



HAL
open science

Modeling solvent selection for biorefinery application

Gabrielly Miyazaki

► **To cite this version:**

Gabrielly Miyazaki. Modeling solvent selection for biorefinery application. Thermics [physics.class-ph]. Université Paris sciences et lettres, 2023. English. NNT : 2023UPSLM067 . tel-04573560

HAL Id: tel-04573560

<https://pastel.hal.science/tel-04573560v1>

Submitted on 13 May 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



THÈSE DE DOCTORAT

DE L'UNIVERSITÉ PSL

Préparée à Mines Paris

Modélisation pour la sélection de solvants pour des applications bioraffinerie

Modeling solvent selection for biorefinery application

Soutenue par

Gabrielly MIYAZAKI

Le 20 décembre 2023

Ecole doctorale n° 621

**Ingénierie des Systèmes,
Matériaux, Mécanique,
Énergétique**

Spécialité

**Energétique et génie des
procédés**



Composition du jury :

Patrice PARICAUD Professeur, ENSTA Paris	<i>Président</i>
Nicolas FERRÉ Professeur, Université Aix-Marseille	<i>Rapporteur</i>
Christelle GOUTAUDIER Professeure, Université Lyon 1	<i>Rapporteur</i>
Nicolas FERRANDO Docteur, IFP Energies Nouvelles	<i>Examineur</i>
Céline HOURIEZ Chargée de recherche, Mines Paris - PSL	<i>Examineur</i>
Elise EL AHMAR Maître assistant, HDR, Mines Paris - PSL	<i>Examineur</i>
Carlo ADAMO Professeur, Chimie ParisTech - PSL	<i>Directeur de thèse</i>
Christophe COQUELET Professeur, IMT Mines Albi	<i>Directeur de thèse</i>

Contents

Contents	1
List of figures.....	5
List of tables	10
Abstract	13
Introduction.....	14
Résumé	15
Environmental context and issues	18
Chapter 1.....	26
Résumé.....	27
1.1. Biorefinery.....	30
1.1.1. Lignocellulosic biomass	33
1.1.2. Pre-treatment Lignocellulosic biomass.....	37
1.1.3. Conversion technologies of lignocellulosic biomass	38
1.1.4. Value-added chemicals: platform molecules.....	41
1.2. Separation and purification process	46
1.2.1. Distillation.....	47
1.2.2. Liquid-liquid extraction.....	47
1.2.3. Other separation processes	48
1.3. Solvent selection.....	48
1.3.1. Physico-chemical properties.....	49
1.3.2. Economic and environmental aspects.....	50
1.3.3. Efficacy	50
1.4. Deep eutectic solvents	51
1.4.1. Physico-chemicals properties.....	53
1.4.2. Applying DES in a processes.....	54
1.5. Conclusion	56
Chapter 2.....	57
Résumé.....	58
2.1. Introduction	60
2.2. Materials and methods.....	62
2.3. Synthesis of Deep Eutectic Solvent.....	65
2.3.1. Pre-treatment.....	65

Contents

2.3.2.	DES Synthesis	65
2.4.	Mixture preparation	66
2.5.	Characterization of Deep Eutectic Solvent and mixtures	66
2.5.1.	Density	66
2.5.2.	Viscosity	67
2.5.3.	Refractive index	69
2.5.4.	Vapor-liquid Equilibrium for pure compounds	70
2.5.5.	Vapor-liquid equilibrium for binary mixtures	73
2.5.6.	Apparent Henry's law constant and infinite dilution coefficient activity	75
2.6.	Uncertainty	77
2.7.	Conclusion	79
	Chapter 3	80
	Résumé	81
3.1.	Introduction	83
3.2.	Thermodynamic fundamentals of phase equilibrium	84
3.3.	Fugacity and Activity	85
3.4.	Calculation of thermodynamic properties	87
3.4.1.	The γ - ϕ approach	88
3.4.2.	The ϕ - ϕ approach	89
3.4.3.	PT-Flash algorithm	89
3.5.	Activity coefficient	91
3.6.	Activity coefficient models	92
3.6.1.	Scatchard-Hildebrand model	96
3.6.2.	WILSON	97
3.6.3.	UNIFAC	98
3.7.	COSMO-based model	99
3.7.1.	Sigma-profile	100
3.7.2.	COSMO-SAC – activity coefficient prediction	102
3.8.	Conclusion	103
	Chapter 4	105
	Résumé	106
4.1.	Introduction	109
4.2.	Theoretical Background	110
4.2.1.	Slater determinants	113
4.3.	Density Functional Theory	113
4.3.1.	Hohenberg-Kohn Theorems	115
4.3.2.	Kohn-Sham method	116

4.3.3.	The atomic basis set	118
4.4.	The Exchange-Correlation energy	121
4.4.1.	Local density approximation	123
4.4.2.	Generalized gradient approximation.....	123
4.4.3.	Hybrid functionals	124
4.4.4.	Range-separated hybrid functionals	126
4.5.	Modeling solvent effect.....	127
4.5.1.	Explicit solvation models	127
4.5.2.	Implicit Solvation models.....	127
4.5.3.	The polarizable continuum model	133
4.3.5.	Calculation using COSMO model	134
4.6.	Conclusion.....	135
Chapter 5.....		137
Résumé		138
5.1.	Pure properties	140
5.1.1.	Density and Viscosity.....	140
5.1.2.	Vapor pressure	142
5.2.	Vapor-liquid Equilibrium data.....	146
5.2.1.	Data treatment	146
5.2.2.	Results and discussion	146
5.3.	COSMO-SAC predictions	150
5.4.	Infinite Dilution Activity Coefficient (IDAC) of DES	150
5.5.	Conclusion.....	152
Chapter 6.....		154
Résumé		155
6.1.	Introduction	157
6.2.	PSL sigma-profile	158
6.2.1.	Setting the basis set.....	158
6.2.2.	Setting the level of theory.....	159
6.3.	IDAC predictions in conventional solvents	161
6.4.	Phase equilibrium predictions in conventional solvents	165
Chapter 7.....		175
Résumé		176
7.1.	Geometry optimization	179
7.2.	Surface charge density	186
7.3.	Approaches to generate the sigma-profile of DES	189
7.4.	Prediction of infinite dilution activity coefficient in DES.....	190

Contents

7.4.1.	Investigation of sigma-profiles on COSMO-SAC predictions: AB versus A+B approaches	192
7.4.2.	Investigation of combinatorial contribution.....	195
7.4.3.	Investigation of dispersive contribution.....	198
7.5.	COSMO-SAC-DES model.....	205
7.5.1.	Dispersive, eccentricity, and solubility contribution.....	205
7.5.2.	Investigation of the DES contribution.....	207
7.5.3.	Vapor-liquid equilibrium predictions.....	209
7.6.	Solvents selection	215
7.7.	Conclusion.....	217
Chapter 8.....		218
8.1.	Conclusion (Français).....	219
8.2.	Perspective (Français).....	222
Conclusion and perspectives (English).....		224
8.3.	Conclusion (English)	224
8.4.	Perspectives (English).....	226
References.....		228
A.1.	Liquid-liquid equilibrium.....	239
A.1.1.	Cox Herington plot for coexistence data	239

List of figures

Figure 1 – Long-term historic energy transition: Primary energy consumption by source (data from Our World in Data ² and prediction from International Energy Agency ³).	18
Figure 2 – Long-term trend of global carbon emissions and temperature. (Data from Our World in Data ² and prediction from International Energy Agency ³).	19
Figure 3 – Global greenhouse gas emissions by sector for the year 2016 (total of 49.4 billion tones CO ₂ eq). (Data from Our World in Data ⁵).	21
Figure 1.1 – Schematic diagram of a promising closed-loop sustainable waste reutilization to develop value-added green products from biomass.	30
Figure 1.2 – Cellulose chemical structure.....	34
Figure. 1.3 – Hemicelluloses monosaccharide.....	35
Figure 1.4 – Lignols.....	35
Figure 1.5 – Structure of lignocellulosic biomass. ³⁴	36
Figure 1.6 – Biomass pyrolysis scheme. ⁶⁰	39
Figure. 1.7 – Dehydration of fructose and hemicellulose.....	42
Figure 1.8 - 5-Hydroxymethylfurfural (HMF) platform molecules. ⁶⁶	43
Figure 1.9 – Furfural-derived platform molecules and biofuels. ⁶⁷	44
Figure 1.10. Examples of thermal separation processes (higher energy use) and nonthermal separation processes (lower energy use). ⁸⁸	46
Figure 1.11 -. Number of publications with DES as key word per year.	51
Figure 1.12 – Choline Chloride:Urea (1:2) phase diagram.....	52
Figure 2.1. Frutaline, Glycaline and Ethalyne.....	66
Figure 2.2 – Digital density meter: Anton Paar DSA 5000M (modified from Anton Paar ⁹⁷).	67
Figure 2.3 – Rolling-ball principle showing the main forces acting on the descending ball. F _G : Effective portion of gravitational force, F _B : Effective portion of buoyancy force, and F _V : Viscous force. ⁹⁷	68
Figure 2.4 – Flow diagram of the synthetic apparatus: DAU: data acquisition unit; DS: degassed solution; DT: displacement transducer; EC: equilibrium cell; LB: liquid bath; PN: pressurized nitrogen; PP: Platinum probe; PT: pressure transducer; S: stirrer; SD: stirring device; TR: thermal regulator; Vi: valve; VP: vacuum pump; VVCS: variable volume cylinder (modified from Coquelet et al. (2020) ¹⁰⁴	70
Figure 2.5 – Low-pressure numerical reference standard model 24610 from Desgranges & Huot.	71
Figure 2.5 – Flowchart to predict VLE using the static synthetic method.	75
Figure 2.6 – Flow diagram of the equipment: BF, bubble flow meter; C, gas chromatograph; D, dilutor; d.a.s., data acquisition system; He, helium cylinder; E1, E2, heat exchangers; FE, flow meter electronic; FR, flow regulator; L, sampling loop; LB, liquid bath; O, O-ring; PP, platinum resistance thermometer probe; S, saturator; SI, solute injector; Sp, septum; SV, sampling valve; TR, temperature regulator; VSS, variable speed stirrer (from Coquelet et al. ¹⁰⁷)	76
Figure 3.1 – Approaches to modeling DES (modified from Alkhatib et al. ¹⁰⁸).....	83

Figure 3.2 – Flash tank.....	89
Figure 3.3 – Scheme of successive substitution algorithm for the PT flash problem for phi-phi approach.	91
Figure 3.4 – Several publications about the COSMO model, including articles, conference papers, book chapters, reviews, and conference reviews covering the 1999 – 2022 periods. Green: COSMO-SAC and yellow: COSMO-RS.	96
Figure 3.5 – COSMO surface of (a) Water, (b) Furan, and (c) n-butanol color-coded by the screening charge density.	99
Figure 3.6 – Sigma-profile of Furan for different exchange-correlation Functionals using 6-311G(2df,p) as a basis set in comparison with the VT 2005 Sigma Profile Database.	101
Figure 3.7 – COSMO-SAC steps	103
Figure 4.1 – A molecular coordinate system: i and j = electrons; α and β = nuclei.	111
Figure 4.2 – (a) Schematic representation of electrons denoted by black circles with arrows illustrating Coulomb interactions between some electrons. (b) Illustration of a mean-field DFT approach, where DFT quasiparticles are depicted as white circles moving within a background field generated by the electric field of other particles.	114
Figure 4.3 – Graphical representation of the Jacob’s ladder of chemical accuracy. For each rung in the ladder an indication of the accuracy is given according to the heat map shown. Adapted from reference ¹³⁷	122
Fig. 4.4 – Representations of (a) Explicit and (b) Implicit solvation models.	127
Figure 4.5 – Thermodynamic cycle depicting the solvation free energy decomposition into contributions.	129
Figure 4.6 – Solute Cavity. Blue region: dielectric medium (ϵ_{outer}). White region: solute region ($\epsilon_{\text{inner}}=1$). Yellow: solute-solvent interface (Γ).	130
Figure 4.7 – 2D example of a SAS in green and a SES solute cavity in yellow, with Solvent Probe in orange, and vdW surfaces and sphere in black and gray, respectively.	132
Figure 4.8 – COSMO surface of (a) Water, (b) Furan, and (c) n-butanol color-coded by the screening charge density.	134
Figure 4.9 – Steps of the DFT calculation.....	135
Figure 5.1 – Density in function of the temperature of ChCl:EG (left) and ChCl:Gly (right).	141
Figure 5.2 – Viscosity in function of the temperature of ChCl:EG (left) and ChCl:Gly (right).	142
Figure 5.3 – Vapor pressure of Ethaline. Black symbols: experiments, green symbols: regressed, dash lines area: error range.	144
Figure 5.4 – Vapor pressure of Glycaline. Black symbols: experiments, green symbols: regressed, dash line area: error range.	145
Figure 5.5 – VLE isotherms at (●)303 K, (■) 313 K, (▲) 323 K, and (◆) 333 K. for Ethaline (1) and Ethanol (2) mixture. Symbols: Experimental data. Black lines: Wilson. Dash lines: NRTL.	147
Figure 5.6 – VLE isotherms at (●)303 K, (■) 313 K, (▲) 323 K, and (◆) 333 K of Ethaline (1) and Isopropanol (2). Symbols: Experimental data. Black lines: Wilson. Dash lines: NRTL.	148
Figure 5.7 – VLE isotherms at (●)303 K, (■) 313 K, (▲) 323 K of Glycaline (1) and Ethanol (2). Symbols: Experimental data. Black lines: Wilson. Dash lines: NRTL.	149
Figure 5.8 – VLE isotherms at (●)303 K, (■) 313 K, (▲) 323 K of Glycaline (1) and Isopropanol (2). Symbols: Experimental data. Black lines: Wilson. Dash lines: NRTL.	149

Figure 5.9 – VLE isotherms at 303 K of Glycaline (1) and Isopropanol (2). Symbols: Experimental data. Black lines: Wilson. Dash lines: NRTL and green lines: COSMO-SAC-dps.....	150
Figure 5.9 – Peak area as function of time at different flowrates (a: 22.5, b: 34.5 and c: 47.2 mL/min) and at 323 K for Ethaline (1) and Ethanol (2) mixture. Symbols: Experimental data. Black line: linear regression.	151
Figure 5.10 – Peak area as function of time at different flowrates (a: 47.4 and b: 61.2 mL/min) and at 323 K for Ethaline (1) and Pentanol (2). Symbols: Experimental data. Black line: linear regression.	152
Figure 6.1 – PSL sigma-profile.	157
Figure 6.2 – Mean absolute deviation MAD for infinite dilution activity coefficients of conventional mixture combinations for COSMO-SAC-dps (a) and modified UNIFAC(Do) (b).....	162
Figure 6.3 – IDAC of Benzene in n-Alkanes using COSMO-SAC with different sigma-profiles (grey: UD, orange: PSL-Klamt and green: PSL-UFF).	164
Figure 6.4 – IDAC of Toluene in n-Alkanes using COSMO-SAC with different sigma-profiles (grey: UD, orange: PSL-Klamt and green: PSL-UFF).	164
Figure 6.5 – Experimental <i>vs.</i> predicted activity coefficient of aromatics in <i>n</i> -alkanes using (▲) NRTL and (◆) UNIFAC Do. The dash line represents a margin of error of 15%.	167
Figure 6.6 – Experimental <i>vs.</i> predicted activity coefficient of aromatics in <i>n</i> -alkanes using (■) Scatchard-Hildebrand (green: no combinatorial term; deep blue: with combinatorial term of Staverman-Guggenheim). The dash line represents a margin of error of 15%.	168
Figure 6.7 – Experimental <i>vs.</i> predicted activity coefficient of aromatics in <i>n</i> -alkanes using (●) COSMO-SAC 2010 with different sigma-profiles database (grey: UD; orange: PSL (Klamt radii) and green: PSL (UFF radii)). The dash line represents a margin of error of 15%.	169
Figure 6.8 – Experimental <i>vs.</i> predicted activity coefficient of aromatics in <i>n</i> -alkanes using different models.●: for COSMO-SAC 2010 using sigma-profiles database (grey: UD; orange: PSL (Klamt radii) and green: PSL (UFF radii)); ▲: for NRTL and ◆: for UNIFAC(Do) and ■: for Scatchard-Hildebrand (green: no combinatorial term; deep blue: with combinatorial term of Staverman-Guggenheim). The dash line represents a margin of error of 15%.	170
Figure 6.9 – Evolution of activity coefficient in function of the mole fraction for different thermodynamic models: grey lines: COSMO-SAC (UD); orange lines: COSMO-SAC (PSL-Klamt); green lines: COSMO-SAC (PSL-UFF); blue lines: NRTL; light blue lines: UNIFAC(Do); light green lines: SH and(deep blue lines: SH+SG. ●:experimental values.	172
Figure 7.1 – A potential energy curve for the covalent bond in a H ₂ molecule. The distance <i>r</i> is the distance between the nuclei of the two H atoms ¹⁸⁰	180
Figure 7.2 – Potential energy surface of the ozone molecule varying the O-O distance (from 1.0 to 1.4 Å) and the O-O-O angle (from 55° to 125°) ¹⁸¹	180
Figure 7.3 – 1-butyl-3-methylimidazolium (BMIM ⁺) conformers and their five favorable regions for chloride ion.	181
Figure 7.4 – Possible geometries for 1-butyl-3-methylimidazolium chloride:glycerol (ration 2:1) - BMIMCl:Gly.....	185
Figure 7.5 – Surface charge of the [1-butyl-3-methylimidazolium] ⁺ ion.	187
Figure 7.6 – Surface charge of the two conformers of 1-butyl-3-methylimidazolium and the five positions of the chloride ion.	188
Figure 7.7 – Surface charge of the BMIMCl: Gly (ratio 2:1) (geometry 4).	188

Figure 7.8 – Sigma-profile of the BMIMCl: Gly (ratio 2:1). (a) AB approach and (b) A+B approach. Black lines: non-hydrogen bond profile, and grey lines: hydrogen bond profile.	190
Figure 7.9 – Experimental <i>vs.</i> predicted IDAC of (a) <i>n</i> -alkanes and (b) <i>n</i> -alkenes in BMIMCl: Gly (ratio 2:1). Solid symbols represent results using sigma-profile AB, and hollow symbols sigma-profile A+B. Dash lines represent 15% of the error from bisector.	193
Figure 7.10 – Experimental <i>vs.</i> predicted IDAC of (a) cyclo-hydrocarbons and (b) aromatics hydrocarbons in BMIMCl: Gly (ratio 2:1). Solid symbols represent results using sigma-profile AB, and hollow symbols sigma-profile A+B. Dash lines represent 15% of the error from bisector.	194
Figure 7.11 – Experimental <i>vs.</i> predicted IDAC of alcohols in BMIMCl: Gly (ratio 2:1). Solid symbols represent results using sigma-profile AB, and hollow symbols sigma-profile A+B. Dash lines represent 15% of the error from bisector.	194
Figure 7.12 – Experimental <i>vs.</i> predicted IDAC of alkanes in DES. Black symbols represent the results of COSMO-SAC-dps with combinatorial term optimized, and blue symbols COSMO-SAC-dps(DES0). Dash lines represent 15% of the error from bisector.	200
Figure 7.13 – Experimental <i>vs.</i> predicted IDAC of alcohols in DES. Black symbols represent the results of COSMO-SAC-dps with combinatorial term optimized, and blue symbols COSMO-SAC-dps(DES0). Dash lines represent 15% of the error from bisector.	201
Figure 7.14 – Experimental <i>vs.</i> predicted IDAC of aromatics in DES. Black symbols represent the results of COSMO-SAC-dps with combinatorial term optimized, and blue symbols COSMO-SAC-dps(DES0). Dash lines represent 15% of the error from bisector.	202
Figure 7.15 – Experimental <i>vs.</i> IDAC of cyclohydrocarbons in DES. Black symbols represent the results of COSMO-SAC-dps with combinatorial term optimized, and blue symbols COSMO-SAC-dps(DES0). Dash lines represent 15% of the error from bisector.	203
Figure 7.16 – Experimental <i>vs.</i> IDAC of ketones in DES. Black symbols represent the results of COSMO-SAC-dps with combinatorial term optimized, and blue symbols COSMO-SAC-dps(DES0). Dash lines represent 15% of the error from bisector.	204
Figure 7.17 – Left: Experimental <i>vs.</i> predicted IDAC of conventional solute in BMIMCl:Gly (ratio 2:1). Dash lines represent 15% of the error. Right: zoomed-in view.	208
Figure 7.18 – Experimental <i>vs.</i> predicted IDAC of conventional solute in BMIMCl:Gly (ratio 2:1). Dash lines represent 15% of the error from bisector.	209
Figure 7.19 – VLE isotherms at (●)303 K, (■) 313 K, (▲) 323 K of Glycaline (1) and Ethanol (2). Symbols: Experimental data. Black lines: Wilson. Dash lines: NRTL. Orange: COSMO-SAC-DES: parameters optimized with IDAC data (dash orange lines) and optimized with all gammas from VLE data predicted with NRTL (orange lines).	210
Figure 7.20 – VLE isotherms at (●)303 K, (■) 313 K, (▲) 323 K of Glycaline (1) and Isopropanol (2). Symbols: Experimental data. Black lines: Wilson. Dash lines: NRTL. Orange: COSMO-SAC-DES: parameters optimized with IDAC data (dash orange lines) and optimized with all gammas from VLE data predicted with NRTL (orange lines).	211
Figure 7.21 – VLE isotherms at (●)303 K, (■) 313 K, (▲) 323 K of Glycaline (1) and Ethanol (2). Symbols: Experimental data. Black lines: Wilson. Dash lines: NRTL. Blue: 2-parameters COSMO-SAC-DES.	212
Figure 7.22 – VLE isotherms at (●)303 K, (■) 313 K, (▲) 323 K of Glycaline (1) and Isopropanol (2). Symbols: Experimental data. Black lines: Wilson. Dash lines: NRTL. Blue: 2-parameters COSMO-SAC-DES.	213
Figure 7.23 –Margules 2-parameters optimization: Illustrating the evolution of prediction error with increasing optimization points (optimal 10 points).	214

List of figures

Fig. 7.24 – Solvation power predicted and experimental of BMIMCl:Gly for conventional solutes.....	216
Figure 7.25 - Solvation power values (S_p) of several DES for conventional solutes.	216
Figure A1 – (a) SLE and (b) activity coefficient of Phenol (1)/n-Hexane (2) at atmospheric pressure. (a) Black circle: experiments values (this work), Black dot-line: NRTL, Orange dot-line: COSMO-SAC. (b) black: NRTL and orange: COSMO-SAC (2010) and PSL sigma-profile (Klamt radii).	241
Figure A2 – LLE of Phenol (1)/n-Hexane at atmospheric pressure. Black circle: experiments values (this work), white circle: critical point, back line: Cox-Herington, black dot-line: NRTL, Orange dot-line: COSMO-SAC.....	241

List of tables

Table 1.1 – Comparison of biomass materials to produce value-added products. ²³	32
Table 1.2 – The major chemical composition of various LC biomasses ³⁴	34
Table 1.3 – Summary of lignocellulosic pretreatment methods. ³⁵	37
Table 1.4 – Conversion technologies. ⁶⁰	38
Table 1.5 – Catalytic dehydration of hexoses to HMF in various DESs.....	45
Table 1.6 – Catalytic dehydration of pentoses to FF in various DESs.....	45
Table 2.1 – Chemical used for experimental work.	63
Table 2.2 – Pure properties measurement.	64
Table 2.3 – Pseudo binary mixture properties measurement.	64
Table. 2.4 – DES nomenclature.	65
Table 3.1 – Parameters used in the sigma-profile calculation.	100
Table 4.1 – Fundamental constants in atomic unit. ¹²⁵	111
Table 4.2 – Nomenclature for basis sets.	121
Table 4.3 – meta-GGA functionals.....	124
Table 4.4 – Atomic radius in angstrom.	132
Table 5.1 – DES properties.....	141
Table 5.2 – Vapor pressure Ethaline.	143
Table 5.3 -Vapor pressure Glycaline.....	143
Table 5.4 – Hildebrand solubility parameter at 298 K.	145
Table 5.5 – IDAC of DES and Ethanol for different flowrate at 323 K.....	152
Table 6.1 – Pople's nomenclature for basis sets.....	159
Table 6.2 – MAD of IDAC predictions of furan in toluene.....	159
Table 6.3 – IDAC of furan in toluene, experimental and predicted with COSMO-SAC-dps ¹²³	160
Table 6.4 – IDAC of furan in cyclohexane, experimental and predicted with COSMO-SAC-dps ¹²³	161
Table 6.5 – Main family structure studied.....	161
Table 6.6 – Literature data of aromatic and n-alkane binary system.....	165
Table 6.7 – Obtained NRTL parameters.	166
Table 6.8. AAE% values of activity coefficient for different thermodynamic models.	171
Table 7.1 – Optimized structures of 1-butyl-3-methylimidazolium chloride, calculated at PBE0/6-311G(d,f) level.	181
Table 7.2 – Interaction energy in kJ/mol.....	183
Table 7.3 – Difference between 6-31G(d) and 6-31++G(d,f) in terms of interaction energy in kJ/mol.	184

List of tables

Table 7.4 – Interaction energy in kJ/mol.....	186
Table 7.5 – Cavity volume and surface area of BMIMCl:Cly (2:1).....	190
Table 7.6 – Bibliographic references of IDAC data in DES.....	191
Table 7.7 – ARD of IDAC of conventional solute in [BMIMCl + Glycerol] and its parameter a_{DES0} ..	192
Table 7.8 – ARD of IDAC of conventional solute in DES and its parameter $p0$ and $q0$	197
Table 7.9 – ARD of IDAC of conventional solutes in different DESs and its parameter a_{DES0}	199
Table 7.10 – Constants for pure component.....	206
Table 7.11 – ARD of IDAC in DES using COSMO-SAC-DES models.	207
Table B1 – NRTL parameters for $\alpha_{ij0} = 0.3$	242

Résumé

Dans le contexte des bioraffineries, le choix d'un solvant approprié est crucial pour des processus de séparation durables et économiquement viables. Une approche globale intégrant des critères tels que l'Analyse du Cycle de Vie, l'analyse de la toxicité, la régénération énergétiquement efficace du solvant, des pertes minimales de solvant et une haute sélectivité est nécessaire. Cependant, le choix devient complexe lorsqu'il s'agit des solvants eutectiques profonds (Deep Eutectic Solvents - DES) en raison du manque de données expérimentales complètes, en particulier concernant les propriétés thermodynamiques et physiques telles que l'équilibre de phases, la masse volumique, la viscosité, la capacité thermique, la conductivité thermique, la solubilité, et autres. Pour combler cette lacune, cette étude propose de générer des données expérimentales essentielles (telles que la masse volumique, la viscosité et les propriétés d'équilibre de phases) pour optimiser les modèles théoriques. De plus, ce travail propose une approche de sélection de solvant basée sur la modélisation des propriétés thermodynamiques en utilisant le modèle COSMO-SAC (COnductor-like Screening MOdel segment activity coefficient). L'étude vise à améliorer le modèle COSMO-SAC en examinant les variables computationnelles, établissant une base de données de profils sigma PSL et affinant les prédictions grâce aux contributions enthalpiques, entropiques et intermoléculaires. Malgré les défis rencontrés dans la prédiction précise des coefficients d'activité en dilution infinie (IDAC) pour les systèmes DES, une approche d'optimisation réduit considérablement les écarts, offrant ainsi une voie prometteuse pour la sélection précise du solvant dans les processus de bioraffinerie.

Abstract

In the context of biorefineries, selecting an appropriate solvent is crucial for sustainable and economically viable separation processes. A comprehensive approach integrating criteria like Life Cycle Assessment, toxicity analysis, energy-efficient solvent regeneration, minimal solvent losses, and high selectivity is required. However, the choice becomes challenging when considering Deep Eutectic Solvents (DES) due to a lack of comprehensive experimental data, particularly regarding thermodynamic and physical properties like phase equilibrium, density, viscosity, heat capacity, thermal conductivity, solubility, and more. To bridge this gap, this study proposes generating essential experimental data (such as density, viscosity and phase equilibrium properties) to optimize theoretical models. Moreover, this work proposes a solvent screening approach based on modeling thermodynamic properties using the (COnductor-like Screening MOdel segment activity coefficient (COSMO-SAC) model. The study aims to enhance the COSMO-SAC model by investigating computational variables, establishing a PSL sigma-profile database, and refining predictions through enthalpic, entropic, and intermolecular contributions. Despite encountering challenges in accurately predicting activity coefficients at infinite dilution (IDAC) for DES systems, an optimization approach significantly reduces deviations, offering a promising route for precise solvent selection in biorefinery processes.

Introduction

Environmental context and issues

Contexte environnemental et problématiques

Résumé

Dans le contexte environnemental et des problèmes liés à la consommation d'énergie et aux émissions de gaz à effet de serre, on observe clairement une transition de la consommation d'énergie à l'échelle mondiale. Cette transformation se traduit par un déplacement de la domination du pétrole et du charbon vers un intérêt croissant pour les sources d'énergie renouvelable, une évolution motivée par les préoccupations environnementales et le désir de réduire les émissions de gaz à effet de serre. La hausse alarmante des niveaux de dioxyde de carbone (CO₂) et sa corrélation historique avec les variations de température mondiale et changement climatique sont élucidées.

De multiples conséquences résultant du changement climatique, liées à l'augmentation des températures, sont abordées : incluant l'augmentation du niveau de la mer, les phénomènes météorologiques extrêmes, le dégel du permafrost, entre autres. Nous soulignons la nécessité d'entreprendre des efforts d'atténuation et d'adaptation pour réduire au minimum les impacts variés sur l'environnement et la société.

De plus, la répartition des émissions mondiales de gaz à effet de serre par secteur est présentée en mettant l'accent sur le rôle dominant du secteur de l'énergie et l'importance des innovations dans divers secteurs pour réduire les émissions de CO₂ et atteindre la neutralité carbone. L'engagement de l'Union Européenne dans la lutte contre le changement climatique et ses politiques "Fit for 55" visant à réduire les émissions de gaz à effet de serre est également présenté.¹

Cette mise en avant de la question environnementale nous amène à souligner l'importance croissante des matières premières à base de biomasse en tant qu'alternative renouvelable et durable aux ressources fossiles. Enfin, nous introduisons le concept de bioraffinerie, en mettant l'accent sur l'efficacité et la durabilité de l'utilisation des ressources, présentant cette approche comme pivot dans le traitement de la biomasse. Nous soulignons également l'importance de la sélection de solvants respectueux de l'environnement dans les processus de bioraffinerie, pour garantir la durabilité environnementale et la viabilité économique.

Dans ce contexte, l'objectif général de cette thèse est le suivant :

- Développer une méthodologie de sélection de solvants basée sur le calcul des propriétés thermodynamiques des fluides pour soutenir la sélection de solvants dans la bioraffinerie, en

mettant l'accent sur les solvants verts, en particulier les solvants eutectiques profonds (DES). Le modèle choisi à cette fin était le modèle COSMO-SAC (COnductor like Screening MOdel segment activity coefficient).

Le modèle COSMO-SAC utilise les résultats de calculs de mécanique quantique sur les molécules individuelles pour prédire les propriétés thermodynamiques des mélanges de ces molécules, par exemple le coefficient d'activité. Pour comparer les résultats obtenus et optimiser le modèle, les objectifs spécifiques de cette thèse sont, les suivants :

- Étudier le rôle des variables de calcul en mécanique quantique sur les performances du modèle COSMO-SAC, en combinant une investigation théorique et expérimentale.
- Créer une base de données (nommée « PSL sigma-profile ») de nombreux composants chimiques pour alimenter le modèle COSMO-SAC.
- Optimiser les prédictions du modèle COSMO-SAC en étudiant les contributions enthalpiques et entropiques et introduisant une contribution intermoléculaire basée sur le modèle de coefficient d'activité de Margules.
- Examiner les solvants eutectiques profonds pour l'extraction de composés furaniques (5-hydroxyméthylfurfural (HMF), furfural (FF) et acide lévulinique (LA)) produits par déshydratation du fructose et de la xylose.

Ce manuscrit, est organisé en sept chapitres :

Chapitre 1 présentera le concept de la bioraffinerie ainsi que les avantages et les inconvénients de certaines matières premières à base de biomasse. La discussion se concentre sur la biomasse lignocellulosique. Une revue de la littérature sur le prétraitement de la biomasse sera présentée. Ensuite, une discussion sur l'application industrielle des produits chimiques à valeur ajoutée (HMF et FF) issus de la pyrolyse de la biomasse sera présentée. Une vue d'ensemble des principaux procédés de séparation utilisés dans l'industrie sera discutée, en particulier, l'extraction liquide-liquide est les facteurs clés dans la sélection des solvants. La dernière partie se concentre sur les DES et leurs propriétés.

Chapitre 2 couvrera la partie expérimentale de la thèse et décrira le protocole de synthèse et de caractérisation des DES et des mélanges. Les équipements expérimentaux et les protocoles seront discutés en détail mettant l'accent sur l'acquisition de données de coefficient d'activité à dilution infinie (IDAC) par la méthode de « gas stripping » et de données d'équilibre liquide-vapeur par la méthode synthétique.

Chapitre 3 se concentrera sur les fondements de la modélisation thermodynamique commençant par la fugacité et la définition de l'activité et décrivant les deux approches pour le calcul de l'équilibre en deux phases dans un système binaire (approches γ - ϕ et ϕ - ϕ) et l'algorithme PT-Flash. Une revue sera présentée des modèles de coefficient d'activité existants développés au cours des dernières décennies. La discussion se centre en particulier sur le modèle COSMO-SAC.

Chapitre 4 présentera les principales méthodes de mécanique quantique non relativiste utilisées tout au long de cette thèse. L'équation de Schrödinger et l'approximation de Born-Oppenheimer seront introduites en premier suivies des méthodes de la théorie de la fonctionnelle de la densité (DFT). La dernière partie présentera les modèles de solvation en continu en mettant l'accent sur l'approche COSMO.

Chapitre 5 présentera une analyse expérimentale des DES en incluant leurs propriétés pures telles que la masse volumique, la viscosité et la tension de vapeur. De plus, nous avons exploré le comportement de l'équilibre liquide-vapeur des DES dans les alcools en utilisant la méthode synthétique. Les résultats obtenus ont été comparés aux résultats prédictifs obtenus avec le modèle COSMO-SAC-dps proposé par Hsieh (2014) révélant que le modèle n'est pas adapté aux DES. De plus, une étude le coefficient d'activité à dilution infinie (IDAC) des DES a été réalisée en utilisant la méthode de « gas stripping ». Cependant, il convient de noter que cette méthode n'était pas optimale pour l'étude de l'IDAC dans les fluides à haute viscosité tels que les DES.

Chapitre 6 sera consacré à l'exploration de l'impact des variables informatiques sur les performances du modèle COSMO-SAC. Un processus de benchmarking approfondi (englobant diverses méthodes théoriques et variables informatiques) a été mené pour comprendre leur influence sur la précision des prédictions à l'aide du modèle COSMO-SAC. En tant que résultat significatif de cette étude, nous avons développé la base de données de profil sigma PSL. Le modèle COSMO-SAC a été appliqué en utilisant le profil sigma PSL pour les prédictions de coefficient d'activité. De plus, ces prédictions ont été comparées à des modèles de coefficient d'activité établis, notamment NRTL, UNIFAC et Scatchard-Hildebrand.

Chapitre 7 examinera la contribution combinatoire et de dispersion au sein du modèle COSMO-SAC pour la prédiction de l'IDAC dans les DES. Les enseignements tirés de cette étude ont servi de base à l'amélioration des capacités prédictives du modèle COSMO-SAC grâce à des contributions innovantes inspirées du modèle Scatchard-Hildebrand. Celles-ci comprennent les contributions de Dispersion, d'Excentricité et de Solubilité spécifiques aux interactions DES. Ces contributions novatrices ont été intégrées dans le modèle nouvellement proposé connu sous le nom de modèle COSMO-SAC-DES, qui promet d'offrir une précision prédictive accrue des interactions soluté-solvant dans les systèmes DES.

Environmental context and issues

Energy is a key source of economic growth, and its consumption is an important thermometer of economic development. Without energy, a country fails to run all types of activities to produce goods and services like heat or climatization, transportation, etc. In 1965, Global energy consumption^{2,3} (Figure 1) was dominated by petroleum and coal as primary sources (42% and 37%, respectively). Nearly 55 years later, by 2020, petroleum was still the first primary source, followed by natural gas, renewables were third and coal dropped to fourth place. An increasing interest in renewable energy has grown from 20% to 28% of global energy consumption from 2011 to 2021 due to the need to reduce greenhouse gas emissions contributing to global climate change. However, renewable energies have not replaced fossil energies but have been added. Long-term projections for 2050 from International Energy Outlook⁴ estimated that renewable sources will represent over 30% of global energy consumption.

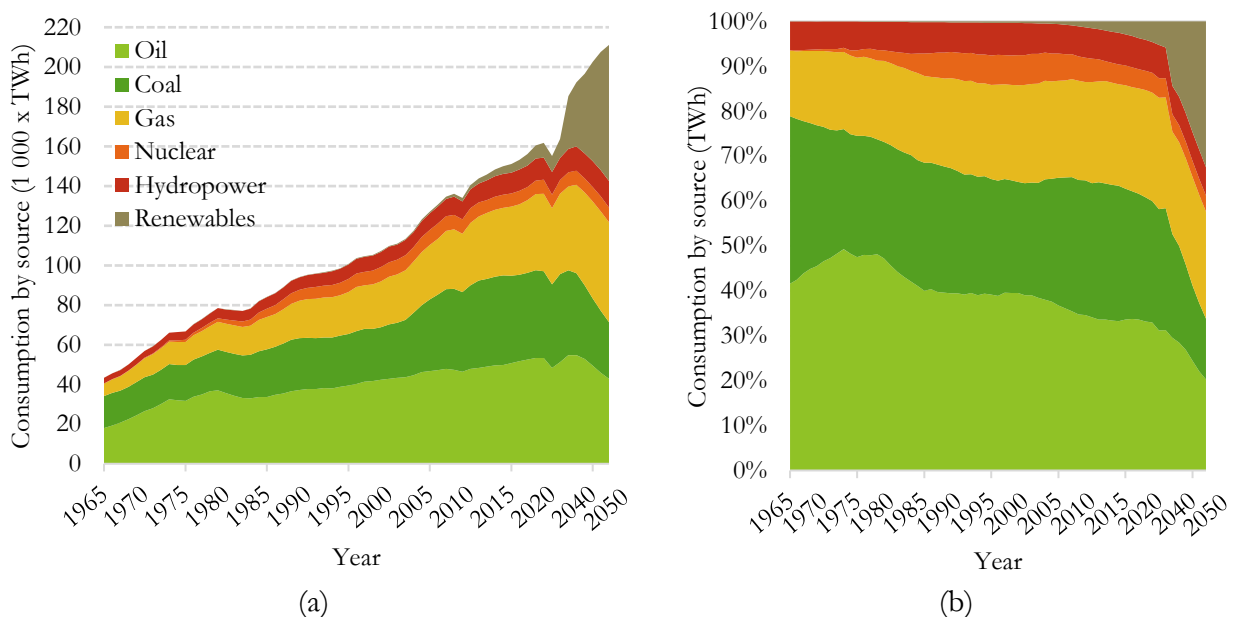


Figure 1 – Long-term historic energy transition: Primary energy consumption by source (data from Our World in Data² and prediction from International Energy Agency³).

Greenhouse gases (GHG) absorb infrared radiation, trapping heat in the atmosphere and warming the planet. The world emits around 50 billion tons of GHG each year⁵. At the global scale, the key GHG directly emitted by human activities include carbon dioxide (CO₂), methane (CH₄), nitrous oxide (N₂O), and several trace gases such as fluorine-containing halogenated substances (HFCs, PFCs, SF₆, and NF₃) (pie chart in figure 2). In which, the primary GHG emitted is the CO₂, representing 75% of the total of GHG emitted by the world in 2020. Followed by methane with 17%. Global CO₂ emissions have significantly increased and show an average decadal growth rate of 1%, 3%, and 2% for the last three

decades (Figure 2). CO₂ levels today are higher than at any point in human history and are currently at nearly 412 ppm and rising.

The last time CO₂ amounts were so high was during the interval between 3.35 and 3.15 million years ago (the mid-Pliocene) when the global surface temperature was about 2 to 3°C warmer globally on average⁶ and the sea level was about 20 meters higher⁷ than during the pre-industrial era. Consequently, the mid-Pliocene warm period serves as a valuable reference point for understanding the potential consequences of climate change, particularly in terms of sea level rise, and it helps scientists refine climate models and assess the impacts on the ecosystems. By studying this geological period, researchers can gain critical insights into how our planet’s climate system responds to higher temperatures and greenhouse gas concentrations⁸⁻¹¹.

As temperatures rise, more water from the Earth's surface (oceans, rivers, lakes, and land) evaporates into the atmosphere. Warmer air can hold more water vapor, so this leads to an increase in the amount of water vapor in the atmosphere. The increased water vapor in the atmosphere enhances the greenhouse effect, leading to further warming of the planet. This, in turn, can lead to more evaporation, perpetuating the cycle.

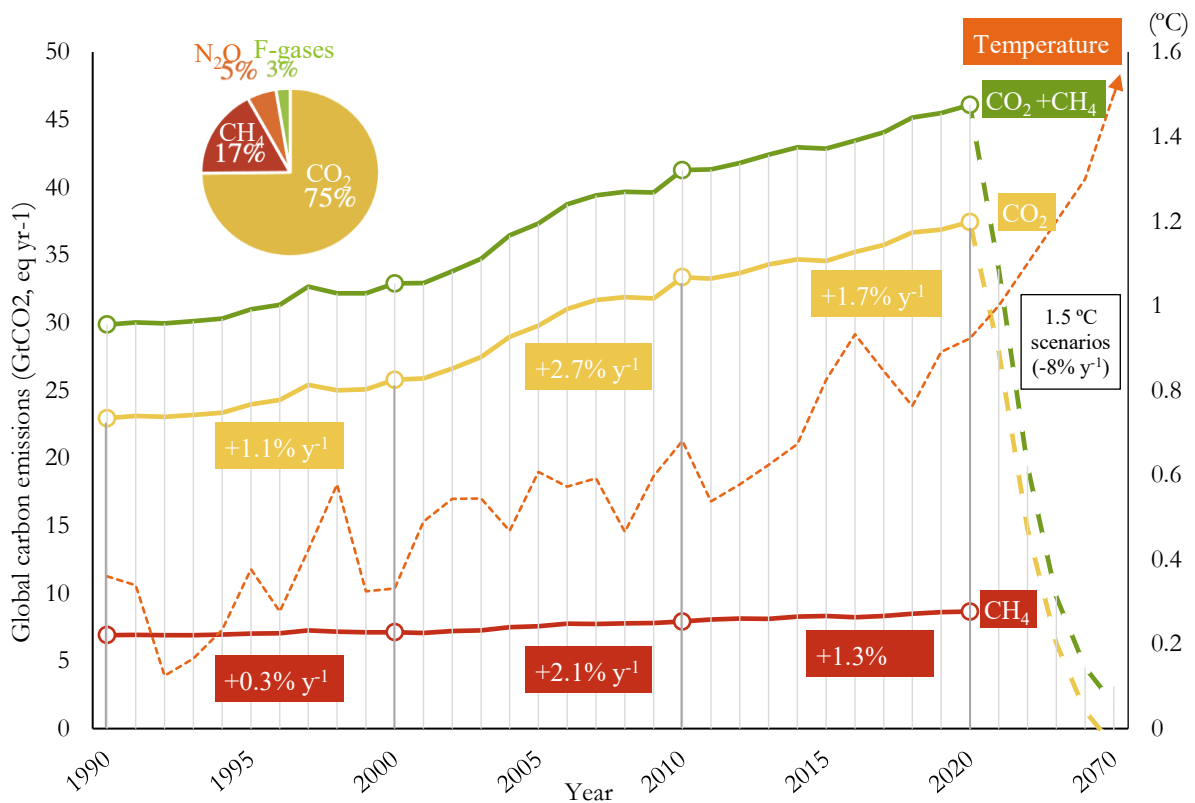


Figure 2 – Long-term trend of global carbon emissions and temperature. (Data from Our World in Data² and prediction from International Energy Agency³).

Some actions are required to control global warming and prevent severe climate change, such as the need to rapidly reduce global GHG. In 2015, 196 Parties adopted the climate change treaty, called by Paris Agreement¹². One of the Paris Agreement's goals is to limit the temperature increase to well below 2 °C, preferably to 1.5 °C, compared to pre-industrial levels. According to Net Zero Emissions by 2050 Scenario report (NZE)¹³, the carbon intensity of primary energy is projected to decrease globally by about 8% yr⁻¹ between 2020 to 2050 in modeled scenarios that limit warming to 1.5 °C.

Higher temperatures, driven by climate change, have a wide range of significant effects. They include melting polar ice, rising sea levels, more extreme weather events, ocean acidification, disrupted ecosystems, heatwaves, impacts on agriculture and water resources, increased disease spread, loss of biodiversity, economic cost, migration and displacement, national security concerns, threats to water resources social inequality and cultural heritage loss¹⁴⁻¹⁷. Extreme weather events, such as hurricanes, storms, heat waves, floods, and droughts, is one of the most notable effects of higher temperature¹⁸. Another one is the thawing of permafrost. As the global temperature rise, the permafrost in many regions is gradually warming and thawing. This process can release large amounts of greenhouse gases that have been trapped in the frozen ground for thousands of years¹⁹. Some diseases, like those transmitted by mosquitoes (e.g., malaria, dengue, Zika), can be influenced by temperature because it affects the breeding and survival of the vector (the mosquito in this case)¹⁵. Addressing climate change through mitigation and adaptation efforts is crucial to minimize these impacts on the environment and society.

To figure out how it can effectively reduce emissions and what emissions can and can't be eliminated with the current technologies. Firstly, it is important to know where the emissions are coming from. Figure 3 shows the breakdown of global emissions of 2016 by sector in one diagram of three levels. Global GHG emissions can be roughly traced back to four categories (first level of the diagram in figure 3): energy, agriculture, industry, and waste. Each one contributes to CO₂ emissions in 73.2%, 18.4%, 5.2%, and 3.2%, respectively⁵.

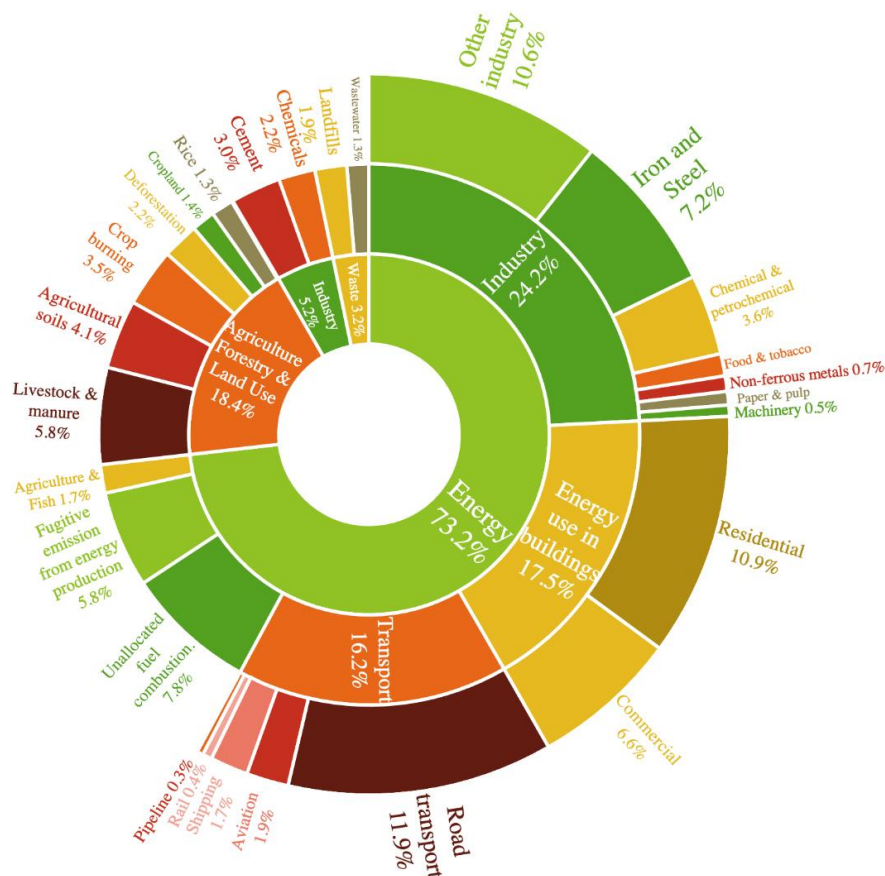


Figure 3 – Global greenhouse gas emissions by sector for the year 2016 (total of 49.4 billion tones CO₂ eq). (Data from Our World in Data⁵).

The energy sector is responsible for emitting around three-quarters of GHG today due to the use of fossil fuels (oil, coal, and natural gas). The energy sector can be divided into six categories (second level of the diagram in figure 3): agriculture and fish, fugitive emission from energy production, transport, buildings, and industry-related energy consumption, which the last one can account for 24.2% of global GHG emissions.

The third level in figure 3, within industry-related energy consumption, can be divided into 7 subcategories, such as iron and steel (7.2%), chemical and petrochemical (3.6%), food and tobacco (1%), non-ferrous metals (0.7%), paper and pulp (0.6%), machinery (0.5%) and other industry (10.6%). It is clear from figure 3 that to reduce CO₂ emissions and reach net-zero emissions the world needs innovations across many sectors.

In Transport related to energy consumption, innovations to achieve efficient engines and the deployment of electric vehicles will transform cars and trucks to zero or low carbon emissions. In the Energy sector, an aggressive energy efficiency strategy combined with the ramping up of renewables to replace fossil fuels will rapidly shift towards zero or low-carbon sources of energy by 2030. Renewable technologies, including

solar, wind, hydro, biofuels, and others, are at the center of the transition to a less carbon-intensive and more sustainable energy system.

Following the climate targets set in the Paris Agreement, the European Union aims to lead in climate action, working towards becoming the first climate-neutral continent by 2050. The European Green Deal (EGD)¹, established at the end of 2019, represents the first strategy to combat climate change and achieve this ambitious goal. Initial targets were proposed to be achieved by 2030. On July 14, 2021, they introduced "Fit for 55," a comprehensive set of policies aimed at implementing the EGD. "Fit for 55" targets a substantial reduction of greenhouse gas emissions by at least 55% by 2030, in comparison to levels recorded in 1990. President Macron emphasized, in September of this year, that this requires an acceleration of efforts, describing the need to progress "twice as fast" as in previous years. He also affirmed that France plans to halt operations and transition the two-remaining coal-burning plants to biomass energy, generated by burning wood, plants, and organic materials, by the year 2027.²⁰

In this context of reducing or eliminating the reliance on fossil fuel, biomass feedstocks are indeed gaining increasing attention as a renewable and sustainable alternative to fossil resources for various purposes, including the production of biofuels²¹ (such as, biodiesel, ethanol, butanol, and hydrogen), high-value-added chemicals^{22,23} (like furfural, sugar alcohols), and various bioproducts. Biomass feedstocks utilized for these productions include energy crops, agricultural biomass residues, forest biomass, and food-based biomass waste²⁴. This shift towards biomass utilization is driven by environmental concerns, the desire to reduce greenhouse gas emissions, and the need for more sustainable resource management.

One way to make biomass processing more profitable in the economic aspect is by applying the biorefinery concept, i.e., many products are produced instead of one²⁵. In another words, biorefinery concept is a sustainable approach to resource utilization and industrial processes that takes inspiration from the traditional petroleum refinery but uses biological and renewable feedstocks instead of fossil fuels. It aims to maximize the value extracted from biomass materials, such as agricultural residues, energy crops, forest biomass, and food-based waste, by converting them into a variety of products including not only biofuels, but also a range of high-value chemicals, materials, and bioproducts, such as biodegradable plastics, and bio-based materials for construction and textiles. This approach enhances resource efficiency, reduces waste, and supports sustainability goals while contributing to more diverse and less carbon-intensive economy. Biorefineries play a crucial role in the transition to a bio-based, circular economy that minimizes environmental impacts and reduces dependence on fossil resources.

Chemical and thermochemical conversion processes in biorefineries encompass the transformation of lignocellulosic biomass²⁴. These methods, such as pyrolysis, hydrothermal conversion, gasification,

combustion and/or liquefaction, play a crucial role in efficiently extracting energy and materials from biomass. Subsequently, separation and purification process steps are employed to refine and extract the desired products from the mixtures generated during the conversion, ensuring the quality, purity, and suitability of these products for various applications²⁶.

The solvent selection used in the separation process design for biorefinery application plays an important role to define the environmental sustainability and economic viability of the process. Generally, 'green' solvents are attractive due to their low toxicity, economic value, and biodegradability²⁷.

In this context, the general objective of this thesis is:

- Develop a methodology for solvent screening based on calculating thermodynamic properties of mixed fluids to support solvent selection in the biorefinery, focusing on green solvents, specifically Deep Eutectic Solvents (DES). The model chosen for this aim was the COSMO-SAC (COnductor like Screening MOdel segment activity coefficient).

The COSMO-SAC model uses results from quantum mechanical calculations on individual molecules to predict thermodynamic properties of mixtures of these molecules, for example, activity coefficient. To compare the results obtained here, and optimize the model, the specific objectives of this thesis are:

- Study the role of computational variables in quantum mechanical calculations on the performance of the COSMO-SAC model, combining a theoretical and experimental investigation.
- Create a free database (named by PSL sigma-profile database) of quantum mechanical calculations of many components to feed the COSMO-SAC model.
- Optimize the COSMO-SAC predictions by studying the enthalpic and entropic contributions, and introducing an intermolecular contribution based on Margules activity coefficient model.
- Screen Deep Eutectic Solvent for the extraction of furanic compounds (5-hydroxymethylfurfural (HMF) and furfural (FF), and levulinic acid (LA)) produced by dehydration of fructose and xylose.

This manuscript is organized into seven chapters:

Chapter 1 will present the biorefinery concept, advantages, and disadvantages of some biomass feedstocks. The discussion is centered on the lignocellulosic biomass. A literature review of the pre-treatment of the biomass will be presented. Then, the industrial application of the value-added chemicals (HMF and FF) issue of the pyrolysis of the biomass, studied here in this thesis, will be presented. An overview of the main separation processes used in the industry will be discussed. In particular, the key

factors in the selection of the solvents for the liquid-liquid extraction. The last part focuses on the Deep Eutectic Solvents (DES) and their properties.

Chapter 2 covers the experimental part of the thesis and will describe the synthesis protocol and characterization of DES and mixtures. All the experimental equipment and protocols will be discussed in detail: focusing on acquisition of limiting activity coefficient data through gas stripping method and vapor liquid equilibrium data through synthetic method.

Chapter 3 will focus on the fundamentals of thermodynamic modeling, starting from the fugacity and activity definition and describing the two approaches to the calculation of two-phase equilibrium in a binary system: γ - ϕ and ϕ - ϕ approaches and PT-Flash algorithm. A review will be presented of some existing activity coefficient models developed in the last decades. The discussion is centered particularly on the COSMO-SAC model.

Chapter 4 will overview the main non-relativistic Quantum Mechanical methods used throughout this thesis. Schrödinger equation and the Born-Oppenheimer approximation will be introduced first, followed by the Density Functional Theory (DFT) methods. The last part will present Continuum solvation models, focusing on the COSMO approach.

Chapter 5 will present an experimental analysis of Deep Eutectic Solvents (DES), encompassing their inherent properties such as density, viscosity, and vapor pressure. Additionally, we have explored the behavior of vapor-liquid equilibrium of DES in alcohols, employing a synthetic method. The results were compared with the predictive ones obtained using the COSMO-SAC-dps model proposed by Hsieh (2014), revealing that the model is not suitable for deep eutectic solvents. Furthermore, an examination of the IDAC of DES was conducted using the gas stripping method. However, it should be noted that this method was not optimally suited for studying IDAC in fluids with high viscosity, such as DES.

Chapter 6 is dedicated to exploring the impact of computational variables on the performance of the COSMO-SAC model. An extensive benchmarking process, encompassing various theoretical methods and computational variables, has been conducted to comprehend their influence on the accuracy of predictions using the COSMO-SAC model. As a significant outcome of this study, we have developed the PSL sigma-profile database. The COSMO-SAC model has been applied using the PSL sigma-profile for activity coefficient predictions. Furthermore, these predictions have been compared with established activity coefficient models, including NRTL, UNIFAC, and Scatchard-Hildebrand.

Chapter 7 investigates the combinatorial and dispersion contribution within the COSMO-SAC model for predicting IDAC in DES. The insights derived from this study served as the foundation for enhancing the predictive capabilities of the COSMO-SAC model through innovative contributions inspired by the Scatchard-Hildebrand model. These include the Dispersive, Eccentricity, and Solubility contributions specific to DES interactions. These novel contributions were integrated into the newly proposed model known as the COSMO-SAC-DES model, which promises to offer enhanced predictive accuracy of solute-solvent interactions in DES systems.

Chapter 1

Separation process in the biorefinery

Processus de séparation dans la bioraffinerie

Résumé

Les installations de bioraffinerie (similaires aux raffineries pétrolières) utilisent des sources de biomasse au lieu de sources fossiles pour produire de l'énergie, carburant et produits chimiques. La pyrolyse est un processus clé pour transformer les déchets en produits à valeur ajoutée impliquant la décomposition de la biomasse organique à haute température en produits liquides (bio-huile), gazeux (syngaz) et solides (biochar). Différentes sources de biomasse incluent les cultures énergétiques, les résidus agricoles, les déchets alimentaires et la biomasse forestière. La biomasse est une source d'énergie renouvelable majeure en France, contribuant à réduire la dépendance aux combustibles fossiles.

À l'horizon 2050, le gisement de gaz renouvelable injectable estimé 460 TWh pourrait couvrir entièrement la demande de gaz en France. Cette évaluation est basée sur une étude réalisée et publiée par l'Agence de la Transition Écologique (ADEME), Gaz Réseau Distribution France (GRDF) et Gaz Réseau Transport (GRTgaz)²⁸. Cette étude a exploré la faisabilité technico-économique d'un gaz 100 % renouvelable. La pyrogazéification de la biomasse lignocellulosique issue des forêts est une option prometteuse.

La biomasse lignocellulosique (LC) est le matériau végétal le plus abondant qui peut servir de substitut aux combustibles fossiles dans la production et l'approvisionnement durable en biocarburants liquides. La biomasse LC est composée principalement de trois biopolymères (environ 40 % de cellulose, 25 % d'hémicellulose et 25 % de lignine) et d'environ 10 % d'autres éléments et de composants mineurs.

Pour accéder à tous ces polysaccharides très appréciés dans la production de biocarburants et produits chimiques à valeur ajoutée, un processus de séparation est nécessaire. En raison de plusieurs facteurs tels que la structure complexe des constituants de la paroi cellulaire dans la biomasse LC, l'hétérogénéité structurale, la cristallinité de la cellulose et le degré de lignification, la séparation de la cellulose et de la lignine est un grand défi. Pour surmonter ces facteurs, le processus de prétraitement devient une étape significative pour réduire le degré de polymérisation, dégrader l'hémicellulose et rompre la gaine de lignine, perturbant ainsi la structure récalcitrante de la lignocellulose. Cela peut réduire la consommation réactive chimique, les coûts et la demande énergétique et améliorer la production de monomères de sucre lors de l'hydrolyse.

Les techniques de prétraitement des diverses sources lignocellulosiques peuvent être regroupées en quatre catégories principales : physiques, chimiques, physico-chimiques, et biologiques. Le choix de la méthode

de prétraitement dépend du type de biomasse LC utilisé, puisque la composition en cellulose, hémicellulose et lignine varie. Certaines méthodes de prétraitement et leurs effets sur la séparation des composants complexes de différentes sources lignocellulosiques seront discutées.

Les technologies de conversion des biomasses lignocellulosiques en produits de valeur et sources d'énergie se répartissent en deux catégories principales : la conversion thermochimique et la conversion biochimique. Ce chapitre se concentrera sur la technologie thermochimique : la pyrolyse.

La pyrolyse est un terme couramment utilisé pour décrire la dégradation thermique de la matière première en l'absence d'oxygène, généralement à des températures comprises entre 300 et 600 °C. Ce processus crée du charbon solide (également appelé biochar), du bio huile, et de la vapeur de pyrolyse. La distribution et les propriétés de ces produits de pyrolyse sont influencées par plusieurs facteurs physiques et chimiques, notamment le type de biomasse, la taille des particules, la température, la pression, la vitesse de chauffage, le temps de résidence, etc.

D'un point de vue thermique, la pyrolyse de la biomasse se déroule en quatre étapes distinctes : séchage, dévolatilisation de la biomasse, dévolatilisation étendue et dévolatilisation secondaire des biohuiles.²⁹⁻³¹ La pyrolyse peut être classée en cinq types principaux (la classification étant déterminée par les combinaisons particulières de taux de chauffage, de températures et de durée du processus): pyrolyse lente, pyrolyse rapide, pyrolyse humide (ou carbonisation hydrothermale), pyrolyse hydrique (ou liquéfaction hydrothermale) et pyrolyse douce (torréfaction). Elles seront discutées en détail dans ce chapitre.

Actuellement, le processus de pyrolyse est principalement utilisé pour produire la fraction liquide, c'est-à-dire le bio huile, qui sert de carburant pour la production d'énergie via des cogénérateurs ainsi que des produits chimiques à valeur ajoutée.

Les molécules plates désignent des composés chimiques clés obtenus à partir de la bio-huile. Ces molécules servent de blocs de construction fondamentaux pour la synthèse d'une variété de produits chimiques à valeur ajoutée tels que des plastiques, des résines, des solvants et des produits pharmaceutiques. Deux molécules plates essentielles (le 5-hydroxyméthylfurfural et le furfural) sont obtenues à partir de la déshydratation du fructose et de la xylose et constituent des intermédiaires cruciaux dans la synthèse de nombreux produits chimiques industriels.

Des étapes de séparation et de purification seraient nécessaires pour obtenir tous les produits d'intérêt. Le choix du solvant utilisé dans la conception du processus de séparation pour l'application en bioraffinerie joue un rôle important dans la définition de la durabilité environnementale et de la viabilité économique

du processus. Dans ce cadre, cette thèse explore les progrès et le développement des solvants verts en particulier les solvants eutectiques profonds pour l'extraction de composés furaniques (FF, 5-hydroxyméthylfurfural (HMF) et acide lévulinique (LA)) produits par déshydratation du fructose et du xylose.

Les bioraffineries utilisent divers procédés de séparation notamment la distillation, l'extraction liquide-liquide, l'adsorption, l'absorption et les séparations par membrane. Le choix du solvant est crucial pour l'extraction liquide-liquide et ses propriétés physico-chimiques telles que la solubilité, la viscosité, la masse volumique, etc. influent sur l'efficacité de la séparation. Le recyclage des solvants peut réduire les coûts énergétiques, et favoriser la durabilité.

L'efficacité d'un solvant dépend de son coefficient d'activité à dilution infinie, de sa capacité de solvation et de sa sélectivité. Le choix du solvant peut considérablement affecter l'efficacité globale du processus de séparation. Les solvants alternatifs plus respectueux de l'environnement suscitent un intérêt croissant en tenant compte de l'impact de la pollution, de la consommation d'énergie et des préoccupations environnementales.

Dans ce contexte, cette thèse se concentre sur les DES, qui sont une catégorie de solvants formés en mélangeant divers composants, généralement constitués d'un donneur de liaison hydrogène (tel qu'un sel d'ammonium quaternaire) et d'un accepteur de liaison hydrogène (comme un chlorure métallique). Lorsqu'ils sont mélangés en proportions spécifiques, ces composants créent un mélange eutectique qui se comporte comme un liquide à des températures relativement basses souvent en dessous des points de fusion des composants individuels. Les DES suscitent plus un intérêt en tant qu'alternatives durables et respectueuses de l'environnement que les solvants organiques traditionnels et les liquides ioniques.

Fondamentalement, en se concentrant sur les propriétés physico-chimiques des DES, il convient de noter que la conception et l'amélioration du processus de séparation dépendent de propriétés thermodynamiques et de transport précises et facilement disponibles.

En résumé, le Chapitre 1 aborde les aspects de la bioraffinerie, examine les avantages et les inconvénients des matières premières à base de biomasse, en mettant l'accent sur la biomasse lignocellulosique. Il offre une revue de la littérature sur le prétraitement de la biomasse. De plus, il présente l'application industrielle des produits chimiques à valeur ajoutée (HMF et FF) obtenus par pyrolyse de la biomasse. Il explore également les principaux procédés de séparation industrielle, en mettant l'accent sur les facteurs déterminants dans le choix des solvants pour l'extraction liquide-liquide. Enfin, il se penche sur les DES et leurs caractéristiques.

Separation process in the biorefinery

1.1. Biorefinery

The biorefinery facilities are analogous to the oil refineries, which instead of using fossil sources (oil & gas) to convert into energy, fuel, and chemicals, use biomass sources. In a linear economy, the products are used by human activities, and in the end, wastes are disposed of in the environment, or incinerated. To reduce waste and its impact on our environment, the biorefinery concept came as a sustainable method of recycling and reusing waste by a process called by pyrolysis and turning it into value-added products (Figure 1.1). Pyrolysis is a thermochemical conversion process in which organic materials, such as biomass residue and organic waste, are subjected to high temperature in an inert atmosphere with limited oxygen. Pyrolysis can also be done by hydrothermal conversion, also known as hydrothermal carbonization (HTC) or wet pyrolysis. In this method, organic materials, are subjected to elevated temperatures and pressures in the presence of water. As result, of pyrolysis processes, the organic material breaks down into various products, including liquid (bio-oil), gaseous fractions (syn-gas) and solid (biocarbon or biochar), with different proportions of the obtained gaseous, liquid and solid phases, depending on the conditions of pyrolysis and on the nature of biomass. These products can be used directly or after processing as fuel, electricity, and heat. Biocarbon and bio-oil can also be utilized to produce chemicals and value-added products.

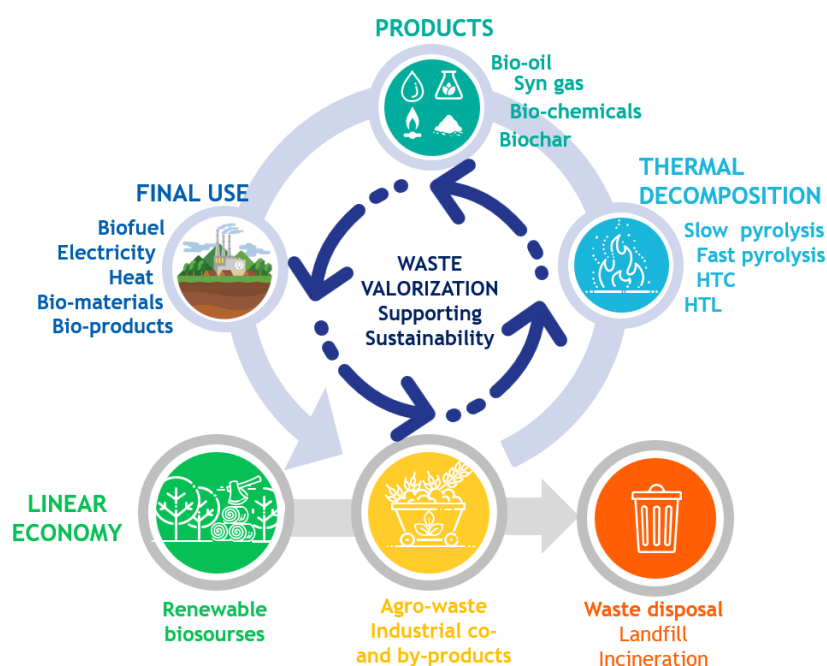


Figure 1.1 – Schematic diagram of a promising closed-loop sustainable waste reutilization to develop value-added green products from biomass.

Biomass can be obtained or produced from almost unlimited renewable sources. In the next section, various types of biomass materials will be compared.

1.1.1. Comparison of biomass feedstock

Energy crops are specifically grown to produce biofuels or directly exploited for their energy contents. Renewable sources can be grains of food crops, such as corn, sugarcane, sugar beet, sweet sorghum, etc. Nonfood crops can be used to make biofuel production more sustainable because they are not in direct competition with food production. Examples of nonfood crops are poplar trees, switchgrass, miscanthus, kenaf, etc. Most of the current biofuel production uses so-called conventional feedstocks, such as corn, and sugarcane to produce bioethanol. The United States of America is the world's largest producer of ethanol, having produced over 15.4 billion gallons in 2022. Brazil is the second one, with an output of 7.5 billion gallons that same year. Together, the United States and Brazil produce 82% of the world's ethanol. European Union is the third one, having produced about 1.5 billion gallons in the same year. France is the largest producer of fuel ethanol in European Union. In 2022, the country had a production volume of over 369.5 million gallons, i.e., 25% of the European Union production. The U.S.A. ethanol production majority uses corn as feedstock, while Brazil primarily uses sugarcane. In Europe, the main crops to produce bioethanol are starch crops (such as common wheat) and sugar beet.⁴

Agricultural biomass residues from production chains are important sources of low-cost feedstock. Agricultural biomass includes bagasse, straw, stem, stalk, leaves, husk, shell, peel, pulp, stubble, etc. These can provide a substantial amount of biomass to produce a wide range of value-added products without increasing the amount of land used for agriculture.

Food biomass waste from post-industrial, such as residues and byproducts, and post-consumer can be used as a biomass energy source. Food waste can include oil residues, nuts hulls, peanuts shells, corn stover, rice hull/husk, and other grain biomass residues.

Forest biomass is the largest source of lignocellulosic (LC) biomass that can be a good candidate to replace fossil fuels in the production of energy and other value-added products at relatively low cost. Forest biomass includes trunks, crowns, and branches, unused wood from forests and wood manufacturing, and processing residues. The composition of the LC biomass depends on the type of the tree, deciduous (hardwood) or coniferous trees (softwood). The wastes from wood industries are sawdust, off-cuts, trims, and shavings. The high costs of harvesting and transportation are the main challenges to use these biomass materials as feedstocks. On the other hand, forest biomass has been widely used directly to produce heat by the combustion process.

Table 1.1 summarizes the advantages and disadvantages of biomass materials for utilization as raw materials to produce value-added products.

Table 1.1 – Comparison of biomass materials to produce value-added products.²³

Biomass Material	Advantages	Disadvantages
Switchgrass	<ul style="list-style-type: none"> - Abundant and widely grown. - Requires less fertilizer, water, and energy for production. - Perennial nature, reducing the need for replanting each year. - Potential for use in bioenergy, bioproducts, and as a feedstock for various industrial applications. 	<ul style="list-style-type: none"> - Lower energy content. - Requires suitable soil and climate conditions. - May compete with food crops for land.
Miscanthus	<ul style="list-style-type: none"> - High biomass productivity (more biomass than switchgrass). - Requires less fertilizer, water, and energy for production. - Perennial nature, reducing the need for annual replanting. - Versatile uses in paper production, fiberboards, textiles, and animal bedding. 	<ul style="list-style-type: none"> - Initial establishment costs can be relatively high. - May be invasive in some regions. - Regional suitability and adaptability may vary.
Kenaf	<ul style="list-style-type: none"> - Rapid growth (in tropical and subtropical regions). - Requires less water, fertilizer, and pesticides for production. - Versatile uses in paper production, fiberboards, textiles, and biodegradable products. 	<ul style="list-style-type: none"> - Not well-suited for temperate or cold climates. - Limited commercialization and awareness. - Requires suitable soil conditions and management for optimal growth.
Corn	<ul style="list-style-type: none"> - Abundant and well-established crop. - Relatively low-cost source of starch, primarily used for ethanol production. 	<ul style="list-style-type: none"> - Competition with food markets. - High water and fertilizer requirements. - Seasonal nature, requiring replanting each year.
Sugarcane	<ul style="list-style-type: none"> - High sugar content, making it a valuable source for bioethanol and other bio-based products. - Efficient photosynthesis and low fertilizer requirements. - Utilizes waste product (bagasse) for additional value-added products. 	<ul style="list-style-type: none"> - Restricted to tropical and subtropical regions. - Concerns about deforestation and land-use change. - Water-intensive crop, requiring substantial irrigation.
Sweet Sorghum	<ul style="list-style-type: none"> - Requires less fertilizer compared to some other crops. - High sugar content in the juice, suitable for bioethanol and other value-added products. - Potential for dual-use: grain production and value-added use of the stalks for bioenergy. 	<ul style="list-style-type: none"> - Limited flexibility in harvesting due to the need to process the juice immediately. - Requires suitable climate conditions for optimal growth. - Challenges in transportation due to the rapid deterioration of harvested stalks.
Poplar Trees	<ul style="list-style-type: none"> - Fast growth rates and high biomass productivity. - Suitable for bioenergy, pulp and paper, and construction. - Erosion control and carbon sequestration benefits. 	<ul style="list-style-type: none"> - Long growth cycle before commercial harvest (3-10 years). - High initial investment and land preparation costs. - Potential environmental concerns if not managed sustainably.
Cereal Straw	<ul style="list-style-type: none"> - Abundant and widely available as a byproduct of cereal crops like rice, wheat, and corn. - Provides an additional revenue stream for farmers and reduces waste from primary crop harvests. - Suitable for use in bioenergy, animal bedding, and soil amendment applications. 	<ul style="list-style-type: none"> - Lower energy content compared to dedicated energy crops. - Requires efficient collection and handling to minimize losses and maintain quality. - Seasonal availability tied to primary crop harvests.
Corn Stover	<ul style="list-style-type: none"> - Abundant and widely available as a byproduct of corn cultivation. - Increases overall economic value for farmers by utilizing the entire corn plant. - Suitable for use in bioenergy, animal bedding, and soil amendment applications. 	<ul style="list-style-type: none"> - May compete with other uses of corn stover (e.g., livestock bedding). - Requires efficient collection and handling to minimize losses and maintain quality. - Seasonal availability tied to corn harvests.
Food Biomass Wastes	<ul style="list-style-type: none"> - Large and consistent supply of biomass generated from food processing and waste. - Potential for conversion into bioenergy, biochar, and other value-added products. - Addresses waste management and environmental concerns associated with food waste. 	<ul style="list-style-type: none"> - Challenges in sorting and processing to remove contaminants and enhance purity. - Variability in composition and quality of food waste biomass. - Potential concerns related to food security and resource allocation.
Forest Biomass	<ul style="list-style-type: none"> - Largest source of lignocellulosic biomass, providing abundant raw material for bioenergy and bioproducts. - Carbon-neutral or carbon-negative if managed sustainably. - Potential for improved forest management and waste utilization to reduce wildfire risks. 	<ul style="list-style-type: none"> - High costs associated with harvesting, transportation, and processing. - Ecosystem and biodiversity concerns if not managed sustainably. - Limited suitability for certain bioconversion processes like gasification and fermentation.

In France, biomass is the main source of renewable energy in France, representing over 55% of final energy production and significantly contributing to reducing our consumption of fossil fuels.³² Biomass produces energy for various uses such as heat, electricity, biogas, or fuels. This biomass comes from the forest, agriculture (dedicated crops, crop residues, intermediate crops, and livestock effluents), waste (green waste; household biowaste; waste from catering, distribution, agri-food industries, and fishing; wood waste; sludge from wastewater treatment plants; etc).

In the projected scenario for the year 2050, the gas demand of 276 to 361 TWh can be met with renewable gas.²⁸ This assessment is based on a study that explored the techno-economic viability of 100% renewable gas, conducted and published by the Agency for Ecological Transition (ADEME), Gaz Réseau Distribution France (GRDF), and Gaz Réseau Transport (GRTgaz). According to this study, the potential of renewable gas of 460 TWh, has the capability to completely meet the gas requirements of France by the year 2050. The study examines three primary renewable pathways for renewable gas production: methanization (30% of the potential), pyrolysis (40%), and power-to-gas (30%). It's important to note that the technical evaluations are rooted in available resources that do not interfere with food and raw material usage. Among non-edible biomass, forest biomass is a promising feedstock. The French Forest is the 4th largest forest in Europe, growing from 1.8 to 2.8 billion cubic meters between 1985 and today, the volumes of wood in our forests have also increased significantly over the past 30 years.³³

The perspective presented here leads into the broader discussion on biomass, particularly related to lignocellulosic biomass within the context of forests, as elaborated in the subsequent section. Lignocellulosic biomass refers to plant-based biomass primarily composed of three major components: cellulose, hemicellulose, and lignin. These components are essential structural elements found in the cell walls of plants, including trees, grasses, agricultural residues, and other woody materials. In the upcoming section, a detailed examination of lignocellulosic biomass will be provided.

1.1.2. Lignocellulosic biomass

LC biomass is the most abundant plant material that can be a substitute for fossil fuels in the production and sustainable supply of liquid biofuels. The compositions of some LC biomass types are presented in Table 1.2.

Table 1.2 – The major chemical composition of various LC biomasses³⁴.

Compound	Cellulose (%)	Hemicellulose (%)	Lignin (%)
Hardwood	40-55	24-40	18-25
Softwood	45-50	25-35	25-35
Sugarcane Straw	33-40	21-32	19-32
Sugarcane Bagasse	36-45	25-35	19-32
Wheat Straw	33-40	20-25	15-20
Rice Straw	28-36	23-28	12-14
Wastepaper	60-65	4-13	1-2

LC biomass is composed roughly of three biopolymers, such as about 40% of cellulose, 25% of hemicellulose, 25% of lignin, and about 10% of other elements and some minor components. All these three biopolymers, cellulose, hemicellulose, and lignin are the main constituents of plant cell walls and will be better described above.

Cellulose is a polysaccharide linear homopolymer. It consists of several hundred to many thousands of β -1,4-glycosidic linked D-glucopyranose units (Figure 1.2). Cellulose has been recognized as being a highly useful polymer since it can be depolymerized into D-glucose monosaccharides units. D-glucose (C6) contains six carbon atoms and an aldehyde group and can be used to produce alcohol or other chemicals using different chemical methods.

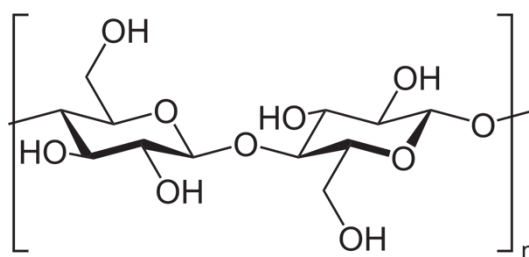


Figure 1.2 – Cellulose chemical structure.

Hemicellulose is a polymer consisting of short linear chains and branched saccharide polymers. Unlike cellulose, hemicelluloses are composed of diverse five-carbon sugars (D-xylose and D-arabinose), six-carbon sugars (D-glucose, D-galactose, D-mannose), and sugar acids. In contrast, cellulose is derived exclusively from D-glucose (Figure 1.3).

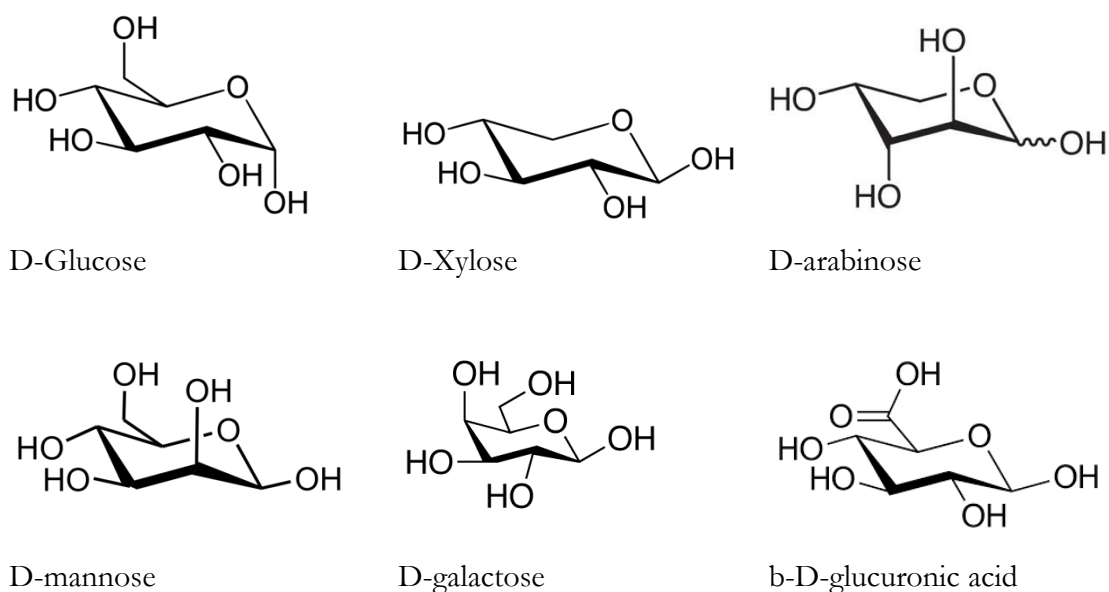


Figure. 1.3 – Hemicelluloses monosaccharide.

Lignin is a collection of highly heterogeneous polymers derived from a handful of precursor monolignols namely p-coumaryl alcohol, coniferyl alcohol, and sinapyl alcohol (Figure 1.4).

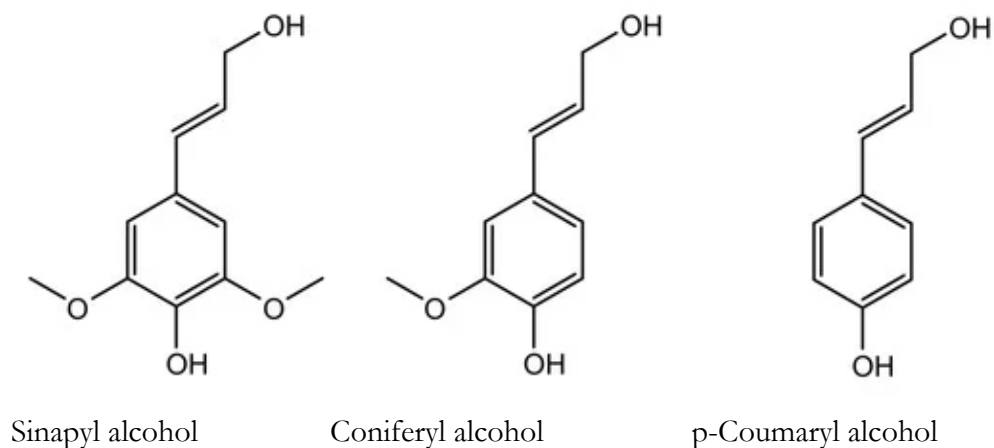


Figure 1.4 – Lignols.

The lignin and hemicelluloses form a cross-linked matrix that covers the cellulose fibers, as depicted in figure 1.5, and thereby forms a tight and compact structure providing the plant cell walls strength and hydrophobic and indigestible properties.

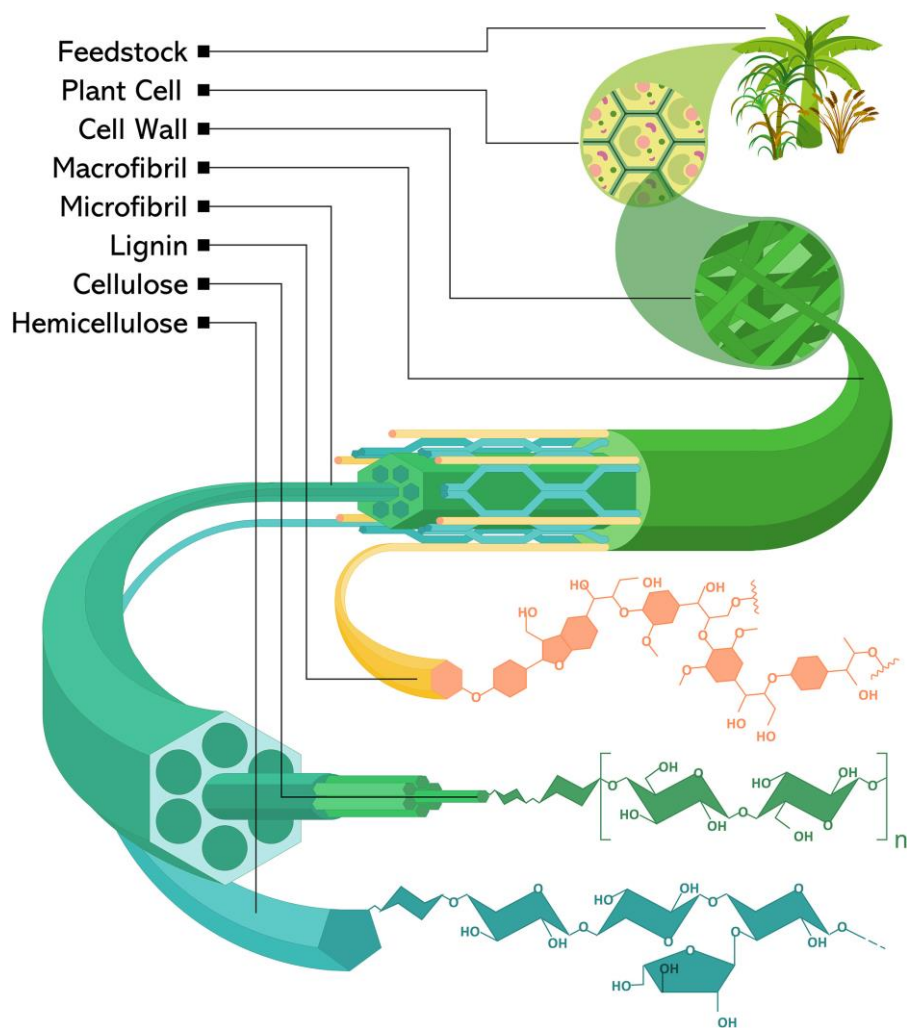


Figure 1.5 – Structure of lignocellulosic biomass.³⁴

To access all these polysaccharides which are highly valued in the production of biofuels such as bioethanol and biogas, various organic acids, phenols, and aldehydes, a separation process is required. However, the presence of lignin in LC biomass is the main obstacle to the biomass being resistant to biological and chemical breakdown and releasing their sugars. Due to multiple factors, such as the complex structure of cell wall constituents in LC biomass, structural heterogeneity, crystallinity of cellulose, and degree of lignification, the separation of cellulose and lignin is a big challenge. To overcome these factors, the pretreatment process becomes a significant step to the reduced degree of polymerization, degradation of hemicellulose, and breakage of lignin sheath, upsetting the recalcitrant structure of lignocellulose. Which, it can reduce chemical reactive consumption, cost, and energy demand, and enhance the production of monomers of sugar during hydrolysis.

1.1.3. Pre-treatment Lignocellulosic biomass

Pretreatment techniques of various lignocellulosic sources can be divided into four major categories: physical, chemical, physicochemical, and biological. Among the several pretreatments processes the preference for the convenient one depends on the type of LC biomass used as the composition of cellulose, hemicellulose, and lignin varies. Some of the pretreatment and its effects on the separation of the complex components of various lignocellulosic sources are discussed in Table 1.3.

Table 1.3 – Summary of lignocellulosic pretreatment methods.³⁵

Pre-treatment method	Process	Principles	Critical view point	Ref.
Physical	Milling, chipping, grinding	Easier handling and increased surface area to volume ratio which facilitates reactivity. Decreased degree of lignocellulose crystallinity.	First step in multiple pre-treatment processes. High power consumption.	36–38
	Irradiation	The energy breaks the hydrogen bonds of the cellulose crystalline structure and makes it prone to enzymatic digestion	Efficient in the presence of lignin. Expensive and not convenient for large scale application	39,40
	Hydro-thermal	Uses water at high temperature and pressure that dissolves most of the lignin and hemicelluloses which in turn facilitates the hydrolysis of the cellulosic fraction	Generates acetic acid and other organic acids	41,42
	Pyrolysis	Use of high temperature to disrupt the lignocelluloses	Efficient when carried out in the presence of oxygen	43
Physio-chemical	Explosion e.g. steam explosion, ammonia fiber explosion, CO ₂ explosion	Alters the structure of cellulosic biomass to make it more accessible. Exposes biomass to high temperature and pressure followed by a sudden pressure fall to make an explosive decomposition	Steam explosion is the most commonly used method for the pretreatment of lignocellulosic biomass	44–47
Chemical	Acid	Increases the biomass porosity by removing the hemicelluloses and altering the lignin structure, and this facilitates enzymatic digestibility	Accompanied with aldehyde formation. Consumes significant energy for pretreatment and product recovery	41,47
	Alkali	Removes lignin, acetyl and various uronic acid substitutions from the biomass by saponification that breaks the intermolecular ester bonds. This improves enzymatic digestibility of the biomass	Utilize lower temperatures and pressures compared to other pretreatment methods	48–50
	Wet oxidation	Treatment of biomass in the presence oxygen/air and water at high temperature and pressure opens the crystalline structure of cellulose	All biomass fractions are affected. Hemicelluloses degrade substantially	51–53
	Ozonolysis	Targets lignin degradation by attacking and cleavage of aromatic rings structures	The cellulose and hemicellulose fractions remain intact	54,55
	Solvent extraction	Use solvents to remove lignin and some hemicelluloses to facilitate enzymatic hydrolysis of lignocellulosic biomass. Often accomplished at moderately high temperature	Requires removal of the solvent from the treated biomass	56–58
Biological	Micro- biological	Micoorganisms (often fungi) degrades lignin and hemicelluloses	Long process but low in energy consumption and requirement.	59

Biomass valorization is the process of converting dried plant waste into a variety of valuable bio-chemicals. This conversion process involves two main phases. The first phase is biomass pretreatment, discussed in this section. It aimed at overcoming the complex lignocellulosic structure and facilitating the conversion of biomass into valuable products like fuel, food, and chemical building blocks. The second phase involves either biological (metabolic) or thermochemical (gasification, liquefaction, pyrolysis, and combustion) processing.

1.1.4. Conversion technologies of lignocellulosic biomass

Conversion technologies to transform lignocellulosic biomass into valuable products and energy sources encompass two primary categories, as presented in Table 1.4: thermochemical and biochemical conversion. Thermochemical processes, which include combustion, pyrolysis, gasification, and hydrothermal conversions, are perceived to be more efficient in converting waste biomass to energy and value-added products than biochemical processes. This section will focus on the thermochemical technology: the pyrolysis.

Table 1.4 – Conversion technologies.⁶⁰

	Conversion Technology	Corresponding Products
Thermochemical Conversion	Combustion	Heat and Electricity
	Pyrolysis	Bio-oil, Syngas and Char
	Gasification	Syngas, liquid fuels and Char
	Hydrothermal conversions	Liquid Fuels, Chemicals, Syngas
Biochemical Conversion	Anaerobic Digestion	Biogas, Liquid fuels, Electricity
	Fermentation	Ethanol, Butanol

1.1.4.1. Pyrolysis

Pyrolysis is a term commonly used to describe the thermal degradation of feedstock in the absence of oxygen, typically occurring at temperatures within the range of 300 to 600 °C.²⁹ The typical scheme of the biomass pyrolysis process is illustrated in figure 1.6.

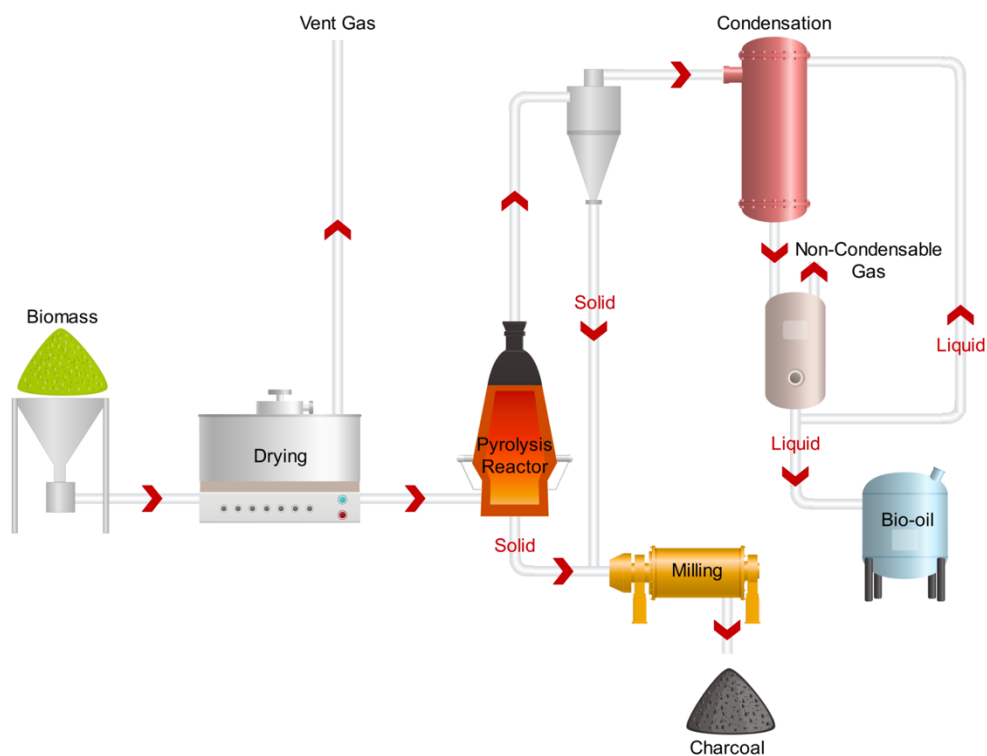


Figure 1.6 – Biomass pyrolysis scheme. ⁶⁰

This process leads to the creation of solid char, also known as biochar, tar, or bio-oil, and pyrolysis vapor. The pyrolysis vapor is then condensed to obtain bio-oil and producer gas. The distribution and properties of these pyrolysis products are influenced by several physical and chemical factors, including the type of biomass, particle size, temperature, pressure, heating rate, residence time, and more. The impact of temperature on the distribution, yield, and quality of pyrolysis products stands as the most notable factor among those mentioned. ³¹

From a thermal perspective, biomass pyrolysis proceeds through four distinct stages. Initially, in the drying stage, occurring at low temperatures, unbound moisture within the biomass evaporates. Subsequently, in the second stage, known as biomass devolatilization, a gradual process occurs within the temperature range of 100 to 200 °C, releasing low-molecular-weight gases like N₂, CO, CO₂, and small quantities of acetic acid and methanol. The third stage is characterized by extensive devolatilization at higher temperatures ranging from 200 to 600 °C. During this stage, large biomass molecules decompose into char, condensable gases (such as ketones, aldehydes, phenols, esters), and non-condensable gases. Finally, in the last stage, secondary devolatilization of tars takes place, converting them into secondary char and non-condensable gases (such as H₂, CO, CH₄, CO₂, etc.) at temperatures typically between 300 to 900 °C. ³⁰

Pyrolysis can be categorized into five primary types, with classification determined by the particular combinations of heating rates, temperatures, and process duration: slow or conventional pyrolysis, fast pyrolysis, wet pyrolysis (referred as hydrothermal carbonization), hydrous pyrolysis (or hydrothermal liquefaction), and mild pyrolysis (also known as torrefaction).

Slow pyrolysis is a thermal decomposition process where biomass materials are gradually heated, usually at temperatures around 500 °C, without the presence of oxygen or with limited oxygen supply.^{61,62} This method involves a relatively moderate heating rate compared to faster pyrolysis processes. During slow pyrolysis, the biomass undergoes extended residence times due to the moderate temperature, resulting in the formation of three main fractions: biochar (a solid carbon-rich residue), bio-oil (in liquid form), and gas. These fractions are typically obtained in approximately equal proportions. Biochar, a stable carbon-rich solid product, is a significant output of slow pyrolysis. The bio-oil produced during this process may vary in composition and properties, often requiring further refining or upgrading before it can be used as a fuel or chemical feedstock.

Fast pyrolysis is a rapid thermal decomposition process involving the swift heating of biomass materials at high temperatures, typically within the range of 500 to 650 °C, in an environment with limited or no oxygen.⁶¹ This method employs significantly higher heating rates and shorter residence times compared to slow pyrolysis, allowing for the quick gasification of the biomass. A prominent outcome of fast pyrolysis is the substantial production of bio-oil, constituting a significant portion, usually around 60-70%, of the initial biomass input. In addition to bio-oil, fast pyrolysis also yields char and gases, with the gas fraction typically accounting for about 20% of the biomass input. The efficiency and higher energy density of the bio-oil make fast pyrolysis an attractive technology for biomass conversion.

Wet pyrolysis or Hydrothermal carbonization (HTC) is an advanced thermochemical process that transforms wet biomass and organic waste materials into a carbon-rich substance called hydrochar.⁶³ This conversion takes place in an aqueous environment at elevated temperatures (usually between 180 and 250 °C) and high pressures (typically between 10 and 60 bar). Water is a vital component in HTC, acting as both a reactant and a medium for the reactions that occur. The organic compounds in the biomass undergo a series of complex reactions, including hydrolysis, dehydration, polymerization, and carbonization, resulting in the formation of hydrochar.

Hydrous pyrolysis or Hydrothermal liquefaction (HTL) is an advanced thermochemical process that transforms various types of biomass or organic feedstocks into a liquid-like substance, often referred to as biocrude or bio-oil.⁶⁴ This conversion occurs under elevated temperatures, typically ranging from 150 to 450 °C in the absence of oxygen under high pressures (>1 MPa) and in the presence of water. During

the HTL process, the biomass undergoes liquefaction, resulting in the production of a bio-oil that shares characteristics with crude oil. This bio-oil can serve as a valuable renewable liquid fuel and holds potential for further refining to obtain high-quality transportation fuels. One of the primary objectives of HTL is the efficient production of bio-oil, making it a promising avenue in renewable energy.

Mild pyrolysis or Torrefaction (also known as carbonization), represents a thermochemical process that involves heating biomass in the absence of oxygen, typically within the temperature range of 200 to 300°C. This process is designed to partially carbonize the biomass, resulting in a dry, energy-dense, and stable product known as torrefied biomass or biochar.⁶⁰ Notably, torrefaction occurs at lower temperatures and shorter residence times compared to conventional pyrolysis, thus maintaining the solid structure of the biomass. The process includes initial drying to remove moisture content, followed by the release of volatile components as the temperature rises. The remaining solid biomass is then transformed into a stable, carbon-rich material with improved energy content and reduced reactivity. The resulting torrefied biomass, or biochar, exhibits enhanced properties such as higher energy content, improved grindability, hydrophobicity, and extended shelf life compared to the original biomass.

Currently, the pyrolysis process is primarily utilized to produce the liquid fraction, i.e., bio-oil, which serves as fuel for energy production through co-generators, and value-added chemicals.^{60,65}

1.1.5. Value-added chemicals: platform molecules

Platform molecules refer to key chemical compounds that are obtained from bio-oil, a product of the pyrolysis process. These molecules act as fundamental building blocks or intermediates for the synthesis of a diverse range of value-added chemicals. They serve as starting points for intricate chemical reactions and processes, enabling the creation of a wide spectrum of valuable substances such as plastics, resins, solvents, pharmaceuticals, and more.

In Figure 1.7, the dehydration process of fructose and xylose is illustrated. This process leads to the production of essential value-added derived chemicals commonly referred to as platform molecules, namely 5-Hydroxymethylfurfural and furfural. The conversion of fructose and hemicellulose into these platform molecules is a significant step in biomass transformation.^{66–68} These platform molecules, 5-Hydroxymethylfurfural and furfural, serve as crucial intermediates in the synthesis of a wide array of value-added chemicals with diverse industrial applications.

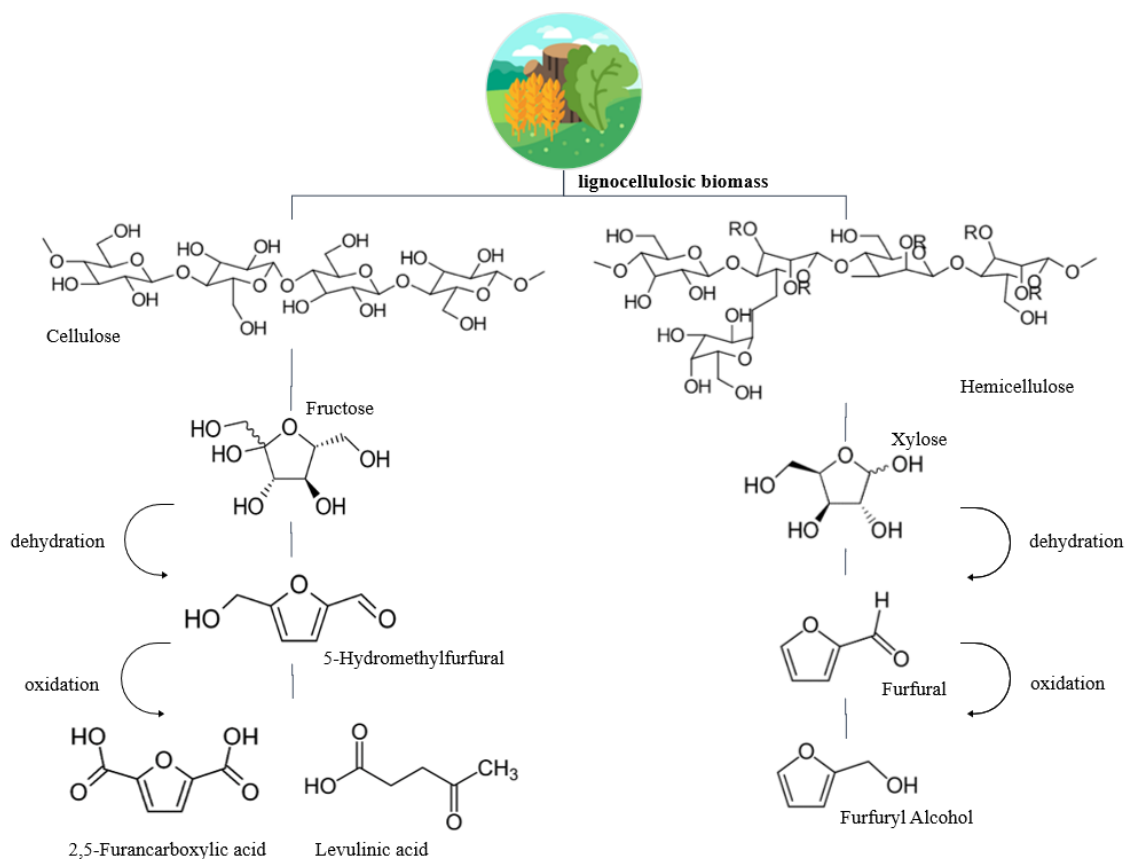


Figure. 1.7 – Dehydration of fructose and hemicellulose.

The multistep acid-catalyzed dehydration of hexoses is an effective method for the preparation of 5-Hydroxymethylfurfural (HMF).⁶⁷ Hexoses, including cellulose, glucose, fructose, sucrose, starch, inulin, and various raw biomass materials like wheat straw, serve as active substrates in this process. Carbohydrates such as cellulose can undergo a three-step reaction to yield HMF. The process involves the hydrolysis of cellulose to glucose, isomerization of glucose to fructose. Subsequent, the dehydration of fructose by removing three water molecules to obtain HMF (as illustrated in left side of figure 1.7). It's worth noting that different types of hexoses exhibit varying levels of selectivity in HMF formation. Fructose and other ketoses demonstrate the highest selectivity for HMF production, while glucose and other aldoses exhibit a lower selectivity in this regard.

5-Hydroxymethylfurfural (HMF) derived from fructose, is a top value-added biomass-derived chemical, which can be transformed into various important chemicals and fuels because of the presence of C=O, C-O, and furan ring functional groups. One of the chemicals that are obtained from the HMF is the 2,5-Furancarboxylic acid (FDCA) by the oxidation of the C=O and C-O groups of the side chains. FDCA (furan dicarboxylic acid) is the key building block for the biopolymer polyethylene furanoate (PEF). Another derivative of the HMF is the 2,5-dimethylfuran (DMF) obtained by the hydrogenolysis of the

C=O and C-O groups of the side chains. DMF has similar properties to commercial gasoline. As well as ethoxymethylfurfural, furfuryl alcohol, dimethylfuran, and 2,5-diformylfuran. These chemicals serve as the building blocks of diverse commodities including pharmaceuticals, polymers, resins, solvents, fungicides, and biofuels.

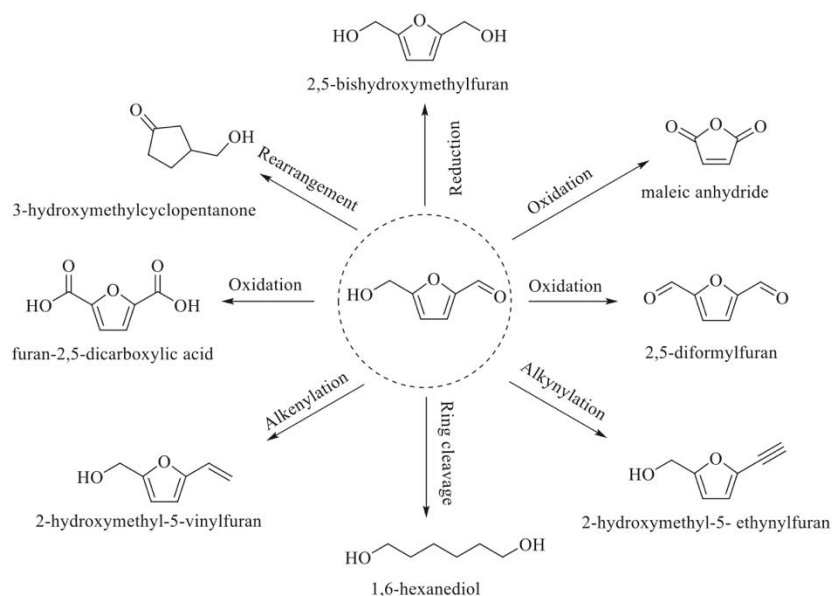


Figure 1.8 - 5-Hydroxymethylfurfural (HMF) platform molecules.⁶⁶

In contrast to cellulose, characterized by its crystalline structure and robust resistance to hydrolysis, hemicellulose possesses an amorphous, disordered structure with lower strength. It is susceptible to hydrolysis, particularly under the influence of dilute acid or base. When aiming to obtain furfural from hemicellulose, a multistep process involving an acidic catalyst is employed.⁶⁶ In this multistep process, the acidic catalyst initiates the breakdown of pentose sugars, making xylans within hemicellulose more susceptible to hydrolysis. This process transforms xylans into xylose through hydrolysis. The resulting xylose, formed during acid hydrolysis, can then undergo cyclodehydration to obtain furfural. The acidic catalyst not only facilitates the hydrolysis of xylans but also promotes the subsequent conversion of xylose into furfural, an important compound widely used in various industrial applications.

Furfural (FF) is an important organic chemical. Furfural itself has many applications, such as oil refining, as a bonding agent in grinding and abrasive wheels, in pharmaceuticals, and the manufacture of phenolic resins. FF has been addressed as one of the most important biomass-derived chemicals because the two functionalities (aldehyde group and the aromatic ring) that make it high chemical reactive. Biofuels and Furfural resins can be derived either directly from FF or from any of the chemicals obtained from it (see

figure 1.9). Furfuryl alcohol (FFOL) is the most important chemical derived from FF by hydrogenation and is primarily used to produce resins.

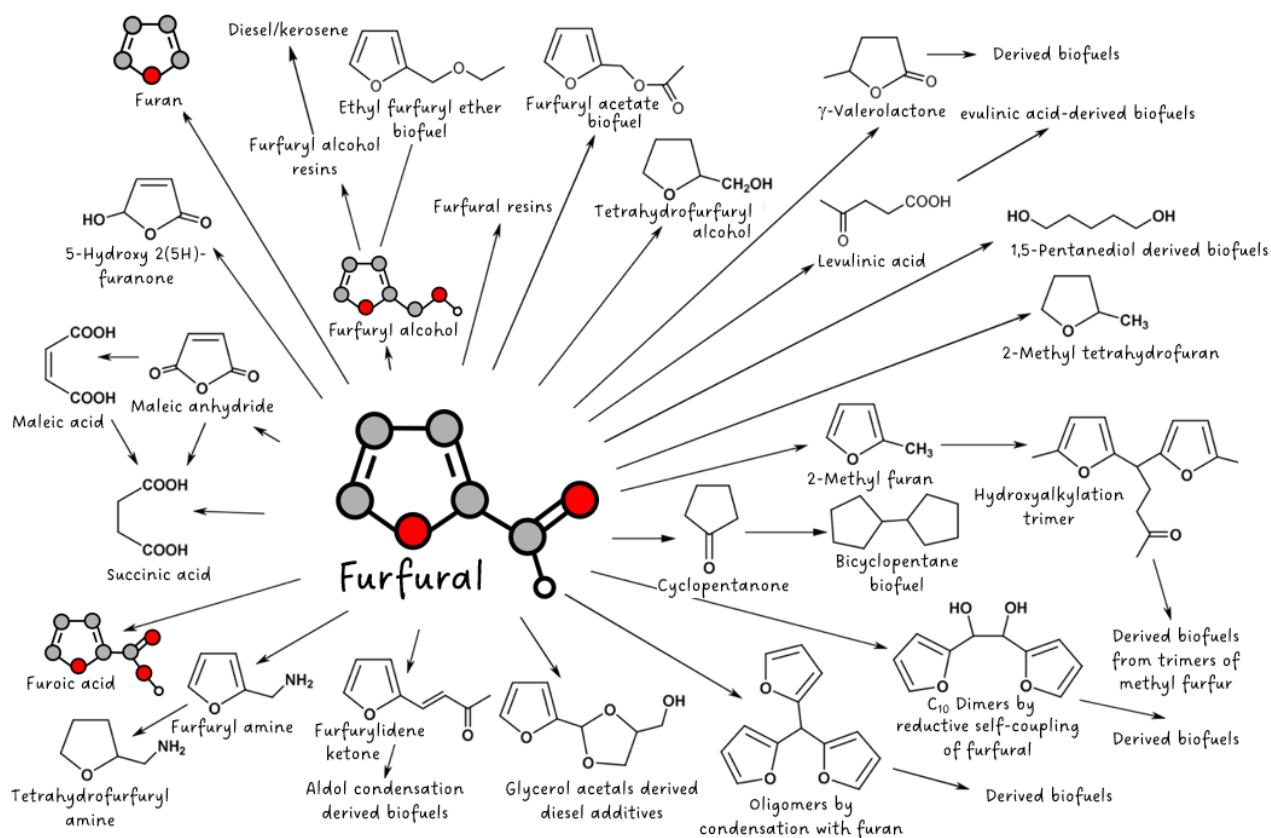


Figure 1.9 – Furfural-derived platform molecules and biofuels.⁶⁷

Separation and purification process steps would be required to obtain all the products of interest. The solvent selection used in the separation process design for biorefinery application plays an important role to define the environmental sustainability and economic viability of the process. In this framework, this thesis explores the progress and development of green solvents, specifically the Deep Eutectic Solvents, for the extraction of furanic compounds (FF, 5-hydroxymethylfurfural (HMF), and levulinic acid (LA)) produced by dehydration of fructose and xylose.

In recent years, DESs have emerged as promising media for the dehydration of hexoses (such as fructose) and pentoses (such as xylose) to furanic derivatives like 5-hydroxymethylfurfural (HMF) and furfural.^{69–73} These furanic derivatives are valuable compounds used in various industrial applications. Several strategies have been explored to effectively convert hexoses into HMF using DESs, shown in Table 1.5.

Table 1.5 – Catalytic dehydration of hexoses to HMF in various DESs.

DES	Catalyst	Carbohydrate	T (°C)	HMF Yield (%)	Reference
ChCl/urea	FeCl ₃ , AlCl ₃ , CrCl ₃	fructose	100	<30	74
ChCl/citric acid	-	fructose	80	70-80	
ChCl/citric acid	-	fructose	120	62	75
ChCl/oxalic acid	-	fructose	120	59	
ChCl/citric acid	-	inulin	80	51	76
ChCl/oxalic acid	-	inulin	80	56	
ChCl/p-TSA		fructose	80	90	77
ChCl/fructose		-	100	40-70	
ChCl/glucose	FeCl ₃ , CrCl ₂ , CrCl ₃ , p-TSA,	-	100	15-45	78
ChCl/inulin	Sc(OTf) ₃ , amberlyst-15	-	100	35-60	
ChCl/sucrose	montmorillonite K-10	-	100	25-60	
ChCl/fructose	(C ₆ H ₁₅ O ₂ N ₂) ₂ HPW ₁₂ O ₄₀	-	110	92	79
ChCl/fructose	CO ₂ (4 MPa)	-	120	73	80
ChCl/inulin	CO ₂ (4 MPa)	-	120	41	81

Hu and Ilgen^{74,78} studied the catalytic conversion of fructose to HMF within DESs, particularly ChCl and urea based. However, initial attempts showed low HMF yields, mainly due to side reactions between fructose and urea. Further research improved HMF yields by using DESs based on ChCl citric and oxalic acid. These variations significantly enhanced the yield of HMF, reaching up to 76% when the reaction was heated at 80°C for 1 hour. Similar success was achieved in promoting fructose dehydration to HMF using DESs composed of ChCl and p-toluene sulfonic acid (p-TSA).⁷⁷ Such DESs could dissolve fructose and facilitate its dehydration to HMF at 80°C with a 90% yield after 1 hour of reaction. These findings were especially promising when using a low content of fructose (2.5 wt%), with the HMF yield dropping as the fructose content increased.

Furthermore, researchers extended this concept to the production of furfural from xylan or xylose within DES, presented in Table 1.6.

Table 1.6 – Catalytic dehydration of pentoses to FF in various DESs.

DES	Catalyst	Carbohydrate	T (°C)	FF Yield (%)	Reference
ChCl/oxalic acid	AlCl ₃ ·6 H ₂ O	xylan	140	38	82
ChCl/ oxalic acid	AlCl ₃ ·6 H ₂ O	xylose	140	60	
ChCl/citric acid	AlCl ₃ ·6 H ₂ O	xylan	140	54	83
ChCl/citric acid	AlCl ₃ ·6 H ₂ O	xylose	140	59	
ChCl/formic acid	SnCl ₄ ·5H ₂ O, AlCl ₃ , CeCl ₃ ·7H ₂ O, ZrCl ₄	xylose	120	32-60	84
ChCl/ glycerol	water	xylose	120	60	85
ChCl/ ethylene glycol	water	xylose	180	76	86
ChCl/ malic acid	water	xylan	150	75	87

The reactions within ChCl /citric acid containing $\text{AlCl}_3 \cdot 6 \text{H}_2\text{O}$ conducted at $140 \text{ }^\circ\text{C}$ for 15 to 25 minutes resulted in furfural yields of 54% to 59% from xylan and xylose, respectively.⁸³

The use of DESs for biomass conversion offers significant advantages. They can produce highly concentrated solutions of carbohydrates, sometimes even liquid at near-room temperature. DESs' unique ability to stabilize furanic derivatives enhances the conversion of carbohydrates to HMF and furfural with high selectivity. However, extracting furanic derivatives from DESs remains a challenge due to unfavorable partition coefficients.

Despite this challenge, DESs offer a novel strategy to convert furanic derivatives in situ to chemicals that can be more easily extracted through distillation, precipitation, or liquid–liquid phase extraction for optimal utilization in the biorefinery industry.

1.2. Separation and purification process

Separation and purification processes in biorefineries typically use traditional technologies derivatives from the classic refinery (i.e., petrochemical industry), such as distillation, extraction, adsorption, absorption, membranes, crystallization, etc. Some of these techniques, such as distillation, can indeed be energy-intensive due to component vaporization, while others, like membrane filtration, can be less energy-intensive. While many of these traditional processes can be energy-intensive, it's important to note that not all of them rely solely on the vaporization enthalpy of chemical components. Some of these techniques may use other mechanisms for separation. Emerging technologies, as illustrated in Figure 1.10, are generally more energy-efficient, particularly when they avoid energy-intensive vaporization processes.

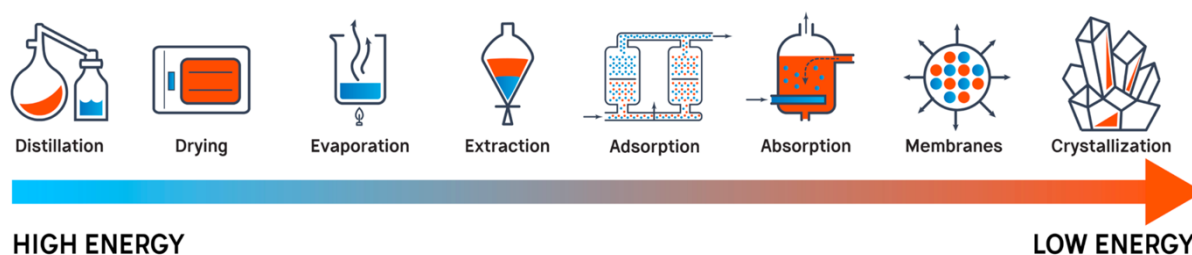


Figure 1.10. Examples of thermal separation processes (higher energy use) and nonthermal separation processes (lower energy use).⁸⁸

Many separations are still being performed by using distillation (reported to account for 80% of industrial separations), which is the most energy intensive. Distillation in the bio(refineries) can account for 40-50% of the total operation cost in the industry⁸⁹ and generate large waste streams that are expensive to manage.

This cost can represent in terms of energy consumption 10–15% of the world's energy consumption⁹⁰. For those reasons, a successful design, and operation improvement of these processes are economically critical to developing a sustainable chemical facility and can reduce energy use.

This section gives an overview of the separation processes in biorefineries, conveniently grouped here according to energy consumption.

1.2.1. Distillation

Distillation is a process for separating chemical compounds of a liquid mixture by boiling and collecting condensed vapors. Two prerequisites must be satisfied simultaneously to ensure that distillation can take place: the relative volatility as a driving force and the separation equipment. In the biorefinery, the distillation processes are used for the separation and dehydration of alcohols (bioethanol and biobutanol), biodiesel purification, isolation of volatility organic compounds (essential oils) and phytochemicals from biomass (extract), the concentration of chemicals in pyrolysis oil and separated in various fractions (alcohols, aldehydes, ketones, acids, phenolics, and sugar).

1.2.2. Liquid-liquid extraction

Liquid-liquid extraction, also known as solvent extraction, is a common method used in chemistry and industry. Its main purpose is the separation of a solute from a mixture. In this process, the solute can be originally dissolved in a selective way in a solvent, the **diluent**. Subsequently, the solute can be transferred into another solvent, the **extractant**, to accomplish the separation. This process typically involves three steps. First, an immiscible or partially miscible extractant is added to the diluent, allowing the solute to dissolve in the extractant. The second step involves thorough mixing of these two solvents to facilitate the transfer of the solute from the diluent to the extractant phase. The ability of a solute to preferentially dissolve in one of two immiscible or partially miscible solvents (the extractant and diluent) based on differences in its solubility in these solvents is called by selectivity. The final step involves layering the two solvents, allowing for the physical separation of the solute by isolating the extractant from the diluent. After this separation, the solute can be recovered from the extractant using various methods such as distillation or simple evaporation. In the biorefinery, Liquid-liquid extraction is used for separating biofuels (bioalcohols) and chemicals (carboxylic acid) from dilute mixtures (fermentation broths), extraction of acetic acid from biomass hydrolysates using mixed solvents, extraction of HMF from an aqueous reaction solution using methyl isobutyl ketone as a solvent, chemicals extraction (aqueous extraction, or extraction using hydrophobic-polar solvent) of fast pyrolysis bio-oils.²⁶ One of the main advantages of solvent

extraction is the recovery of components that are boiling at high temperatures than water boiling temperature.

1.2.3. Other separation processes

Other separation processes can be used in biorefineries, such as adsorption, absorption, and membrane separations.

Adsorption is a technique where solid adsorbents, such as activated carbon or zeolites, are used to selectively capture specific components from a mixture. It is effective as purification method, for removing impurities and can be employed in biorefineries to separate and purify valuable compounds. In the biodiesel production process, adsorption can play a crucial role in removing impurities from the raw biodiesel. This includes the removal of glycerol, methanol, free fat acids, soap, catalyst, metals, water, and glycerides. These impurities must be effectively eliminated to meet biodiesel quality standards and ensure the fuel's optimal performance.

Absorption involves the transfer of one or more components from a gas phase into a liquid phase. This process is often used for gas purification and the capture of volatile compounds, such as CO₂, from biogas or fermentation off-gases. In the production of methanol and diesel fuel from syngas the presence of impurities such as H₂S, CO₂, and NH₃ can poison catalysts and negatively impact product quality. Absorption units are utilized to selectively capture these impurities, ensuring a cleaner syngas feed for further conversion process.

Membranes separation processes, including microfiltration, ultrafiltration, nanofiltration, and reverse osmosis, rely on semi-permeable membranes to separate components based on size, charge, or other properties. These processes are widely used in biorefineries for the separation of proteins, enzymes, and other biomolecules, as well as the concentration and purification of bio-based products. Moreover, membrane separation finds specific applications in biorefineries such as the separation of hemicelluloses from biomass hydrolysates, lignin recovery, biomass pre-hydrolysis, and the separation and purification of biodiesel, exemplifying its versatility and importance in the bio-based industry.

1.3. Solvent selection

The solvent selection is the first and basic consideration for achieving an optimal liquid-liquid extraction.⁹¹

1.3.1. Physico-chemical properties

Solubility data is a critical piece of information that guides the selection of a suitable solvent for a given application. It enables engineers and chemists to make informed decisions about the solvent's ability to dissolve the solute effectively and achieve the desired process outcomes. In simpler terms, solubility represents the ability of a solvent to dissolve a given solute or substance efficiently, which is fundamental in various chemical and industrial processes.

Dynamic and kinematic viscosity influences the transfers of mass, heat, and momentum within fluids, significantly impacting these processes. High-viscosity solvents have higher viscosity values, which directly affects the Reynolds number (Re) through its presence in the Re equation. The increased viscosity results in greater viscous resistance, making it more challenging to move the fluid, which requires more energy for mixing, pumping, and maintaining desired flow rates. Furthermore, higher viscosity can lead to changes in the mass transfer rate, affecting the efficiency of processes like mixing and mass transport within the fluid. In the domain of fluid dynamics and mass transfer analyses, Sherwood number (Sh) is a relationship commonly used to describe the mass transfer rate. It is expressed as $Sh = f(Re, Sc)$, with Sc representing the Schmidt number. The Schmidt number is a measure of the relative importance of momentum transport and depends on the viscosity, density, and diffusivity of the solute in the fluid.

Density can influence phase separation in extraction or separation processes. Differences in density between solvent phases can enhance or hinder separation.

Interfacial tension also can influence mass transfer properties, in which low values can allow the disruption of solvent droplets with low agitation and is also associated with emulsion formation in two-phase liquid systems.

Boiling point defines the physical state of the solvent at the temperature used for the extraction. The boiling point should be suitable for the process conditions. Solvents with low boiling points are easily removed by distillation, while high-boiling solvents may require more energy for recovery.

Flash Point represents the minimum temperature at which the solvent's vapors can ignite when exposed to an open flame or spark. A low flash point could pose fire hazards. Choosing solvents with appropriate flash points helps prevent fire hazards and ensures the safety of personnel and the environment.

Chemical compatibility ensures that the solvent is chemically compatible with the solute and other materials in the process, such as equipment and catalysts. By choosing a chemically compatible solvent,

engineers and chemists can avoid potential complications, unwanted side reactions that could degrade the solvent, the generation of impurities, and oxidation reactions. Solvents with minimal or no reactivity with the solute are highly desired to maintain the purity and integrity of the intended chemical or industrial process, thereby ensuring the success of the desired process and the quality of the final product.

Polarity can enhance the efficiency and selectivity of the separation process by matching the polarity of the solvent to the solute. Polar solvents dissolve polar compounds, while nonpolar solvents are better suited for nonpolar compounds. Some processes may require a solvent with intermediate polarity.

Toxicity of the solvent is another relevant aspect that is considered in solvent selection to ensure the safety of workers, protect the environment, comply with regulations, and promote responsible and sustainable chemical practices. Substituting hazardous solvents with less harmful alternatives is a key strategy in enhancing safety and minimizing health and environmental risks.

1.3.2. Economic and environmental aspects

The choice of solvent could considerably reduce the overall energy requirements and the separation process costs. Solvent recyclability can significantly impact the overall cost and sustainability of a process. Alternative solvents, as a more environmentally friendly substitute to petrochemical solvents, have attracted interest from researchers due to a growing awareness of the impact of solvents on pollution, energy usage, and contributions to air quality and climate change. Green solvents represent a significant portion of research in alternative solvents. Providing benefits in terms of resource efficiency, minimal health & safety concerns, and being environmentally friendly.

1.3.3. Efficacy

The Activity coefficient at Infinite Dilution (IDAC, γ_i^∞) is a parameter that describes the behavior of one solute molecule surrounded by solvent molecules, showing the higher nonideality of the mixture. Therefore, is an important parameter used in chemical engineering to design of the separation process, helping the selection of the solvent by its capacity and selectivity. The selectivity is represented by the solvation power^{92,93}, S_p , and can be obtained by

$$S_p = \frac{1}{\gamma_i^\infty} \frac{M_i}{M_{solv}} \quad (1.1)$$

where M_i and M_{solv} are the molar mass of solute and solvent, respectively.

Solvation power has been used in the development of the methodology to select solvents, here in this work.

1.4. Deep eutectic solvents

Deep Eutectic Solvents (DESs) have emerged as a groundbreaking class of solvents, characterized by their formation from a mix of two or more solid or liquid components. Typically, this mix involves a hydrogen bond donor (HBD), such as a quaternary ammonium salt, and a hydrogen bond acceptor (HBA), like a metal chloride. When these components are combined in precise proportions, they form a eutectic mixture that remains liquid at temperatures often below the melting points of the individual components. This unique characteristic has positioned DESs as sustainable and environmentally friendly alternatives to traditional organic solvents and ionic liquids (IL), heralding a new era in green chemistry.

The literature investigation into DESs, although relatively nascent compared to classical ILs research, has seen a significant surge in interest. The Scopus search reveals approximately 6,000 journal articles on DESs, indicating a robust and growing body of research since the first paper was published in 2001. This literature spans a diverse array of fields, including chemistry, chemical engineering, physics and astronomy, engineering, material science and computer science, reflecting the interdisciplinary appeal of DESs (see figure 1.11). Notably, the past decade has witnessed an exponential growth in DES publications, with research efforts concentrated on metal processing and synthesis media. DESs have been identified as promising alternative media for metals that are traditionally challenging to plate or process and for synthesis applications where environmentally hazardous processes are involved.

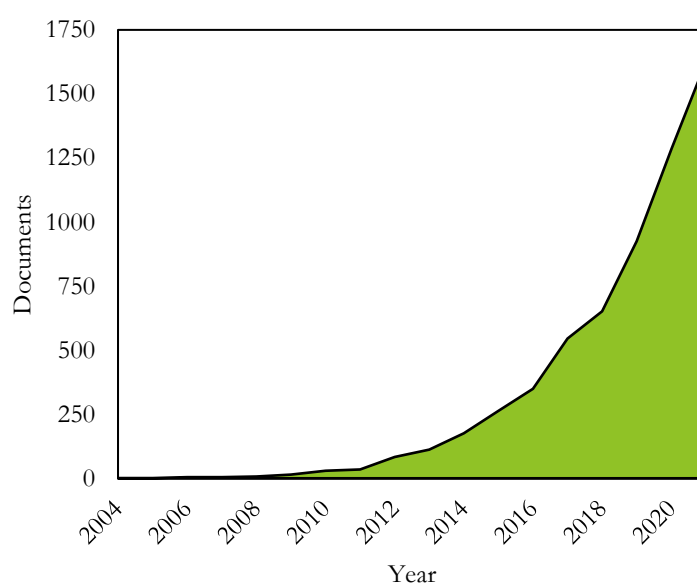


Figure 1.11 -. Number of publications with DES as key word per year.

In the investigation of DES, exemplified by the Reline system comprising choline chloride as a hydrogen bond acceptor (HBA) and urea as a hydrogen bond donor (HBD) in a 1:2 ratio, the phase diagram assumes a central role to elucidate their unique characteristics. This phase diagram presented in the Figure 1.12 offers a pivotal depiction of the intricate interplay between temperature and composition within the mixture. At the heart of this representation lies the eutectic point, a critical feature that captures the conditions under which the DES undergoes a phase transition. The chosen composition, with choline chloride and urea in a 1:2 ratio, gives rise to the specific DES named Reline.

The phase diagram presented in the Figure 1.12 serves as a comprehensive tool, elucidating both ideal and real scenarios. In the ideal case (gray line), the eutectic point represents the composition at which the mixture achieves its theoretically melting point. The hydrogen bonding interactions, with choline chloride acting as the hydrogen bond acceptor and urea as the hydrogen bond donor, play a pivotal role in charge delocalization, contributing to the reduction in the melting point of the mixture.

Conversely, real-world scenarios (black line) introduce complexities, such as non-ideal behavior, which deviate from the idealized case. The real eutectic point, reflecting the actual conditions at which the DES undergoes a phase transition. This nuanced understanding, coupled with the distinctive characteristics of the Reline system, is crucial for tailoring DES formulations to specific applications, contributing to advancements in the effective utilization of Deep Eutectic Solvents across diverse fields.

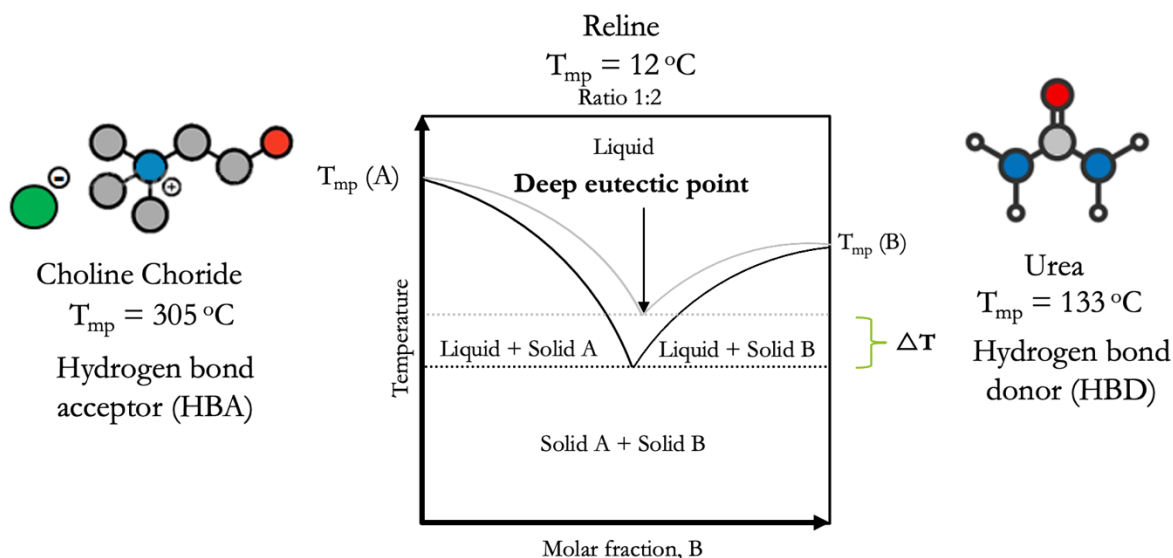


Figure 1.12 – Choline Chloride:Urea (1:2) phase diagram.

DESs represent a pivotal innovation in green chemistry, offering a sustainable alternative to traditional solvents. Common DES formulation, such as the mixture of choline chloride with urea, ethylene glycol,

and glycerol, exemplify the versatility and environmental compatibility of the solvents. Their applications span a vast array of fields, from green chemistry to pharmaceutical and nutraceutical industries, and even into the realms of metal processing and biomass conversion.

In the context of green chemistry, DESs are celebrated for their non-toxic and biodegradable nature, providing an eco-friendly medium for various chemical reactions, catalysis, and extraction processes. For instance, DESs have been employed in extracting bioactive compounds from natural sources such as plants, algae, and marine organisms, showcasing their potential in pharmaceutical and nutraceutical industries.

In metal processing, DESs offer considerable advantages over conventional solvents. Their application in metal plating, for instance, has demonstrated not only improved plating quality but also a reduction in the environmental impacts typically associated with such processes. This aligns with the broader goals of reducing hazardous waste and enhancing sustainability within the industry.

Furthermore, the role of DESs in biomass processing and conversion is increasingly recognized. Their effectiveness in biomass pretreatment, cellulose dissolution, and catalytic conversion processes contributes significantly to the development of sustainable bio-based industries. This includes facilitating more efficient and environmentally friendly methods for converting biomass into valuable chemicals and fuels.

The versatility and potential of DESs across these diverse applications underscore their capacity to driving innovation towards more sustainable and eco-friendly practices in various sectors. As research and development in the field of DESs continue to expand, their integration into industrial process is poised to offer significant advancements in environmental sustainability and green chemistry.

1.4.1. Physico-chemicals properties

Physical characteristics of DES can vary depending on their composition and the specific components used to create them. However, some common physical characteristics and properties of DES include:

Lower Melting Point compared to the individual components used to create them. This property allows them to remain in liquid form at relatively low temperatures.

Viscosity can vary depending on their composition. Some DESs have low viscosity, like that of traditional organic solvents, while others may have higher viscosities.

Solubility of specific substances in a DES may vary depending on the nature of the solute and the DES composition. DESs can dissolve a wide range of organic and inorganic compounds.

Polarity can have variable polarity depending on the choice of components. This tunability of polarity makes them suitable for a range of applications, from polar to nonpolar solvents.

Density can vary depending on the composition, and it may be different from that of individual components.

Thermal stability at moderate temperatures, although this can vary depending on the specific components used.

Environmentally Friendly than traditional solvents due to their lower toxicity, reduced volatility, and potential for recycling.

Refractive Index of a DES may vary depending on its composition and can affect optical properties and applications.

Interesting Phase Behavior, including the ability to form biphasic systems with water or other solvents, which is useful in certain extraction and separation processes. (some DESs can exhibit it).

1.4.2. Applying DES in a processes

When considering the use of a DES in a process, it's crucial to gather the following important properties and information. This data serves as the foundational framework for informed decision-making and optimized procedural outcomes.

First and foremost, an understanding of the chemical composition of the DES is important. These solvents are typically composed of a hydrogen bond donor (HBD) and a hydrogen bond acceptor (HBA). To establish a robust process design, it is essential to discern the precise identity of these components and their respective molar ratios. Additionally, the eutectic composition of the DES is of great significance. This represents the specific ratio of HBD to HBA that yields the lowest melting point and is often critical for achieving the desired properties.

Thermal properties, such as stability, melting point, and boiling point must be gathered to delineate the permissible operating temperature range of the process. Understanding the solubility and selectivity of the DES is crucial, particularly in relation to the targeted solutes or reactants. DESs exhibit unique solvation

properties, necessitating a thorough comprehension of the compounds amenable to effective dissolution or reaction.

Physico-chemicals properties, such as viscosity and density are also important. The viscosity of the DES should be measured as it can significantly influence mixing, flow dynamics, and handling throughout the process. Elevated viscosity may necessitate specialized equipment or adjustments to process parameters. Density, too, warrants determination as it can impact volume measurements, especially when transitioning to larger-scale processes.

DESs are not inherently hygroscopic. Whether a particular DES exhibits hygroscopic properties depends on its specific chemical composition. Some DESs may contain components that are hygroscopic, while others may not. Hygroscopicity is generally associated with substances that have a strong affinity for water and readily absorb moisture from the surrounding environment. If a DES includes hygroscopic components, it may exhibit hygroscopic behavior to some extent. However, the overall hygroscopicity of a DES will depend on the balance of its constituent components and their interactions. It's essential to assess the hygroscopic properties of a specific DES by considering its chemical composition and conducting experiments or tests to determine its moisture absorption characteristics under the relevant environmental conditions.

Investigate the compatibility of the DES with the reactants or compounds integral to the process. Ensuring that reactions progress as anticipated and that undesirable side reactions are averted is paramount.

The environmental implications of DES usage must be considered, encompassing aspects such as biodegradability and toxicity. Compliance with environmental regulations may be mandatory. The acquisition of safety data is essential for a comprehensive risk assessment. Information pertaining to flammability, toxicity, and safe handling procedures should be obtained, with the implementation of suitable safety measures. Verify compliance with regulatory requirements and any restrictions pertinent to DES utilization within the designated application. This is particularly pertinent in industries subject to rigorous regulations such as food and pharmaceuticals.

Conducting a comprehensive cost analysis is advisable to gauge the economic feasibility of employing DESs in the process. This entails factoring in expenses related to raw materials, recycling, and disposal.

When transitioning from lab-scale to industrial-scale processes, evaluate how the properties and behavior of the DES may evolve at larger volumes. This assessment is essential to ensure seamless scalability.

Finally, preliminary experiments or lab-scale trials should be conducted to validate the suitability of the DES for the specific process, providing valuable insights into its performance characteristics.

1.5. Conclusion

In summary, this chapter emphasizes the significant dependence on crude oil for energy and highlights the need to explore sustainable alternatives for energy production. Biomass feedstocks have emerged as a promising renewable and sustainable option. The chapter has elucidated the essence of the biorefinery concept, presenting its potential advantages while acknowledging the limitations associated with various biomass feedstocks.

Lignocellulosic biomass, as a pivotal component, has been thoroughly discussed, highlighting its central role in transitioning to sustainable energy resources. The discussion extensively covered the essential separation and purification processes necessary for effective biomass conversion. Traditional technologies rooted in conventional refinery practices were introduced, focusing notably on the extraction process.

The chapter emphasizes solvent selection as a critical factor in achieving optimal liquid-liquid extraction, addressing key considerations such as physico-chemical properties, economic aspects, and selectivity. Additionally, focused attention was given to green solvents, particularly deep eutectic solvents (DES), revealing their substantial potential through a thorough examination of their physico-chemical properties. In essence, focusing on the physico-chemical properties of Deep Eutectic Solvents (DES), it's important to highlight that the design and optimization of the separation process rely on accurate and accessible thermodynamic and transport properties data. However, a significant challenge arises from the lack of experimental data for DES systems. A critical objective of this thesis is to address this scarcity by generating experimental data to refine and optimize theoretical models. The subsequent chapter will comprehensively describe the experimental methods applied to achieve this objective, providing valuable insights into the methodologies utilized.

In conclusion, this work represents a pioneering initiative within our research group, utilizing the expertise and technology available at the CTP lab to explore the realm of DES. Overcoming the challenges posed by the high viscosity properties inherent to DES, adapting laboratory equipment for testing purposes was a significant obstacle that we effectively addressed. This study can serve as a valuable contribution advancing knowledge in this field, highlighting the significance of innovation in addressing emerging scientific challenges. Moving forward, our findings provide a solid foundation for further exploration and advancement in the understanding and application of deep eutectic solvents.

Chapter 2

Experimental study

Étude expérimentale

Résumé

Le manque de données thermodynamiques sur les mélanges contenant des solvants eutectiques profonds (DES) constitue un défi important. La complexité des compositions des DES ainsi que leurs propriétés contribuent à ce problème. Les DES sont généralement formés en mélangeant deux ou plusieurs composants qui peuvent inclure un donneur de liaison hydrogène (HBD) et un accepteur de liaison hydrogène (HBA). Ces composants peuvent varier en type, concentration et ratio conduisant un grand nombre de formulations DES possibles. Chaque combinaison peut présenter des comportements thermodynamiques différents. La complexité des compositions DES rend impraticable la réalisation de mesures expérimentales pour chaque combinaison possible.

Pour résoudre partiellement ce problème et combler le manque d'information, les chercheurs ont employé des approches théoriques telles que les simulations moléculaires et les modèles de contribution de groupe pour prédire le comportement de phase des systèmes d'intérêt. Cependant, il est important de souligner que la fiabilité de ces prédictions dépend de la similarité du système avec la base de données utilisée pour ajuster les nombreux paramètres d'interaction de groupe. En d'autres termes, plus le système que l'on étudie est semblable à la base de données qui ont servi à paramétrer le modèle, plus les prédictions sont fiables. D'un autre côté, ne s'appuyant que sur des méthodes de prédiction (telles que la théorie de la fonctionnelle de la densité (DFI) ou une équation d'état) sans vérification expérimentale peut conduire à une conception peu fiable qui peut être coûteuse pour l'industrie chimique.

Une partie importante de cette thèse est dédiée à la génération des données expérimentales essentielles pour optimiser les modèles théoriques. À cette fin, nous avons choisi trois DES en basant sur leur application pour la déshydratation d'hexoses (comme le fructose) et de pentoses (comme la xylose) en dérivés furaniques tels que le 5-hydroxyméthylfurfural (HMF) et le furfural.

Une partie importante de cette thèse est dédiée à la génération de données expérimentales essentielles. À cette fin, nous avons choisi

Ces DES sont:

- Chlorure de choline : Éthylène Glycol (1 :2), également connu sous le nom en anglais « 'Ethaline ».

- Chlorure de choline : Glycérine (1 :2), également appelé en anglais « Glycaline ».
- Chlorure de choline : Fructose (2 :1), connu sous le nom en anglais de « Fructaline ».

Fructaline a fait l'objet d'études de la part de certains auteurs dans le processus de déshydratation du fructose en DMF. De plus, Ethylene et Glycaline ont été examinés pour leur rôle dans la déshydratation de la xylose en furfural. Ces DES ont montré des résultats prometteurs pour faciliter la conversion des sucres en composés furaniques et leur potentiel d'utilisation dans ces processus de déshydratation est exploré par divers chercheurs.

Ce chapitre présente les aspects pratiques de ce travail. Nous décrivons le protocole de synthèse des DES ainsi que l'équipement et les protocoles expérimentaux utilisés dans ce travail pour les caractériser les DES ainsi que leurs mélanges. Couvrant des propriétés clés telles que la masse volumique, la viscosité dynamique, l'indice de réfraction, et notamment, l'acquisition de données de coefficient d'activité à dilution infinie par « gas stripping » ainsi que de données d'équilibre liquide-vapeur par une méthode synthétique.

Experimental study

2.1. Introduction

A successful design and optimization of separation processes in the chemical industry rely on the availability and accuracy of thermodynamic and transport properties data. These two aspects are fundamental for the development of thermodynamic models.

Thermodynamic data, such as Solid-Liquid Equilibrium (SLE), Vapor-Liquid Equilibrium (VLE), Liquid-Liquid Equilibrium (LLE), Vapor-Liquid-Liquid Equilibrium (VLLE), infinite dilution activity coefficients (IDAC) data, provide essential information about the composition and phase behavior of mixtures at different conditions. This data is crucial for designing separation processes like distillation, liquid-liquid extraction, and crystallization. Specific heat capacity (C_p) data helps in understanding how a substance's temperature changes as heat is added or removed. It's vital for calculating the heat requirements or heat removal in separation processes, especially in processes involving temperature changes, like distillation and heat exchangers. Furthermore, the enthalpy of mixing is crucial for understanding the energy changes that occurs when different components are mixed. This data is essential in processes like mixing, dissolution, and crystallization. Enthalpy of Vaporization quantifies the amount of energy required to convert a liquid into vapor at given temperature and pressure. It is a fundamental property in vaporization processes, such as distillation and evaporation.

Transport properties, such as density, diffusion coefficients, and viscosity, play an important role in the design of separation processes, such as the sizing of distillation columns, reactors, pumps, and pipes. Accurate data are necessary to ensure that equipment operates as intended and meets safety and environmental standards.

One of the difficulties often faced by chemical and process engineers is the scarcity of experimental data in the literature for specific systems. While these data are easily accessible for conventional solvents, and well known in the literature, for unconventional solvents the databases are limited. Moreover, even when experimental data are available, it often lacks important information about the experimental procedure applied, uncertainty, and consistency. These details are critical for assessing the reliability of the data and for comparing them with data from other sources.

The lack of thermodynamic data on mixtures containing DES is a significant challenge, and the complexity of DES compositions, as well as their unique properties, contribute to this issue. DESs are typically formed by mixing two or more components, which can include a hydrogen bond donor (HBD) and a hydrogen bond acceptor (HBA). These components can vary in type, concentration, and ratio, leading to a vast number of possible DES formulations. Each combination may exhibit different thermodynamic behaviors. The complexity of DES compositions makes it impractical to conduct experimental measurements for every possible combination. Additionally, DESs are known for their intricate network of hydrogen bonds and other non-covalent interactions. These interactions can lead to the formation of various phases, including solid precipitates or multiple liquid phases, depending on the composition and conditions. Predicting the thermodynamic behavior of multicomponent DES mixtures requires a detailed understanding of these interactions, which can be challenging to model accurately. Furthermore, DESs often have higher viscosities compared to conventional solvents, which can impede the measurement of thermodynamic properties. High viscosity can affect the accuracy of experimental techniques and may require specialized equipment and methodologies.

To (partially) resolve this problem and fill the information gap, industries have employed theoretical approaches, such as molecular simulations and group contribution models, to predict phase behavior for systems of interest, and more details will be explained in chapter 3. However, it is important to highlight that the reliability of these predictions depends on the similarity of the system with the database used to adjust hundreds of group interaction parameters. On the other hand, relying only on prediction methods, such as density functional theory (DFT) or equation of state, without experimental verification can lead to an unreliable design which can be costly for the chemical industry concerned.

One important part of this thesis is dedicated to bridging the gap between theory and reality by generating essential experimental data to optimize theoretical models. This chapter presents the practical aspects of this work. Within this chapter, we will describe the synthesis protocol for DES. Additionally, we will describe the experimental equipment and protocols used in this work to characterize them and their mixtures. Covering key properties such as density, dynamic viscosity, refractive index, and notably, the acquisition of activity coefficient at infinite dilution data through gas stripping, and vapor-liquid equilibrium data via a synthetic method.

Three DES were selected based on the recent tendency of their application for the dehydration of hexoses (such as fructose) and pentoses (such as xylose) to furanic derivatives like 5-hydroxymethylfurfural (HMF) and furfural.

These DES are:

- Choline chloride: Ethylene Glycol (1:2), also known as Ethaline.
- Choline chloride: Glycerol (1:2), also referred to as Glycaline.
- Choline chloride: Fructose (2:1), known as Fructaline.



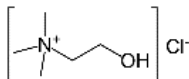
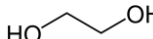
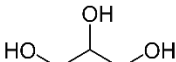
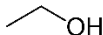

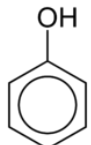

Fructaline has been the subject of investigation by certain researchers for its role in the dehydration of fructose in dimethylformamide (DMF). Furthermore, Ethaline and Glycaline have been explored for their potential in the dehydration of xylose into furfural.

Fructaline has been investigated by some researchers for its effectiveness in dehydrating fructose in HMF.^{79–81} Additionally, Ethaline⁸⁶ and Glycaline⁸⁵ have been explored for the dehydration of xylose in furfural. These DES have demonstrated promise in facilitating the conversion of sugars into valuable furanic compounds, and their potential applications in these dehydration processes are being explored by various researchers.^{69–73}

2.2. Materials and methods

All the chemicals used during this work are listed in table 2.1.

Table 2.1 – Chemical used for experimental work.

Chemical name	Formula	CAS	Molar Mass (g/mol)	Purity (GC) (%wt)	Supplier	Refractive index	
						Measured ^a at 298.15 K and 101.325 kPa	Lit. ^b
Furan (stabilized)		110-00-9	68.08	≥ 99.0	SIGMA ALDRIC H	1.4217	1.417 8
Cyclohexane		110-82-7	84.16	≥ 99.0	SIGMA ALDRIC H	1.4263	1.423 5
Choline Chloride		67-48-1	139.62	≥ 99.0	SIGMA ALDRIC H	-	-
Ethylene Glycol		107-21-1	62.07	≥ 99.0	SIGMA ALDRIC H	1.4428	1.438 5
Glycerol		56-81-5	92.09	≥ 99.0	SIGMA ALDRIC H	1.4705	1.472 2
Ethanol		64-17-5	46.07	≥ 99.0	SIGMA ALDRIC H	1.3745	1.361 4
Isopropanol		67-63-0	60.10	≥ 99.0	SIGMA ALDRIC H	1.3826	1.377 2
Phenol		108-95-2	94.11	≥ 99.0	SIGMA ALDRIC H	-	-
n-Hexane		101-54-3	86.18	≥ 99.0	Alfa Aesar	1.3725	1.372 3

GC: Gas Chromatograph

^a Apparatus: Anton Paar ABBEMAT 300 Refractometer ($\lambda=589$ nm). Standard uncertainties: refractive index $u(n_D) = 6 \times 10^{-5}$, temperature $u(T) = 0.02$ K and pressure $u(P) = 0.029$ kPa)

^b Value from Simulis® thermodynamics⁹⁵.

In this work, two DESs based on choline chloride have been studied. The physical and thermodynamic properties of two pure DES have been performed and summarized in table 2.2.

Table 2.2 – Pure properties measurement.

DES	Short name	Measurement	Technique/Apparatus	T (K)
Choline chloride: Ethylene Glycol (1:2)	Ethaline	Density	Oscillating U-tube densimeter	293 - 333
		Viscosity	Rolling ball viscosimeter	293 - 333
Choline chloride: Glycerol (1:2)	Glycaline	Vapor pressure	Synthetic method	303 - 333

Therefore, VLE, LLE, SLE and IDAC measurements of pseudo binary mixture have been studied and are summarized in table 2.3. To prepare the mixtures, an empty 20 cm³ glass bottle is air-tight closed with a septum and then put under vacuum using a vacuum pump wherein a needle is introduced through the septum. The empty bottle is weighed, and then the less volatile component, freshly degassed, is introduced by means of a syringe. After weighing the bottle loaded with the first component, the more volatile one is added similarly and then the bottle is weighed again. All weightings are performed using an analytical balance with 10⁻⁴ g accuracy and hence the uncertainty is estimated to be lower than 2x10⁻⁵ for mole ratio.

Table 2.3 – Pseudo binary mixture properties measurement.

Solvent	Solute	Measurement	Technique/Apparatus	condition
Ethaline	Ethanol	Density	Oscillating U-tube densimeter	393 K – 333 K
		Viscosity	Rolling ball viscosimeter	293 K – 333 K
	Isopropanol	Phase equilibrium (VLE)	Synthetic method	303 K – 333 K
Glycaline	Ethanol	Density	Oscillating U-tube densimeter	303 K – 333 K
		Viscosity	Rolling ball viscosimeter	303 K – 333 K
	Isopropanol	Phase equilibrium (VLE)	Synthetic method	303 K – 333 K
Fructaline		Gas stripping (IDAC)	Exponential Dilutor technique	303 K – 333 K

In the following sections, the different experimental techniques used in this work are described.

2.3. Preparation of Deep Eutectic Solvent

The preparation process for mixing the DES based on choline chloride is described in detail, emphasizing the critical steps and conditions for achieving a high-quality solvent.

2.3.1. Pre-treatment

Before beginning the DES synthesis, it is essential to pre-treat hygroscopic solids, such as choline chloride, to remove any water content. This is achieved by drying the solids at 110°C overnight to make it water free. Alternatively, a shorter drying period of at least 6 hours at 80°C may be used, although the overnight drying at 110°C is the preferred method for achieving optimal dryness.

2.3.2. DES Preparation

The DES mixing process begins with the dried choline chloride and the chosen hydrogen-bond donor (HBD) component. The mixing process is done by combining the two components in the desired ratio based on the specific DES formulation, as presented in table 2.4 and figure 2.1. The synthesis should be conducted under the following conditions:

- Maintain continuous stirring at approximately 400 revolutions per minute (rpm). This ensures the complete mixing of the components during the entire process.
- The temperature for the synthesis should be set at 80°C, which is a critical parameter for achieving the desired eutectic mixture.
- Continue stirring and heating for a duration of 6 hours or until a homogeneous and stable liquid is formed.

The formation of a homogeneous liquid with no visible precipitates is a key indicator of a successful DES synthesis.

Table. 2.4 – DES nomenclature.

HBA	HBD	Molar ratio	Abbreviation	Name
Choline Chloride	Fructose	2:1	ChCl-Fru	Frutaline
Choline Chloride	Glycerol	1:2	ChCl-Gly	Glycaline
Choline Chloride	Ethylene glycol	1:2	ChCl-EG	Ethaline



Figure 2.1. Frutaline, Glycaline and Ethalyne.

2.4. Mixture preparation

This section provides a detailed procedure for preparing mixtures of DES and solvents to ensure precise and accurate measurements of the components' weights and mole fractions. All the mixture were prepared in the empty 20 cm³ glass bottle airtight sealed with a septum. Before introducing the components, the glass bottle was put under vacuum using a vacuum pump wherein a needle is introduced through the septum, to eliminate any traces of air or gas. The empty bottle was weighed, and the initial weight was recorded. Then, the less volatile component, freshly degassed, was introduced using a syringe. The bottle was weighed again with the less volatile component added and this weight was recorded. Similarly, the more volatile component was introduced into the bottle using a syringe. The bottle was weighted once more with both components now present. Like the previous component, it was ensured that it was also freshly degassed. All weight measurements were conducted using a high-precision analytical balance, providing an accuracy of 10⁻⁴ g accuracy. Correspondingly, the uncertainty in determining the mole fractions was estimated to be less than 2x10⁻⁵ (details are presented in appendix A).

2.5. Characterization of Deep Eutectic Solvent and mixtures

2.5.1. Density

The density, ρ , was measured as a function of temperature with a digital density meter known as the Anton Paar DSA 5000M,⁹⁶ which operates based on the oscillating U-tube principle. The heart of the equipment, represented in figure 2.2 is composed of a gas cylinder filled with gas, a piezo element, a reference oscillator, a U-shaped borosilicate glass tube, and optical pick-ups. To measure the density, the samples were inserted into a U-shaped tube. The tube was then electronically excited to oscillate at its characteristic frequency, causing it to move up and down. During this oscillation, optical pick-ups detected and recorded period of these oscillations.

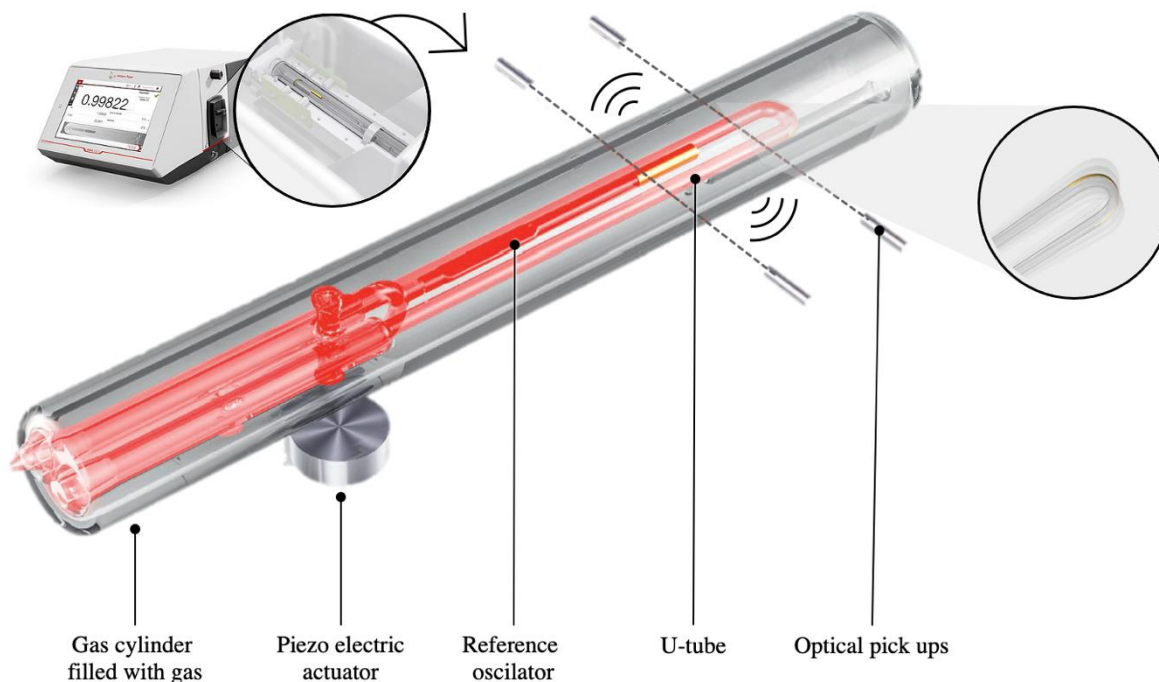


Figure 2.2 – Digital density meter: Anton Paar DSA 5000M (modified from Anton Paar⁹⁷).

The oscillating period or frequency changes depending on the density of the filled sample. This data was used to calculate the density of the samples at various temperatures. Eq. 2.1 correlate the density, ρ , of the fluid with the corresponding oscillating period, τ (1 divided by frequency of oscillation, f).

$$\rho = A + B\tau^2 \quad (2.1)$$

where A and B are constants to be adjusted. A calibration has been done by using two reference fluids with known densities: bi-distilled and degassed water, and dry air at 293.15 K.

Uncertainty on measured density is estimated to be lower than $10^{-5} \text{ g.cm}^{-3}$ after calibration. One platinum resistance thermometer with 0.01 K accuracy is inside of the equipment for temperature measurements. The sample densities are then measured at thermal equilibrium for various temperatures (293 K to 363 K). This density meter permits also the measurement of the speed of sound with an accuracy of 0.5 m.s^{-1} .

2.5.2. Viscosity

The viscosity is an important property in the design of separation process, as it provides insights into the thickness and flow behavior of mixtures. To obtain viscosity data, a rolling ball viscosimeter, specifically the Anton Paar LOVIS 2000 ME,⁹⁸ was used. Figure 2.3 illustrates the operational principle of a falling

ball viscometer. This instrument allows for the accurate determination of viscosity by measuring the rate at which a ball falls through a sample.

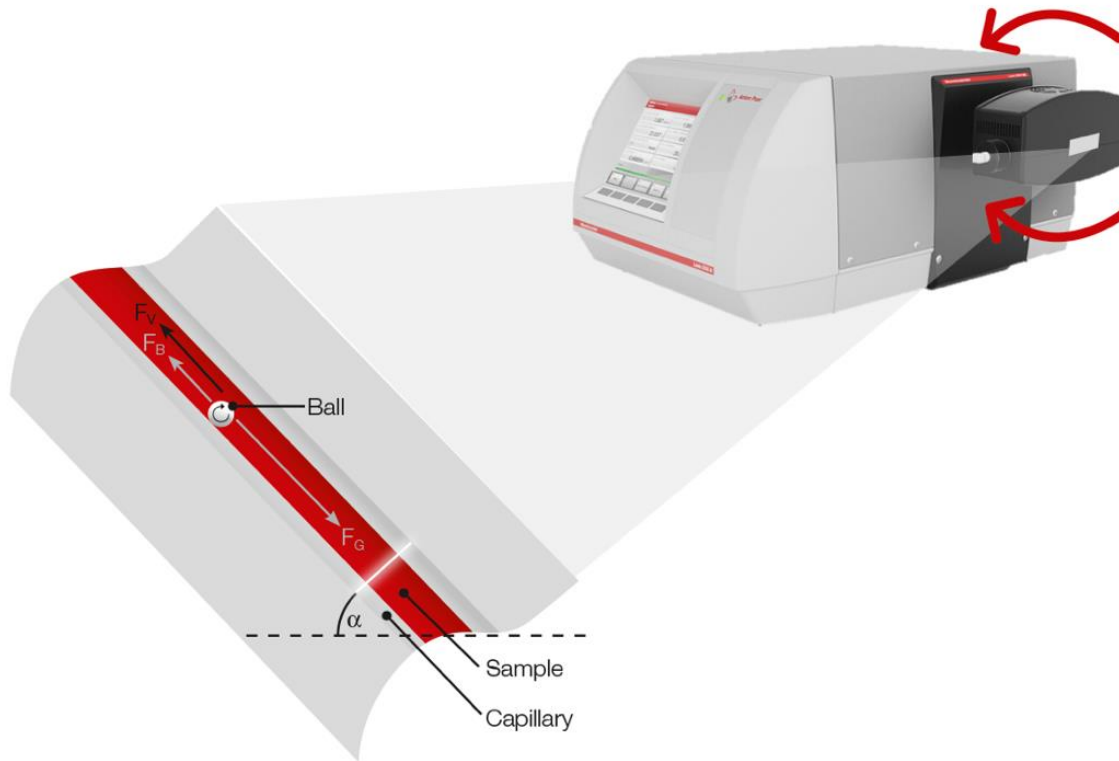


Figure 2.3 – Rolling-ball principle showing the main forces acting on the descending ball. F_G : Effective portion of gravitational force, F_B : Effective portion of buoyancy force, and F_V : Viscous force.⁹⁷

The equipment used for viscosity measurements is composed of a glass capillary tube, which comes in two diameters, 1.59 mm and 1.8 mm, into which the sample is introduced. The capillary tube is equipped with a steel ball that is placed in a temperature-controlled capillary block. The block can be adjusted to vary its inclination angle, according to the calibration performed.

The experimental method is based on Höppler’s falling ball principle⁹⁹, where the steel ball rolls through the capillary tube filled with the sample at a defined angle. Three inductive sensors precisely measure the rolling time it takes for the ball to roll between predefined marks. The determination of sample density involves considering the net forces acting on the sphere, ensuring a balance between the gravitational force (F_G), buoyancy (F_B), and fluid drag force (F_D) equal to zero. Based on the formula for drag force $F_D = 6\pi\eta r_b v_b$, gravitational force $F_G = V_b \rho_b g$ and buoyancy $F_B = V_b \rho_s g$, and considering the velocity of the ball $v_b = s/t$, the dynamic viscosity η of a liquid is calculated as Eq. 2.2.

$$\eta = \frac{2}{9s} r_b^2 g (\rho_b - \rho_s) t \quad (2.2)$$

where η is the dynamic viscosity of the sample solution (mPa.s); r_b is the radius of the ball (cm); s is the distance traveled (cm); V_b is the volume of the ball; g is the acceleration of the gravity; ρ_b is the ball density (g.cm^{-3}); ρ_s is the sample density (g.cm^{-3}), measured by the instrument; and t is the rolling time in seconds.

The elements in front of the parenthesis can be combined to form a constant. The drag force used in this equation is based on Stokes' law¹⁰⁰, which assumes an infinitely extended vessel where the wall of the vessel has no impact on the sphere's sinking. However, in practice, the vessel is not infinitely extended, and the wall of the vessel does influence the sphere's motion. To account for this, the manufacturer corrects the constant for each sphere as part of a calibration process. The modified equation, incorporating the ball constant (K) obtained from the manufacturer's test protocol, is Eq. 2.3.

$$\eta = K(\rho_b - \rho_s)t \quad (2.3)$$

The estimate of viscosity uncertainty depends on the uncertainty of density. The calculation of the uncertainty is given by Eq. 2.4.

$$u = \sqrt{\left(\frac{u_B}{\sqrt{3}}\right)^2 + \left(\frac{s}{\sqrt{N}}\right)^2} \quad (2.4)$$

u_B , linked to the accuracy of the instrument, is equal to the 0.5% given by the Manual of Lovis. The second term is the standard deviation (s) of the average of three different measurement cycles (considered 3 cycles). When just one cycle of measurements is performed, the uncertainty of the experimental data corresponds to the uncertainty of the instrument used (u_{RB}). More details concerning the estimation of uncertainty are given in section 2.6 of this chapter.

2.5.3. Refractive index

The refractive index of the sample was determined using a refractometer from Anton Paar ABBEMAT 300¹⁰¹ with a wavelength of 589 nm. In this method, based on the critical angle method, the sample is in direct contact with the surface of the refractometer prism, which possesses a precisely known refractive index. A light source located within the refractometer emits light, which then illuminates the interface between the prism and the sample. By carefully controlling the angle at which light enters the sample and monitoring the intensity of the transmitted light, the critical angle at which total internal reflection occurs is measured. This critical angle, along with the known refractive index of the prism, is used to calculate the refractive index of the sample. The accuracy of measurement equal to +/- 0.0001.

All measurement was conducted at a temperature of 293.15 K. Additionally, atmospheric pressure during the measurement was determined using a GE Druck DPI 142 Barometric Indicator, with an uncertainty of 0.029 kPa.

2.5.4. Vapor-liquid Equilibrium for pure compounds

The variable volume cell technique, as illustrated in Figure 2.4, based on “static-synthetic” method^{102–104} was used for the determination of equilibrium properties of the pure DESs. The equilibrium cell is equipped with a high-temperature pressure transmitter (35XHTC series from Keller 0-5 bar). Temperature is given to the Pt100 Platinum Resistance Thermometer probe located at the top of the cell. An internal stirring system with an external motor reduced the time required to reach equilibrium. Fig. 2.4 presents the equilibrium cell.

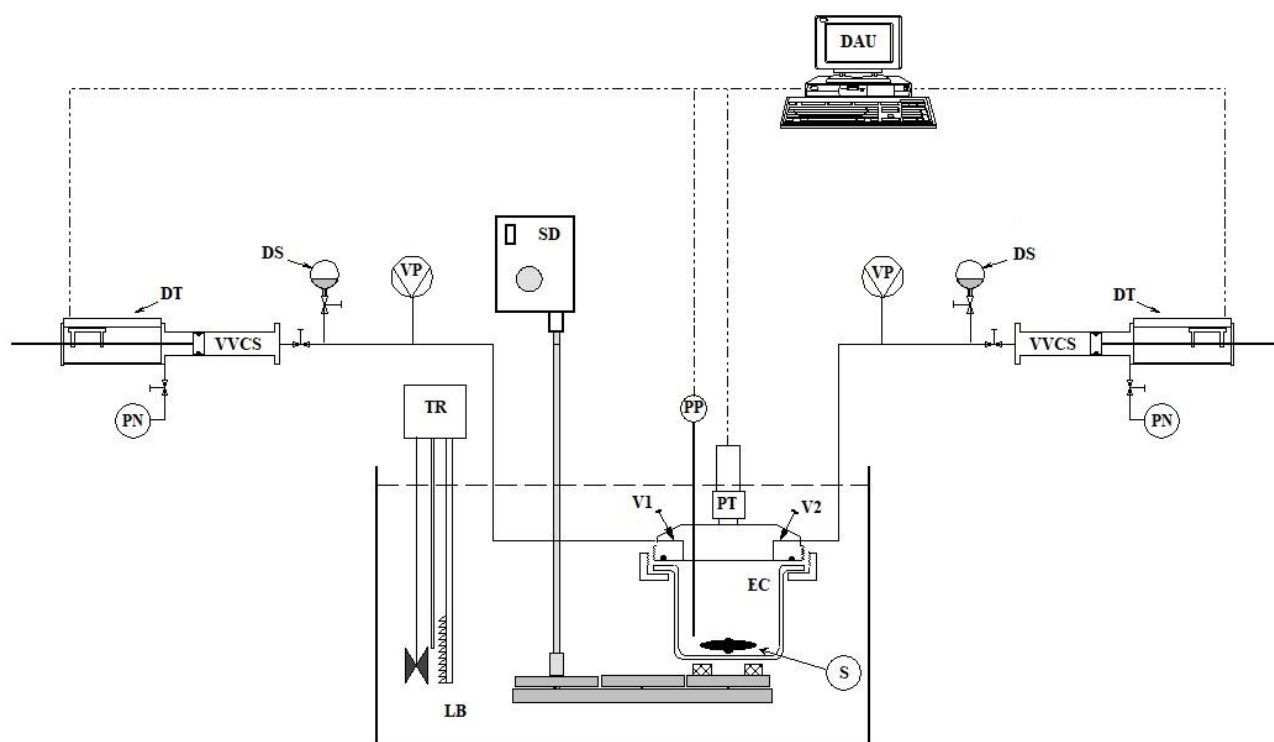


Figure 2.4 – Flow diagram of the synthetic apparatus: DAU: data acquisition unit; DS: degassed solution; DT: displacement transducer; EC: equilibrium cell; LB: liquid bath; PN: pressurized nitrogen; PP: Platinum probe; PT: pressure transducer; S: stirrer; SD: stirring device; TR: thermal regulator; Vi: valve; VP: vacuum pump; VVCS: variable volume cylinder (modified from Coquelet et al. (2020)¹⁰⁴).

Calibration of temperature sensors: The equilibrium cell temperature is measured using a Pt100 Platinum Resistance Thermometer (PRT) probe, calibrated against a reference PRT, specifically the STHP-BP 25 Ohms model from TINSLEY. It's noteworthy that the reference PRT has been calibrated by LNE (Laboratoire national de métrologie et d'essais), further ensuring the accuracy and traceability of temperature measurements. The accuracy for temperature is estimated not higher than ± 0.02 °C by the temperature range.

Calibration of pressure transducer: The pressure transducer has been calibrated against a low-pressure numerical reference standard model 24610 from Desgranges & Huot (Figure 2.5), covering the range from 0 to 5 absolute bars. The principle of operation is based on a fundamental principle of pressure measurement, often referred to as the piston-cylinder principle. The core of the device consists of a piston and cylinder arrangement. The piston is positioned inside the cylinder, and there is no friction between them. When pressure is applied to the system, it exerts a force on the piston. This force is directly proportional to the pressure and the surface area of the piston by the fundamental equation $P = F/A$. (The instrument has two chambers, one under vacuum and the other pressurized to the desired calibration pressure, with the piston assembly located in between. By measuring the force and knowing the piston's surface area, it calculates and displays pressure digitally. Atmospheric pressure has to be taken into account. This mechanical measurement principle provides superior reliability and stability over time, eliminates most risks of measurement error, and reduces calibration uncertainties. The expanded standard uncertainty in pressure measurements was estimated to be ± 0.1 mbar.

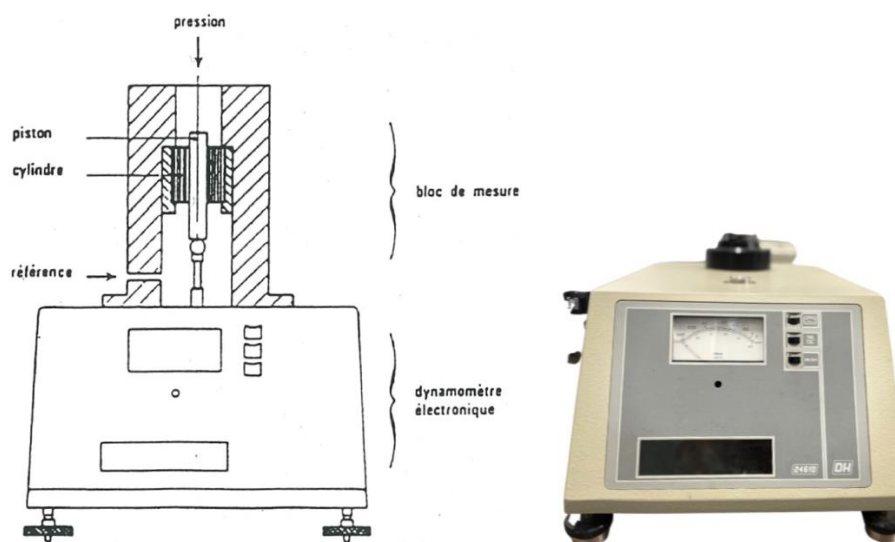


Figure 2.5 – Low-pressure numerical reference standard model 24610 from Desgranges & Huot.

a. Viscous component

The equilibrium cell (Figure 2.4) is weighed empty before assembling it on the apparatus. Then, the viscous component degassed is introduced directly to the cell using a syringe and weighed again. The mole number of the viscous component is calculated by Eq. 2.5.

$$n_{\text{viscous component}} = \frac{m}{\rho(T)} \quad (2.5)$$

Once the equilibrium cell has been loaded with the viscous component, it is then assembled onto the apparatus. Subsequently, the cell undergoes a degassing process achieved by applying a vacuum to remove the air inside of the cell. This precautionary step taken few seconds to minimize the risk of unintentionally removing the main component. Once the liquid and vapor of the pure viscous component achieve equilibrium, the temperature and vapor pressure of the pure viscous component is read by a data acquisition unit and recorded. The equilibrium cell, once installed on the apparatus, undergoes a brief degassing process for a few minutes to prevent the inadvertent removal of any components.

b. Non-viscous component

The component is stored and compressed in a volumetric pump, composed of a variable volume cell (VVC) with an internal diameter of 30.04 ± 0.01 mm, a piston connected to one optical-electronic displacement transducer (model LS 406 C, digital display ND 221 from Heidenhain) with a precision of ± 0.003 mm and a pressurizing circuit by nitrogen. The volumetric pump allows controlled injections of the component into the equilibrium cell. A vacuum is obtained inside the equilibrium cell before loading the component. An amount of the component is introduced into the equilibrium cell, and the temperature of the volumetric pump and the piston displacement is read by a data acquisition unit and recorded. The mole number of components introduced is calculated by Eq. 2.6.

$$n_{\text{solvent}} = \rho(T) \pi r^2 \Delta l \quad (2.6)$$

Where Δl is the piston displacement (accuracy ± 0.01 mm), and r is the radius of the piston. Once the liquid and vapor of the component achieve equilibrium, the temperature and vapor pressure of the pure component is recorded.

2.5.5. Vapor-liquid equilibrium for binary mixtures

The same apparatus (Figure 2.4) as previously was used for the determination of the VLE of a mixture of different DESs with alcohols. Calculation is based on a mass balance. Once the cell is loaded with the first component following the procedure explained before, the second component is introduced into the cell through successive loadings. For each injection of the second component, the piston displacement and temperature of the volumetric pump are recorded. The mole number of the second component is calculated by Eq. 2.6. Having the mole number of the first and second components, the global composition, z_i , is determined using Eq. 2.7.

$$z_i = \frac{n_i}{\sum_{i=1}^{NC} n_i} \quad (2.7)$$

Considering no excess volume, the volume of the liquid (Eq. 2.8) phase is obtained by the mass of compounds and their density at the temperature of measurement.

$$V^L = \sum_{i=1}^{NC} \frac{m_i}{\rho_i(T)} \quad (2.8)$$

Consequently, the volume of the vapor phase (Eq. 2.9) is calculated by the difference between the total volume of the cell and the volume of the liquid phase.

$$V^V = V^T - (V^L + V^{stirrer}) \quad (2.9)$$

Volume of liquid phase can also be estimated visually by Eq. 2.10.

$$V^L = \pi r_{cell}^2 h_{liq} - V^{stirrer} \quad (2.10)$$

where r_{cell} is the radius of the equilibrium cell, h_{liq} is the measured level of the vapor-liquid interface and $V^{stirrer}$ is the volume of the magnetic stirrer.

Assuming the global composition z_i is equal to the liquid composition x_i , as a first guess, and using a gamma-phi approach, the bubble pressure is predicted using Eq. 2.21.

$$P^{cal} = \sum_{i=1}^{NC} x_i \gamma_i \phi_i^{Sat} P_i^{Sat} P_{Oy_i} / \phi_i^{Vap} \quad (2.11)$$

where γ_i is the activity coefficient of compounds i in the liquid phase, P_i^{Sat} is the vapor pressure of the pure compound, ϕ_i^{Sat} is the fugacity coefficient at the saturated conditions, and the $P_{\text{oy}i}$ is the Poynting correction which considers the pressure effect, here assumed to be equal one. The vapor phase is considered as an ideal gas, and consequently the fugacity coefficients are assumed to be equal to unity. An activity coefficient model (non-random two-liquid, NRTL, or Wilson) was applied with interaction parameters to represent the liquid phase. The parameters of the model are minimized by the objective function of the pressure (Eq. 2.12).

$$OF = \frac{1}{N_{\text{data}} - N_{\gamma \text{ parameters}}} \sum_{k=1}^{NP} \left(\frac{|P_k^{\text{cal}} - P_k^{\text{exp}}|}{P_k^{\text{exp}}} \right)^2 \quad (2.12)$$

Number of mols of vapor phase is given by Eq. 2.13 and the molar volume of vapor phase is given by Eq. (2.14).

$$n^V = V^V \rho^V(T_{\text{cell}}, P) = \frac{V^V R T_{\text{cell}}}{P_{\text{solute}}} \quad (2.13)$$

$$v^V = \frac{P}{RT_{\text{cell}}} \quad (2.14)$$

And we consider Eq. 2.15 to calculate the liquid molar volume.

$$v^L = \sum_{i=1}^{NC} x_i v_i^L \quad (2.15)$$

Eq. 2.16 is solved to calculate the vapour phase mole number.

$$0 = V_{\text{cell}} - [v^L(n^T - n^V) + v^V n^V] \quad (2.16)$$

The mole number in the liquid phase is determined by considering Eq. 2.17.

$$n^L = n^T - n^V \quad (2.17)$$

Considering the composition of each phase at equilibrium (bubble point), the mole number of each species in the liquid phase is calculated by considering the global composition of the mixture (z) and the composition of the vapor phase (y), Eq. 2.18.

$$n_i^L = z_i n^T - y_i n^V \quad (2.18)$$

The solubility is determined with Eq. 2.19.

$$x_i = \frac{n_i}{\sum n_j} \tag{2.19}$$

The calculation ends when a non-evolution of composition of liquid phase is observed. Fig. 2.5 summarizes the calculation to predict VLE in a flowchart.

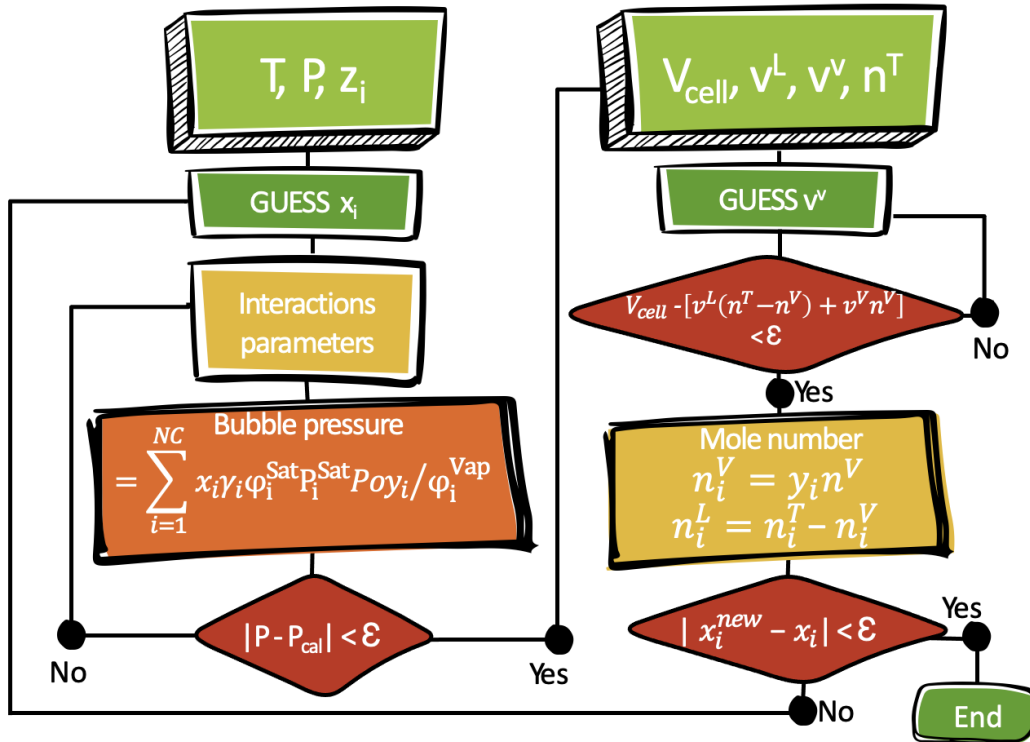


Figure 2.5 – Flowchart to predict VLE using the static synthetic method.

2.5.6. Apparent Henry’s law constant and infinite dilution coefficient activity

Apparent Henry’s law constant and infinite dilution activity coefficient (IDAC) were measured using a gas stripping method using a dilutor and a saturator cell (Figure 2.6). The experimental procedure, fully described by Richon et al.^{105,106}, is based on the variation of vapor phase composition when the highly diluted solute of the liquid mixture in an equilibrium cell is stripped from the solution by a flow of inert gas (helium). The composition of the gas leaving the cell is periodically sampled and analyzed by gas chromatography. The peak area, S, of the solute decreases exponentially with the volume of inert gas flowing out from the cell.

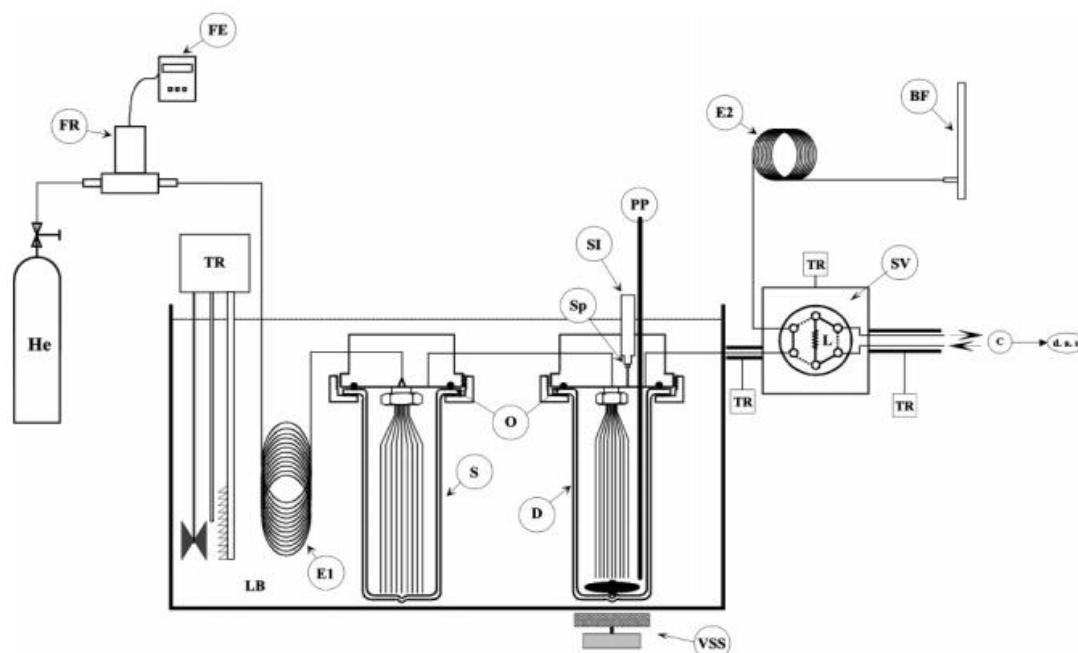


Figure 2.6 – Flow diagram of the equipment: BF, bubble flow meter; C, gas chromatograph; D, dilutor; d.a.s., data acquisition system; He, helium cylinder; E1, E2, heat exchangers; FE, flow meter electronic; FR, flow regulator; L, sampling loop; LB, liquid bath; O, O-ring; PP, platinum resistance thermometer probe; S, saturator; SI, solute injector; Sp, septum; SV, sampling valve; TR, temperature regulator; VSS, variable speed stirrer (from Coquelet et al.¹⁰⁷)

In this method, two cells (Dilutor :D and Saturator :S) are immersed inside a liquid bath (LB) regulated within 0.01 K. The carrier gas from the helium cylinder (He) gets into the system with a flow rate settled by a flow regulator (FR) and measured by a flowmeter (FE). Then, it passes through a heat exchanger (E1) to control to the experimental temperature. The temperature is controlled by liquid bath (LB). The flowrate of the carrier gas is additionally measured by a soap bubble flowmeter (BF) placed in the end of the apparatus. A platinum probe (PP), in contact with the liquid phase of the dilutor cell (D) connected to an electronic display, is used for temperature readings. Temperature accuracy was estimated to be ± 0.2 K.

In this experiment, 40 cm³ of Fructaline was introduced into the “saturator cell”, while about 50 cm³ of the solvent were introduced into the dilutor cell and few mL of solute (n-Hexane) were also added. Both cells are fitted with thin stainless-steel capillaries though the carrier gas is introduced into the cells. A constant stripping gas “helium” flow adjusted to a given value by means of a mass flow regulator was bubbled through the stirred liquid phase and stripped the volatile solute into the vapor phase. The D cell contains a PP, a septum (Sp) and a magnetic stirrer driven by a variable speed stirrer (VSS). The composition of the gas leaving the dilutor cell was periodically sampled and analyzed by gas chromatography using a gas sampling valve (SV).

Equilibrium must be reached between the gas leaving the cell and the liquid phase in the cell. This can be checked by verifying that the measured activity coefficient value does not depend on the eluting gas flowrate. The peak area of the solute decreases exponentially with time if the analysis is made in the linearity range of the detector.

The Apparent Henry's Law coefficient, H_i (Pa), of solute i is calculated by means of Eq 2.20 with the assumption that the equilibrium has been reached between the gas leaving the cell and the liquid phase residing inside the cell.

$$P \cdot y_i = H_i \cdot x_i \quad (2.20)$$

Eq. 2.21 is obtained considering mass balance around the equilibrium cell concerning the solute.

$$H_i = -\frac{1}{t} \ln \left(\frac{S_i}{(S_i)_{t=0}} \right) \cdot \frac{RTN}{\frac{D}{1 - \frac{P_{solv}^{sat}}{P}} + \frac{V_G}{t} \ln \left(\frac{S_i}{(S_i)_{t=0}} \right)} \quad (2.21)$$

where D is the carrier gas flow rate ($\text{m}^3 \cdot \text{s}^{-1}$); N is the total number of moles of solvent inside the dilutor cell; V_G (m^3) is the volume of the vapor phase inside the dilutor cell; S_i is the chromatograph solute i peak area; t (s) is the time; T (K) is the temperature inside the cell; P is the pressure inside the cell (around atmospheric pressure); P_{solv}^{sat} (Pa) is the saturation pressure of the solvent; and R ($\text{J} \cdot \text{mol}^{-1} \cdot \text{K}^{-1}$) is the ideal gas constant. Uncertainty concerning the Henry's law coefficient is estimated to be within 15 %. This estimation comes from propagation of errors on the uncertainty of the solute i peak area determination, the uncertainties on the flow, the uncertainties related to the temperature and pressure, number of moles of solvent and accuracy of the approach (mass balance and hypothesis, see Hajiw et al. (2017)⁹³).

IDAC, γ_i^∞ , is calculated through Eq. 2.22:

$$\gamma_i^\infty = \frac{H_i}{P_{solv}^{sat}} \quad (2.22)$$

2.6. Uncertainty

There are two main sources of uncertainties: uncertainties due to calibration and uncertainty due to repeatability. The uncertainty due to calibration is of Type B. It means that after calibration we do not directly get the uncertainty but the accuracy. A statistic distribution must be chosen. In term of probability, it is common to consider a rectangular statistic distribution. The rectangular distribution is given as

$u_{calib}(\theta) = \frac{b}{\sqrt{3}}$. b is defined as the half-width between the upper and lower error limits (so the value of the accuracy). This value is obtained after polynomial regression between value given by the sensor and the reference value. For a normal distribution, $\pm u$ encompasses about 68 % of the distribution; for a rectangular distribution, $\pm u$ encompasses about 58 % of the distribution.

The electronic balance used in this work for weigh the chemical products was the Mettler Toledo model XS205 with an uncertainty of $u_{calib} = 0.01$ mg.

The platinum probe Pt100 Ω was calibrated using a device composed by a reference probe 25 Ohms Hart Scientific on a HP 34420A central. The uncertainty after calibration was 0.01 K.

During an experiment, to determine the value of one quantity (T, P or mole numbers) we have to consider all values acquired. It follows a calculation of average values and standard deviation δ . The averaging of repeated readings yields a mean θ_{avg} with a standard deviation δ . As before, this can be statistically converted to an uncertainty due to repeatability of the measurements, via Eq. 2.23

$$u_{rep}(\theta) = \frac{\delta}{\sqrt{n}} \sqrt{\frac{1}{n(n-1)} \cdot \sum_{i=1}^n (\theta_i - \theta_{avg})^2} \quad (2.23)$$

where, $\theta_{avg} = \frac{1}{n} \sum_{i=1}^n \theta_i$ and n is the number of repeated quantity measurements. A Gaussian type of distribution is the likely behavior here, since the repeated readings are likely to fall close to the mean (with maybe one or two values falling from the mean). This is known as a type A evaluation (systematic uncertainty), where only statistical methods are required to interpret the uncertainty. Eq. (2.24) presents the combined standard temperature uncertainty, u_c (T).

$$u_c(\theta) = \sqrt{u_{calib}(\theta)^2 + u_{rep}(\theta)^2} \quad (2.24)$$

with subscripts calib, rep denoting that of calibration, repeatability.

The determination of the uncertainty of the composition required the uncertainty of each mole numbers. The uncertainty of the mole fraction is determined after calibration of the GC detectors (Equations 2.25 and 2.26).

$$x_i = \frac{n_i}{\sqrt{\sum_i^{ncomp} n_i}} \quad (2.25)$$

$$u(x_i) = \sqrt{\sum_i^{ncomp} \left(\frac{\partial x_i}{\partial n_i}\right)_{i \neq j}^2 u^2(n_i)} \quad (2.26)$$

For example, for a binary system, one can calculate $u(x_1) = x_1(1 - x_1)\sqrt{\left(\frac{u(n_1)}{n_1}\right)^2 + \left(\frac{u(n_2)}{n_2}\right)^2}$.

Uncertainty on apparent Henry's law constant can be calculated from uncertainty on mole fraction (both vapor and liquid phases) and uncertainty on pressure. Eq. 2.27 details the expression.

$$u(H_i) = H_i \cdot \sqrt{\left(\frac{u(P)}{P}\right)^2 + \left(\frac{u(x_i)}{x_i}\right)^2 + \left(\frac{u(y_i)}{y_i}\right)^2} \quad (2.27)$$

2.7. Conclusion

In conclusion, this chapter explores the experimental study, covering the synthesis and characterization of DES and mixtures. It outlines the methodology, materials, and procedures involved in creating these solvents, emphasizing essential properties such as density, viscosity, refractive index, and vapor-liquid equilibrium. The chapter also highlights the critical aspect of uncertainty in the data obtained, laying the foundation for a comprehensive exploration of DES properties and behaviors.

In the described methodology for obtaining Vapor-Liquid Equilibrium (VLE) data through the static-synthetic method, an activity coefficient model should be employed, with interaction parameters to represent the liquid phase. Fine-tuning these parameters with the experimental data is crucial to improve the accuracy of the model. Activity coefficient models thus become indispensable tools, enabling the prediction of thermodynamic properties and aiding optimization of separation processes.

Within this study, a specific activity coefficient model will serve as the central focus of the methodology proposed for solvent selection. The forthcoming chapter will delve deeply into the realm of thermodynamic modeling, placing specific emphasis on activity coefficient models. Moreover, it will offer a comprehensive overview of the model chosen for this study: the COSMO-based models (COnductor-like Screening MOdel).

Chapter 3

Thermodynamic modeling

Modélisation thermodynamique

Résumé

Lorsqu'il s'agit de traiter les DES et du défi d'obtenir des données expérimentales, diverses approches théoriques peuvent être employées pour modéliser et prédire les propriétés de ces solvants. Ces approches peuvent contribuer à combler le manque d'information et à fournir des perspectives précieuses pour les applications scientifiques et en ingénierie. Une grande variété d'approches de modélisation a été développée pour répondre à ce besoin et elles sont largement classées en quatre catégories distinctes en fonction des principes directeurs de leurs cadres théoriques. Ces catégories sont les modèles empiriques (c'est-à-dire les corrélations), la thermodynamique classique (c'est-à-dire la méthode de contribution de groupe, les équations d'état classiques et les modèles de coefficients d'activité), la thermodynamique statistique (c'est-à-dire la simulation moléculaire et les équations d'état moléculaires) et les modèles basés sur la chimie quantique (c'est-à-dire le modèle COSMO).

Le choix de l'approche de modélisation dépend des objectifs de recherche spécifiques, des ressources informatiques disponibles et de l'objectif de représenter et de prédire avec précision les comportements, les propriétés ou les phénomènes souhaités du système.

Avant d'introduire les approches de modélisation qui sont basées sur des modèles thermodynamiques, ce chapitre aborde les notions fondamentales essentielles de la thermodynamique de l'équilibre des phases. Les équations fondamentales de l'équilibre des phases, exprimées en termes de potentiel chimique (μ_i) qui joue un rôle central dans la compréhension des conditions d'équilibre, seront présentées. Les deux grandeurs thermodynamiques proposées par Lewis et appelées fugacité (f) et activité (a) seront présentées dans ce chapitre. La fugacité sert à prendre en compte les écarts par rapport à l'état de gaz idéal dans la phase vapeur tandis que l'activité permet de tenir compte de la non-idéalité d'un mélange. Ces deux grandeurs, fournissent ainsi une base théorique solide pour la modélisation et la compréhension des équilibres de phases grâce à leur coefficient de fugacité et celui d'activité.

De plus, ce chapitre explique deux approches classiques de modélisation des mélanges réels en équilibre de phases : l'approche résiduelle (basée sur le coefficient de fugacité - Φ) et l'approche d'excès (basée sur le coefficient d'activité - γ). L'approche résiduelle quantifie les écarts par rapport au comportement du gaz idéal en utilisant des équations d'état tandis que l'approche d'excès se concentre sur les fonctions d'excès et en particulier pour les mélanges liquides ayant des interactions moléculaires fortes.

Ces méthodes sont utilisées pour représenter à la fois les phases gazeuses et liquides : approche γ - ϕ et approche ϕ - ϕ . L'approche γ - ϕ combine l'utilisation des coefficients de fugacité pour décrire la phase vapeur (γ) et les coefficients d'activité pour décrire la phase liquide (ϕ). Les relations d'équilibre sont présentées et des équations sont données pour calculer les coefficients de fugacité et d'activité. Ces coefficients sont cruciaux pour résoudre des problèmes d'équilibre et sont particulièrement précis pour les calculs à basse et moyenne pression. En revanche, dans l'approche ϕ - ϕ , les phases vapeur et liquide sont décrites à l'aide des coefficients de fugacité.

Ce chapitre se concentre sur les modèles de coefficients d'activité qui sont classés en trois catégories : modèles empiriques, semi-prédictifs et prédictifs. Un aperçu historique du développement des modèles de coefficients d'activité sera présenté. Ce chapitre fournit des descriptions approfondies des modèles importants y compris les modèles de Scatchard-Hildebrand, de Wilson, de NRTL, d'UNIFAC et ceux basés sur COSMO (COnductor-like Screening Model)

Les méthodes basées sur COSMO reposent sur le calcul chimique quantique et la thermodynamique statistique. Il s'agit notamment de COSMO-RS (COSMO for real solvents) et de COSMO-SAC (COSMO segment activity coefficient). Dans ces modèles, les éléments clés représentant le lien entre le calcul chimique quantique et la thermodynamique statistique sont décrits par le « sigma-profile ». L'accent de ce chapitre est mis sur une explication détaillée de la partie de thermodynamique statistique de ces modèles.

Thermodynamic modeling

3.1. Introduction

As previously mentioned, when dealing with DESs and the challenge of obtaining experimental data, various theoretical approaches can be employed to model and predict the properties of these solvents. These approaches can help to fill the information gap and to provide valuable insights for engineering and scientific applications. Over the years, a large variety of modeling approaches has been developed to address this need and broadly classified into four distinct categories based on the governing principles of their theoretical frameworks. These categories are empirical models (i.e., machine learning and correlations), classical thermodynamics (i.e., group contribution method, classical equation of states (EoSs) and activity coefficients models), statistical thermodynamics (i.e., molecular simulation and molecular EoSs), and quantum chemistry-based models (i.e., COSMO-model). Figure 3.1 shows these categories, types of thermodynamics models, and their contribution to modeling pure DESs and their multicomponent mixtures.

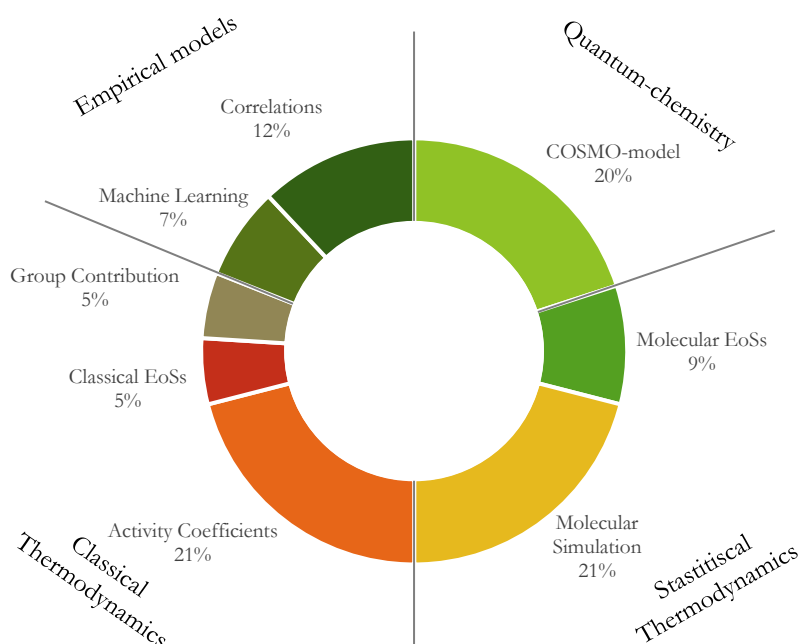


Figure 3.1 – Approaches to modeling DES (modified from Alkhatib et al.¹⁰⁸).

The choice of modeling approach depends on the specific research objectives, available computational resources, aiming to accurately represent and predict the desired system behaviors, properties, or phenomena. In this chapter, we will explore two important categories of theoretical modeling techniques for DESs: classical thermodynamics and quantum chemistry-based models, with a focus on COSMO-based approaches. Classical thermodynamics, rooted in equilibrium principles, offer insights into DES

behavior through EoS and phase equilibrium models, enabling predictions of phase behavior, solubility, and thermodynamic interactions. On the other hand, COSMO-model is a widely used continuum solvation model that combines quantum chemistry with classical thermodynamics to describe the solvation behavior of molecules in solution.

3.2. Thermodynamic fundamentals of phase equilibrium

In a system consisting of both liquid and one vapor phase, component i in the liquid phase exhibits a tendency to escape into the vapor phase. Simultaneously, the same component i in the vapor phase tends to condense into the liquid phase. The driving force mass behind these mass transfer processes is the chemical potential (μ_i). When the chemical potential is higher in one phase than in other, the component will tends to move from the phase with the higher chemical potential to the one with the lower chemical potential. The system reaches a state of equilibrium a given temperature when these two mass transfers processes become equal, resulting in the establishment of a constant vapor pressure known as the equilibrium pressure. The fundamental equation for describing this phase equilibrium phenomenon is expressed in terms of the equality of the chemical potential between the coexisting phases ($\mu_i^L = \mu_i^V$). The chemical potential can be described in terms of partial molar Gibbs energy using (Eq. 3.1):

$$\mu_i = \bar{g}_i = \left[\frac{\partial G}{\partial n_i} \right]_{T, P, n_{j \neq i}} \quad (3.1)$$

where n_i , and n_j represent the mole numbers.

The partial differential of molar Gibbs energy is related to the temperature and pressure of a given substance i by (Eq. 3.2):

$$dG_i = V_i dP - S_i dT + \sum_{i=1} \mu_i dn_i \quad (3.2)$$

where V_i and S_i are the molar volume and the molar entropy of pure substance i , respectively.

Using the definition for internal energy (U) and chemical potential, the fundamental equation in terms of U is represented by Eq. 3.3.

$$\mu_i = dU + \sum_{i=1} \mu_i dn_i = TdS - PdV + \sum_{i=1} \mu_i dn_i \quad (3.3)$$

By integration at constant temperature and composition the Eq. 3.3, and then differentiation of this equation gives the Eq. 3.4.

$$\mu_i = TdS + SdT - PdV - VdP + \sum_{i=1} \mu_i dn_i + \sum_{i=1} n_i d\mu_i \quad (3.4)$$

The combination of Eq. 3.3 and 3.4, leads to the Gibbs-Duhem equation (Eq. 3.5) which is fundamental in thermodynamics of solutions.

$$SdT - VdP + \sum_{i=1} n_i d\mu_i = 0 \quad (3.5)$$

By integrating the Gibbs-Duhem equation, the chemical potential can therefore be expressed from a reference state by Eq. 3.6.

$$\mu_i(T, P) = \mu_i(T^0, P^0) - \int_{T^0}^T \bar{s}_i dT + \int_{P^0}^P \bar{v}_i dP \quad (3.6)$$

Using the ideal gas law for an isothermal pressure variation from P^0 to P , we obtain the Eq. 3.7.

$$\mu_i(T, P) = \mu_i(T^0, P^0) + RT \cdot \ln \frac{P}{P^0} \quad (3.7)$$

Where the superscript, ⁰ is used to identify a selected reference state or ideal gas.

Since Eq. 3.7 is true only for a pure and ideal gas, Lewis^{109,110} proposed two new thermodynamic quantities, called fugacity and activity. The fugacity is to consider the deviation from the ideal gas state in the vapor phase and the activity is to consider the non-ideality of a mixture. These two concepts will be presented in the following section and are essential tools in thermodynamics for characterizing deviations from ideal behavior in real systems, providing a solid theoretical basis for modeling and understanding phase equilibria.

3.3. Fugacity and Activity

Approximately about 120 years ago, Lewis stated that for any chemical component i in any phase (either a gas, a liquid or a solid) pure, or mixed, ideal, or non-ideal, the change in chemical potential during an isothermal transition from fugacity f_i^0 to fugacity f_i is written as Eq. 3.8.

$$\mu_i^{IM}(T, P, x_i) - \mu_i^0(T, P, x_i) = RT \cdot \ln \frac{f_i^{IM}(T, P, x_i)}{f_i^0(T, P)} \quad (3.8)$$

The fugacity is equal to the pressure for pure ideal gas: $f_i^0 = P$. For an ideal mixture, the fugacity of i is equal to the partial pressure of this gas: $f_i^{IM} = x_i P$. Thus, the ratio is represented by Eq. 3.9.

$$\frac{f_i^{IM}(T, P, x_i)}{f_i^0(T, P)} = x_i \quad (3.9)$$

Since both pure ideal gases and ideal mixtures approach the behavior of ideal gases at very low pressures, the following limits (Eq. 3.10 and Eq. 3.11) are adopted to complete the definition of fugacity:

$$\lim_{P \rightarrow 0} \frac{f(T, P)}{P} = 1 \quad (\text{for ideal gas}) \quad (3.10)$$

and

$$\lim_{P \rightarrow 0} \frac{f_i^{IM}(T, P, x_i)}{P \cdot x_i} = 1 \quad (\text{for ideal mixtures}) \quad (3.11)$$

Finally, for a non-ideal mixture (noted *real*), it is necessary to correct the partial pressure by a factor quantifying the non-ideality (or deviation from the ideal). While considering ideal gas as reference, the Eq. 3.8 becomes 3.12a and it becomes 3.12b while considering ideal mixture as reference.

$$\mu_i^{real}(T, P, x_i) - \mu_i^0(T, P) = RT \cdot \ln \frac{f_i^{real}(T, P, x_i)}{f_i^0(T, P)} \quad (3.12a)$$

$$\mu_i^{real}(T, P, x_i) - \mu_i^{IM}(T, P, x_i) = RT \cdot \ln \frac{f_i^{real}(T, P, x_i)}{f_i^{IM}(T, P, x_i)} \quad (3.12b)$$

The factor quantifying the non-ideality was also introduced by Lewis and is called fugacity coefficient (ϕ) and activity (a). In which fugacity coefficient is defined as the fugacity ratio of the real fugacity with respect to the ideal gas one, and activity is the ratio of real fugacity with respect to the ideal mixture, as presented in Eq. 3.13.

$$\frac{f_i^{real}(T, P, x_i)}{f_i^0(T, P)} = \phi_i \quad \text{and} \quad \frac{f_i^{real}(T, P, x_i)}{f_i^{IM}(T, P, x_i)} = a_i \quad (3.13)$$

The following equation (Eq. 3.14) shows that the activity allows a difference in chemical potential to be expressed at a fixed temperature (isothermal):

$$\mu_i^{real}(T, P, x_i) - \mu_i^{IM}(T, P, x_i) = RT \cdot \ln a_i \quad (3.14)$$

In the case of an ideal mixture, it takes the form of Eq. 3.15.

$$a_i = x_i \quad (3.15)$$

In the case of a real mixture, the Eq. 3.15 will be corrected, and presented as Eq. 3.16 with a new parameter that is called activity coefficient (γ_i).

$$a_i = x_i \cdot \gamma_i \quad (3.16)$$

Substitution of Eq. 3.16 into Eq. 3.14, for a non-ideal mixture gives Eq. 3.17.

$$\mu_i - \mu_i^o = RT \cdot \ln x_i \gamma_i \quad (3.17)$$

3.4. Calculation of thermodynamic properties

Two classical approaches to model a real mixture in phase equilibrium:

- Residual approach or phi: the ideal gas is selected as the reference state,
- Excess approach or gamma: the ideal mixture is selected as the reference state.

In the first approach, deviations from ideal gas behavior are typically quantified using an EoS, and adjustments of the fugacity are made using the fugacity coefficient. This method is extensively employed for representing fluids, both gas and liquid phase (Eq. 3.18):

$$f^V(T, P, x_i) = P \cdot x_i^V \cdot \phi_i^V \quad \text{or} \quad f^L(T, P, x_i) = P \cdot x_i^L \cdot \phi_i^L \quad (3.18)$$

Where superscript V and L denoted gas and liquid phase, respectively.

Conversely, the second approach centers on the calculation of deviations from ideality using excess functions, primarily proposed for modeling liquid mixtures characterized by strong molecular interactions. The activity coefficient is used to characterize deviations from ideality, as in Eq. 3.19.

$$f^L(T, P, x_i) = f_i^{oL} \cdot x_i^L \cdot \gamma_i^L \quad (3.19)$$

where f_i^{oL} is the fugacity in liquid phase of pure component i .

The fugacity function has been introduced because its relation to the Gibbs energy makes it useful in phase equilibrium calculation. Phase equilibrium is expressed by the equality of fugacity, at constant temperature and pressure in the two phases (Eq. 3.20):

$$f^I(T, P, x_i, y_i) = f^{II}(T, P, x_i, y_i) \quad (3.20)$$

Since this equation follows directly the equality of molar Gibbs energy in each phase equilibrium, it can be used as criteria for equilibrium. This equilibrium problem can be solved by two different approaches: “gamma-phi” or “phi-phi”.

3.4.1. The γ - ϕ approach

In the classical approach, called “gamma-phi”, the definition of the fugacity coefficient is used to describe the vapor phase, and the activity coefficient is used to describe the liquid phase. The following equilibrium relation can be written as Eq. 3.21.

$$\phi_i^V \cdot P \cdot y_i = \gamma_i \cdot x_i \cdot f_i^{0L} \quad (3.21)$$

Integration of Eq. 3.13 at a constant temperature from vapor pressure P^{sat} to pressure P leads to Eq. 3.22.

$$f_i^{0L} = P_i^{sat} \cdot \phi_i^{sat} \cdot \exp\left(\int_{P^{vap}}^P \left(\frac{\bar{V}_i^L}{RT}\right) dP\right) \quad (3.22)$$

Where ϕ_i^{sat} is the fugacity coefficient at the saturated conditions at the same T and P . The exponential on the right side is called the Poynting factor (Poy). The integral in the Poynting factor is simplified if we make the approximation that \bar{V}_i^L is independent of pressure. We obtain then the approximate relation by Eq. 3.23.

$$f_i^{0L} = P_i^{sat} \cdot \phi_i^{sat} \cdot \frac{1}{RT} \exp\left(\frac{\bar{V}_i^L (P - P_i^{sat})}{RT}\right) \quad (3.23)$$

Substitution of Eq. 3.23 into 3.21, leads to Eq. 3.24.

$$(\phi_i^V \cdot P \cdot y_i)^V = (\gamma_i \cdot x_i \cdot P_i^{sat} \cdot \phi_i^{sat} \cdot Poy)^L \quad (3.24)$$

The fugacity coefficient can be calculated by an equation of state and the activity coefficient by a model of solution. This method can be applied to a wide variety of mixtures and is accurate for low to moderate pressure calculations. The “gamma-phi” approach has been used to solve most of the equilibrium problems of this study. Furthermore, activity coefficient models have been extensively studied and will be presented in the section 3.4.

For liquid-liquid equilibria, where the same standard state is used in both phases, we obtain the Eq. 3.25.

$$(\gamma_i \cdot x_i)^{L1} = (\gamma_i \cdot x_i)^{L2} \quad (3.25)$$

3.4.2. The ϕ - ϕ approach

The “phi-phi” approach, for every component i , the equation for vapor-liquid equilibrium is represented as Eq. 3.26.

$$y_i \cdot \phi_i^V P = x_i \cdot \phi_i^L P \quad (3.26)$$

and for liquid-liquid equilibrium is presented as Eq. 3.27.

$$(x_i \cdot \phi_i^L)^{L1} = (x_i \cdot \phi_i^L)^{L2} \quad (3.27)$$

3.4.3. PT-Flash algorithm

Flash calculations are used for processes with VLE. A typical process that requires flash calculations is when a feed stream F (with composition z_i) is separated into a vapor product V (with composition y_i) and liquid product L (with composition x_i); see figure 3.2.

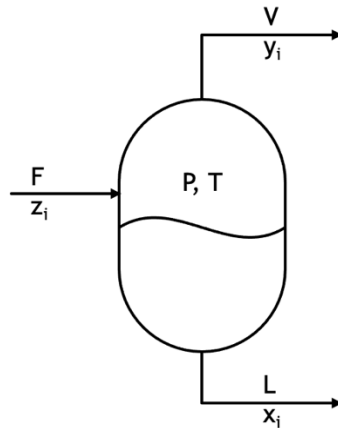


Figure 3.2 – Flash tank.

A successive substitution algorithm has been implemented for flash calculation. For each of the N_c components, we can write a component material balance as Eq. 3.28.

$$y_i \cdot \beta + x_i \cdot (1 - \beta) = z_i \quad (3.28)$$

In addition, the vapor and liquid are assumed to be equal to Eq. 3.29.

$$y_i = K_i \cdot x_i \quad (3.29)$$

With K_i equilibrium constant function of T , P , x and y .

Once, the approach is defined to solve VLE problems, two possibilities are encountered: either a bubble point or a dew point must be calculated. Then, K -factors were calculated based on fugacity coefficients as expressed in the Eq. 3.30 and used as a convergency criterium.

$$K_i = \frac{\varphi_i^l}{\varphi_i^v} \quad \text{or} \quad K_i = \frac{\gamma_i^l}{\varphi_i^v} \quad (3.30)$$

Substitution of Eq. 3.29 into 3.28, and solution with respect to x_i gives Eq. 3.31.

$$x_i = \frac{z_i}{1 - \beta + \beta \cdot K_i} \quad (3.31)$$

Here, we cannot directly calculate x_i because β is not known. To get β , we may use the relationship Eq. 3.32 which will lead to the so-called Rachford-Rice equation (Eq. 3.33)

$$\sum_{i=1}^{Nc} y_i - x_i = 0 \quad (3.32)$$

$$\sum_{i=1}^{Nc} \frac{z_i \cdot (K_i - 1)}{1 - \beta + \beta \cdot K_i} = 0 \quad (3.33)$$

The successive substitution algorithm for the PT flash problem is summarized in the scheme presented in the figure 3.3.

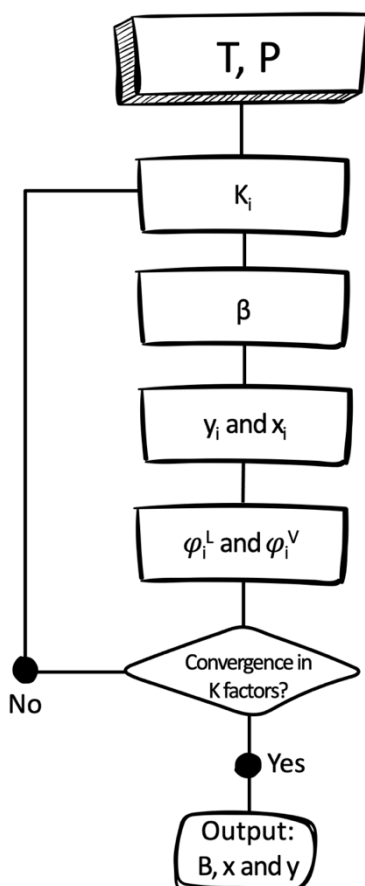


Figure 3.3 – Scheme of successive substitution algorithm for the PT flash problem for phi-phi approach.

3.5. Activity coefficient

From a mathematical point of view, the activity coefficient is a factor to account for deviations from the ideal mixing of chemical substance. Ideal behavior in a solution occurs when the interactions between molecules are the same (or very close) for molecules of the same component and for molecules of different components. For example, this can be the case for two hydrocarbons like n-pentane and n-hexane when they are in equilibrium at fixed temperature and pressure. In other words, in an ideal solution, there are no variation in the strength of forces between the components, including neither van-der-Waals (dispersive) nor Coulomb forces. This uniformity of interactions characterizes the ideal behavior. Under this condition, the activity coefficient is equal to 1, and the solution follows Raoult's law.

However, when dealing with non-ideal solution, various forces act on mixtures of chemical substances. For example, intermolecular forces between solute and solvent molecules may be stronger or weaker than those between similar molecules. In these non-ideal situations, the activity coefficient deviates from 1. Typically, polar substance solutions exhibit higher activity coefficient values. In contrast, activity coefficients are close to 1 in a mixture of non-polar substances, such as a mixture of light hydrocarbons.

The relatively weak intermolecular forces and similar molecular structures in non-polar mixtures lead to a scenario where the actual behavior closely aligns with ideal expectations.

In the presence of significant molecular associations or interactions, the activity coefficient can indeed be less than 1, a characteristic frequently observed in solutions with molecules exhibit strong attractions or bonds. This deviation indicates a reduction in the effective concentration or activity of the substance compared to what would be expected in an ideal solution scenario. In the case of electrolyte solutions, this behavior is especially evident due to the pronounced solvation of ions by surrounding water molecules. The strong ion-solvent interactions markedly decrease the effective ion concentration, resulting in an activity coefficient below 1.

The activity coefficient is mainly temperature and composition-dependents. For simplification to a single compositional point, the activity coefficient at diluted concentration can be considered. At this point, the activity coefficient is known as the infinite dilution activity coefficient (IDAC) and describes the behavior of one solute molecule surrounded by solvent molecules, showing the higher nonideality of the mixture. Therefore, IDAC can be used to find the maximum separation performance of a solvent and investigate solvents in terms of capacities and selectivity. Whereas these values are required to screen potential solvent to separation processes and to design them, such as extractive distillation and liquid-liquid extraction methods.

Measurements of IDAC are quickly conducted for a wide range of solutes by gas chromatographic¹¹¹, ebulliometer¹¹², or gas stripping methods¹¹³ (well described in the chapter 2, section 2.3.2). On the other hand, it is hard to predict them, and usually, the models predict by extrapolation to composition equal to zero. IDAC values can be used to evaluate the performance of activity coefficient models.

3.6. Activity coefficient models

Activity coefficients are obtained from excess Gibbs energy. In thermodynamics, the relationship between the excess enthalpy and the excess entropy is given in eq. 3.34. The excess enthalpy considers energetic interactions called residual (dispersive, dipole, association) and the excess entropy considers the entropic contribution (size of molecules) called configurational.

$$G^E = H^E - TS^E \quad (3.34)$$

Three cases arise based on the comparison of excess entropy and excess enthalpy:

Regular Solution is referred to when the excess entropy is significantly smaller than the excess enthalpy. Energetic interactions (like dispersive, dipole, or association forces) dominate over the entropic contributions (arising from molecular size or configurational aspects).

Athermal Solution is referred to when the excess enthalpy is notably smaller than the excess entropy. The entropic contributions, particularly associated with molecular size or configuration, play a more significant role compared to energetic interactions.

Balanced Contributions are scenarios where neither excess enthalpy nor excess entropy can be neglected. Both energetic interactions and entropic effects are equally significant, and their combined influence on the system cannot be ignored.

There are three categories of thermodynamic activity coefficient models: empirical, semi-predictive, and predictive models.

Empirical correlations are mathematical or statistical models derived from empirical data and observations rather than being based on theoretical principles or underlying scientific laws. These models are built to describe and represent the relationships observed in the data, without a direct link to the fundamental mechanisms governing the phenomena.

Semi-predictive Models require adjustable parameters that are typically determined through regression or fitting to experimental data. These parameters are adjusted to match model predictions with experimental observations with a theoretical basis.

Predictive Models aim to calculate activity coefficients without the need for adjustable parameters derived from experimental data. These models are based on fundamental principles and molecular interactions based on molecular models, or quantum chemistry.

More than one hundred years ago, most chemical engineering thermodynamics was based on experimental data and some “blind” empirical correlation without any physical meaning. Few correlations were done on a theoretical basis till that time, if any. Just in 1895, Margules¹¹⁴ correlated isothermal binary VLE with a power series of the mole fractions to represent the natural logarithm of the activity coefficient of component 1. Then, the activity coefficient of component 2 is obtained from the Gibbs-Duhem equation without adding any parameter. Margules correlation is better known as the Margules activity coefficient model.

In 1910 van Laar¹¹⁵ derived his excess model from van der Waals's equation of state, assuming that the isothermal entropy of the mixing at constant volume is equal to that for an ideal solution. Therefore, van Laar considered the excess entropy and excess volume is equal to zero.

About 20 years after van Laar, Scatchard, and Hildebrand¹¹⁶ (independently) also derived their equation for excess Gibbs energy like that of van Laar. However, instead of van der Waals's a -parameter, they introduced the concept of cohesive energy density, which is the energy required to vaporize a liquid per unit liquid volume. The square root of this cohesive energy density is the solubility parameter δ , which has been determined for many solvents, see Hansen¹¹⁷. Flory-Huggins proposed their equation for segmented molecules, otherwise like the Scatchard-Hildebrand equation.

Until about 1964, chemical engineering applications of activity coefficient were either based on Margules or the van Laar activity coefficient model. However, these models often have some limitations due to their temperature-dependent parameters. Moreover, it gives unsatisfactory predictions for complex systems containing strongly polar and hydrogen bonding components, e.g., alcohol and hydrocarbons (mainly, when applied to multicomponent LLE systems). Furthermore, extending them to multicomponent systems is difficult, in which these models are applicable up to ternary systems. These limitations are problematic in solving difficulties at the industrial level, e.g., in multicomponent distillations, where the distillation is typically performed at constant pressure and varying temperatures throughout the column (from plate to plate). Thus, employing a model with temperature-independent parameters or an in-built temperature dependency (such as the local composition models that will be presented later) is advantageous.

The revolution in activity coefficient models or solution theories came in 1964 when Wilson proposed the concept of local compositions or non-random mixing. Wilson originally derived his model based on the Flory-Huggins equation using local composition fractions instead of the traditional segment or volume fractions used in the Flory-Huggins model. Despite, Wilson's model still has (like Margules and van Laar) two parameters having theoretical significance. More importantly, the model has an in-built temperature dependency. The Wilson model was the first of its kind and more satisfactory models using the same local-composition concept were proposed later, e.g., NRTL and UNIQUAC. In those models, the data are expressed in a mathematical function suggested by classical thermodynamics and correlated to the experimental data.

Comparing UNIQUAC with Wilson model, both models have a combinatorial and a residual term. In Wilson, it is a Flory-Huggins term based on volume fractions, while in UNIQUAC, it is a modified Flory-Huggins based on segment and surface area fractions. The second part of the combinatorial term of

UNIQUAC (Staverman-Guggenheim term) does not usually contribute much to the activity coefficient compared to the Flory-Huggins term. The residual terms of the two models are similar, but the Wilson one is based on volume fractions, while the UNIQUAC one is based on surface area fractions. Both models have two interaction parameters that must be estimated from experimental phase equilibrium data. Wilson's equation only applies to fully miscible systems, but NRTL and UNIQUAC are applied to both fully and partially miscible systems.

In 1968, Renon and Prausnitz proposed NRTL model. Unlike Wilson, NRTL has three parameters. This includes the non-randomness parameter, denoted as α , which has values between 0.20 and 0.47 in most cases. In instances where the data is lacking, it can be arbitrarily set equal to 0.3. When the non-randomness parameter is zero, the mixture is completely random, and the NRTL reduces to the two-suffix (one parameter) Margules equation.

In 1975, Fredenslund and Prausnitz published one of the most successful activity coefficient models based on group contribution methods, the UNIFAC. This model is highly parameterized using a large amount of experimental data, such as VLE, to fine-tune their group interaction parameters. Consequently, their accuracy depends on the availability and similarity of experimental data with the concerned system.

In the mid-1990s, a valuable (and semi-empirical) alternative to these models was developed by applying quantum mechanics (QM) to calculate thermodynamic properties. Klamt developed a model called COnductor-like Screening MOdel for real solvents (COSMO-RS). The COSMO segment activity coefficient model (COSMO-SAC) proposed by Lin and Sandler¹¹⁸ is a reimplement of COSMO-RS, filling the information gaps in the equation that were not fully described in the original model proposed by Klamt in 1995¹¹⁹. Both models use surface charges for molecules that depend on the atom and the identity of other atoms in the same molecule. That method has attracted significant attention due to its simplicity and low computational requirements.

Figure 3.4 presents the annual number of publications about the COSMO model, including articles, conference papers, book chapters, reviews, and conference reviews covering the 1999 – 2022 periods. The keyword employed for the literature search was COSMO-RS, or COSMO-SAC, and it was performed using the Scopus search. For that period, 3196 works were published, and the main subject areas where they had been classified were chemistry, chemical engineering, physics and astronomy, engineering, material science, and computer science. A growing number of research publications have been noticed in the last ten years, going from 100 in 2012 to 488 in 2022.

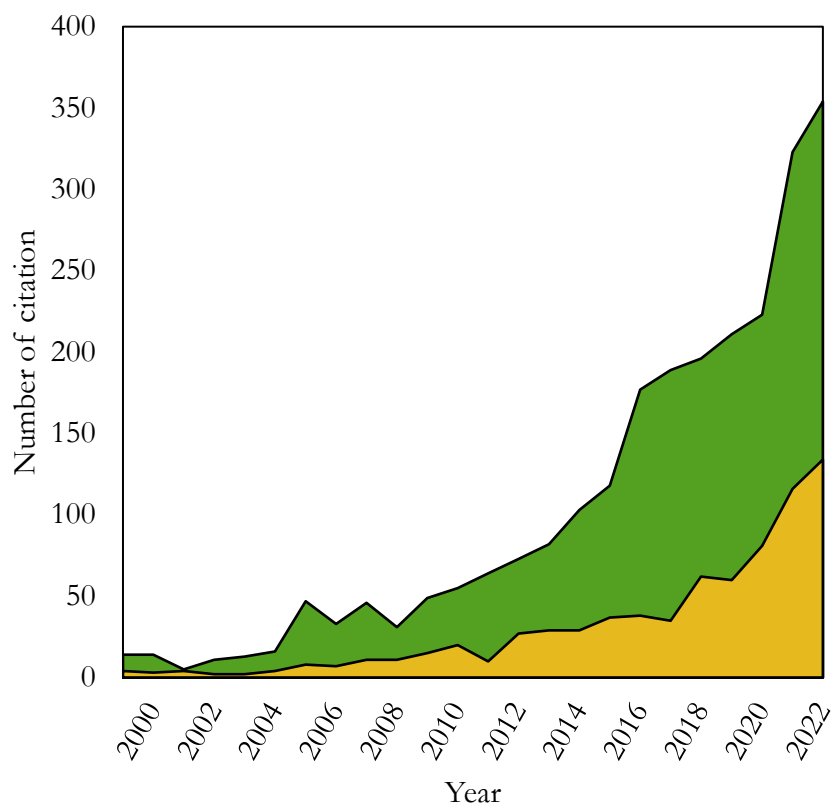


Figure 3.4 – Several publications about the COSMO model, including articles, conference papers, book chapters, reviews, and conference reviews covering the 1999 – 2022 periods. Green: COSMO-SAC and yellow: COSMO-RS.

In the following sections, important models for understanding this manuscript will be well described, such as Scatchard-Hildebrand, Wilson, NRTL, UNIFAC, and COSMO models.

3.6.1. Scatchard-Hildebrand model

The regular solution model developed by Scatchard and Hildebrand introduced the concept of cohesive energy density, leading to Eq. 3.35.

$$RT \ln \gamma_1 = v_1 \Phi_1^2 (\delta_1 - \delta_2) \quad (3.35)$$

Where v_i is the molar volume of the subcooled liquid pure compound i , Φ_1 is the volume fraction of i defined by: $\Phi_1 = x_1 v_1 / x_1 v_1 + x_1 v_2$, and δ_i is called the solubility parameter.

The solubility parameter (Eq. 3.36) is defined by the relationship between the enthalpy of vaporization and the molar liquid fraction.

$$\delta_i = \left(\frac{\Delta H_i^{vap} - RT}{V_m} \right)_i^{1/2} \quad (3.36)$$

This model does not require any parameters other than the solubility parameters, which can be calculated or easily found in databases.

3.6.2. WILSON

Wilson introduced the local composition concept in relation to thermodynamic excess free energy function to develop his thermodynamic model. He suggested that the local fraction x_{ji} of type i molecules and the local fraction x_{ij} of type j molecules which are in the immediate vicinity of a type i molecule can be related in terms of the overall mole fractions and two Boltzmann factors as Eq. 3.37.

$$\frac{x_{ji}}{x_{ii}} = \frac{x_j e^{-\frac{g_{ji}}{RT}}}{x_i e^{-\frac{g_{ii}}{RT}}} \quad (3.37)$$

where the $-g_{ji}$ and $-g_{ii}$ are energies of interaction between molecules of types j and i , and two types i molecules, respectively.

The generalized form of Wilson's expression for a binary mixture is Eq. 3.38.

$$\frac{g^E}{RT} = - \sum_i x_i \ln \left(\sum_j x_j \Lambda_{ij} \right) \quad (3.38)$$

Where Λ_{ij} is:

$$\Lambda_{ij} = \frac{v_2}{v_1} e^{-\frac{g_{ij}-g_{ji}}{RT}} \quad (3.39)$$

It is important to highlight that Wilson model is primarily used for describing non-ideal behavior in VLE and is not suitable for predict LLE.

3.1.1. NRTL

The NRTL model is described by equation (3.40).

$$\ln \gamma_i = \frac{\sum_{j=1}^c (x_j \tau_{ji} G_{ji})}{\sum_{k=1}^c (G_{kj} x_k)} + \sum_{j=1}^c \left[\frac{(x_j G_{ij})}{\sum_{k=1}^c (G_{kj} x_k)} \left(\tau_{ij} - \frac{\sum_{k=1}^c (x_k \tau_{kj} G_{kj})}{\sum_{k=1}^c G_{kj} x_k} \right) \right] \quad (3.40)$$

With $G_{ij} = \exp(-\alpha_{ij} \tau_{ij})$ and $\tau_{ij} = \frac{g_{ij}-g_{jj}}{RT}$, where $g_{ij} - g_{jj}$ is the binary interaction parameter. The temperature dependence is evaluated according to the following expressions (Eq. 3.41 and Eq. 3.42) included in the Simulis Thermodynamic™ software:

$$g_{ij} - g_{jj} = C_{ij}^0 + C_{ij}^T (T - 273.15) \quad (3.41)$$

$$\alpha_{ij} = \alpha_{ij}^0 + \alpha_{ij}^T (T - 273.15) \quad (3.42)$$

The parameter α_{ij} are the non-randomness parameters. In this work, α_{ij}^0 can be set from 0.2, 0.3, 0.4, and 0.47 as recommended by Renon and Prausnitz (1968)¹²⁰ and compared with an optimized one.

The binary interaction parameters of the NRTL are adjusted with the experimental data using the following objective function Eq. 3.43 for a Flash(P, T) type calculation.

$$OF = \sum_{k=1}^N \left(\frac{x_k^{exp} - x_k^{cal}}{u(x)} \right)^2 + \sum_{k=1}^N \left(\frac{y_k^{exp} - y_k^{cal}}{u(y)} \right)^2 \quad (3.43)$$

Where x_k^{exp} and x_k^{cal} are the experimental and calculated mole fractions of liquid (and y for the mol fraction of vapor) for each component over the coexistence line for the k th value, respectively. N is the number of points in the data set. $u(x)$ corresponds to the uncertainty of measured composition.

3.6.3. UNIFAC

The activity coefficient is represented by two contributions as in Eq. 3.44.

$$\ln \gamma_{i/S} = \ln \gamma_{i/S}^{comb} + \ln \gamma_{i/S}^{res} \quad (3.44)$$

The combinatorial term (Eq. 3.45), given by the Staverman-Guggenheim^{121,122} contribution, is represented by the entropic effects from the molecular volume and surface area (obtained in the first step of the model).

$$\gamma_{i/S}^{comb} = \ln \frac{\Phi_i}{x_i} + \frac{z}{2} q_i \ln \frac{\theta_i}{\Phi_i} + l_i - \frac{\Phi_i}{x_i} \sum_j x_j l_j \quad (3.45)$$

The molecular volume parameter is given by $\Phi_i = r_i x_i / \sum_j r_j x_j$, the surface area fraction is given by $\theta_i = q_i x_i / \sum_j q_j x_j$, and $l_i = z/2 (r_i - q_i) - (r_i - 1)$.

The residual contribution is represented by Eq. 3.46.

(3.46)

$$\gamma_{i/s}^{res} = q_i \left[1 - \ln \left(\sum_j \theta_j \tau_{ji} \right) - \sum_j \frac{\theta_j \tau_{ji}}{\sum_k \Phi_k \tau_{ji}} \right]$$

3.7. COSMO-based model

In the COSMO-based methods (COSMO-RS and COSMO-SAC), the key elements representing the link between the quantum chemical computations and the thermodynamic quantities are described by the so-called **sigma-profile**. The process to obtain the sigma-profile begins with a Quantum Chemistry (QC) calculation, which usually is performed using Density Functional Theory (DFT) method. In this calculation the electronic structure of the molecule in the presence of an isotropic solvent represented by a polarizable continuum media is determined. The electrostatic contribution to the solvation free energy is then computed using apparent surface charges located at the interface between the molecular cavity and the dielectric media. These charges are generated by the polarization induced by the molecule (solute) on the dielectric media (solvent). The result of this calculation is the screening charge density. Chapter 4 will provide a detailed explanation of DFT methods.

The screening charge density ($\sigma(s)$) is a perfect local qualitative measure of the molecular polarity and can be used to quantify and color-code molecular polarity on the surface. Figure 3.5 shows the color-coded sigma surfaces of (a) water, (b) furan, and (c) n-butanol which red areas represent strongly negative parts of the molecular surface and hence a strongly positive value of σ . On the other hand, deep blue areas represent strongly positive surface regions, and a strongly negative value of σ and green areas represent the nonpolar surface.

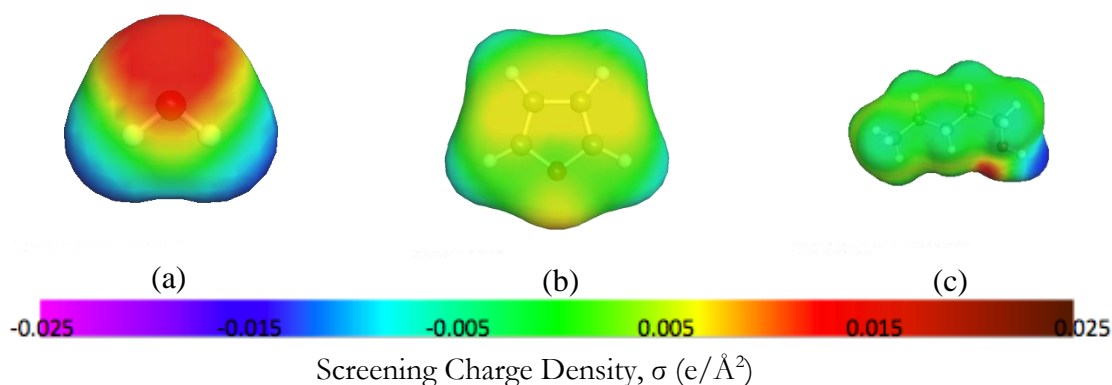


Figure 3.5 – COSMO surface of (a) Water, (b) Furan, and (c) n-butanol color-coded by the screening charge density.

3.7.1. Sigma-profile

Starting from the polarizable continuum model calculations, the screening charge is averaged by Eq. 3.47.

$$\sigma_m = \frac{\sum_n \sigma_n^* \frac{r_n^2 r_{eff}^2}{r_n^2 + r_{eff}^2} \exp\left(-f_{decay} \frac{d_{mn}^2}{r_n^2 + r_{eff}^2}\right)}{\sum_n \frac{r_n^2 r_{eff}^2}{r_n^2 + r_{eff}^2} \exp\left(-f_{decay} \frac{d_{mn}^2}{r_n^2 + r_{eff}^2}\right)} \quad (3.47)$$

Where σ_n^* is the original surface charge of the n th segment and r_n is radius of the segment n , both taken from the COSMO file. The d_{mn} is the distance (in Å) between the centers of the surface segments n and m calculated by Eq 3.48. All the parameters used to the sigma-profile construction are defined in table 3.1.

$$d_{mn} = \sqrt{(x_m - x_n)^2 + (y_m - y_n)^2 + (z_m - z_n)^2} \quad (3.48)$$

Position data of all segments are given in atomic units and can be converted to the Angstrom by multiplying with the conversion factor 0.52917721067 Å/a.u.

Table 3.1 – Parameters used in the sigma-profile calculation.

Parameters		
a_{eff}	7.5 Å ²	effective area for averaging
r_{eff}	$= \sqrt{a_{eff}/\pi}$	effective radius for averaging
f_{decay}	3.57	empiric parameter

Once the screening charge averaged has been obtained for each segment m , the sigma-profile, $p_i(\sigma)$, is created. This involves considering the probability of finding a specific segment with a given value of σ , multiplied by the surface area of the molecule i , as defined by Eq. 3.49.

$$p_i(\sigma) = \frac{A_i(\sigma)}{A_i} \quad (3.49)$$

Where $A_i(\sigma)$ represents the surface cavity area with charge density σ_m of the i molecule, A_i is the total cavity surface area. Lin and Sandler have defined $A_i(\sigma) = a_{eff} n_i(\sigma)$, where a_{eff} is the effective surface area of a standard surface segment. This effective surface represents the contact area between different molecules, e.g. a theoretical bonding site. Klamt set this adjustable parameter to 7.1 Å². In this work we set to 7.5 Å², as indicated in Table 3.1.

Once the sigma-profile is established for pure compounds, it is then assessed for mixtures by calculating the weighted sums of contributions from all pure components (Eq. 3.50).

$$p_s(\sigma) = \frac{\sum_i x_i n_i p_i(\sigma)}{\sum_i x_i n_i} = \frac{\sum_i x_i A_i p_i(\sigma)}{\sum_i x_i A_i} \quad (3.50)$$

The values of σ vary in a range of -0.025 to $+0.025$ $e/\text{\AA}^2$ and in increments of 0.001 $e/\text{\AA}^2$, forming a set of 51 points.

This histogram is called sigma-profile (σ -profile) (Figure 3.6) and describes three interaction regions; hydrogen bond (HB) acceptor, hydrogen bond donor, and nonpolar region, and it could be a qualitative measure of the molecular polarity. The non-polar region shows peaks between -0.01 and $+0.01$ $e/\text{\AA}^2$, the HB donor region shows peaks on the extreme left, and the HB acceptor shows peaks on the extreme right of figure 3.7.

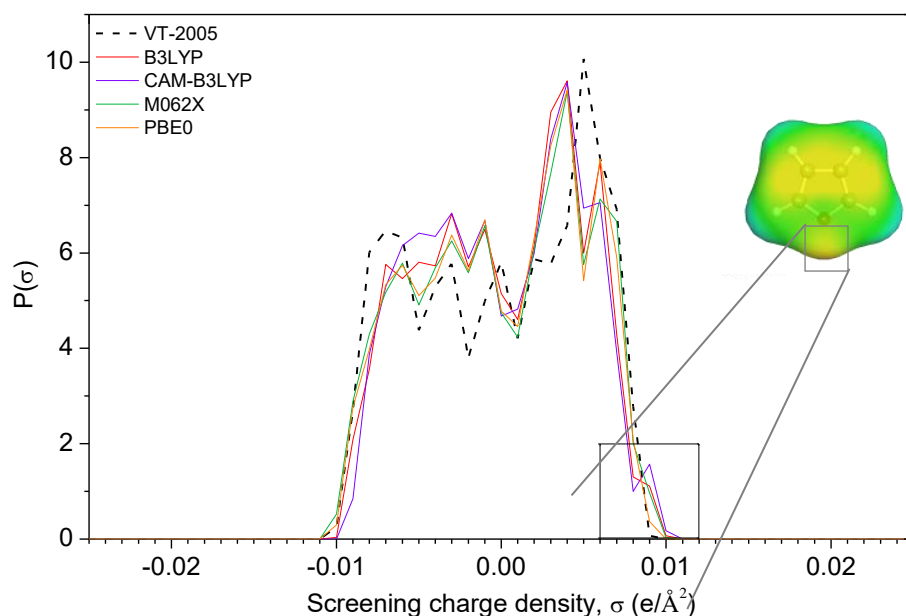


Figure 3.6 – Sigma-profile of Furan for different exchange-correlation Functionals using 6-311G(2df,p) as a basis set in comparison with the VT 2005 Sigma Profile Database.

From the σ -profile, the chemical potentials of species in solution can be predicted, which can be used, through a thermodynamic approach, for evaluating several properties,, including IDACs.

3.7.2. COSMO-SAC – activity coefficient prediction

In this second step of the model, the activity coefficient is calculated by combining the two contributions, as in the UNIFAC and modified UNIFAC, the combinatorial and the restoring contribution (Eq. 3.51).

$$\ln\gamma_{i/S} = \ln\gamma_{i/S}^{comb} + \ln\gamma_{i/S}^{res} \quad (3.51)$$

The combinatorial term, given by the Staverman-Guggenheim^{121,122} contribution (Eq. 3.52), is represented by the entropic effects from the molecular volume and surface area (obtained in the first step of the model).

$$\gamma_{i/S}^{comb} = \ln \frac{\Phi_i}{x_i} + \frac{z}{2} q_i \ln \frac{\theta_i}{\Phi_i} + l_i - \frac{\Phi_i}{x_i} \sum_j x_j l_j \quad (3.52)$$

The molecular volume parameter is given by $\Phi_i = r_i x_i / \sum_j r_j x_j$, the surface area fraction is given by $\theta_i = q_i x_i / \sum_j q_j x_j$, and $l_i = z/2 (r_i - q_i) - (r_i - 1)$. Where x_i is the molar fraction of component i , r_i and q_i are the volume and surface area for i normalized to a standard volume of 66.69 \AA^3 ,¹²³ and surface area of 79.53 \AA^2 ,¹²³ respectively. z is the coordination number, usually taken to be 10 and the summation is over all the species in the mixture.

The restoring free energy of a solute molecule i from the ideal conductor to the real solvent is obtained as the summation over the segment activity coefficients (Eq. 3.53 and 3.54).

$$\frac{\Delta G_{i/S}^{*res}}{RT} = n \sum_{\sigma_m} p(\sigma_m) \ln \Gamma_S(\sigma_m) \quad (3.53)$$

$$\ln \Gamma_S(\sigma_m) = -\ln \left\{ \sum_{\sigma_n} p_s(\sigma_n) \Gamma_S(\sigma_n) \exp \left[\frac{\Delta W(\sigma_m, \sigma_n)}{kT} \right] \right\} \quad (3.54)$$

ΔW is the exchange energy, and k is the Boltzmann constant.

$n = A_i/a_{\text{eff}}$ is the number of segments in the molecule, which is the ratio of the total surface area of a single molecule i (A_i) to the area of the standard surface of segments (a_{eff}).

$p(\sigma_m)$ is a two-dimensional histogram.

Finally, the combinatorial sum and the restoring contribution are used to obtain the activity coefficient, as shown in Eq. 3.51 Figure 3.7 summarizes the steps of the model COSMO-based approach.

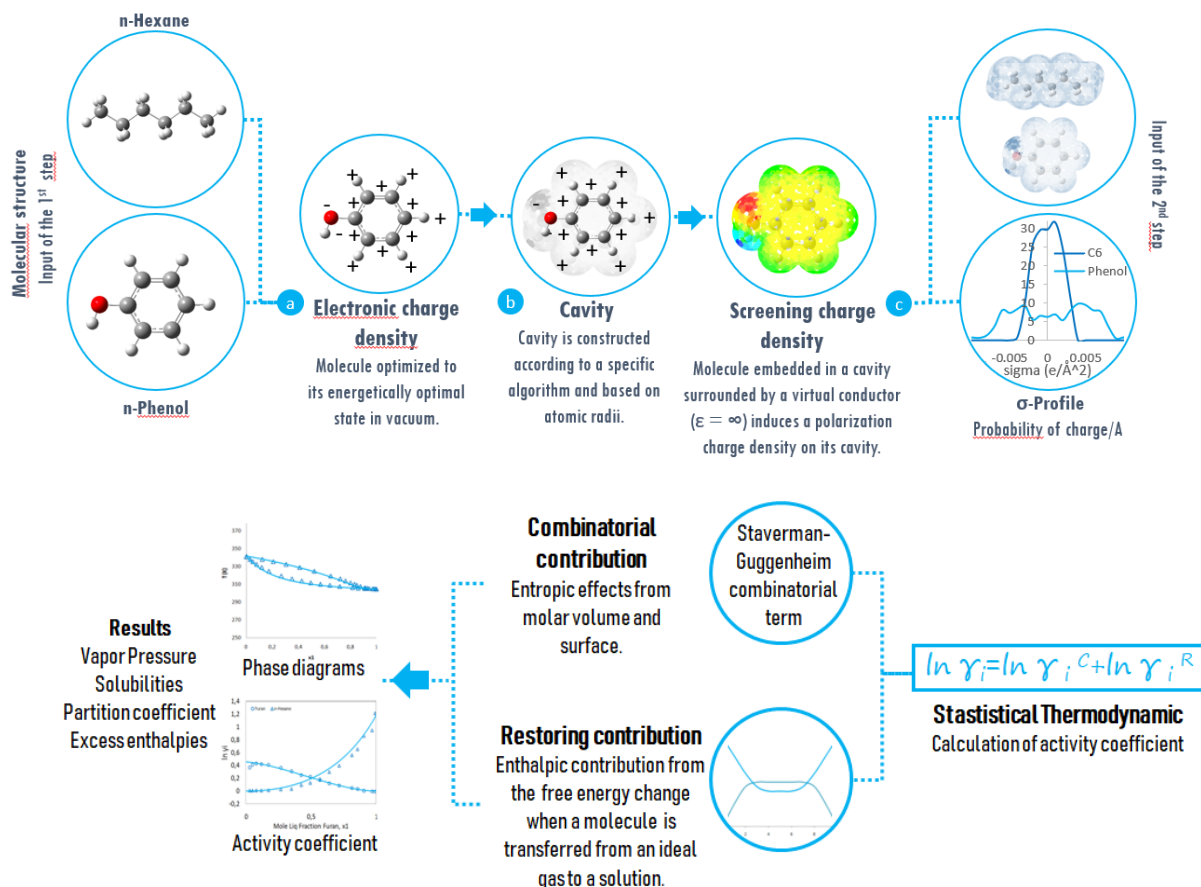


Figure 3.7 – COSMO-SAC steps

3.8. Conclusion

In summary, this chapter has provided a comprehensive exploration of thermodynamic modeling in the context of phase equilibrium and thermodynamic properties calculation. Furthermore, the chapter delved into the concept of activity coefficient and its models. Several widely used activity coefficient models (including the Scatchard-Hildebrand model, WILSON, NRTL, UNIFAC and COSMO-based model) were elucidated. The detailed explanation of the COSMO-SAC model, encompassing its sigma-profile and how the model predicts activity coefficient, enriched the understanding of this model, focus of these these.

At the heart of the COSMO-based model are the sigma-profiles, acting as a crucial link between quantum chemical calculations and thermodynamic properties. These profiles are generated through quantum chemical calculations, typically using Density Functional Theory (DFT). The upcoming chapter will

elucidate the foundational principles of quantum chemistry approaches and their integration with COSMO-based models.

Chapter 4

COSMO Solvation model and the Quantum Chemistry

Le modèle de solvation COSMO et la chimie quantique

Résumé

Ce chapitre se concentre sur les modèles basés sur COSMO (en particulier COSMO-SAC) et leurs intégrations avec des approches basées sur la chimie quantique. Au cœur du modèle basé sur COSMO se trouve le profil sigma, un lien essentiel qui relie les calculs de chimie quantique aux propriétés thermodynamiques. Ce profil provient des calculs de chimie quantique, généralement effectués à l'aide de la théorie de la fonctionnelle de la densité (Density Functional Theory : DFT).

Dans le cadre de la DFT, les caractéristiques d'un système à plusieurs électrons sont plutôt déterminées par la densité électronique que par les fonctions d'onde des électrons individuels ; simplifiant ainsi les calculs et fournissant des résultats précis pour une large gamme de systèmes. Cette approche facilite la prédiction et l'analyse de diverses propriétés (notamment les aspects électroniques, structuraux et thermodynamiques des matériaux et des molécules). La DFT est souvent considérée comme une méthode précise pour résoudre le problème de la mécanique quantique à plusieurs corps indépendants du temps.

Avant d'examiner la structure formelle de DFT, nous introduirons d'abord les principes fondamentaux de la chimie quantique (Quantum Chemistry : QC). La chimie quantique est dédiée à la description des systèmes atomiques, moléculaires et étendus en résolvant l'équation de Schrödinger. À mesure que le nombre d'électrons augmente, la complexité de l'équation de Schrödinger augmente, limitant la disponibilité de solutions analytiques aux cas plus simples tels que l'hydrogène et l'hélium. Pour les systèmes multiélectroniques, des approximations deviennent nécessaires telle que l'approximation de Born-Oppenheimer qui permet de séparer la dynamique électronique et nucléaire. L'hamiltonien électronique (dérivé de l'hamiltonien à plusieurs corps) caractérise la fonction d'onde électronique et l'énergie.

Pour construire des fonctions d'onde multiélectroniques, on utilise des orbitales moléculaires. Les électrons, en tant que fermions, respectent le principe d'exclusion de Pauli qui exige une antisymétrie dans leurs fonctions d'onde. Alors que le produit de Hartree ne parvient pas à satisfaire cette exigence, l'introduction du Déterminant de Slater résout ce problème satisfaisant mathématiquement le principe d'antisymétrie. Il existe deux principales catégories de méthodes pour résoudre l'équation électronique : la méthode de Hartree-Fock (HF) et la DFT. HF est basée sur la théorie des orbitales moléculaires et repose sur la construction d'une fonction d'onde approchée en utilisant un unique déterminant de Slater. Tandis que la DFT est une approche qui se concentre sur la densité électronique elle-même, en reformulant le

problème des nombreux corps en termes de densité électronique. Cette dernière utilise les théorèmes de Hohenberg, Kohn and Sham pour déterminer les propriétés de l'état fondamental du système. Suivant la méthode de Kohn-Sham, la densité est exprimée le plus souvent en fonction d'un déterminant de Slater d'orbitales moléculaires (monoélectroniques). Ces orbitales de Kohn-Sham ne sont pas identiques aux orbitales HF : ce sont celle d'un système fictif d'électrons sans interaction, mais possédant la même densité que le système réel. Dans cette étude, nous nous basons principalement sur des méthodes en DFT.

Les fondements de la DFT reposent sur les Théorèmes de Hohenberg-Kohn. Le premier théorème établit que la densité électronique dicte de manière unique l'Hamiltonien, englobant le potentiel et l'ensemble des propriétés du système. Cette relation fonctionnelle de la densité est exprimée à travers la fonctionnelle de l'énergie (un opérateur mathématique qui associe la densité électronique à l'énergie totale du système). Le deuxième théorème (connu sous le nom de théorème de variation) facilite la détermination de l'énergie électronique de l'état fondamental. En commençant par une densité électronique approximative, on peut appliquer des principes de variation pour trouver la densité qui minimise l'énergie de l'état fondamental. Pour calculer l'énergie dans le cadre de la DFT, la méthode de Kohn-Sham est couramment employée. Elle répartit l'énergie totale en composantes d'énergie cinétique, d'interaction électron-électron, d'interaction nucléaire-électron et d'énergie d'échange-corrélation. L'énergie d'échange-corrélation est un élément clé de la DFT et sa formulation précise reste un défi redoutable. Diverses approximations sont utilisées notamment les approximations de la densité locale, celles corrigées par gradient et les fonctionnelles hybrides, chacune offrant des niveaux de précision variables. De plus, le choix de la base d'orbitales joue un rôle essentiel dans les calculs de la DFT.

Les bases sont des ensembles de fonctions utilisées pour décrire les orbitales électroniques, et elles peuvent être classées en bases minimales et en bases à valence séparée, avec ajout de fonction de polarisation ou de fonctions diffuses. La taille de la base influe considérablement sur la précision des calculs, les ensembles plus volumineux offrant une plus grande précision au prix de ressources computationnelles accrues.

En chimie computationnelle, la prise en compte des effets du solvant est essentielle pour comprendre comment les solvants influencent la structure électronique et les caractéristiques d'un système. Deux approches principales sont utilisées : les modèles de solvation explicites et implicites. Les modèles explicites incluent la représentation directe des molécules de solvant dans le cadre computationnel, offrant une description détaillée des interactions soluté-solvant spécifiques, au prix de ressources computationnelles substantielles. Les modèles implicites considèrent le solvant comme un milieu continu avec une constante diélectrique, approximant sa polarisabilité. Ces modèles divisent l'énergie de solvation en composantes électrostatiques et non électrostatiques, incluant la répulsion de Pauli, la dispersion, la

liaison hydrogène et l'énergie de cavitation. Les modèles implicites utilisent l'équation de Poisson pour calculer le potentiel électrostatique et font appel à des surfaces de cavité, souvent définies par des sphères, pour décrire les interactions du soluté avec le solvant. Les cavités de soluté peuvent être classées comme empiriques ou non empiriques avec diverses techniques pour construire des surfaces de cavité.

Une approche courante consiste à utiliser des sphères centrées sur les atomes en se basant sur les rayons de van der Waals. Le Modèle de Screening de Type Conducteur (COSMO) ou le modèle du Continuum Polarisable type Conducteur (C-PCM) est une méthode de solvation implicite largement utilisée, incorporant des facteurs d'ajustement tels que les rayons de Klamt pour différents éléments. Dans les calculs, la surface de la cavité est subdivisée en de plus petits éléments de surface appelés « tesselles », qui fournissent une approximation de la surface moléculaire.

Ce qui distingue le modèle COSMO des autres modèles est son traitement unique du milieu diélectrique comme un conducteur. Les conducteurs se caractérisent par une permittivité relative infinie ($\epsilon = +\infty$), simplifiant le problème électrostatique en éliminant les gradients de potentiel au sein du milieu. Dans le modèle COSMO, la densité de charge de surface ($\sigma^\infty(s)$) sur la surface de la cavité est directement déterminée par le potentiel électrostatique généré par la densité de charge du soluté, puis elle est ajustée à l'aide d'une fonction basée sur la permittivité relative réelle du solvant ($f(\epsilon)$) pour obtenir la densité de charge appropriée ($\sigma(s)$).

La procédure du modèle COSMO commence par optimiser la géométrie de la molécule dans le vide, puis la place dans une cavité entourée d'un conducteur virtuel. Ce modèle décompose la surface de la cavité moléculaire en segments plus petits et la densité électronique du soluté à l'intérieur de la cavité induit une polarisation dans le milieu diélectrique, dans le continuum et dans le solvant. En conséquence, le soluté lui-même se polarise, ce qui modifie sa distribution électronique. La densité de charge résultant de cette polarisation sur les segments de surface est ensuite calculée, permettant l'analyse de la polarité moléculaire et sa conversion en profil sigma.

COSMO Solvation model and the Quantum Chemistry

4.1. Introduction

The COSMO-based models, particularly COSMO-SAC, constitute a central focus of this work. These models, including COSMO-SAC and COSMO-RS have demonstrated their efficacy in predicting phase equilibrium and thermodynamic properties through a two-step calculation process that intricately combines quantum chemistry (QC) and statistical thermodynamics. This chapter is dedicated to investigating the integration of COSMO-based model with quantum chemistry-based approaches.

At the heart of the COSMO-based model lies the σ -profile, a critical link connecting quantum chemical calculations to thermodynamic properties. This profile originates from QC computations and is typically conducted using Density Functional Theory (DFT). These computations focus on the molecular electronic structure within a polarizable continuum, simulating an isotropic solvent using a solvation model.

The COSMO solvation model, which represents the QC step, is implemented in several quantum chemical computational packages, such as Gaussian. Belonging to the family of dielectric continuum solvation models (CSM)s, the COSMO solvation model is based on an implicit approach. It does not explicitly introduce solvent molecules into the chemical system being analyzed. Instead, it views the solvent as a homogeneous, polarizable medium characterized by its dielectric constant (ϵ). The solute is placed within a cavity created within this medium. The electron density of the solute inside the cavity induces polarization in the dielectric, continuum, and the solvent, in turn, polarizes the solute, modifying its electron distribution. The electrostatic contribution to the solvation Gibbs energy is evaluated, in this reciprocal polarization process.

In DFT, the properties of a many-electron system are determined by the electron density rather than the wave functions of individual electrons, simplifying than the calculations while providing accurate results for a wide range of systems. This approach allows the prediction and analysis of various properties (including electronic, structural, and thermodynamic aspects of materials and molecules). DFT is often considered as an exact approach to solve the time-independent many-body problem. Before reviewing the formal framework of DFT and CSMs, we introduce the many-body problem.

4.2. Theoretical Background

Quantum chemistry (QC) aims to provide comprehensive descriptions of atomic, molecular, and extended systems, enabling the prediction of their properties from fundamental principles. This is accomplished by solving the non-relativistic, time-independent Schrödinger equation¹²⁴ for a given system containing N electron (Eq. 4.1).

$$\hat{H}\Psi(x_1, \dots, x_N) = E\Psi(x_1, \dots, x_N) \quad (4.1)$$

Where \hat{H} is the Hamiltonian operator, $\Psi(x_1, \dots, x_N)$ is the many-body wave function, which contains all information on the quantum state of the system, and E is the total energy of the system in state Ψ .

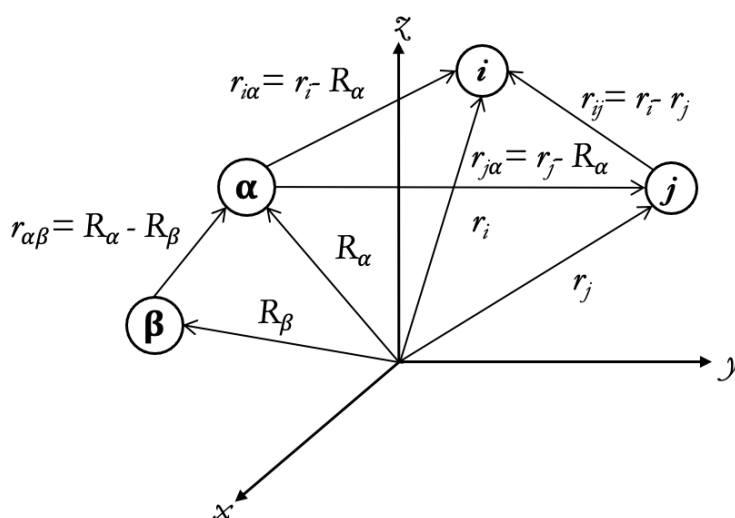
For a system of M nuclei and N electrons, the non-relativistic Hamiltonian operator is written as a sum of kinetic and potential energies (Eq. 4.2).

$$\hat{H} = \hat{T}_e + \hat{T}_N + \hat{V}_{Ne} + \hat{V}_{ee} + \hat{V}_{NN} \quad (4.2)$$

Where \hat{T}_e and \hat{T}_N are the operators for the kinetic energy of electrons and nuclei, respectively. Terms denoted by \hat{V} are the electrostatic term, representing the Coulomb attraction between electrons and nuclei (\hat{V}_{Ne}), the electron-electron repulsion (\hat{V}_{ee}), and internuclear repulsion (\hat{V}_{NN}). If we assume the nuclei and electrons to be point masses and if we neglect spin-orbit and other relativistic interactions, the atomic or molecular Hamiltonian can be written by Eq. 4.3.

$$\hat{H} = \sum_i^N \left(\frac{-\hbar^2}{2m_e} \nabla_i^2 \right) + \sum_\alpha^M \left(\frac{-\hbar^2}{2M_\alpha} \nabla_\alpha^2 \right) - \sum_\alpha^M \sum_i^N \frac{eZ_\alpha}{4\pi\epsilon_0 r_{i\alpha}} + \sum_i^{N-1} \sum_{j>i}^N \frac{e^2}{4\pi\epsilon_0 r_{ij}} + \sum_\alpha^{M-1} \sum_{\beta>\alpha}^M \frac{e^2 Z_\alpha Z_\beta}{4\pi\epsilon_0 r_{\alpha\beta}} \quad (4.3)$$

Where subscripts i and j (α and β) run over all electrons (nuclei); q and m_e (Z and M) are the charge and mass of an electron (nucleus); r is the inter-particle distance as shown in figure 4.1; \hbar is the Plank's constant divided by 2π and ∇^2 is the Laplacian differential operator. The restriction $j > i$ ($\beta > \alpha$) avoids the double counting of the same interelectronic (internuclear) repulsion and the counting of the self-repulsion.

Figure 4.1 – A molecular coordinate system: i and j = electrons; α and β = nuclei.

Atomic unit (a.u.) is system of natural units commonly used in quantum physics and chemistry. When dealing with expressions involving several small-valued constants, such as the electron mass (m_e) and reduced Planck's constant (\hbar), this unit system simplifies calculations by setting a selected group of constants equal to 1. Table 4.1. summarize the constants expressed in atomic units (a.u.) and their respective values in the international system (SI) are provided for reference.

Table 4.1 – Fundamental constants in atomic unit.¹²⁵

Quantity Set Equal to 1	Constant Expression	Value in International system unit (SI)
Mass of an Electron	$m_e = 1$	$9.1093835 \times 10^{-31}$ kg
Elementary Charge	$e = 1$	$1.6021765 \times 10^{-19}$ C
Reduced Planck's Constant (\hbar)	$\hbar = 1$	$1.0545716 \times 10^{-34}$ J·s
Coulomb Constant	$\frac{e^2}{4\pi\epsilon_0} = 1$	8.98755179×10^9 N·m ² /C ²
Bohr Radius	$a_0 = \frac{4\pi\epsilon_0\hbar^2}{m_e e^2} = 1$	$5.2917721 \times 10^{-11}$ m
Hartree Energy	$\frac{m_e e^4}{(4\pi\epsilon_0)^2 \hbar^3} = 1$	$4.3597446 \times 10^{-18}$ J
Time for one electron travel one period in the first Bohr orbit	$\frac{(4\pi\epsilon_0)^2 \hbar^3}{m_e e^4} = 1$	$2.4188843 \times 10^{-17}$ s
Permittivity of a vacuum times 4π	$4\pi \epsilon_0 = 1$	1.112650×10^{-12} C ² /(J·m)

This choice simplifies mathematical expressions, allowing for a more compact and elegant form of the Hamiltoniana (Eq. 4.4).

$$\hat{H} = - \sum_i^N \left(\frac{1}{2} \nabla_i^2 \right) - \sum_\alpha^M \left(\frac{1}{2M_\alpha} \nabla_\alpha^2 \right) - \sum_\alpha^M \sum_i^N \frac{Z_\alpha}{r_{i\alpha}} + \sum_i^{N-1} \sum_{j>i}^N \frac{1}{r_{ij}} + \sum_\alpha^{M-1} \sum_{\beta>\alpha}^M \frac{Z_\alpha Z_\beta}{r_{\alpha\beta}} \quad (4.4)$$

The complexity of the Schrödinger equation increases dramatically with the number of electrons in the system, and thus the analytical solution of this equation is only possible for hydrogen and helium atoms and simple molecules such as H₂. However, it can be solved for multi-electron systems by applying some approximations.

The many-body Hamiltonian in Eq. 4.3 describes both the motion of the electrons and that of the nuclei. However, electrons and nuclei dynamics happening on a very different timescale. In fact, due to the smaller mass electrons move much faster than nuclei, we can solve only the electronic Schrödinger equation for a given fixed nuclear geometry thus separating out the nuclear wave function, by virtue of the so-called Born-Oppenheimer (BO) approximation¹²⁶. The many-body Schrödinger electronic equation can thus be written as follows:

$$\hat{H}_{el} \Psi_{el} = E_{el} \Psi_{el} \quad (4.5)$$

where \hat{H}_{el} is the electronic Hamiltonian, Ψ_{el} is the electronic wave function, and E_{el} is the total electronic energy.

Therefore, the Hamiltonian to be solved reduces to the electronic Hamiltonian is expressed by Eqs. 4.6 and 4.7.

$$\hat{H} = - \sum_i^N \left(\frac{1}{2} \nabla_i^2 \right) - \sum_\alpha^M \sum_i^N \frac{Z_\alpha}{r_{i\alpha}} + \sum_i^{N-1} \sum_{j>i}^N \frac{1}{r_{ij}} \quad (4.6)$$

$$\hat{H} = \hat{T}_e + \hat{V}_{Ne} + \hat{V}_{ee} \quad (4.7)$$

The total energy of the system is thus expressed as the sum of the electronic and the nuclear repulsion energy (Eq. 4.8).

$$E_{tot} = E_{el} + E_{NN} \quad (4.8)$$

From the Eq. 4.7, the two first terms (the electronic kinetic energy, and nuclear repulsion) can be calculated directly. While the third one, the electron-electron repulsion, cannot be solved exactly for the many-electron system. Over the years, different methods have been developed to solve Schrödinger's electronic equation, referring to different types of approximations. These methods can be divided into two categories methods based on Hartree-Fock¹²⁷ and methods following the DFT. The studies presented in this study have been conducted using methods rooted in the second.

4.2.1. Slater determinants

The molecular orbitals are used to construct the multi-electronic wave function of the system. Since the electrons are fermions particles, they obey the Pauli exclusion principle. According to this principle, a wave function describing a collection of fermions must exhibit anti-symmetry when exchanging the coordinates of two electrons. The Hartree Product, which is a product of single-electron wave functions, fails to satisfy this anti-symmetry requirement. In response to this limitation, Fock and Slater introduced a solution^{128,129}: they replaced the Hartree Product with the Slater Determinant. The Slater Determinant, defined by Eq. 4.9, forms the polyelectronic wavefunction and mathematically represents the anti-symmetry required by the Pauli principle. The polyelectronic wavefunction (Ψ) for a system of N electrons is written as a determinant of N mono-electronic wavefunctions (φ):

$$\Psi_{Slater}(x_1, x_2, \dots, x_N) = \frac{1}{\sqrt{N!}} \begin{bmatrix} \varphi_1(x_1) & \dots & \varphi_N(x_1) \\ \vdots & \ddots & \vdots \\ \varphi_1(x_N) & \dots & \varphi_N(x_N) \end{bmatrix} \quad (4.9)$$

The mono-electronic functions expressed in Eq. 4.10 are called spin-orbitals.

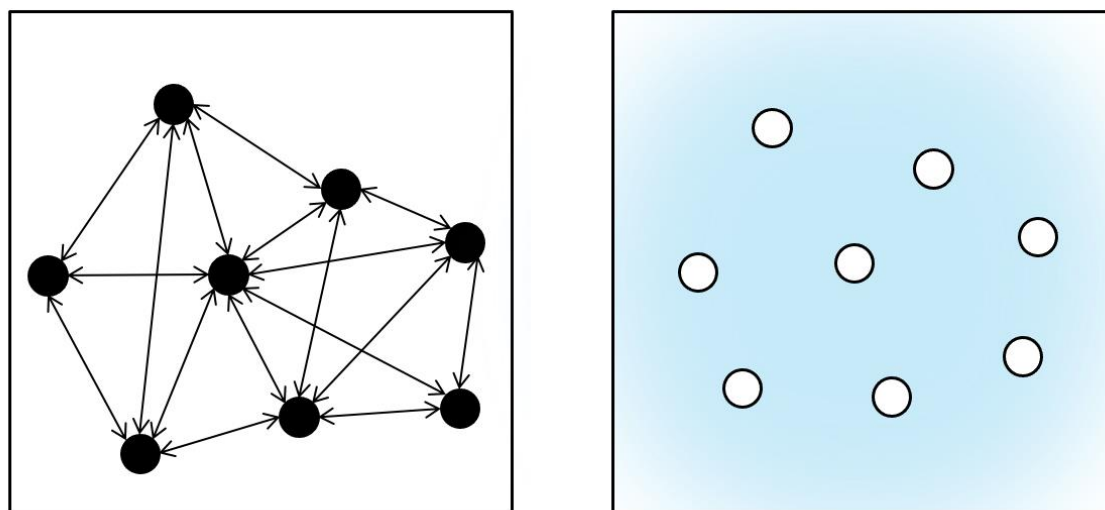
$$\varphi_i(x) = \varphi_i(r)\sigma(s) \quad (4.10)$$

which depends on both the spatial coordinates (r) and the spin coordinates (s) of an electron. The spatial orbital, $\varphi(r)$, describes the spatial distribution of an electron, while the spin function, $\sigma(s)$, indicates the electron's spin (which can be either alpha or beta). It's important to note that, in principle, a wavefunction could be composed of a linear combination of several Slater determinants using multi-reference methods. However, many commonly used methods, especially those discussed here, rely on single-reference methods. In this approach, the ground state is described by a single Slater determinant.

4.3. Density Functional Theory

In a system composed of N particles, the Schrödinger equation can be exactly written considering the Coulombic interactions between electrons (figure 4.2a.) The dimensionality of the quantum state space increases dramatically by the number of particles. For each particle N , we need three coordinates (x, y, z) to describe its position in three dimensions. Therefore, the dimensionality of the entire system's wavefunction is $3N$. Consequently, finding an analytical solution to this general N -particle Schrödinger equation is exceedingly challenging from a computational perspective. This is where approximations like DFT become essential to allow the description of the electronic structure of molecular systems.

In this theory, instead of dealing with wavefunctions, the energy and the properties of every N electron (N : total number of electrons) system are described using electronic density as the main variable. This is demonstrated in figure 4.2b. This methodology simplifies a problem dependent on $3N$ cartesian coordinates for a system of N electrons into a unique electronic density problem, defined by only three spatial coordinates (x,y,z) represented by r .



(a)

(b)

Figure 4.2 – (a) Schematic representation of electrons denoted by black circles with arrows illustrating Coulomb interactions between some electrons. (b) Illustration of a mean-field DFT approach, where DFT quasiparticles are depicted as white circles moving within a background field generated by the electric field of other particles.

The electronic density $\rho(r)$ and the wave function are linked by Eq. 4.11.

$$\rho(r) = \Psi^*(r)\Psi(r) = |\Psi^2(r)| \quad (4.11)$$

N is directly correlated to the electron density by the integration of the above defined density over all space (Eq. 4.12):

$$N = \int \rho(r)dr \quad (4.12)$$

The rigorous formulation for this theory came from Hohenberg and Kohn in 1964. Their theorems provide the mathematical consistency which has contributed to confer DFT's position as one of the most used approaches in theoretical chemistry.

4.3.1. Hohenberg-Kohn Theorems

Theorem 1: “Two N-electron systems defined by two Hamiltonians with external potentials differing by more than one constant cannot have ground states with the same electronic density, i.e., the external potential is defined by the electron density of the ground state.”

In this first theorem, also known as the existence theorem, Hohenberg, and Kohn¹³⁰ demonstrated that the electron density of an N-electron system, with a given electronic interaction, uniquely determines the Hamilton operator, i.e., the potential and thus all properties of the system. The electronic energy of a chemical system can be expressed (Eq. 4.13) as a functional of the density. Functional in quantum mechanics is a mathematical operator that takes a function (e.g., a density function or a wave function) as input and maps it to a scalar, typically representing a physical quantity. For instance, the electron density function (which gives the probability of finding an electron at each point in space) is transformed by the functional into the total energy of the system (which is a function of the electron density function).

$$E[\rho(r)] = T_e[\rho(r)] + V_{ee}[\rho(r)] + V_{Ne}[\rho(r)] \quad (4.13)$$

$T_e(\rho)$, $V_{ee}(\rho)$, and $V_{Ne}(\rho)$ are the kinetic energy, the electron-electron repulsion and the nucleus-electron attraction potential energies, respectively.

The potential generated by the nuclei can be considered as an external potential, v_{ext} , from which the electrons are affected. Therefore, it is possible to generalize the Eq. 4.13 in the form of the Eq. 4.14.

$$E[\rho(r)] = F_{HK}[\rho(r)] + \int v_{ext}(r)\rho(r)dr \quad (4.14)$$

$F_{HK}(\rho)$ represents the Hohenberg-Kohn functional and incorporates both the kinetic energy of the electrons and their interactions. The last term is the contribution from the external potential generated by the nuclei. This term represents the potential energy due to the interaction of electrons with the nuclei. The Hohenberg-Kohn functional is often referred to as “universal functional” because it is independent from the external potential and therefore, it depends only on the electrons.

Theorem 2 (called the variational theorem): “Any approximate trial electronic density $\rho(r)$, which defines completely the v_{ext} and a trial wave function Ψ , can be used to determine variationally the electronic density that minimizes the energy of the ground state”.

This second theorem applies the variational principle of quantum mechanics as a resolution procedure to calculate the ground state electronic energy, $E_0[\rho]$. Starting with an approximate density, its own external

potential $v_{ext}(\mathbf{r})$ can be determined and hence, its associated wavefunction Ψ . The variational principle allows the selection of the electronic density ($E_0[\rho]$) that closely approximates the true density by minimizing the total energy of the system. This selection aims to satisfy the following property (Eq. 4.15).

$$\langle \psi(\rho) | \hat{H} | \psi(\rho) \rangle = F_{HK}[\rho(r)] + \int v_{ext}(r)\rho(r)dr = E[\rho(r)] > E_0[\rho_0(r)] \quad (4.15)$$

Where ground state energy (E_0) corresponding to the Hamiltonian in Eq. 4.7 can be mathematically expressed as Eq. 4.16.

$$E_0[\rho_0(r)] = \min_{\psi} \langle \psi(\rho) | \hat{H} | \psi(\rho) \rangle \quad (4.16)$$

The energy of a system is a priori exact thanks to the knowledge of universal functional (F_{HK}) and of the electronic density only. However, is not possible to know the analytic formula of the universal functional, as the exact formula of the electronic kinetic energy ($T_e(\rho)$) and electron-electron repulsion potential energies ($V_{ee}(\rho)$) are not known.

To overcome the principal drawback of expressing the universal functional, Kohn and Sham proposed an exact method to determine the ground state electron density of a system and consequently, the energy.

4.3.2. Kohn-Sham method

In the DFT, the Kohn and Sham¹³¹ energy functional (E^{HK-KS}) can be expressed as the sum of four terms given in Eq. 4.17.

$$E^{HK-KS}[\rho] = T_{ns}[\rho] + (T[\rho] - T_{ns}[\rho]) + V_{ee}[\rho] + (V_{ee}[\rho] - J[\rho]) + V_{ne}[\rho] \quad (4.17)$$

T_{ns} is the kinetic energy of a non-interacting electron of a given density can be exactly expressed analytically by Eq. 4.18

$$T_{ns}[\rho(r)] = \sum_i^N \langle \psi_i | \frac{1}{2} \nabla_i^2 | \psi_i \rangle \quad (4.18)$$

The subscript ns means non-interacting system. The electron-electron interaction term is represented by the classical electrostatic term plus a non-classical one J (Eq. 4.19).

$$V_{ee}[\rho(r)] = \sum_i^N \langle \psi_i | \int \frac{\rho(r')}{|r-r'|} dr' | \psi_i \rangle + J[\rho(r)] \quad (4.19)$$

And the external potential is given by Eq. 4.20.

$$V_{Ne}[\rho(r)] = \sum_i^N \langle \psi_i | \int \frac{Z_a}{|r - R_a|} dr | \psi_i \rangle \quad (4.20)$$

The use of molecular orbitals, as introduced in the section before, is essential for construct the multi-electronic wavefunction of a non-interacting system. This wavefunction is trivially represented by a Slater determinant, as described in Eq. 4.21.

$$\Psi_{Slater}(r_1, r_2, \dots, r_N) = \frac{1}{\sqrt{N!}} \begin{bmatrix} \psi_1(r_1) & \dots & \psi_N(r_1) \\ \vdots & \ddots & \vdots \\ \psi_1(r_N) & \dots & \psi_N(r_N) \end{bmatrix} \quad (4.21)$$

In Eq. 4.21 where the single particle orbitals are a set of orthonormal orbitals, each one is a solution of a set of N Schrödinger equation (Eq. 4.22).

$$H_{KS}\psi_i = \varepsilon_i\psi_i \quad (4.22)$$

The electronic Hamiltonian is written as mono-electronic Hamiltonian (Eq. 4.23).

$$H_{KS} = -\frac{1}{2}\nabla_i^2 + \int \frac{\rho(r')}{|r - r'|} dr' - \int \frac{Z_a}{|r - R_a|} dr + v_{XC}[\rho(r)] \quad (4.23)$$

The first term of the KS Hamiltonian (Eq. 4.23) accounts for the kinetic energy, the second represents the coulomb interaction between the electrons and the third term is the static coulomb repulsion between the nuclei and the electron. Finally, the last term (v_{XC}) represents the exchange correlation potential, which is gathering all the corrections that are not known and need approximation to be estimated. It can be defined as the exchange correlation energy functional derivative to the electron density given by Eq. 4.24.

$$v_{XC}[\rho(r)] = \frac{\delta E_{XC}[\rho(r)]}{\delta \rho(r)} \quad (4.24)$$

The exchange correlation energy functional accounts for the kinetic energy correction of non-interacting electrons and considers non-classical interactions between electrons (Eq. 4.25).

$$E_{XC}[\rho(r)] = (T[\rho(r)] - T_s[\rho(r)]) + (V_{ee}[\rho(r)] - J[\rho(r)]) \quad (4.25)$$

Thus, Eq. 4.17 can be rewritten (Eq. 4.26):

$$E^{HK-KS}[\rho(r)] = T_{ns}[\rho(r)] + V_{ee}[\rho(r)] + V_{ne}[\rho(r)] + E_{XC}[\rho(r)] \quad (4.26)$$

Kohn and Sham have thus devised a method for calculating the energy of a system in its ground state. The exact ground state energy can be calculated if the true expression of the E_{xc} functional was known. Since

the exact exchange-correlation functional expression is not known, different approximations have been developed to estimate it. Consequently, the accuracy of DFT calculations depends on the accuracy of exchange-correlation functional. However, compared to the other energetic terms, it is relatively small, thus, yielding quite realistic results.

Before introducing the exchange-correlation functional to perform DFT calculations on molecules, it is necessary to define the space where the molecular wavefunction extends. This is done by the conversion of the canonical differential equations into algebraic equations introducing a set of variable functions, called atomic basis set.

4.3.3. The atomic basis set

The molecular orbitals are built by a Linear Combination of Atomic Orbitals (LCAO) of mono-electronic wave functions (Eq. 4.27), called basis functions^{125,132} (Atomic Orbitals: AO) and centered on each atom.

$$\phi_i(r) = \sum_{q=1}^K c_{iq} \chi_q(r) \quad (4.27)$$

$\phi(r)$, is the molecular orbital, c_{iq} is the coefficient representing the weight of the atomic orbital $\chi_q(r)$. The set of $\chi_q(r)$ atomic orbitals constitutes the atomic basis set. Decomposed on these set of functions, the electronic Schrödinger equation assumes a matrix representation, and can be solved by linear-algebraic matrix techniques. Two main types of functions are used to describe the AO, either Slater-type orbitals (STO) or Gaussian-type orbitals (GTO).^{125,132}

Slater-Type Orbitals (STO): are expressed by the following function obtained from the exact solutions of the Schrödinger equation for the hydrogen atom and have an exponential form (Eq. 4.28).

$$\chi_q^{STO}(x, y, z) = N_{\alpha q} x^{a_{q,x}} y^{a_{q,y}} z^{a_{q,z}} e^{-\alpha_q |r-r_q|} \quad (4.28)$$

Where x , y , and z are the molecular cartesian coordinates. $N_{\alpha q}$ is the normalization constant, a_q is the orbital angular moment, r_q is the atomic center and α_q is the exponent. STOs have the advantage that they closely mimic the orbital shape of a hydrogen atom thanks to the linear dependence on the distance at the exponent. However, the computation of the two-electron integrals at different centers cannot be performed analytically.

Gaussian-Type Orbitals (GTO): use a linear combination of GTO to reproduce as close as possible the overall form of a given STO. Since they are simple to integrate, they are commonly employed, and their general form is given by Eq. 4.29.

$$\chi_q^{GTO}(x, y, z) = N_{a,q} x^{a_{q,x}} y^{a_{q,y}} z^{a_{q,z}} e^{-\alpha_q |r-r_q|^2} \quad (4.29)$$

Due to the quadratic dependence in the exponential term $|r-r_q|^2$, gaussian functions are less similar to the orbital shape of a hydrogen atom. They do not peak at the nuclear center, and they decay more rapidly. To overcome these limitations, it is typically used to retain some of the Gaussian Primitives (PGTO) in fixed linear combinations, and this procedure is called contraction.

Contracted Gaussian-Type Orbitals (CGTO): is a linear combination of primitive GTO used for building Contracted Gaussian-Type Orbitals (CGTO) according to Eq. 4.30.

$$\chi_q^{GTO}(x, y, z) = \sum_{k=1}^{N_g} d_{q,k} \chi_q^{GTO} \quad (4.30)$$

These primitive functions are combined using contraction coefficients, denoted as $d_{d,k}$, allowing control on the overall shape of the CGTO. Different contraction procedures exist^{125,132}: a general contraction involves combining multiple primitive Gaussian functions (each centered on the same atom or nucleus) into a single contracted orbital. This contraction is achieved by linearly combining the primitive Gaussians with contraction coefficients. By using general contractions, the electronic structure of molecules can be described in a computationally efficient manner retaining a high level of accuracy in representing the wave functions of the system.

The selection of the atomic basis set holds significant importance. A larger basis set provides a more accurate description of the system but comes at the cost of increased computational time. Hence, choosing an appropriate basis set involves addressing the trade-off between accuracy and computational effort. Basis sets can be categorized into:

Minimal basis set represents the smallest set of functions required to describe the electron orbitals for an atom adequately. Each orbital essential for describing the atom's electrons is represented by a single function in this basis set. For instance, in a hydrogen atom, only one orbital is required, constituting the minimal basis set for hydrogen. On the other hand, a carbon atom demands five orbitals (1s, 2s, 3px, 2py, and 2pz) to accurately depict its electron orbitals, even with one of the p orbitals remaining unoccupied. Consequently, the minimal basis set for carbon comprises five functions to represent these orbitals.

The most common minimal basis set is called Slater Type Orbitals approximated by n Gaussian functions (STO- n G). Where n is an integer that represents the contraction length and can vary from 2 to 6. Despite their tendency to yield less accurate results, these sets are widely used due to their low computational cost, making them valuable for quick qualitative insights. The most common is the **STO-3G**, which accounts for three Gaussian primitives per basis function. As we seek to improve accuracy without excessive computational cost, we arrive at the Split-valence basis set.

Split-valence basis set: In this basis set, the core electrons are approximated using a single contracted basis function containing n Gaussian primitives. On the other hand, valence orbitals are described using two (double- ζ), or three (triple- ζ) basis functions, each one characterized by Gaussian primitives denoted as a , b , c . The symbol " ζ " (zeta) signifies the exponent determining the spatial decay and extent of Gaussian or Slater-type orbitals. The categorization of these basis sets follows Pople's nomenclature.¹³³ In practical application, consider the **3-21G** basis set for a carbon atom:

- 1s Orbital: Described by simple- ζ with $n = 3$ Gaussian primitives.
- 2s, 2px, 2py, and 2pz Orbitals: Described by $a = 2$ Gaussian primitives (first STO of the double- ζ), and $b = 1$ Gaussian function in the second STO double- ζ .

The split valence basis set allows orbitals to adapt in size without altering their shape. Advancements have led to the development of polarized basis sets.

Polarized basis sets address the limitation of shape by introducing functions with higher angular momentum than valence electrons. These higher angular momentum functions are denoted by specific letters:

- d** Representing higher angular momentum functions for elements of the second row.
- f** Indicating higher angular momentum functions for heavy atoms.
- p** signifying higher angular momentum functions for hydrogen.

Another critical enhancement emerges to describe electron behavior, particularly when electrons are relatively far from the nucleus.

Diffuse functions allow orbitals to occupy larger regions of space, providing a more accurate representation of electron density distribution. This is especially important for systems involving lone pairs, negatively charged species, excited states, or those with low ionization potential. These functions are symbolized by a "+", denoting the addition of extra functions. The first "+" indicates that extra functions

are added to heavy atoms, and the second “+” adds extra functions to hydrogen atoms. This addition expands the functionality of the basis set, enabling a more comprehensive and precise description of electron behavior.

The notation of a basis set can be represented in a compact form using the following equation (Eq. 4.31).

$$n\text{-}abc\text{+}G(\text{df}, \text{p}) \quad (4.31)$$

This equation provides a representation of the basis set notation following Pople's nomenclature, including the presence of diffuse and polarized functions when applicable. Table 4.2 provides a comprehensive overview of the nomenclature regarding these basis sets, defining its symbols and presenting common values.

Table 4.2 – Nomenclature for basis sets.

Basis set	Value	Description
n	3 or 6	Gaussians primitive in the core orbital
a	2 or 3	Gaussians primitives in the inner valence shell
b	1	Gaussian primitive in the outer valence shell (valence orbital duplicated)
c	0 or 1	Gaussian primitive in the outer valence shell (valence orbital triplicate)
+		Diffuse functions on heavy atoms or
++		Diffuse functions in both, heavy and hydrogen atoms
d	0 to 3d	d-type polarization functions
f	0 to 3f	f-type polarization functions on heavy atoms
p	0 to 3p	p-type polarization function on hydrogen atoms

4.4. The Exchange-Correlation energy

As mentioned previously, the main problem in the DFT is finding the appropriate form of the exchange-correlation energy (EXC) which contains all information about the many-body system of interacting electrons. Currently, we do not know the exact form of this functional, however, many approximations have been proposed. These approximate models can roughly be divided into three classes:

Local Density Approximations (LDA) consider the exchange-correlation energy at each point in space as a functional of the local electron density at that point.

Gradient-Corrected Approximations (GGA) incorporate not only the local electron density but also its gradient.

Hybrid functionals combine a fraction of Hartree-Fock exchange with exchange and correlation terms from either LDA or GGA.

Perdew¹³⁴ proposed a general classification of the different approximations called “Jacob’s ladder of density functional approximation” (see figure 4.3). Where the bottom represents the Hartree world, the top corresponds to the “heaven” of chemical accuracy with the highest level of accuracy. The level of density accuracy is intricately linked to the chosen model approximation, particularly the exchange-correlation (XC) functional which is a key component in DFT. Different XC functionals make varying approximations to the true exchange and correlation effects within the electron system. Each rung of the ladder represents the XC functional. As you climb the "Jacob's ladder of density functional approximation" and use more sophisticated XC functionals, hence their accuracy. The circles in figure 4.3 represent the relative density differences with respect to coupled cluster singles and doubles (CCSD)^{135,136} for the beryllium atom across the four different groups of density functionals. The beryllium atom is a convenient and representative test case to evaluate the accuracy and performance of density functionals due to its well-known properties, sensitivity to different approximations, computational feasibility, and relevance in understanding fundamental principles of electronic structure.

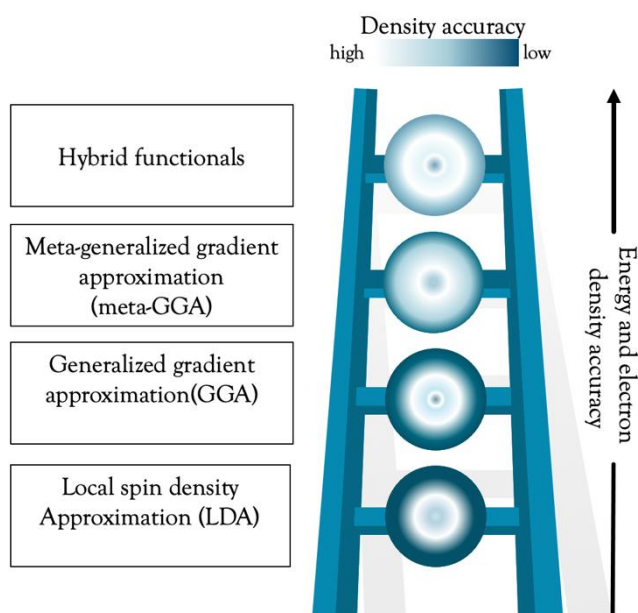


Figure 4.3 – Graphical representation of the Jacob’s ladder of chemical accuracy. For each rung in the ladder an indication of the accuracy is given according to the heat map shown. Adapted from reference¹³⁷.

4.4.1. Local density approximation

Local Density Approximation (LDA)¹³¹ is the simplest approach to approximate XC functional, in which approximates the real electron density to a homogeneous electron gas. In other words, LDA assumes that the electron density varies very slowly with the position r , thus the density can be considered constant anywhere in the space. Furthermore, LDA considers the value of the electron density for a given point to compute the energy density at that point, therefore the functional is local (Eq. 4.32):

$$E_{XC}^{LDA}[\rho(r)] = \int \rho(r) \cdot E_{XC}[\rho(r)] dr \quad (4.32)$$

$E_{XC}[\rho]$ is the exchange-correlation energy per electron that has a constant density ρ . Furthermore, it can be separate (Eq. 4.33) into two contributions: exchange and correlation.

$$E_{XC}[\rho(r)] = E_X[\rho(r)] + E_C[\rho(r)] \quad (4.33)$$

The exchange energy (EX) can be computed analytically by the Dirac equation (Eq. 4.34).

$$E_{XC}[\rho(r)] = -\frac{3}{4} \sqrt{\frac{3\rho(r)}{\pi}} \quad (4.34)$$

On the other hand, the correlation energy (E_C) is not known exactly. Therefore, an approximation expression can be obtained based on a complex function depending on parameters fitted by Monte-Carlo simulations for the energy of the homogeneous electron.

This functionals LDA can describe systems characterized by slowly varying electronic density (i.e., metals) with certain accuracy. On the other hand, inefficient for materials with many highly delocalized electrons.

4.4.2. Generalized gradient approximation

The electron density in atoms, molecules, or metals is typically not homogeneous and varies significantly from one point to another. To overcome the limitation in the use of LDA functionals, a second generation of approximate functionals introduced a dependency not only on the local density value but also on its gradient and called Generalized gradient approximation (GGA)¹³⁸. Usually, these functionals are constructed by adding a correction term to the LDA functional, in the form of Taylor series expansion (Eq. 4.35).

$$E_{XC}^{GGA}[\rho(r)] = E_{XC}^{LDA}[\rho(r)] + \left[\frac{\nabla\rho}{\rho^{\frac{4}{3}}(r)} \right] \quad (4.35)$$

GGA can be improved considering also the second derivative of the electron density (∇^2) known as the kinetic energy density. Those functionals are called meta-GGA functionals¹³⁹ and takes the form given by Eq. 4.36.

$$E_{XC}^{meta-GGA}[\rho(r)] = \int \rho(r) \cdot E_{XC}[\rho(r), \nabla\rho, \nabla^2\rho] dr \quad (4.36)$$

These functionals can be divided into two classes: empirical, and semi-empirical functionals and presented in Table 4.3. Empirical functionals are parameter-free functionals and are known for their general applicability giving accurate results for a wide range of systems, such as **BP86**, **PW91** developed by Perdew and Wang and PBE by Perdew, Burke, and Ernzerhof. Instead, semi-empirical functionals use adjusted parameters by fitting experimental data, such as atomic energies. Some examples are **LYP** by Lee, Yang, and Par, or **BLYP** when including Becke exchange functional with the **LYP**. Where the last one, **BLYP**, containing only one parameter, fits the exact Hartree-Fock exchange energies of a wide variety of atomic systems with remarkable accuracy, surpassing the performance of previous functionals containing two parameters or more.

Table 4.3 – meta-GGA functionals.

Type	Name	Description	Reference
Empirical	BP86	General-purpose, accurate for a wide range of systems	Perdew and Wang
	PW91	General-purpose, provides accurate results	Perdew and Wang
	PBE	General-purpose, widely used and accurate	Perdew, Burke, and Ernzerhof ¹³⁸
Semi-Empirical	LYP	Uses adjusted parameters based on fitting	Lee, Yang, and Par
	BLYP	Adjusted parameters with Becke exchange	Becke and Lee

More accurate functionals are obtained when including fractions of the exact Hartree-Fock Exchange in a GGA functional and are called hybrid functionals.

4.4.3. Hybrid functionals

LDA and GGA are affected by different sources of error. Among the known problems, the self-interaction error comes from the fact that the residual self-interaction in the Coulomb part and in the exchange part do not cancel each other exactly. Calculations performed with these methods tend to over-delocalize the

electronic structure to minimize this unphysical self-interaction. To overcome this drawback and remove the self-interaction error, the strategy is to incorporate a portion of exact exchange from the Hartree-Fock theory with the DFT one. This inclusion may be justified through the so-called adiabatic connection; a fundamental concept that establishes a link between the non-interacting and interacting systems of electrons in a many-body quantum system. The basic idea behind this connection relies on the fact that when the Slater determinant is built with the occupied Kohn-Sham orbitals, rather than the density, can give an exact solution in the case of the non-interacting system. Since the exchange-correlation energy arises from electron-electron interaction, a functional that accounts for this interaction should include a part of the exchange energy.

Drawing this connection between the interacting and non-interacting systems forms the basis for what are known as hybrid density functionals. These two systems can be expressed by the exchange-correlation potential as a function of parameter λ (Eq. 4.37). This results in a polynomial function of degree $n-1$ and depends on λ . This parameter is varying from zero to one and the electron-electron interaction is switched off when λ is equal to zero and switched on when λ is equal to one. The integer n controls the speed with which the correction brought to DFT is canceled when λ tends toward the unit.

$$v_{XC}^{\lambda}[\rho(r)] = E_{XC}^{DFT}[\rho(r)] + (E_X^{HF} + E_X^{DFT}[\rho(r)])(1 - \lambda)^{n-1} \quad (4.37)$$

Integration of the equation 4.37 over the interval $\lambda \in [0,1]$ gives (Eq. 4.38):

$$E_{XC}[\rho(r)] = \int_0^1 v_{XC}^{\lambda}[\rho(r)]d\lambda = E_{XC}^{DFT}[\rho(r)] + \frac{1}{n}(E_X^{HF} + E_X^{DFT}[\rho(r)]) \quad (4.38)$$

One can derive that exchange-correlation energy can be computed by eq. 4.39.

$$E_{XC}[\rho(r)] = \int_0^1 \langle \psi(\lambda) | V_{ee}(\lambda) | \psi(\lambda) \rangle - J[\rho(r)] \quad (4.39)$$

B3LYP is an example of this group of functionals, which contains additional empirical parameters to improve the computation of the total exchange-correlation energy. **B3LYP** incorporates 20% of exact exchange from HF theory (with $a=0.20$), and 72% of GGA exchange functional developed by Becke, E_X^{B88} (with $b=0.72$), and 81% of GGA correlation functional developed by Lee, Yang, and Par, E_C^{LYP} (with $c=0.81$), combining with the LDA exchange-correlation functional according to the Eq. 4.40.

$$E_{XC}^{B3LYP} = E_X^{LDA} + aE_X^{HF} + bE_X^{B88} + (1 - c)E_C^{LDA} + cE_C^{LYP} \quad (4.40)$$

PBE0 proposed by Adamo et al. in 1999 is another hybrid functional widely known. This functional is constructed using the parameter-free PBE, combining 25% of the HF exchange. PBE0 functional (Eq. 4.41) is considered a parameter-free global hybrid functional, since the Hartree-Fock exchange was proposed by Perdew on purely theoretical considerations and the PBE does not contain empirical parameters.

$$E_{XC}^{\text{PBE0}} = E_X^{\text{PBE}} + E_C^{\text{PBE}} + \frac{1}{4}(E_X^{\text{HF}} - E_X^{\text{PBE}}) \quad (4.41)$$

4.4.4. Range-separated hybrid functionals

Range-separated hybrid (RSH) functionals are a subgroup of hybrid functionals in which the ratio of HF exchange varies with the inter-electronic distance. These functionals have been developed since at short distances the dominant interaction is the dynamic correlation, whereas at large distances, the non-Coulombic interaction is dominant (well described with HF approach). Therefore, the ratio of HF exchange should not be considered constant. Thus, in the RSH functional the Coulomb operator split into short- and long-range components by Eq. 4.42.

$$\frac{1}{r_{12}} = \frac{1 - [\alpha + \beta \text{erf}(\omega r_{12})]}{r_{12}} + \frac{\alpha + \beta \text{erf}(\omega r_{12})}{r_{12}} \quad (4.42)$$

r_{12} is the inter-electronic distance, α and β are parameters determined empirically. ω is a constant that determines how the Coulomb operator is separated into short-range and long-range components as a function of the inter-electronic distance (r_{12}).

As in the case of global hybrids, the usual procedure is fitting a standard benchmark set of molecules to experimental thermochemical data and the obtained values are then used for all systems. The first term accounts for the short-range interaction, while the second the long-range one. The error functions (erf) link the short- and the long-range of the Coulomb operator under an attenuation factor, ω .

CAM-B3LYP is an example of a range-separated hybrid functional. This functional combines the Coulomb-attenuating method (CAM) and the global hybrid B3LYP. Whereas at long-range, it considers 65% of exact HF exchange and at short-range, it considers 19% of inter-electronic interactions. This functional is thus qualified for long-range corrected functional. The parameters of the Coulomb operator (Eq. 4.42), α , β and ω are set to 0.19, 0.46 and 0.33 respectively.

4.5. Modeling solvent effect

In computational chemistry, modeling the solvent effect is crucial for understanding how a solvent influences the electronic structure (at the molecular level) and overall properties of a system (at the macroscopic level). This is important in studying chemical reactions, molecular interactions, and various processes that occur in a solvent environment.

To describe the interaction between solvent and solute in the context of quantum chemistry, two main approaches are used to account for solvent effect: explicit solvation models and implicit (or continuum) models.

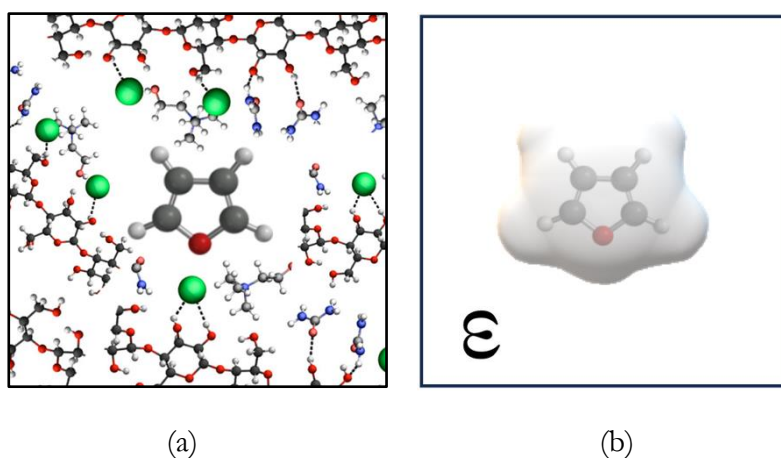


Fig. 4.4 – Representations of (a) Explicit and (b) Implicit solvation models.

4.5.1. Explicit solvation models

Explicit models involve representing the solvent molecules explicitly in the computational model. Solvent molecules are typically represented by a set of atoms or molecules, within a simulation box, interacting directly with the solute, as illustrated in figure 4.4a. In this approach, the solute interacts directly with the explicitly included solvent molecules, allowing for a detailed description of specific solute-solvent interactions, such as hydrogen bonds. As the position of solvent molecules are explicitly defined, this approach provides a more realistic representation of the polarity around the solute.

4.5.2. Implicit Solvation models

Implicit models describe the solvent molecules as a continuum medium rather than representing them individually, as illustrated in figure 4.4b. Unlike explicit models where solvent molecules are explicitly considered, implicit models approximate the solvent's effect on the solute through a continuum description. This continuum is often represented by a dielectric constant (ϵ), which allows the evaluation

of how much the solvent is polarized due to the solute presence. In turn, the electronic cloud of the solute is affected by the solvent's polarization, enabling an understanding of their mutual influence without representing each solvent molecule explicitly, as ϵ implicitly represents this effect. The dielectric constant (ϵ) is a dimensionless quantity equal to the electric permittivity relative to the vacuum and ranging from around 2 for nonpolar solvents like benzene and hexane to higher values like 78 for water and even up to 110 for formamide.

In implicit solvation models, the solvation energy ΔG_{sol} of a solute in a fixed conformation is usually separated between two contributions: ΔG_{el} (the electrostatic contribution) and ΔG_{ne} (the non-electrostatic contribution) (Eq. 4.43).

$$\Delta G_{\text{sol}} = \Delta G_{\text{el}} + \Delta G_{\text{ne}} \quad (4.43)$$

The electrostatic contribution to solvation energy is a result of the work needed to polarize the solute-solvent system. At both electronic and nuclear levels, the solute's charge density induces a polarization effect in the surrounding solvent. This polarization, known as the Reaction-Field (RF), creates an electric field in the solvent. RF is responsible for the initial polarization of the solute charge density, reflecting short-range and long-range electrostatic responses of the solvent to the solute. Furthermore, this initial polarization initiates a self-consistent process of mutual polarization between the solute and the solvent, known as the Self-Consistent Reaction Field (SCRf).

On the other hand, the non-electrostatic contribution to the solvation energy in implicit solvent models includes a diverse range of physical phenomena beyond charge interactions. It consists of several components that account for different types of interactions and effects between the solute and the solvent. These components are as follows:

- ΔG_{rep} - Short-range Pauli repulsion: accounts for the energy associated with the repulsive interactions between electrons in the solute and the surrounding solvent due to the Pauli exclusion principle.
- ΔG_{dis} - Dispersion interactions: arise from van der Waals forces and account for the attractive forces between a solute and surrounding solvent molecules.
- ΔG_{hb} - Hydrogen-bonding: specific interactions between a solute and solvent molecule involving the sharing of a hydrogen atom.
- ΔG_{cav} - Cavitation energy: accounts for the energy required to form a cavity when a solute is introduced into the solvent.

Each of these components contributes uniquely to the overall non-electrostatic solvation energy, as schematically shown in Fig. 4.2 and expressed by the Eq. 4.44.

$$\Delta G_{ne} = \Delta G_{rep} + \Delta G_{dis} + \Delta G_{hb} + \Delta G_{cav} \quad (4.44)$$

In the figure 4.5, W_{el} stands for the electrostatic work needed to polarize the dielectric media. ΔG_{cav} has been considered separate from ΔG_{ne} (gray) to point out the cavity (yellow) formation within the dielectric media (blue).

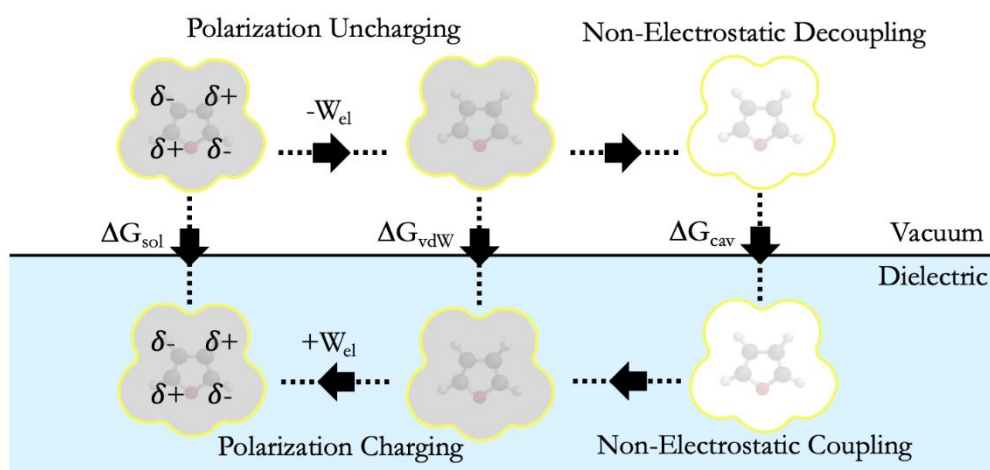


Figure 4.5 – Thermodynamic cycle depicting the solvation free energy decomposition into contributions.

4.5.2.1. The solution of the electrostatic problem

The main assumption of most implicit solvation models is that the electric behavior of a medium can be simplified and described by a macroscopic polarization defined by a dipole density, denoted as $P(\mathbf{r})$. In the presence of a dielectric medium, the conventional electric field $E(\mathbf{r})$ in vacuum is replaced by the electric displacement field, denoted as $D(\mathbf{r})$. The electric displacement field accounts for the influence of the dielectric medium, altering how electric forces are distributed and experienced within the medium by the Eq. 4.45.

$$D(\mathbf{r}) = \varepsilon(\mathbf{r})E(\mathbf{r}) = E(\mathbf{r}) + 4\pi P(\mathbf{r}) \quad (4.45)$$

In Eq. 4.45, the permittivity $\varepsilon(\mathbf{r})$ is presented as a scalar-valued function rather than a simple dielectric constant. Given a charge density $\rho(\mathbf{r})$, including both nuclei and electrons for the solute, Maxwell's equation for the displacement field $D(\mathbf{r})$ is deduced accordingly Eq. 4.46.

$$\nabla \cdot D(\mathbf{r}) = 4\pi\rho(\mathbf{r}) \quad (4.46)$$

This equation can be reformulated in the familiar form of Poisson's by considering the relation between the electric field and the negative gradient of the potential ($E(\mathbf{r}) = -\nabla\phi(\mathbf{r})$) as Eq. 4.47.

$$\nabla[\epsilon(\mathbf{r}) \cdot \nabla\phi(\mathbf{r})] = -4\pi\rho(\mathbf{r}) \quad (4.47)$$

When explicit solvent molecules are replaced with a dielectric medium, the clear boundary between the solute and the solvent, often resulting from Pauli repulsion, disappears. To address this, an arbitrarily defined surface denoted as Γ is introduced. This surface acts as the interface between the region where atoms of the solute are explicitly considered, denoted as Ω_{inner} , and the region where the solvent is substituted with the dielectric medium, denoted as Ω_{outer} , as illustrated in figure 4.6. The introduction of this surface requires defining $\epsilon(\mathbf{r})$ as a step function. Specifically, it takes on the values ϵ_{inner} and ϵ_{outer} when \mathbf{r} is situated within the inner Ω_{inner} or outer Ω_{outer} region, according to relationship in Eq. 4.48.

$$\epsilon(\mathbf{r}) = \begin{cases} \epsilon_{\text{inner}} & \text{if } \mathbf{r} \text{ is } \in \Omega_{\text{inner}} \\ \epsilon_{\text{outer}} & \text{if } \mathbf{r} \text{ is } \in \Omega_{\text{outer}} \end{cases} \quad (4.48)$$

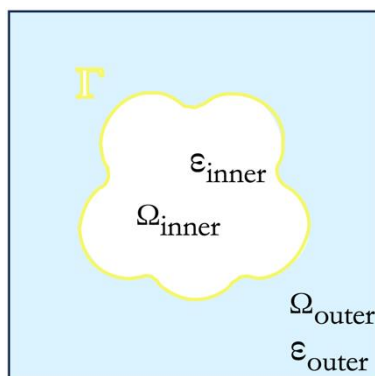


Figure 4.6 – Solute Cavity. Blue region: dielectric medium (ϵ_{outer}). White region: solute region ($\epsilon_{\text{inner}}=1$). Yellow: solute-solvent interface (Γ).

In QM calculations, a typical practice is to set the relative permittivity ϵ_{inner} to 1 for the inner region Ω_{inner} (representing vacuum), while for the outer region Ω_{outer} (representing the solvent), ϵ_{outer} is assigned the relative permittivity of the solvent.

Poisson's equation is foundational in continuum electrostatics, relating the electric potential $\phi(\mathbf{r})$ to charge density $\rho(\mathbf{r})$. In this context, starting with $\rho(\mathbf{r})$ from electronic structure calculations, Poisson's equation is solved to find $\phi(\mathbf{r})$ across space. The potential is then separated into two parts (Eq. 4.49):

$$\phi(\mathbf{r}) = \phi_{\rho}(\mathbf{r}) + \phi_{RF}(\mathbf{r}) \quad (4.49)$$

Where $\phi_p(r)$ is the potential of the charge density of the solute $\rho(r)$, originally responsible for the polarization of the dielectric media and the appearance of an additional charge term $\rho_{pol}(r)$ acting as source term of the reaction-field potential $\phi_{RF}(r)$. Once $\rho(r)$ is derived from Schrödinger's equation and $\phi(r)$ is obtained by solving Eq. 4.47, the electrostatic solvation energy can be expressed as Eq. 4.50.

$$E_{RF} = \frac{1}{2} \int \phi_{RF}(r) \rho(r) dr = \frac{1}{2} \int \phi_p(r) \rho_{pol}(r) dr \quad (4.50)$$

The presence of the factor 1/2 in Eq. 4.50 signifies that the interaction energy is decreased by exactly half its initial value. This adjustment considers the energy needed to polarize the surrounding environment.

The reaction-field potential $\phi_{RF}(r)$ is a fundamental component shared by all Apparent Surface Charge (ASC) methods. ASC methods effectively utilize this potential to model the polarization effects induced by charges originating from the solute within a solvent by Eq. 4.51.

$$\phi_{RF}(r) = \int_{s \in \Gamma} \frac{\sigma(s)}{|s-r|} ds \quad (4.51)$$

This equation (Eq. 4.51) describes the potential at a point r inside the volume, as a result of the integration overall points s on the cavity surface. To calculate the potential at a specific point within the volume $\phi_{RF}(r)$ it is only necessary to discretize the cavity surface Γ rather than the whole of three-dimensional space and should be consider the statement (Eq. 4.52).

$$\begin{cases} \phi_{inner}(r) = \phi_{outer}(r) \\ \epsilon_{inner} \mathbf{n} \cdot \nabla \phi(r)|_{inner} = \epsilon_{outer} \mathbf{n} \cdot \nabla \phi(r)|_{outer} \end{cases} \quad (4.52)$$

4.5.2.2. Cavity definition

The solute cavities can be roughly categorized into empirical cavities, which are based on empirical van der Waals radii or the molar volume of the solute, and non-empirical cavities, which rely on an isodensity surface derived from the electronic density of the solute obtained through quantum mechanical (QM) calculations. In the context of initial electrostatic models developed by Born, Onsager, Kirkwood, and others, a common and straightforward solute cavity type is a simple spherical cavity surrounding the solute charge density. The radius for such cavity was suggested to be (Eq. 4.53):

$$R = \left(\frac{3V_m}{4\pi N_a} \right)^{1/3} \quad (4.53)$$

where V_m is the molar volume of the solute and N_a is the Avogadro's constant.

Figure 4.7 is an illustrative representation of various approaches to construct cavity surfaces. The foundational method considers a union of atom-centered spheres, commonly denoted as a van der Waals (vdW) cavity surface. These spheres are defined using empirically obtained radii. The most popular is the one defined by Bondi.¹⁴⁰ The vdW-based solute cavity serves as the initial step in defining another surface, which shares a similar premise: simplifying the solvent molecule into a sphere with a volume equivalent to the van der Waals volume of the solvent by using a scaling factor. These constructed surfaces are known as Solvent-Accessible Surfaces (SAS). Both the vdW surface and the SAS exhibit characteristic cusps at the intersections of atomic spheres. These cusps are effectively eliminated in the solvent-excluded surface (SES), which is generated as a probe sphere rolls over the vdW surface, as illustrated in Figure 4.7.

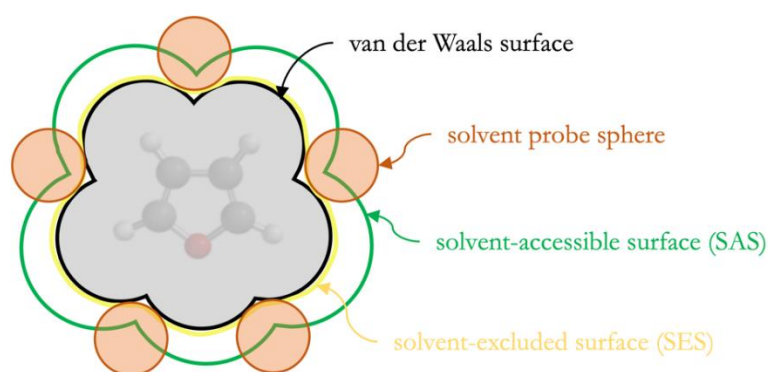


Figure 4.7 – 2D example of a SAS in green and a SES solute cavity in yellow, with Solvent Probe in orange, and vdW surfaces and sphere in black and gray, respectively.

The Conductor-like Screening Model (COSMO) or C-PCM (C for conductor) are ASC method. In COSMO model, a scaling factor of $f = 1.17$ is employed and commonly used to obtain the Klamt radii. Subsequently, factors around $f \approx 1.1 - 1.2$ have become widely accepted, as the one from UFF universal forcefield, presented in Table 4.4.

Table 4.4 – Atomic radius in angstrom.

Element	Bondi	UFF	Klamt
C	1.70	1.925	2.000
C-(H)	1.20	1.443	1.300
O	1.52	1.750	1.720
N	1.55	1.729	1.830
P	1.80	1.487	2.106
F	1.47	1.552	1.720
Cl	1.75	1.697	2.050
Br	1.85	1.905	2.160

4.5.2.3. The tessellation of the cavity surface

Once the set of spheres that define the chosen surface have been selected, the cavity surface is divided into a grid of surface elements. This grid construction employs methodologies, such as the Geometria Polihedro (GEPOL) algorithm¹⁴¹. The goal of GEPOL is to discretize the molecular cavity, transforming it into a polyhedral shape that can be more easily handled for computational calculations. GEPOL algorithm approximates a sphere to a polyhedron with 60 triangular faces, such as an icosahedron. This polyhedron is circumscribed around the molecule, providing a discrete representation of the cavity. The surface of this polyhedron is divided into smaller triangular surface elements, referred to as *tesserae*, utilizing a tessellation process. Each one of these triangles approximates a portion of the molecular surface. The areas of these triangular surface elements are calculated based on the surface of the polyhedron and the number of faces it has. These areas are then used for subsequent calculations and analyses.

4.5.3. The polarizable continuum model

The ASC has become nearly synonymous with the term Polarizable Continuum Model (PCM) in contemporary computational chemistry. The PCM represents the most frequently employed approach for modeling the solute-solvent interaction, and encompasses such variants as the COSMO, also known as C-PCM (C for conductor), as well as the integral equation formalism (IEF-PCM). The solute molecules are placed into a cavity with a continuous dielectric medium representing the solvent. The shape of the cavity must reproduce as well as possible the solvent molecular shape and it is constructed by placing interlocked spheres centered on each heavy (non-hydrogen) atom of the molecule. These spheres define a solvent-accessible surface and a solvent-excluded surface. The cavity surface is split into small elements called *tesserae*, at the center of which there are point charges representing the electrostatic polarization of the solute electron density.

4.5.3.1. COSMO

The COSMO or C-PCM (C for conductor) are ASC method that stand in sharp contrast to other implicit solvation models as they replace the dielectric media with a conductor. The advantage of this approach is that conductors are characterized by an infinite relative permittivity $\epsilon = +\infty$ and thus, no potential exists within the media simplifying the electrostatic problem due to the lack of boundary conditions.

The surface charge density over the cavity surface of the conductor $\sigma^\infty(s)$ is determined directly from the electrostatic potential generated by the solute charge density and is later scaled down by a function of the real relative permittivity of the solvent $f(\epsilon)$ to obtain the proper charge density $\sigma(s)$ (Eq. 4.54).

$$\sigma(\mathbf{s}) = f(\epsilon)\sigma^\infty(\mathbf{s}) \quad (4.54)$$

The scaling function $f(\epsilon)$ is empirically determined as Eq. 4.55.

$$f(\epsilon) = \frac{\epsilon - 1}{\epsilon + x} \quad (4.55)$$

Where ϵ is the macroscopic dielectric constant and the empirical scaling x is equal to 0.5 in the original publication of COSMO.

4.3.5. Calculation using COSMO model

The main task of the model is solving the non-relativistic time-independent Schrödinger equation through an exact procedure that bypasses the calculation of the n -electron wavefunction by using the electron density. One first calculation is done to optimize the molecule geometry in a vacuum (Fig 4.9a). Then, the molecule optimized is embedded in a molecule-shaped cavity surrounded by a virtual conductor ($\epsilon = \infty$) (Fig 4.9b). The cavity is constructed based on the element-specific atomic radii, typically in the 1.2 times Bondi radii range. In such an environment, the surface of the molecular cavity is discretized in finite surface segments with an area (called *tesserae*). The molecule induced a polarization charge density on its cavity surface segments (Fig 4.9c). These finite segments are small enough to consider the polarization charge density almost constant inside each segment. As a result, one obtained the three-dimensional geometric distribution of the polarization charges on the surface, also called screening charge density, $\sigma(\mathbf{s})$ (Fig 4.9d).

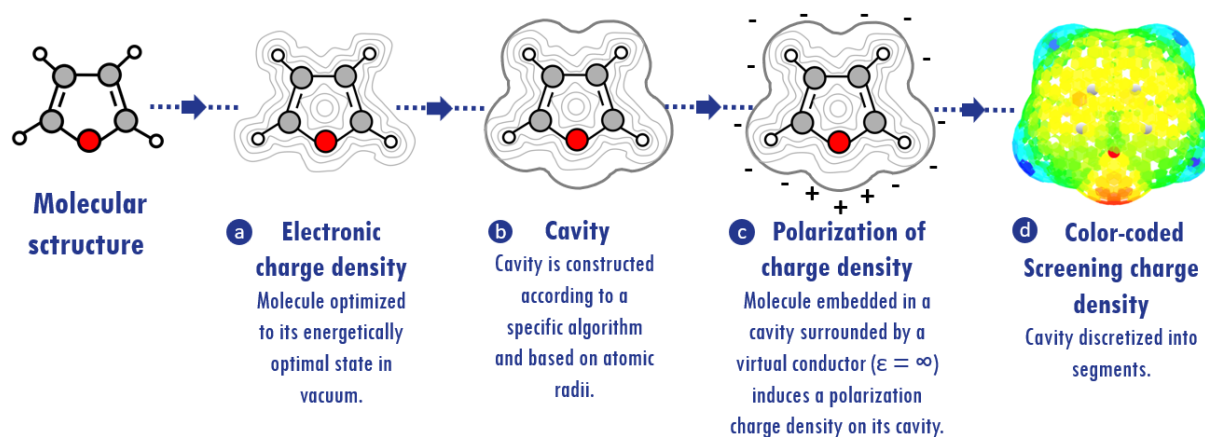


Figure 4.9 – Steps of the DFT calculation

4.6. Conclusion

DFT emerged as a central theoretical framework, offering a pathway to efficiently predict molecular properties. We elucidated the Hohenberg-Kohn Theorems and the Kohn-Sham method, providing essential insights into the theoretical underpinnings critical for accurate electronic structure calculations. The discussion extended to the atomic basis set, a fundamental component in electronic structure modeling. Addressing approximations of the exchange-correlation energy, we explored various levels of approximation, including the local density approximation, generalized gradient approximation, hybrid functionals, and range-separated hybrid functionals.

Understanding the influence of solvation on molecular systems is important. Among of the modeling approaches presented, the PCM stood out as a powerful tool to account for solvation effects in a computationally efficient manner.

As for all the PCM models, the cavity plays a significant role in the COSMO approach. It is built starting from interlocking atomic spheres, whose radii *de facto* represent a computational parameter. Indeed, cavity size modulates the strengths of the solute/solvent interactions: large cavity corresponding to weaker reaction fields and vice-versa. Therefore, it is not surprising that several recipes for their evaluation have been proposed to improve PCM performances. Typically, these radii are determined upon error minimization of some thermodynamics quantities, such as free energy of solvation (ΔG_{solv}). In the original COSMO-RS model, the optimization of the model parameters, including atomic radii, was done considering a set of 217 molecules and six thermodynamic properties, including free energy of hydration,

ΔG_{hydro} , vapor pressure, and some partition coefficients. IDACs were, therefore, not part of the training set. However, the COSMO-SAC variant optimized most of the model parameters and retains the same optimized radii of the first COSMO-RS parameterization.

It should be noticed that UFF radii are significantly larger than those used by Klamt or even traditional vdW values. These radii are then scaled by a factor of 1.1. Instead, the standard COSMO procedure reported in Gaussian uses the Klamt radii 17% (± 2) larger than the Bondi radii, another standard widely used in chemical engineering). These radii, Klamt or UFF, are then used to compute a van der Waals cavity as the sum of the interlocked spheres, using the GEPOL algorithm¹⁴¹, as in standard COSMO-based approaches.

Chapter 5

Experimental results

Résultats expérimentaux

Résumé

Dans ce chapitre, l'accent est principalement mis sur la présentation des résultats expérimentaux relatifs aux DES, avec une attention particulière portée à deux types de DES à base de chlorure de choline (ChCl) l'Ethaline (composé de ChCl et d'éthylène glycol) et le Glycaline (composé de ChCl et de glycérol). Le chapitre englobe une exploration des propriétés essentielles telles que la masse volumique, la viscosité, la tension de vapeur.

De plus, cette étude explore comment les DES interagissent avec des alcools en examinant leur comportement à l'équilibre entre la phase liquide et la phase vapeur. Cette étude repose sur une méthode synthétique (détaillé dans le chapitre 2) comprenant la mesure de la pression d'équilibre d'un mélange multiphasique avec une composition globale connue à température constante. Cette méthode facilite la détermination des compositions de phases grâce à l'application des principes de bilan de matière et à l'utilisation de modèles thermodynamiques tels que NRTL et Wilson. Le mélange d'Ethaline et d'isopropanol présente une forte concordance entre les données expérimentales et les prévisions du modèle à 303 et 313 K, le modèle Wilson montre une meilleure cohérence avec l'augmentation de la température. En revanche, dans le cas du mélange de Glycaline et d'éthanol, les données se situent constamment en dessous des prévisions de tous les modèles, notamment à 323 et 333 K. De plus, les données obtenues pour le mélange de Glycaline et d'isopropanol soulignent l'importance d'une modélisation précise, le modèle Wilson montre une performance remarquable à des températures élevées.

Le chapitre conclut que le modèle Wilson démontre une concordance supérieure, en particulier avec l'augmentation de la température, ce qui signifie son efficacité dans la modélisation du comportement des DES. Cependant, une observation cruciale est faite concernant les mesures de pression. Les faibles valeurs de pression relevées au cours des expériences introduisent des sources potentielles d'incertitude dans les résultats. Pour remédier à ce problème, la recommandation est formulée d'utiliser des capteurs de pression spécifique étalonnés pour les mesures de basse pression, améliorant ainsi la précision et la résolution des données de pression. Cette suggestion revêt une importance particulière lorsqu'il s'agit de systèmes à basse pression.

De plus, le chapitre examine la compatibilité du modèle COSMO-SAC-dps (tel que proposé par Hsieh en 2014) avec les DES. Il est constaté que le modèle n'est pas adéquat pour décrire ces solvants puisqu'il prédit des coefficients d'activité inférieurs à un pour le mélange de Glycaline et d'isopropanol, impliquant

une déviation par rapport à la loi de Raoult. Néanmoins, les résultats expérimentaux contredisent ces prédictions, révélant une disparité entre les valeurs modélisées et les valeur expérimentale.

Ce chapitre offre des perspectives précieuses sur le comportement des DES à base de ChCl et souligne l'importance cruciale de la précision dans les techniques de modélisation et de mesure pour leur application efficace dans les processus de bioraffinerie. De plus, le chapitre explore l'utilisation de la méthode de « gas stripping » pour prédire le IDAC de l'Ethaline. Bien que la méthode de « gas stripping » soit une technique largement utilisée offrant des informations pratiques pour estimer le IDAC et comprendre les interactions soluté-solvant, il est noté que cette méthode pourrait ne pas être idéale pour étudier le IDAC dans des fluides à haute viscosité. Des méthodologies alternatives ou modifiées pourraient être nécessaires pour garantir des résultats fiables et significatifs dans de tels contextes.

Experimental results

This chapter is dedicated to presenting the experimental results for the properties of DES, both in the pure form and when mixed with alcohols. We have investigated two Choline Chloride-based (ChCl-based) DES. The ChCl and Ethylene Glycol (Ethaline) and ChCl and Glycerol (Glycaline). The properties examined include density, viscosity, vapor pressure, and VLE. These properties are fundamental in assessing the behavior and suitability of DES as solvents in biorefinery processes. Understanding how DES behave under different conditions is essential for making informed decisions regarding their potential as eco-friendly and efficient alternatives in biorefinery processes. The experimental measurements and subsequent analysis of density, viscosity, and vapor pressure are imperative for identifying optimal DES candidates for specific applications within the biorefinery domain. The VLE isotherms, solubility parameters and activity coefficients obtained in this study are foundational elements that will play a crucial role in shaping our solvent selection methodology. This methodology will be introduced and comprehensively discussed in the final chapter.

5.1. Pure properties

The pure properties of DES are fundamental due to their direct impact on a wide array of applications and processes, as elaborated in Chapter 1. A thorough understanding and characterization of these properties offer crucial insights into the behavior and applicability of DES. Specifically, the density, viscosity, and vapor pressure of Ethaline and Glycaline will be presented in the subsequent section. We will illustrate how this data enables us to predict the solubility parameter, a key factor in determining solvation power and the compatibility of the DES in various applications. This prediction will significantly contribute to the optimization of the COSMO-SAC model.

5.1.1. Density and Viscosity

In Table 5.1, data for distinct batches of DES is presented, each formulated with different molar ratio of ChCl in combination with either Ethylene Glycol (EG) or Glycerol (Gly). The molar ratio represents the ratio of moles of ChCl (a hydrogen-bond acceptor, HBA) to moles of the respective co-solvent (EG or Gly) acting as the hydrogen bond donor (HBD).

Table 5.1 – DES properties.

	HBA	HBD	Batch	Molar ratio	Molar mass	$u(x_i)$
Ethaline	ChCl	Ethylene Glycol	#1	1: 2.04	87.5794	0.0839
			#2	1: 2.03	87.6807	0.0056
			#3	1: 2.00	87.9162	0.0023
			#4	1: 2.28	85.6786	0.0112
			#5	1: 2.01	87.8279	0.0077
			#6	1: 2.00	87.9302	0.0002
			#7	1: 2.00	87.9089	0.0014
Glycaline	ChCl	Glycerol	#1	1: 2.00	107.9257	0.0041
			#2	1: 2.00	107.9248	0.0041

Figure 5.1 illustrates the experimentally measured densities of ethaline (left) and glycaline (right) in function of the temperature. As expected, an evident reduction in density is noted for both solvents with increasing temperature. This inverse relationship between temperature and density underscores a fundamental principle: higher temperatures result in lower densities for ethaline and glycaline.

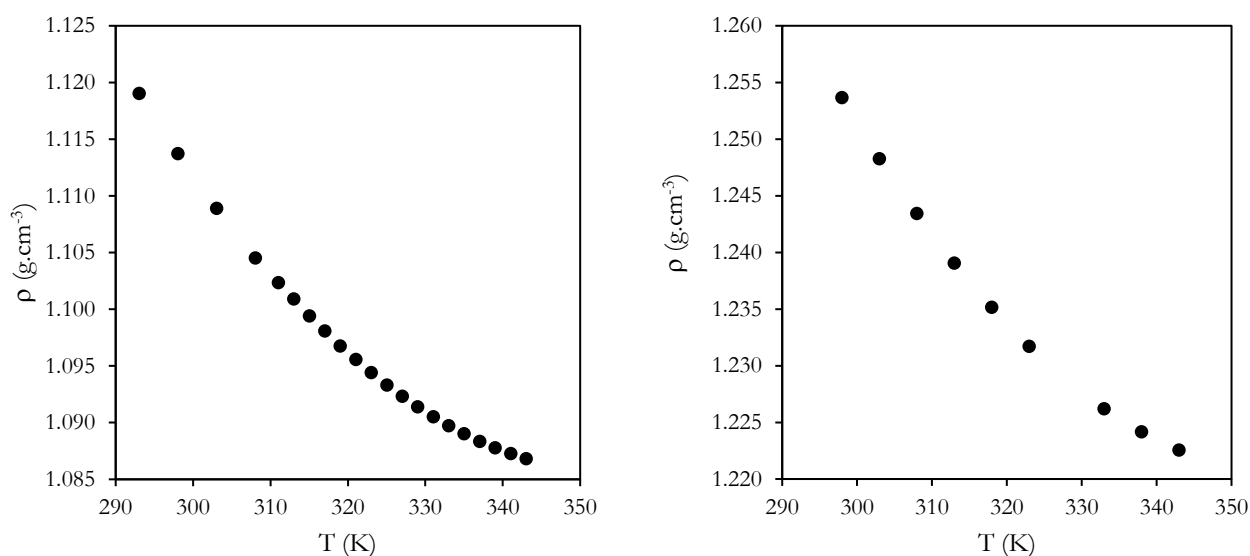


Figure 5.1 – Density in function of the temperature of ChCl:EG (left) and ChCl:Gly (right).

Conversely, in examining the viscosity trend for both ethaline and glycaline as a function of temperature, there is a noticeable decrease in viscosity for both solvents, as illustrated in Fig. 5.2. This outcome aligns with the general behavior of liquids, where higher temperatures lead to reduced viscosity.

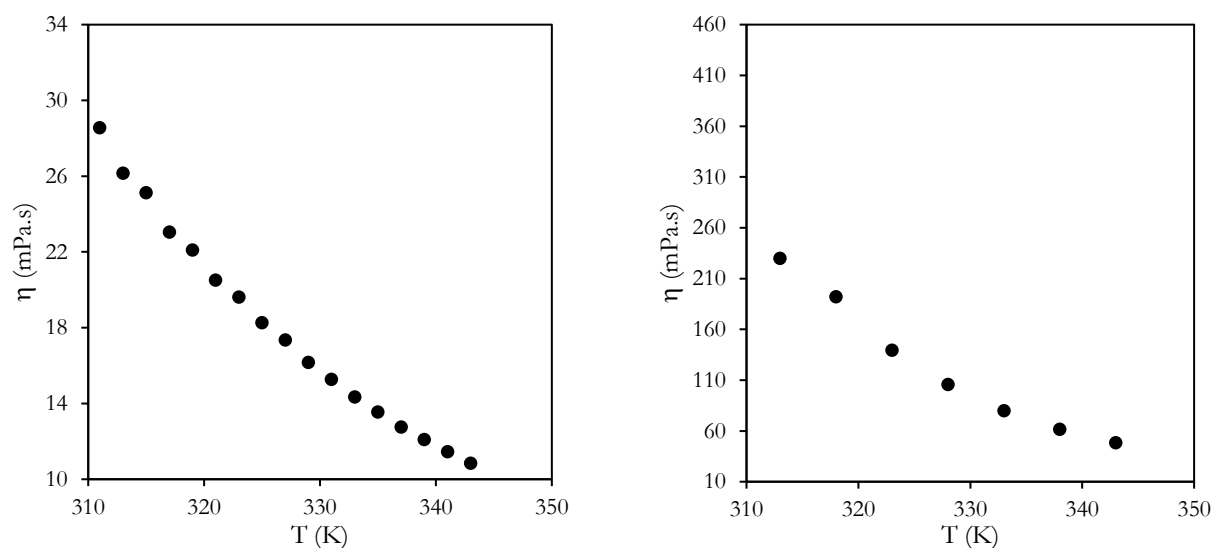


Figure 5.2 – Viscosity in function of the temperature of ChCl:EG (left) and ChCl:Gly (right).

5.1.2. Vapor pressure

The vapor pressure data was obtained based on the experimental measurements using the variable volume cell technique, as detailed in Chapter 2. This well-established and reliable technique involves a controlled, closed system with a known volume, where pressure and temperature are monitored. The method involved loading the equilibrium cell with the viscous components and ensuring that the liquid and vapor phases of the pure viscous component achieve equilibrium prior to data acquisition. The equilibrium cell was carefully assembled within the experimental apparatus. To facilitate accurate measurements and prevent any potential loss of components, the cell was degassed for a few minutes. Upon achieving equilibrium, the temperature and vapor pressure of the pure DES were meticulously measured using a data acquisition unit. The experimental data, representing the equilibrium state, was recorded, and presented in table 5.3 and 5.4. These data were further subjected to regression analysis. This regression analysis allowed for the formulation of the Clausius-Clapeyron equation (Eq. 5.1), a fundamental relation that links vapor pressure to temperature and the enthalpy of vaporization. The equation is expressed in a linear form as:

$$\ln P = -\frac{\Delta H_{vap}}{RT} + b \quad (5.1)$$

Where ΔH_{vap} is the enthalpy of vaporization and represents the energy required to transform a substance from the liquid phase to the vapor phase at a given temperature; P is the vapor pressure; R is the universal

gas constant (8.314J/mol K); and T is the temperature in Kelvin. In this linear form, the slope of the line in a plot of $\ln P$ versus $1/T$ is $-\frac{\Delta H_{vap}}{R}$.

For ethaline and glycaline, the estimated ΔH_{vap} values are 42.53 kJ/mol and 133.17 kJ/mol, respectively.

Clausius-Clapeyron correlation can be derived from the regression analysis enhanced the dataset, enabling the prediction of vapor pressure at temperatures beyond those experimentally measured. This expansion of the dataset through correlation gives an accuracy of around 15% and the parameters are presented in table 5.4.

Table 5.2 – Vapor pressure Ethaline.

Experiments			Regressed	
Batch	T (K)	P (bar)	T (K)	P (bar)
#1	303.25	3.48	303	3.70
#2	303.26	4.27	304	3.91
#3	303.23	2.58	306	4.37
#1	313.26	6.85	309	5.13
#2	313.27	10.60	312	6.02
#3	313.26	4.18	315	7.04
#1	323.27	24.78	318	8.20
#2	323.20	19.30	321	9.53
#3	323.27	9.49	324	11.05
#1	333.31	11.73	327	12.77
#2	333.30	19.54	330	14.72
#3	333.29	8.70	333	16.93

$u(T)=0.02$ K, $u(P)=0.1$ mbar, accuracy = 15%

Table 5.3 -Vapor pressure Glycaline.

Experiments			Regressed	
Batch	T (K)	P (bar)	T (K)	P (bar)
#1	303.23	0.30	303	0.34
#2	303.24	0.22	307	0.68
#1	313.28	3.05	311	1.32
#2	313.23	4.22	315	2.55
#1	323.30	6.99	319	4.82
#2	323.27	6.53	323	8.97

$u(T)=0.02$ K, $u(P)=0.1$ mbar, accuracy = 15%

Figure 5.3 presents the vapor pressure of Ethaline as a function of temperature. The black symbols represent the experimental data points, measured at various temperatures. The green symbols signify the

regressed data obtained through a regression analysis, allowing for predictions of vapor pressure beyond the experimentally measured. The area inside of the dash lines outlines the error range associated with the data points.

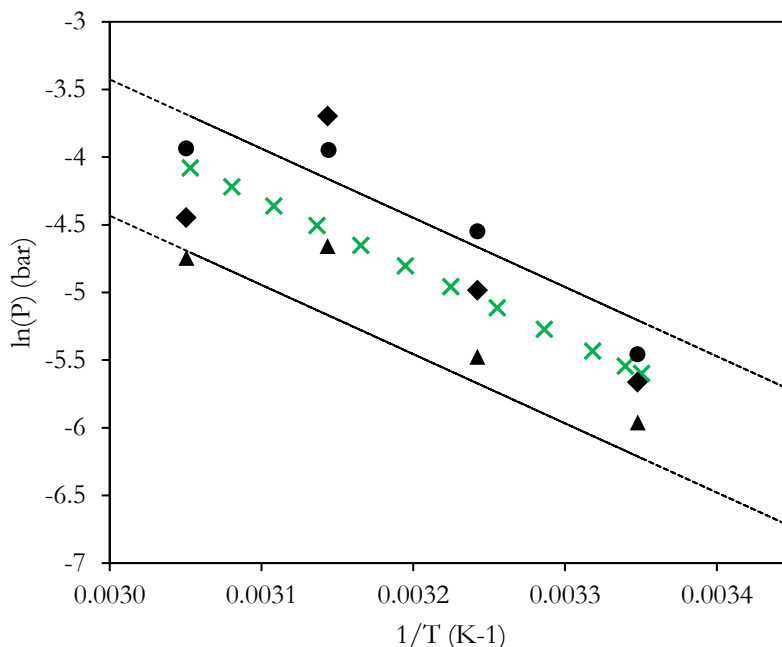


Figure 5.3 – Vapor pressure of Ethaline. Black symbols: experiments, green symbols: regressed, dash lines area: error range.

Figure 5.4 presents the vapor pressure of glycaline as function of the temperature. The black symbols effectively capture the experimental data points, measured across a range of temperatures. In addition, the green symbols represent regressed data, extrapolated through the regression analysis. The area enclosed by the dash lines signifies the error range associated with the data points, further emphasizing the precision and reliability of both experimental and regressed results.

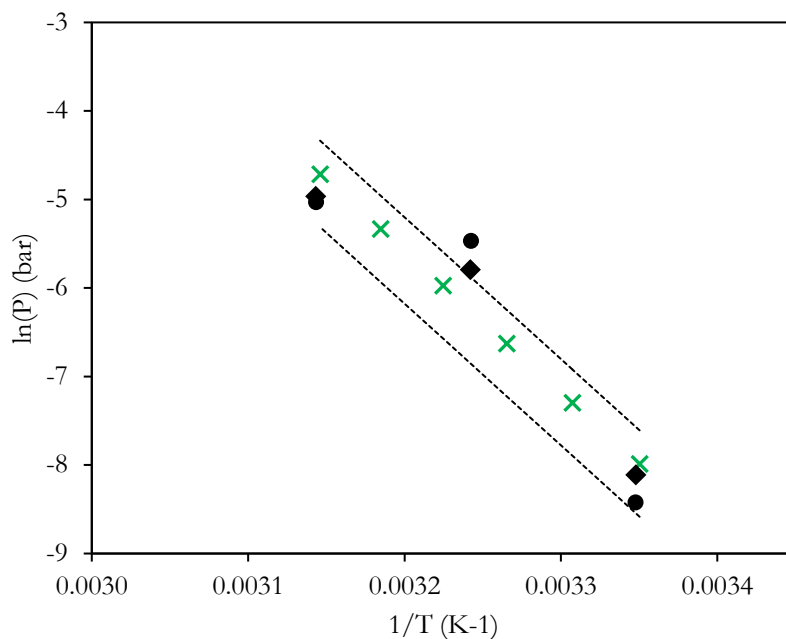


Figure 5.4 – Vapor pressure of Glycaline. Black symbols: experiments, green symbols: regressed, dash line area: error range.

The Hildebrand solubility parameter¹¹⁶ (δ) is a measure of the cohesive energy density of a substance, providing insights into material compatibility and miscibility. It can be estimated using the ΔH_{vap} obtained through the Clausius-Clapeyron equation. The Hildebrand solubility parameter is defined as the square root of the cohesive energy density and is related to ΔH_{vap} by the equation:

$$\delta = (E_{\text{coh}})^{1/2} = \left(\frac{\Delta H_{\text{vap}} - RT}{V_m} \right)^{1/2} \quad (5.2)$$

where V_m the molar volume of the solvent, which can be calculated using the molar mass and the density.

In table 5.4, we present the obtained Hildebrand solubility parameter at 298 K for ethaline and glycaline, along with the molar volume and the ΔH_{vap} .

Table 5.4 – Hildebrand solubility parameter at 298 K.

	molar volume	ΔH_{vap}	b	δ (MPa ^{0.5})	
	cm ³ /mol	kJ/mol	kJ/mol	This work	literature ¹⁴²
Ethaline	78.7	42.5	0.0113	22.6	24.5
Glycaline	86.1	133.2	0.0449	39.0	26.8

5.2. Vapor-liquid Equilibrium data

The vapor pressure of the mixture with ethaline and glycaline was measured using the synthetic method. This involves measuring the equilibrium pressure of a multiphase mixture with a known overall composition at constant temperature which allows deducing phase compositions using material balance principles. Without phase sampling, the static-synthetic method relies on thermodynamic models to carry out material balance calculations, and the selection of these models can potentially impact the outcomes of the experimental measurements. This calculation will be featured in the subsequent section.

5.2.1. Data treatment

Upon establishing the global composition, to initiate the data treatment, an initial assumption is made, in which the global composition (z_i) is presumed to be equivalent to the liquid composition (x_i). This assumption sets the first guess for predicting the bubble pressure using Equation 5.3.

$$P^{cal} = \sum_{i=1}^{NC} x_i \gamma_i P_i^{Sat} \quad (5.3)$$

where γ_i is the activity coefficient of component i in the liquid phase, P_i^{Sat} is the vapor pressure of the pure compound. An activity coefficient model (such as non-random two-liquid, NRTL, or Wilson) is selected to represent the liquid phase. The parameters of the model are minimized by the objective function of the pressure (Eq. 5.4).

$$OF = \frac{1}{N_{data} - N_{\gamma \text{ parameters}}} \sum_{k=1}^{NP} \left(\frac{|P_k^{cal} - P_k^{exp}|}{P_k^{exp}} \right)^2 \quad (5.4)$$

The entire procedure was presented in detail in chapter 2, section 2.5.5. The procedure presented in the flowchart is repeated until the change in both the vapor and liquid phases' mole numbers (or composition) are below the tolerance. In this work, a deviation lower than 10^{-8} is applied.

5.2.2. Results and discussion

In Figure 5.5, VLE data at temperatures of 303, 313, 323 and 333 K for ethaline (1) and ethanol (2) mixture are illustrated. The experimental data is represented by these symbols, providing validation that the predictions align well with the experimental data. The solid black lines correspond to the Wilson model, while the dashed lines correspond to the NRTL model. A zoomed-in view at low concentrations is

presented to provide a detailed analysis of this critical section of the isotherms. This focused perspective is crucial in evaluating the accuracy of the predictions.

The two isotherms at the bottom of Figure 5.5 correspond to the temperatures of 303 K and 313 K. At these temperatures, both the NRTL and Wilson models demonstrate a similar level of predictive accuracy. However, as the temperature increases, a slight deviation from the experimental values is observed for both models.

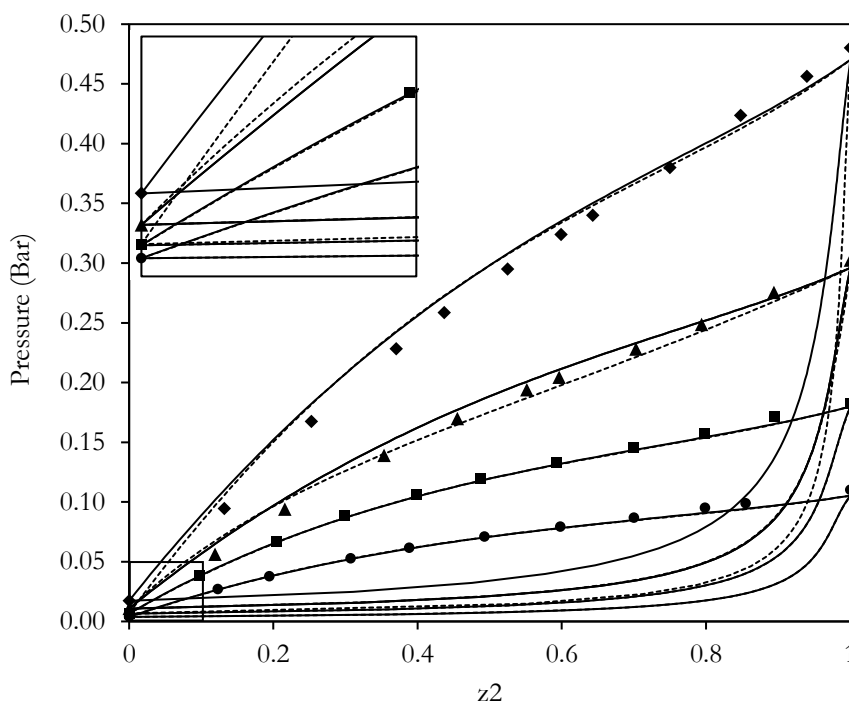


Figure 5.5 – VLE isotherms at (●)303 K, (■) 313 K, (▲) 323 K, and (◆) 333 K. for Ethaline (1) and Ethanol (2) mixture. Symbols: Experimental data. Black lines: Wilson. Dash lines: NRTL.

In Figure 5.6, VLE data at 303, 313, 323 and 333 K of ethaline (1) and isopropanol (2) are presented. Upon a more detailed analysis, it becomes apparent that at lower temperatures, the isotherms present a good agreement with the experimental data. Interestingly, as the temperature increase, particularly at 323 K and 333 K, the Wilson model demonstrates better agreement with the experimental data, showcasing its enhanced performance compared to the NRTL model.

Figure 5.7 presents VLE data for glycaline (1) and ethanol (2) mixture at 303, 313 and 323K. It is noteworthy to emphasize that at all studied temperatures, a significant observation is that all computational models consistently demonstrate a systematic tendency to underestimate the pressure when compared with the experimental data.

Figure 5.8, VLE data at 303, 313 and 323 K for glycaline (1) and isopropanol (2) mixture are illustrated. At 303 K, both models predict with remarkable accuracy, showcasing a good match with the experimental data. However, as the temperature increases to 313 and 323 K, the performance of the Wilson model is better than NRTL. However, it consistently underestimates the pressure values compared to the experimental data. Conversely, the isotherms prediction for this system using the NRTL model exhibits the least accurate results among the isotherms presented.

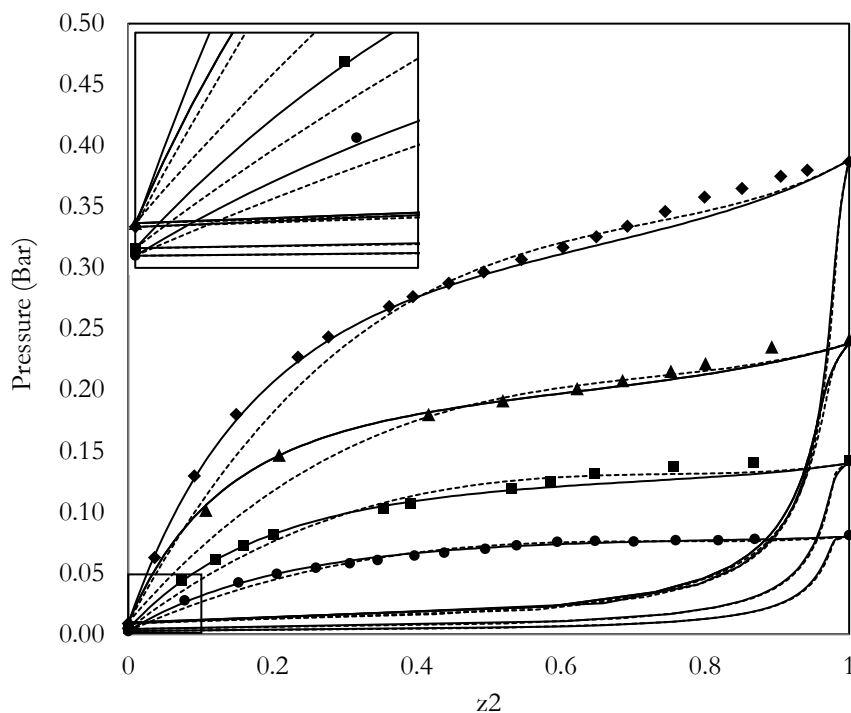


Figure 5.6 – VLE isotherms at (●)303 K, (■) 313 K, (▲) 323 K, and (◆) 333 K of Ethaline (1) and Isopropanol (2). Symbols: Experimental data. Black lines: Wilson. Dash lines: NRTL.

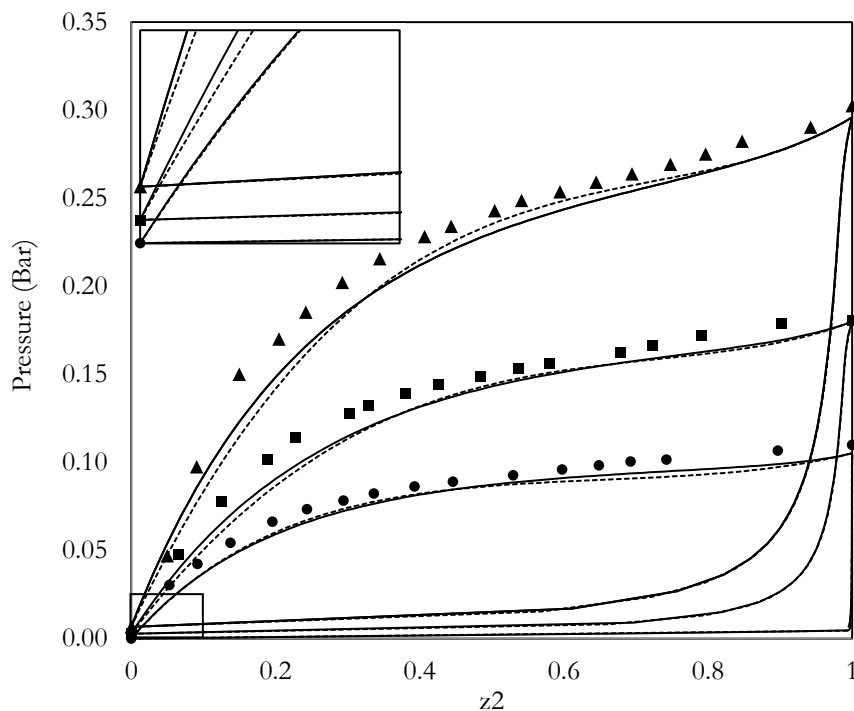


Figure 5.7 – VLE isotherms at (●)303 K, (■) 313 K, (▲) 323 K of Glycaline (1) and Ethanol (2). Symbols: Experimental data. Black lines: Wilson. Dash lines: NRTL.

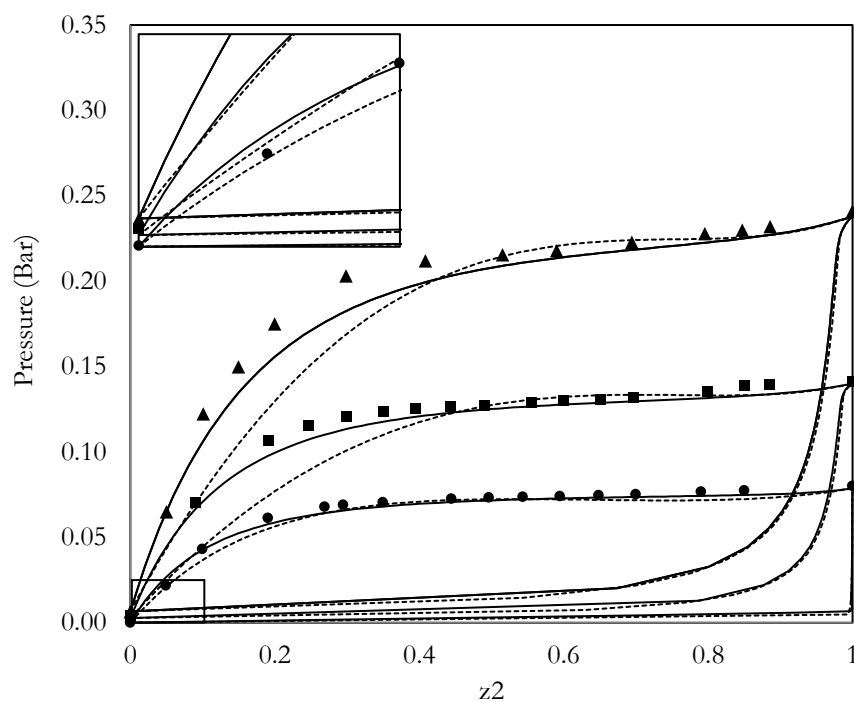


Figure 5.8 – VLE isotherms at (●)303 K, (■) 313 K, (▲) 323 K of Glycaline (1) and Isopropanol (2). Symbols: Experimental data. Black lines: Wilson. Dash lines: NRTL.

5.3. COSMO-SAC predictions

The experimental results obtained in this study were compared with predictions made using the COSMO-SAC-dps model proposed by Hsieh (2014).¹²³ It's important to highlight that the COSMO-SAC-dps model is not well-suited for making predictions involving deep eutectic solvents. Moreover, the COSMO-SAC model predicts activity coefficients lower than one for the system under investigation, specifically for the Glycaline and Isopropanol mixture (figure 5.9). The negative deviation from Raoult's law, as predicted by the COSMO-SAC model, was not observed in the experiments performed.

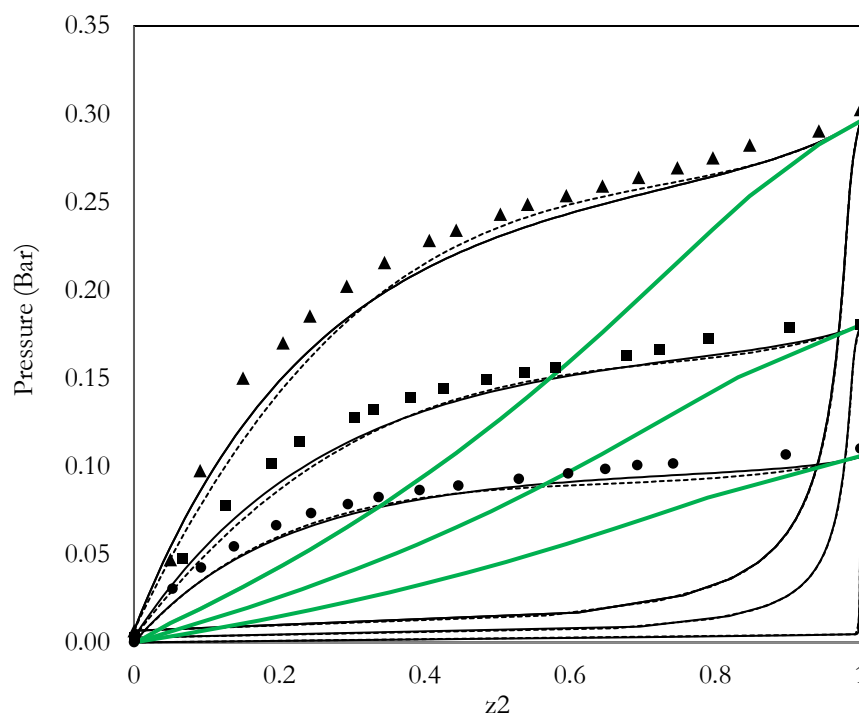


Figure 5.9 – VLE isotherms at 303 K of Glycaline (1) and Isopropanol (2). Symbols: Experimental data. Black lines: Wilson. Dash lines: NRTL and green lines: COSMO-SAC-dps.

5.4. Infinite Dilution Activity Coefficient (IDAC) of DES

We used the gas stripping method to predict the IDAC of Ethaline, aiming to provide valuable insights into its behavior under different conditions. The gas stripping method, a widely employed technique, offers a practical approach to estimate IDAC and understand the solute-solvent interactions in this specific context.

Figure 5.9 presents results (peak area as function of time) at 323 K for Ethaline (1) and Ethanol (2) mixture at various flow rates, aiming to determine the IDAC of DES. The experimental data is represented by symbols, and a black line describes the linear regression. However, the results remain inconclusive due to the significant viscosity of the solvent. The gas stripping method employed was not specifically designed

for highly viscous solvents, contributing to the challenges in obtaining conclusive data for the IDAC in this scenario.

In this context, the regression obtained in the three graphs, each varying the flow rate, is expected to the same slope. This slope is crucial as it allows us to determine the Apparent Henry's Law coefficient (H_i). The relationship is represented by Equation 2.21 in chapter 2, where the slope of the regression lines is a fundamental parameter in calculating H_i .

According to Equation 2.22 in chapter 2, the IDAC is calculated using the H_i . Ideally, at the same temperature and varying flow rates, the obtained IDAC values should remain consistent and unchanged. However, as observed in Table 5.5, if the IDAC values differ significantly. It's essential to investigate and rectify these inconsistencies to ensure accurate and reliable results in the study.

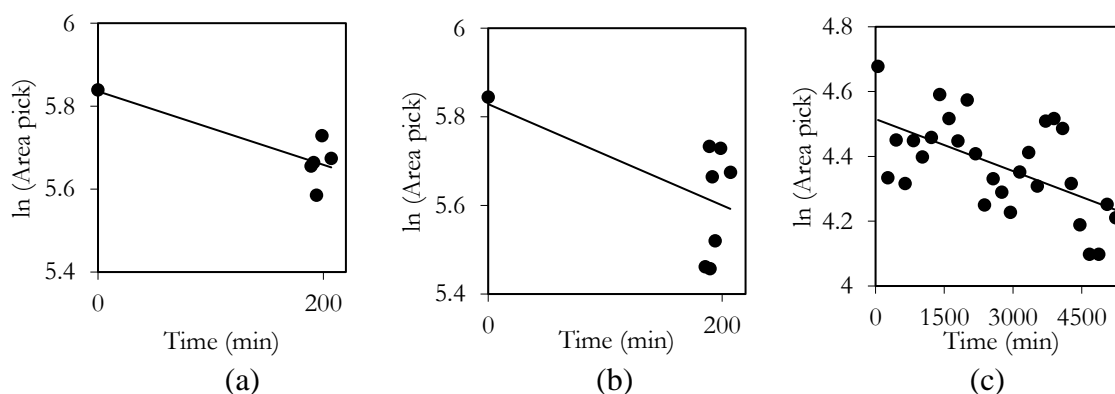


Figure 5.9 – Peak area as function of time at different flowrates (a: 22.5, b: 34.5 and c: 47.2 mL/min) and at 323 K for Ethaline (1) and Ethanol (2) mixture. Symbols: Experimental data. Black line: linear regression.

The observed discrepancies in IDAC values (detailed in Table 5.4) are directly influenced by the high viscosity of the solvent. High viscosity in a fluid implies a resistance to flow or movement due to internal friction between its molecules and is linked with number of Reynolds. In the context of the gas stripping method applied, this high viscosity significantly alters the dynamics of bubble formation and behavior. The increased resistance to diffusion resulting from the elevated viscosity, and related to the Sherwood number, leads to low mass transfer. Moreover, it was observed the formation of smaller bubbles during the gas stripping process. In which, this viscosity-induced effect can cause irregularities in the breakage and coalescence of bubbles, leading to unpredictable vapor liberation patterns. These combined effects contribute to the observed variations in the IDAC values obtained.

Table 5.5 – IDAC of DES and Ethanol for different flowrate at 323 K.

Flowrate (mL/min)	IDAC	ln(IDAC)
22.5	14.5	2.67
34.5	12.8	2.55
47.2	1.88	0.634

The similar observations regarding bubble dynamics due to high viscosity were also evident in the experiments with ethaline in pentanol, illustrated in Figure 5.10. In which present the isotherms at 323 K under various flow rates. The distinctive features associated with high-viscosity solvents persisted, influencing bubble formation, size, and behavior during the gas stripping process. These effects were reflected in the irregularities and variations seen in the isotherms.

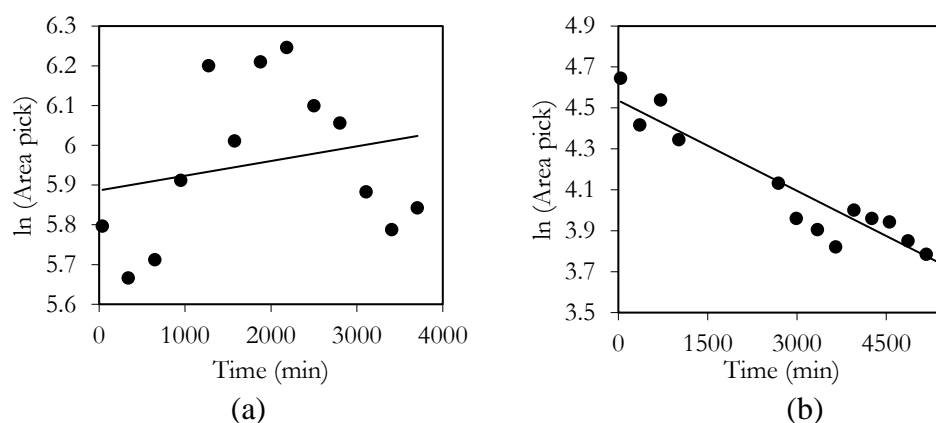


Figure 5.10 – Peak area as function of time at different flowrates (a: 47.4 and b: 61.2 mL/min) and at 323 K for Ethaline (1) and Pentanol (2). Symbols: Experimental data. Black line: linear regression.

In conclusion, the findings from the presented experiments, notably highlighted in Figures 5.9 to 5.10, bring attention to the challenges posed by high viscosity when utilizing the gas stripping method. The observed irregularities and variations in isotherms, as well as discrepancies in IDAC values, underscore the limitations imposed by high viscosity in this context. Consequently, it can be concluded that the gas stripping method may not be optimally suited for studying IDAC in fluids with high viscosity. Alternative or modified methodologies may be necessary to ensure reliable and meaningful results in such scenarios.

5.5. Conclusion

In summary, figures 5.6 to 5.9 offer a comprehensive insight into the VLE data for 4 studied mixtures, encompassing various temperature and substance combinations. The isotherms for ethaline and ethanol validate the precision of predictive models, as evidenced by the close correspondence between model predictions and experimental data. It's noteworthy that the Wilson model showcases better agreement as temperatures increase, underscoring its efficacy compared to the NRTL model.

Analyzing the experiments reveals a critical aspect concerning pressure measurements. The observed low-pressure values present a potential source of uncertainty in the results. Given that the pressure transducer is calibrated for a range of 1 to 5 bar, employing a specialized transducer designed for low-pressure measurements could enhance the accuracy and resolution of pressure data. This consideration is crucial, especially when dealing with systems where the pressure values fall within a very low range.

In the isotherms for ethaline and isopropanol mixture, good agreement of the experimental data with the predictions at temperatures of 303 and 313 K. As temperatures rise, the Wilson model exhibits better agreement than NRTL. For glycaline and ethanol mixture, the data are underestimated across all models, particularly at temperatures of 323 and 333K. Lastly, the data obtained for glycaline and isopropanol mixture reaffirm the importance of accurate modeling, with the Wilson model performing notably well at higher temperatures.

The study compared experimental data with predictions made using the COSMO-SAC-dps model, confirmed that the model is not suitable for deep eutectic solvents, and predicted activity coefficients below one for the Glycaline and Isopropanol mixture. This suggested a negative deviation from Raoult's law. However, experimental results did not confirm this deviation, highlighting a discrepancy between the predictions and experimental data. In the upcoming chapter, a more in-depth exploration of predictions using a COSMO-based model will be undertaken.

Chapter 6

Phase equilibrium prediction

Prédiction de l'équilibre de phase

Résumé

Ce chapitre se concentre sur l'équilibre des phases des DES, en particulier sur l'équilibre liquide-vapeur (ELV) et sur l'application du modèle COSMO-SAC-dps proposé par Hsieh et al. (2014). Ce modèle qui comprend la contribution dispersive est présenté au chapitre 5. Cependant, il est noté que ce modèle sous-estime souvent la pression d'ELV pour les mélanges de DES. Pour explorer l'influence des variables de calcul sur la performance du modèle, une étude préliminaire a été menée sur des solvants conventionnels avant d'aborder les mélanges de DES. Cette étude sert de base à l'établissement de la base de données des profils sigma PSL. Elle effectue un benchmark complet des méthodes théoriques et des variables de calcul pour comprendre leur impact sur la précision des prédictions du modèle COSMO-SAC.

L'étude commence avec l'utilisation du furane dans un solvant apolaire comme référence pour évaluer l'influence de la base sur la précision du modèle COSMO-SAC dans la prédiction de l'IDAC du furane dans le toluène. Les calculs de DFT utilisent la fonctionnelle hybride PBE0 et neuf bases de valence fractionnée de complexité variable sont testées. Les résultats mettent en évidence la sensibilité des prédictions de l'IDAC à la complexité de la base. Les bases plus avancées, telles que la base triple- ζ : 6-311G(2df,p), permettent d'obtenir une précision prédictive accrue, une déviation de la moyenne absolue (MAD) plus faible et un pourcentage MAD (%MAD) plus bas. L'inclusion de fonctions de polarisation supplémentaires contribue à une meilleure précision, tandis que les bases de diffusion entraînent une MAD plus élevée. En conséquence, la base 6-311G(2df,p) est choisie comme la plus adaptée sur la base de cette étude.

Après avoir sélectionné la base 6-311G(2df,p), une étude ultérieure vise à évaluer l'impact de différentes fonctionnelles de DFT sur la précision des prédictions de l'IDAC. Cette étude utilise le furane dans un solvant apolaire comme référence et inclut des fonctionnelles de DFT telles que B3LYP, PBE0, M06-2X et CAM-B3LYP, représentant les familles de fonctionnelles hybrides générales et hybrides à séparation de portée. PBE0 et M06-2X se distinguent avec les valeurs de %MAD les plus faibles (respectivement 0,4% et 0,5%), ce qui indique leur grande précision prédictive pour les prédictions de l'IDAC concernant le furane dans le toluène. Cette étude s'étend à la comparaison de l'IDAC du furane dans le cyclohexane avec les prédictions à l'aide de l'approche COSMO-SAC. Notamment, PBE0 est reconnue comme la fonctionnelle la plus performante parmi les options sélectionnées, malgré une déviation aussi élevée que 30%. La deuxième meilleure performeuse, M06-2X, présente une erreur significative de 44%.

Sur la base des connaissances acquises grâce à ces études, la fonctionnelle PBE0 et la base 6-311G(2df,p) sont choisies pour effectuer les calculs de DFT nécessaires à l'établissement de la base de données des profils sigma PSL. Cette base de données est appliquée dans le modèle COSMO-SAC pour prédire l'IDAC dans des solvants conventionnels et ces prédictions sont comparées avec le modèle UNIFAC(Do) modifié. L'étude examine les mélanges binaires de 46 composants classés en 10 combinaisons principales de familles. Un total de 356 points de données a été analysé en utilisant la MAD comme principale métrique. La conclusion générale est que le modèle UNIFAC(Do) surpasse généralement le COSMO-SAC-dps dans la prédiction de l'IDAC pour la plupart des systèmes, avec des prédictions hautement précises observées uniquement pour les solvants aromatiques dans les alcanes et les alcools dans les cycloalcanes.

L'étude examine également la définition d'une cavité dans le contexte des profils sigma, impliquant la spécification des rayons de cavité. Les profils sigma UD utilisent généralement des rayons Klamt, tandis que d'autres logiciels comme Gaussian utilisent des rayons UFF basés sur le modèle de force universelle (UFF). De manière significative, les résultats montrent que le COSMO-SAC avec la base PSL (rayons Klamt) offre de meilleures performances que la base UD et même la base PSL (rayons UFF) surpasse légèrement les rayons Klamt, bien qu'elle sous-estime légèrement les valeurs de l'IDAC.

Le chapitre présente également une évaluation des prédictions des coefficients d'activité en utilisant différents modèles et bases de profils sigma. Les résultats montrent que le modèle UNIFAC(Do) atteint une bonne précision pour la plupart des systèmes, tandis que les performances du COSMO-SAC varient en fonction de la base de profils sigma utilisée. Enfin, l'étude discute de l'évolution des coefficients d'activité pour trois systèmes représentatifs et compare la performance de différents modèles pour chacun d'entre eux. Le choix du modèle le plus performant dépend du système spécifique et de la température.

Phase equilibrium prediction

6.1. Introduction

The phase equilibrium of DESs (i.e., VLE) were performed, analyzed and predicted using the COSMO-SAC-dps proposed by Hsieh et al. (2014)¹²³ including dispersive contribution. These results have been presented in Chapter 5. Nevertheless, the COSMO-SAC-dps model underestimated the VLE for the DES mixture. The precision of thermodynamic calculations in COSMO-based models heavily depends on the molecule-specific sigma-profiles. To investigate and comprehend how computational variables (used to generate these profiles) influence the performance of the model, we conducted an initial study using conventional solvents before examining the DES mixture. This study forms the foundation for establishing the PSL sigma-profile database (Fig. 6.1). PSL refers to Université Paris Sciences & Letters and is named following the nomenclature of other profiles such as VT-2005 Sigma-profile developed by Mullins under the supervision of Liu's group at the University of Virginia Tech and the UD database, created by Sandler's group at the University of Delaware.

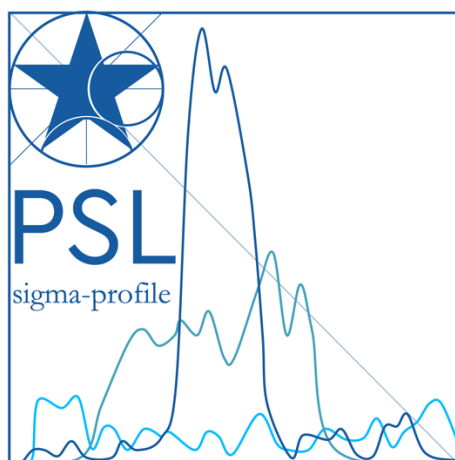


Figure 6.1 – PSL sigma-profile.

A sigma-profile (derived from QM calculations using solvation models), represents the molecular-specific distribution of surface charge density. It provides a detailed representation of how the charge is distributed over a molecule's surface. This crucial information forms the core of the model and enables precise calculations and predictions of phase equilibria and various properties. However, it's important to note that QM calculations, particularly DFT, constitute the most time-consuming and computationally expensive aspect of COSMO-based models.

In this chapter, we introduce the PSL sigma-profile database (an open accessible to both academic and non-commercial users). A comprehensive benchmarking of various theoretical methods and

computational variables has been conducted to comprehend their impact on prediction accuracy using the COSMO-SAC model. Furthermore, these predictions have been compared with established activity coefficient models including NRTL, UNIFAC, and Scatchard-Hildbrand for a comprehensive assessment.

6.2. PSL sigma-profile

6.2.1. Setting the basis set

In this section, we initiate our investigation by using furan in an apolar solvent as a benchmark. Our primary objective is to assess the basis set dependence of the accuracy of COSMO-SAC model to predict IDAC of furan in toluene¹⁴³. To achieve this, DFT calculations were conducted following the procedure already described in chapter 4, section 4.3.5. Specifically, the hybrid functional PBE0, proposed by Adamo et al. (1999)¹⁴⁴, was employed as a theoretical level to assess the influence of nine different split-valence basis set. This study has been published (see the article of G. Miyazaki *et al.*¹⁴⁵ entitled “Role of Computational Variables on the Performances of COSMO-SAC Model: A Combined Theoretical and Experimental Investigation”).

Errors were evaluated using the mean absolute deviation (MAD) between the experimental and theoretical logarithm of IDACs:

$$MAD = \frac{1}{N} \sum_{i=1}^N |\ln(\gamma_{i,exp}^{\infty}) - \ln(\gamma_{i,th}^{\infty})| \quad (6.1)$$

The %MAD is then defined as

$$\%MAD = \sum_{i=1}^N \left| \frac{\ln(\gamma_{i,exp}^{\infty}) - \ln(\gamma_{i,th}^{\infty})}{\ln(\gamma_{i,exp}^{\infty})} \right| \quad (6.2)$$

The nine selected split-valence basis sets vary in complexity, ranging from the fundamental 6-31G(d) to the most sophisticated 6-311++G(2df,p), illustrating a diverse range of basis set variations. As explained previously, a split-valence basis set is a type of basis, where the basis functions for each atomic center are divided into two groups: core and valence. Common split-valence basis sets include the popular family of basis sets developed by Pople and co-workers^{146–148} such as the 6-31G(d) and 6-311G(d) basis sets. The notation "6-31G(d)" signifies that each inner shell (1s orbital) is represented using a linear combination of 6 primitive Gaussians and the valence shell is split into an inner part (described by 3 primitives) and an outer part (described by 1 primitive). Additionally, d-type polarization functions are included to improve accuracy in describing electronic behavior in the valence shell. Table 6.1. summarizes Pople's nomenclature

for basis sets(detailing the number and types of functions for different electronic shells and polarization functions).

Table 6.1 – Pople's nomenclature for basis sets.

Basis set	Description
6	Gaussians primitive in the core orbital
3	Gaussians primitives in the inner valence shell
1	Gaussian primitive in the outer valence shell (valence orbital duplicated)
1	Gaussian primitive in the outer valence shell (valence orbital triplicate)
+ or ++	Diffuse functions on heavy atoms or Diffuse functions in both, heavy and hydrogen atoms
2df	2 additional d-type polarization functions and one f-type functions on heavy atoms
p	One set of p-type polarization function on hydrogen atoms

Table 6.2 exemplifies the significant impact of basis sets on accurately predicting IDAC for furan in toluene. The results demonstrate the sensitivity of IDAC predictions to the complexity of the basis set, where employing a more sophisticate the basis set leads to lower MAD and %MAD. There is a notable enhancement in predictive accuracy, attributed to the inclusion of additional basis functions such as polarization functions.

Table 6.2 – MAD of IDAC predictions of furan in toluene.

basis set	MAD	%MAD
6-31G(d)	0.012	1.1
6-311G(d)	0.005	0.4
6-31G(d,p)	0.012	1.1
6-31G(2df,p)	0.003	0.3
6-311++G(d)	0.007	6.4
6-31++G(d,p)	0.036	3.5
6-31++G(2df,p)	0.026	2.5
6-311G(2df,p)	0.004	0.4
6-311++G(2df,p)	0.017	1.7

All data were computed using PBE0.

The addition of the diffusion basis sets led to higher MAD. Specifically, the split-valence double- ζ basis set 6-31G(d,p),the triple- ζ basis sets 6-311G(d) and 6-311G(2df,p) exhibited superior results. Consequently, based on this study, the triple- ζ basis set with lower MAD(6-311G(2df,p)) was selected.

6.2.2. Setting the level of theory

Once the basis set 6-311G(2df,p) was chosen, a secondary investigation was conducted to assess the influence of the theoretical level on the accuracy of IDAC predictions. This assessment utilized furan in an apolar solvent as the benchmark. In chapter 4, section 4.2.3, DFT functionals were presented, each

distinguished by its exchange-correlation (XC) contribution to the total energy. Among these options, specific functionals were selected, including B3LYP¹⁴⁹, PBE0¹⁴⁴, M06-2X¹⁵⁰, and CAM-B3LYP¹⁵¹. The findings from this investigation have been published in our paper¹⁴⁵. These chosen functionals represent the most significant and commonly used families of DFT functionals, namely the general hybrid functionals family (B3LYP, PBE0, M06-2X) and the Range-separated hybrid functionals family (CAM-B3LYP).

Table 6.3 summarizes the MADs of computed IDAC for furan in toluene across various temperatures using different functionals. Among the functionals listed, the two with the lowest % MAD are PBE0 (0.4%) and M06-2X (0.5%) indicating that they have the best predictive accuracy for the IDAC predictions of furan in toluene.

Table 6.3 – IDAC of furan in toluene, experimental and predicted with COSMO-SAC-dps¹²³.

T (K)	IDAC experimental ¹⁴³	B3LYP	CAM-B3LYP	M06-2X	PBE0
293.15	1.037	1.078	1.044	1.039	1.030
298.15	1.033	1.075	1.042	1.037	1.029
303.16	1.034	1.072	1.039	1.036	1.027
308.19	1.025	1.070	1.037	1.034	1.026
313.17	1.022	1.067	1.036	1.033	1.025
318.12	1.028	1.064	1.034	1.031	1.023
323.09	1.026	1.062	1.032	1.030	1.022
MAD		0.040	0.008	0.005	0.004
%MAD		3.9	0.8	0.5	0.4

All data were computed using 6-311G(2df,p) basis set.

Once the first two computational parameters were validated, IDAC of furan in cyclohexane were compared with those predicted using the COSMO-SAC approach. The comparison results are summarized in Table 6.4. Notably, PBE0 is identified as the top performer among the selected functionals and displayed a deviation as high as 30%. On the other hand, the second-best performer (M06-2X) exhibited a substantial error of 44%.

Table 6.4 – IDAC of furan in cyclohexane, experimental and predicted with COSMO-SAC-dps¹²³.

T	IDAC experimental ¹⁴⁵ .	M062X	PBE0
283.59	2.284	3.15	2.894
288.55	2.257	3.043	2.801
293.55	2.068	2.943	2.715
298.2	2.017	2.856	2.64
303.35	1.973	2.767	2.562
308.48	1.916	2.684	2.490
313.41	1.835	2.610	2.426
318.36	1.763	2.540	2.366
MAD		0.897	0.598
%MAD		44.5	29.7

Based on the findings from these studies, the PBE0 functional and the triple-zeta basis set 6-311G(2df,p) were selected for conducting the DFT calculations necessary to build the PSL sigma-profile database.

6.3. IDAC predictions in conventional solvents

The performance assessment of COSMO-SAC model was conducted by studying the IDAC predictions for conventional solvents, comparing them with the group contribution model modified UNIFAC(Do). In this evaluation, binary mixtures of 46 components were investigated and systematically categorized into 10 primary family combinations (as outlined in Table 6.5). This comprehensive study resulted in the examination of 356 data points¹⁵². To quantify the accuracy of each prediction method, the MAD was calculated and will be discussed in detail below.

Table 6.5 – Main family structure studied.

Family	members	examples
n-Alkanes	16	n-pentane, n-hexadecane, n-eicosane
R-Alkanes	3	methyl cyclopentane, methyl cyclohexane, ethylcyclohexane
Alkenes	5	2-pentene, 1-hexene, 1-octene
Alcohols	6	methanol, ethanol, n-propanol
Cycloalkanes	2	cyclohexane, cycloheptane
Cycloalkenes	4	cyclopentene, cyclohexene, cycloheptene, cyclooctene
Carboxylic Acid	2	acetic acid, proionic acid
Ketone	4	acetone, 2-butanone, 3-pentanone, 5-nonanone
Amide	1	dimethylformamide
Aromatic	3	benzene, toluene, phenol

Figure 6.2a illustrates the Mean Absolute Deviation (MAD) for various combinations of solvent families and solutes using the COSMO-SAC-dps model. The horizontal axis represents solvent families, and the vertical axis represents solutes. Different colors in the figure represent the range of MAD values, and only

main families with available data are presented. Figure 6.2b does the same analysis but focuses on the modified UNIFAC (Do) model.

The general conclusion drawn from these figures is that, overall, the modified UNIFAC(Do) model outperforms the COSMO-SAC-dps model in predicting IDAC for most systems. Only two combinations of solvent families and solutes showed highly accurate predictions with a slight advantage for the COSMO-SAC-dps model. These combinations were aromatic solvents in alkanes and alcohols in cycloalkanes.

Several other combinations produced good predictions with the COSMO-SAC-dps model, while others yielded reasonably accurate predictions. However, some combinations, particularly involving n-alkanes in alcohols, carboxylic acids as solutes, and mixtures containing amides, resulted in less accurate predictions according to the COSMO-SAC-dps model. These include alcohols in alkanes, cycloalkanes, R-cycloalkanes in aromatics, and n-alkanes and cycloalkanes in carboxylic acids. However, some combinations, particularly involving alkanes in ketones and aromatics, ketones in alkenes and alcohols, cycloalkanes in alcohols, and cycloalkenes in amides, resulted in less accurate predictions according to the COSMO-SAC-dps model. including.

These figures (Figure 2a and Figure 2b) collectively offer a detailed assessment of the COSMO-SAC-dps model's performance in predicting IDAC and highlight the specific scenarios where it excels or falls short in representing experimental data.

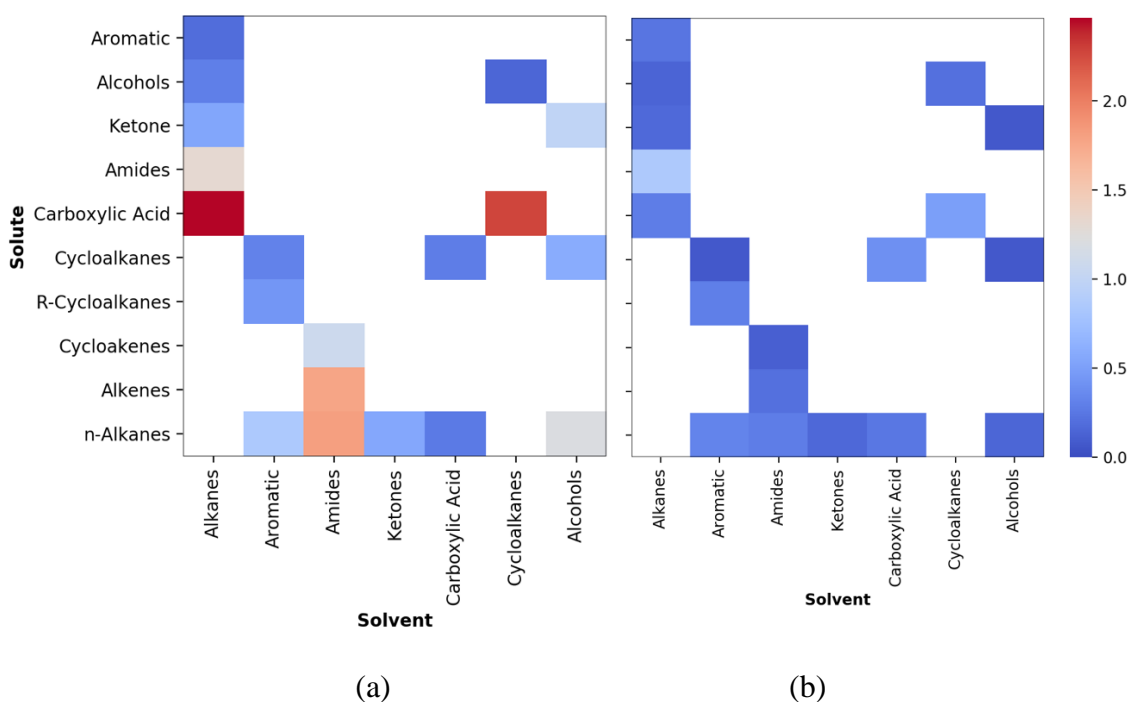


Figure 6.2 – Mean absolute deviation MAD for infinite dilution activity coefficients of conventional mixture combinations for COSMO-SAC-dps (a) and modified UNIFAC(Do) (b).

Due to the significant number of representative data points, the combination of aromatic solvents in alkanes, where the COSMO-SAC-dps model demonstrated superior predictive performance compared to UNIFAC, was chosen for a thorough examination of another computational parameter: cavity radii. The results of this analysis will be detailed in the upcoming section. This analysis will be presented in the subsequent section.

6.3.1. Cavity radii investigation

In this study, we compared the IDAC predictions obtained using COSMO-SAC-dps with the PSL sigma-profile against those utilizing the UD sigma-profile. It's important to highlight that the UD sigma-profile is an extension derived from the openly accessible VT-2005 database. The computational approach of the UD sigma-profile involves utilizing the Generalized Gradient Approximation (GGA), the VWN-BP¹⁵³ functional (which combines the Vosko, Wilk and Nusair exchange functional (VWN)¹⁵³ and the Becke exchange functional (BP)¹⁴⁹). The basis set is a double- ζ with polarization functions (DNP). In contrast, for the PSL sigma-profile, we utilized a hybrid functional (PBE0) and a split valence triple- ζ basis set with polarization functions (6-311G(2df,p)).

The definition of a cavity in the context of sigma-profiles involves specifying the radii of the cavity. In the UD sigma-profiles, the typical cavity radii are based on the Klamt radii. These radii originate from the traditional Bondi's van der Waals radii scaled by a factor of 1.17. Conversely, in computer codes like Gaussian that implement PCM-like models, different radii (referred to as UFF radii) are often used. These UFF radii are defined based on the Universal Force Field (UFF); a widely used parameterization for molecular mechanics calculations. They are easily generated for a wide range of elements, following straightforward rules. Particularly, in implementations of PCM like that in the Gaussian code, these UFF radii are used.

Figures 6.3 and 6.4 show benzene and toluene's IDAC in n-alkanes using different sigma-profiles, respectively. The results obtained using PSL database (Klamt radii) are better than UD database. Moreover, PSL database (UFF radii) are better than Klamt radii even if IDAC values is underestimated.

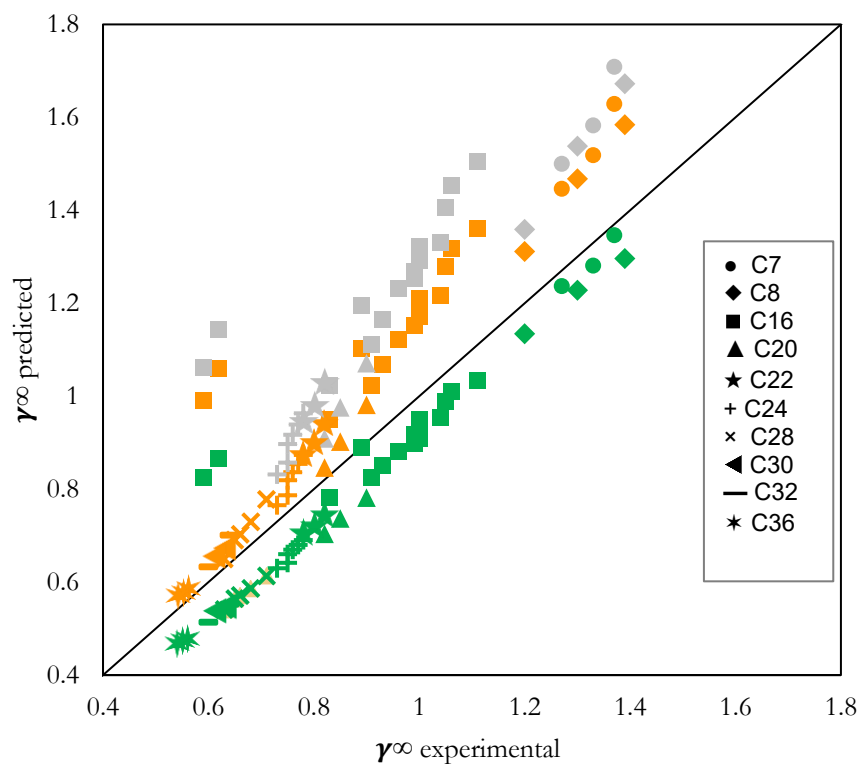


Figure 6.3 – IDAC of Benzene in n-Alkanes using COSMO-SAC with different sigma-profiles (grey: UD, orange: PSL-Klamt and green: PSL-UFF).

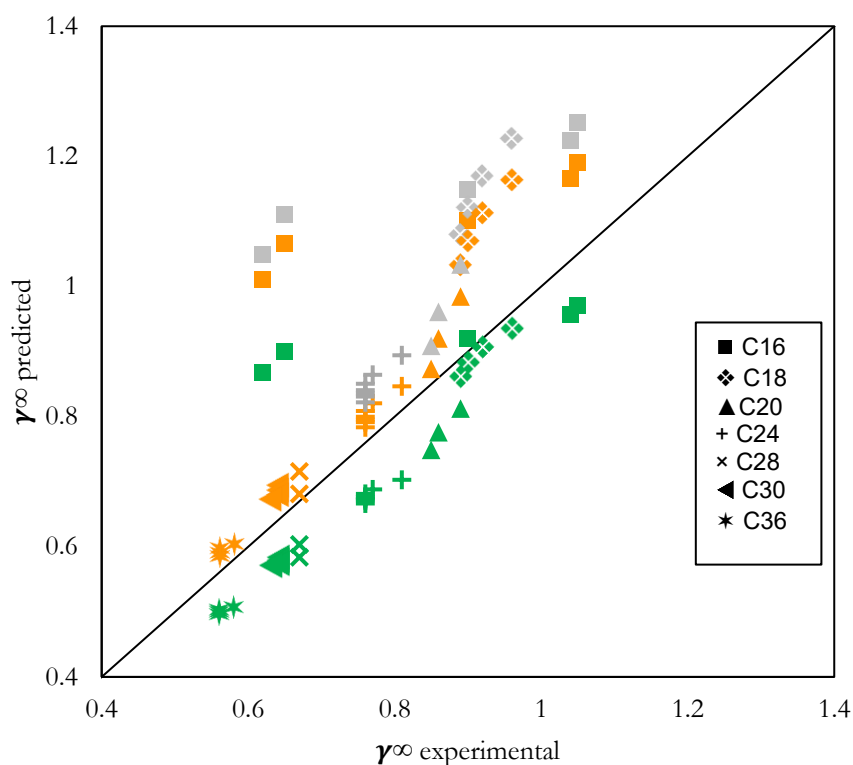


Figure 6.4 – IDAC of Toluene in n-Alkanes using COSMO-SAC with different sigma-profiles (grey: UD, orange: PSL-Klamt and green: PSL-UFF).

6.4. Phase equilibrium predictions in conventional solvents

Activity coefficients for aromatics + n-hydrocarbons have been correlated from VLE isotherms data at low pressure (below 500 kPa) (Table 6.6) using the thermodynamic equation (6.3) and assuming the vapor phase as an ideal gas and neglecting the Poynting correction factor.

$$x_i \gamma_i P_i^{sat} = y_i P \quad (6.3)$$

Where x_i is the mole fraction in the liquid phase, γ_i is the activity coefficient, P_i^{sat} is the vapor pressure, y_i is the mole fraction in the vapor phase, and P is the total equilibrium pressure. The pure component vapor pressure is correlated from the DIPPR (Design Institute for Physical Properties) database.¹⁵⁴ The correlation and the parameters used are presented in Table 6.7.

Table 6.6 – Literature data of aromatic and n-alkane binary system.

(1)	System (2)	Number of points	Temperature range (K)		Mole fraction range (1)		reference
			max.	min.	max.	min.	
o-xylene	<i>n</i> -heptane	57	348	368	0.054	0.993	155–157
o-xylene	<i>n</i> -octane	13	308.09		0.050	0.950	157
p-xylene	<i>n</i> -heptane	28	348.13		0.052	0.958	157
m-xylene	<i>n</i> -heptane	27	348.13		0.082	0.972	155
m-xylene	<i>n</i> -decane	33	373	393	0.004	0.995	158
toluene	<i>n</i> -heptane	159	298	362	0.047	0.971	159–162
toluene	<i>n</i> -octane	18	333.15		0.047	0.953	163
toluene	<i>n</i> -decane	34	373	393	0.047	0.957	158,164
benzene	<i>n</i> -hexane	65	333	343	0.028	0.972	165–167
benzene	<i>n</i> -heptane	182	293	488	0.021	0.994	168–176
benzene	<i>n</i> -octane	66	328	348	0.183	0.957	177
benzene	<i>n</i> -dodecane	38	333	353	0.122	0.890	178

Activity coefficients correlated from VLE experimental data have been used to compare the accuracy of four models: NRTL, Scatchard-Hildebrand, modified UNIFAC(Do) and COSMO-SAC-dps.

The Simulis thermodynamic software developed by Prosim (France) was used to correlate the data using NRTL and fit the NRTL model parameters. The PSL sigma-profiles were obtained based on these DFT/COSMO calculations considering two types of radii: Klamt and UFF.

Table 6.7 – Obtained NRTL parameters.

System		T (K)	Number of points	α_{ij}^0 regressed	AAE% calculated by a series of α_{ij}^0				
					0.2	0.3	0.4	0.47	regressed
o-xylene	<i>n</i> -heptane	348.13	23	0.502	0.30	0.16	0.13	0.13	0.07
o-xylene	<i>n</i> -heptane	358.08	11	0.823	0.51	0.41	0.36	0.32	0.24
o-xylene	<i>n</i> -heptane	348.08	12	0.800	0.79	0.62	0.72	0.54	0.51
o-xylene	<i>n</i> -heptane	368.28	11	0.349	0.56	0.47	0.45	0.43	0.43
o-xylene	<i>n</i> -octane	308.09	13	3.335	0.16	0.35	0.34	0.34	0.09
p-xylene	<i>n</i> -heptane	348.13	28	1.410	0.34	0.33	0.31	0.27	0.11
m-xylene	<i>n</i> -heptane	348.13	27	0.810	0.25	0.21	0.16	0.14	0.06
m-xylene	<i>n</i> -decane	348.13	11	0.247	1.24	1.25	1.25	1.25	1.25
m-xylene	<i>n</i> -decane	393.67	10	0.832	0.97	0.95	0.95	0.94	0.92
m-xylene	<i>n</i> -decane	383.57	12	0.350	1.60	1.60	1.60	1.60	1.60
toluene	<i>n</i> -heptane	298.14	49	0.72	0.71	0.70	0.72	0.69	0.72
toluene	<i>n</i> -heptane	303.14	44	0.70	0.67	0.65	0.64	0.63	0.70
toluene	<i>n</i> -heptane	313.14	25	0.53	0.58	0.49	0.49	0.47	0.53
toluene	<i>n</i> -heptane	322.99	15	0.79	0.79	0.77	0.76	0.52	0.79
toluene	<i>n</i> -heptane	332.98	15	0.82	0.78	0.74	0.70	0.75	0.82
toluene	<i>n</i> -heptane	362.98	14	8.56	8.43	8.31	8.21	7.44	8.56
toluene	<i>n</i> -heptane	347.98	17	0.81	0.78	0.74	0.71	0.57	0.81
toluene	<i>n</i> -heptane	302.99	13	1.06	1.05	1.01	0.99	0.85	1.06
toluene	<i>n</i> -octane	333.15	18	0.39	0.35	0.32	0.29	0.13	0.39
toluene	<i>n</i> -decane	373.47	12	1.60	1.61	1.60	1.60	1.60	1.60
toluene	<i>n</i> -decane	383.57	10	2.48	2.47	2.46	2.46	2.46	2.48
toluene	<i>n</i> -decane	393.67	10	3.53	3.41	3.46	3.51	3.41	3.53
benzene	<i>n</i> -hexane	333.13	32	0.605	1.60	1.89	1.80	1.71	1.66
benzene	<i>n</i> -hexane	343.12	7	0.417	0.23	0.19	0.21	0.23	0.24
benzene	<i>n</i> -heptane	333.12	15	0.994	0.25	0.96	0.87	0.79	0.73
benzene	<i>n</i> -heptane	353.12	26	1.146	0.27	0.64	0.59	0.54	0.51
benzene	<i>n</i> -heptane	348.12	8	1.414	0.20	1.07	1.02	0.96	0.92
benzene	<i>n</i> -heptane	293.14	22	1.562	0.25	0.81	0.68	0.60	0.57
benzene	<i>n</i> -heptane	488.11	9	0.629	7.34	7.59	7.53	7.47	7.43
benzene	<i>n</i> -heptane	458.11	10	0.652	0.59	0.60	0.59	0.59	0.59
benzene	<i>n</i> -heptane	473.11	10	0.639	0.62	0.63	0.63	0.62	0.62
benzene	<i>n</i> -heptane	443.11	10	3.354	0.26	0.54	0.53	0.53	0.52
benzene	<i>n</i> -heptane	428.11	9	2.720	0.40	0.83	0.82	0.80	0.79
benzene	<i>n</i> -heptane	413.12	9	1.621	0.30	0.59	0.57	0.55	0.53
benzene	<i>n</i> -heptane	383.12	17	0.647	0.96	1.14	1.11	1.07	1.04
benzene	<i>n</i> -heptane	318.14	15	0.860	0.34	1.07	0.97	0.86	0.78
benzene	<i>n</i> -octane	328.14	27	1.491	0.59	1.58	1.49	1.41	1.35
benzene	<i>n</i> -octane	348.13	13	1.510	0.49	0.90	0.86	0.82	0.79
benzene	<i>n</i> -octane	338.13	26	1.731	0.48	1.57	1.51	1.44	1.40
benzene	<i>n</i> -dodecane	333.13	11	0.721	2.23	2.24	2.23	2.23	2.23
benzene	<i>n</i> -dodecane	313.14	8	0.669	2.37	2.54	2.50	2.47	2.44
benzene	<i>n</i> -dodecane	353.13	16	0.572	2.46	2.46	2.46	2.46	2.46

The results of predicted activity coefficient compared with experimental values using the semi-empirical NRTL and the predictive model UNIFAC(Do) are presented in Fig. 6.5. Comparing to predictions obtained using NRTL, it is evident that UNIFAC(Do) demonstrates notable accuracy across the systems and temperature investigated in this study, outperforming predictions obtained using NRTL. Notably, for all systems studied, the errors (AAE%) remain under 15, with global errors of 3.06% and 4.05% for NRTL and UNIFAC(Do) respectively. It observed that the NRTL model fails the prediction near the infinite dilution, giving activity coefficient values larger than 10^{15} or tending towards infinity. This discrepancy can be attributed to the correlation of the binary interaction parameters with experimental data from the VLE, without considering IDAC. Consequently, NRTL proves insufficient for accuracy predicting IDAC.

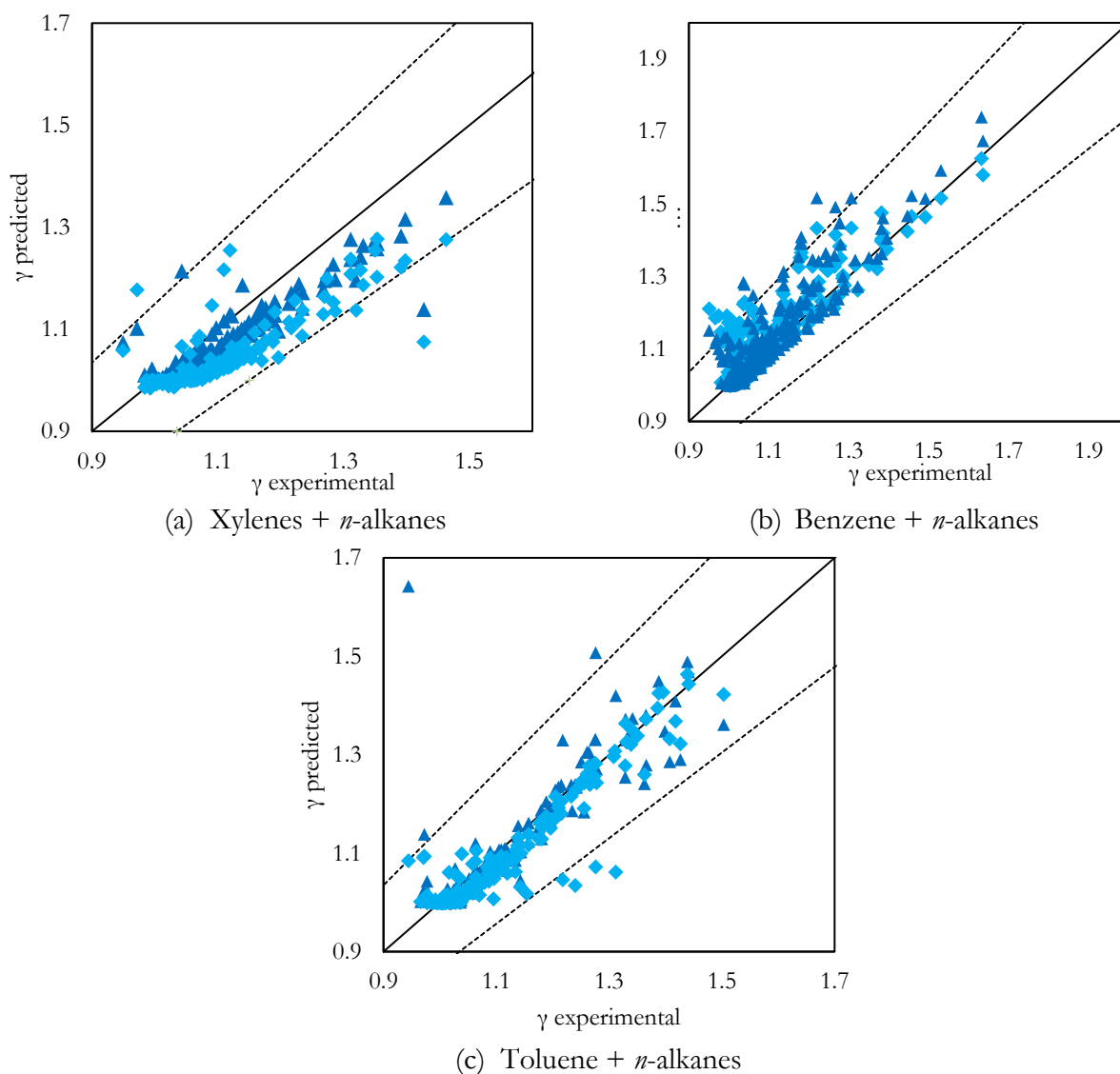


Figure 6.5 – Experimental vs. predicted activity coefficient of aromatics in *n*-alkanes using (▲) NRTL and (◆) UNIFAC Do. The dash line represents a margin of error of 15%.

Two approaches were considered regarding the prediction of activity coefficient using the Scatchard-Hildebrand model (SH). The first one, activity coefficients were predicted considering the contribution of the SH term. The second one has included the combinatorial term given by the Staverman-Guggenheim equation (SG) like UNIFAC and COSMO-SAC model. The volume and surface area were taken from the DFT calculation performed in this work. Figure 6.6 compares the experimental activity coefficient values with the predicted one using the SH model and SH model + SG term. After including the combinatorial term, the SH model was improved for benzene and toluene systems and performed equally for xylene systems. Moreover, the global errors obtained in those predictions were 3.97% and 3.82% for the SH model and SH+SG, respectively.

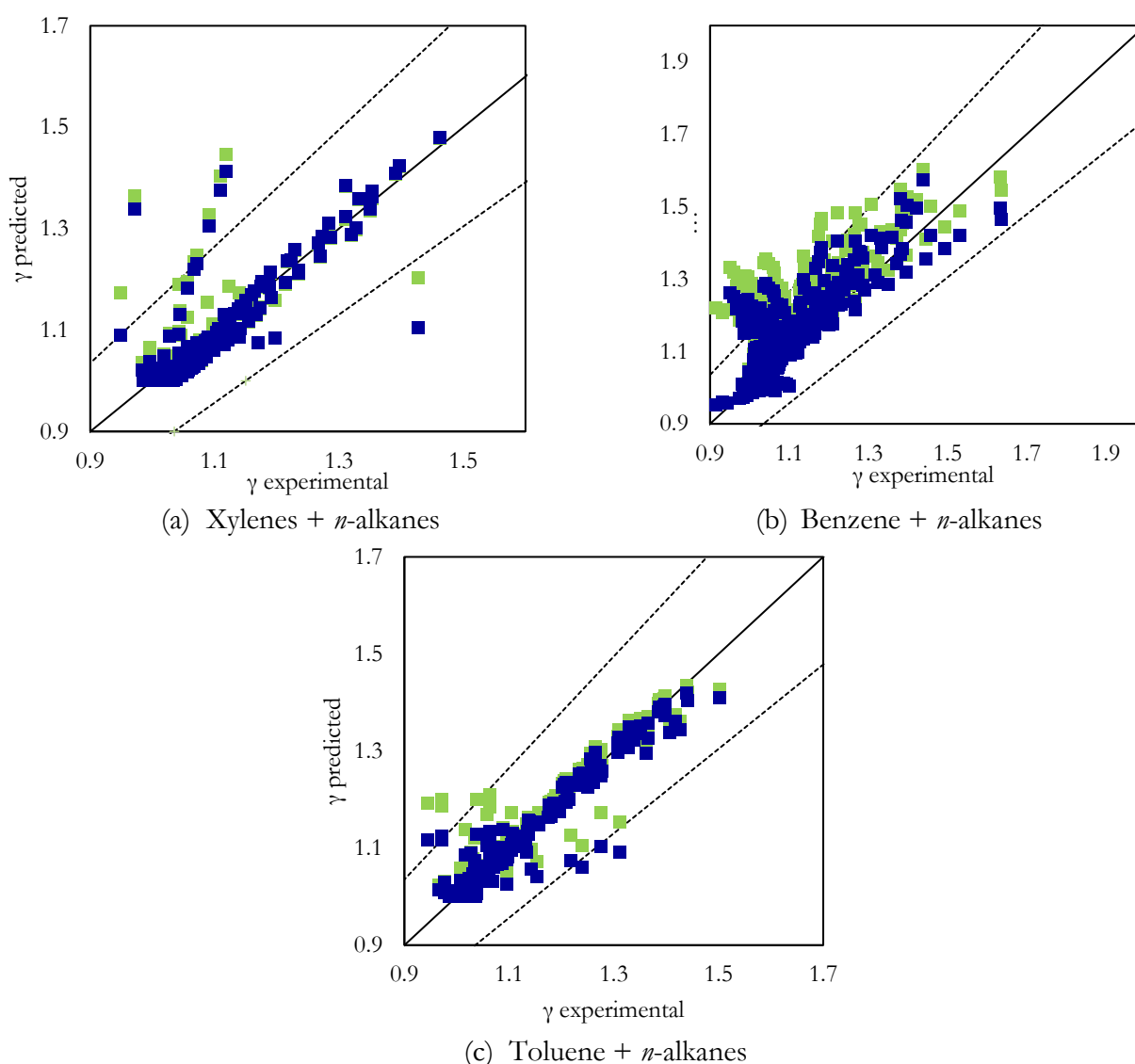


Figure 6.6 – Experimental vs. predicted activity coefficient of aromatics in *n*-alkanes using (■) Scatchard-Hildebrand (green: no combinatorial term; deep blue: with combinatorial term of Staverman-Guggenheim). The dash line represents a margin of error of 15%.

Concerning the COSMO-SAC predictions, three databases of σ -profiles were considered: UD database¹⁷⁹ (an extension of the freely available VT database) and two databases generated in this study consider two types of radii: Klamt and UFF (called PSL-Klamt and PSL-UFF databases). Detailed procedure of generating the PSL σ -profiles database by DFT/COSMO calculations were described in our previous paper¹⁴⁵. Figure 6.7 compares the experimental activity coefficient values with the predicted one using the COSMO-SAC model with UD, PSL-Klamt, and PSL-UFF σ -profile database. For almost all systems, COSMO-SAC using the PSL-UFF database performs better than using the UD database for the studied systems. Except for xylene systems, COSMO-SAC using the PSL-Klamt database performs better than PSL-UFF.

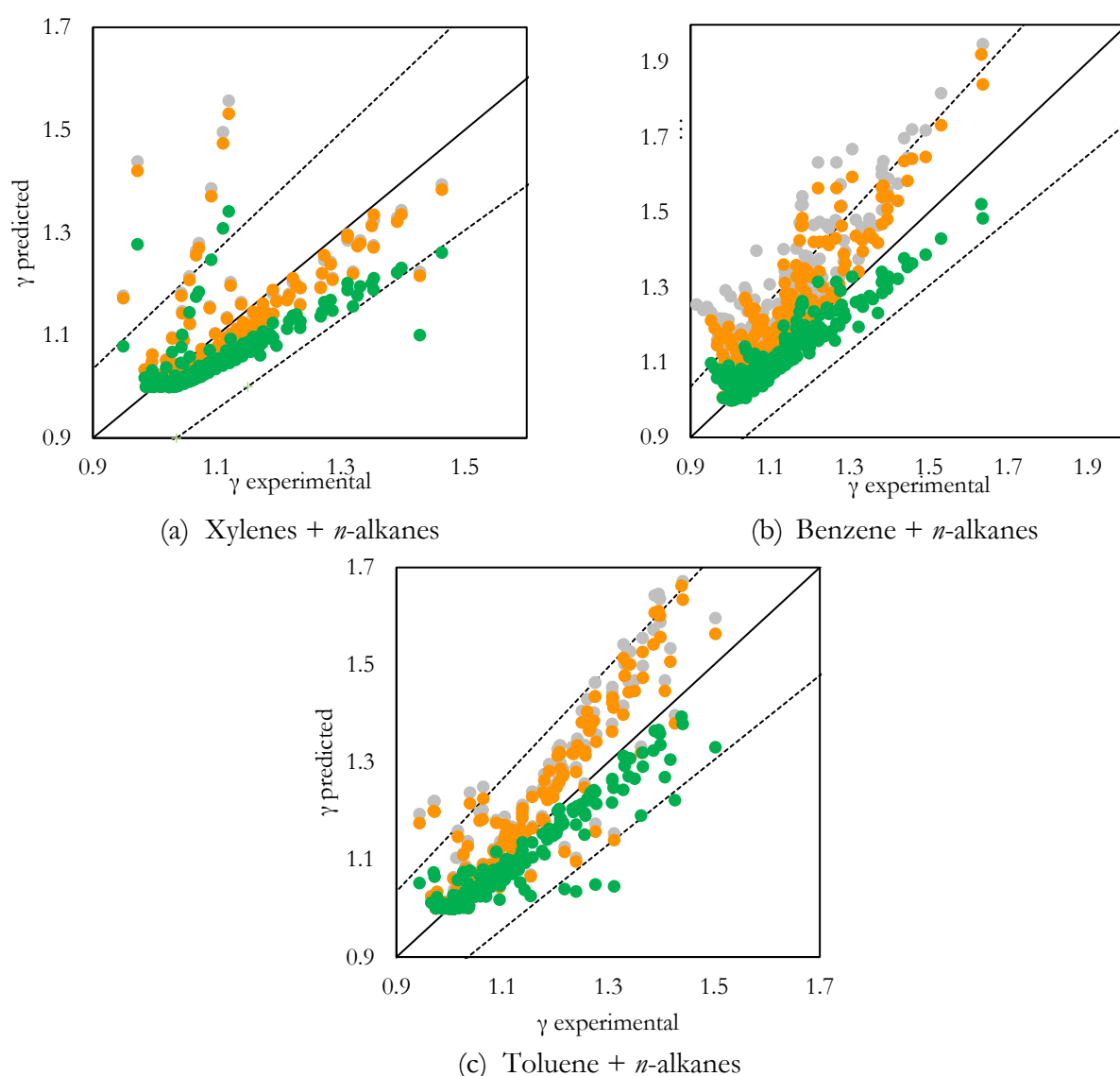


Figure 6.7 – Experimental *vs.* predicted activity coefficient of aromatics in *n*-alkanes using (●) COSMO-SAC 2010 with different sigma-profiles database (grey: UD; orange: PSL (Klamt radii) and green: PSL (UFF radii)). The dash line represents a margin of error of 15%.

However, the global error obtained in those predictions were 5.41%, 4.68% and 4.07% for UD, PSL-Klamt, and PSL-UFF databases, respectively. Thus, we can conclude that the level of theory considered in the UD database, in contrast to the level of theory and basis set considered in the PSL database (this study), impacts activity coefficient predictions using the COSMO-SAC model. Moreover, comparing the COSMO-SAC model with a semi-empirical (i.e., NRTL model), we can conclude that this model has very good accuracy for all the systems studied in this study.

A more general representation of experimental activity coefficient values with all models used to predict it in this work is presented in figure 6.8.

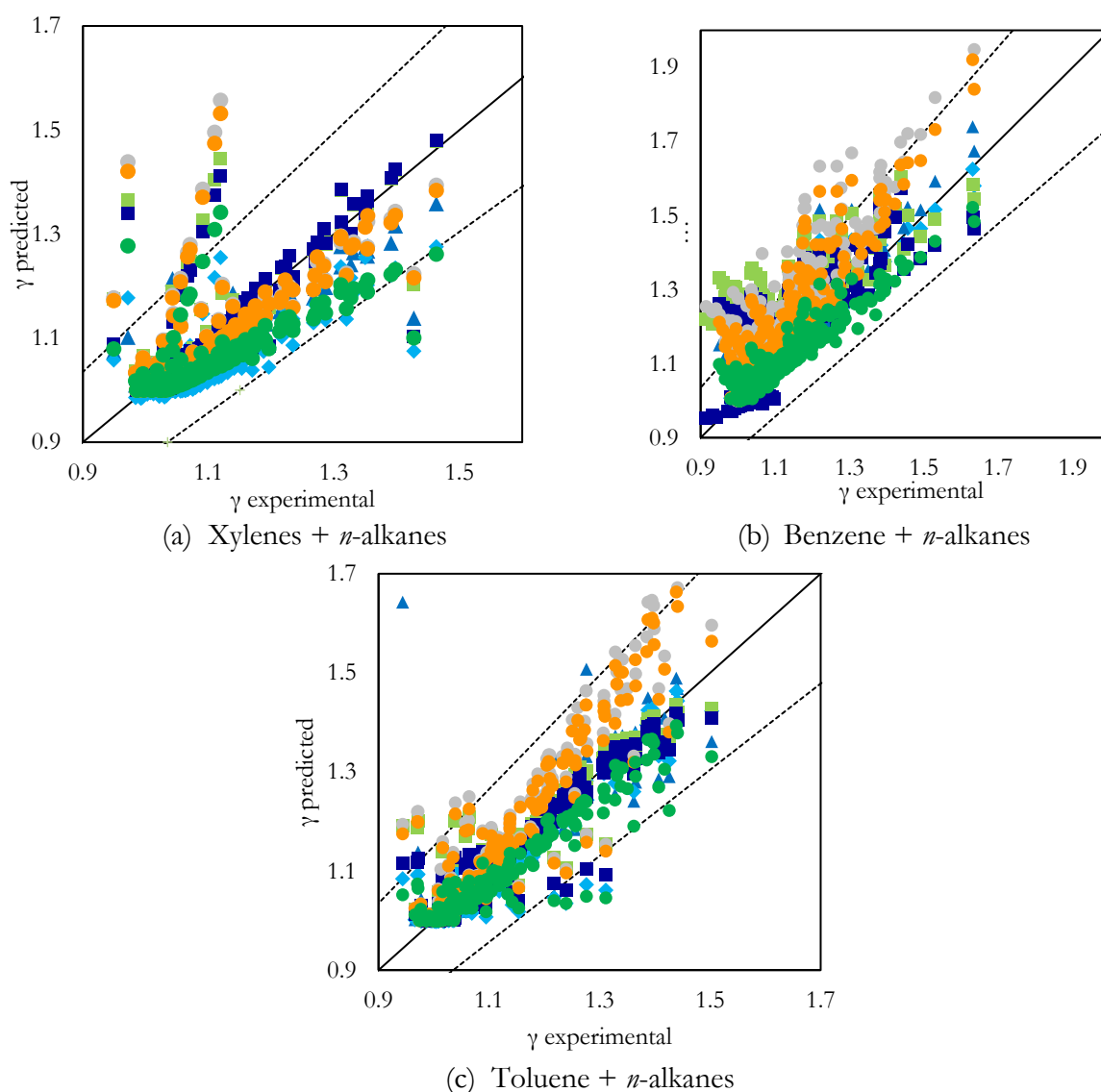


Figure 6.8 – Experimental vs. predicted activity coefficient of aromatics in *n*-alkanes using different models. ●: for COSMO-SAC 2010 using sigma-profiles database (grey: UD; orange: PSL (Klamt radii) and green: PSL (UFF radii)); ▲: for NRTL and ◆: for UNIFAC(Do) and ■: for Scatchard-Hildebrand (green: no combinatorial term; deep blue: with combinatorial term of Staverman-Guggenheim). The dash line represents a margin of error of 15%.

In the evaluation of activity coefficient predictions for several binary liquid-liquid systems, Table 6.8 presents Average Absolute Error percentage (AAE%) values for different thermodynamic models. High errors for all models were found for the system *o*-xylene + *n*-octane. The source of those errors is probably from the accuracy of experimental data, which shows the unexpected fluctuation of the activity coefficient among the molar fractions. This same behavior was observed for the systems xylene + *n*-decane, benzene + *n*-dodecane, and toluene + *n*-decane, which is notorious the impact on the errors found for all models.

In the context of the *o*-xylene + *n*-heptane system, the SH+SG model provided activity coefficient predictions with an error of 2.44%, closely followed by the SH model with an error of 2.34%. The NRTL model demonstrated superior accuracy with an error of 1.31%. UNIFAC Do. reported an error of 3.77%, while COSMO-SAC (PSL-Klamt) and COSMO-SAC (PSL-UFF) exhibited errors of 2.02% and 2.90%, respectively. Lastly, COSMO-SAC (UD) had an error of 1.77%.

For the benzene + *n*-heptane system, the NRTL model outperformed other models, achieving an error of 3.25%. In comparison, the SH+SG and SH models both had errors of 4.14%. COSMO-SAC (PSL-UFF) resulted in an error of 3.77%, UNIFAC had an error of 3.16%, COSMO-SAC (PSL-Klamt) reported an error of 4.63%, and COSMO-SAC (UD) exhibited the highest error at 6.97%.

In the case of toluene + *n*-heptane, the NRTL model once again provided the most accurate activity coefficient predictions with an error of 1.82%. Following closely, COSMO-SAC (PSL-Klamt) reported an error of 4.22%, COSMO-SAC (UD) had an error of 3.86%, SH yielded an error of 1.92%, SH+SG had an error of 1.82%, COSMO-SAC (PSL-UFF) exhibited an error of 2.82%, and UNIFAC Do. showed an error of 2.84%.

These values in Table 6.8 offer critical insights into the relative performance of different thermodynamic models in predicting activity coefficients for specific binary systems.

Table 6.8. AAE% values of activity coefficient for different thermodynamic models.

		Number of points	COSMO- SAC UD	COSMO- SAC PSL Klamt	COSMO- SAC PSL UFF	NRTL	UNIFAC	SH	SH+SG
<i>o</i> -xylene	<i>n</i> -heptane	57	2.68%	2.16%	4.23%	2.22%	6.47%	3.39%	3.44%
	<i>n</i> -octane	13	13.32%	11.17%	7.01%	1.22%	4.03%	9.94%	9.18%
<i>p</i> -xylene	<i>n</i> -heptane	28	1.77%	2.02%	2.90%	1.91%	3.77%	2.34%	2.44%
	<i>n</i> -heptane	27	4.18%	3.11%	4.48%	2.86%	5.24%	3.26%	3.29%
<i>m</i> - xylene	<i>n</i> -decane	33	5.03%	4.76%	6.00%	5.27%	7.61%	4.80%	5.83%
	global	158	5.01%	4.39%	4.78%	3.02%	5.29%	4.38%	4.59%
toluene	<i>n</i> -heptane	192	3.86%	4.22%	2.82%	1.98%	2.84%	1.92%	1.82%
	<i>n</i> -octane	18	5.58%	3.98%	2.00%	1.20%	1.80%	0.82%	1.32%
	<i>n</i> -decane	32	7.72%	7.06%	7.91%	5.93%	8.08%	6.83%	6.45%

	global	242	4.60%	4.27%	2.85%	2.11%	2.88%	2.10%	1.95%
	<i>n</i> -hexane	39	3.90%	3.29%	3.64%	2.07%	2.28%	3.81%	3.31%
	<i>n</i> -heptane	160	6.97%	4.63%	3.77%	3.25%	3.16%	5.13%	4.14%
benzene	<i>n</i> -octane	66	5.39%	4.69%	2.64%	2.26%	1.97%	2.93%	2.27%
	<i>n</i> -dodecane	35	11.09%	14.56%	14.73%	14.79%	14.69%	14.88%	17.68%
	global	300	5.42%	4.20%	3.35%	2.53%	2.47%	3.96%	3.24%

The evolution of the activity coefficient as a function of the composition is presented in figure 6.9 for three representative systems: *o*-xylene + *n*-heptane at 348.13 K, benzene + *n*-heptane at 353.12 K, and toluene + *n*-heptane at 347.98 K.

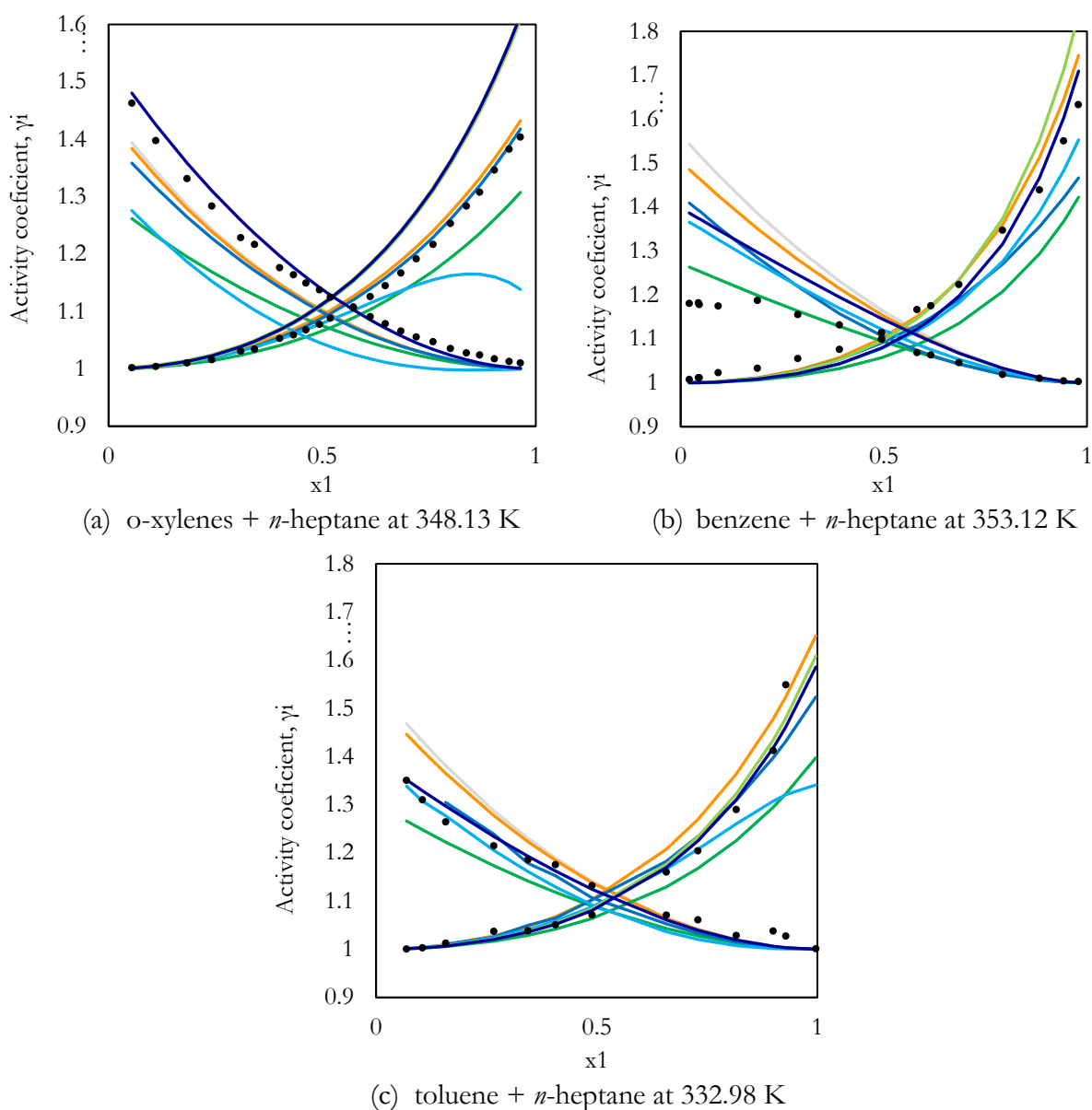


Figure 6.9 – Evolution of activity coefficient in function of the mole fraction for different thermodynamic models: grey lines: COSMO-SAC (UD); orange lines: COSMO-SAC (PSL-Klamt); green lines: COSMO-SAC (PSL-UFF); blue lines: NRTL; light blue lines: UNIFAC(Do); light green lines: SH and (deep blue lines: SH+SG).

●: experimental values.

6.4. Conclusion

This study serves as the foundation for creating the PSL sigma-profile database, initiated by investigating conventional solvents to understand how computational factors impact model performance. The investigation began with a benchmark using furan in an apolar solvent to assess the basis set's impact on the accuracy of predicting IDAC. The choice of basis set and DFT functional was found to significantly affect accuracy. A sophisticated basis set, 6-311G(2df,p), and the PBE0 functional were identified as the most suitable options.

Using these insights, the PSL sigma-profile database was established. This chapter presented predictions of IDAC in conventional solvents compared to the UNIFAC(Do) model. Overall, UNIFAC(Do) outperformed the COSMO-SAC-dps model, especially for most systems. COSMO-SAC-dps demonstrated high accuracy only in predicting IDAC for two specific combinations.

The chapter also explored different sigma-profile databases and models for predicting activity coefficients. UNIFAC(Do) outperformed NRTL, which was found unsuitable for predicting Infinite Dilution Activity Coefficients (IDAC). The Scatchard-Hildebrand model was also considered, and the inclusion of a combinatorial term (Staverman-Guggenheim) improved its predictions for certain systems.

Finally, the COSMO-SAC model was compared using different sigma-profile databases. The PSL-UFF database yielded better results compared to PSL-Klamt and UD for most systems. The chapter provided a comprehensive assessment of the models' performance in predicting IDAC.

Chapter 7

Sigma-profile and COSMO-SAC approach for solvent selection

Profil Sigma et approche COSMO-SAC pour la sélection de solvants

Résumé

Le chapitre commence par aborder la création de la base de données des profils sigma PSL en mettant l'accent sur les DES composés de deux ou trois substances. L'optimisation de la géométrie a été réalisée dans le but d'identifier les structures moléculaires les plus stables pour les systèmes DES, en mettant particulièrement l'accent sur le système BMIMCl : Gly dans un ratio de 2:1. Le choix de la géométrie optimale repose sur les énergies d'interaction, et l'étude souligne le rôle essentiel du choix des méthodes théoriques et des ensembles de base dans ces prédictions. Deux approches distinctes pour la génération de profils sigma sont explorées : les méthodes AB et A+B. Dans l'approche AB, les profils sigma sont créés en prenant en compte la géométrie impliquant l'ensemble des molécules. Dans l'approche A+B, les profils sigma sont dérivés en additionnant les sigmas en fonction du ratio de chaque composé pour former le profil sigma du DES lui-même.

En résumé, l'étude fournit des informations précieuses sur l'importance du choix des méthodes théoriques et des ensembles de base appropriés pour les calculs de mécanique quantique, en particulier lors de la génération de profils sigma pour les DES. Elle met également en évidence l'importance de l'identification de géométries stables pour les structures DES et la distance entre les ions émergeant comme un facteur critique.

Un modèle de solvant implicite appelé le modèle COSMO (également appelé modèle CPCM) a été utilisé pour obtenir une compréhension approfondie de la géométrie moléculaire optimisée dans un solvant. Ce modèle implique le placement de molécules (considérées comme solutés) dans une cavité d'un milieu diélectrique continu qui représente le solvant. En utilisant la forme de la cavité et les rayons atomiques, le modèle calcule des charges ponctuelles à la surface de la cavité. Les informations clés extraites de ce calcul comprennent le volume et la surface de la cavité, la position des noyaux, la densité de charge et la position des éléments de la cavité.

Un aspect crucial de cette approche est la représentation des résultats à l'aide d'une densité de charge de surface représentée par un code couleur. Par exemple, la densité de charge de surface de l'ion 1-butyl-3-méthylimidazolium est représentée avec des régions rouges indiquant des zones fortement négatives (surfaces acceptant l'hydrogène), du bleu profond indiquant des surfaces fortement positives (surfaces donnant de l'hydrogène) et du vert représentant des régions neutres (surfaces non polaires). L'analyse de cette densité de charge de surface est essentielle pour comprendre l'énergie d'interaction calculée dans les sections précédentes et pour sélectionner les positions les plus stables pour les ions.

L'étude explore également le calcul de la densité de charge de surface pour différents conformères et identifie les régions favorables pour l'ion chlorure de 1-butyl-3-méthylimidazolium. Cette analyse contribue à déterminer la position la plus stable pour l'ion chlorure dans le système de solvant.

De plus, l'étude décrit la génération de profils sigma pour les DES. Ce processus implique la normalisation des données à partir du fichier de résultats COSMO et le calcul des distances par paires entre les noyaux. Les profils sigma générés permettent de représenter la polarité moléculaire, et deux approches distinctes, AB et A+B, sont utilisées pour créer des profils sigma pour un DES spécifique BMIMCl : Gly dans un ratio de 2 :1.

Les défis computationnels liés au calcul de la densité de charge de surface sont également abordés et souligne que le calcul de mécanique quantique est l'aspect le plus intensif en calcul du modèle COSMO-SAC.

Enfin, l'étude plonge dans la prédiction des IDAC dans les DES. Cette propriété thermodynamique est cruciale pour les applications industrielles. Les données expérimentales d'IDAC sont utilisées pour optimiser le modèle COSMO-SAC, en mettant principalement l'accent sur la contribution de dispersion basée sur le modèle de Margules à un paramètre. Différentes approches et versions du modèle COSMO-SAC pour prédire l'IDAC dans différents systèmes DES sont utilisées fournissant des références pour les sources de données d'IDAC.

En résumé, l'impact des différentes approches de profil sigma (AB et A+B) sur les prédictions de l'IDAC dans les solvants eutectiques profonds en utilisant le modèle COSMO-SAC est examiné. Le modèle COSMO-SAC traditionnel n'est pas adapté aux prédictions DES en raison des propriétés structurales uniques des DES. Cependant, une version améliorée, COSMO-SAC-dps(DES0), réduit considérablement la déviation relative moyenne (ARD), en particulier pour les alcanes et les alcools. L'approche de profil sigma A+B est préférée pour son rapport coût-efficacité et sa précision modérée. L'optimisation des paramètres pour la contribution combinatoire dans les modèles basés sur COSMO est faite soulignant l'importance de paramètres personnalisés pour différentes familles de solvants. De plus, l'impact de la contribution de dispersion (aDES0) sur les prédictions de l'IDAC dans diverses combinaisons de solvants DES est examiné montrant des niveaux de précision variables (le modèle prédit bien les IDAC pour les aromatiques mais rencontre des défis pour prédire les IDAC pour les alcanes et les alcools). La dépendance à la taille en matière de précision est mise en évidence et l'introduction d'une nouvelle contribution pour résoudre les complexités des interactions des alcanes dans les systèmes DES est suggérée.

Le modèle COSMO-SAC-DES est considéré comme une variante du modèle COSMO-SAC-dps121, spécialement conçu pour prédire les IDAC dans les DESs. Ce modèle intègre les profils sigma A+B obtenus grâce à des calculs de la DFT à partir de la base de données des profils sigma PSL.

Sigma-profile and COSMO-SAC approach for solvent selection

Thermodynamic calculation based on COSMO-based models relies on the sigma-profiles specific to each molecule. A sigma-profile is a molecular-specific distribution of the surface charge density, which is obtained by quantum mechanical (QM) calculation based on solvation models (see sections 3.4.5.2 and 4.3.4 from Chapters 3 and 4, respectively). This information is the heart of the model and is used to calculate and predict phase equilibria and other properties, through a solvation thermodynamic approach. Whereas the QM calculation represents the most time-consuming and computationally expensive part of the COSMO-based models.

The commonly used QM calculation is Density Functional Theory (DFT), with its theoretical method and basis set. The main challenge of this calculation is to find theoretical method and basis set that gives a good balance between desired accuracy – as high as possible - and computational effort. For that, a benchmark of different levels of theoretical methods, and their computational variables have been investigated to understand their role in the performance of the predictions using the COSMO-SAC model. Based on this benchmark, PSL sigma-profile database has been created. Considering DES, which is generally composed of two or three substances, two distinct approaches to obtain the sigma-profile are investigated in this work.

Two approaches to generate the sigma-profile for DES, the geometry optimization in the gas phase is described.

7.1. Geometry optimization

The way how the total energy of a system varies with small changes in its structure is described by its potential energy. The potential energy of a system of two atoms depends only on the distance between them and is described by a 2D graph based on the Lennard-Jones potential model (Figure 6.1). At large distances, the energy is zero, which means that there is no interaction between them. At short distances, attractive forces are dominant, whereas at very close distances repulsive forces appear, causing the energy to rise.

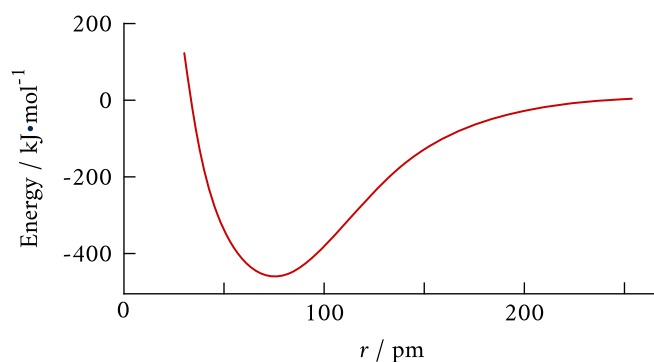


Figure 7.1 – A potential energy curve for the covalent bond in a H₂ molecule. The distance r is the distance between the nuclei of the two H atoms¹⁸⁰.

On the other hand, for large systems, the potential energy has as many dimensions as there are degrees of freedom in the molecule. The plots of the potential energy create a surface that can be used to theoretically explore the properties of structures, such as the potential energy of the ozone molecule represented by Fig 7.2, where each point corresponds to different values for the various bond distances, bond angles, and dihedral angles. The global energy minimum shows the lowest energy point anywhere on the potential surface and represents the equilibrium structure of a molecular system. Geometry optimization usually attempts to locate minima on the potential energy surface, thereby predicting the most stable geometry of a molecular system. In this work, all the geometries have been first optimized using the GAUSSIAN 16 program package. A vibrational analysis was performed on all calculations to ensure the absence of negative vibrational frequencies and verify the existence of a true minimum.

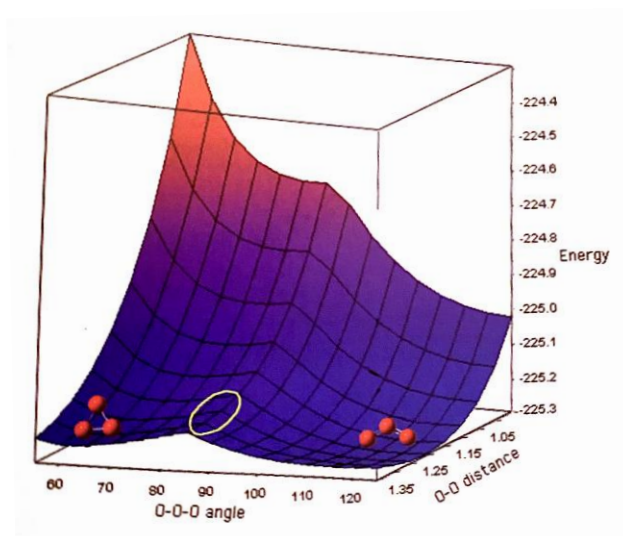


Figure 7.2 – Potential energy surface of the ozone molecule varying the O-O distance (from 1.0 to 1.4 Å) and the O-O-O angle (from 55° to 125°)¹⁸¹

The geometry optimization of the DES, which is generally composed of two or three substances, has been optimized considering a geometry including the ensemble of molecules. In this work, the geometry of the molecules giving the deep eutectic solvent with the chemical name of (BMIMCl: Gly, ration 2:1) have been studied. The structure of the hydrogen-bond acceptor compound, BMIMCl, was considered as an ion-par in a meta geometry. Two conformers (C1 and C2), and five favorable regions for the ion Cl⁻ at the neighborhood of the imidazolium ring have been considered (Fig. 7.3)

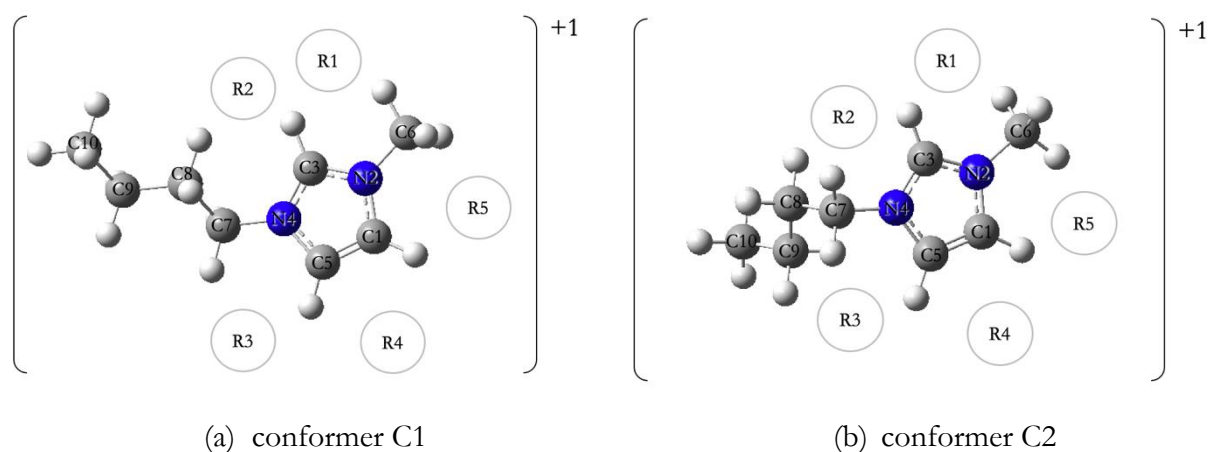


Figure 7.3 – 1-butyl-3-methylimidazolium (BMIM⁺) conformers and their five favorable regions for chloride ion.

Table 7.1 – Optimized structures of 1-butyl-3-methylimidazolium chloride, calculated at PBE0/6-311G(d,f) level.

Favorable regions for Cl ⁻ ion and their ion pairs' name				
1	2	3	4	5
BMIMCl_C1R1	BMIMCl_C1R2	BMIMCl_C1R3	BMIMCl_C1R4	BMIMCl_C1R5
BMIMCl_C2R1	BMIMCl_C2R2	BMIMCl_C2R3	BMIMCl_C2R4	BMIMCl_C2R5

In this study, the total energy has been calculated to find the optimal geometry. The first row of Table 7.1 shows the optimized conformers C1 geometry, and the second row shows the conformers C2.

The interaction energy has been considered to compare the different levels of theoretical methods and basis set.

$$\Delta E = \Delta E_{BMIMCl} - (\Delta E_{BMIM^+} + \Delta E_{Cl^-}) \quad (7.1)$$

The interaction energies are shown in Table 7.2. Large absolute values of the interaction energies indicate a favorable affinity of chloride ion with the cation BMIM⁺, resulting in more stable geometries. Instead, low absolute values of the interaction energies indicate unfavorable interaction of the ions. Generally, the structure where the chloride ion⁻ interacts with the two hydrogen atoms attached to the imidazolium ring (favorable region 4) is the least stable. On the other hand, interactions where the chloride ion interacts with the C3-H (in regions 1 and 2) exhibit relatively lower energies, indicating the most stable geometries.

Furthermore, the different conformations have little effect on the interaction energies of the ion pairs. For example, no difference was found in the energy between the BMIMCl_C1R1 and BMIMCl_C2R1 at the PBE0/6-311G(2df,p) level, and only 6.22 kJ/mol at the PBE0/6-31++G(d,p) level. On the other hand, the position of the chloride ion significantly affects the interaction energies. For example, comparing the energy from regions 1 and 4, for the C1 conformer the difference found is 63.63 kJ/mol, and for C2, the difference is 58.61 kJ/mol. Moreover, when we observe the interactions of the chloride ion in region 1, the hydrogen bonds that it forms is more stable than the one that it forms in region 4. This is due to the position of the group butyl that envelops the ion (stabilizing it). This conclusion is completely in agreement with the results presented in Table 7.2, where the interaction energy of the BMIMCl_C1R1 is much lower than the BMIMCl_C1R4. Based on the results obtained from this study, we can select region 1 as the most active region, resulting in more stable geometries.

Table 7.2 – Interaction energy in kJ/mol.

Ion pairs	PBE0	PBE0	PBE0	B3LYP	B3LYP	B3LYP	B3LYP
BMIMCl	6-31G(d)	6-31++G(d,p)	6-311G(2df,p)	6-31G(d)	6-31++G(d,p)	6-31G(d) ^(a)	6-31++G(d,p) ^(a)
C1R1	-383.34	-376.45	-388.36	-375.25	-368.08	-405.86	-452.44
C1R2	-382.29	-375.79	-388.08	-374.05	-367.93	-405.23	-452.55
C1R3	-352.19	-345.07	-357.63	-344.33	-337.39	-374.65	-421.51
C1R4	-315.72	-345.07	-357.63	-309.83	-337.39	-340.97	-392.54
C1R5	-346.24	-340.83	-352.76	-339.34	-334.29	-370.11	-418.85
C2R1	-383.34	-376.45	-388.36	-375.25	-368.08	-409.45	-450.13
C2R2	-379.15	-379.38	-384.83	-370.24	-364.14	-403.97	-446.85
C2R3	-350.48	-343.50	-355.18	-341.94	-334.80	-374.46	-416.67
C2R4	-350.48	-343.50	-338.57	-307.07	-334.80	-341.51	-388.15
C2R5	-345.30	-339.78	-351.67	-337.54	-332.25	-370.50	-414.50

(a) From Wang et. al. (2005).

According to all results presented in Table 7.2, the predictions of interaction energies are slightly sensitive to the level of the theoretical method used. The average difference between PBE0 and B3LYP levels is about 5.36 kJ/mol. For the same geometry, PBE0 always found the local minimum energy lower than the B3LYP. On the other hand, we can see that the predictions using either PBE0 or B3LYP as a level of the theoretical method are highly sensitive to the basis set used. Globally, the interactions energy predicted at the 6-31G(d) basis set are smaller than that predicted by 6-31++G(d,f) and 6-311++G(d,f) basis set. It is important to emphasize that the basis set 6-311G(df,p), 6-31++G(d,f) and 6-31G(d) has 483, 303, and 199 basis functions to describe the orbitals, respectively. It can be observed that the average difference between 6-31++(d,p) and 6-311G(df,p) using PBE0 as a level of theoretical method is 86.67 kJ/mol. This difference can be explained by the fact that a bigger basis set can envelop better the structure using diffused orbitals, in turn, can describe better hydrogen bonds (e.g., this structure). These results illustrate the importance of the selection of a basis set for the prediction of the interaction energy using DFT and for the study of the ion pairs, for which in turn a large basis set is essential.

The interaction energy predicted in this work using the same level of theoretical method and basis set is bigger than the results obtained by Wang et. al. (2005). However, it was noticed that he found a large difference in changing the basis set. For example, comparing the C1 conformer with the chloride in region 4, he obtained -340.97 kJ/mol for 6-31G(d) basis set and -392.54 kJ/mol for 6-31++G(d,f), arising in a difference of 51.57 kJ/mol in magnitude. On the other hand, repeating the same calculation we obtained a difference of 6.94 kJ/mol (Table 7.3). It is noteworthy that the calculation presented here follows a trend

with a low difference between the 6-31G(d) and 6-31++G(d,f) basis set for most of the positions of the chloride ion, except for position 4 (least stable geometry). The same behavior was not observed in the results of Wang et. al. (2005). Probably they carried out single point calculations at the 6-31G(d) structures.

Table 7.3 – Difference between 6-31G(d) and 6-31++G(d,f) in terms of interaction energy in kJ/mol.

Ion pairs	$\Delta [6-31G(d) - 6-31++G(d,f)]$		
	PBE0	B3LYP	B3LYP ^(a)
BMIMCl			
C1R1	6.89	7.17	-46.58
C1R2	6.5	6.12	-47.32
C1R3	7.12	6.94	-46.86
C1R4	-29.35	-27.56	-51.57
C1R5	5.41	5.05	-48.74
C2R1	6.89	7.17	-40.68
C2R2	-0.23	6.1	-42.88
C2R3	6.98	7.14	-42.21
C2R4	6.98	-27.73	-46.64
C2R5	5.52	5.29	-44

(a) From Wang et. al. (2005).

In summary, the interaction patterns between the chloride ion and imidazolium ring were studied in detail. The large interaction energies between the chloride ion and imidazolium ring suggest that the hydrogen bond is by far the most important contribution to the total energy in DES. Furthermore, results obtained allow us to investigate not just the favorable region of the ion, but the influence of the level of theoretical method and basis set in the predictions of interaction energy.

To conclude, considering all the key factors discussed here, the C1 conformer and the favorable region 1 have been selected as the more representative geometry, which shows relatively lower energies, in turn, gives the most stable geometry.

The next step for geometry optimization considering DES is to include the hydrogen bond donor (HBD) compound. In the case of the BMIMCl:Glycerol (ratio 2:1), first the Glycerol has been optimized by itself.

And then, in a second step, this optimized geometry is inserted into the HBA-optimized one. For that, some probable regions have been considered (Fig. 7.4).

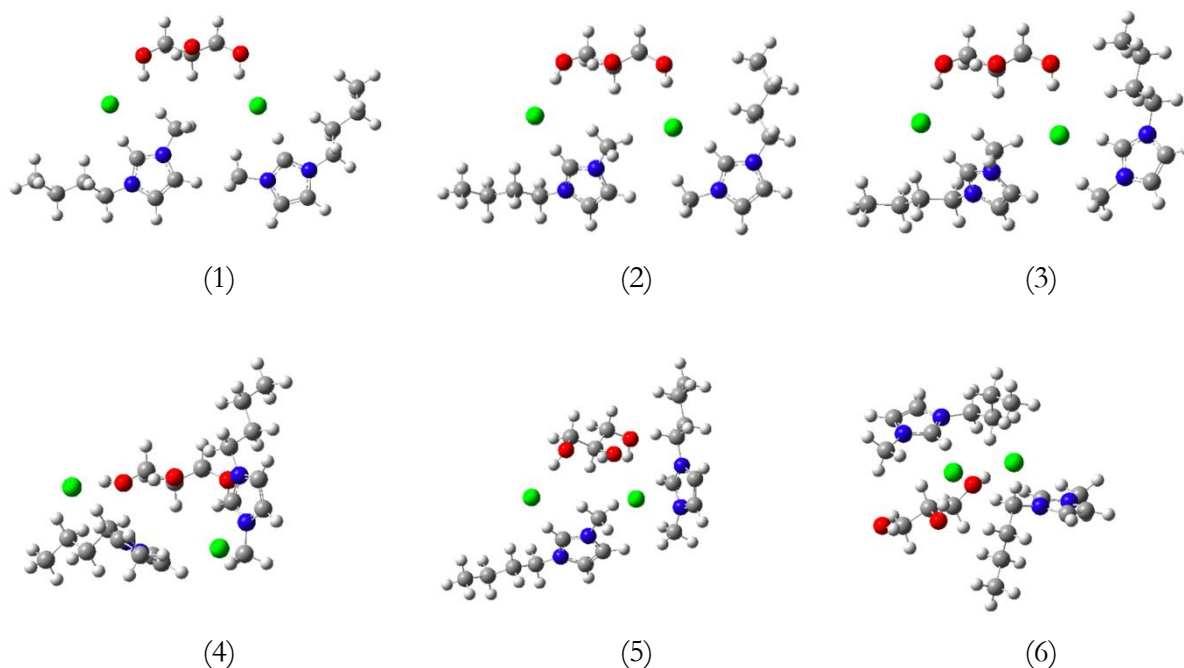


Figure 7.4 – Possible geometries for 1-butyl-3-methylimidazolium chloride:glycerol (ration 2:1) -BMIMCl:Gly.

In geometry 1 and 2, the two ions-pairs BMIMCl are placed in the same plan of the glycerol. In geometry 3, 4 and 5, the molecules are placed in a 3D arrange. And in geometry 6, the molecules are presented in three layers, where the BMIMCl ion-pair is on the first layer, the glycerol on the middle, and the second BMIMCl on the third layer, as a sandwich.

According to results presented in Table 7.4, the bigger is the distance between the ions, the more stable is the structure. Considering, of course, that the electrostatic attraction of the ions with the glycerol and BMIM^+ cations is favorable. In agreement with this point of view, the most stable structure was found considering the bigger distance between the ions, studied here, i.e., geometry 4. Furthermore, it was observed that even with a big distance between the chloride ions, the molecule placed in the same plan is less stable than the one in a 3D arrangement. Considering geometry 6, the total energy of the structure is bigger than the sum of the total energy of each ion. This result shows the second less stable geometry, giving a positive interaction energy. This means that geometry 6 requires more energy to be in that structured shape than separated as ions. The same behavior was observed for geometry 5, however, because this geometry shows smaller distances between the chloride ions.

Table 7.4 – Interaction energy in kJ/mol.

Geometry	Chloride ions distance (angstroms)	Interaction energy (kJ/mol)
1	7.488	-720.18
2	6.728	-915.51
3	6.275	-913.19
4	7.495	-937.42
5	6.309	248.24
6	6.540	3.28

It is important to point out that, optimizing this type of metastructure is computationally expensive. To solve the calculation for geometry 4, it takes 9 days 1 hour, and 27 minutes. In contrast, to solve the calculation for the BMIMCL_C1R1, it takes 5 hours and 21 minutes, and for glycerol, it takes 33 minutes.

7.2. Surface charge density

Now that, the geometry is optimized in the gas phase, it's time to examine its behavior on the solvent. In this step, an implicit solvent model is applied to obtain the surface charge density, the COSMO model, also known as the CPCM model. As explained in Chapter 4, the solute molecules are placed into a cavity with a continuous dielectric medium representing the solvent. The shape of the cavity is constructed based on the cavity algorithm and the atom radii, set on the calculation. Then, the cavity surface is split into small elements called *tesserae*, where the point charges are calculated. Each one represents the electrostatic polarization of the solute electron density. This calculation generates the COSMO result file, and the important information that can be extracted from there are:

- Volume and surface area of the cavity.
- Position of all nuclei.
- Charge density, area, and location of each element that forms the cavity.

The surface charge densities can be represented in a color-coded surface, e.g., surface charge density of the 1-butyl-3-methylimidazolium ion studied here (Fig 7.5). The areas colored in red represent the strongly negative parts of the ion, and hence hydrogen-accepting surfaces. On the other hand, areas colored in deep blue represent strongly positive surface regions and hydrogen-donating surfaces. While areas colored green represent the neutral regions, i.e., nonpolar surfaces. Considering the favorable regions for the chloride ion suggested in this work, the surface charge density of the BMIM⁺ presented in Fig. 7.5 cannot give a conclusive clue about the more stable position of the chloride ion. In contrast, the interaction energy

calculated for each region allow us to select the most stable position of the chloride ion. The area around the C-H bonded to the two nitrogen of the ring represents the more negative surface of the ion. Around light blue areas we can find the hydrogens located adjacent to the nitrogen, and they give positive surfaces, hence negative values of surface charge density. These regions are hydrogen-donating surfaces. The hydrogens found close to these light blue areas and nearby the area colored in red (R1 and R2) can do stronger bonds with the chloride ion than the hydrogens nearby the area colored in deep blue (R3 and R5). Around R4, a neutral surface is found and it is expected that hydrogen bonds there be quite weak, resulting in an unfavorable region to find the chloride ion.

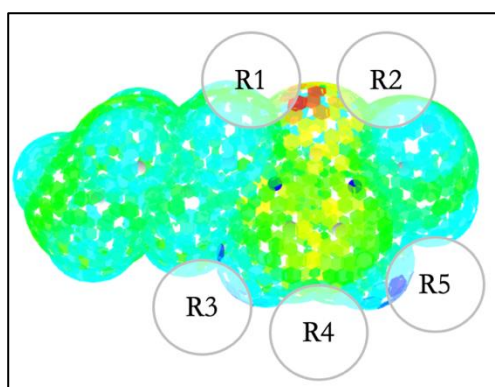


Figure 7.5 – Surface charge of the [1-butyl-3-methylimidazolium]⁺¹ ion.

Analyzing the surface charge density of the two conformers and their five favorable regions of the 1-butyl-3-methylimidazolium chloride, we can observe the behavior of the surface charge density when we include the chloride ion and change its position (Figure 7.6).

In summary, the surface charge density can give us a local qualitative measurement of the molecular polarity and allow us to understand the interaction energy calculated in the previous sections, arising on the same conclusion that the region R1 and R2 are more favorable to finding the chloride ion.

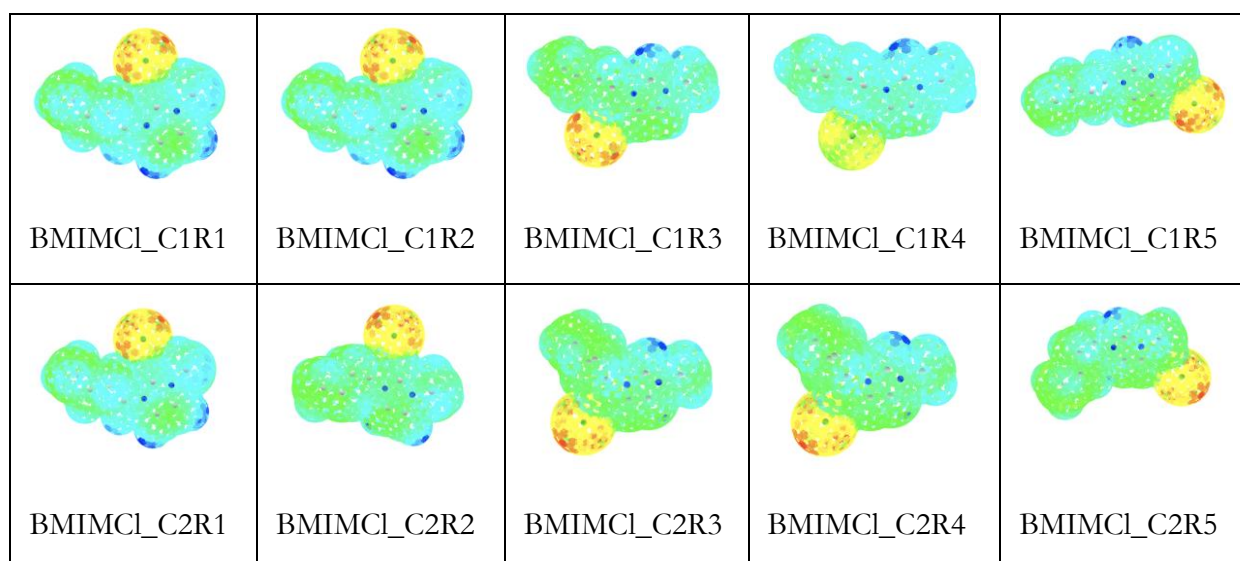


Figure 7.6 – Surface charge of the two conformers of 1-butyl-3-methylimidazolium and the five positions of the chloride ion.

The surface charge density of the DES has been calculated considering the same structure of the BMIMCl: Gly (ratio 2:1) studied on the gas phase and presented in the previous section. The color-coded surface charge density (Figure 7.7) shows more green regions than when the BMIMCl ion is alone (Figure 7.6), due to the neutral regions formed when all the ions were considered and interacts with each other. Strong negative parts, and hence positive surface charge density are observed nearby the chloride ions.

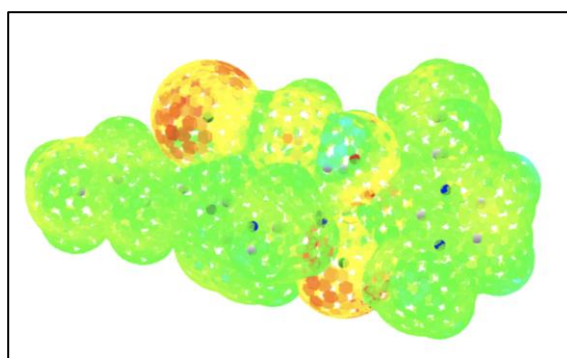


Figure 7.7 – Surface charge of the BMIMCl: Gly (ratio 2:1) (geometry 4).

In general, the surface charge density calculated using CPCM model converges faster than using IEFPCM model. The more complex structure, geometry 4 of the BMIMCl: Gly (ratio 2:1) converged in 2 hours and 40 minutes (in cpu time). In contrast, to solve the calculation for the BMIMCl_C1R1 and for glycerol, it takes 45 minutes, and 14 minutes, respectively. Globally, the quantum mechanical calculation is so far the more computationally expensive part of the COSMO-SAC model. At least, these calculations just need to be done once.

As the COSMO result file is a text file of non-standardized format containing all the results of this calculation, a normalization should be done, resulting in the sigma-profile that can be stored in a database.

However, building a sigma-profile database of DES considering a meta-structure as presented before will be extensive and computationally expensive. This is due to the complexity of the structure of the DES and the many combinations that can be proposed (changing the HBD:HBA ratio, the HBA molecule, and even the family of the HBD). To overcome this issue, an approach to obtain the sigma-profile of DES has been proposed and will be present in the next section.

7.3. Approaches to generate the sigma-profile of DES

To normalize the data from the COSMO result file to build the sigma-profile, first the position of all nuclei has been converted to the Ångstrom scale by multiplying with the conversion factor $0.52917721067 \text{ \AA}$ (Bohr radius). Then, the pairwise distance between each pair of nuclei m and n , in Å, is calculated using Eq. (3.39) cited in chapter 3, and the screening charge is averaged by Eq. (3.38). Once the screening charge averaged has been obtained for each segment m , the sigma-profile, $p_i(\sigma)$, is generated considering the probability of finding a given segment with a specified value of σ multiplied by the surface area of the molecule i . Generally, the values of σ vary in a range of -0.025 e/\AA^2 to 0.025 e/\AA^2 , considering increments of 0.001 e/\AA^2 , the sigma-profile is formed in a set of 51 points. Two approaches have been studied to generate the sigma-profile of deep eutectic solvents: AB and A+B (Fig. 7.8).

The sigma-profile of the DES: BMIMCl: Gly (ratio 2:1) following the AB approach is presented in Fig 7.8a and has been generated considering the geometry including the ensemble of molecules. Where the black line represents the non-hydrogen bond profile, and the grey is the hydrogen bond profile.

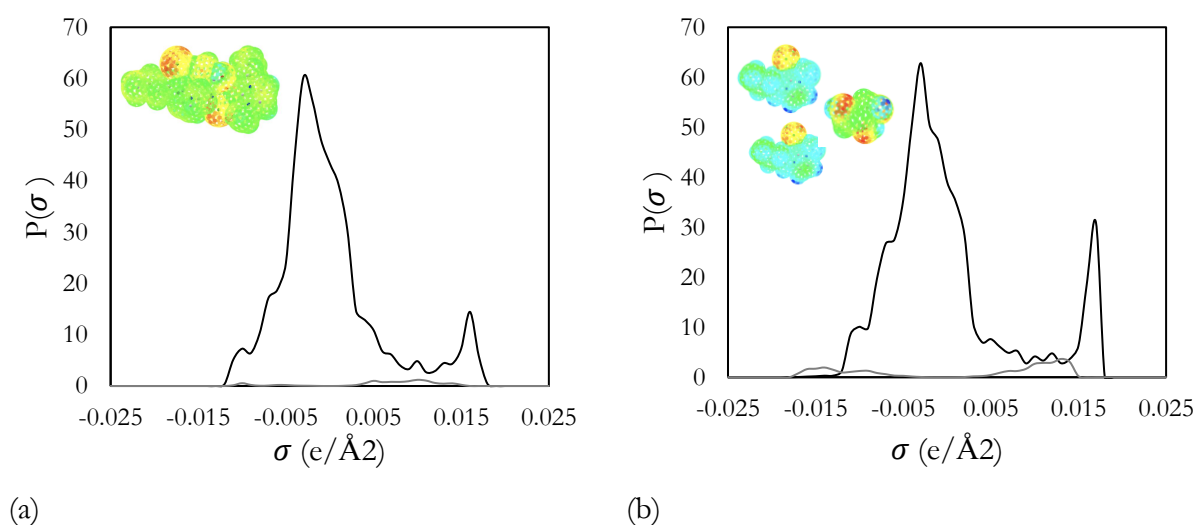


Figure 7.8 – Sigma-profile of the BMIMCl: Gly (ratio 2:1). (a) AB approach and (b) A+B approach. Black lines: non-hydrogen bond profile, and grey lines: hydrogen bond profile.

The sigma-profile of 1-butyl-3-methylimidazolium chloride and glycerol have been generated separately. Then, following the A +B approach, a sum of the sigmas in a ratio of two BMIMCl to one glycerol has been considered to generate the sigma-profile of the DES itself: BMIMCl: Gly (ratio 2:1) presented in Fig 7.8b. Where the black line represents the non-hydrogen bond profile, and the grey is the hydrogen bond profile.

In the two profiles, we can observe the same big central pick with two small shoulders on the left side, which represents the nonpolar part of the molecule. On the other hand, the pick on the right side is higher in the A+B profile. These picks represent the hydrogen bond due to the three hydroxyl groups of the glycerol. This difference was expected because, in the case of AB profile, part of the electric charge of hydroxyl groups was neutralized inside the molecule. The same behavior was found for the hydrogen bond profile in grey, which is more pronounced in the A+B profile, for the same reason explained here.

The cavity volume and surface area are bigger in the case of the A+B profile than the AB profile, in which a part of the surface of the BMIMCl and the glycerol is hidden inside of the cavity created around the meta structure, as observed in Tab 7.5.

Table 7.5 – Cavity volume and surface area of BMIMCl:Cly (2:1).

BMIMCl:Gly	Cavity volume (\AA^3)	Surface area(\AA^2)
A+B (2:1)	547.93	609.29
AB (2:1)	535.06	528.43

7.4. Prediction of infinite dilution activity coefficient in DES

IDAC are important thermodynamic properties and are commonly used for pre-screening solvents for industrial applications. IDAC measurements are both difficult and expensive. Nevertheless, activity coefficient data are helpful for testing the applicability and development of the theoretical models describing interactions on a physical basis.

Few papers have reported experimental data of IDAC comprising DES. In this work, these data have been used to optimize the COSMO-SAC model including the parameter A_{DES0} to the dispersion contribution of the COSMO-SAC model based on the one-parameter Margules model.

$$\frac{G^{E,dsp}}{RT} = A_{DES0}x_1x_2 \quad (7.2)$$

And

$$\ln \gamma_1^{dps} = A_{DES0}x_2^2 \quad (7.3)$$

$$\ln \gamma_2^{dps} = A_{DES0}x_1^2 \quad (7.4)$$

With

$$A_{DES0} = A \cdot a_{DES0} \quad (7.5)$$

Which A is original dispersive term from COSMO-SAC-dps model, and is expressed by:

$$A = \varepsilon_r \cdot w \cdot \left[\frac{1}{2}(\varepsilon_1 + \varepsilon_2) - \sqrt{\varepsilon_1 \varepsilon_2} \right] \quad (7.6)$$

The average relative deviations (ARDs) minimized as objective functions (OFs) were used to obtain the adjustable parameters a_{DES0} of dispersion contribution.

$$OF = \min \left\{ \frac{1}{N} \sum_i^N \left| \frac{\gamma_i^{\infty,exp} - \gamma_i^{\infty,cal}}{\gamma_i^{\infty,exp}} \right| \right\} \quad (7.7)$$

Two approaches to obtain the sigma-profile and different versions of COSMO-SAC cite have been used to predict IDAC in different DESs, listed in Tab. 7.6.

Table 7.6 – Bibliographic references of IDAC data in DES.

	HBA	HBD	ratio	Abbv.	Ref.
DES1	1-butyl-3-methylimidazolium chloride	glycerol	(2:1)	BMIMCl:Gly	182
DES2	1-butyl-2,3-dimethylimidazolium chloride	ethylene glycol	(1:3)	BDMIMCl:EG	183
DES3	Choline chloride	glycerol	(1:1)	ChCh:Gly_1	184
DES4	Choline chloride	glycerol	(1:2)	ChCl:Gly_2	184
DES5	Tetramethylammonium chloride	1,6 hexanediol	(1:1)	TMACl:C6ol	185
DES6	Tetramethylammonium chloride	ethylene glycol	(1:2)	TMACl:EG	186

7.4.1. Investigation of sigma-profiles on COSMO-SAC predictions: AB versus A+B approaches

Predictions of IDAC in BMIMCl: Gly (2:1), hereafter referred to as DES1, have been used as reference to investigation of the different approaches to obtain the sigma-profile and their impact in the accuracy of the COSMO-SAC based model. The table 7.7 presents the ARD of the predictions using COSMO-SAC-dps(DES0) models. COSMO-SAC-dps(DES0) represents the COSMO-SAC including the Margules term described in the equation 7.3 and its parameter a_{DES0} .

Table 7.7 – ARD of IDAC of conventional solute in [BMIMCl + Glycerol] and its parameter a_{DES0} .

		COSMO-SAC-dps(DES0)	
		ARD	a_{DES0}
AB	Alkanes	57.3%	18.05
	Cyclo	20.4%	20.52
	Aromatics	37.7%	18.30
	Alcohols	7.7%	-2.83
A+B	Alkanes	49.1%	15.45
	Cyclo	20.4%	16.91
	Aromatics	25.3%	16.63
	Alcohols	46.1%	12.51

COSMO-SAC-based model is not particularly suitable for calculations with DES. When the COSMO-SAC model is applied to predict IDAC for the solute (alkanes, cycloalkanes, aromatics and alcohol) in DES systems (such as butylmethylimidazolium-based), the ARD between the prediction and experimental data is high as 99%. This indicates that if the parameters in COSMO-SAC model obtained from conventional compound systems are directly extended to predict the thermodynamic properties of DES systems, large deviations between experiments and predictions often arise. A significant contributing factor to the substantial disparities in the predictions considering DES can be attributed to the formation of nano-segregated structures in DES (polar networks interspersed with non-polar domains), already observed in ionic liquids¹⁸⁷, which is completely different from the solvation properties of molecular solutes in conventional solvents at the microscopic scale. Parametrization of the model based on experimental data of DES is essential to improve the predictions and will be better discussed in the next section.

Nonetheless, a significant improvement in the predictions using the COSMO-SAC-dps(DES0) is observed, in which the ARD between the prediction and experimental data decreased to 57% in the worst

case (alkanes), and surprisingly 8% to alcohols, in the better case. These predictions have been used to evaluate the sigma-profiles approaches (AB and A+B) and is presented in detail on the Fig. 7.9, 7.10, and 7.11.

The Fig. 7.9 compares the predicted and experimental IDAC of *n*-alkanes and *n*-alkenes in DES1 using COSMO-SAC-dps(DES0) model using two approaches to obtain the sigma-profile (AB and A+B). The Fig. 7.10 shows the results for cyclo and aromatics hydrocarbons. In all these cases, the COSMO-SAC-dps(DES0) model overpredicts the IDAC, and A+B approach shows better results. The only exception is observed in the Fig. 7.11, in which represent the results for alcohols. This deviation is explained by the fact that with the AB sigma-profile, certain intermolecular interactions are hidden within the cavity, unlike the A+B sigma-profile where all interaction sites are explicitly considered.

Based on the results presented in this section and balancing the computer time consumption of the AB and A+B sigma-profiles and their accuracy, the A+B sigma-profile was chosen as less expensive and giving a moderate accuracy.

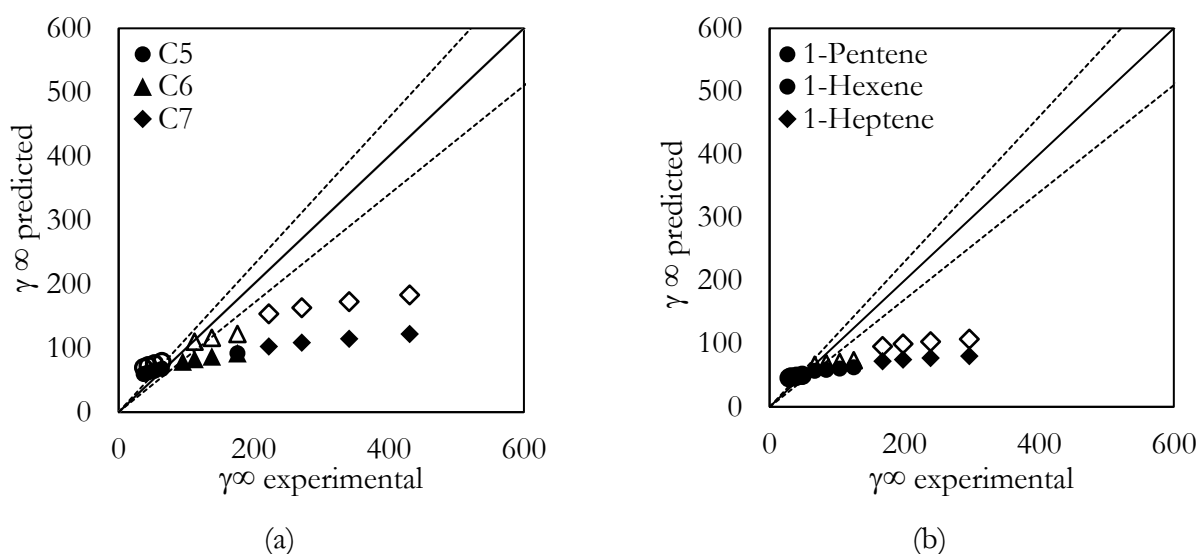


Figure 7.9 – Experimental *vs.* predicted IDAC of (a) *n*-alkanes and (b) *n*-alkenes in BMIMCl: Gly (ratio 2:1). Solid symbols represent results using sigma-profile AB, and hollow symbols sigma-profile A+B. Dash lines represent 15% of the error from bisector.

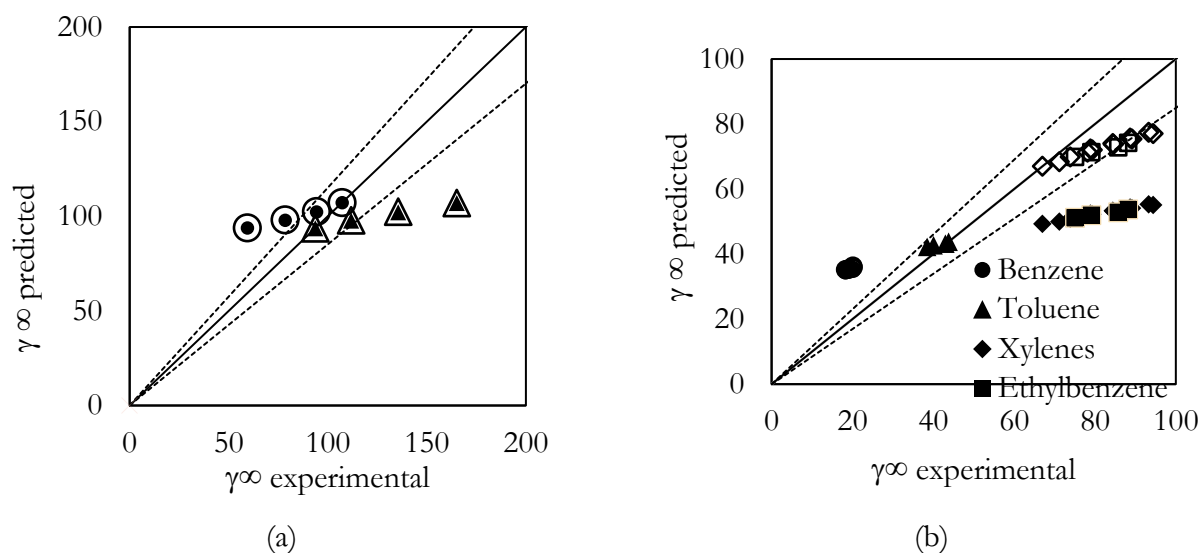


Figure 7.10 – Experimental *vs.* predicted IDAC of (a) cyclo-hydrocarbons and (b) aromatics hydrocarbons in BMIMCl: Gly (ratio 2:1). Solid symbols represent results using sigma-profile AB, and hollow symbols sigma-profile A+B. Dash lines represent 15% of the error from bisector.

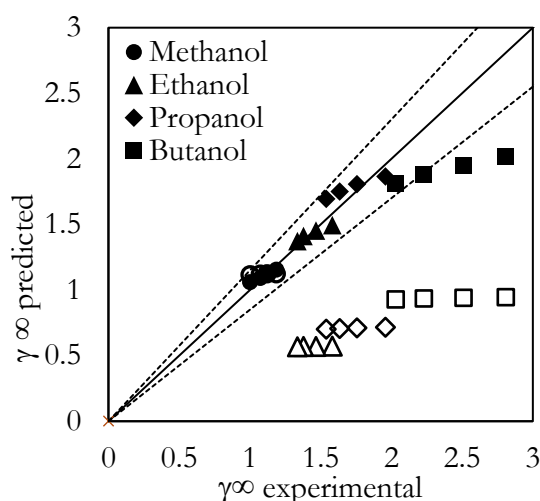


Figure 7.11 – Experimental *vs.* predicted IDAC of alcohols in BMIMCl: Gly (ratio 2:1). Solid symbols represent results using sigma-profile AB, and hollow symbols sigma-profile A+B. Dash lines represent 15% of the error from bisector.

The total computational cost to generate the A+B sigma-profile takes about 6 hours in contrast to 9 days to obtain the AB sigma-profile. Furthermore, using the A+B sigma-profile can give the user the flexibility to combine different HBD and HBA and easily change the composition ratio. This A+B approach has been used in this work to build the PSL sigma-profile and all the further calculations presented in this work were based on it.

7.4.2. Investigation of combinatorial contribution

In COSMO-based models, the activity coefficient is given by the product of two contributions (combinatorial and residual) and can also include a third one (dispersive) (see Chapter 3). It is widely recognized that the residual contribution can be orders of magnitude larger than the combinatorial contribution. Since the model is based on the two contributions, poor performance in the prediction of the combinatorial term can directly affect the accuracy of the model.

The combinatorial contribution should consider differences in size, shape, and free volume. Once the COSMO quantum mechanical step calculation generates the cavity comprising the molecule, the size, and the shape of the molecule can be obtained without any additional calculation.

In the COSMO-RS proposed by Klamt *et al.* the combinatorial contribution is represented by the simple expression (Eq.7.8).

$$\ln \gamma_{i/S}^{comb} = \lambda \ln \frac{A_s}{A_i} \quad (7.8)$$

where λ is a parameter estimated by the authors as 0.14, A_s is the mean cavity surface area of all components, and A_i is the cavity surface area of component i .

In the COSMO-SAC 2002 proposed by Lin and Sandler, and in several COSMO-SAC implementations^{123,188,189}, the combinatorial term is given by the Staverman-Guggenheim^{121,122} (SG) equation (Eq.7.9).

$$\ln \gamma_i^{comb} = \ln \Phi_i + 1 - \Phi_i - \frac{z}{2} q_i \left(\ln \frac{\Phi_i}{\theta_i} + 1 - \frac{\Phi_i}{\theta_i} \right) \quad (7.9)$$

The molecular volume parameter is given by $\Phi_i = r_i x_i / \sum_j r_j x_j$, the surface area fraction is given by $\theta_i = q_i x_i / \sum_j q_j x_j$. Where x_i is the molar fraction of component i ; z is the coordination number, usually taken to be 10; $r_i = V_i/r$ and $q_i = A_i/q$ are the normalized volume and surface area, respectively; A_i is the cavity surface area and V_i is the cavity volume; q and r are universal parameters of the model.

Analyzing the equations above, it is clear that Φ_i and θ_i are independent of the parameters q and r . Moreover, the SG equation is independent of r . Assuming the coordination number equal to 10, the only adjustable parameter of the model is q .

A modified SG equation has been proposed by Kikic et al.¹⁹⁰ based on the research of Donohue and Prausnitz¹⁹¹. According to Zhong¹⁹², this is the one of the most important improvements in this equation. Specifically, Kikic et al.¹⁹⁰ modified the Flory-Huggins part of the SG combinatorial equation by replacing the volume fraction Φ_i' with the Eq.7.10.

$$\Phi_i' = \frac{r_i^{p_i} x_i}{\sum_j x_j r_j^{p_j}} \quad (7.10)$$

Where p_i is a parameter and has the value between 0 and 1.

Thus, SG expression modified, which is used in modified UNIFAC implementations with a single value for $p_i = p$, is given by (Eq.7.11).

$$\ln \gamma_i^{comb} = \ln \Phi_i' + 1 - \Phi_i' - \frac{z}{2} q_i \left(\ln \frac{\Phi_i}{\theta_i} + 1 - \frac{\Phi_i}{\theta_i} \right) \quad (7.11)$$

It is well-known that the combination of SQ expression modified with the residual part of group contribution methods, such as UNIFAC, resulted in much improved predictions of activity coefficient in mixtures containing saturated hydrocarbons. In this work the performance of modified SQ equation is investigated. A component-independent exponent $p\theta$ is considered, likewise adopted in modified UNIFAC. Thus, there are only two parameters to be optimized: $p\theta$ and $q\theta$.

The experimental data group consisting of 358 data points of IDAC of conventional solutes in DES was taken on as the training set to optimize the two adjustable parameters in the combinatorial term.

Table 7.8 presents the parameters optimized based on the experimental data of the 6 types of conventional solutes studied in this work.

Table 7.8 – ARD of IDAC of conventional solute in DES and its parameter p_0 and q_0 .

	n points	p_0	$q_0/\text{\AA}^2$	ARD
Alkanes	130	1.0000	79.53	77.1%
		0.6370	124.00	74.0%
		0.2774	79.53	72.5%
		0.2331	23.87	72.3%
Alcohols	85	1.0000	79.53	73.5%
		0.6370	124.00	93.1%
		1.3330	79.53	63.7%
		1.6216	4.54801	57.2%
Aromatics	82	1.0000	79.53	78.3%
		0.6370	124.00	70.5%
		0.4228	79.53	69.1%
		0.7004	1.692	55.2%
Cyclo	30	1.0000	79.53	86.7%
		0.6370	124.00	114.2%
		1.4440	79.53	72.8%
		1.4664	1.617	72.0%
Ketones	25	1.0000	79.53	138.8%
		0.6370	124.00	196.1%
		1.8374	79.53	83.0%
		1.8321	1428.98	80.3%
all data	358	1.0000	79.53	81.0%
		0.6370	124.00	90.2%
		1.2685	79.53	78.9%
		1.2982	32.39	78.9%

Accordingly with the results presented in the Table 7.8, it can be concluded that the performance of the COSMO-SAC model can be improved by changing the parameters p_0 and q_0 of the combinatorial contribution. In these calculations, the residual and dispersive contributions have been considered as the one implemented in COSMO-SAC-dps. The results presented in the first line of Table 7.8 ($p_0=1$ and $q_0=79.53 \text{ \AA}^2$) are the values suggested by Wang and Sandler (2007)¹⁹³, and implemented in most COSMO-SAC models. The results presented in the second line ($p_0=0.637$ and $q_0=124.00 \text{ \AA}^2$) are the values optimized by Soares (2011)¹⁹⁴ based on experimental data for the logarithm of infinite dilution activity

coefficients of linear and branched alkanes ranging from C_4 to C_{36} . As expected, the ARD of IDAC for alkanes as solutes indicates an improvement in model accuracy when utilizing the values recommended by Soares (2011). In contrast, the ARD of IDAC of alcohols, cyclo-alkanes, and ketones as solutes were higher when Soares (2011) parameters were employed. Basically, because these parameters were not optimized considering these classes of solvent. Comparing these two predictions and considering that the second one was optimized considering only one class of solvent, it is evident the dependence of these parameters on the accuracy of the model. Furthermore, the data set used to optimize the parameters cannot be neglected and the results suggest that personalized parameters for each solvent family improve much more the accuracy of the model.

Comparing the optimization of the combinatorial term done in this work, shown in the third and fourth lines of tab. 7.8, superior results were achieved when p_0 and q_0 were optimized simultaneously. In contrast, the results for a scenario where p_0 was optimized while maintaining a fixed q_0 value, similar to the approach used in the COSMO-SAC 2010 implementation, are presented. The fourth row of the table 7.8 indicates that IDAC predictions are more accurate compared to those based on Soares's parameters but are on par or slightly less accurate than when p_0 and q_0 are optimized simultaneously.

7.4.3. Investigation of dispersive contribution

In this section, an investigation into the dispersive term, based on equation 7.3 and considering the a_{DES0} parameter is conducted to better understand its role in predicting infinite dilution activity coefficients (IDAC) within deep eutectic solvents (DES) using the COSMO-SAC model. Table 7.9 provides a comprehensive overview of the average relative deviation (ARD) for various DES-solute combinations, offering valuable insights into the accuracy and reliability of the COSMO-SAC model's predictions in comparison to experimental data. Each a_{DES0} value is specifically tailored to interactions between a particular DES and a solvent family, reflecting the unique strengths and natures of dispersion forces involved in different combinations. The ARD values, ranging from approximately 3% to over 98%, highlight the varying predictive capabilities of the model for different systems. Lower ARD values signify a closer alignment between model predictions and experimental data, emphasizing the critical role of precise a_{DES0} determination to enhance the accuracy of predicting IDAC. This analysis emphasizes the necessity for a comprehensive understanding of dispersion contribution and its role in accurately modeling the behavior of DES.

Table 7.9 – ARD of IDAC of conventional solutes in different DESs and its parameter a_{DES0} .

	n			n			n		
	points	a_{DES0}	ARD	points	a_{DES0}	ARD	points	a_{DES0}	ARD
	Alkanes			Alcohols			Aromatics		
BMIMCl:Gly	20	11.36	59.0%	16	4.950	56.5%	24	14.184	24.3%
BDMIMCl:EG	50	1.208	48.3%	19	-19.07	26.7%	30	0.688	27.4%
	-								
ChCl:Gly	6	0.6121	45.9%	10	26.01	57.3%	4	2.535	27.6%
ChCl	3	10.25	32.9%	5	5.552	70.9%	2	14.891	12.3%
TMAC:C6diol	16	12.56	13.8%	14	1.237	45.4%	12	7.452	11.6%
TMAC:EG	16	11.08	48.8%	16	8.566	60.8%	8	1.000	11.7%
	Cyclo			Ketones			all data		
BMIMCl:Gly	8	21.123	20.5%	4	255.322	3.0%	88	15.7994	48.0%
	-								
BDMIMCl:EG	10	0.072	34.0%	5	357.385	22.4%	114	2.94395	57.8%
ChCl:Gly	2	4.704	27.9%	-	-	-	22	18.4326	44.2%
ChCl	-	-	-	-	-	-	10	7.11719	65.4%
TMAC:C6diol	4	13.769	7.6%	8	85.223	47.5%	54	-0.8607	75.6%
TMAC:EG	4	8.145	1.3%	8	515.536	38.2%	52	7.23418	63.5%

The ARD of IDAC for alkanes range from 13.8% to 59.0% across various DES. For instance, interactions with BMIMCl:Gly, BDMIMCl:EG and TMAC:EG show relatively high average ARD (59.0%, 48.3% and 48.8%, respectively) show up challenges in accurately predicting alkane behavior. Conversely, the TMAC:C6diol exhibits a lower ARD (13.8%).

IDAC values for the specified DES exhibit distinct magnitude range: 60 to 1200 for BMIMCl:Gly, 6 to 440 for BDMIMCl:EG, 300 to 1400 to ChCl:Gly, 200 to 1800 for TMAC:EG and 50 to 130 for TMAC:C6diol. The observed trend of increasing IDAC values with the size of the alkane molecule is consistent with expectations, considering larger molecules tend to have weaker interactions within a solvent. However, the model underestimates the high IDAC values for alkanes. This discrepancy could be due the model not fully capturing the intricate molecular interactions leading to higher IDAC values, potentially related to dispersion forces or other molecular complexities present in larger alkane molecules. Fig. 7.12 presents a comparative analysis between the experimental versus predicted IDAC using the COSMO-SAC(DES0) model, incorporating the optimized dispersion contribution discussed in this

section, and the COSMO-SAC-dps p0 q0 model with the combinatorial contribution optimized in the preceding section (7.4.2 of this chapter) .

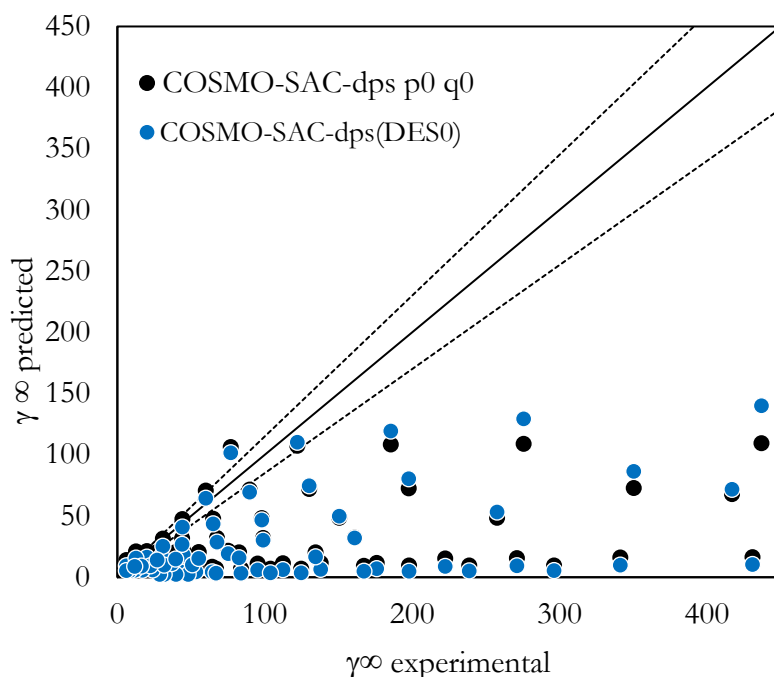


Figure 7.12 – Experimental *vs.* predicted IDAC of alkanes in DES. Black symbols represent the results of COSMO-SAC-dps with combinatorial term optimized, and blue symbols COSMO-SAC-dps(DES0). Dash lines represent 15% of the error from bisector.

The ARD of IDAC for alcohols range from 26.7% to 70.9% across various DES. For instance, interactions with ChCl and TMAC:EG show relatively high average relative deviation (ARD) percentages (70.9% and 60.8%, respectively) underscore challenges in accurately predicting alcohols behavior. Conversely, the BDMIMCl:EG exhibits a lower ARD (26.7%). Despite the majority of IDAC values for alcohols falling within the range of 0.1 to 3 (as illustrated in Fig. 7.13), the COSMO-SAC-dps(DES0) model demonstrates limitations in accurately predicting these values.

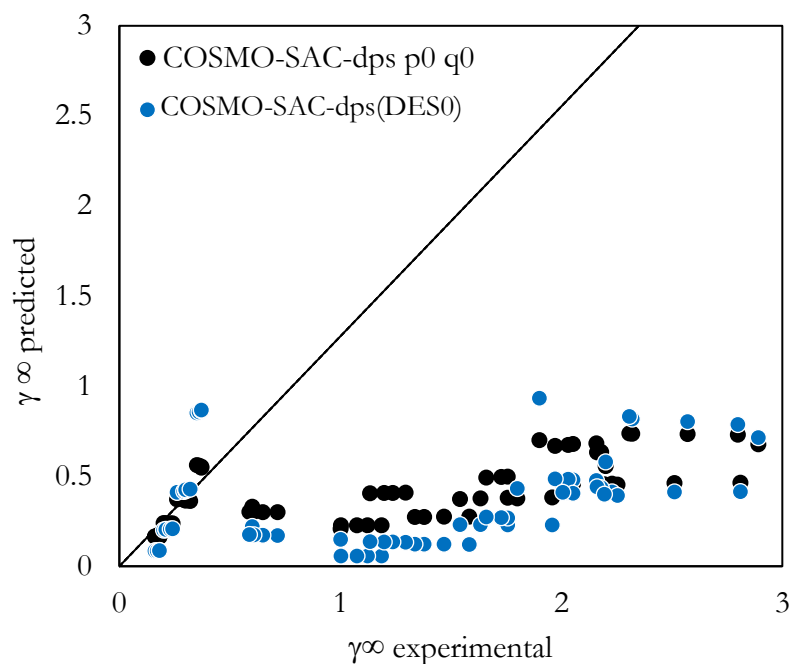


Figure 7.13 – Experimental *vs.* predicted IDAC of alcohols in DES. Black symbols represent the results of COSMO-SAC-dps with combinatorial term optimized, and blue symbols COSMO-SAC-dps(DES0). Dash lines represent 15% of the error from bisector.

Among the various solvent families, the COSMO-SAC model demonstrated its best predictive performance for the aromatic. This was evident from the ARD values, ranging from 11.60% to 27.60%. These results indicate that the COSMO-SAC model, with the optimized dispersive contribution with the a_{DES0} parameter, was more accurate in predicting IDAC for aromatics. In general, the COSMO-SAC model was recognized in the chapter 3 and 4 for its relatively good accuracy in predicting thermodynamic properties when applied to aromatic solvents. Fig. 7.14 illustrates a comparison between experimental and predicted IDAC for aromatic solutes in DES.

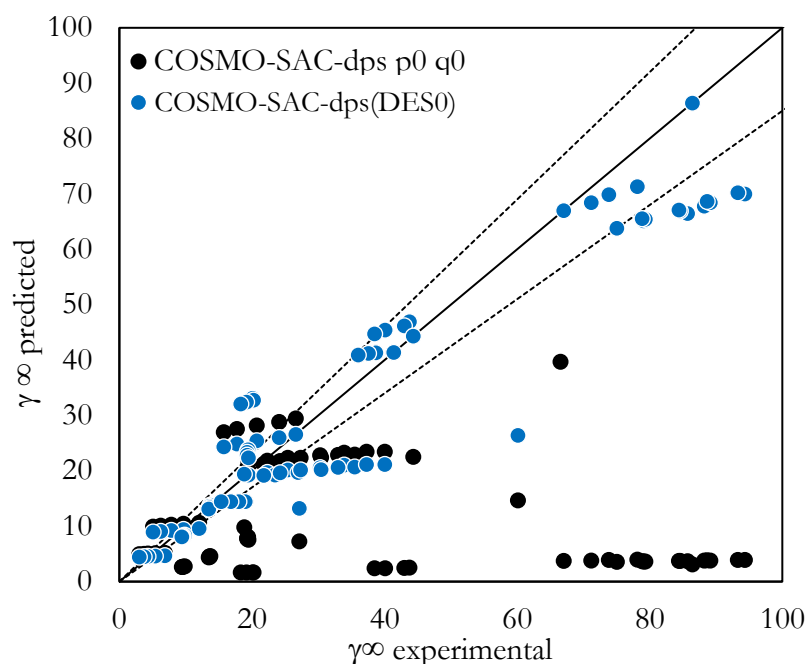


Figure 7.14 – Experimental *vs.* predicted IDAC of aromatics in DES. Black symbols represent the results of COSMO-SAC-dps with combinatorial term optimized, and blue symbols COSMO-SAC-dps(DES0). Dash lines represent 15% of the error from bisector.

In the context of cycloalkanes in DES, the provided data shows the dispersion contribution (a_{DES0}) and associated accuracy of the COSMO-SAC model in IDAC. Notably, the model displayed varying levels of precision across different DES-cycloalkane combinations. For instance, an ARD of 20.5% is showcased in BMIMCl:Gly, indicating a moderate deviation from experimental data, while in BDMIMCl:EG, a higher ARD (34.0%) is revealed, suggesting notable discrepancies. On the other hand, a low ARD (1.3%) is obtained in TMAC:EG, demonstrating a relatively accurate prediction. The absence of data for ChCl makes it challenging to assess its predictive performance. Figure 7.15 provides a clear illustration of the significant improvement achieved by optimizing the dispersion contribution compared to predictions where only the combinatorial term was optimized. The refined dispersion contribution, integrated into the COSMO-SAC model, notably enhances the accuracy of predicted thermodynamic properties, minimizing the deviation from experimental data.

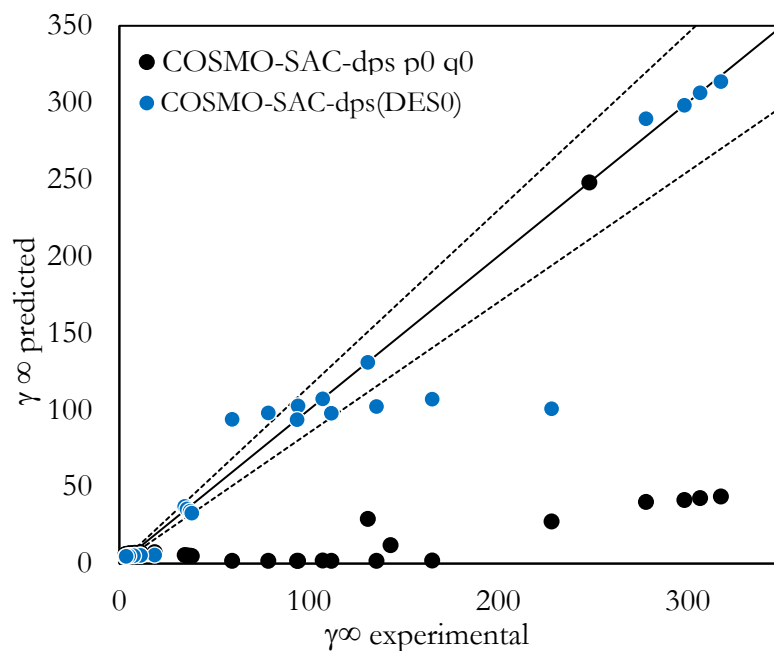


Figure 7.15 – Experimental *vs.* IDAC of cyclohydrocarbons in DES. Black symbols represent the results of COSMO-SAC-dps with combinatorial term optimized, and blue symbols COSMO-SAC-dps(DES0). Dash lines represent 15% of the error from bisector.

The ARD of IDAC of ketones, notably in BMIMCl:Gly and BMIMCl:EG exhibited lower values 3.0% and 22.4% respectively, indicating a relatively closer agreement between the model and experimental data. On the contrary, a higher ARD (47.5%) is obtained in TMAC:C6diol, suggesting a notable deviation from experimental results. In Figure 7.16, a graphical representation is provided, comparing experimental IDAC with predicted values for ketones within DES.

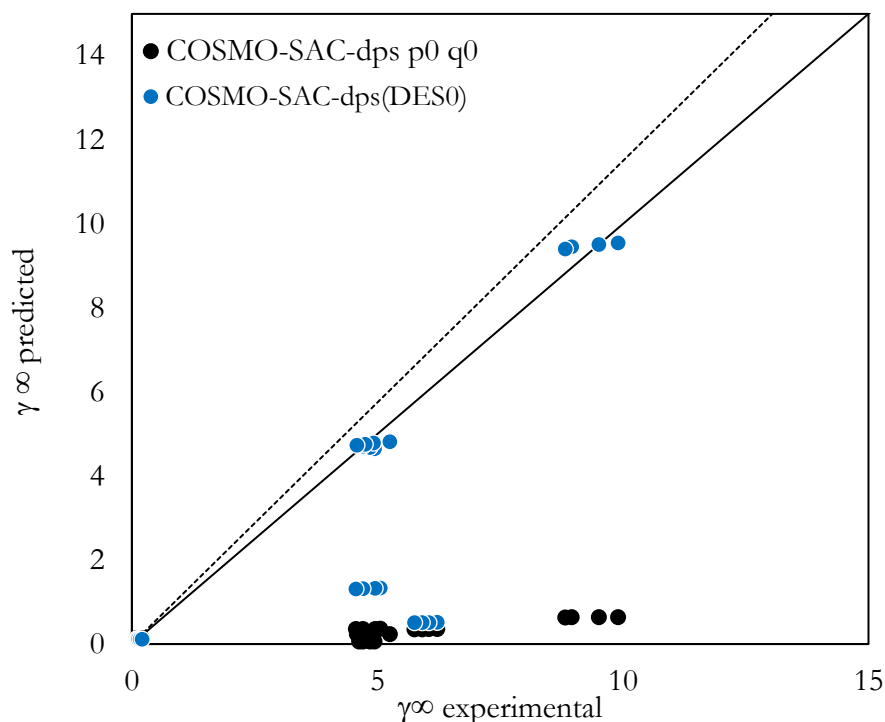


Figure 7.16 – Experimental *vs.* IDAC of ketones in DES. Black symbols represent the results of COSMO-SAC-dps with combinatorial term optimized, and blue symbols COSMO-SAC-dps(DES0). Dash lines represent 15% of the error from bisector.

In summary, this section conducted an in-depth investigation into the role of the dispersion contribution within the COSMO-SAC model for predicting IDAC in DES. The challenges encountered in accurately predicting IDAC for alkanes and alcohols highlighted the complexity of these solutes within different DES. Notably, the model excelled in predicting IDAC for aromatic solvents, underscoring its precision and applicability within this solute family. This provides valuable insights for refining the model and enhancing IDAC predictions across diverse DES systems. Nevertheless, despite the optimization, the COSMO-SAC model continues to have difficulties in accurately characterizing interactions involving alkanes. The observed relationship between the error and molecular size suggests that the model encounters challenges in predicting interactions proportional to the size of the alkane molecule and the affinity between the molecules. In this regard, a new contribution has been proposed and will be introduced.

7.5. COSMO-SAC-DES model

The model proposed in this work, called by COSMO-SAC-DES, is a new variant of COSMO-SAC-dps¹²³, that extends from our prior examination of how each parameter affects the precision of predicting IDAC within DESs. Additionally, the model utilizes the A+B sigma-profile available in the PSL sigma-profile database. These profiles were generated through a meticulous study conducted at the quantum mechanical level using DFT calculations, presented in the preceding section. All parameters, except for the residual contribution, have been meticulously optimized using experimental data specific to DES. The essence of this model lies in tailoring the predictions for DES by introducing a novel contribution, which will be thoroughly elucidated in the upcoming section.

7.5.1. Dispersive, eccentricity, and solubility contribution

To enhance the predictive power of the COSMO-SAC model, a contribution inspired by the Scatchard-Hildebrand model has been integrated. The Scatchard-Hildebrand model, as comprehensively elucidated in Chapter 3, is a widely recognized tool used for estimating the interaction energy or affinity between a solute and a solvent in a range of systems. Furthermore, this contribution incorporates the acentric factor (ω), a representation of a molecule's polarity and shape. Within this contribution, two parameters, a_{DES0} and a_{DES1} , are included, and their optimization is a key aspect of this enhancement.

The parameter A_{DES} , introduced in the preceding section, now takes the form of a function involving the dispersive term from the original COSMO-SAC-dps, the solubility parameters of each compound, (δ_i) and ω . This relationship is expressed by Equation 7.12.

$$A_{DES} = A_D \cdot A_{ES} \quad (7.12)$$

Where A_D is the original dispersive term suggested by Hsied and A_{ES} is

$$A_{ES} = (a_{DES0} + \omega \cdot a_{DES1}) \cdot (\delta_1 - \delta_2) \quad (7.13)$$

In this context, the newly introduced contribution, named henceforth as the Contribution of Dispersive, Eccentricity, and Solubility or simply DES contribution, is expressed as follows:

$$\ln \gamma_1^{des} = A_{DES} x_2^2 \quad (7.14)$$

$$\ln \gamma_2^{des} = A_{DES} x_1^2 \quad (7.15)$$

Finally, the COSMO-SAC-DES predicts the activity coefficient using the following expression:

$$\ln\gamma_{ij} = \ln\gamma^{comb} + \ln\gamma^{res} + \ln\gamma^{des} \quad (7.16)$$

where the DES contribution considers dispersive interactions, solute eccentricity through the acentric factor, and molecular affinity via the solubility parameter.

Parameters a_{DES0} and a_{DES1} were optimized using the experimental data of 84 data points of IDAC of conventional solutes in BMIMCl:Gly (ratio 2:1). The optimization was carried out based on the defined objective function (Eq. 7.7).

The δ_i , was obtained from Abboud *et al.* (1999)¹⁹⁵, while ω of the solute was obtained from NIST database¹⁹⁶ and is presented in Table 7.10.

Table 7.10 – Constants for pure component.

	n points		δ MPa ^{1/2}	ω
Alkanes	36	<i>n</i> -pentane	14.5	0.2506
		<i>n</i> -hexane	14.9	0.3003
		<i>n</i> -heptane	15.2	0.3497
		<i>n</i> -nonane	15.6	0.4482
		<i>n</i> -decane	15.8	0.4878
		1-pentene	14.5	0.2359
		1-hexene	15	0.2885
		1-heptene	15.3	0.3419
		1-decene	15.8	0.4717
Cyclo	8	Cyclohexene	17.3	0.2198
		Cyclohexane	16.8	0.2101
Aromatics	24	Benzene	18.8	0.2101
		Toluene	18.2	0.265
		Ethylbenzene	18	0.3035
		m-xylene	18	0.3271
		p-xylene	18	0.3232
		o-xylene	18.4	0.3125
Alcohols	16	Methanol	29.4	0.5614
		Ethanol	26.5	0.6427
		1-Propanol	24.6	0.6157
		1-Butanol	23.2	0.5948
		Glycerol	34.2	1.51

Experimentally determining solubility parameters for non-volatile solvents like ionic liquids and DESs, can be challenging, and only a few experimental values are available in the literature^{197,198}. As a result, some researchers use Molecular Dynamics (MD) simulations to compute these parameters^{142,199}. The solubility parameter of BMIMCl was estimated to be 35 MPa^{1/2} by Abbott et al.²⁰⁰ By considering the contribution of the solubility parameter of each molecule that comprises the DES, and weighting them by the ratio, the solubility parameter for BMIMCl:Gly (2:1) was estimated to be 34.4 MPa^{1/2} in this work.

7.5.2. Investigation of the DES contribution

The investigation in this section explores the DES contribution within the COSMO-SAC-DES model, aiming to enhance prediction accuracy for a variety of solute families. Tab. 7.11 provides a comprehensive view of the Average Relative Deviation (ARD) and the fine-tuned parameters a_{DES0} and a_{DES1} for different solute groups. Notably, the ARDs obtained for alkanes, cycloalkanes, aromatics, and alcohols are significantly reduced compared to the previous one-parameter a_{DES0} optimization. Specifically, alkanes demonstrate an ARD of 18.2% (previously 49.1%), cycloalkanes showcase 14.2% (previously 20.4%), aromatics achieve 4.0% (previously 25.3%), and alcohols exhibit a 20% ARD (previously 46.1%).

Table 7.11 – ARD of IDAC in DES using COSMO-SAC-DES models.

Conventional Solute	COSMO-SAC-dps(DES0)		COSMO-SAC-DES		
	ARD	a_{DES0}	ARD	a_{DES0}	a_{DES1}
Alkanes	49.1%	15.45	18.2%	0.1571	2.2159
Cyclo	20.4%	16.91	14.2%	2.6241	-7.7351
Aromatics	25.3%	16.63	4.0%	0.5993	1.4187
Alcohols	46.1%	12.51	20%	-97.87	178.8

Figure 7.17 illustrates the three optimizations studied in this work. The first optimization, considering the parameters p_0 and q_0 in the combinatorial contribution, is referred to as COSMO-SAC-dps* and is represented by the black symbols. The second optimization, incorporating the one-parameter a_{DES0} in the dispersive term, is referred to as COSMO-SAC-dps-DES0 and is depicted by the blue circles. Finally, the optimization presented in this section, considering the novel contribution that considers dispersive, eccentricity, and solubility in two parameters, a_{DES0} and a_{DES1} , is referred to as COSMO-SAC-DES, and is represented by green stars. The zoomed-in view on the right of Fig. 7.17 provides a closer look at the comparison, revealing the improved accuracy achieved through optimization.

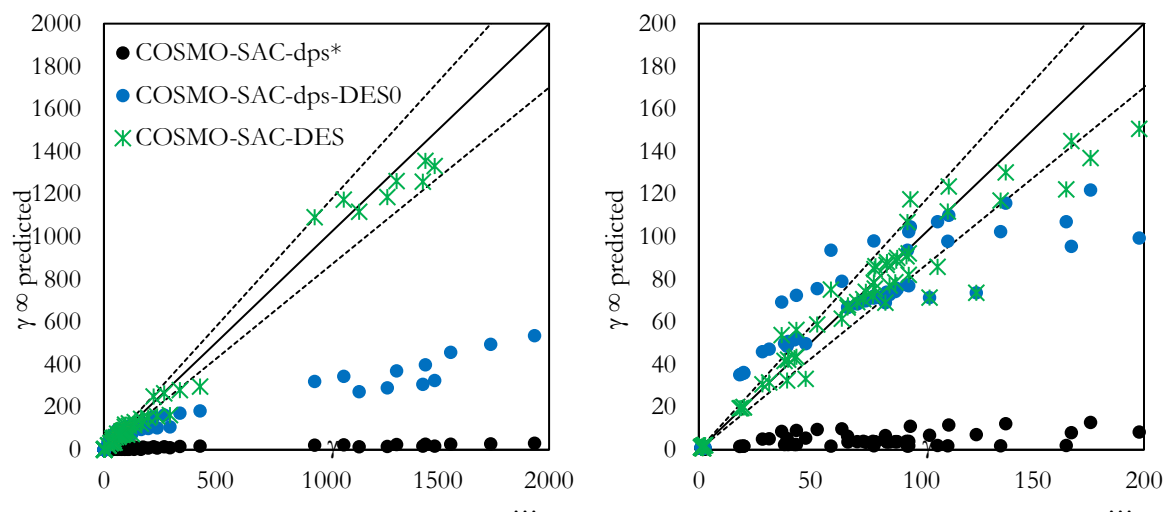
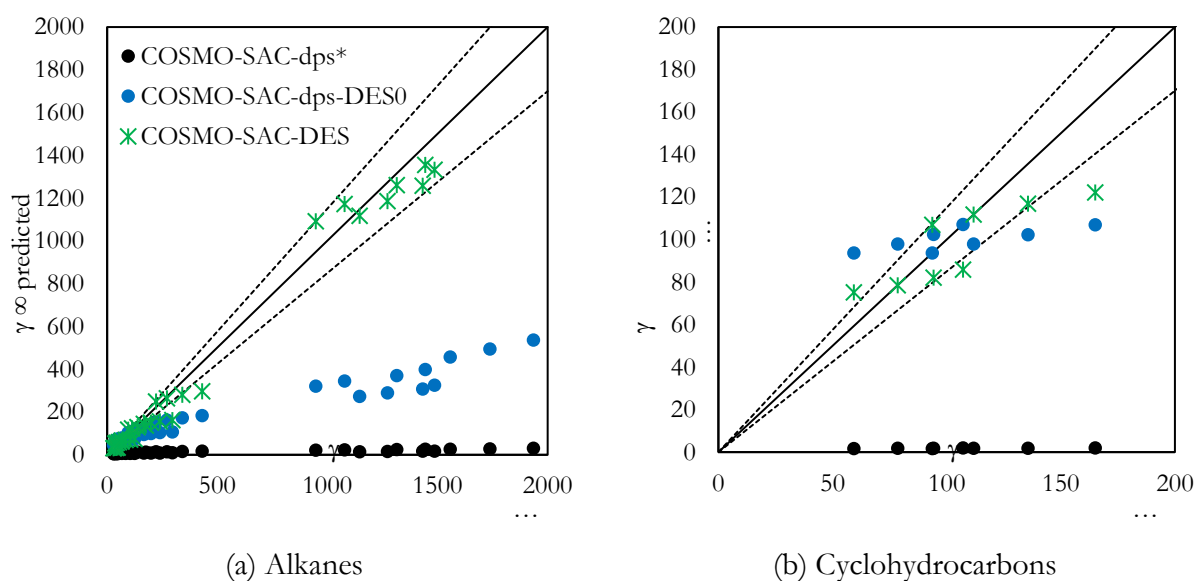


Figure 7.17 – Left: Experimental vs. predicted IDAC of conventional solute in BMIMCl:Gly (ratio 2:1). Dash lines represent 15% of the error. Right: zoomed-in view.

Figure 7.18 categorizes the obtained results by solute group. In Fig. 7.18a, predicted and experimental IDACs of alkanes in BMIMCl:Gly are represented, while Fig. 7.18b,c and d showcase these IDACs of cyclohydrocarbons, aromatics, and alcohols, respectively.



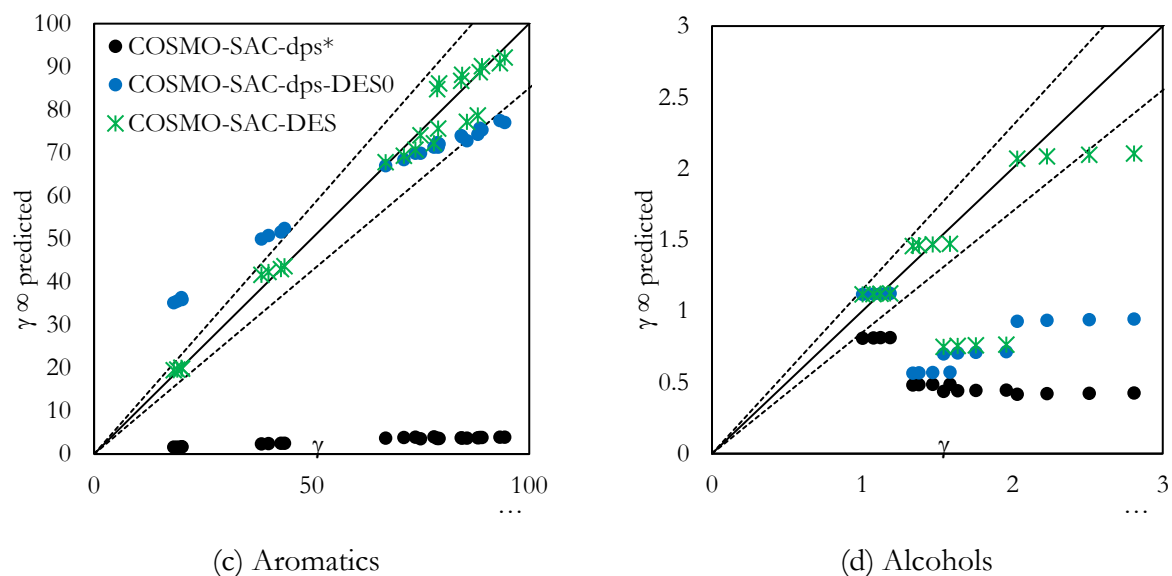


Figure 7.18 – Experimental *vs.* predicted IDAC of conventional solute in BMIMCl:Gly (ratio 2:1). Dash lines represent 15% of the error from bisector.

In conclusion, this section was dedicated to a detailed investigation of the dispersive, eccentricity, and DES solubility contribution within the COSMO-SAC-DES model, aiming to refine predictions of IDAC for diverse solute families in DESs. Through meticulous parameter optimization and refinement, the model demonstrated enhanced accuracy in predicting IDAC across various solute groups, including alkanes, cyclohydrocarbons, aromatics, and alcohols. The tailored parameters a_{DES0} and a_{DES1} , optimized based on experimental data, underscored the model's precision and its ability to account for the dispersion interactions and molecular interactions specific to different solute-solvent combinations. The comparative analysis of the three optimization approaches, COSMO-SAC-dps*, COSMO-SAC-dps-DES0, and COSMO-SAC-DES, showcased notable improvements in prediction accuracy, emphasizing the significance of the refined DES contribution in the model. Overall, this section significantly contributes to advancing the understanding and predictive capabilities of the COSMO-SAC-DES model for a wide array of solutes in DES. In the next section, the focus will be on evaluating Vapor-Liquid Equilibrium (VLE) predictions using the COSMO-SAC-DES model.

7.5.3. Vapor-liquid equilibrium predictions

Figures 7.19 and 7.20 illustrate the vapor-liquid equilibrium (VLE) isotherms for the mixture of Glycaline (1) in alcohols, particularly Ethanol (2) and Isopropanol (2), across various temperatures. Three specific temperatures are examined: 303 K, 313 K, and 323 K. In these figures, the symbols represent experimental data points, providing a direct comparison with the model predictions.

For a comprehensive evaluation of the model's predictions, both figures incorporate various lines and color codes. The black lines correspond to the Wilson model. Additionally, the dashed lines represent predictions made using the NRTL model. These results are covered in chapter 5 and are contrasted with the original predictions generated by the COSMO-SAC-dps model. It's crucial to highlight that the original COSMO-SAC-dps model predicted the activity coefficient of this mixture to be lower than 1, which can have significant implications for the understanding of the mixture's behavior.

To enhance predictions, when dealing with DES, COSMO-SAC-DES has been evaluated. This model is assessed through two different approaches, each with its own significance. First, the dash orange lines represent predictions made with parameters optimized exclusively using IDAC data. Second, the orange lines signify predictions made with parameters fine-tuned using all gamma values derived from VLE data predicted using the NRTL model. The contrast between these two sets of predictions showcases the impact of parameterization on the accuracy of the COSMO-SAC-DES model.

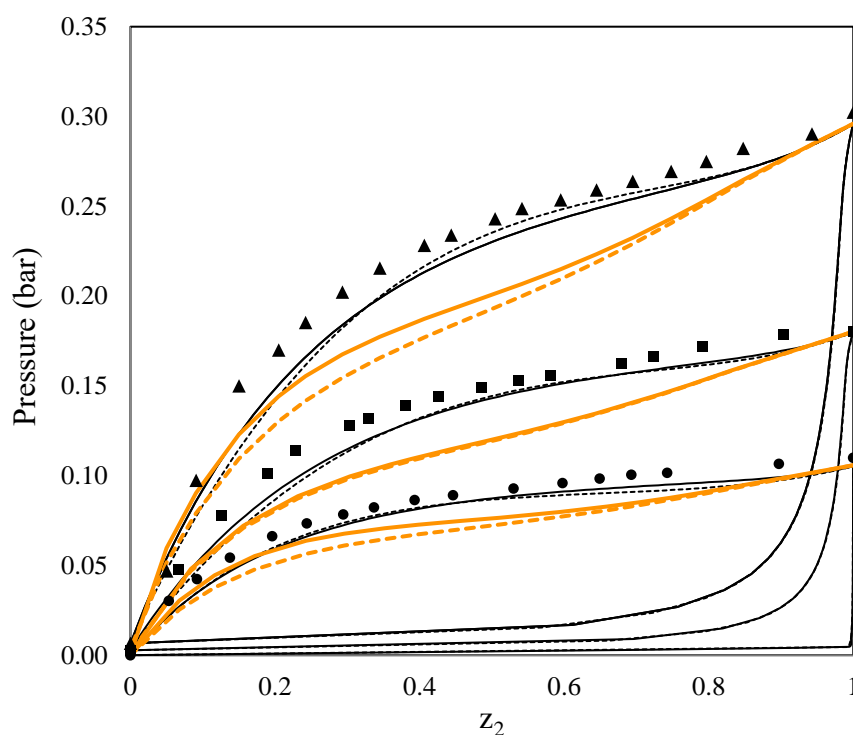


Figure 7.19 – VLE isotherms at (●)303 K, (■) 313 K, (▲) 323 K of Glycaline (1) and Ethanol (2). Symbols: Experimental data. Black lines: Wilson. Dash lines: NRTL. Orange: COSMO-SAC-DES: parameters optimized with IDAC data (dash orange lines) and optimized with all gammas from VLE data predicted with NRTL (orange lines).

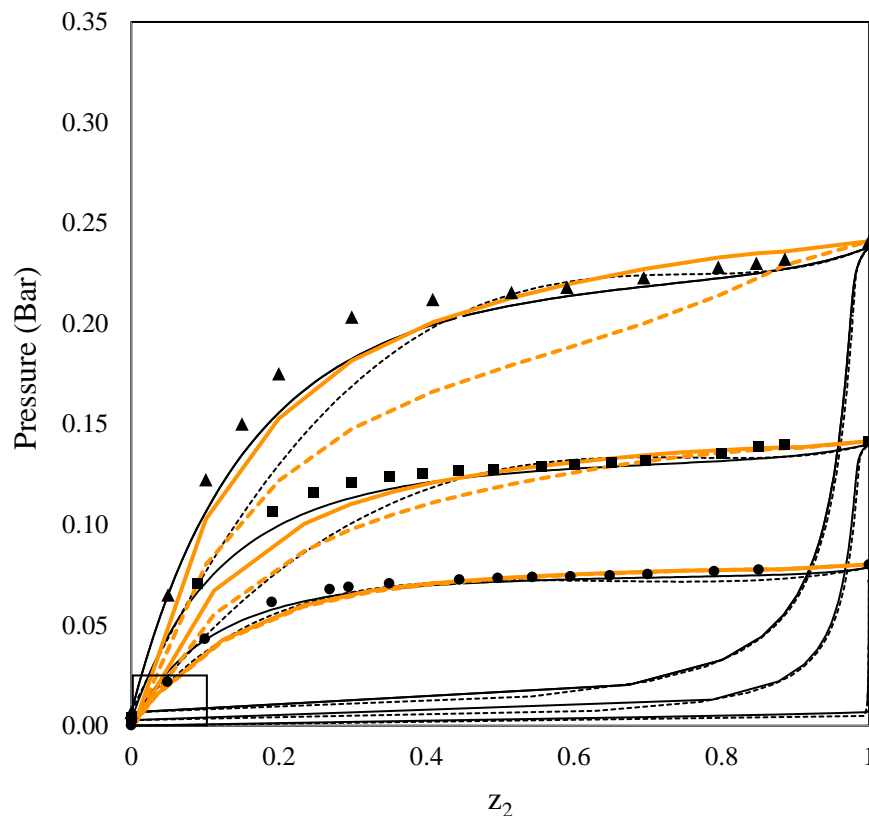


Figure 7.20 – VLE isotherms at (●)303 K, (■) 313 K, (▲) 323 K of Glycaline (1) and Isopropanol (2). Symbols: Experimental data. Black lines: Wilson. Dash lines: NRTL. Orange: COSMO-SAC-DES: parameters optimized with IDAC data (dash orange lines) and optimized with all gammas from VLE data predicted with NRTL (orange lines).

In summary, Figures 7.19 and 7.20 provide a comprehensive visual representation of VLE isotherms for Glycaline in alcohols at different temperatures. It allows for a comparative assessment of model predictions, highlights the influence of parameterization on the COSMO-SAC-DES model.

7.5.3.1. Margules two-parameters for enhanced predictions

Recognizing the need for continual improvement in predicting VLE isotherms, a two-parameter Margules model has been explored to enhance accuracy. This model specifically incorporates the DES contribution (presented in the sections 7.5.2) with two parameters. The tuning of these parameters follows a two-step approach, with the first parameter, A_{DES} in the Eq. (7.17) is fine-tuned exclusively using data at infinite dilution. The DES contribution takes the form of Eq. (7.17).

$$\ln \gamma_{ij}^{des} = (a_{DES2} + 2 \cdot (A_{DES} - a_{DES2}) \cdot x_1) \cdot x_2^2 \quad (7.17)$$

Figures 7.21 and 7.22 are presented in a manner analogous to the preceding illustrations (7.19 and 7.20), however using the 2-margules DES contribution. These figures focus on VLE isotherms for the mixture

of Glycaline (1) in two different alcohols: Ethanol (2) in Figure 7.21 and Isopropanol (2) in Figure 7.22. Three distinct temperatures are examined, 303 K, 313 K, and 323 K.

Symbols in both figures represent experimental data points, providing a valuable benchmark for evaluating the model predictions. Where black lines denote the Wilson model, while dashed lines signify predictions using the NRTL model. The Margules 2-parameters model is introduced in this context, with its predictions portrayed through blue lines.

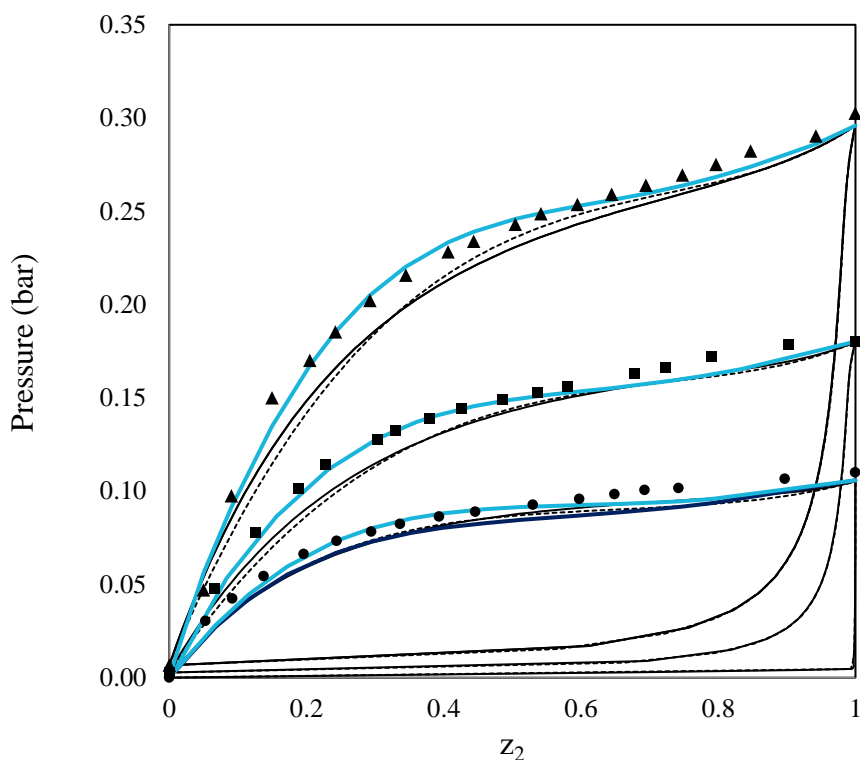


Figure 7.21 – VLE isotherms at (●)303 K, (■) 313 K, (▲) 323 K of Glycaline (1) and Ethanol (2). Symbols: Experimental data. Black lines: Wilson. Dash lines: NRTL. Blue: 2-parameters COSMO-SAC-DES.

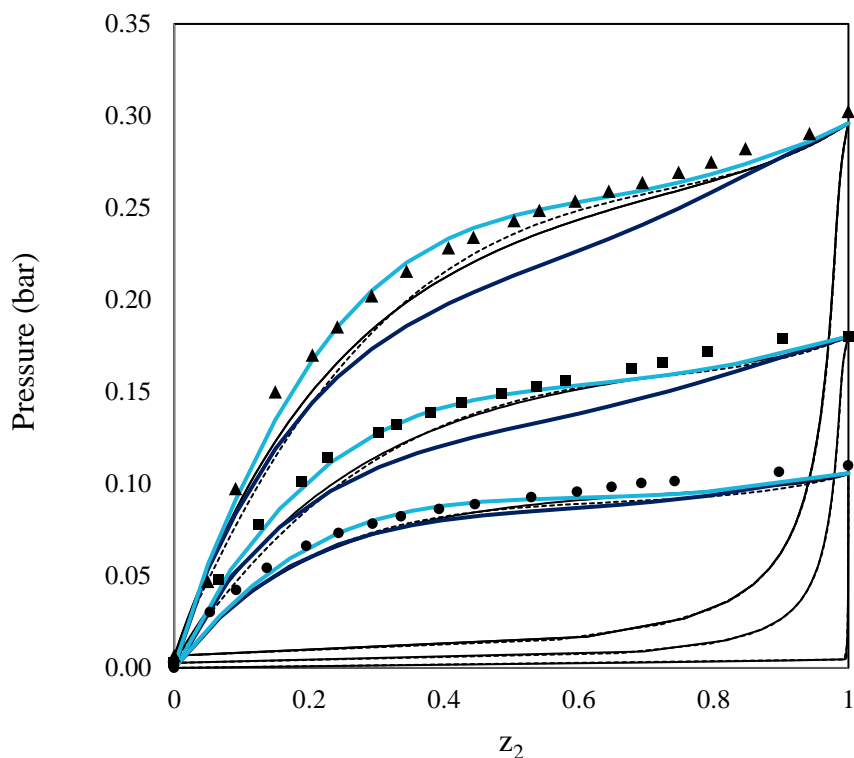


Figure 7.22 – VLE isotherms at (●)303 K, (■) 313 K, (▲) 323 K of Glycaline (1) and Isopropanol (2). Symbols: Experimental data. Black lines: Wilson. Dash lines: NRTL. Blue: 2-parameters COSMO-SAC-DES.

It is imperative to note the significance of the Margules 2-parameters model in addressing the limitations of the original COSMO-SAC-dps model, which predicted activity coefficients below 1. This refinement is crucial for a more accurate understanding of the mixture's behavior and has broader implications in practical applications.

To optimize the second parameter of the Margules 2-parameters model, a meticulous process was undertaken. The optimization involved utilizing 20 points from the VLE dataset, and the errors were evaluated for optimization conducted with 2 until 20 points. The results of this optimization process are elucidated in Figure 7.23.

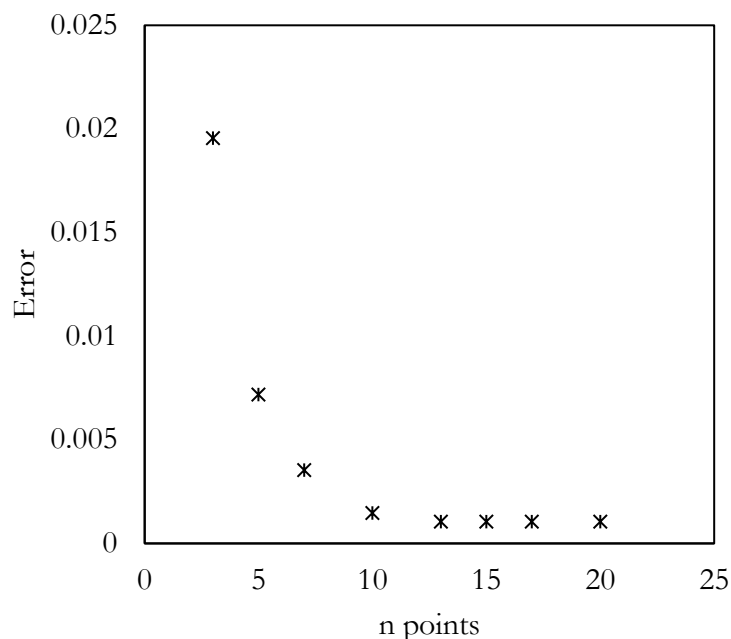


Figure 7.23 –Margules 2-parameters optimization: Illustrating the evolution of prediction error with increasing optimization points (optimal 10 points).

Figure 7.23 provides a clear representation of the errors associated with different numbers of points used for optimization. Notably, the graph reveals that using 10 points from the VLE dataset yields a favorable level of accuracy. This finding is instrumental in determining an optimal balance between computational efficiency and predictive precision.

In conclusion, Figures 7.21 and 7.22 offer an insightful visual representation of VLE isotherms for Glycaline in alcohols at different temperatures. The inclusion of the Margules 2-parameters model allows for a thorough comparative assessment of model predictions, shedding light on the influence of parameterization on its accuracy. The optimization process, as demonstrated in Figure 7.23, underscores the practicality of utilizing 10 points for achieving a favorable level of accuracy in the Margules 2-parameters model.

The methodology presented in this section, utilizing IDAC and VLE predicted with the COSMO-SAC-DES model, holds promise for screening solvents in the context of biorefinery applications. Particularly, it offers a targeted approach for selecting suitable DESs. The subsequent section will elaborate on how this methodology can be effectively applied to optimize DES selection.

7.6. Solvents selection

As described in Chapter 1, selectivity stands as a fundamental parameter for the effective selection of a solvent, based on its affinity with the solute. This critical aspect of selectivity is reflected in the solvation power (S_p), as defined by Equation 1.1 introduced in the initial chapter. To reiterate, the equation is presented as follows:

$$S_p = \frac{1}{\gamma_i^\infty} \frac{M_i}{M_{solv}} \quad (7.18)$$

where M_i and M_{solv} are the molar mass of solute and solvent, respectively.

The S_p value provides a qualitative measure of the affinity between the solvent and the solute. A higher solvation power indicates a stronger affinity. It serves as a rapid and efficient tool for screening solvents based on their interactions with the solute, allowing for a quick and effective assessment of potential solvents.

In Figure 7.24, the solvation power values predicted by the COSMO-SAC-DES model and the corresponding experimental values of BMIMCl:Gly for conventional solvents are illustrated. Remarkably, the predicted and experimental results are in close agreement, underscoring the accuracy and reliability of the COSMO-SAC-DES model. Upon qualitative analysis of the results, a trend emerges where most of the studied solutes exhibit low solvation power values, indicating a poor affinity of BMIMCl:Gly towards them. Lower solvation power values may suggest immiscibility, highlighting the limited interaction and compatibility between BMIMCl:Gly and these solutes.

Exceptionally high solvation power values were observed for alcohols, e.g. 649% for ethanol and 402% for isopropanol (not represented in the figure), suggesting a notably strong affinity between BMIMCl:Gly and this solvent family. These elevated solvation power values indicate a potent compatibility and interaction between BMIMCl:Gly and alcohols, indicating their potential suitability for various applications where a strong affinity with BMIMCl:Gly is desirable.

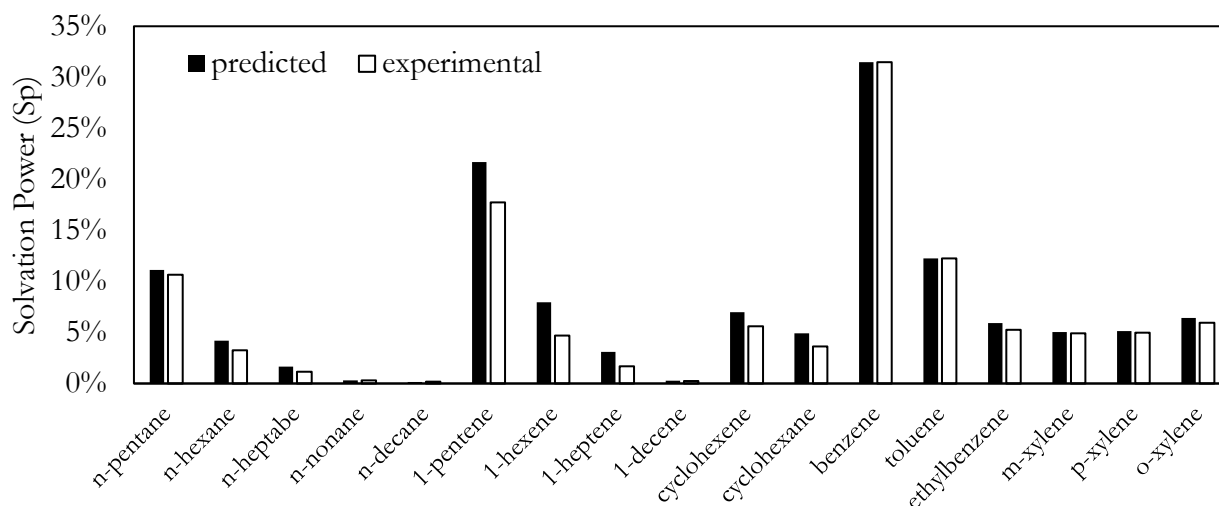


Fig. 7.24 – Solvation power predicted and experimental of BMIMCl:Gly for conventional solutes.

In summary, Figure 7.25 presents the solvation power of all the studied DESs, color-coded indicates the affinity between solvent and solute (red color is denoting lower affinity and green color is indicating higher affinity). This color scheme helps visualizing the different levels of compatibility between DES and the studied solute.

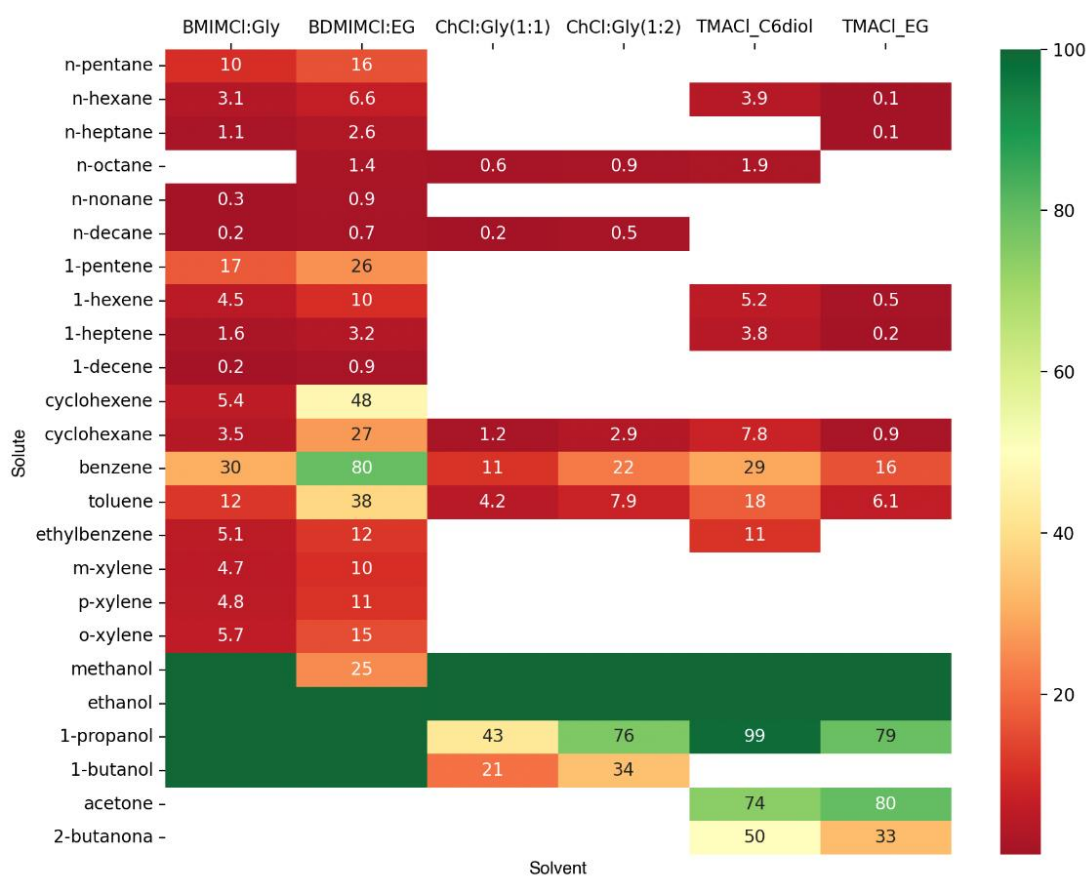


Figure 7.25 - Solvation power values (Sp) of several DES for conventional solutes.

7.7. Conclusion

In summary, this chapter investigated the role of the dispersion contribution within the COSMO-SAC model for predicting IDAC in DES. The challenges faced in accurately predicting IDAC for alkanes and alcohols elucidate the complexities of these solutes within different DES. Notably, the model displayed remarkable competence in predicting IDAC for aromatic solvents, underscoring its precision and suitability for this solute family. These findings offer valuable insights for refining the model and improving IDAC predictions in various DES systems. Nevertheless, despite optimization efforts, the COSMO-SAC-dps model continues to face challenges in accurately characterizing interactions involving alkanes. The observed relationship between errors and molecular size suggests that the model encounters difficulties in predicting interactions proportional to the size of the alkane molecule and the affinity between molecules. To address this, a new contribution has been proposed.

A comprehensive investigation of the dispersive, eccentricity, and DES solubility contributions within the COSMO-SAC-DES model, aiming to enhance predictions of IDAC for diverse solute families in DESs. Through meticulous parameter optimization and refinement, the model demonstrated improved accuracy in predicting IDAC across various solute groups, including alkanes, cyclohydrocarbons, aromatics, and alcohols. The two parameters a_{DES0} and a_{DES1} , optimized based on experimental data, underscored the model's precision and its ability to account for dispersion interactions and molecular interactions specific to different solute-solvent combinations. The comparative analysis of the three optimization approaches, COSMO-SAC-dps*, COSMO-SAC-dps-DES0, and COSMO-SAC-DES, highlighted significant improvements in prediction accuracy, emphasizing the importance of the refined DES contribution in the model.

Furthermore, the evaluation of VLE predictions using the COSMO-SAC-DES model was conducted. Two methodologies were presented, which use IDAC or VLE data to fine-tune parameters to predict activity coefficient with the COSMO-SAC-DES model. This approach shows great potential for the efficient screening of solvents in the context of biorefinery applications. It provides a targeted approach for selecting suitable DESs. This methodology provides a targeted and efficient means of selecting suitable DESs. A methodology for effectively optimizing DES selection was presented, grounded in the solvation power of all the studied DESs. This color-coded scheme aids in visualizing the diverse levels of compatibility between DES and the studied solute.

Chapter 8

Conclusion and perspectives

Conclusion et perspectives

Conclusion et perspectives (Français)

8.1. Conclusion (Français)

Cette thèse débute en mettant en évidence la dépendance mondiale au pétrole brut pour l'énergie et la nécessité de trouver des alternatives durables. Les matières premières issues de la biomasse sont présentées comme une option renouvelable, et le concept de bioraffinerie est exploré, notant ses avantages et limites. Le rôle central de la biomasse lignocellulosique dans l'énergie durable est souligné, en mettant l'accent sur les processus de séparation et de purification pour la conversion de la biomasse. Le choix du solvant, une décision cruciale dans l'extraction liquide-liquide en tant que partie du processus de séparation, est reconnu comme essentiel. Il prend en compte divers facteurs, y compris les propriétés physico-chimiques, les considérations économiques et la possibilité d'atteindre la sélectivité. Dans ce contexte, les solvants respectueux de l'environnement ou "verts" sont particulièrement attrayants en raison de leurs qualités telles que la faible toxicité, le coût abordable et la biodégradabilité. Par conséquent, ce travail s'est principalement concentré sur l'exploration des liquides ioniques profonds (DES) à ces fins.

Ce travail explore les aspects expérimentaux, se concentrant spécifiquement sur la synthèse et la caractérisation des liquides ioniques profonds (DES) et de leurs mélanges, la pression de vapeur, les données d'équilibre liquide-vapeur à l'aide de la méthode statique-synthétique, et le coefficient d'activité à dilution infinie (IDAC) par la méthode de stripping gazeux.

En résumé, les données d'équilibre liquide-vapeur pour quatre mélanges étudiés à différentes températures. Les isothermes pour l'éthaline dans l'éthanol valident l'exactitude des modèles prédictifs, montrant une étroite correspondance entre les prédictions du modèle et les données expérimentales. Notamment, le modèle Wilson montre une meilleure concordance à mesure que les températures augmentent, surpassant les performances du modèle NRTL. Une idée cruciale issue des expériences est la source potentielle d'incertitude due aux mesures de basse pression. Les valeurs de basse pression enregistrées pourraient avoir un impact sur l'exactitude des résultats. Pour atténuer cela, l'utilisation d'un transducteur spécialisé conçu pour les mesures de basse pression est suggérée, en particulier dans les cas où les valeurs de pression se situent dans une plage très basse.

Pour le mélange d'éthaline et d'isopropanol, une bonne concordance est observée entre les données expérimentales et les prédictions à des températures de 303 et 313 K. À mesure que les températures augmentent, le modèle Wilson continue à surpasser le modèle NRTL. Cependant, pour le mélange de

glycaline et d'éthanol, les données se situent constamment en dessous des prévisions de tous les modèles, en particulier à des températures de 323 et 333 K. Dans le cas du mélange de glycaline et d'isopropanol, une modélisation précise est mise en avant comme étant cruciale, le modèle Wilson se comportant particulièrement bien à des températures élevées.

L'étude a également comparé les données expérimentales aux prédictions obtenues en utilisant le modèle COSMO-SAC-dps. Les résultats indiquent que le modèle n'est pas adapté aux liquides ioniques profonds (DES).

Dans cette étude, les bases pour la création de la base de données de profil sigma PSL ont été posées par le biais d'une enquête sur les solvants conventionnels, dans le but de comprendre comment les facteurs de calcul influencent les performances des modèles prédictifs. L'enquête a commencé par une étude de référence en utilisant le furane dans un solvant apolaire pour évaluer l'influence de l'ensemble de base sur la précision de la prédiction de l'IDAC. Il a été découvert que le choix de l'ensemble de base et du niveau de théorie dans les calculs DFT affecte significativement la précision des prédictions. Les options les plus adaptées ont été identifiées comme l'ensemble de base 6-311G(2df,p) et la fonctionnelle PBE0.

Sur la base de ces connaissances, la base de données de profil sigma PSL a été établie. Elle a présenté des prédictions de l'IDAC dans les solvants conventionnels et les a comparées au modèle UNIFAC(Do). Dans l'ensemble, UNIFAC(Do) a surpassé le modèle COSMO-SAC-dps, en particulier pour la plupart des systèmes. COSMO-SAC-dps a démontré une grande précision uniquement dans la prédiction de l'IDAC pour des combinaisons spécifiques. Nous avons également exploré différentes bases de données de profils sigma et modèles pour la prédiction des coefficients d'activité.

Cette étude a examiné cinq modèles de coefficients d'activité, NRTL, UNIFAC(Do), Scatchard-Hildebrand (SH), et SH avec terme combinatoire (Staverman-Guggenheim) et le modèle COSMO-SAC, pour prédire l'équilibre liquide-vapeur et l'IDAC des alcanes dans les solvants conventionnels. Il a été observé que le modèle UNIFAC(Do) surpasse le modèle NRTL, qui s'est révélé inadapté pour la prédiction de l'IDAC. Le modèle SH a également été évalué, et ses prédictions ont été améliorées pour certains systèmes en incluant un terme combinatoire, le modèle Staverman-Guggenheim. De plus, l'étude a comparé les performances du modèle COSMO-SAC en utilisant différentes bases de données de profils sigma. La base de données PSL-UFF a donné des résultats supérieurs par rapport à PSL-Klamt et UD pour la plupart des systèmes.

En résumé, l'étude de recherche a présenté plusieurs conclusions et contributions clés :

Importance des méthodes théoriques et des ensembles de base : L'étude a souligné le rôle crucial de la sélection de méthodes théoriques appropriées et d'ensembles de base pour les calculs de mécanique quantique lors de la génération de profils sigma pour les liquides ioniques profonds (DES).

Création de la base de données de profil sigma PSL : L'étude a décrit la création de profils sigma pour les DES, essentiels pour représenter la polarité moléculaire. Elle a utilisé deux approches distinctes, AB et A+B, pour créer des profils sigma pour des systèmes DES spécifiques.

Prédictions de l'IDAC : La recherche s'est concentrée sur la prédiction de l'IDAC dans les liquides ioniques profonds (DES) en utilisant le modèle COSMO-SAC. Elle a constaté qu'un modèle COSMO-SAC traditionnel n'est pas adapté aux prédictions pour les DES, mais une version améliorée, le COSMO-SAC-DES, montre des perspectives. L'approche A+B pour les profils sigma a été préférée pour son coût abordable et sa précision modérée.

Modèle COSMO-SAC-DES : L'étude a introduit le modèle COSMO-SAC-DES, adapté pour prédire l'IDAC dans les DES. Il a incorporé diverses contributions, telles que les interactions dispersives, l'excentricité du soluté et l'affinité moléculaire, pour améliorer les prédictions. Les paramètres ont été optimisés sur la base de données expérimentales.

Prédictions d'équilibre liquide-vapeur (VLE) : La recherche a étendu la méthodologie pour prédire le comportement de l'équilibre liquide-vapeur (VLE) dans les DES, en mettant l'accent sur l'influence de la paramétrisation sur le modèle COSMO-SAC-DES.

Méthodologie de sélection des solvants : L'étude a analysé le pouvoir de solvation de divers DES pour des solutés conventionnels, fournissant des informations sur l'affinité entre les solvants et les solutés.

En conclusion, cette étude contribue à une meilleure compréhension des systèmes DES, de leurs interactions moléculaires et de la prédiction de propriétés thermodynamiques importantes comme l'IDAC. Le modèle COSMO-SAC-DES offre une approche prometteuse pour un criblage efficace des solvants, en particulier dans les applications de bioraffinerie, avec le potentiel d'améliorer la sélection et la conception de systèmes DES pour divers processus industriels.

8.2. Perspective (Français)

Les perspectives pour les futures recherches consistent à étendre l'application de la méthodologie COSMO-SAC-DES développée à un large éventail de solvants. Alors que ce travail s'est principalement concentré sur les DES, il existe un potentiel pour élargir cette approche à divers types de solvants. Cette expansion peut être cruciale pour répondre à divers besoins industriels et scientifiques. Elle implique d'adapter le modèle et les paramètres pour convenir aux caractéristiques spécifiques de différents solvants, qu'il s'agisse de liquides ioniques, de solvants organiques ou d'autres familles.

La simplicité de l'approche proposée, qui combine des calculs de chimie théorique et l'IDAC expérimental pour évaluer les corrections et faciliter la génération de diagrammes de phases ou la sélection de solvants, offre un grand potentiel pour les applications industrielles. La transition de l'université à des scénarios industriels pratiques est une réalisation significative. En perspective, cette intégration implique une collaboration avec des partenaires industriels et l'application de la méthodologie à des processus industriels. Cela peut fournir des solutions plus durables et plus efficaces pour la sélection de solvants, l'optimisation des processus et la réduction de l'impact environnemental.

Au-delà de l'IDAC, il existe un potentiel pour élargir le champ d'application afin de prédire un éventail d'autres propriétés thermophysiques à l'équilibre. Cela comprend les propriétés de transport, telles que la diffusivité et la viscosité, ainsi que les densités et la tension superficielle.

Pour illustrer, les hypothèses selon lesquelles les phases bulk et de surface sont en équilibre, et la surface partielle molaire du composant i est la même que la surface molaire de i , conduisent à l'équation de Sprow et Prausnitz^{201,202} pour la prédiction de la tension superficielle. Dans cette équation, la tension superficielle est déterminée en fonction du coefficient d'activité de chaque composant. Une alternative intéressante consiste à utiliser COSMO-SAC-DES pour prédire le coefficient d'activité, plutôt que de s'appuyer sur des modèles comme UNIFAC.

Étant donné que les processus industriels impliquent souvent des systèmes multi-composants, il est important d'évaluer la faisabilité et la précision de la méthode proposée dans de tels scénarios. Cela implique de comprendre comment la méthode se comporte dans la prédiction du comportement des phases, notamment la solubilité et la sélectivité, lorsque plusieurs composants sont impliqués.

En résumé, les perspectives de ce travail incluent l'expansion de la méthodologie à un éventail plus large de solvants, son intégration dans les pratiques industrielles, la prédiction des propriétés thermophysiques et la résolution des défis posés par les systèmes multi-composants. Ces orientations ont le potentiel

d'améliorer notre compréhension des équilibres de phase et de contribuer à des processus plus durables et plus efficaces dans divers domaines scientifiques et industriels.

Conclusion and perspectives (English)

8.3. Conclusion (English)

This thesis begins by highlights the global dependence on crude oil for energy and the need for sustainable alternatives. Biomass feedstocks are introduced as a renewable option, and the concept of biorefinery is explored, noting its advantages and limitations. Lignocellulosic biomass's central role in sustainable energy is emphasized, with a focus on separation and purification processes for biomass conversion.

The process of solvent selection, a crucial decision in liquid-liquid extraction as part of the separation process, is recognized as essential. It considers various factors, including physico-chemical properties, economic considerations, and the ability to achieve selectivity. In this context, environmentally friendly or 'green' solvents are particularly attractive due to their qualities such as low toxicity, cost-effectiveness, and biodegradability. As a result, this work primarily centered around the exploration of DES for these purposes.

This work explores the experimental aspects, specifically focusing on the synthesis and characterization of Deep Eutectic Solvents (DES) and their mixtures, vapor pressure, and VLE data using the static-synthetic method, and IDAC by gas stripping method.

In summary, VLE data for four studied mixtures across various temperature. The isotherms for ethaline in ethanol validate the accuracy of predictive models, showing close alignment between model predictions and experimental data. Notably, the Wilson model demonstrates better agreement as temperatures increase, surpassing the NRTL model's performance. A crucial insight from the experiments is the potential source of uncertainty due to low-pressure measurements. The recorded low-pressure values could impact result accuracy. To mitigate this, the use of a specialized transducer designed for low-pressure measurements is suggested, especially in cases where pressure values fall within a very low range.

For the ethaline and isopropanol mixture, good agreement is observed between experimental data and predictions at temperatures of 303 and 313 K. As temperatures increase, the Wilson model continues to outperform the NRTL model. However, for the glycaline and ethanol mixture, the data consistently falls below the predictions across all models, particularly at temperatures of 323 and 333 K. In the case of the gycaline and isopropanol mixture, accurate modeling is highlighted as crucial, with the Wilson model performing notably well at higher temperatures.

The study also compared the experimental data with predictions obtained using the COSMO-SAC-dps model. The findings indicate that the model is not suitable for DES.

In this study, the foundation for the creation of the PSL sigma-profile database was laid through an investigation into conventional solvents, aiming to understand how computational factors impact the performance of predictive models. The investigation began with a benchmark study using furan in an apolar solvent to evaluate the influence of the basis set on the accuracy of predicting IDAC. It was discovered that the choice of basis set and the level of theory in DFT calculations significantly affects the accuracy of predictions. The most suitable options were identified as the 6-311G(2df,p) basis set and the PBE0 functional.

Based on these insights, the PSL sigma-profile database was established. It was presented predictions of IDAC in conventional solvents and compared them to the UNIFAC(Do) model. Overall, UNIFAC(Do) outperformed the COSMO-SAC-dps model, especially for most systems. COSMO-SAC-dps demonstrated high accuracy only in predicting IDAC for specific combinations. We also explored different sigma-profile databases and models for predicting activity coefficients.

This study examined five activity coefficient models, NRTL, UNIFAC(Do), Scatchard-Hildebrand (SH), and SH with combinatorial term (Staverman-Guggenheim) and COSMO-SAC model, for predict VLE and IDAC of alkanes in conventional solvents. It was observed that UNIFAC(Do) was found to outperform NRTL, which was deemed unsuitable for predicting IDAC. The SH model was also evaluated, and its predictions were improved for certain systems by including a combinatorial term, the Staverman-Guggenheim model. Furthermore, the study compared the performance of the COSMO-SAC model using different sigma-profile databases. The PSL-UFF database yielded superior results compared to PSL-Klamt and UD for most systems.

In summary, the research study presented several key findings and contributions:

Importance of theoretical methods and basis sets: The study emphasized the critical role of selecting appropriate theoretical methods and basis sets for quantum mechanical calculations when generating sigma-profiles for DESs.

Sigma-Profile PSL database generation: The study described the generation of sigma-profiles for DES, which are vital for representing molecular polarity. It employed two distinct approaches, AB and A+B, for creating sigma-profiles for specific DES systems.

IDAC predictions: The research focused on predicting IDAC in DES using the COSMO-SAC model. It found that a traditional COSMO-SAC model is not well-suited for DES predictions, but an improved version, COSMO-SAC-DES, showed promise. The A+B sigma-profile approach was preferred for its cost-effectiveness and moderate accuracy.

COSMO-SAC-DES model: The study introduced the COSMO-SAC-DES model, adapted for predicting IDAC in DES. It incorporated various contributions, such as dispersive interactions, solute eccentricity, and molecular affinity, to improve predictions. Parameters were optimized based on experimental data.

VLE predictions: The research extended the methodology to predict VLE behavior in DES, emphasizing the influence of parameterization on the COSMO-SAC-DES model.

Methodology to select solvent: The study analyzed the solvation power of various DESs for conventional solutes, providing insights into the affinity between solvents and solutes.

In conclusion, this study contributes to a better understanding of DES systems, their molecular interactions, and the prediction of important thermodynamic properties like IDAC. The COSMO-SAC-DES model offers a promising approach for efficient solvent screening, particularly in biorefinery applications, with the potential to improve the selection and design of DES systems for various industrial processes.

8.4. Perspectives (English)

As perspectives for future research is expand the application of the developed methodology COSMO-SAC-DES to a wide type of solvents. While this work has primarily focused on DES, there is potential to extend this approach to encompass various types of solvents. This expansion can be crucial for addressing diverse industrial and scientific needs. It involves adapting the model and parameters to suit the specific characteristics of different solvents, be they ionic liquids, organic solvents, or any other family.

The simplicity of the proposed approach, which combines theoretical chemistry calculations and experimental IDAC to assess corrections and facilitate phase diagram generation or solvent selection, holds significant promise for industrial applications. The transition from the academia to practical industrial scenarios is a significant achievement. As perspective, this integration involves collaborating with industry partners and applying the methodology to industrial processes. This can provide more sustainable and efficient solutions for solvent selection, process optimization, and environmental impact reduction.

Beyond IDAC, there is potential to expand the scope to predict a range of other thermophysical properties at equilibrium. These include transport properties, such as diffusivity and viscosity, as well as densities and surface tension.

To illustrate, the very assumptions that bulk and surface phases are in equilibrium, and the partial molar area of the component i is the same as the molar area of i lead to the Sprow and Prausnitz^{201,202} equation for prediction of surface tension. In this equation the surface tension is determined based on the activity coefficient of each component. An interesting alternative is to use COSMO-SAC-DES to predict the activity coefficient, rather than relying on models like UNIFAC.

As industrial processes often involve multicomponent systems, it's important to assess the feasibility and precision of the proposed method in such scenarios. This involves understanding how the method performs in predicting phase behavior, including solubility and selectivity, when multiple components are involved.

In summary, the perspectives for this work include expanding the methodology to a wider range of solvents, integrating it into industrial practices, predicting thermophysical properties, and addressing the challenges of multicomponent systems. These directions hold the potential to further enhance our understanding of phase equilibria and contribute to more sustainable and efficient processes in various scientific and industrial domains.

References

- (1) Commission, E.; for Research, D.-G.; Innovation. *European Green Deal – Research & Innovation Call*; Publications Office of the European Union, 2021. <https://doi.org/doi/10.2777/33415>.
- (2) Ritchie, H.; Roser, M.; Rosado, P. Our World in Data: Energy. *Our World in Data* **2022**.
- (3) modified from IEA. Based on Data from International Energy Agency. *Net Zero by 2050* **2021**.
- (4) U.S. Energy Information Administration. *International Energy Outlook 2021*. https://www.eia.gov/outlooks/ieo/pdf/IEO2021_Narrative.pdf (accessed 2023-01-16).
- (5) Ritchie, H.; Roser, M.; Rosado, P. Our World in Data: CO₂ and Greenhouse Gas Emissions. *Our World in Data* **2022**.
- (6) Robinson, M. M.; Dowsett, H. J.; Chandler, M. A. Pliocene Role in Assessing Future Climate Impacts. *Eos, Transactions American Geophysical Union* **2008**, *89* (49), 501. <https://doi.org/10.1029/2008EO490001>.
- (7) Dumitru, O. A.; Austermann, J.; Polyak, V. J.; Fornós, J. J.; Asmerom, Y.; Ginés, J.; Ginés, A.; Onac, B. P. Constraints on Global Mean Sea Level during Pliocene Warmth. *Nature* **2019**, *574* (7777), 233–236. <https://doi.org/10.1038/s41586-019-1543-2>.
- (8) Haywood, A. M.; Dowsett, H. J.; Dolan, A. M. Integrating Geological Archives and Climate Models for the Mid-Pliocene Warm Period. *Nat Commun* **2016**, *7* (1), 10646. <https://doi.org/10.1038/ncomms10646>.
- (9) Rosenbloom, N. A.; Otto-Bliesner, B. L.; Brady, E. C.; Lawrence, P. J. Simulating the Mid-Pliocene Warm Period with the CCSM4 Model. *Geosci Model Dev* **2013**, *6* (2), 549–561. <https://doi.org/10.5194/gmd-6-549-2013>.
- (10) Fedorov, A. V.; Brierley, C. M.; Lawrence, K. T.; Liu, Z.; Dekens, P. S.; Ravelo, A. C. Patterns and Mechanisms of Early Pliocene Warmth. *Nature* **2013**, *496* (7443), 43–49. <https://doi.org/10.1038/nature12003>.
- (11) Haywood, A. M.; Hill, D. J.; Dolan, A. M.; Otto-Bliesner, B. L.; Bragg, F.; Chan, W.-L.; Chandler, M. A.; Contoux, C.; Dowsett, H. J.; Jost, A.; Kamae, Y.; Lohmann, G.; Lunt, D. J.; Abe-Ouchi, A.; Pickering, S. J.; Ramstein, G.; Rosenbloom, N. A.; Salzmann, U.; Sohl, L.; Stepanek, C.; Ueda, H.; Yan, Q.; Zhang, Z. Large-Scale Features of Pliocene Climate: Results from the Pliocene Model Intercomparison Project. *Climate of the Past* **2013**, *9* (1), 191–209. <https://doi.org/10.5194/cp-9-191-2013>.
- (12) UNFCCC. *The Paris Agreement*. <https://unfccc.int/process-and-meetings/the-paris-agreement/the-paris-agreement> (accessed 2023-03-09).
- (13) International Energy Agency. *Net Zero Emissions by 2050 Scenario (NZE)*. <https://www.iea.org/reports/global-energy-and-climate-model/net-zero-emissions-by-2050-scenario-nze#abstract> (accessed 2023-03-09).
- (14) Huber, D. G.; Gullede, J. *Extreme Weather and Climate Change: Understanding the Link, Managing the Risk*; Pew Center on Global Climate Change Arlington, 2011.
- (15) Aadil Gulzar; Tajamul Islam; Ruquia Gulzar; Tabasum Hassan. Climate Change and Impacts of Extreme Events on Human Health: An Overview. *Indonesian Journal of Social and Environmental Issues (IJSEI)* **2021**, *2* (1), 68–77. <https://doi.org/10.47540/ijsei.v2i1.180>.
- (16) Easterling, D. R.; Evans, J. L.; Groisman, P. Y.; Karl, T. R.; Kunkel, K. E.; Ambenje, P. Observed Variability and Trends in Extreme Climate Events: A Brief Review. *Bull Am Meteorol Soc* **2000**, *81* (3), 417–425. [https://doi.org/10.1175/1520-0477\(2000\)081<0417:OVATIE>2.3.CO;2](https://doi.org/10.1175/1520-0477(2000)081<0417:OVATIE>2.3.CO;2).
- (17) Hardy, J. T. *Climate Change: Causes, Effects, and Solutions*; John Wiley & Sons, 2003.
- (18) Kron, W.; Löw, P.; Kundzewicz, Z. W. Changes in Risk of Extreme Weather Events in Europe. *Environ Sci Policy* **2019**, *100*, 74–83. <https://doi.org/10.1016/j.envsci.2019.06.007>.
- (19) Lu, Z.; Wu, C.; Wu, N.; Lu, H.; Wang, T.; Xiao, R.; Liu, H.; Wu, X. Changes in Natural Gas Hydrates in Permafrost on the Qinghai-Tibet Plateau under the Background of Global Warming and Their

- Impacts on Carbon Emissions. *China Geology* **2022**, *5* (3), 1–35. <https://doi.org/10.31035/cg2022034>.
- (20) Emmanuel Macron. *Conclusion du Conseil de planification écologique*. <https://www.elysee.fr/emmanuel-macron/2023/09/25/conseil-de-planification-ecologique> (accessed 2023-10-04).
- (21) Malode, S. J.; Gaddi, S. A. M.; Kamble, P. J.; Nalwad, A. A.; Muddapur, U. M.; Shetti, N. P. Recent Evolutionary Trends in the Production of Biofuels. *Mater Sci Energy Technol* **2022**, *5*, 262–277. <https://doi.org/10.1016/j.mset.2022.04.001>.
- (22) Gallezot, P. Conversion of Biomass to Selected Chemical Products. *Chem. Soc. Rev.* **2012**, *41* (4), 1538–1558. <https://doi.org/10.1039/C1CS15147A>.
- (23) Irmak, S. Biomass as Raw Material for Production of High-Value Products. In *Biomass Volume Estimation and Valorization for Energy*; InTech, 2017. <https://doi.org/10.5772/65507>.
- (24) Tursi, A. A Review on Biomass: Importance, Chemistry, Classification, and Conversion. *Biofuel Research Journal* **2019**, *6* (2), 962–979. <https://doi.org/10.18331/BRJ2019.6.2.3>.
- (25) Yang, S.; Yu, M. Integrated Biorefinery for Sustainable Production of Fuels, Chemicals, and Polymers. In *Bioprocessing Technologies in Biorefinery for Sustainable Production of Fuels, Chemicals, and Polymers*; Wiley, 2013; pp 1–26. <https://doi.org/10.1002/9781118642047.ch1>.
- (26) Kiss, A. A.; Lange, J. P.; Schuur, B.; Brilman, D. W. F.; van der Ham, A. G. J.; Kersten, S. R. A. Separation Technology—Making a Difference in Biorefineries. *Biomass Bioenergy* **2016**. <https://doi.org/10.1016/j.biombioe.2016.05.021>.
- (27) New, E. K.; Tnah, S. K.; Voon, K. S.; Yong, K. J.; Procentese, A.; Yee Shak, K. P.; Subramonian, W.; Cheng, C. K.; Wu, T. Y. The Application of Green Solvent in a Biorefinery Using Lignocellulosic Biomass as a Feedstock. *J Environ Manage* **2022**, *307*, 114385. <https://doi.org/10.1016/j.jenvman.2021.114385>.
- (28) Agency for Ecological Transition (ADEME); Gaz Réseau Distribution France (GRDF); Gaz Réseau Transport (GRTgaz). *Un Mix de Gaz 100% Renouvelable En 2050?*; 2018.
- (29) Soria-Verdugo, A.; Rubio-Rubio, M.; Goos, E.; Riedel, U. On the Characteristic Heating and Pyrolysis Time of Thermally Small Biomass Particles in a Bubbling Fluidized Bed Reactor. *Renew Energy* **2020**, *160*, 312–322. <https://doi.org/10.1016/j.renene.2020.07.008>.
- (30) Fatehi, H.; Weng, W.; Li, Z.; Bai, X.-S.; Aldén, M. Recent Development in Numerical Simulations and Experimental Studies of Biomass Thermochemical Conversion. *Energy & Fuels* **2021**, *35* (9), 6940–6963. <https://doi.org/10.1021/acs.energyfuels.0c04139>.
- (31) Perera, S. M. H. D.; Wickramasinghe, C.; Samarasinghe, B. K. T.; Narayana, M. Modeling of Thermochemical Conversion of Waste Biomass – a Comprehensive Review. *Biofuel Research Journal* **2021**, *8* (4), 1481–1528. <https://doi.org/10.18331/BRJ2021.8.4.3>.
- (32) Ministère de la Transition écologique et de la Cohésion des territoires. *Biomasse énergie*. <https://www.ecologie.gouv.fr/biomasse-energie> (accessed 2023-10-13).
- (33) Ministère de la Transition écologique et de la Cohésion des territoires. *Penser la forêt française de demain*. <https://www.ecologie.gouv.fr/penser-foret-francaise-demain> (accessed 2023-10-13).
- (34) Hasanov, I.; Raud, M.; Kikas, T. The Role of Ionic Liquids in the Lignin Separation from Lignocellulosic Biomass. *Energies (Basel)* **2020**, *13* (18), 4864. <https://doi.org/10.3390/en13184864>.
- (35) Mamo, G.; Faryar, R.; Karlsson, E. N. Microbial Glycoside Hydrolases for Biomass Utilization in Biofuels Applications. In *Biofuel Technologies*; Springer Berlin Heidelberg: Berlin, Heidelberg, 2013; pp 171–188. https://doi.org/10.1007/978-3-642-34519-7_7.
- (36) Tassinari, T.; Macy, C. Differential Speed Two Roll Mill Pretreatment of Cellulosic Materials for Enzymatic Hydrolysis. *Biotechnol Bioeng* **1977**, *19* (9), 1321–1330. <https://doi.org/10.1002/bit.260190906>.
- (37) Cadoche, L.; López, G. D. Assessment of Size Reduction as a Preliminary Step in the Production of Ethanol from Lignocellulosic Wastes. *Biological Wastes* **1989**, *30* (2), 153–157. [https://doi.org/10.1016/0269-7483\(89\)90069-4](https://doi.org/10.1016/0269-7483(89)90069-4).

- (38) Galbe, M.; Zacchi, G. Pretreatment of Lignocellulosic Materials for Efficient Bioethanol Production. In *Biofuels*; Springer Berlin Heidelberg: Berlin, Heidelberg, 2007; pp 41–65. https://doi.org/10.1007/10_2007_070.
- (39) Kumakura, M.; Kaetsu, I. Effect of Radiation Pretreatment of Bagasse on Enzymatic and Acid Hydrolysis. *Biomass* **1983**, *3* (3), 199–208. [https://doi.org/10.1016/0144-4565\(83\)90012-4](https://doi.org/10.1016/0144-4565(83)90012-4).
- (40) Kumakura, M.; Kojima, T.; Kaetsu, I. Pretreatment of Lignocellulosic Wastes by Combination of Irradiation and Mechanical Crushing. *Biomass* **1982**, *2* (4), 299–308. [https://doi.org/10.1016/0144-4565\(82\)90015-4](https://doi.org/10.1016/0144-4565(82)90015-4).
- (41) MOSIER, N. Features of Promising Technologies for Pretreatment of Lignocellulosic Biomass. *Bioresour Technol* **2005**, *96* (6), 673–686. <https://doi.org/10.1016/j.biortech.2004.06.025>.
- (42) Negro, M. J.; Manzanares, P.; Ballesteros, I.; Oliva, J. M.; Cabañas, A.; Ballesteros, M. Hydrothermal Pretreatment Conditions to Enhance Ethanol Production from Poplar Biomass. *Appl Biochem Biotechnol* **2003**, *105* (1–3), 87–100. <https://doi.org/10.1385/ABAB:105:1-3:87>.
- (43) Shafizadeh, F.; Bradbury, A. G. W. Thermal Degradation of Cellulose in Air and Nitrogen at Low Temperatures. *J Appl Polym Sci* **1979**, *23* (5), 1431–1442. <https://doi.org/10.1002/app.1979.070230513>.
- (44) Grous, W. R.; Converse, A. O.; Grethlein, H. E. Effect of Steam Explosion Pretreatment on Pore Size and Enzymatic Hydrolysis of Poplar. *Enzyme Microb Technol* **1986**, *8* (5), 274–280. [https://doi.org/10.1016/0141-0229\(86\)90021-9](https://doi.org/10.1016/0141-0229(86)90021-9).
- (45) Brownell, H. H.; Yu, E. K. C.; Saddler, J. N. Steam-Explosion Pretreatment of Wood: Effect of Chip Size, Acid, Moisture Content and Pressure Drop. *Biotechnol Bioeng* **1986**, *28* (6), 792–801. <https://doi.org/10.1002/bit.260280604>.
- (46) Emmel, A.; Mathias, A. L.; Wypych, F.; Ramos, L. P. Fractionation of Eucalyptus Grandis Chips by Dilute Acid-Catalysed Steam Explosion. *Bioresour Technol* **2003**, *86* (2), 105–115. [https://doi.org/10.1016/S0960-8524\(02\)00165-7](https://doi.org/10.1016/S0960-8524(02)00165-7).
- (47) Kumar, P.; Barrett, D. M.; Delwiche, M. J.; Stroeve, P. Methods for Pretreatment of Lignocellulosic Biomass for Efficient Hydrolysis and Biofuel Production. *Ind Eng Chem Res* **2009**, *48* (8), 3713–3729. <https://doi.org/10.1021/ie801542g>.
- (48) Kassim, E. A.; El-Shahed, A. S. Enzymatic and Chemical Hydrolysis of Certain Cellulosic Materials. *Agricultural Wastes* **1986**, *17* (3), 229–233. [https://doi.org/10.1016/0141-4607\(86\)90097-1](https://doi.org/10.1016/0141-4607(86)90097-1).
- (49) Fox, D. J.; Gray, P. P.; Dunn, N. W.; Marsden, W. L. Comparison of Alkali and Steam (Acid) Pretreatments of Lignocellulosic Materials to Increase Enzymic Susceptibility: Evaluation under Optimised Pretreatment Conditions. *Journal of Chemical Technology & Biotechnology* **2007**, *44* (2), 135–146. <https://doi.org/10.1002/jctb.280440205>.
- (50) MacDonald, D. G.; Bakhshi, N. N.; Mathews, J. F.; Roychowdhury, A.; Bajpai, P.; Moo-Young, M. Alkali Treatment of Corn Stover to Improve Sugar Production by Enzymatic Hydrolysis. *Biotechnol Bioeng* **1983**, *25* (8), 2067–2076. <https://doi.org/10.1002/bit.260250815>.
- (51) Palonen, H.; Thomsen, A. B.; Tenkanen, M.; Schmidt, A. S.; Viikari, L. Evaluation of Wet Oxidation Pretreatment for Enzymatic Hydrolysis of Softwood. *Appl Biochem Biotechnol* **2004**, *117* (1), 01–18. <https://doi.org/10.1385/ABAB:117:1:01>.
- (52) Varga, E.; Klinke, H. B.; Réczey, K.; Thomsen, A. B. High Solid Simultaneous Saccharification and Fermentation of Wet Oxidized Corn Stover to Ethanol. *Biotechnol Bioeng* **2004**, *88* (5), 567–574. <https://doi.org/10.1002/bit.20222>.
- (53) Martín, C.; Klinke, H. B.; Thomsen, A. B. Wet Oxidation as a Pretreatment Method for Enhancing the Enzymatic Convertibility of Sugarcane Bagasse. *Enzyme Microb Technol* **2007**, *40* (3), 426–432. <https://doi.org/10.1016/j.enzmictec.2006.07.015>.
- (54) Neely, W. C. Factors Affecting the Pretreatment of Biomass with Gaseous Ozone. *Biotechnol Bioeng* **1984**, *26* (1), 59–65. <https://doi.org/10.1002/bit.260260112>.
- (55) Euphrosine-Moy, V.; Lasry, T.; Bes, R. S.; Molinier, J.; Mathieu, J. Degradation of Poplar Lignin with Ozone. *Ozone Sci Eng* **1991**, *13* (2), 239–248. <https://doi.org/10.1080/01919519108552429>.

- (56) Pan, X.; Arato, C.; Gilkes, N.; Gregg, D.; Mabee, W.; Pye, K.; Xiao, Z.; Zhang, X.; Saddler, J. Biorefining of Softwoods Using Ethanol Organosolv Pulping: Preliminary Evaluation of Process Streams for Manufacture of Fuel-Grade Ethanol and Co-Products. *Biotechnol Bioeng* **2005**, *90* (4), 473–481. <https://doi.org/10.1002/bit.20453>.
- (57) Pan, X.; Gilkes, N.; Kadla, J.; Pye, K.; Saka, S.; Gregg, D.; Ehara, K.; Xie, D.; Lam, D.; Saddler, J. Bioconversion of Hybrid Poplar to Ethanol and Co-Products Using an Organosolv Fractionation Process: Optimization of Process Yields. *Biotechnol Bioeng* **2006**, *94* (5), 851–861. <https://doi.org/10.1002/bit.20905>.
- (58) Araque, E.; Parra, C.; Freer, J.; Contreras, D.; Rodríguez, J.; Mendonça, R.; Baeza, J. Evaluation of Organosolv Pretreatment for the Conversion of Pinus Radiata D. Don to Ethanol. *Enzyme Microb Technol* **2008**, *43* (2), 214–219. <https://doi.org/10.1016/j.enzmictec.2007.08.006>.
- (59) Kurakake, M.; Ide, N.; Komaki, T. Biological Pretreatment with Two Bacterial Strains for Enzymatic Hydrolysis of Office Paper. *Curr Microbiol* **2007**, *54* (6), 424–428. <https://doi.org/10.1007/s00284-006-0568-6>.
- (60) Tursi, A. A Review on Biomass: Importance, Chemistry, Classification, and Conversion. *Biofuel Research Journal* **2019**, *6* (2), 962–979. <https://doi.org/10.18331/BRJ2019.6.2.3>.
- (61) Waheed, Q. M. K.; Nahil, M. A.; Williams, P. T. Pyrolysis of Waste Biomass: Investigation of Fast Pyrolysis and Slow Pyrolysis Process Conditions on Product Yield and Gas Composition. *Journal of the Energy Institute* **2013**, *86* (4), 233–241. <https://doi.org/10.1179/1743967113Z.00000000067>.
- (62) Al Arni, S. Comparison of Slow and Fast Pyrolysis for Converting Biomass into Fuel. *Renew Energy* **2018**, *124*, 197–201. <https://doi.org/10.1016/j.renene.2017.04.060>.
- (63) Ischia, G.; Fiori, L. Hydrothermal Carbonization of Organic Waste and Biomass: A Review on Process, Reactor, and Plant Modeling. *Waste Biomass Valorization* **2021**, *12* (6), 2797–2824. <https://doi.org/10.1007/s12649-020-01255-3>.
- (64) Gollakota, A. R. K.; Kishore, N.; Gu, S. A Review on Hydrothermal Liquefaction of Biomass. *Renewable and Sustainable Energy Reviews* **2018**, *81*, 1378–1392. <https://doi.org/10.1016/j.rser.2017.05.178>.
- (65) Lebaka, V. R. Potential Bioresources as Future Sources of Biofuels Production: An Overview. In *Biofuel Technologies*; Springer Berlin Heidelberg: Berlin, Heidelberg, 2013; pp 223–258. https://doi.org/10.1007/978-3-642-34519-7_9.
- (66) Kong, Q.-S.; Li, X.-L.; Xu, H.-J.; Fu, Y. Conversion of 5-Hydroxymethylfurfural to Chemicals: A Review of Catalytic Routes and Product Applications. *Fuel Processing Technology* **2020**, *209*, 106528. <https://doi.org/10.1016/j.fuproc.2020.106528>.
- (67) Mamman, A. S.; Lee, J.-M.; Kim, Y.-C.; Hwang, I. T.; Park, N.-J.; Hwang, Y. K.; Chang, J.-S.; Hwang, J.-S. Furfural: Hemicellulose/Xylo-derived Biochemical. *Biofuels, Bioproducts and Biorefining* **2008**, *2* (5), 438–454. <https://doi.org/10.1002/bbb.95>.
- (68) Chen, H.; Liu, J.; Chang, X.; Chen, D.; Xue, Y.; Liu, P.; Lin, H.; Han, S. A Review on the Pretreatment of Lignocellulose for High-Value Chemicals. *Fuel Processing Technology* **2017**, *160*, 196–206. <https://doi.org/10.1016/j.fuproc.2016.12.007>.
- (69) Loow, Y.-L.; New, E. K.; Yang, G. H.; Ang, L. Y.; Foo, L. Y. W.; Wu, T. Y. Potential Use of Deep Eutectic Solvents to Facilitate Lignocellulosic Biomass Utilization and Conversion. *Cellulose* **2017**, *24* (9), 3591–3618. <https://doi.org/10.1007/s10570-017-1358-y>.
- (70) Zuo, M.; Jia, W.; Feng, Y.; Zeng, X.; Tang, X.; Sun, Y.; Lin, L. Effective Selectivity Conversion of Glucose to Furan Chemicals in the Aqueous Deep Eutectic Solvent. *Renew Energy* **2021**, *164*, 23–33. <https://doi.org/10.1016/j.renene.2020.09.019>.
- (71) Vigier, K. D. O.; Chatel, G.; Jérôme, F. Contribution of Deep Eutectic Solvents for Biomass Processing: Opportunities, Challenges, and Limitations. *ChemCatChem* **2015**, *7* (8), 1250–1260. <https://doi.org/10.1002/cctc.201500134>.
- (72) Pena-Pereira, F.; Namieśnik, J. Ionic Liquids and Deep Eutectic Mixtures: Sustainable Solvents for Extraction Processes. *ChemSusChem* **2014**, *7* (7), 1784–1800. <https://doi.org/10.1002/cssc.201301192>.

- (73) Zhang, X.; Zhu, P.; Li, Q.; Xia, H. Recent Advances in the Catalytic Conversion of Biomass to Furfural in Deep Eutectic Solvents. *Front Chem* **2022**, *10*. <https://doi.org/10.3389/fchem.2022.911674>.
- (74) Hu, S.; Zhang, Z.; Zhou, Y.; Han, B.; Fan, H.; Li, W.; Song, J.; Xie, Y. Conversion of Fructose to 5-Hydroxymethylfurfural Using Ionic Liquids Prepared from Renewable Materials. *Green Chemistry* **2008**, *10* (12), 1280. <https://doi.org/10.1039/b810392e>.
- (75) To, L. N.; Nguyen, T. H.; Nguyen, T. P.; Phan, H. B.; Nguyen, L. H. T.; Doan, T. L. H.; Dang, C. Van; Tran, P. H. Efficient Conversion of Monosaccharides into 5-Hydroxymethylfurfural Using Acidic Deep Eutectic Solvents. *Catalysts* **2023**, *13* (8), 1216. <https://doi.org/10.3390/catal13081216>.
- (76) Hu, S.; Zhang, Z.; Zhou, Y.; Song, J.; Fan, H.; Han, B. Direct Conversion of Inulin to 5-Hydroxymethylfurfural in Biorenewable Ionic Liquids. *Green Chemistry* **2009**, *11* (6), 873. <https://doi.org/10.1039/b822328a>.
- (77) Assanosi, A. A.; Farah, M. M.; Wood, Joseph.; Al-Duri, B. A Facile Acidic Choline Chloride–p-TSA DES-Catalysed Dehydration of Fructose to 5-Hydroxymethylfurfural. *RSC Adv.* **2014**, *4* (74), 39359–39364. <https://doi.org/10.1039/C4RA07065H>.
- (78) Ilgen, F.; Ott, D.; Kralisch, D.; Reil, C.; Palmberger, A.; König, B. Conversion of Carbohydrates into 5-Hydroxymethylfurfural in Highly Concentrated Low Melting Mixtures. *Green Chemistry* **2009**, *11* (12), 1948. <https://doi.org/10.1039/b917548m>.
- (79) Zhao, Q.; Sun, Z.; Wang, S.; Huang, G.; Wang, X.; Jiang, Z. Conversion of Highly Concentrated Fructose into 5-Hydroxymethylfurfural by Acid–Base Bifunctional HPA Nanocatalysts Induced by Choline Chloride. *RSC Adv.* **2014**, *4* (108), 63055–63061. <https://doi.org/10.1039/C4RA10121A>.
- (80) Liu, F.; Barrault, J.; De Oliveira Vigier, K.; Jérôme, F. Dehydration of Highly Concentrated Solutions of Fructose to 5-Hydroxymethylfurfural in a Cheap and Sustainable Choline Chloride/Carbon Dioxide System. *ChemSusChem* **2012**, *5* (7), 1223–1226. <https://doi.org/10.1002/cssc.201200186>.
- (81) Wu, S.; Fan, H.; Xie, Y.; Cheng, Y.; Wang, Q.; Zhang, Z.; Han, B. Effect of CO₂ on Conversion of Inulin to 5-Hydroxymethylfurfural and Propylene Oxide to 1,2-Propanediol in Water. *Green Chemistry* **2010**, *12* (7), 1215. <https://doi.org/10.1039/c002553d>.
- (82) Zhang, L.-X.; Yu, H.; Yu, H.-B.; Chen, Z.; Yang, L. Conversion of Xylose and Xylan into Furfural in Biorenewable Choline Chloride–Oxalic Acid Deep Eutectic Solvent with the Addition of Metal Chloride. *Chinese Chemical Letters* **2014**, *25* (8), 1132–1136. <https://doi.org/10.1016/j.ccllet.2014.03.029>.
- (83) Zhang, L.; Yu, H. Conversion of Xylan and Xylose into Furfural in Biorenewable Deep Eutectic Solvent with Trivalent Metal Chloride Added. *Bioresources* **2013**, *8* (4), 6014–6025.
- (84) Yu, Q.; Bai, R.; Wang, F.; Zhang, Q.; Sun, Y.; Zhang, Y.; Qin, L.; Wang, Z.; Yuan, Z. A Sustainable System for Maleic Acid Synthesis from Biomass-derived Sugar. *Journal of Chemical Technology & Biotechnology* **2020**, *95* (3), 751–757. <https://doi.org/10.1002/jctb.6260>.
- (85) Rusanen, A.; Lappalainen, K.; Kärkkäinen, J.; Lassi, U. Furfural and 5-Hydroxymethylfurfural Production from Sugar Mixture Using Deep Eutectic Solvent/MIBK System. *ChemistryOpen* **2021**, *10* (10), 1004–1012. <https://doi.org/10.1002/open.202100163>.
- (86) Chen, Z.; Wan, C. A Novel Deep Eutectic Solvent/Acetone Biphasic System for High-Yield Furfural Production. *Bioresour Technol Rep* **2019**, *8*, 100318. <https://doi.org/10.1016/j.biteb.2019.100318>.
- (87) Morais, E. S.; Freire, M. G.; Freire, C. S. R.; Coutinho, J. A. P.; Silvestre, A. J. D. Enhanced Conversion of Xylan into Furfural Using Acidic Deep Eutectic Solvents with Dual Solvent and Catalyst Behavior. *ChemSusChem* **2020**, *13* (4), 784–790. <https://doi.org/10.1002/cssc.201902848>.
- (88) National Academies of Sciences Engineering and Medicine. *A Research Agenda for Transforming Separation Science*; National Academies Press: Washington, D.C., 2019. <https://doi.org/10.17226/25421>.

- (89) Kiss, A. A. Design, Control and Economics of Distillation. In *Advanced Distillation Technologies*; John Wiley & Sons, Ltd: Chichester, UK, 2013; pp 37–65. <https://doi.org/10.1002/9781118543702.ch2>.
- (90) Sholl, D. S.; Lively, R. P. Seven Chemical Separations to Change the World. *Nature* **2016**, *532* (7600), 435–437. <https://doi.org/10.1038/532435a>.
- (91) Byrne, F. P.; Jin, S.; Paggiola, G.; Petchey, T. H. M.; Clark, J. H.; Farmer, T. J.; Hunt, A. J.; Robert McElroy, C.; Sherwood, J. Tools and Techniques for Solvent Selection: Green Solvent Selection Guides. *Sustainable Chemical Processes* **2016**, *4* (1), 7. <https://doi.org/10.1186/s40508-016-0051-z>.
- (92) Scilipoti, J. A.; Cismondi, M.; Andreatta, A. E.; Brignole, E. A. Selection of Solvents with A-UNIFAC Applied to Detoxification of Aqueous Solutions. *Ind Eng Chem Res* **2014**, *53* (44), 17051–17058. <https://doi.org/10.1021/ie500469f>.
- (93) Hajiw, M.; Valtz, A.; Ahmar, E. El; Coquelet, C. Apparent Henry's Law Constants of Furan in Different n-Alkanes and Alcohols at Temperatures from 293 to 323 K. *J Environ Chem Eng* **2017**, *5* (1), 1205–1209. <https://doi.org/10.1016/j.jece.2017.02.001>.
- (94) Abbott, A. P.; Capper, G.; Davies, D. L.; Munro, H. L.; Rasheed, R. K.; Tambyrajah, V. Preparation of Novel, Moisture-Stable, Lewis-Acidic Ionic Liquids Containing Quaternary Ammonium Salts with Functional Side Chains. *Chemical Communications* **2001**, No. 19, 2010–2011. <https://doi.org/10.1039/b106357j>.
- (95) ProSim. Simulis® Thermodynamics.
- (96) DSA 5000M - Instruction Manual and Safety Information. <https://www.anton-paar.com/> (accessed 2023-01-30).
- (97) Wiki Anton Paar. <https://wiki.anton-paar.com/en/> (accessed 2023-01-30).
- (98) Lovis 2000 M/ME - Instruction Manual and Safety Information. Anton Paar. <https://www.anton-paar.com/> (accessed 2023-01-30).
- (99) Höppler, F. Der Exzentrische Fall von Kugeln in Hohlzylindern Mit Flüssigkeiten Oder Gasen. *Zeitschrift für Technische Physik* **1933**, *14*, 165–169.
- (100) Stokes, G. ~G. On the Effect of the Internal Friction of Fluids on the Motion of Pendulums. *Transactions of the Cambridge Philosophical Society* **1851**, *9*, 8.
- (101) Anton Paar. ABBEMAT 300 - Instruction Manual and Safety Information. <https://www.anton-paar.com/> (accessed 2023-01-30).
- (102) Meskel-Lesavre, M.; Richon, D.; Renon, H. New Variable Volume Cell for Determining Vapor-Liquid Equilibriums and Saturated Liquid Molar Volumes by the Static Method. *Industrial & Engineering Chemistry Fundamentals* **1981**, *20* (3), 284–289. <https://doi.org/10.1021/i100003a017>.
- (103) Fontalba, F.; Richon, D.; Renon, H. Simultaneous Determination of Vapor–Liquid Equilibria and Saturated Densities up to 45 MPa and 433 K. *Review of Scientific Instruments* **1984**, *55* (6), 944–951. <https://doi.org/10.1063/1.1137870>.
- (104) Coquelet, C.; Valtz, A.; Théveneau, P. Experimental Determination of Thermophysical Properties of Working Fluids for ORC Applications. In *Organic Rankine Cycles for Waste Heat Recovery - Analysis and Applications*; IntechOpen, 2020. <https://doi.org/10.5772/intechopen.87113>.
- (105) Richon, D.; Renon, H. Infinite Dilution Henry's Constants of Light Hydrocarbons in n-Hexadecane, n-Octadecane, and 2,2,4,4,6,8,8-Heptamethylnonane by Inert Gas Stripping. *J Chem Eng Data* **1980**, *25* (1), 59–60. <https://doi.org/10.1021/je60084a028>.
- (106) Richon, D.; Antoine, P.; Renon, H. Infinite Dilution Activity Coefficients of Linear and Branched Alkanes from C₁ to C₉ in n-Hexadecane by Inert Gas Stripping. *Industrial & Engineering Chemistry Process Design and Development* **1980**, *19* (1), 144–147. <https://doi.org/10.1021/i260073a025>.
- (107) Coquelet, C.; Boonaert, E.; Valtz, A.; Huang, S. Effect of Methane, CO₂, and H₂S on the Solubility of Methyl and Ethyl Mercaptans in a 25 Wt % Methyl-diethanolamine Aqueous Solution at 333 and 365 K. *J Chem Eng Data* **2021**, *66* (11), 4000–4017. <https://doi.org/10.1021/acs.jced.1c00160>.
- (108) Alkhatib, I. I. I.; Bahamon, D.; Llovel, F.; Abu-Zahra, M. R. M.; Vega, L. F. Perspectives and Guidelines on Thermodynamic Modelling of Deep Eutectic Solvents. *J Mol Liq* **2020**, *298*, 112183. <https://doi.org/10.1016/j.molliq.2019.112183>.

- (109) Lewis, G. N. The Law of Physico-Chemical Change. *Proceedings of the American Academy of Arts and Sciences* **1901**, 37 (3), 49. <https://doi.org/10.2307/20021635>.
- (110) Lewis, G. N. Das Gesetz Physiko-Chemischer Vorgänge. *Zeitschrift für Physikalische Chemie* **1901**, 38U (1), 205–226. <https://doi.org/10.1515/zpch-1901-3816>.
- (111) Nkosi, N.; Tumba, K.; Ramsuroop, S. Measurements of Activity Coefficient at Infinite Dilution for Organic Solutes in Tetramethylammonium Chloride + Ethylene Glycol Deep Eutectic Solvent Using Gas-Liquid Chromatography. *Fluid Phase Equilib* **2018**. <https://doi.org/10.1016/j.fluid.2018.01.019>.
- (112) Delcros, S.; Grolier, J.-P. E.; Dohnal, V.; Fenclová, D. Infinite-Dilution Activity Coefficients by Comparative Ebulliometry: Measurements and Group Contribution Calculations for Some Binary Mixtures Ether + n-Alkane and Ether + Alcohol. *Chem Eng Sci* **1995**, 50 (18), 2957–2962. [https://doi.org/10.1016/0009-2509\(95\)00137-T](https://doi.org/10.1016/0009-2509(95)00137-T).
- (113) Zin, R. M.; Coquelet, C.; Valtz, A.; Abdul Mutalib, M. I.; Sabil, K. M. Measurement of Henry's Law Constant and Infinite Dilution Activity Coefficient of Isopropyl Mercaptan and Isobutyl Mercaptan in (Methyldiethanolamine (1) + Water (2)) with $w_1 = 0.25$ and 0.50 at Temperature of (298 to 348) K Using Inert Gas Stripping Method. *J Chem Thermodyn* **2016**, 93, 193–199. <https://doi.org/10.1016/j.jct.2015.10.005>.
- (114) Margules, M. Über Die Zusammensetzung Der Gesättigten Dämpfe von Mischungen. *Sitzungsber. Akad. Wiss. Wien, math.-naturwiss. Klasse* **1895**, 104, 1243–1278.
- (115) Van Laar, J. J. Vapour Pressure of Binary Mixtures. *Z. phys. Chem.* **1910**, 72, 723–751.
- (116) Hildebrand, J. H.; Scott, R. S. The Solubility of Nonelectrolytes. *J Phys Chem* **1951**, 55 (4), 619–620. <https://doi.org/10.1021/j150487a027>.
- (117) Hansen, C. M. *Hansen Solubility Parameters*; CRC Press, 2007. <https://doi.org/10.1201/9781420006834>.
- (118) Lin, S.-T.; Sandler, S. I. A Priori Phase Equilibrium Prediction from a Segment Contribution Solvation Model. *Ind Eng Chem Res* **2002**, 41 (5), 899–913. <https://doi.org/10.1021/ie001047w>.
- (119) Klamt, A. Conductor-like Screening Model for Real Solvents: A New Approach to the Quantitative Calculation of Solvation Phenomena. *J Phys Chem* **1995**, 99 (7), 2224–2235. <https://doi.org/10.1021/j100007a062>.
- (120) Renon, H.; Prausnitz, J. M. Local Compositions in Thermodynamic Excess Functions for Liquid Mixtures. *AIChE Journal* **1968**, 14 (1), 135–144. <https://doi.org/10.1002/aic.690140124>.
- (121) Staverman, A. J. The Entropy of High Polymer Solutions. Generalization of Formulae. *Recueil des Travaux Chimiques des Pays-Bas* **1950**. <https://doi.org/10.1002/recl.19500690203>.
- (122) Edward Armand Guggenheim. *Mixtures: The Theory of the Equilibrium Properties of Some Simple Classes of Mixtures, Solutions and Alloys*; Clarendon Press, Ed.; Universidade de Michigan, 1952.
- (123) Hsieh, C.-M.; Lin, S.-T.; Vrabec, J. Considering the Dispersive Interactions in the COSMO-SAC Model for More Accurate Predictions of Fluid Phase Behavior. *Fluid Phase Equilib* **2014**, 367, 109–116. <https://doi.org/10.1016/j.fluid.2014.01.032>.
- (124) Schrödinger, E. Quantisierung Als Eigenwertproblem. *Ann Phys* **1926**, 384 (4), 361–376. <https://doi.org/10.1002/andp.19263840404>.
- (125) Szabo, A.; Ostlund, N. S. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*; Dover Books on Chemistry; Dover Publications, 1996.
- (126) Born, M.; Oppenheimer, R. Zur Quantentheorie Der Molekeln. *Ann Phys* **1927**, 389 (20), 457–484. <https://doi.org/10.1002/andp.19273892002>.
- (127) Hartree, D. R. The Wave Mechanics of an Atom with a Non-Coulomb Central Field. Part I. Theory and Methods. *Mathematical Proceedings of the Cambridge Philosophical Society* **1928**, 24 (1), 89–110. <https://doi.org/10.1017/S0305004100011919>.
- (128) Fock, V. Näherungsmethode Zur Lösung Des Quantenmechanischen Mehrkörperproblems. *Zeitschrift für Physik* **1930**, 61 (1–2), 126–148. <https://doi.org/10.1007/BF01340294>.
- (129) Slater, J. C. Self-Consistent Field for Molecules and Solids. McGraw Hill Book Company 1974.

- (130) Hohenberg, P.; Kohn, W. Inhomogeneous Electron Gas. *Physical Review* **1964**, *136* (3B), B864–B871. <https://doi.org/10.1103/PhysRev.136.B864>.
- (131) Kohn, W.; Sham, L. J. Self-Consistent Equations Including Exchange and Correlation Effects. *Physical Review* **1965**, *140* (4A), A1133–A1138. <https://doi.org/10.1103/PhysRev.140.A1133>.
- (132) Cramer, C. J. *Essentials of Computational Chemistry: Theories and Models*; Wiley, 2005.
- (133) Ditchfield, R.; Hehre, W. J.; Pople, J. A. Self-Consistent Molecular-Orbital Methods. IX. An Extended Gaussian-Type Basis for Molecular-Orbital Studies of Organic Molecules. *J Chem Phys* **1971**, *54* (2), 724–728. <https://doi.org/10.1063/1.1674902>.
- (134) Perdew, J. P. Jacob's Ladder of Density Functional Approximations for the Exchange-Correlation Energy. In *AIP Conference Proceedings*; AIP, 2001; pp 1–20. <https://doi.org/10.1063/1.1390175>.
- (135) Scuseria, G. E.; Janssen, C. L.; Schaefer, H. F. An Efficient Reformulation of the Closed-Shell Coupled Cluster Single and Double Excitation (CCSD) Equations. *J Chem Phys* **1988**, *89* (12), 7382–7387. <https://doi.org/10.1063/1.455269>.
- (136) Čížek, J. On the Use of the Cluster Expansion and the Technique of Diagrams in Calculations of Correlation Effects in Atoms and Molecules; LeFebvre, R., Moser, C., Eds.; Wiley, 1969; Vol. 14, pp 35–89. <https://doi.org/10.1002/9780470143599.ch2>.
- (137) Hammes-Schiffer, S. A Conundrum for Density Functional Theory. *Science (1979)* **2017**, *355* (6320), 28–29. <https://doi.org/10.1126/science.aal3442>.
- (138) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys Rev Lett* **1996**. <https://doi.org/10.1103/PhysRevLett.77.3865>.
- (139) Tao, J.; Perdew, J. P.; Staroverov, V. N.; Scuseria, G. E. Climbing the Density Functional Ladder: Nonempirical Meta-Generalized Gradient Approximation Designed for Molecules and Solids. *Phys Rev Lett* **2003**, *91* (14), 146401. <https://doi.org/10.1103/PhysRevLett.91.146401>.
- (140) Bondi, A. Van Der Waals Volumes and Radii. *J Phys Chem* **1964**, *68* (3), 441–451. <https://doi.org/10.1021/j100785a001>.
- (141) Pascual-ahuir, J. L.; Silla, E.; Tuñón, I. GEPOL: An Improved Description of Molecular Surfaces. III. A New Algorithm for the Computation of a Solvent-Excluding Surface. *J Comput Chem* **1994**, *15* (10), 1127–1138. <https://doi.org/10.1002/jcc.540151009>.
- (142) Salehi, H. S.; Ramdin, M.; Moulton, O. A.; Vlucht, T. J. H. Computing Solubility Parameters of Deep Eutectic Solvents from Molecular Dynamics Simulations. *Fluid Phase Equilib* **2019**, *497*, 10–18. <https://doi.org/10.1016/j.fluid.2019.05.022>.
- (143) Théveneau, P.; Valtz, A.; Coquelet, C. Vapor Liquid Equilibrium Data for the Furan–Toluene Binary System between 313.02 and 352.99 K. *J Chem Eng Data* **2017**, *62* (4), 1168–1172. <https://doi.org/10.1021/acs.jced.6b00424>.
- (144) Adamo, C.; Barone, V. Toward Reliable Density Functional Methods without Adjustable Parameters: The PBE0 Model. *J Chem Phys* **1999**, *110* (13), 6158–6170. <https://doi.org/10.1063/1.478522>.
- (145) Miyazaki, G.; Tirri, B.; Baudouin, O.; Valtz, A.; Houriez, C.; Coquelet, C.; Adamo, C. Role of Computational Variables on the Performances of COSMO-SAC Model: A Combined Theoretical and Experimental Investigation. *Ind Eng Chem Res* **2021**, *60* (5), 2314–2325. <https://doi.org/10.1021/acs.iecr.0c04276>.
- (146) Binkley, J. S.; Pople, J. A.; Hehre, W. J. Self-Consistent Molecular Orbital Methods. 21. Small Split-Valence Basis Sets for First-Row Elements. *J Am Chem Soc* **1980**, *102* (3), 939–947. <https://doi.org/10.1021/ja00523a008>.
- (147) Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. Self-consistent Molecular Orbital Methods. XX. A Basis Set for Correlated Wave Functions. *J Chem Phys* **1980**, *72* (1), 650–654. <https://doi.org/10.1063/1.438955>.
- (148) Pietro, W. J.; Francl, M. M.; Hehre, W. J.; DeFrees, D. J.; Pople, J. A.; Binkley, J. S. Self-Consistent Molecular Orbital Methods. 24. Supplemented Small Split-Valence Basis Sets for Second-Row Elements. *J Am Chem Soc* **1982**, *104* (19), 5039–5048. <https://doi.org/10.1021/ja00383a007>.

- (149) Becke, A. D. Density-functional Thermochemistry. III. The Role of Exact Exchange. *J Chem Phys* **1993**, *98* (7), 5648–5652. <https://doi.org/10.1063/1.464913>.
- (150) Zhao, Y.; Truhlar, D. G. The M06 Suite of Density Functionals for Main Group Thermochemistry, Thermochemical Kinetics, Noncovalent Interactions, Excited States, and Transition Elements: Two New Functionals and Systematic Testing of Four M06-Class Functionals and 12 Other Functionals. *Theor Chem Acc* **2008**, *120* (1–3), 215–241. <https://doi.org/10.1007/s00214-007-0310-x>.
- (151) Yanai, T.; Tew, D. P.; Handy, N. C. A New Hybrid Exchange–Correlation Functional Using the Coulomb-Attenuating Method (CAM-B3LYP). *Chem Phys Lett* **2004**, *393* (1–3), 51–57. <https://doi.org/10.1016/j.cplett.2004.06.011>.
- (152) Voutsas, E. C.; Tassios, D. P. Prediction of Infinite-Dilution Activity Coefficients in Binary Mixtures with UNIFAC. A Critical Evaluation. *Ind Eng Chem Res* **1996**, *35* (5), 1438–1445.
- (153) Vosko, S. H.; Wilk, L.; Nusair, M. Accurate Spin-Dependent Electron Liquid Correlation Energies for Local Spin Density Calculations: A Critical Analysis. *Can J Phys* **1980**, *58* (8), 1200–1211. <https://doi.org/10.1139/p80-159>.
- (154) Daubert, T. E.; Danner, R.; Sibul, H.; Stebbins, C. L. DIPPR Data Compilation of Pure Compound Properties, Project 801 Sponsor Release. *Design Institute for Physical Property Data* **1993**.
- (155) Peña, M. D.; Compostizo, A.; Crespo Colín, A. Excess Gibbs Free Energy of (n-Heptane+Xylene) at 348.15 K. *J Chem Thermodyn* **1980**, *12* (3), 259–263. [https://doi.org/10.1016/0021-9614\(80\)90045-2](https://doi.org/10.1016/0021-9614(80)90045-2).
- (156) Proust, P. C.; Ramirez, A. L.; Yianatos, J. B. Vapor-Liquid Equilibria for the System n-Heptane-o-Xylene at 348.1, 358.1, and 368.3 K. *J Chem Eng Data* **1980**, *25* (4), 329–331. <https://doi.org/10.1021/jc60087a015>.
- (157) Díaz, C.; Tojo, J. Phase Equilibria Behaviour of N-Heptane with o-Xylene, m-Xylene, p-Xylene and Ethylbenzene at 101.3kPa. *J Chem Thermodyn* **2002**, *34* (12), 1975–1984. [https://doi.org/10.1016/S0021-9614\(02\)00227-6](https://doi.org/10.1016/S0021-9614(02)00227-6).
- (158) Willman, B.; Teja, A. S. Vapor-Liquid Equilibria in Toluene + m-Xylene, Toluene + n-Decane, and n-Decane + m-Xylene Mixtures. *J Chem Eng Data* **1985**, *30* (1), 116–119. <https://doi.org/10.1021/jc00039a034>.
- (159) Surovy, J.; Heinrich, J. No Title. *Sb. Pr. Chem. Fak. SVST* **1966**, *201*, 5–6.
- (160) Markuzin, N. P.; Pavlova, L. M. Vapor-Liquid Equilibrium in an Acetic Acid + Heptane + Toluene System: I. *Zh. Prikl. Khim. (Leningrad)* **1971**, *44*, 311–315.
- (161) Ashcroft, S. J.; Clayton, A. D.; Shearn, R. B. Isothermal Vapor-Liquid Equilibria for the Systems Toluene-n-Heptane, Toluene-Propan-2-Ol, Toluene-Sulfolane and Propan-2-Ol-Sulfolane. *J.Chem.Eng.Data* **1979**, *24*, 195–199.
- (162) Liu, Y.; Xie, R.; Fan, Z.; Yan, J. Studies on the VLE of N-Heptane-Methylbenzene. *Chemical Engineering (China)* **1983**, *6*, 42–50.
- (163) Berro, C.; Lai, X.; Choubi, F.; Rauzy, E. Isothermal (Vapour + Liquid) Equilibria and Excess Volumes of (3-Methylpentane + Heptane), of (3-Methylpentane + Octane), and of (Toluene + Octane). *J Chem Thermodyn* **1994**, *26* (8), 863–869. <https://doi.org/10.1006/jcht.1994.1103>.
- (164) Willman, B.; S. Teja, A. Vapor-Liquid Equilibria in Toluene + m-Xylene, Toluene + n-Decane, and n-Decane + m-Xylene Mixtures. *Journal of Chemical & Engineering Data* **2002**, *30* (1), 116–119. <https://doi.org/10.1021/jc00039a034>.
- (165) Saez, C.; Compostizo, A.; Rubio, R. G.; Crespo Colín, A.; Díaz Peña, M. P. T, x, y Data of Benzene + n-Hexane and Cyclohexane + n-Heptane Systems. *Fluid Phase Equilib* **1985**, *24* (3), 241–258. [https://doi.org/10.1016/0378-3812\(85\)85006-8](https://doi.org/10.1016/0378-3812(85)85006-8).
- (166) Susarev, M. P.; Chen, S. T. No Title. *Zh. Fiz. Khim.* **1963**, *37*, 1739.
- (167) Beyer, W.; Schuberth, H.; Leibnitz, E. No Title. *J. Prakt. Chem.* **1965**, *27*, 276.
- (168) Brown, I.; Ewald, A. Liquid-Vapour Equilibria. II. The System Benzene-n-Heptane. *Aust J Chem* **1951**, *4* (2), 198. <https://doi.org/10.1071/CH9510198>.

- (169) Jean Fu, S.; Lu, B. C.-Y. Binary Vapour-Liquid Equilibria. *Journal of Applied Chemistry* **2007**, *16* (11), 324–326. <https://doi.org/10.1002/jctb.5010161103>.
- (170) Werner, G.; Schubert, H. Das Phasengleichgewicht Flüssig-Flüssig Des Systems Benzol/n-Heptan/Acetonitril Sowie Die Phasengleichgewichte Dampfförmig-Flüssig Der Entsprechenden Binären Systeme Bei 20,0 °C. *J. Prakt. Chem.* **1966**, *31*, 225–239. <https://doi.org/https://doi.org/10.1002/prac.19660310501>.
- (171) Friend, J.; Scheller, W. A.; Weber, J. H. Calculation of Net Deviations from Consistency in Low Pressure Vapor- Liquid Equilibrium Data. *Ind. Eng. Chem. Proc. Des. Dev.* **1979**, *9*, 144.
- (172) Butcher, K. L.; Ramasubramanian, K. R.; Medani, M. S. Thermodynamic Properties of the Benzene and N-Heptane System at Elevated Temperatures. *Journal of Applied Chemistry and Biotechnology* **2007**, *22* (11), 1139–1155. <https://doi.org/10.1002/jctb.5020221103>.
- (173) A. Palmer, D.; D. Smith, B. Thermodynamic Excess Property Measurements for Acetonitrile-Benzene-n-Heptane System at 45.Deg. *Journal of Chemical & Engineering Data* **2002**, *17* (1), 71–76. <https://doi.org/10.1021/jc60052a037>.
- (174) Treszczanowicz, T.; Kehiaian, H. Vapour-Liquid Equilibria in 1,2-Dimethoxyethane+benzene and Di-n-Propyl Ether+benzene Binary Systems. *V. Bull. Acad. Pol. Sci.-Chim* **1973**, *21* (2), 97–105.
- (175) Yarym-Agaev, N. L.; Kalinichenko, V. P. Determination and Estimation of the Reliability of Liquid-Vapor Equilibrium Parameters in the Benzene + Heptane System. *Zh. Prikl. Khim. (Leningrad)* **1982**, *55*, 2612.
- (176) Wiśniewska, B.; Gregorowicz, J.; Malanowski, S. Development of a Vapour-Liquid Equilibrium Apparatus to Work at Pressures up to 3 MPa. *Fluid Phase Equilib* **1993**, *86*, 173–186. [https://doi.org/10.1016/0378-3812\(93\)87174-Y](https://doi.org/10.1016/0378-3812(93)87174-Y).
- (177) Elshayal, I. M.; Lu, B. C.-Y. J. Isothermal Vapor-Liquid Equilibria for the Binary System Benzene-n-Octane Octane. *J. Appl. Chem* **1968**, *18*, 277–280. <https://doi.org/http://dx.doi.org/10.20381/ruor-11692>.
- (178) Messow, U.; Quitzsch, K. Z. P. Thermodynamics of Solvent + Paraffin Systems. *hys. Chem. (Leipzig)* **1976**, *257*, 121.
- (179) Bell, I. H.; Mickoleit, E.; Hsieh, C. M.; Lin, S. T.; Vrabec, J.; Breitkopf, C.; Jäger, A. A Benchmark Open-Source Implementation of COSMO-SAC. *J Chem Theory Comput* **2020**. <https://doi.org/10.1021/acs.jctc.9b01016>.
- (180) LibreTexts. *The Potential-Energy Surface Can Be Calculated Using Quantum Mechanics*. [https://chem.libretexts.org/Bookshelves/Physical_and_Theoretical_Chemistry_Textbook_Maps/Physical_Chemistry_\(LibreTexts\)/30%3A_Gas-Phase_Reaction_Dynamics/30.10%3A_The_Potential-Energy_Surface_Can_Be_Calculated_Using_Quantum_Mechanics](https://chem.libretexts.org/Bookshelves/Physical_and_Theoretical_Chemistry_Textbook_Maps/Physical_Chemistry_(LibreTexts)/30%3A_Gas-Phase_Reaction_Dynamics/30.10%3A_The_Potential-Energy_Surface_Can_Be_Calculated_Using_Quantum_Mechanics) (accessed 2023-07-04).
- (181) Foresman, J. B.; Frisch, A. *Exploring Chemistry with Electronic Structure Methods*, 3rd Edition.; Gaussian, Inc.: Wallingford, 2015.
- (182) Kabane, B.; Redhi, G. G. Thermodynamic Properties and Activity Coefficients at Infinite Dilution for Different Solutes in Deep Eutectic Solvent: 1-Butyl-3-Methylimidazolium Chloride + Glycerol. *J Mol Liq* **2020**, *311*, 113216. <https://doi.org/10.1016/j.molliq.2020.113216>.
- (183) Kabane, B.; Redhi, G. G. Thermodynamic Properties and Activity Coefficients at Infinite Dilution for Different Solutes in Deep Eutectic Solvent: 1-Butyl-3-Methylimidazolium Chloride + Glycerol. *J Mol Liq* **2020**. <https://doi.org/10.1016/j.molliq.2020.113216>.
- (184) Verevkin, S. P.; Sazonova, A. Yu.; Frolkova, A. K.; Zaitsau, D. H.; Prikhodko, I. V.; Held, C. Separation Performance of BioRenewable Deep Eutectic Solvents. *Ind Eng Chem Res* **2015**, *54* (13), 3498–3504. <https://doi.org/10.1021/acs.iecr.5b00357>.
- (185) Nkosi, N.; Tumba, K.; Ramsuroop, S. Activity Coefficients at Infinite Dilution of Various Organic Solutes in the Deep Eutectic Solvent (Tetramethylammonium Chloride + 1,6 Hexanediol in the 1:1 Molar Ratio). *S Afr J Chem Eng* **2019**, *27*, 7–15. <https://doi.org/10.1016/j.sajce.2018.11.003>.
- (186) Nkosi, N.; Tumba, K.; Ramsuroop, S. Measurements of Activity Coefficient at Infinite Dilution for Organic Solutes in Tetramethylammonium Chloride + Ethylene Glycol Deep Eutectic Solvent

- Using Gas-Liquid Chromatography. *Fluid Phase Equilib* **2018**, *462*, 31–37. <https://doi.org/10.1016/j.fluid.2018.01.019>.
- (187) Han, J.; Dai, C.; Yu, G.; Lei, Z. Parameterization of COSMO-RS Model for Ionic Liquids. *Green Energy & Environment* **2018**, *3* (3), 247–265. <https://doi.org/10.1016/j.gee.2018.01.001>.
- (188) Hsieh, C.-M.; Sandler, S. I.; Lin, S.-T. Improvements of COSMO-SAC for Vapor–Liquid and Liquid–Liquid Equilibrium Predictions. *Fluid Phase Equilib* **2010**, *297* (1), 90–97. <https://doi.org/10.1016/j.fluid.2010.06.011>.
- (189) Xiong, R.; Sandler, S. I.; Burnett, R. I. An Improvement to COSMO-SAC for Predicting Thermodynamic Properties. *Ind Eng Chem Res* **2014**, *53* (19), 8265–8278. <https://doi.org/10.1021/ie404410v>.
- (190) Kikic, I.; Alessi, P.; Rasmussen, P.; Fredenslund, A. On the Combinatorial Part of the UNIFAC and UNIQUAC Models. *Can J Chem Eng* **1980**, *58* (2), 253–258. <https://doi.org/10.1002/cjce.5450580218>.
- (191) Donohue, M. D.; Prausnitz, J. M. Combinatorial Entropy of Mixing Molecules That Differ in Size and Shape. A Simple Approximation for Binary and Multicomponent Mixtures. *Can J Chem* **1975**, *53* (11), 1586–1592. <https://doi.org/10.1139/v75-224>.
- (192) Zhong, C.; Sato, Y.; Masuoka, H.; Chen, X. Improvement of Predictive Accuracy of the UNIFAC Model for Vapor-Liquid Equilibria of Polymer Solutions. *Fluid Phase Equilib* **1996**, *123* (1–2), 97–106. [https://doi.org/10.1016/S0378-3812\(96\)90015-1](https://doi.org/10.1016/S0378-3812(96)90015-1).
- (193) Wang, S.; Sandler, S. I.; Chen, C.-C. Refinement of COSMO-SAC and the Applications. *Ind Eng Chem Res* **2007**, *46* (22), 7275–7288. <https://doi.org/10.1021/ie070465z>.
- (194) Soares, R. de P. The Combinatorial Term for COSMO-Based Activity Coefficient Models. *Ind Eng Chem Res* **2011**, *50* (5), 3060–3063. <https://doi.org/10.1021/ie102087p>.
- (195) Abboud, J.-L. M.; Notari, R. Critical Compilation of Scales of Solvent Parameters. Part I. Pure, Non-Hydrogen Bond Donor Solvents. *Pure and Applied Chemistry* **1999**, *71* (4), 645–718. <https://doi.org/10.1351/pac199971040645>.
- (196) ASPEN Properties. NIST ThermoData Engine.
- (197) Yoo, B.; Afzal, W.; Prausnitz, J. M. Solubility Parameters for Nine Ionic Liquids. *Ind Eng Chem Res* **2012**, *51* (29), 9913–9917. <https://doi.org/10.1021/ie300588s>.
- (198) Weerachanchai, P.; Wong, Y.; Lim, K. H.; Tan, T. T. Y.; Lee, J.-M. Determination of Solubility Parameters of Ionic Liquids and Ionic Liquid/Solvent Mixtures from Intrinsic Viscosity. *ChemPhysChem* **2014**, *15* (16), 3580–3591. <https://doi.org/10.1002/cphc.201402345>.
- (199) Mohan, M.; Huang, K.; Pidatala, V. R.; Simmons, B. A.; Singh, S.; Sale, K. L.; Gladden, J. M. Prediction of Solubility Parameters of Lignin and Ionic Liquids Using Multi-Resolution Simulation Approaches. *Green Chemistry* **2022**, *24* (3), 1165–1176. <https://doi.org/10.1039/D1GC03798F>.
- (200) Abbott, S.; Hansen, C. M.; Yamamoto, H.; Valpey III, R. S. *Hansen Solubility Parameters in Practice: Complete with Software, Data, and Examples*; Hansen-Solubility.com, 2013.
- (201) Sprow, F. B.; Prausnitz, J. M. Vapour-Liquid Equilibria and Surface Tensions for the Nitrogen-Argon-Methane System at 90.67° K. *Cryogenics (Guildf)* **1966**, *6* (6), 338–340. [https://doi.org/10.1016/0011-2275\(66\)90131-7](https://doi.org/10.1016/0011-2275(66)90131-7).
- (202) Sprow, F. B.; Prausnitz, J. M. Surface Tensions of Simple Liquid Mixtures. *Transactions of the Faraday Society* **1966**, *62*, 1105. <https://doi.org/10.1039/tf9666201105>.
- (203) Cox, J. D.; Herington, E. F. G. The Coexistence Curve in Liquid-Liquid Binary Systems. *Transactions of the Faraday Society* **1956**, *52*, 926. <https://doi.org/10.1039/tf9565200926>.
- (204) Inozemtsev, P. P.; Liakumovich, A. G.; Gracheva, Z. D. Heats of Fusion of Sterically Hindered Phenols Based on Differential Thermal Analysis Data. *Russian Journal of Physical Chemistry* **1972**, *46* (6), 1594–1596.

Appendix A

A.1. Liquid-liquid equilibrium

The LL equilibrium was correlated to obtain the activity coefficient solving a set of two liquid-liquid equilibrium equations at a fixed temperature and pressure:

$$\begin{cases} x_1^{L1} \cdot \gamma_1^{L1} = x_1^{L2} \cdot \gamma_1^{L2} & (7) \\ (1 - x_1^{L1}) \cdot \gamma_2^{L1} = (1 - x_1^{L2}) \cdot \gamma_2^{L2} & (8) \end{cases}$$

Where x_1^{L1} and x_1^{L2} are the mole fractions of component 1 in each of the two liquid phases in equilibrium and γ_i is the activity coefficient of component i in a liquid phase.

To prediction of LLE, it was considered a liquid mixture with molar fraction x_i across the entire range composition can be split into two coexistence liquid phases in equilibrium (L1 and L2). The LL compositions necessary to correlate the experimental data obtained were obtained by application of the Cox and Herington equation²⁰³. The phase-equilibrium problem can be solved numerically by Eq. 10 and 11 given initial estimates for x_1^{L1} and x_1^{L2} until a convergence criterion is reached.

A.1.1. Cox Herington plot for coexistence data

Cox and Herington showed that the two arms of the coexistence data curves could be represented over a large range of temperature by the equations:

$$(T - T_c)^\beta = A^{L1} \log \frac{x_1^{L1}}{1 - x_1^{L1}} + B^{L1} \text{ for } x_1^{L1} > x_1^c \quad (9)$$

$$(T - T_c)^\beta = A^{L2} \log \frac{x_1^{L2}}{1 - x_1^{L2}} + B^{L2} \text{ for } x_1^{L2} > x_1^c \quad (10)$$

The superscripts L1 and L2 denote the two arms of the coexistence curves less and greater than the critical composition (x_1^c), respectively. Where x_i is molar composition, and the two constants A and B are two correlating parameters. In this equation $B=A \log\left(\frac{1-x^c}{x^c}\right)$, so that an expansion of logarithm near-critical

composition leads the approximate form for representation of equilibrium composition as a function of temperature:

$$x_1^{L1} - x_1^c = K^{L1}(T - T_c)^\beta \quad (11)$$

$$x_1^{L2} - x_1^c = K^{L2}(T - T_c)^\beta \quad (12)$$

Where the exponent β justifies the linearity of the Cox-Herington plot and usual is considered equal 1/3. However, it can be varied to the best fit of data. The parameters obtained are represented in table 4.

A.2 Solid-liquid equilibrium

The solid-liquid equilibrium was obtained by the thermodynamic relationship between solubility and activity coefficients, which was not considered the difference of the molar heat capacities of the pure component due to the temperature range of the solid-liquid equilibrium for phenol and n-hexane is not so far from the melting point of pure components.

$$\ln(x_i \gamma_i) = \left[\frac{\Delta H_{m,i}}{RT_{m,i}} \cdot \left(1 - \frac{T_{m,i}}{T} \right) \right] \quad (13)$$

where $T_{m,i}$ is the melting point of the pure component and $\Delta H_{m,i}$ is the molar enthalpy of melting. The $\Delta H_{m,i}$ of phenol is 11,510.0 J/mol, which was taken from Inozemtsev et al. (1972)²⁰⁴. Activity coefficients were predicted using two activity coefficient models: NRTL model and COSMO-SAC (2010). First the activity coefficients were obtained considering a melting temperature sufficiently close to that of the optimal solution. Then, given those activity coefficients and the composition, the temperature was calculated using the Eq. 13 and minimized by the square of the absolute deviation between the experimental temperature and the calculated one.

SLE and LLE have been determined for phenol and n-hexane at atmospheric pressure. Results were correlated with two different models (i.e., NRTL, PR EoS) determining the activity coefficients. Moreover, all results were compared with the predictive model COSMO-SAC. Predictions of activity coefficient using COSMO-SAC can accurately represent SLE and. For LLE diagrams, COSMO-SAC shows a shift in the phenol-rich phase and seems to be a competitive model for representing the LLE diagram.

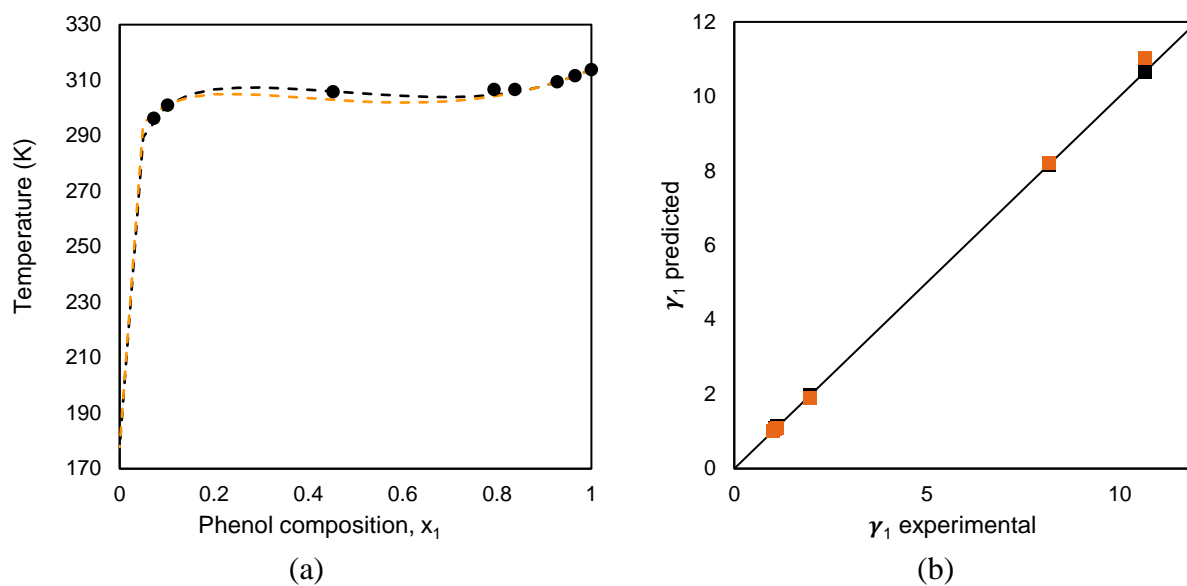


Figure A1 – (a) SLE and (b) activity coefficient of Phenol (1)/n-Hexane (2) at atmospheric pressure. (a) Black circle: experiments values (this work), Black dot-line: NRTL, Orange dot-line: COSMO-SAC. (b) black: NRTL and orange: COSMO-SAC (2010) and PSL sigma-profile (Klamt radii).

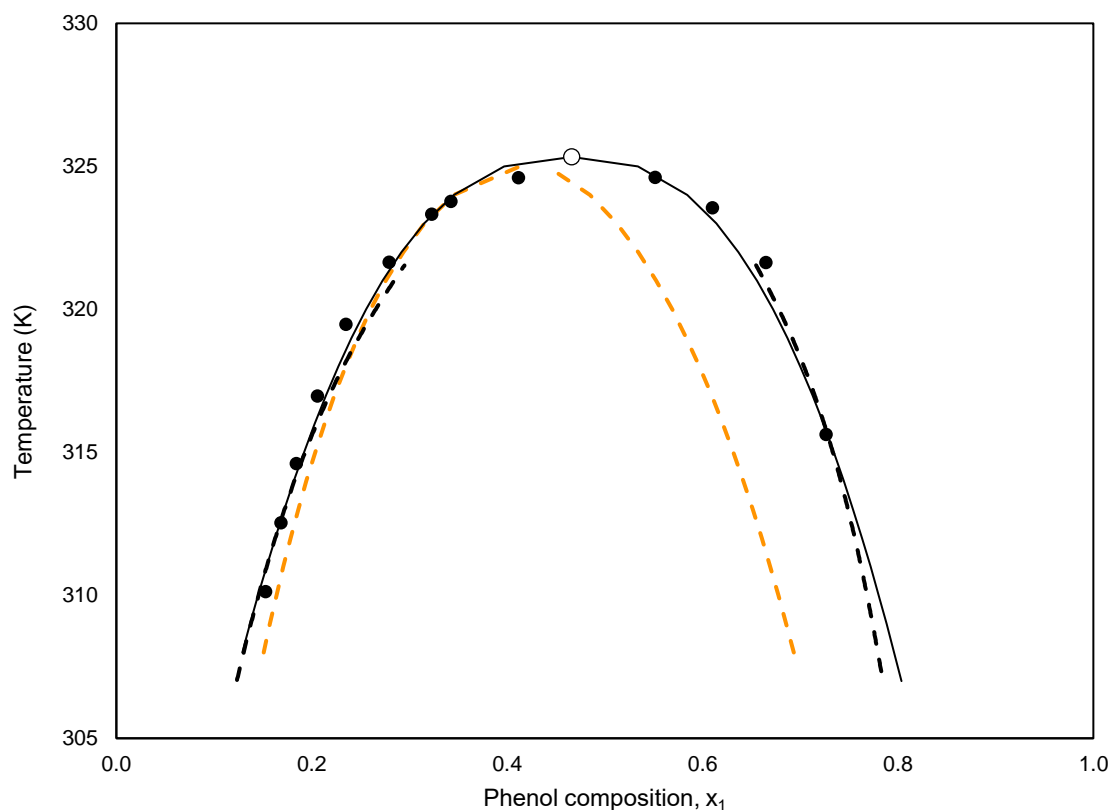


Figure A2 – LLE of Phenol (1)/n-Hexane at atmospheric pressure. Black circle: experiments values (this work), white circle: critical point, back line: Cox-Herrington, black dot-line: NRTL, Orange dot-line: COSMO-SAC.

Appendix B

Table B1 – NRTL parameters for $\alpha_{ij}^0 = 0.3$.

index des constituants		NP	T (K)	cij0	cji0	AAE%
o-xylene	n-heptane	23	348.1	260.6	1.6	0.16%
o-xylene	n-heptane	11	358.1	253.1	2.1	0.41%
o-xylene	n-heptane	12	348.1	247.7	0.0	0.62%
o-xylene	n-heptane	11	368.3	139.5	97.0	0.47%
o-xylene	n-octane	13	308.1	85.9	0.0	0.35%
p-xylene	n-heptane	28	348.1	197.5	0.0	0.33%
m-xylene	n-heptane	27	348.1	213.3	1.8	0.21%
m-xylene	n-decane	11	348.1	104.6	0.9	1.25%
m-xylene	n-decane	10	393.7	174.4	0.1	0.95%
m-xylene	n-decane	12	383.6	64.8	2.7	1.60%
toluene	n-heptane	49	298.1	297.5	16368.9	0.58%
toluene	n-heptane	44	303.1	280.6	23139.0	0.77%
toluene	n-heptane	25	313.1	293.0	19631.9	0.58%
toluene	n-heptane	15	323.0	220.9	23521.2	0.79%
toluene	n-heptane	15	333.0	280.8	54301.7	0.78%
toluene	n-heptane	14	363.0	209.3	27334.4	8.43%
toluene	n-heptane	17	348.0	259.2	28751.1	0.78%
toluene	n-heptane	13	303.0	244.1	21578.8	1.05%
toluene	n-octane	18	333.2	228.6	55996.8	0.35%
toluene	n-decane	12	373.5	21.0	26964.4	1.61%
toluene	n-decane	10	383.6	110.6	35129.2	2.47%
toluene	n-decane	10	393.7	495.3	244124.0	3.41%
benzene	n-hexane	32	333.1	349.8	50083.6	2.99%
benzene	n-hexane	7	343.1	306.0	20016.4	0.21%
benzene	n-heptane	15	333.1	330.8	40768.7	0.67%
benzene	n-heptane	26	353.1	281.6	52350.5	0.64%
benzene	n-heptane	8	348.1	283.2	54301.7	1.02%
benzene	n-heptane	22	293.1	419.3	27250.8	0.68%
benzene	n-heptane	9	488.1	133.0	48461.9	7.53%
benzene	n-heptane	10	458.1	110.2	54400.6	0.59%
benzene	n-heptane	10	473.1	132.8	56005.6	0.63%
benzene	n-heptane	10	443.1	145.1	45688.2	0.53%
benzene	n-heptane	9	428.1	165.0	45524.1	0.82%
benzene	n-heptane	9	413.1	248.7	48927.7	0.87%
benzene	n-heptane	17	383.1	281.2	165137.2	1.11%
benzene	n-heptane	15	318.1	343.7	33359.5	0.97%
benzene	n-octane	27	328.1	306.1	33345.4	1.49%
benzene	n-octane	13	348.1	243.7	33763.5	0.86%
benzene	n-octane	26	338.1	265.6	33352.7	1.51%
benzene	n-dodecane	11	333.1	38.7	33342.9	2.23%
benzene	n-dodecane	8	313.1	158.5	33342.4	2.50%
benzene	n-dodecane	16	353.1	0.0	1777806831378	2.52%

Appendix B

Table 4. VLE of Ethaline (1) and Ethanol (2).

T = 303 K									T = 323 K								
z2	P exp/bar	x2	y2 wilson	Pcal /bar Wilson	y2 nrtl	Pcal /bar NRTL	ERR% (P) Wilson	ERR% (P) NRTL	z2	P exp/bar	x2	y2 wilson	Pcal/bar Wilson	y2 nrtl	Pcal /bar NRTL	ERR % (P) Wilson	ERR % (P) NRTL
0.000	0.003	0.000	0.000	0.005	0.000	0.005	11.69	11.69	0.000	0.025	0.000	0.000	0.025	0.000	0.025	0.00	0.00
0.123	0.027	0.123	0.846	0.027	0.846	0.027	0.01	0.01	0.119	0.056	0.119	0.681	0.069	0.688	0.071	4.98	6.61
0.195	0.038	0.195	0.898	0.038	0.898	0.038	0.00	0.00	0.216	0.094	0.216	0.808	0.102	0.803	0.101	0.73	0.55
0.307	0.053	0.307	0.935	0.052	0.935	0.052	0.02	0.01	0.354	0.139	0.354	0.887	0.146	0.877	0.138	0.23	0.00
0.388	0.061	0.388	0.949	0.061	0.949	0.061	0.02	0.02	0.455	0.170	0.455	0.919	0.175	0.910	0.164	0.10	0.14
0.493	0.071	0.493	0.961	0.070	0.961	0.070	0.01	0.02	0.552	0.194	0.552	0.940	0.200	0.932	0.187	0.12	0.14
0.598	0.079	0.598	0.969	0.078	0.969	0.078	0.02	0.04	0.596	0.204	0.596	0.947	0.211	0.941	0.197	0.11	0.13
0.700	0.087	0.700	0.976	0.085	0.976	0.084	0.04	0.07	0.703	0.228	0.703	0.962	0.235	0.960	0.222	0.10	0.06
0.799	0.095	0.799	0.982	0.091	0.982	0.090	0.18	0.23	0.794	0.248	0.794	0.973	0.254	0.974	0.244	0.05	0.03
0.855	0.099	0.855	0.986	0.095	0.986	0.094	0.19	0.23	0.894	0.275	0.894	0.985	0.273	0.987	0.269	0.00	0.05
1.000	0.110	1.000	1.000	0.105	1.000	0.105	0.19	0.19	1.000	0.303	1.000	1.000	0.296	1.000	0.296	0.05	0.05
T = 313 K									T = 333 K								
0.000	0.007	0.000	0.000	0.004	0.000	0.004	14.26	14.26	0.000	0.012	0.000	0.000	0.007	0.000	0.0067	18.27	18.27
0.098	0.038	0.098	0.898	0.038	0.898	0.038	0.00	0.02	0.132	0.094	0.132	0.945	0.114	0.948	0.1083	4.34	2.16
0.204	0.067	0.204	0.948	0.067	0.948	0.067	0.01	0.00	0.253	0.167	0.253	0.972	0.180	0.970	0.1506	0.55	0.99
0.299	0.089	0.299	0.964	0.088	0.964	0.088	0.01	0.01	0.370	0.228	0.370	0.981	0.229	0.979	0.2040	0.00	1.12
0.398	0.106	0.398	0.973	0.105	0.973	0.105	0.01	0.01	0.437	0.259	0.437	0.985	0.254	0.983	0.2364	0.03	0.74
0.488	0.119	0.488	0.978	0.118	0.978	0.118	0.01	0.01	0.525	0.295	0.525	0.988	0.286	0.986	0.2746	0.09	0.48
0.593	0.133	0.593	0.982	0.131	0.982	0.131	0.02	0.02	0.599	0.324	0.599	0.990	0.312	0.989	0.3117	0.13	0.14
0.700	0.145	0.700	0.986	0.143	0.986	0.142	0.03	0.04	0.643	0.340	0.643	0.991	0.328	0.990	0.3483	0.13	0.06
0.799	0.157	0.799	0.990	0.153	0.990	0.153	0.06	0.08	0.750	0.380	0.750	0.993	0.367	0.993	0.3855	0.12	0.02
0.895	0.172	0.895	0.994	0.165	0.994	0.165	0.16	0.17	0.848	0.423	0.848	0.996	0.405	0.996	0.4078	0.19	0.14
1.000	0.182	1.000	1.000	0.180	1.000	0.180	0.01	0.01	1.000	0.480	1.000	1.000	0.443	1.000	0.470	0.006	0.000

u(T) = 0.02K u(z)=0.01 u(P)=0.002 bar

Table 5. VLE of Ethaline(1)+Isopropanol(2).

T = 303 K								
z2	P exp/bar	x2	y2 wilson	Pcal /bar Wilson	y2 nrtl	Pcal /bar NRTL	ERR% (P) Wilson	ERR% (P) NRTL
0.000	0.0026	0.009	0.000	0.0066	0.000	0.0026	0.0041	0.0000
0.078	0.0277	0.063	0.908	0.0765	0.887	0.0213	0.0488	0.0064
0.153	0.0423	0.126	0.945	0.1247	0.938	0.0361	0.0824	0.0062
0.206	0.0493	0.216	0.956	0.1510	0.953	0.0450	0.1018	0.0043
0.260	0.0545	0.225	0.963	0.1728	0.962	0.0527	0.1183	0.0017
0.307	0.0581	0.242	0.966	0.1883	0.967	0.0583	0.1302	0.0003
0.346	0.0606	0.269	0.969	0.1991	0.970	0.0623	0.1385	0.0017
0.397	0.0643	0.278	0.971	0.2112	0.973	0.0665	0.1469	0.0022
0.439	0.0666	0.289	0.973	0.2196	0.975	0.0693	0.1530	0.0026
0.495	0.0698	0.298	0.974	0.2294	0.977	0.0722	0.1595	0.0024
0.539	0.0727	0.307	0.976	0.2358	0.978	0.0738	0.1631	0.0010
0.595	0.0755	0.316	0.977	0.2431	0.979	0.0751	0.1676	0.0004
0.647	0.0765	0.322	0.978	0.2490	0.979	0.0757	0.1725	0.0008
0.701	0.0760	0.329	0.978	0.2547	0.979	0.0760	0.1787	0.0000
0.759	0.0770	0.337	0.979	0.2607	0.979	0.0760	0.1837	0.0010
0.818	0.0770	0.345	0.980	0.2670	0.980	0.0762	0.1900	0.0008
0.869	0.0780	0.394	0.981	0.2730	0.980	0.0764	0.1950	0.0016
1.000	0.0810	0.394	1.000	0.2960	1.000	0.0810	0.2150	0.0000

T = 313 K								
0.000	0.0042	0.000	0.000	0.004	0.000	0.0042	0.0000	0.0000
0.074	0.0439	0.073	0.909	0.043	0.888	0.0350	0.0007	0.0089
0.121	0.0611	0.119	0.938	0.061	0.928	0.0519	0.0003	0.0092
0.160	0.0725	0.158	0.950	0.073	0.944	0.0645	0.0000	0.0080
0.201	0.0820	0.199	0.957	0.083	0.955	0.0764	0.0006	0.0056
0.354	0.1030	0.352	0.970	0.106	0.972	0.1087	0.0032	0.0057
0.392	0.1068	0.391	0.971	0.110	0.974	0.1143	0.0032	0.0075
0.531	0.1195	0.531	0.975	0.119	0.979	0.1272	0.0000	0.0077
0.585	0.1248	0.585	0.977	0.122	0.980	0.1298	0.0028	0.0050
0.647	0.1317	0.647	0.978	0.125	0.981	0.1317	0.0070	0.0000
0.757	0.1375	0.757	0.980	0.129	0.981	0.1330	0.0089	0.0045
0.868	0.1405	0.868	0.984	0.133	0.983	0.1346	0.0075	0.0059
1.000	0.1420	1.000	1.000	0.140	1.000	0.1401	0.0019	0.0019

Table 5. VLE of Ethaline(1)+Isopropanol(2) (continuation).

T = 323 K								
0.000	0.0095	0.000	0.000	0.0095	0.009	0.0095	0.0000	0.0000
0.108	0.1016	0.100	0.914	0.1011	0.877	0.0705	0.0005	0.0311
0.209	0.1466	0.204	0.944	0.1466	0.120	0.1203	0.0000	0.0263
0.417	0.1799	0.413	0.960	0.1837	0.180	0.1797	0.0037	0.0003
0.520	0.1912	0.517	0.964	0.1932	0.196	0.1957	0.0021	0.0045
0.623	0.2012	0.621	0.967	0.2012	0.206	0.2061	0.0000	0.0049
0.686	0.2076	0.685	0.969	0.2060	0.211	0.2108	0.0016	0.0032
0.753	0.2153	0.752	0.972	0.2114	0.215	0.2153	0.0038	0.0000
0.801	0.2217	0.800	0.975	0.2157	0.219	0.2185	0.0060	0.0032
0.893	0.2354	0.893	0.982	0.2254	0.226	0.2265	0.0100	0.0089
1.000	0.2413	1.000	1.000	0.2380	1.000	0.2380	0.0034	0.0034
T = 333 K								
0.000	0.0087	0.000	0.000	0.0087	0.000	0.0087	0.0000	0.0000
0.036	0.0628	0.035	0.866	0.0628	0.819	0.0468	0.0000	0.0160
0.092	0.1296	0.089	0.936	0.1262	0.919	0.0986	0.0034	0.0310
0.150	0.1800	0.215	0.966	0.2165	0.963	0.1933	0.0365	0.0133
0.235	0.2269	0.232	0.968	0.2246	0.965	0.2033	0.0023	0.0236
0.278	0.2433	0.274	0.971	0.2424	0.970	0.2266	0.0009	0.0167
0.362	0.2683	0.359	0.975	0.2695	0.976	0.2644	0.0012	0.0039
0.395	0.2764	0.392	0.976	0.2779	0.978	0.2764	0.0015	0.0000
0.445	0.2874	0.443	0.978	0.2891	0.980	0.2922	0.0017	0.0048
0.493	0.2966	0.491	0.979	0.2980	0.981	0.3042	0.0014	0.0076
0.545	0.3069	0.544	0.980	0.3069	0.983	0.3153	0.0000	0.0084
0.603	0.3169	0.602	0.982	0.3156	0.984	0.3250	0.0013	0.0081
0.649	0.3255	0.649	0.983	0.3224	0.985	0.3316	0.0031	0.0061
0.692	0.3342	0.692	0.984	0.3288	0.986	0.3373	0.0054	0.0031
0.745	0.3462	0.745	0.985	0.3370	0.987	0.3440	0.0092	0.0022
0.800	0.3580	0.799	0.987	0.3450	0.988	0.3500	0.0130	0.0080
0.851	0.3650	1.002	1.000	0.3943	1.000	0.3943	0.0293	0.0293
0.905	0.3750	1.002	1.000	0.3945	1.000	0.3945	0.0195	0.0195
0.942	0.3800	1.002	1.000	0.3945	1.000	0.3945	0.0145	0.0145
1.000	0.3868	1.000	1.000	0.3936	1.000	0.3936	0.0068	0.0068

$u(T) = 0.02K$ $u(z)=0.01$ $u(P)=0.002$ bar

Table 6. VLE of Glycaline(1) + Ethanol(2).

T = 303 K								
z ₂	P exp bar	x ₂	y ₂ Wilson	Pcal /bar Wilson	y ₂ NRTL	Pcal /bar NRTL	ERR% Wilson	ERR% NRTL
0.000	0.000	0.000	0.000	0.0000	0.0000	0.0000	0.0000	0.0000
0.053	0.030	0.068	0.999	0.0275	0.9990	0.0274	0.0029	0.0030
0.092	0.042	0.116	0.999	0.0420	0.9994	0.0423	0.0004	0.0000
0.138	0.054	0.172	1.000	0.0548	0.9995	0.0556	0.0004	0.0013
0.196	0.066	0.241	1.000	0.0664	0.9996	0.0680	0.0001	0.0017
0.244	0.073	0.296	1.000	0.0733	0.9997	0.0746	0.0000	0.0012
0.294	0.078	0.352	1.000	0.0789	0.9997	0.0799	0.0006	0.0015
0.337	0.082	0.398	1.000	0.0825	0.9997	0.0831	0.0002	0.0008
0.393	0.086	0.458	1.000	0.0864	0.9997	0.0862	0.0000	0.0001
0.446	0.089	0.512	1.000	0.0890	0.9998	0.0881	0.0000	0.0009
0.530	0.093	0.595	1.000	0.0922	0.9998	0.0902	0.0006	0.0026
0.598	0.096	0.660	1.000	0.0940	0.9998	0.0914	0.0018	0.0045
0.649	0.098	0.707	1.000	0.0953	0.9998	0.0923	0.0031	0.0061
0.693	0.101	0.746	1.000	0.0962	0.9998	0.0931	0.0043	0.0074
0.742	0.102	0.790	1.000	0.0972	0.9998	0.0942	0.0043	0.0074
1.000	0.110	1.000	1.000	0.0160	1.0000	0.0160	0.0940	0.0940
T = 313 K								
0.000	0.003	0.000	0.000	0.0028	0.000	0.0028	0.0000	0.0000
0.066	0.048	0.083	0.948	0.0485	0.942	0.0440	0.0008	0.0037
0.126	0.078	0.156	0.969	0.0779	0.967	0.0731	0.0000	0.0048
0.189	0.101	0.231	0.978	0.1006	0.977	0.0973	0.0008	0.0041
0.228	0.114	0.304	0.982	0.1171	0.982	0.1157	0.0031	0.0017
0.303	0.128	0.361	0.984	0.1274	0.984	0.1272	0.0003	0.0005
0.330	0.132	0.390	0.985	0.1318	0.985	0.1321	0.0003	0.0000
0.380	0.139	0.444	0.986	0.1390	0.987	0.1399	0.0000	0.0009
0.427	0.144	0.492	0.987	0.1442	0.988	0.1454	0.0003	0.0015
0.485	0.149	0.551	0.988	0.1498	0.989	0.1509	0.0009	0.0020
0.537	0.153	0.602	0.989	0.1538	0.989	0.1545	0.0008	0.0015
0.581	0.156	0.643	0.989	0.1565	0.990	0.1568	0.0006	0.0009
0.679	0.163	0.734	0.991	0.1619	0.991	0.1611	0.0007	0.0015
0.724	0.166	0.774	0.991	0.1642	0.991	0.1630	0.0019	0.0031
0.791	0.172	0.832	0.992	0.1673	0.992	0.1658	0.0049	0.0064
0.903	0.179							
1.000	0.180							

Table 6. VLE of Glycaline(1) + Ethanol(2) (continuation).

T = 323 K								
z ₂	P		y ₂	Pcal /bar		ERR%		ERR%
	exp/bar	x ₂		Wilson	NRTL	Wilson	NRTL	
0.000	0.007	0.000	0.000	0.0066	0.000	0.0066	0.0000	0.0000
0.050	0.047	0.050	0.883	0.0540	0.866	0.0475	0.0070	0.0005
0.091	0.098	0.091	0.929	0.0858	0.921	0.0772	0.0118	0.0204
0.150	0.150	0.150	0.953	0.1231	0.950	0.1147	0.0269	0.0353
0.205	0.170	0.205	0.963	0.1506	0.962	0.1442	0.0194	0.0258
0.242	0.185	0.242	0.968	0.1660	0.967	0.1616	0.0193	0.0237
0.293	0.202	0.293	0.972	0.1838	0.972	0.1821	0.0184	0.0201
0.345	0.216	0.345	0.975	0.1987	0.976	0.1996	0.0169	0.0161
0.407	0.228	0.407	0.978	0.2133	0.979	0.2166	0.0150	0.0117
0.444	0.234	0.444	0.979	0.2206	0.980	0.2249	0.0134	0.0091
0.504	0.243	0.504	0.981	0.2308	0.982	0.2360	0.0123	0.0071
0.542	0.249	0.542	0.982	0.2362	0.983	0.2415	0.0126	0.0072
0.595	0.254	0.595	0.983	0.2430	0.984	0.2481	0.0107	0.0056
0.645	0.259	0.645	0.984	0.2487	0.985	0.2530	0.0104	0.0061
0.695	0.264	0.695	0.985	0.2540	0.985	0.2573	0.0099	0.0066
0.748	0.269	0.748	0.986	0.2595	0.986	0.2616	0.0099	0.0078
0.797	0.275	0.797	0.987	0.2646	0.987	0.2656	0.0105	0.0094
0.847	0.282	0.847	0.989	0.2703	0.989	0.2705	0.0120	0.0118
0.943	0.290	0.943	0.994	0.2840	0.993	0.2838	0.0064	0.0066
1.000	0.303	1.000	1.000	0.2960	1.000	0.2960	0.0065	0.0065

u(T) = 0.02K u(z)=0.01 u(P)=0.002 bar

Table 7. VLE of Glycaline (1)+ Isopropanol (2).

T = 303 K								
z2	P exp/bar	x2	y2	Pcal /bar		ERR%	ERR%	ERR%
				Wilson	y2			
0.000	0.000	0.000	0.000	0.0000	0.000	0.000	0.0002	0.0002
0.049	0.022	0.034	0.999	0.0203	0.998	0.016	0.0016	0.0061
0.100	0.043	0.124	0.999	0.0486	0.999	0.043	0.0054	0.0003
0.191	0.061	0.233	1.000	0.0626	1.000	0.061	0.0011	0.0000
0.269	0.068	0.323	1.000	0.0679	1.000	0.069	0.0000	0.0008
0.295	0.069	0.351	1.000	0.0691	1.000	0.070	0.0000	0.0011
0.350	0.071	0.411	1.000	0.0709	1.000	0.072	0.0002	0.0015
0.445	0.073	0.510	1.000	0.0727	1.000	0.073	0.0001	0.0006
0.497	0.073	0.562	1.000	0.0734	1.000	0.073	0.0000	0.0001
0.544	0.074	0.608	1.000	0.0739	1.000	0.073	0.0000	0.0007
0.595	0.074	0.657	1.000	0.0743	1.000	0.073	0.0001	0.0014
0.648	0.075	0.706	1.000	0.0747	1.000	0.073	0.0000	0.0021
0.700	0.075	0.752	1.000	0.0751	1.000	0.073	0.0004	0.0028
0.790	0.077	0.831	1.000	0.0762	1.000	0.074	0.0007	0.0032
0.850	0.078	0.881	1.000	0.0765	1.000	0.074	0.0011	0.0032
1.000	0.080	1.000	1.000	0.0793	1.000	0.079	0.0009	0.0009
T = 313 K								
0.000	0.0042	0.000	0.000	0.0028	0.000	0.0028	0.0015	0.0015
0.090	0.0704	0.111	0.967	0.0762	0.950	0.0495	0.0058	0.0209
0.192	0.1066	0.233	0.978	0.1066	0.975	0.0871	0.0000	0.0195
0.247	0.1157	0.296	0.980	0.1145	0.979	0.1015	0.0012	0.0141
0.300	0.1208	0.355	0.981	0.1195	0.982	0.1123	0.0012	0.0085
0.351	0.1238	0.411	0.982	0.1229	0.984	0.1202	0.0009	0.0036
0.396	0.1255	0.459	0.983	0.1251	0.985	0.1255	0.0004	0.0000
0.444	0.1268	0.509	0.983	0.1269	0.986	0.1295	0.0000	0.0027
0.491	0.1275	0.556	0.983	0.1283	0.987	0.1322	0.0008	0.0047
0.556	0.1288	0.619	0.984	0.1297	0.987	0.1342	0.0009	0.0054
0.601	0.1298	0.662	0.984	0.1306	0.987	0.1349	0.0008	0.0051
0.652	0.1308	0.709	0.985	0.1317	0.988	0.1350	0.0008	0.0042
0.697	0.1318	0.750	0.985	0.1324	0.988	0.1350	0.0006	0.0032
0.800	0.1352	0.839	0.986	0.1341	0.987	0.1350	0.0011	0.0002
0.850	0.1387	0.881	0.987	0.1354	0.988	0.1355	0.0033	0.0032
0.886	0.1397	0.910	0.988	0.1363	0.988	0.1359	0.0034	0.0038
1.000	0.1415	1.000	1.000	0.1418	1.000	0.1418	0.0003	0.0003

Table 7. VLE of Glycaline (1)+ Isopropanol (2) (continuation).

T = 323 K

0.000	0.007	0.000	0.000	0.0066	0.000	0.0066	0.0001	0.0001
0.050	0.065	0.124	0.951	0.1223	0.936	0.0921	0.0573	0.0271
0.101	0.122	0.353	0.973	0.1951	0.974	0.1879	0.0727	0.0655
0.150	0.150	0.471	0.976	0.2086	0.979	0.2119	0.0586	0.0619
0.200	0.175	0.579	0.978	0.2158	0.981	0.2229	0.0408	0.0479
0.299	0.203	0.652	0.979	0.2195	0.982	0.2262	0.0163	0.0230
0.409	0.212	0.747	0.980	0.2233	0.982	0.2273	0.0114	0.0154
0.515	0.216	0.835	0.982	0.2277	0.983	0.2288	0.0122	0.0132
0.591	0.218	0.879	0.983	0.2300	0.984	0.2300	0.0120	0.0120
0.695	0.223	0.909	0.985	0.2315	0.985	0.2311	0.0086	0.0082

RÉSUMÉ

Dans le contexte des bioraffineries, le choix d'un solvant approprié est crucial pour des processus de séparation durables et économiquement viables. Une approche globale intégrant des critères tels que l'Analyse du Cycle de Vie, l'analyse de la toxicité, la régénération énergétiquement efficace du solvant, des pertes minimales de solvant et une haute sélectivité est nécessaire. Cependant, le choix devient complexe lorsqu'il s'agit des solvants eutectiques profonds (Deep Eutectic Solvents - DES) en raison du manque de données expérimentales complètes, en particulier concernant les propriétés thermodynamiques et physiques critiques telles que l'équilibre de phase, la densité, la viscosité, la capacité thermique, la conductivité thermique, la solubilité, et autres. Pour combler cette lacune, cette recherche propose de générer des données expérimentales essentielles (telles que la densité, la viscosité et les propriétés d'équilibre de phase) pour optimiser les modèles théoriques. De plus, ce travail propose une approche de sélection de solvant basée sur la modélisation des propriétés thermodynamiques en utilisant le modèle COSMO-SAC (COnductor-like Screening MOdel segment activity coefficient). L'étude vise à améliorer le modèle COSMO-SAC en examinant les variables computationnelles, en établissant une base de données de profils sigma PSL et en affinant les prédictions grâce aux contributions enthalpiques, entropiques et intermoléculaires. Malgré les défis rencontrés dans la prédiction précise des coefficients d'activité en dilution infinie (IDAC) pour les systèmes DES, une approche d'optimisation réduit considérablement les écarts, offrant ainsi une voie prometteuse pour la sélection précise du solvant dans les processus de bioraffinerie.

MOTS CLÉS

Bioraffinerie, Modèle COSMO, Solvant Eutectique Profond, DFT.

ABSTRACT

In the context of biorefineries, selecting an appropriate solvent is crucial for sustainable and economically viable separation processes. A comprehensive approach integrating criteria like Life Cycle Assessment, toxicity analysis, energy-efficient solvent regeneration, minimal solvent losses, and high selectivity is required. However, the choice becomes challenging when considering Deep Eutectic Solvents (DES) due to a lack of comprehensive experimental data, particularly regarding critical thermodynamic and physical properties like phase equilibrium, density, viscosity, heat capacity, thermal conductivity, solubility, and more. To bridge this gap, this research proposes generating essential experimental data (such as density, viscosity, and phase equilibrium properties) to optimize theoretical models. Moreover, this work proposes a solvent screening approach based on modeling thermodynamic properties using the (COnductor-like Screening MOdel segment activity coefficient (COSMO-SAC) model. The study aims to enhance the COSMO-SAC model by investigating computational variables, establishing a PSL sigma-profile database, and refining predictions through enthalpic, entropic, and intermolecular contributions. Despite encountering challenges in accurately predicting activity coefficients at infinite dilution (IDAC) for DES systems, an optimization approach significantly reduces deviations, offering a promising route for precise solvent selection in biorefinery processes.

KEYWORDS

Biorefinery, COSMO Model, Deep Eutectic Solvent, DFT.

