N}(A^{Cl-N}_{mn})_2+d^2_{Cl-S}(A^{Cl-S}_{mn})_2+d^2_{S-C}(A^{S-C}_{mn})_2+d^2_{S-O}(A^{S-O}_{mn})_2+d^2_{S-N}(A^{S-N}_{mn})_2+d^2_{S-Cl}(A^{S-Cl}_{mn})_2+d^2_{S-S}(A^{S-S}_{mn})_2}_{2nd-order\,interaction}\\
&+e^2_{C-C}(A^{C-C}_{mn})_3+e^2_{C-O}(A^{C-O}_{mn})_3+e^2_{C-N}(A^{C-N}_{mn})_3+e^2_{C-Cl}(A^{C-Cl}_{mn})_3+e^2_{C-S}(A^{C-S}_{mn})_3+e^2_{O-C}(A^{O-C}_{mn})_3+e^2_{O-O}(A^{O-O}_{mn})_3+e^2_{O-N}(A^{O-N}_{mn})_3\\
&+e^2_{O-Cl}(A^{O-Cl}_{mn})_3+e^2_{O-S}(A^{O-S}_{mn})_3+e^2_{N-C}(A^{N-C}_{mn})_3+e^2_{N-O}(A^{N-O}_{mn})_3+e^2_{N-N}(A^{N-N}_{mn})_3+e^2_{N-Cl}(A^{N-Cl}_{mn})_3+e^2_{N-S}(A^{N-S}_{mn})_3+e^2_{Cl-C}(A^{Cl-C}_{mn})_3\\
&+\underbrace{e^2_{Cl-O}(A^{Cl-O}_{mn})_3+e^2_{Cl-N}(A^{Cl-N}_{mn})_3+e^2_{Cl-S}(A^{Cl-S}_{mn})_3+e^2_{S-C}(A^{S-C}_{mn})_3+e^2_{S-O}(A^{S-O}_{mn})_3+e^2_{S-N}(A^{S-N}_{mn})_3+e^2_{S-Cl}(A^{S-Cl}_{mn})_3+e^2_{S-S}(A^{S-S}_{mn})_3}_{3rd-order\,interaction}
\end{aligned}
$$

$$
\begin{aligned}
c_{mn} &= b^3_{C-C}(A^{C-C}_{mn})_0+b^3_{C-O}(A^{C-O}_{mn})_0+b^3_{C-N}(A^{C-N}_{mn})_0+b^3_{C-Cl}(A^{C-Cl}_{mn})_0+b^3_{C-S}(A^{C-S}_{mn})_0+b^3_{O-C}(A^{O-C}_{mn})_0+b^3_{O-O}(A^{O-O}_{mn})_0+b^3_{O-N}(A^{O-N}_{mn})_0\\
&+b^3_{O-Cl}(A^{O-Cl}_{mn})_0+b^3_{O-S}(A^{O-S}_{mn})_0+b^3_{N-C}(A^{N-C}_{mn})_0+b^3_{N-O}(A^{N-O}_{mn})_0+b^3_{N-N}(A^{N-N}_{mn})_0+b^3_{N-Cl}(A^{N-Cl}_{mn})_0+b^3_{N-S}(A^{N-S}_{mn})_0+b^3_{Cl-C}(A^{Cl-C}_{mn})_0\\
&+\underbrace{b^3_{Cl-O}(A^{Cl-O}_{mn})_0+b^3_{Cl-N}(A^{Cl-N}_{mn})_0+b^3_{Cl-S}(A^{Cl-S}_{mn})_0+b^3_{S-C}(A^{S-C}_{mn})_0+b^3_{S-O}(A^{S-O}_{mn})_0+b^3_{S-N}(A^{S-N}_{mn})_0+b^3_{S-Cl}(A^{S-Cl}_{mn})_0+b^3_{S-S}(A^{S-S}_{mn})_0}_{0th-order\,interaction}\\
&+c^3_{C-C}(A^{C-C}_{mn})_1+c^3_{C-O}(A^{C-O}_{mn})_1+c^3_{C-N}(A^{C-N}_{mn})_1+c^3_{C-Cl}(A^{C-Cl}_{mn})_1+c^3_{C-S}(A^{C-S}_{mn})_1+c^3_{O-C}(A^{O-C}_{mn})_1+c^3_{O-O}(A^{O-O}_{mn})_1+c^3_{O-N}(A^{O-N}_{mn})_1\\
&+c^3_{O-Cl}(A^{O-Cl}_{mn})_1+c^3_{O-S}(A^{O-S}_{mn})_1+c^3_{N-C}(A^{N-C}_{mn})_1+c^3_{N-O}(A^{N-O}_{mn})_1+c^3_{N-N}(A^{N-N}_{mn})_1+c^3_{N-Cl}(A^{N-Cl}_{mn})_1+c^3_{N-S}(A^{N-S}_{mn})_1+c^3_{Cl-C}(A^{Cl-C}_{mn})_1\\
&+\underbrace{c^3_{Cl-O}(A^{Cl-O}_{mn})_1+c^3_{Cl-N}(A^{Cl-N}_{mn})_1+c^3_{Cl-S}(A^{Cl-S}_{mn})_1+c^3_{S-C}(A^{S-C}_{mn})_1+c^3_{S-O}(A^{S-O}_{mn})_1+c^3_{S-N}(A^{S-N}_{mn})_1+c^3_{S-Cl}(A^{S-Cl}_{mn})_1+c^3_{S-S}(A^{S-S}_{mn})_1}_{1st-order\,interaction}\\
&+d^3_{C-C}(A^{C-C}_{mn})_2+d^3_{C-O}(A^{C-O}_{mn})_2+d^3_{C-N}(A^{C-N}_{mn})_2+d^3_{C-Cl}(A^{C-Cl}_{mn})_2+d^3_{C-S}(A^{C-S}_{mn})_2+d^3_{O-C}(A^{O-C}_{mn})_2+d^3_{O-O}(A^{O-O}_{mn})_2+d^3_{O-N}(A^{O-N}_{mn})_2\\
&+d^3_{O-Cl}(A^{O-Cl}_{mn})_2+d^3_{O-S}(A^{O-S}_{mn})_2+d^3_{N-C}(A^{N-C}_{mn})_2+d^3_{N-O}(A^{N-O}_{mn})_2+d^3_{N-N}(A^{N-N}_{mn})_2+d^3_{N-Cl}(A^{N-Cl}_{mn})_2+d^3_{N-S}(A^{N-S}_{mn})_2+d^3_{Cl-C}(A^{Cl-C}_{mn})_2\\
&+\underbrace{d^3_{Cl-O}(A^{Cl-O}_{mn})_2+d^3_{Cl-N}(A^{Cl-N}_{mn})_2+d^3_{Cl-S}(A^{Cl-S}_{mn})_2+d^3_{S-C}(A^{S-C}_{mn})_2+d^3_{S-O}(A^{S-O}_{mn})_2+d^3_{S-N}(A^{S-N}_{mn})_2+d^3_{S-Cl}(A^{S-Cl}_{mn})_2+d^3_{S-S}(A^{S-S}_{mn})_2}_{2nd-order\,interaction}\\
&+e^3_{C-C}(A^{C-C}_{mn})_3+e^3_{C-O}(A^{C-O}_{mn})_3+e^3_{C-N}(A^{C-N}_{mn})_3+e^3_{C-Cl}(A^{C-Cl}_{mn})_3+e^3_{C-S}(A^{C-S}_{mn})_3+e^3_{O-C}(A^{O-C}_{mn})_3+e^3_{O-O}(A^{O-O}_{mn})_3+e^3_{O-N}(A^{O-N}_{mn})_3\\
&+e^3_{O-Cl}(A^{O-Cl}_{mn})_3+e^3_{O-S}(A^{O-S}_{mn})_3+e^3_{N-C}(A^{N-C}_{mn})_3+e^3_{N-O}(A^{N-O}_{mn})_3+e^3_{N-N}(A^{N-N}_{mn})_3+e^3_{N-Cl}(A^{N-Cl}_{mn})_3+e^3_{N-S}(A^{N-S}_{mn})_3+e^3_{Cl-C}(A^{Cl-C}_{mn})_3\\
&+\underbrace{e^3_{Cl-O}(A^{Cl-O}_{mn})_3+e^3_{Cl-N}(A^{Cl-N}_{mn})_3+e^3_{Cl-S}(A^{Cl-S}_{mn})_3+e^3_{S-C}(A^{S-C}_{mn})_3+e^3_{S-O}(A^{S-O}_{mn})_3+e^3_{S-N}(A^{S-N}_{mn})_3+e^3_{S-Cl}(A^{S-Cl}_{mn})_3+e^3_{S-S}(A^{S-S}_{mn})_3}_{3rd-order\,interaction}
\end{aligned}
\qquad \text{(C.1)}
$$

$$
\begin{aligned}
a_{nm} &= \overline{b^1_{C-C}}(A^{C-C}_{nm})_0+\overline{b^1_{C-O}}(A^{C-O}_{nm})_0+\overline{b^1_{C-N}}(A^{C-N}_{nm})_0+\overline{b^1_{C-Cl}}(A^{C-Cl}_{nm})_0+\overline{b^1_{C-S}}(A^{C-S}_{nm})_0+\overline{b^1_{O-C}}(A^{O-C}_{nm})_0+\overline{b^1_{O-O}}(A^{O-O}_{nm})_0+\overline{b^1_{O-N}}(A^{O-N}_{nm})_0\\
&+\overline{b^1_{O-Cl}}(A^{O-Cl}_{nm})_0+\overline{b^1_{O-S}}(A^{O-S}_{nm})_0+\overline{b^1_{N-C}}(A^{N-C}_{nm})_0+\overline{b^1_{N-O}}(A^{N-O}_{nm})_0+\overline{b^1_{N-N}}(A^{N-N}_{nm})_0+\overline{b^1_{N-Cl}}(A^{N-Cl}_{nm})_0+\overline{b^1_{N-S}}(A^{N-S}_{nm})_0+\overline{b^1_{Cl-C}}(A^{Cl-C}_{nm})_0\\
&+\underbrace{\overline{b^1_{Cl-O}}(A^{Cl-O}_{nm})_0+\overline{b^1_{Cl-N}}(A^{Cl-N}_{nm})_0+\overline{b^1_{Cl-S}}(A^{Cl-S}_{nm})_0+\overline{b^1_{S-C}}(A^{S-C}_{nm})_0+\overline{b^1_{S-O}}(A^{S-O}_{nm})_0+\overline{b^1_{S-N}}(A^{S-N}_{nm})_0+\overline{b^1_{S-Cl}}(A^{S-Cl}_{nm})_0+\overline{b^1_{S-S}}(A^{S-S}_{nm})_0}_{0th-order\,interaction}\\
&+\overline{c^1_{C-C}}(A^{C-C}_{nm})_1+\overline{c^1_{C-O}}(A^{C-O}_{nm})_1+\overline{c^1_{C-N}}(A^{C-N}_{nm})_1+\overline{c^1_{C-Cl}}(A^{C-Cl}_{nm})_1+\overline{c^1_{C-S}}(A^{C-S}_{nm})_1+\overline{c^1_{O-C}}(A^{O-C}_{nm})_1+\overline{c^1_{O-O}}(A^{O-O}_{nm})_1+\overline{c^1_{O-N}}(A^{O-N}_{nm})_1\\
&+\overline{c^1_{O-Cl}}(A^{O-Cl}_{nm})_1+\overline{c^1_{O-S}}(A^{O-S}_{nm})_1+\overline{c^1_{N-C}}(A^{N-C}_{nm})_1+\overline{c^1_{N-O}}(A^{N-O}_{nm})_1+\overline{c^1_{N-N}}(A^{N-N}_{nm})_1+\overline{c^1_{N-Cl}}(A^{N-Cl}_{nm})_1+\overline{c^1_{N-S}}(A^{N-S}_{nm})_1+\overline{c^1_{Cl-C}}(A^{Cl-C}_{nm})_1\\
&+\underbrace{\overline{c^1_{Cl-O}}(A^{Cl-O}_{nm})_1+\overline{c^1_{Cl-N}}(A^{Cl-N}_{nm})_1+\overline{c^1_{Cl-S}}(A^{Cl-S}_{nm})_1+\overline{c^1_{S-C}}(A^{S-C}_{nm})_1+\overline{c^1_{S-O}}(A^{S-O}_{nm})_1+\overline{c^1_{S-N}}(A^{S-N}_{nm})_1+\overline{c^1_{S-Cl}}(A^{S-Cl}_{nm})_1+\overline{c^1_{S-S}}(A^{S-S}_{nm})_1}_{1st-order\,interaction}\\
&+\overline{d^1_{C-C}}(A^{C-C}_{nm})_2+\overline{d^1_{C-O}}(A^{C-O}_{nm})_2+\overline{d^1_{C-N}}(A^{C-N}_{nm})_2+\overline{d^1_{C-Cl}}(A^{C-Cl}_{nm})_2+\overline{d^1_{C-S}}(A^{C-S}_{nm})_2+\overline{d^1_{O-C}}(A^{O-C}_{nm})_2+\overline{d^1_{O-O}}(A^{O-O}_{nm})_2+\overline{d^1_{O-N}}(A^{O-N}_{nm})_2\\
&+\overline{d^1_{O-Cl}}(A^{O-Cl}_{nm})_2+\overline{d^1_{O-S}}(A^{O-S}_{nm})_2+\overline{d^1_{N-C}}(A^{N-C}_{nm})_2+\overline{d^1_{N-O}}(A^{N-O}_{nm})_2+\overline{d^1_{N-N}}(A^{N-N}_{nm})_2+\overline{d^1_{N-Cl}}(A^{N-Cl}_{nm})_2+\overline{d^1_{N-S}}(A^{N-S}_{nm})_2+\overline{d^1_{Cl-C}}(A^{Cl-C}_{nm})_2\\
&+\underbrace{\overline{d^1_{Cl-O}}(A^{Cl-O}_{nm})_2+\overline{d^1_{Cl-N}}(A^{Cl-N}_{nm})_2+\overline{d^1_{Cl-S}}(A^{Cl-S}_{nm})_2+\overline{d^1_{S-C}}(A^{S-C}_{nm})_2+\overline{d^1_{S-O}}(A^{S-O}_{nm})_2+\overline{d^1_{S-N}}(A^{S-N}_{nm})_2+\overline{d^1_{S-Cl}}(A^{S-Cl}_{nm})_2+\overline{d^1_{S-S}}(A^{S-S}_{nm})_2}_{2nd-order\,interaction}\\
&+\overline{e^1_{C-C}}(A^{C-C}_{nm})_3+\overline{e^1_{C-O}}(A^{C-O}_{nm})_3+\overline{e^1_{C-N}}(A^{C-N}_{nm})_3+\overline{e^1_{C-Cl}}(A^{C-Cl}_{nm})_3+\overline{e^1_{C-S}}(A^{C-S}_{nm})_3+\overline{e^1_{O-C}}(A^{O-C}_{nm})_3+\overline{e^1_{O-O}}(A^{O-O}_{nm})_3+\overline{e^1_{O-N}}(A^{O-N}_{nm})_3\\
&+\overline{e^1_{O-Cl}}(A^{O-Cl}_{nm})_3+\overline{e^1_{O-S}}(A^{O-S}_{nm})_3+\overline{e^1_{N-C}}(A^{N-C}_{nm})_3+\overline{e^1_{N-O}}(A^{N-O}_{nm})_3+\overline{e^1_{N-N}}(A^{N-N}_{nm})_3+\overline{e^1_{N-Cl}}(A^{N-Cl}_{nm})_3+\overline{e^1_{N-S}}(A^{N-S}_{nm})_3+\overline{e^1_{Cl-C}}(A^{Cl-C}_{nm})_3\\
&+\underbrace{\overline{e^1_{Cl-O}}(A^{Cl-O}_{nm})_3+\overline{e^1_{Cl-N}}(A^{Cl-N}_{nm})_3+\overline{e^1_{Cl-S}}(A^{Cl-S}_{nm})_3+\overline{e^1_{S-C}}(A^{S-C}_{nm})_3+\overline{e^1_{S-O}}(A^{S-O}_{nm})_3+\overline{e^1_{S-N}}(A^{S-N}_{nm})_3+\overline{e^1_{S-Cl}}(A^{S-Cl}_{nm})_3+\overline{e^1_{S-S}}(A^{S-S}_{nm})_3}_{3rd-order\,interaction}
\end{aligned}
$$

$$
\begin{aligned}
b_{nm} &= \overline{b^2_{C-C}}(A^{C-C}_{nm})_0+\overline{b^2_{C-O}}(A^{C-O}_{nm})_0+\overline{b^2_{C-N}}(A^{C-N}_{nm})_0+\overline{b^2_{C-Cl}}(A^{C-Cl}_{nm})_0+\overline{b^2_{C-S}}(A^{C-S}_{nm})_0+\overline{b^2_{O-C}}(A^{O-C}_{nm})_0+\overline{b^2_{O-O}}(A^{O-O}_{nm})_0+\overline{b^2_{O-N}}(A^{O-N}_{nm})_0\\
&+\overline{b^2_{O-Cl}}(A^{O-Cl}_{nm})_0+\overline{b^2_{O-S}}(A^{O-S}_{nm})_0+\overline{b^2_{N-C}}(A^{N-C}_{nm})_0+\overline{b^2_{N-O}}(A^{N-O}_{nm})_0+\overline{b^2_{N-N}}(A^{N-N}_{nm})_0+\overline{b^2_{N-Cl}}(A^{N-Cl}_{nm})_0+\overline{b^2_{N-S}}(A^{N-S}_{nm})_0+\overline{b^2_{Cl-C}}(A^{Cl-C}_{nm})_0
\end{aligned}
$$

$$
\begin{aligned}
&\underbrace{+\overline{b^2_{Cl-O}}(A^{Cl-O}_{nm})_0+\overline{b^2_{Cl-N}}(A^{Cl-N}_{nm})_0+\overline{b^2_{Cl-S}}(A^{Cl-S}_{nm})_0+\overline{b^2_{S-C}}(A^{S-C}_{nm})_0+\overline{b^2_{S-O}}(A^{S-O}_{nm})_0+\overline{b^2_{S-N}}(A^{S-N}_{nm})_0+\overline{b^2_{S-Cl}}(A^{S-Cl}_{nm})_0+\overline{b^2_{S-S}}(A^{S-S}_{nm})_0}_{0th-order\,interaction}\\
&+\overline{c^2_{C-C}}(A^{C-C}_{nm})_1+\overline{c^2_{C-O}}(A^{C-O}_{nm})_1+\overline{c^2_{C-N}}(A^{C-N}_{nm})_1+\overline{c^2_{C-Cl}}(A^{C-Cl}_{nm})_1+\overline{c^2_{C-S}}(A^{C-S}_{nm})_1+\overline{c^2_{O-C}}(A^{O-C}_{nm})_1+\overline{c^2_{O-O}}(A^{O-O}_{nm})_1+\overline{c^2_{O-N}}(A^{O-N}_{nm})_1\\
&+\overline{c^2_{O-Cl}}(A^{O-Cl}_{nm})_1+\overline{c^2_{O-S}}(A^{O-S}_{nm})_1+\overline{c^2_{N-C}}(A^{N-C}_{nm})_1+\overline{c^2_{N-O}}(A^{N-O}_{nm})_1+\overline{c^2_{N-N}}(A^{N-N}_{nm})_1+\overline{c^2_{N-Cl}}(A^{N-Cl}_{nm})_1+\overline{c^2_{N-S}}(A^{N-S}_{nm})_1+\overline{c^2_{Cl-C}}(A^{Cl-C}_{nm})_1\\
&\underbrace{+\overline{c^2_{Cl-O}}(A^{Cl-O}_{nm})_1+\overline{c^2_{Cl-N}}(A^{Cl-N}_{nm})_1+\overline{c^2_{Cl-S}}(A^{Cl-S}_{nm})_1+\overline{c^2_{S-C}}(A^{S-C}_{nm})_1+\overline{c^2_{S-O}}(A^{S-O}_{nm})_1+\overline{c^2_{S-N}}(A^{S-N}_{nm})_1+\overline{c^2_{S-Cl}}(A^{S-Cl}_{nm})_1+\overline{c^2_{S-S}}(A^{S-S}_{nm})_1}_{1st-order\,interaction}\\
&+\overline{d^2_{C-C}}(A^{C-C}_{nm})_2+\overline{d^2_{C-O}}(A^{C-O}_{nm})_2+\overline{d^2_{C-N}}(A^{C-N}_{nm})_2+\overline{d^2_{C-Cl}}(A^{C-Cl}_{nm})_2+\overline{d^2_{C-S}}(A^{C-S}_{nm})_2+\overline{d^2_{O-C}}(A^{O-C}_{nm})_2+\overline{d^2_{O-O}}(A^{O-O}_{nm})_2+\overline{d^2_{O-N}}(A^{O-N}_{nm})_2\\
&+\overline{d^2_{O-Cl}}(A^{O-Cl}_{nm})_2+\overline{d^2_{O-S}}(A^{O-S}_{nm})_2+\overline{d^2_{N-C}}(A^{N-C}_{nm})_2+\overline{d^2_{N-O}}(A^{N-O}_{nm})_2+\overline{d^2_{N-N}}(A^{N-N}_{nm})_2+\overline{d^2_{N-Cl}}(A^{N-Cl}_{nm})_2+\overline{d^2_{N-S}}(A^{N-S}_{nm})_2+\overline{d^2_{Cl-C}}(A^{Cl-C}_{nm})_2\\
&\underbrace{+\overline{d^2_{Cl-O}}(A^{Cl-O}_{nm})_2+\overline{d^2_{Cl-N}}(A^{Cl-N}_{nm})_2+\overline{d^2_{Cl-S}}(A^{Cl-S}_{nm})_2+\overline{d^2_{S-C}}(A^{S-C}_{nm})_2+\overline{d^2_{S-O}}(A^{S-O}_{nm})_2+\overline{d^2_{S-N}}(A^{S-N}_{nm})_2+\overline{d^2_{S-Cl}}(A^{S-Cl}_{nm})_2+\overline{d^2_{S-S}}(A^{S-S}_{nm})_2}_{2nd-order\,interaction}\\
&+\overline{e^2_{C-C}}(A^{C-C}_{nm})_3+\overline{e^2_{C-O}}(A^{C-O}_{nm})_3+\overline{e^2_{C-N}}(A^{C-N}_{nm})_3+\overline{e^2_{C-Cl}}(A^{C-Cl}_{nm})_3+\overline{e^2_{C-S}}(A^{C-S}_{nm})_3+\overline{e^2_{O-C}}(A^{O-C}_{nm})_3+\overline{e^2_{O-O}}(A^{O-O}_{nm})_3+\overline{e^2_{O-N}}(A^{O-N}_{nm})_3\\
&+\overline{e^2_{O-Cl}}(A^{O-Cl}_{nm})_3+\overline{e^2_{O-S}}(A^{O-S}_{nm})_3+\overline{e^2_{N-C}}(A^{N-C}_{nm})_3+\overline{e^2_{N-O}}(A^{N-O}_{nm})_3+\overline{e^2_{N-N}}(A^{N-N}_{nm})_3+\overline{e^2_{N-Cl}}(A^{N-Cl}_{nm})_3+\overline{e^2_{N-S}}(A^{N-S}_{nm})_3+\overline{e^2_{Cl-C}}(A^{Cl-C}_{nm})_3\\
&\underbrace{+\overline{e^2_{Cl-O}}(A^{Cl-O}_{nm})_3+\overline{e^2_{Cl-N}}(A^{Cl-N}_{nm})_3+\overline{e^2_{Cl-S}}(A^{Cl-S}_{nm})_3+\overline{e^2_{S-C}}(A^{S-C}_{nm})_3+\overline{e^2_{S-O}}(A^{S-O}_{nm})_3+\overline{e^2_{S-N}}(A^{S-N}_{nm})_3+\overline{e^2_{S-Cl}}(A^{S-Cl}_{nm})_3+\overline{e^2_{S-S}}(A^{S-S}_{nm})_3}_{3rd-order\,interaction}
\end{aligned}
$$

$$
\begin{aligned}
c_{nm}=&\underbrace{\overline{b^3_{C-C}}(A^{C-C}_{nm})_0+\overline{b^3_{C-O}}(A^{C-O}_{nm})_0+\overline{b^3_{C-N}}(A^{C-N}_{nm})_0+\overline{b^3_{C-Cl}}(A^{C-Cl}_{nm})_0+\overline{b^3_{C-S}}(A^{C-S}_{nm})_0+\overline{b^3_{O-C}}(A^{O-C}_{nm})_0+\overline{b^3_{O-O}}(A^{O-O}_{nm})_0+\overline{b^3_{O-N}}(A^{O-N}_{nm})_0}\\
&+\overline{b^3_{O-Cl}}(A^{O-Cl}_{nm})_0+\overline{b^3_{O-S}}(A^{O-S}_{nm})_0+\overline{b^3_{N-C}}(A^{N-C}_{nm})_0+\overline{b^3_{N-O}}(A^{N-O}_{nm})_0+\overline{b^3_{N-N}}(A^{N-N}_{nm})_0+\overline{b^3_{N-Cl}}(A^{N-Cl}_{nm})_0+\overline{b^3_{N-S}}(A^{N-S}_{nm})_0+\overline{b^3_{Cl-C}}(A^{Cl-C}_{nm})_0\\
&\underbrace{+\overline{b^3_{Cl-O}}(A^{Cl-O}_{nm})_0+\overline{b^3_{Cl-N}}(A^{Cl-N}_{nm})_0+\overline{b^3_{Cl-S}}(A^{Cl-S}_{nm})_0+\overline{b^3_{S-C}}(A^{S-C}_{nm})_0+\overline{b^3_{S-O}}(A^{S-O}_{nm})_0+\overline{b^3_{S-N}}(A^{S-N}_{nm})_0+\overline{b^3_{S-Cl}}(A^{S-Cl}_{nm})_0+\overline{b^3_{S-S}}(A^{S-S}_{nm})_0}_{0th-order\,interaction}\\
&+\overline{c^3_{C-C}}(A^{C-C}_{nm})_1+\overline{c^3_{C-O}}(A^{C-O}_{nm})_1+\overline{c^3_{C-N}}(A^{C-N}_{nm})_1+\overline{c^3_{C-Cl}}(A^{C-Cl}_{nm})_1+\overline{c^3_{C-S}}(A^{C-S}_{nm})_1+\overline{c^3_{O-C}}(A^{O-C}_{nm})_1+\overline{c^3_{O-O}}(A^{O-O}_{nm})_1+\overline{c^3_{O-N}}(A^{O-N}_{nm})_1\\
&+\overline{c^3_{O-Cl}}(A^{O-Cl}_{nm})_1+\overline{c^3_{O-S}}(A^{O-S}_{nm})_1+\overline{c^3_{N-C}}(A^{N-C}_{nm})_1+\overline{c^3_{N-O}}(A^{N-O}_{nm})_1+\overline{c^3_{N-N}}(A^{N-N}_{nm})_1+\overline{c^3_{N-Cl}}(A^{N-Cl}_{nm})_1+\overline{c^3_{N-S}}(A^{N-S}_{nm})_1+\overline{c^3_{Cl-C}}(A^{Cl-C}_{nm})_1\\
&\underbrace{+\overline{c^3_{Cl-O}}(A^{Cl-O}_{nm})_1+\overline{c^3_{Cl-N}}(A^{Cl-N}_{nm})_1+\overline{c^3_{Cl-S}}(A^{Cl-S}_{nm})_1+\overline{c^3_{S-C}}(A^{S-C}_{nm})_1+\overline{c^3_{S-O}}(A^{S-O}_{nm})_1+\overline{c^3_{S-N}}(A^{S-N}_{nm})_1+\overline{c^3_{S-Cl}}(A^{S-Cl}_{nm})_1+\overline{c^3_{S-S}}(A^{S-S}_{nm})_1}_{1st-order\,interaction}\\
&+\overline{d^3_{C-C}}(A^{C-C}_{nm})_2+\overline{d^3_{C-O}}(A^{C-O}_{nm})_2+\overline{d^3_{C-N}}(A^{C-N}_{nm})_2+\overline{d^3_{C-Cl}}(A^{C-Cl}_{nm})_2+\overline{d^3_{C-S}}(A^{C-S}_{nm})_2+\overline{d^3_{O-C}}(A^{O-C}_{nm})_2+\overline{d^3_{O-O}}(A^{O-O}_{nm})_2+\overline{d^3_{O-N}}(A^{O-N}_{nm})_2\\
&+\overline{d^3_{O-Cl}}(A^{O-Cl}_{nm})_2+\overline{d^3_{O-S}}(A^{O-S}_{nm})_2+\overline{d^3_{N-C}}(A^{N-C}_{nm})_2+\overline{d^3_{N-O}}(A^{N-O}_{nm})_2+\overline{d^3_{N-N}}(A^{N-N}_{nm})_2+\overline{d^3_{N-Cl}}(A^{N-Cl}_{nm})_2+\overline{d^3_{N-S}}(A^{N-S}_{nm})_2+\overline{d^3_{Cl-C}}(A^{Cl-C}_{nm})_2\\
&\underbrace{+\overline{d^3_{Cl-O}}(A^{Cl-O}_{nm})_2+\overline{d^3_{Cl-N}}(A^{Cl-N}_{nm})_2+\overline{d^3_{Cl-S}}(A^{Cl-S}_{nm})_2+\overline{d^3_{S-C}}(A^{S-C}_{nm})_2+\overline{d^3_{S-O}}(A^{S-O}_{nm})_2+\overline{d^3_{S-N}}(A^{S-N}_{nm})_2+\overline{d^3_{S-Cl}}(A^{S-Cl}_{nm})_2+\overline{d^3_{S-S}}(A^{S-S}_{nm})_2}_{2nd-order\,interaction}\\
&+\overline{e^3_{C-C}}(A^{C-C}_{nm})_3+\overline{e^3_{C-O}}(A^{C-O}_{nm})_3+\overline{e^3_{C-N}}(A^{C-N}_{nm})_3+\overline{e^3_{C-Cl}}(A^{C-Cl}_{nm})_3+\overline{e^3_{C-S}}(A^{C-S}_{nm})_3+\overline{e^3_{O-C}}(A^{O-C}_{nm})_3+\overline{e^3_{O-O}}(A^{O-O}_{nm})_3+\overline{e^3_{O-N}}(A^{O-N}_{nm})_3\\
&+\overline{e^3_{O-Cl}}(A^{O-Cl}_{nm})_3+\overline{e^3_{O-S}}(A^{O-S}_{nm})_3+\overline{e^3_{N-C}}(A^{N-C}_{nm})_3+\overline{e^3_{N-O}}(A^{N-O}_{nm})_3+\overline{e^3_{N-N}}(A^{N-N}_{nm})_3+\overline{e^3_{N-Cl}}(A^{N-Cl}_{nm})_3+\overline{e^3_{N-S}}(A^{N-S}_{nm})_3+\overline{e^3_{Cl-C}}(A^{Cl-C}_{nm})_3\\
&\underbrace{+\overline{e^3_{Cl-O}}(A^{Cl-O}_{nm})_3+\overline{e^3_{Cl-N}}(A^{Cl-N}_{nm})_3+\overline{e^3_{Cl-S}}(A^{Cl-S}_{nm})_3+\overline{e^3_{S-C}}(A^{S-C}_{nm})_3+\overline{e^3_{S-O}}(A^{S-O}_{nm})_3+\overline{e^3_{S-N}}(A^{S-N}_{nm})_3+\overline{e^3_{S-Cl}}(A^{S-Cl}_{nm})_3+\overline{e^3_{S-S}}(A^{S-S}_{nm})_3}_{3rd-order\,interaction}
\end{aligned}
\qquad \text{(C.2)}
$$

In the above equations, the coeffcients $(A^{xy}_{mn})_i$ can also be calculated using Equations (2.39)-(2.42) like for the Original UNIFAC-CI model. Since there are three sets of GIPs, there are also three sets of AIPs which are labeled 1, 2, 3 that correspond to the a_{mn}, b_{mn} and c_{mn} GIPs respectively. However, the parameter c_{mn} has been set to zero in this work because there were not enough experimental data for taking into account extensive temperature effects. Only VLE data were used to regress the remaining adjustable parameters. The calculations of the missing GIPs are more or less the same as for the Original UNIFAC-CI model, but with one more additional set with respect to b_{mn}. Without the use of experimental data such as the activity coefficient at infinite dilution and excess enthalpy to regress the parameters just like the host modified UNIFAC model, it can found that the prediction of those properties using CI-generated parameters are not reliable.

APPENDIX D

CALCULATIONS OF GROUP INTERACTION PARAMETERS OF CH_2-ACCO AND CH_2-ACNH FROM THE CI-METHOD

In Chapter 7, new groups (ACCO and ACNH) have been created and their corresponding valence connectivity indices have been calculated. In addition, their interaction parameters with other groups present in the case studies are also calculated. Example of the calculations of the GIPs between these new groups with UNIFAC main group CH_2 using the relationships which have been established in Equations (2.37)-(2.42). is shown below:

Calculations of the GIPs between CH_2 and ACCO

1) The stoichiometry and the CI values needed is given in Table D.1

Table D.1: Atom Stoichiometry and CI Values for CH_2-ACCO

Group	n_C	n_O	${}^{v}\chi^{0}$	${}^{v}\chi^{1}$	${}^{v}\chi^{2}$
CH_2	1	0	0.7071	0.0000	0.0000
ACCO	2	1	1.4083	0.4541	0.1021

2) Using Equations (2.37)-(2.42) and the atom interaction parameters (AIPs) presented in Table D.2, the group interaction parameters $a_{CH2\text{-}ACCO}$ and $a_{ACCO\text{-}CH2}$ are calculated as follows:

$$a_{CH_2-ACCO} = b_{C-C}\frac{n_C^{(CH_2)}\chi_0^{(ACCO)} - n_C^{(ACCO)}\chi_0^{(CH_2)}}{\chi_0^{(ACCO)}\chi_0^{(CH_2)}} + c_{C-C}\frac{n_C^{(CH_2)}\chi_1^{(ACCO)} - n_C^{(ACCO)}\chi_0^{(CH_2)}}{\chi_1^{(ACCO)}\chi_0^{(CH_2)}} +$$

$$e_{C-C}\frac{n_C^{(CH_2)}\chi_2^{(ACCO)} - n_C^{(ACCO)}\chi_0^{(CH_2)}}{\chi_2^{(ACCO)}\chi_0^{(CH_2)}} + b_{C-O}\frac{n_C^{(CH_2)}\chi_0^{(ACCO)} - n_O^{(ACCO)}\chi_0^{(CH_2)}}{\chi_0^{(ACCO)}\chi_0^{(CH_2)}} +$$

$$c_{C-O}\frac{n_C^{(CH_2)}\chi_1^{(ACCO)} - n_O^{(ACCO)}\chi_0^{(CH_2)}}{\chi_1^{(ACCO)}\chi_0^{(CH_2)}} + e_{C-O}\frac{n_C^{(CH_2)}\chi_2^{(ACCO)} - n_O^{(ACCO)}\chi_0^{(CH_2)}}{\chi_2^{(ACCO)}\chi_0^{(CH_2)}}$$

$$a_{ACCO-CH_2} = \overline{b_{C-C}} \frac{n_C^{(ACCO)}\chi_0^{(CH_2)} - n_C^{(CH_2)}\chi_0^{(ACCO)}}{\chi_0^{(CH_2)}\chi_0^{(ACCO)}} + \overline{b_{O-C}} \frac{n_O^{(ACCO)}\chi_0^{(CH_2)} - n_C^{(CH_2)}\chi_0^{(ACCO)}}{\chi_0^{(CH_2)}\chi_0^{(ACCO)}}$$

3) By substituting all the AIPs, the atom stoichiometry and the CI values, the GIPs obtained are 974.26 and -23.86 respectively for $a_{CH2\text{-}ACCO}$ and $a_{ACCO\text{-}CH2}$.

Table D.2: Atom Interaction Parameters (AIPs) Needed for the Calculation of GIPs Related to $a_{CH2\text{-}ACCO}$ and $a_{ACCO\text{-}CH2}$

AIPs	Values	AIPs	Values
b_{C-C}	969.1222	c_{C-O}	-15.1636
c_{C-C}	-112.3279	e_{C-O}	69.1201
e_{C-C}	-111.7174	$\overline{b_{C-C}}$	-161.3158
b_{C-O}	-1163.8140	$\overline{b_{O-C}}$	32.5118

Calculations of the GIPs between CH_2 and ACNH

1) The stoichiometry and the CI values needed is given in Table D.3

Table D.3: Atom Stoichiometry and CI Values for CH_2-ACNH

Group	n_C	N_N	${}^{v}\chi^{0}$	${}^{v}\chi^{1}$	${}^{v}\chi^{2}$
CH_2	1	0	0.7071	0.0000	0.0000
ACNH	1	1	1.0000	0.5000	0.0000

2) Using Equations (2.37)-(2.42) and the atom interaction parameters (AIPs) presented in Table D.4, the group interaction parameters $a_{CH2\text{-}ACNH}$ and $a_{ACNH\text{-}CH2}$ are calculated as follows:

$$a_{CH_2-ACNH} = b_{C-C} \frac{n_C^{(CH_2)}\chi_0^{(ACNH)} - n_C^{(ACNH)}\chi_0^{(CH_2)}}{\chi_0^{(ACNH)}\chi_0^{(CH_2)}} + c_{C-C} \frac{n_C^{(CH_2)}\chi_1^{(ACNH)} - n_C^{(ACNH)}\chi_0^{(CH_2)}}{\chi_1^{(ACNH)}\chi_0^{(CH_2)}} +$$

$$b_{C-N} \frac{n_C^{(CH_2)}\chi_0^{(ACNH)} - n_N^{(ACNH)}\chi_0^{(CH_2)}}{\chi_0^{(ACNH)}\chi_0^{(CH_2)}} + c_{C-N} \frac{n_C^{(CH_2)}\chi_1^{(ACNH)} - n_N^{(ACNH)}\chi_0^{(CH_2)}}{\chi_1^{(ACNH)}\chi_0^{(CH_2)}}$$

$$a_{ACNH-CH_2} = \overline{b_{C-C}} \frac{n_C^{(ACNH)}\chi_0^{(CH_2)} - n_C^{(CH_2)}\chi_0^{(ACNH)}}{\chi_0^{(CH_2)}\chi_0^{(ACNH)}} + \overline{b_{N-C}} \frac{n_N^{(ACNH)}\chi_0^{(CH_2)} - n_C^{(CH_2)}\chi_0^{(ACNH)}}{\chi_0^{(CH_2)}\chi_0^{(ACNH)}}$$

3) By substituting all the AIPs, the atom stoichiometry and the CI values, the GIPs obtained are 1885.35 and 50.89 respectively for $a_{CH2\text{-}ACNH}$ and $a_{ACNH\text{-}CH2}$.

Table D.4: Atom Interaction Parameters (AIPs) Needed for the Calculation of GIPs Related to $a_{CH2\text{-}ACNH}$ and $a_{ACNH\text{-}CH2}$

AIPs	Values	AIPs	Values
b_{C-C}	969.1222	c_{C-N}	-652.4220
c_{C-C}	-112.3279	$\overline{b_{C-C}}$	-161.3158
b_{C-N}	-1191.5363	$\overline{b_{N-C}}$	38.4533

APPENDIX E

LIST OF CONFERENCE PRESENTATIONS AND PUBLICATIONS

This appendix list the conference presentations and publications which are related to this PhD project. Part of the contributions of this project have been presented in several conferences as listed in E.1. Furthermore, some results of the project have been published in a international journal and two conference proceedings as listed in E.2.

E.1 Conference Presentations

1. **Mustaffa A. A.**, Kontogeorgis G., and Gani R., "Analysis and Application of GC^{Plus} Models for Property Prediction of Organic Chemical Systems", Type: Poster, Presented at: 12th International Conference on Properties and Phase Equilibria for Product and Process Design, 16-21 May 2010, Suzhou, Jiangsu.

2. **Mustaffa A. A.**, Diaz-Tovar C. A., Hukkerikar A., Quaglia A., Sin G., Kontogeorgis G., Sarup B., and Gani R., "Building a Multilevel Modeling Network for Lipid Processing Systems", Type: Oral, Presented at: 4th International Conference on Modeling, Simulation and Applied Optimization (ICMSAO), 19-21 April 2011, Kuala Lumpur, Malaysia.

3. **Mustaffa A. A.**, Kontogeorgis G., and Gani R., "Application of the UNIFAC-CI Model for Phase Equilibria Predictions of Organic Chemical System", Type: Poster, Presented at: Industrial Use of Molecular Thermodynamics (InMoTher 2012), 19-20 March 2012, Lyon, France.

4. **Mustaffa A. A.**, Kontogeorgis G., Kang J. W., Gani R., "Development and Analysis of Original UNIFAC-CI and Modified UNIFAC-CI Models for Prediction of VLE and SLE Systems", Type: Oral, Presented at: 18th Symposium on Thermo-physical Properties, 26-29 June 2012, Boulder, Colorado USA.

5. Sansonetti S., Conte E., **Mustaffa A. A.**, Crafts P. A., and Gani R., "Verification and Prediction of Solubilities of Active (Pharmaceutical) Ingredients In Solvents and Solvent Mixtures", Type: Oral, Presented by Sascha Sansonetti at 2011 AIChE Annual Meeting, Minneapolis, Minnesota USA.

6. Diaz-Tovar, C. A., **Mustaffa A. A.**, Kontogeorgis G., Gani R. and Sarup B., "Lipid Processing Technology: Shifting From Waste Streams to High-Value Commercial by-ProductsType: Oral, Presented by Rafiqul Gani at 2011 AIChE Annual Meeting, Minneapolis, Minnesota USA.

7. **Mustaffa A. A.**, Gani R., and Kang J. W., "Development and Analysis of Original UNIFAC-CI and Modified (Dortmund) UNIFAC-CI Models for Predictions of VLE and SLE Systems", Type: Poster, Presented by Amol Hukkerikar at 2011 AIChE Annual Meeting, Minneapolis, Minnesota USA.

8. Diaz-Tovar C. A., **Mustaffa A. A.**, Hukkerikar A., Quaglia A., Sin G., Kontogeorgis G., Sarup B., and Gani R., "Lipid Processing Technology: Building a Multilevel Modeling Network" Type: Oral, Presented by Carlos Axel Diaz-Tovar at the 21st European Symposium on Computer Aided Process Engineering (ESCAPE), Chalkidiki, Greece.

9. Dada E. A., **Mustaffa A. A.**, and Gani R., "Production of Dialkly Carbonates Via Reactive-Extractive and Pressure-Swing Distillations Using UNIFAC-CI VLE Model Predictions", Type: Oral, Presented by Emmanuel A. Dada at 2012 AIChE Annual Meeting, Pittsburgh, Pennsylvania, USA.

E.2 Publications

1. **Mustaffa A. A.**, Kontogeorgis G. M., and Gani R., "Analysis and Application of GC^{Plus} Models for Property Prediction of Organic Chemical Systems", *Fluid Phase Equilibr.,* 302, 274-283 (2011).

2. Diaz-Tovar C. A., **Mustaffa A. A.**, Hukkerikar A., Quaglia A., Sin G., Kontogeorgis G., Sarup B., and Gani R., "Lipid Processing Technology: Building a Multilevel Modeling Network", Computer Aided Chemical Engineering, 29, 256–260 (2011), 21st European Symposium on Computer Aided Process Engineering.

3. **Mustaffa A. A.**, Diaz-Tovar C. A., Hukkerikar A., Quaglia A., Sin G., Kontogeorgis G., Sarup B., and Gani R., "Building a Multilevel Modeling Network for Lipid Processing Systems", IEEE Xplore, Pages 237-243, 4th International Conference on Modeling, Simulation and Applied Optimization (ICMSAO).

This PhD-project was carried out at CAPEC, the Computer Aided Product-Process Engineering Center.
CAPEC is committed to research, to work in close collaboration with industry and to participate in educational activities. The research objectives of CAPEC are to develop computer-aided systems for product/process simulation, design, analysis and control/operation for chemical, petrochemical, pharmaceutical and biochemical industries. The dissemination of the research results of CAPEC is carried out in terms of computational tools, technology and application. Under computational tools, CAPEC is involved with mathematical models, numerical solvers, process/operation mathematical models, numerical solvers, process simulators, process/product synthesis/design toolbox, control toolbox, databases and many more. Under technology, CAPEC is involved with development of methodologies for synthesis/design of processes and products, analysis, control and operation of processes, strategies for modelling and simulation, solvent and chemical selection and design, pollution prevention and many more. Under application, CAPEC is actively involved with developing industrial case studies, tutorial case studies for education and training, technology transfer studies together with industrial companies, consulting and many more.

Further information about CAPEC can be found at www.capec.kt.dtu.dk.

Computer Aided Process Engineering Center
Department of Chemical and Biochemical Engineering
Technical University of Denmark
Søltofts Plads, Building 229
DK-2800 Kgs. Lyngby
Denmark

Phone: +45 4525 2800
Fax: +45 4525 4588
Web: www.capec.kt.dtu.dk

ISBN : 978-87-92481-95-5