

**Predictive Solvent and Anti-Solvent Selection Method for
Pharmaceutical and Biological Products and Intermediates.**

by

Suresh Ramsuroop

(MSc.Eng.)

Submitted in fulfilment of the academic requirements of

Doctor of Philosophy

in Chemical Engineering

College of Agriculture, Engineering and Science

University of KwaZulu-Natal

Durban

South Africa

(November 2017)

PREFACE

The research contained in this thesis was completed by the candidate while based in the Discipline of Chemical Engineering, School Engineering of the College of Agriculture, Engineering and Science, University of KwaZulu-Natal, Durban, South Africa.

The contents of this work have not been submitted in any form to another university and, except where the work of others is acknowledged in the text, the results reported are due to investigations by the candidate.

Signed:

Date:

DECLARATION 1: PLAGIARISM

I, Suresh Ramsuroop, declare that:

- (i) The research reported in this dissertation, except where otherwise indicated or acknowledged, is my original work;
- (ii) This dissertation has not been submitted in full or in part for any degree or examination to any other university;
- (iii) This dissertation does not contain other persons' data, pictures, graphs or other information, unless specifically acknowledged as being sourced from other persons;
- (iv) This dissertation does not contain other persons' writing, unless specifically acknowledged as being sourced from other researchers. Where other written sources have been quoted, then:
 - a) Their words have been re-written but the general information attributed to them has been referenced;
 - b) Where their exact words have been used, their writing has been placed inside quotation marks, and referenced;
- (v) Where I have used material for which publications followed, I have indicated in detail my role in the work;
- (vi) This dissertation is primarily a collection of material, prepared by myself, published as journal articles or presented as a poster and oral presentations at conferences. In some cases, additional material has been included;
- (vii) This dissertation does not contain text, graphics or tables copied and pasted from the Internet, unless specifically acknowledged, and the source being detailed in the dissertation and in the References sections.

Signed: Suresh Ramsuroop

Date: 06 December 2017

ABSTRACT

This thesis describes the development of a methodology that targets the synthesis and the operational design for crystallisation processes. The main objective of this work is to develop a solvent selection tool for crystallisation operations specifically targeting the ability to perform feasibility studies during the conceptual design stage of active pharmaceutical ingredient (API) production. In addition, this tool can be used for the optimisation and retrofit of existing API production processes. Rigorous optimization studies require reliable initial estimates to determine the optimum operating conditions. Therefore, there is a need for a simple and fast way to provide feasible and near optimum solutions. These solutions could then be used for screening design and operational alternatives, and eventually be used for further rigorous optimisation studies or pilot plant studies.

To achieve this, a comprehensive solvent selection modelling framework is developed. This framework is successfully interfaced into a commercial simulation software creating a master/slave architecture. This successful interfacing greatly increases the computational potential of the framework. It allows the integrated modules to access the vast database of compounds; an up-to-date selection of predictive thermodynamic models to determine pure and mixture properties, and robust Temperature-Pressure (T-P) and Enthalpy-Pressure (H-P) flash algorithms which are used as a basis for phase change determination. This interfacing has also extended the computational capabilities of the simulator by allowing the automated resetting of input parameters; creating a results database from multiple simulations, and managing the operations of the process simulator. The use of the robust and current predictive thermodynamic models within the simulator also has the potential to greatly improve the crystallisation models predictive ability and reduces the need for extensive experimental data.

The computational capabilities of the solvent selection tool developed include: predicting the various eutectic temperatures and compositions that may exist in the system, developing various types of phase diagrams and solubility curves, and identify separation barriers. Various operations such as heating, cooling, solvent addition, solvent removal, anti-solvent addition, and combinations can be studied to systematically evaluate process alternatives. Financial and environmental performance models have also been included into the computational tool to evaluate the process, economic and environmental performance of a selected process. A flexible user defined solvent ranking system the Normalised Cumulative Weighted Score (NCWS) is proposed in this work. It calculates the weighted cumulative score that accounts for: process performance, economic performance, environmental performance and energy performance by a selected solvent. The weighting of each of the performance criteria is defined by the user. The

algorithms are developed with the intention of providing a feasible and near optimum crystallisation processes. The computational framework developed is demonstrated in a series of applications to predict crystallisation behaviour and the determination of optimal operating conditions for cooling, evaporative and anti-solvent crystallisation.

Some experimental solubility measurements were undertaken to verify the co-solvent / anti-solvent behaviour predicted by the solvent selection tool in some systems. A strong variation of solubility, for the selected APIs Acetylsalicylic acid, 4-Acetaminophenol, and 2-(4-Isobutylphenyl)-propanoic acid, was observed in the ethanol and ethyl acetate binary systems. The experimental results show that the solvent selection tool gives a good qualitative representation of the solubility behaviour of co-solvency and anti-solvency over the binary solvent concentration range. The experimental measurements confirmed that both the UNIFAC and Modified UNIFAC (Do) models generally predict conservative values and are capable of predicting the general behaviour of complex systems.

The reliability and robustness of solvent selection tool was further evaluated against an industrial crystallisation processes. The application is the recovery of the natural flavourant 2,3-Butanedione (Diacetyl) from a process stream containing: acetone, acetaldehyde, ethanol, 2,3-pentanedione, 2,3 butanedione (Diacetyl) and water. The recovery of Diacetyl through distillation is not an option because of the several complex azeotropes that exist in the feed stream. Fractional crystallisation is the process used to recover the Diacetyl. The solvent selection tool was used to evaluate the crystallisation of the Diacetyl in the complex feed. The three thermodynamic models were used in the tool: UNIFAC, Modified UNIFAC (Dortmund), and the Scatchard-Hildebrand model. At the average conditions of feed composition of 63 % Diacetyl and a final cooling temperature of -19.4 °C, a predicted yield of 45.3, 0 and 78 % are predicted by the Mod. UNIFAC (Do), UNIFAC, and Scatchard-Hildebrand models, respectively, compared to the plant average yield of 50.6 % at these conditions. The solvent selection tool was further used to investigate existing plant deviations. The co-solvency effects of acetone and water and anti-solvency effects of acetaldehyde and ethanol were identified. It was confirmed with plant data (1200 batches) that the variations of concentrations of these components in the feed stream contributed significantly to the various plant deviations recorded.

The capability of the solvent selection tool is further illustrated as a conceptual design tool for crystallisation processes in the pharmaceutical industry. Using a solvent ranking system developed in this work, the Normalised Cumulative Weighted Score (NCWS) is used to identify and rank potential solvents that can be use in the API manufacturing process. In addition, the tool is used to evaluate which mode of crystallisation (cooling, evaporative, anti-solvent or

combinations) is the best process option to fulfil a particular production objectives. Lastly the tool is used to determine the best fractional crystallisation processing route for multicomponent API feed streams where some components are considered to be impurities or by-products of the process.

It is extensively demonstrated that the computational tool developed in this work, can be used as a conceptual design tool. It provides a simple and fast way to identify feasible, near optimum solutions that could be used for screening design and process alternatives based on a combination of performance criteria. In addition, it can be used for further rigorous optimisation studies, trouble-shooting existing crystallisation processes, and give direction to experimental and pilot plant studies.

.....
S Ramsuroop

.....
Prof D Ramjugernath

.....
Prof J Rarey

ACKNOWLEDGEMENTS

I would like to take this opportunity to acknowledge and thank those who made this work possible.

- First, my supervisors, Prof. J. Rarey and Prof. D. Ramjugernath for giving me this opportunity to engage in this dynamic field of chemical thermodynamics, and the support, ideas, guidance and motivation during the course of my research. It has been a pleasure and honour to be your research student and be part of the Thermodynamics Research Group.
- Pascal Böwer and Thomas Teusch, the exchange students from University of Oldenburg, for their friendship, advice and support, and for their invaluable assistance during the conceptual formulation and development of the computational framework.
- My fellow PhD students in the Thermodynamics Research Group at the University of KwaZulu-Natal, for their friendship, advice and support. In particular, Brian Satola for his invaluable assistance with the various challenges associated programming.
- My colleagues in the Department of Chemical Engineering at the Durban University of Technology for the continued interest, encouragement and support.
- The Management of the Durban University of Technology for their continued support and encouragement.
- The National Research Foundation (NRF) for the financial support for my sabbatical leave.
- On a personal note my wife Paran and son Jyestha for years of support, and the encouragement they gave me to fulfil my dreams.

Contents

PREFACE.....	2
DECLARATION 1: PLAGIARISM.....	3
ABSTRACT.....	4
ACKNOWLEDGEMENTS	7
LIST OF TABLES	11
LIST OF FIGURES	13
NOMENCLATURE.....	18
CHAPTER 1: INTRODUCTION	20
1.1 Introduction.....	20
1.2 Background.....	24
1.2.1. Overview of Crystallisation	24
1.2.2. Solubility Modeling	25
1.2.3. Synthesis, Modeling and Optimization of Crystallisation Processes	25
1.2.4. Solvent-Selection Methods for Crystallisation	27
1.3. Thesis Motivation	29
1.4. Aims and Contributions of this Thesis.....	29
1.5. Structure of the Thesis	31
CHAPTER 2: CONCEPTUAL DESIGN OF CRYSTALLISATION PROCESSES.....	33
2.1 Solubility and Crystallisation.....	33
2.2 Solubility Models.....	36
2.3 Physical processes during crystallisation.....	38
2.3.1 Nucleation	38
2.3.2. Crystal growth and dissolution	38
2.3.3. Agglomeration and Breakage	39
2.4. Modes of Crystallisation	39
2.4.1. Cooling crystallisation	41
2.4.2. Evaporation Crystallisation	41
2.4.3. Anti-solvent Crystallisation	42
2.4.4. Reactive Crystallisation	42
2.5. Synthesis and Analysis of Crystallisation Processes	42
CHAPTER 3: THERMODYNAMIC FRAMEWORK FOR MULTICOMPONENT, MULTIPHASE EQUILIBRIA.....	49
3.1 Use of Phase Diagrams to analyse crystallisation.....	49
3.2. Development of thermodynamic framework.....	50
3.2.1. Thermodynamic Preliminaries	51
3.2. Solid-liquid Equilibria	57

3.3. Activity Coefficient Models	62
3.3.1. Correlative Models	63
3.3.2. Predictive Methods:	65
3.3.2.1. Group Contribution Methods	65
3.3.2.2. NRTL-SAC	73
3.3.2.3. COSMO-RS and COSMO-SAC	76
3.3.2.4. Functional-Segment Activity Coefficient Model (F-SAC)	77
3.3.2.5. Universal Segment Activity Coefficient Model (UNISAC)	78
3.4 Performance of Predictive and hybrid models.....	80
3.5. Development of multiphase flash calculation equations	83
Concluding remarks.....	86
CHAPTER 4: COMPUTATIONAL METHODS	88
4.1. Introduction.....	88
4.2 Development of a Computational Framework	88
4.2.1. Software selection	88
4.2.2. Computational Framework	90
4.4.3. The VBA Module	92
4.3. Models and their Assumptions	96
4.3.1. Practices in the pharmaceutical industry	96
4.3.2 Solvent Database	98
4.3.3. Performance criteria and the Relevant Equations	98
4.4. Operating Conditions:	104
Concluding Remarks	106
CHAPTER 5: EXPERIMENTAL METHODS.....	108
1.1. Background	108
Analytical and Synthetic Methods	111
5.2. Experimental procedure	111
5.2.1. Samples Preparation	113
5.2.2. Solubility Measurement	113
CHAPTER 6: RESULTS AND DISCUSSION	115
6.1 Model Validation	115
6.1.1. Reliability of Predictive Activity Coefficient models:	116
6.1.2. Reliability to Predict the Performance of Industrial Crystallisation Processes: Production of Natural Flavourant 2,3 Butanedione (Diacetyl).	119
6.2. Case Studies.....	126
6.2.1. General Observations	128
6.2.2 Applications	129
6.2. Concluding remarks.....	157

CHAPTER 7: CONCLUSIONS AND RECOMMENDATIONS	159
CHAPTER 8: RECOMMENDATIONS FOR FUTURE WORK	162
REFERENCES.....	163
APPENDIX A: Alternate SLE Derivation.....	173
APPENDIX B: Operating Manual	175
B.1. Introduction.....	176
B.2. Simulation Preparation.....	177
B.3. Running the Programme	181
B.3.1. Menu 1	181
B.3.2. Menu 2	181
B.3.3. Menu 3: Temperature Menu	182
B.3.4. Menu 4: The Solvent/Anti-solvent Menu	185
APPENDIX C: Solvent Database	196
APPENDIX D: CHEMCAD Unifac Groups	199
APPENDIX E: Sample Code	206
APPENDIX F: Publication (submitted)	217
APPENDIX G: Experimental Solubility Measurements.....	237

LIST OF TABLES

Description	Page
Chapter 2	
Table 2.1. Methods of Supersaturation Generation.	40
Table 2.2. Impact of key basic operations in the crystallisation compositional space.	47
Chapter 3	
Table 3.1. Equilibrium Extrema for specified conditions.	51
Table 3.2. Partial Fugacity Coefficient Expressions for some Equations of State.	56
Table 3.3. Examples of Comparative Studies of Predictive Models for APIs solubility.	82
Chapter 4	
Table 4.1. Equipment Costing Parameters.	101
Table 4.2. Utility Costing Parameters.	102
Chapter 5	
Table 5.1. Chemicals used for ternary systems solubility measurements.	113
Chapter 6	
Table 6.1. Design Feed Composition to the Primary and Secondary Crystallisers	120
Table 6.2. Major Feed Component Properties.	121
Table 6.3. Analysis of Primary Crystalliser Plant Performance Meeting the Design Feed Composition.	122
Table 6.4. Analysis of Secondary Crystalliser Plant Performance Meeting the Design Feed Composition.	125
Table 6.5. Structure and Properties of 2-(4-Isobutylphenyl) Propanoic Acid (Ibuprofen).	131
Table 6.6. Number of Solvents that will Achieve 85 % API Crystal Yield Based on Feed Conditions and Available Cooling Capacity.	135
Table 6.7. Relevant Calculation Selection in the Anti-solvent Menu for Anti-Solvent Selection.	146

Table 6.8.	Summary of Results on the Compatibility and Performance of Shortlisted Anti-Solvents at Ambient Conditions.	147
Table 6.9.	Range of Cooling, Evaporation and Anti-Solvent Addition Conditions Simulated.	148
Table 6.10.	Examples of Operating Conditions for Operations With and Without Constraints.	153
Table 6.11.	Structure and Properties of APIs (NIST).	154
Table 6.12.	Operating Conditions for the Recovery of First API From the Binary Mixture Ranked on Lowest Operating Cost.	157
Appendix		
Table B.1.	Scheme to Create Component List for Simulation.	179
Table E.1.	VBA code showing a technical and semantic bridge between Excel, the VBA module and CHEMCAD.	206

LIST OF FIGURES

Description	Page
Chapter 1	
Figure 1.1. Flow diagram of typical pharmaceutical manufacturing process.	21
Figure 1.2. Typical downstream operations associated with crystallisation process.	22
Chapter 2	
Figure 2.1. Examples of variation in API solubility curves.	34
Figure 2.2. The various regions of interest in crystallisation.	35
Figure 2.3. The influence of supersaturation on growth and nucleation rates (adapted from Movers and Rousseau, 1987).	39
Figure 2.4. Solubility Diagram showing how the different modes of crystallisation influence supersaturation (adapted from Jones, 2002).	40
Figure 2.5. A hierarchy of models for process development (adapted from Ng and Wibowo, 2003).	43
Figure 2.6. Solid-Liquid Equilibrium (SLE) phase diagram with the poly-thermal projection of a system consisting 3 components.	45
Figure 2.7. Examples of phase diagrams for ternary systems, the movements in the compositional space and the resulting operating protocol.	48
Chapter 3	
Figure 3.1. Thermodynamic cycle for the sublimation process of a solid.	59
Figure 3.2. Enthalpy and Entropy changes during the sublimation process of a solid.	59
Figure 3.3. Typical iteration scheme of the Successive Substitution Method (SSM).	87
Chapter 4	
Figure 4.1. Components of the integrated computational framework.	91
Figure 4.2. An Example of A Typical Sub-Routine And Its Communication Pathways.	92

Figure 4.3.	Algorithm for the analysis of crystallisation to determine operating protocols.	93
Figure 4.4.	Computational capabilities of the VBA module developed.	95
Figure 4.5.	Decision Tree for the selection of crystallisation process.	106
 Chapter 5		
Figure 5.1.	Determination of the saturation concentration by isothermal (1) and polythermal (2) methods.	109
Figure 5.2.	Apparatus used for analytical measurements of solid-liquid equilibria.	110
Figure 5.3.	Experimental setup used for determining the solubility data by the polythermal solid-disappearance method.	110
Figure 5.4.	The Differential thermal analysis equipment used for solubility measurements.	112
 Chapter 6		
Figure 6.1.	Comparison of Predicted and Measured Solubility of Acetylsalicylic Acid in a Binary Solvent at 25 °C.	117
Figure 6.2.	Comparison of Predicted and Measured Solubility of 2-(4-Isobutylphenyl) Propionic Acid in a Binary Solvent at 25 °C.	117
Figure 6.3.	Comparison of Predicted and Measured Solubility of 4-Acetaminophenol in a Binary Solvent at 25 °C.	118
Figure 6.4.	Comparison of Predicted and Measured Solubility of 4-Acetaminophenol in Ethanol at 25 °C.	118
Figure 6.5.	Comparison of Predicted and Measured Solubility of 4-Acetaminophenol in Ethyl Acetate at 25 °C.	119
Figure 6.6.	Production Plant Data for Primary Crystalliser for 1200 Production Days.	121
Figure 6.7.	Comparison of Predicted and Actual Plant Performance of the Primary Crystalliser.	122
Figure 6.8.	Effect of Feed Component Composition on the Primary Crystalliser Performance.	124
Figure 6.9.	Production Plant Data for Primary Crystalliser for 1200 Production Days.	125

Figure 6.10.	Comparison of Predicted and Actual Plant Performance of the Primary Crystalliser.	126
Figure 6.11.	Solvent Performance When Cooled Down from NBP of Solvent to $T_f = 10\text{ }^{\circ}\text{C}$.	132
Figure 6.12.	Operational and Economic Performance of the Solvents cooled to $T_f = 10\text{ }^{\circ}\text{C}$.	133
Figure 6.13.	List of Solvents That Will Achieve at Least 85 % Yield of Pure API Crystals at $T_f = 10\text{ }^{\circ}\text{C}$.	134
Figure 6.14.	Fixed Annualised Cost of a 1000kg API/Day Production Plant With $T_f = 10\text{ }^{\circ}\text{C}$.	137
Figure 6.15.	The Operational Cost of 1000kg API/Day Production Using Suggested Solvent for $T_f = 10\text{ }^{\circ}\text{C}$.	137
Figure 6.16.	The Environmental (E) Factor of Suggested Pool of Solvents for $T_f = 10\text{ }^{\circ}\text{C}$.	138
Figure 6.17.	The Energy Consumption (E_c) Factor of Suggested Pool of Solvents for $T_f = 10\text{ }^{\circ}\text{C}$.	138
Figure 6.18.	The Fixed Capital Cost of 200 Ton/Yr Production Facility Associated with Selected Solvent for $T_f = 10\text{ }^{\circ}\text{C}$.	138
Figure 6.19.	Ranking of Solvents based on equal weightings of yield, FAC, E-Factor and E_c -Factor for $T_f = 10\text{ }^{\circ}\text{C}$.	139
Figure 6.20.	Yield – Cooling Temperature Profiles of the Shortlisted Solvents.	140
Figure 6.21.	Effect of Cooling Capacity on the Fixed Annualised Cost and Yield.	141
Figure 6.22.	Effect of Yield on Waste Treatment Cost.	141
Figure 6.23.	Effect of Yield on Solvent Replacement Cost.	141
Figure 6.24.	Effect of Yield on Unrecovered API.	141
Figure 6.25.	Effect of Yield on Solvent Recovery Steam Cost	141
Figure 6.26.	Effect of Yield on Cooling Cost.	141
Figure 6.27.	Effect of Yield on E_c Factor.	142
Figure 6.28.	Effect of Yield on E Factor.	142
Figure 6.29.	Effect of Cooling on Crystallisation Yield.	144
Figure 6.30.	Effect of Evaporation and Cooling on Crystal Yield with no Anti-solvent.	145
Figure 6.31.	Effect of Anti-solvent Addition and Cooling on Crystal Yield with no Evaporation.	145
Figure 6.32.	Identification of Good Solvents and Anti-Solvents at $T_i = 50\text{ }^{\circ}\text{C}$.	147

Figure 6.33.	Addition Curves for Selected Anti-solvents at Ambient Conditions.	147
Figure 6.34.	The Effect of Anti-Solvent Addition and Cooling on Yield.	148
Figure 6.35.	Potential Operating Conditions to Achieve a Minimum Yield of 98 % Pure API Crystals.	149
Figure 6.36.	Band of Operating Conditions for the Combination of Cooling and Anti-solvent Addition to Achieve a Minimum Yield of 98 % Pure API Crystals with 10 % Evaporation.	150
Figure 6.37.	Effect of Cooling and Anti-Solvent Addition on the Fixed Annualised Cost.	152
Figure 6.38.	Effect of Cooling on the Crystallisation of a Binary API Mixture.	155
Figure 6.39.	Potential Operating Conditions to Achieve a Minimum Yield of 98 % Pure API Crystals.	156
Figure 6.40.	Band of Operating Conditions for the Combination of Cooling and Anti-solvent Addition to Achieve a Minimum Yield of 98 % Pure API Crystals With 10 % Evaporation.	156

Appendix B

Figure B.1.	View of CHEMCAD File with list of sub-routines.	177
Figure B.2.	Example of Simulation Component List (Output from CHEMCAD into Excel).	180
Figure B.3.	Start-menu.	181
Figure B.4.	Number of components.	181
Figure B.5.	Feed composition Input Requirements.	181
Figure B.6.	Temperature Menu.	182
Figure B.7.	Typical results obtained from various functions of the Temperature Menu.	183
Figure B.8.	Matrix of Yield results for “Run Maximum Combination” function.	184
Figure B.9.	Matrix of Total Operating Cost (\$) results for “Run Maximum Combination” function.	184
Figure B.10.	Results for the calculation of the solubility in different solvents function as an example.	186
Figure B.11.	Results showing Process, Economic and Environmental Performance for shortlisted solvents obtained from the solubility in different solvents function.	187
Figure B.12.	GUI to calculate ternary diagram.	188

Figure B.13.	Resulting diagram in Excel for Benzene, Cyclohexane and 2-Butanone as an example.	189
Figure B.14.	Results of an example calculation for the Addition Curve function.	190
Figure B.15.	Results of an example calculation for the Composition Curve function.	190
Figure B.16	Examples of typical results for Anti-solvent Functions menu.	193
Figure B.17	Extract of the results of the calculation of all possible combinations.	194
Figure B.18.	Extract of the results of the Economic and Environmental Performance calculations of all possible combinations	194
Figure B.19.	Extract of the filtered results for a minimum yield of 98% of pure API crystals	195

NOMENCLATURE

English Letters	
A	Van der Waals group surface area
A	Helmholtz energy
a	Activity
a	Interaction parameter
B	Viral coefficient
c	Concentration
c_p	Heat capacity
d	Differential or change
E	Environmental factor
E_c	Energy consumption factor
f	Fugacity
G	Gibbs energy
H	Enthalpy
K	Equilibrium constant
N	Number of equations
n	Number of moles [mole or kmole]
N_s	Number of solid phases
P	Pressure [kPa]
P_c	Critical pressure
P_r	Reduced pressure
Q	Group area parameter
q	Component surface area parameter
R	Group volume parameter
R	Universal gas constant
r	Component volume parameter
S	Entropy
S	Supersaturation ratio
T	Temperature [$^{\circ}\text{C}$ or K or $^{\circ}\text{R}$]
T_c	Critical temperature
T_m	Melting point
T_r	Reduced temperature
U	Internal Energy [J/mol]
V	Van der Waals group volume
V	Volume [m^3]
X	Group fraction
x	Mole fraction in the liquid phase
y	Mole fraction in the vapour phase
Z	Compressibility Factor
z	Co-ordination number
Δc	Supersaturation
ΔH^f	Heat of fusion

Greek Letters	
γ	Activity coefficient
γ^{∞}	Activity coefficient at infinite dilution
Γ	Activity coefficient of group
ϕ	Average segment fraction of component

δ	Charge density
μ	Chemical potential
α	Constant
φ	Fugacity coefficient
Θ	Group area fraction
Ψ	Group interaction parameter
λ	Interaction parameter
τ	Interaction parameter
ν	Number of groups
∂	Partial differential
θ	Surface area fraction
Λ	Wilson parameter

Subscripts and Postscripts	
<i>bp</i>	Boiling point
<i>cw</i>	Cooling water
<i>deg</i>	Degradation point
<i>E</i>	Excess
<i>ex</i>	Excess
<i>f</i>	Final
<i>flash</i>	Flash point
<i>i</i>	Initial
<i>i</i>	Component i
<i>i, j, k</i>	Components
<i>l, s, v</i>	Liquid, solid, and vapour phase
<i>mp</i>	Melting point
<i>mimp</i>	Melting point of impurity
<i>msol</i>	Melting point of solvent
(1), (2) ..., π	Number of phases
1, 2 ..., m	Number of species
0	Reference state
<i>sat</i>	Saturated

Abbreviations	
API	Active pharmaceutical ingredient
ASOG	Analytical Solution of Groups Method
COSMO	Conductor-like Screening Model
COSMO-SAC	Conductor-like Screening Model- Segment Activity Coefficient
EOS	Equation of State
F-SAC	Functional segment activity coefficient model
MOSCED	Modified separation of cohesive Energy Density Model
NRTL	Non-random two liquid model
NRTL-SAC	Non-random two liquid – segment activity coefficient model
NTAC	Normalised total annualised cost
SLE	Solid liquid equilibria
SLVE	Solid liquid vapour equilibria
TAC	Total annualised cost
UNIFAC	UNIQUAC functional group method
UNIQUAC	Universal quasi-chemical model
UNISAC	Universal segment activity coefficient model

CHAPTER 1: INTRODUCTION

A key criterion for success in the manufacture of active pharmaceutical ingredients (APIs) is their recovery with the desired crystal morphology, at high purities and yields. To meet these criteria, the preferred unit operation in the pharmaceutical industry is a crystallisation process. Hence, this endeavour seeks to devise a conceptual process that incorporates decision making tools for the development, design and operation of an API crystallisation process.

In this chapter, the context and motivation for the research is presented, and its main contributions are highlighted. The thesis is structured in the order of progression of the key elements used to meet the research objectives, as outlined at the end of the chapter.

1.1 Introduction

During the development phase of a new process, one of the key issues to be established is whether the desired production rates and product specifications will be achieved during the scale up from lab trials to commercial scale. In addition, at a commercial scale it is necessary to ensure the optimal conditions for operational benefit savings, and minimum environmental impact. For these purposes, and in order to ensure design viability, a reliable and precise computational tool is required to simulate the necessary processes.

Particularly during the development phase, an effective computational model facilitates technology transfer and process fitting. It evaluates process alternatives, facilitates process adjustments, and determines the range of optimum processing conditions from the data that is available. It can also be used to identify what additional laboratory and pilot plant trials are required to obtain sufficient data, and to confirm the accuracy of operational conditions and constraints.

In the pharmaceutical industry, the process of developing a new drug compound is lengthy, costly, risky, and extraordinarily complex. This process, which begins with synthesizing the raw drug substance on a laboratory scale, followed by animal testing, human testing, obtaining regulatory approval for clinical use, to full production, could span a decade. Any further changes to the process chemistry or operations, after regulatory approval, may lead to further lengthy trails, resulting in additional financial costs (Crafts, 2007). In this highly competitive, investment intensive, and tightly regulated industry, the primary concerns are reducing the product development time and associated costs, obtaining the production throughput, and optimum use of limited facilities and resources. The application of computational tools can meet these concerns and yield profound economic benefits.

Once reliable computational tools are constructed, key pharmaceutical processes and unit operations can be investigated without the need for extensive experimentation, or disruption to existing operations. These tools can be used at all stages of process development, from conceptual design, through process operation and optimization. Simply going through the process of developing a model can deliver deeper understanding, and critical insights into an issue or problem, and can help to improve decision making. In addition, potential constraints can be challenged, significantly improving the likelihood of an innovative solution, or providing confirmation that a tested solution is correct. An accurate model could not be devised without a thorough understanding of the manufacture of APIs. An overview of the important production considerations will, therefore, be provided. The objective of any pharmaceutical procedure is to ensure a technically efficient and cost effective process. Typically, pharmaceutical production - as shown in Figure 1.1 - is a batch process, consisting of successive reaction/synthesis stages to create the desired API from a range of raw materials and reaction enhancers, followed by a series of separation and purification stages, intended to separate and recover the API at its desired purity. As a result, the performance of the different stages (reaction, separation, and waste disposal) should all be optimized and coordinated to minimize costs and operational time, thereby, maximizing profit (am Ende, 2011).

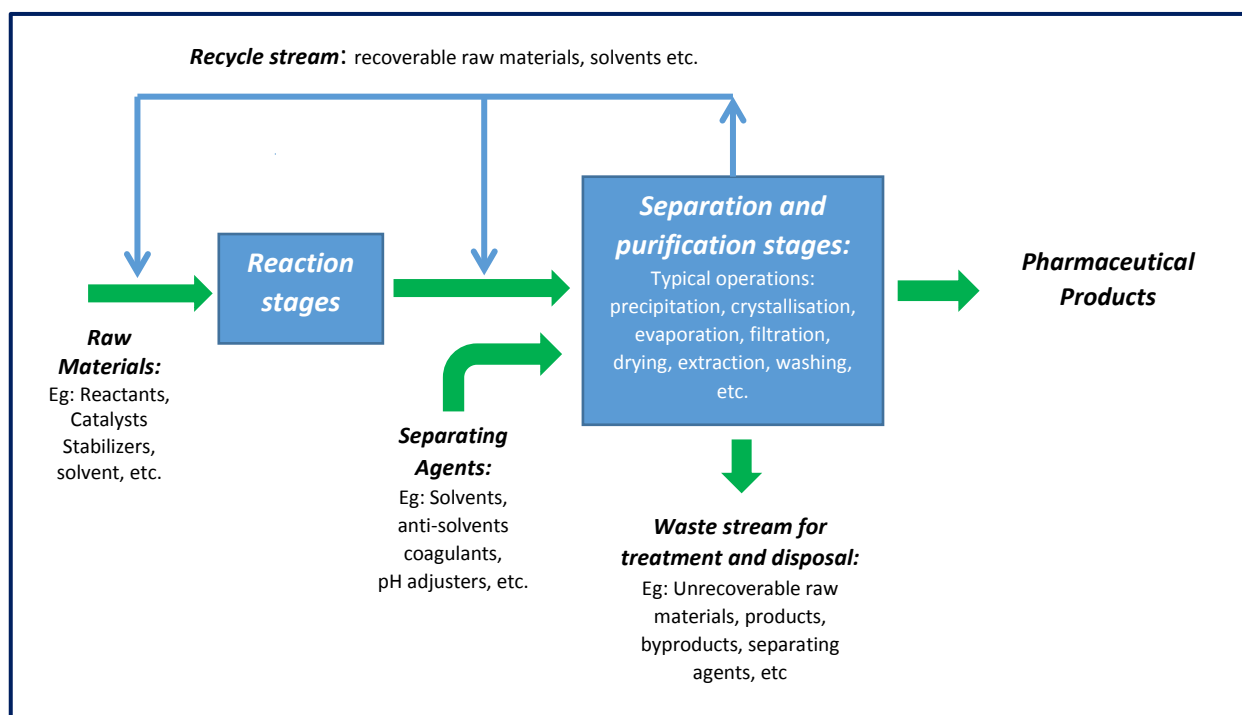


Figure 1.1. Flow diagram of typical pharmaceutical manufacturing process.

The production of pharmaceuticals is generally characterized by a high raw material consumption to API production ratio. The amount of non-product streams can range from 25 kg, to over 100 kg, per kg of desired product (Constable et al., 2007; Sheldon, 2005). Solvents constitute the bulk

of raw materials consumed. Solvents are used in many applications, with their primary functions being: a) as reaction media to facilitate process such as selectivity, solubility and restriction of heat dissipation; b) as reactants; c) as transportation agents, to facilitate the movement of reactants and products around the process, in solid or dissolved state; and d) as separating agents, to facilitate the extraction, separation and recovery of the desired components.

While solvents have a wide range of applications, many solvents raise varying degrees of environmental, health and safety concerns. Hence, the type and amount of the solvents used can contribute significantly to the initial capital investment, the production costs, and the waste treatment disposal costs. Therefore, there is an emphasis on developing cleaner technologies, with lower environmental impact, at competitive manufacturing costs.

A key unit operation, in the production of APIs, is their recovery from the process stream. Crystallisation is the principal technique used to selectively separate out the API as a solid of high purity (Tung , 2009). The three main advantages of using crystallisation for separation are: the production of a high purity product in one process step; with a comparatively low level of energy consumption; and with relatively mild process conditions. The typical operations associated the crystallisation process are shown in Figure 1.2.

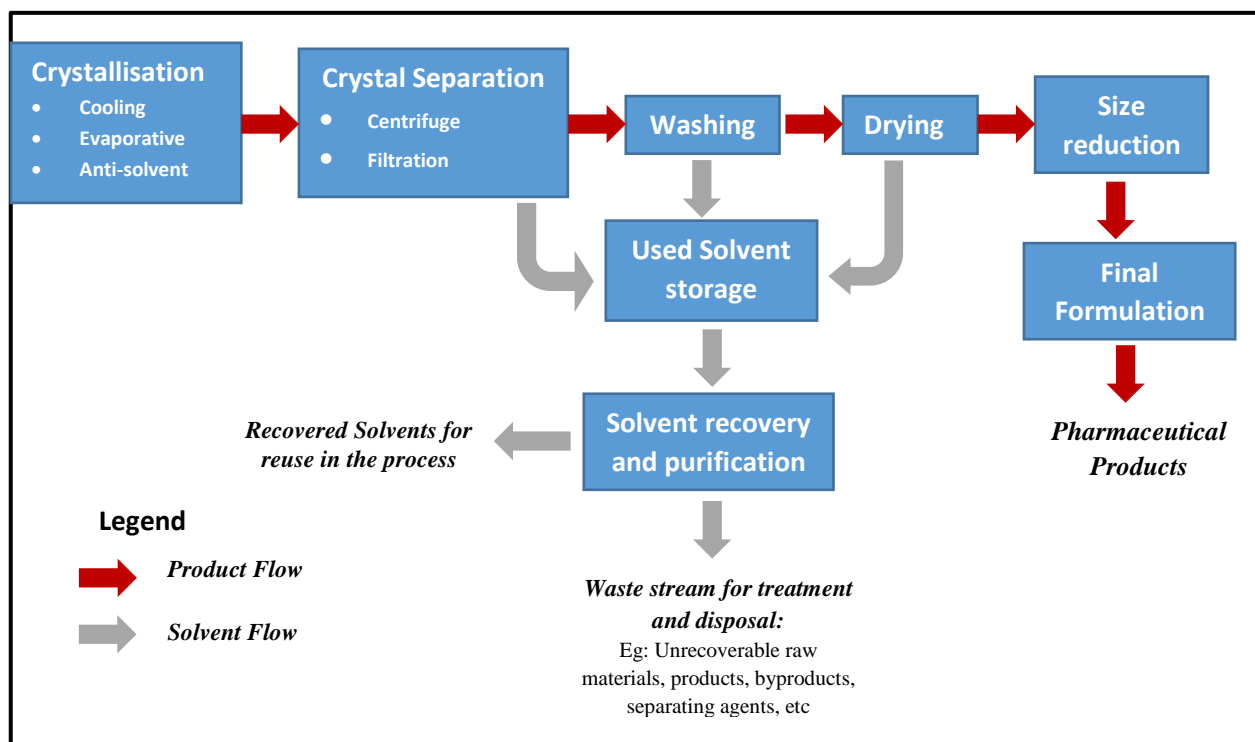


Figure 1.2. Typical downstream operations associated with crystallisation process.

A complexity to be considered in the crystallisation process of pharmaceutical products is the potential presence of stereoisomers, other by-products and impurities, which may also crystallize out with the required product (Tung, 2009; Wibowo and Ng, 2000; Wibowo and Ng, 2002). Hence, a crystallisation process (inclusive of solvent/co-solvent /anti-solvent selection, and resolving agents), for systems with multiple saturation points, may require a series of crystallisation operations, and the manipulation of several variables, to achieve the desired yield and purity of the product. The crystallisation process is also governed by many competing kinetic effects like primary and secondary nucleation, crystal growth, agglomeration, etc. which warrant sophisticated control strategies and long processing times to ensure that the desired, yield, quality and crystal morphology of the API is obtained. Systems with polymorphs may require recrystallisation stages with high levels of control to obtain the desired crystal structure. Notwithstanding these complexities, crystallisation is still the desired process option and forms one of the key operations in API manufacture. It directly impacts the amount and types of solvents that will be used in the manufacture of pharmaceuticals.

Solvent selection is of special importance in process design, since different solvent systems induce different yields, processing volumes, and different downstream processing requirements, for solvent recovery and waste stream treatment (Buxton et al., 1999; Constable, 2007; Karunanithi et al., 2007; Papadopoulos and Linke, 2005; Pistikopoulos and Stefanis, 1998). The ideal goal is to select solvents that result in minimal operational and capital costs, while offering optimal performance in terms of yield of the API, recovery and reuse of solvents, and conformity with regulatory constraints. Hence, an optimal solvent should simultaneously satisfy the objectives of selectivity and solubility, safety constraints and economic and environmental criteria. An added benefit would be to have a minimum inventory of solvents used in the process.

Variankaval et al., (2008) listed the following criteria, in order of priority, to be applied in the development of a process to crystallize API's:

1. A sufficient product purity to meet established quality standards.
2. An isolation of the chosen crystal form, which is typically (with very few exceptions) the most thermodynamically stable form.
3. A specific target particle size distribution (PSD) and crystal shape, as these may affect both bioavailability and processability
4. A high-yield.
5. A good volume productivity, with final slurry concentrations typically targeted for 10 ± 5 wt. %.

6. A reasonable cycle time (generally, 24 h) for the crystallisation, as well as for the associated filtration and drying processes.

Criteria (1) is critical to ensure patient safety, while (2) and (3) are frequently required, depending on their impact on chemical stability and bioavailability. Criteria (2) and (3) can be easily achieved through recrystallization processes, if the initial crystallisation process meets the product requirements. Criteria (4) to (6) have an impact on the plant size, operational costs and efficiency, especially for high-volume drugs.

For the development of a computational tool that models crystallisation, several key issues must be fully reviewed. These include the current process of crystallisation, its thermodynamic limitations, and recent developments in API crystallisation.

1.2 Background

1.2.1. Overview of Crystallisation

The driving force for crystallisation is the difference in chemical potential in the liquid and solid phases. Since chemical potential is hard to measure for systems with solutes, its formulation in terms of species concentrations is successfully used (Mullin, 1993). Supersaturation, which occurs when the solute concentration in solution exceeds the equilibrium concentration, may give rise to crystal formation and gives rise to crystal growth. There are two approaches that can be used to generate the supersaturation required for crystal formation: temperature swing and concentration swing (Wibowo and Ng, 2000). The temperature swing, generally referred to as cooling crystallisation is dependent on there being a significant temperature dependence of the solid-liquid phase equilibrium, and is the preferred mode for API crystallisation. Several studies have evaluated the effects of cooling on yield, purity and morphology. They all, generally, concluded that the rate of cooling greatly influences the quality and quantity of API crystals obtained. It was found that the rate of cooling should be such that the degree of supersaturation is maintained within a narrow zone above the equilibrium solubility, which is referred to as the Meta Stable Zone Width (MSZW). (Jones, 2002; Mullin, 1993; Mersmann, 2001; Tung, 2009).

While the temperature swing approach is the preferred mode for API crystallisation, if the API is temperature sensitive, or if the API's equilibrium solubility does not change significantly with temperature, then a concentration swing technique is required. The two concentration swing techniques that can create the required conditions for supersaturation are, either partial evaporation of the solvent, or the addition of a solvent that decreases the solubility of the API in the solvent. Solvents that decrease solubility are referred to as anti-solvents, and the crystallisation process is called anti-solvent crystallisation. Seeding can also be used to initiate the crystallisation process. By providing a seed of the appropriate type, shape or size, a template for crystallisation

is provide. The primary role of seeding is to promote secondary nucleation and crystal growth, and is generally used to control crystallinity, particle size distribution, purity and a specific polymorph.

Evaporation crystallisation, whilst extensively used in the commodity chemicals industry, has had limited application in the crystallisation of APIs. This has been mainly due to the lack of control of crystal morphology during evaporation processes. However, evaporation coupled with cooling crystallisation, or anti-solvent crystallisation, or a combination of all three modes, can be used to significantly increase the yield of API crystals (Tung et al., 2009).

1.2.2. Solubility Modeling

To calculate the degree of supersaturation, which is linked to the driving force for crystallisation operations, the solubility at different conditions is needed. If the solubility data already exist then a solubility model can be derived. This can either be done empirically or the data can be used to fit the binary interaction parameters of correlative thermodynamic models such as Wilsons, NRTL, UNIQUAC, etc. (Walas, 1986). However, if the solubility data is not known, then the solubility will need to be either measured or estimated. For solvent selection processes, experimental measurements with various solvents, co-solvents and anti-solvent combinations at various conditions can be impractical in terms of cost and experimental time required. The experimental effort for determination of a full phase diagram rapidly increases upon increasing the number of components. Hence, reliable predictive tools are a necessity in the initial screening and selection process. Some of main predictive methods include UNIFAC (Grensemann, 2005.), NRTL-SAC (Chen and Song, 2004), COSMO-RS (Klamt, 2005), COSMO-SAC (Lin and Sandler, 2002); PC-SAFT (Kliener et al., 2009), (F-SAC) (Soares and Gerber, 2013) and UNISAC (Moodley et al. 2015). Several studies have compared solubility results from these predictive techniques to experimental measurements with varying degrees of accuracy when predicting the equilibrium solubility of APIs (Ruether and Sadowski, 2009; Widenski et al., 2010; Sheikholeslamzadeh and Rohani, 2011; Moodley et al. 2015).

1.2.3. Synthesis, Modeling and Optimization of Crystallisation Processes

Literature on crystallisation shows that “synthesis”, “modelling” and “optimization” have three distinct objectives. Synthesis means finding the optimal scheme to be used for the crystallisation in multi-solute systems that require fractional crystallisation. Modelling primarily deals with the simultaneous solution of the population balance models, mass balance and energy balance equations to determine the effect of the process conditions on the particle size distribution (PSD). Optimization primarily focuses on determining the cooling profile / anti-solvent addition profile / seeding profile, to promote crystal growth to maximize yield.

There are two major approaches for the synthesis of crystallisation-based separation. In one approach, phase equilibrium diagrams are used for the identification of separation schemes. With the aid of equilibrium phase diagrams, the effect of some basic operations in crystallisation results in specific movements within the compositional space. These movement vectors are used to determine operations that can be effectively used to achieve the desired targets. While these procedures are easy to understand, and relatively simple to implement, the graphical representation become increasingly complex with the number of solutes present. Several researchers (Fitch, 1970; Berry et al., 1997; Wibowo and Ng, 2000; 2002; Cisternas, 2006) have developed rules and guidelines for synthesising operational protocols, for crystallisation-based separation processes, using this approach. The second strategy is based on the construction of a network flow model to represent the set of potential separation flowsheet structures. The basic idea in this approach is to derive a network superstructure that has embedded all feasible configurations of separation sequences. The overall network is formulated as a nonlinear programming problem, and the optimal separation sequence is obtained by simultaneous optimization using non-linear mathematical programming algorithms (Cisternas et al. 1999, 1998 and 2004).

Chemical engineering modelling processes are designed to ascertain the variation of the key variables, to better understand the behaviour of processes. But, for several reasons, crystallisation modelling is more complicated than that of many other chemical engineering unit operations. First, crystallisation often operates at an unsteady state due to its batch nature. Second, many different types of phenomena, such as primary and secondary nucleation, homogenous and heterogeneous nucleation, crystal growth (which may be through diffusional or surface mechanisms), agglomeration, and attrition, occur simultaneously, and many of these processes are affected by the degree of supersaturation present (Tung, 2009, Jones, 2002; Mullin, 1993).

The primary objective of most modeling studies is to determine the particle size distributions (PSD) obtainable for a given set of operating conditions. The PSD is established by solving the population balance equations that are used to model particulate systems. Establishing the rate constants required to quantify the various phenomena involves making several simplifying assumptions. This then limits the use of these models when scaling-up for commercialization. Despite this limitation, modeling does provide valuable insights into variables that can be manipulated, to achieve desired targets, during the initial investigations.

Optimization studies are a natural extension of modeling studies. Their primary objective is to determine a profile in order to minimize undesirable phenomena (such as primary nucleation), and to maximize desired phenomena, that of crystal growth. It could be a cooling profile, or a

seeding profile, or an anti-solvent addition profile, or a combination of profiles. Some of the typical objective functions used include: maximization of volume mean size, minimization of total nucleation, and specified final volume mean size.

Optimization studies in crystallisation provide operational and control strategies for crystallisation processes. Several optimization studies have shown that the optimal operating strategy is to ensure that the degree of supersaturation is maintained within a narrow zone above equilibrium solubility, which is referred to as the Meta Stable Zone Width (MSZW), (Tung et al., 1999; Sarker et al., 2006; Nowee et al., 2007, Nagy et al., 2008 and Widenski et al.2010).

1.2.4. Solvent-Selection Methods for Crystallisation

Rajgopal et al., (1992) developed systematic procedure for the conceptual design of vapor-liquid-solid processes. It is based on the hierarchical procedures for conceptual design proposed by Douglas (1988). In an evolutionary manner, it leads the user to select process units, to identify the equipment configurations and to determine the important design variables and the associated economic trade-offs. Design variables unique to the generated configuration calculated include: dilution ratio, wash ratio and crystallizer temperature and their effect on overall plant economics. Whilst this procedure was not designed to identify and rank solvents for crystallisation processes, it has the essential components to evaluate the impact of a selected solvent will have on the processing requirements and the subsequent economic impact. It can handle only low-molecular-weight solid particles.

Nass (1994) describes a strategy for choosing crystallization solvents based on equilibrium limits. The solvent selection strategy for choosing crystallization solvents is based upon the determination of equilibrium limits for a given temperature range. The approach utilizes a group-contribution method (UNIFAC) to predict a value for the activity coefficient of the solute in a selected solvent system at the saturation point. This value is then used to calculate the solubility of the solute at a "high" temperature and a "low" temperature. The resulting solubility values determine the maximum theoretical yield for the process which is used to rank order solvents. The solvent selection was limited to single solvent systems and cooling crystallisation applications.

Frank et al.(1999), reviewed strategies for solvent selection for various types of crystallization processes such as cooling crystallization and anti-solvent crystallization. The methods reviewed include the use of Robbin Charts to identify general classes of solvents; Radius of interaction analysis using Hansen solubility parameters to identify potential solvents, and UNIFAC to estimate activity coefficients and solubility. Potential recoveries based on solubility calculations

formed the bases of ranking the solvents. In both of the publications (Nass K. , 1994) (Frank, 1999), the solvent selection task is carried out from a known database of good solvents with solubility being the only criterion for selection.

Karunanithi (2006; 2008) presented a computer-aided molecular design (CAMD) framework for the design and selection of solvents and anti-solvents for crystallization processes. The CAMD technique provides a systematic methodology for identifying/designing a functional chemical to fulfill required/desired properties or performance. Ng et al. (2006) defined the CAMD methodology as: “Given the specifications of a desired product, determine the molecular structures of the chemicals that satisfy the desired product specifications, or, determine the mixtures that satisfy the desired product specifications.” The solvent-selection problem is defined as a CAMD problem, and is formulated as a mixed-integer nonlinear programming model (MINLP). In the CAMD approach for solvent design and selection for crystallization processes other factors can be considered simultaneously in addition to solubility. Many other desired solvent properties such flash point, toxicity, viscosity, normal boiling point, and normal melting point, etc. can also be considered because these properties can be posed as constraints in the MINLP formulation. In the CAMD formulation all the required properties have been estimated using group-contribution-based methods. Solvent design and selection for cooling crystallization and anti-solvent crystallization were presented for the crystallisation of the API ibuprofen. For both case studies the performance of the designed solvents are verified qualitatively through SLE diagrams. The only limitations to the general application to this approach is the challenges faced in solving the complex MINLP problems. Most applications of solvent design use proprietary code developed at Technical University of Denmark (DTU).

Sheikholeslamzadeh et al. (2012) proposed several algorithms for predicting the phase behavior, miscibility testing, and screening of solvents for the crystallization of pharmaceutical components. The SLE behaviour of different APIs was predicted in many common solvents and combination of solvents, using the NRTL–SAC model. The feasible operating temperature range for each crystallization case was calculated. The maximum operating temperature is based on the bubble-point temperature of the solvent mixture and the lower operating temperature is based on the melting point of the model molecule. The batch cooling and anti-solvent crystallization processes was simulated for seven model molecules from the initial temperature to the final temperature and for the different compositions of each solvent. In addition, The NRTL–SAC model was also used to test the miscibility of solvents during the crystallization process. The solvent ranking is based solely on the yield obtained. The hybrid model used requires experimental solubility data for the parameterization of the models prior to its predictive

capability. This limits the methodology as a fast screening conceptual design tool in the absence of any experimental data.

1.3. Thesis Motivation

This project sets out to develop a robust and comprehensive Solvent Selection computational framework to be applied to API crystallisation production processes. To be effective, it must be able to screen solvents, to synthesize operating protocols for fractional crystallisation, and to determine near optimal operating conditions for different types of crystallisation processes.

A review of developments in this research field suggests that this project is well-placed to fill an existing gap. There is evidence of a lack of a framework that can comparatively evaluate the process performance, the economic performance, energy requirement performance, and the environmental performance, of solvent selection for API crystallisation processes, as can this computational framework.

As described earlier, crystallisation is an extremely important unit operation in the pharmaceutical industry. The performance of this separation technique is judged by certain criteria: yield, purity of product and whether the product is of the desired crystal structure and size. These measures are dependent on a multitude of variables, and the selection of the conditions for the optimization of crystallisation is dependent, in turn, on reliable process models. The parameters in these crystallisation models traditionally require extensive experimental data that are complex to obtain and analyse, if they are not already published in the literature. A way to reduce this experimental burden is via first-principles thermodynamic modelling.

First-principles thermodynamic modelling begins by developing predictive thermodynamic solubility models, since the principle phenomena in crystallisation are solubility and degree of supersaturation. Robust predictive solubility models must enable the prediction of the solubility of a solute in pure or mixed solvents, and of the degree of supersaturation at varying temperature and concentration conditions.

1.4. Aims and Contributions of this Thesis

In order to increase expertise in model-based, optimal strategies for crystallisation operations, specifically targeting the pharmaceutical industry, a robust and reliable, model-based Solvent Selection Tool for crystallisation is developed that can predict and optimize the production of crystalline API materials, with the desired yield and purity. This is addressed through the following key outputs:

1. The development and implementation of a comprehensive and coherent framework for modeling crystallisation systems. The modeling framework creates opportunities, not only for finding optimal operating strategies, but also to investigate and develop a comprehensive understanding of the crystallisation process. Specifically, the modeling framework has the following computational capabilities:
 - The optimum operability conditions for the crystallizer can be identified with minimal thermodynamic information on the system, by using multiphase flash calculations. Starting with just the chemical structure, melting point and enthalpy of fusion of the API, and using the predictive thermodynamic property models like UNIFAC, the Solid-Liquid-Vapour Equilibrium (SLVE) phase behaviour can be calculated.
 - The various eutectic temperatures and compositions that exist in the system can be predicted, and data can be generated for developing various types of phase diagrams and solubility curves. These allow a study of the overall composition space, visualize crystallisation regions, and identify the separation barriers to be examined. In particular, the analysis of systems with multiple saturation points is possible.
 - The study of operations such as heating, cooling, solvent addition, and solvent removal, to systematically evaluate process alternatives. This tool enables the user to filter and screen solvents, and evaluate the effects of co-solvents, anti-solvents, other components, and impurities on the solute's solubility, in a specified temperature and composition range.
 - The identification of the operating conditions that give maximum recovery of a desired compound, with a certain solvent or solvent mixture, and the calculation of the percent recoveries, and the total energy requirements (heating/cooling), under various operating conditions. The tool can be used to establish operating strategies, which may involve a combination of "cooling/heating", "co-solvent/anti-solvent addition", and "evaporation" steps to meet the process objectives.
 - The investigation of the process engineering implications of the various solvents. Once the desired production rate is established, the effect each solvent will have on the size of plant required, can be determined, along with the capital expenditure, operational expenses, and environmental impact.
 - The successful interfacing of the developed crystallisation computational tool with a commercial simulation software to improve its robustness and extend its computational capabilities. The computations within the Crystallisation module have access to a full range of thermodynamics models and correlations, a comprehensive database of

compounds - their pure and mixture properties - and rigorous computational algorithms for process calculations and equipment design of the commercial simulator.

2. Experimental work is conducted to validate the simulated optimization results. In particular the behavior of co-solvent and anti-solvent systems.
3. The developed computational tool is also tested and applied to an industrial application that recovers a natural flavourant compound from a complex multicomponent feed-stream.

These computational capabilities allow the modelling framework to be used as a quick screening tool to select the most appropriate solvent(s) and type of crystallisation process/processes for a given application. The modelling framework explores the synergistic combination of multiphase, phase equilibria phenomena, and process systems engineering methods, to develop a decision tool for faster process design and process understanding, during the conceptual phase of an API development, or the retrofit of an existing API manufacturing process. Furthermore, it will greatly reduce the need for experimental solubility data, as compared to the empirical approaches currently used in crystallisation modelling.

1.5. Structure of the Thesis

The thesis is structured as follows:

The first chapter puts into context the motivation for undertaking this research, and highlights the contribution it makes to understanding the impact that solvent selection has, on the economic and processing requirements, in the manufacture of APIs. It also outlines the structure of the thesis.

The second chapter presents the key crystallisation phenomena, including solubility and crystallisation mechanisms. The different types of crystallisation processes are discussed, and the key principles used in the analysis and synthesis of crystallisation processes are presented.

The third chapter presents the key thermodynamic principles, and the development of a thermodynamic framework for multiphase solubility equilibria, which is required for the analysis and synthesis of crystallisation processes. Various predictive phase equilibria models, which can be used for solubility prediction for APIs, are reviewed.

The fourth chapter presents the key features and outlines the development of the crystallisation computational tool, developed during this research work.

The fifth chapter outlines the various methods that can be used for measurement of solubility, and the experimental procedures adopted by this work are presented

The sixth chapter introduces several case studies that reveal the use of the computational framework described in chapters 2, 3 and 4, to determine the optimal operating conditions for evaporative, cooling, and anti-solvent crystallisation.

The seventh chapter presents the conclusions of the dissertation,

The eighth chapter recommends possible future work.

CHAPTER 2: CONCEPTUAL DESIGN OF CRYSTALLISATION PROCESSES

The goal of conceptual design is to identify process flowsheets, equipment sizes and configurations, together with operating conditions that will best meet the desired objectives. To this end, Doherty and Malone (2001) identified two major tasks in conceptual design, namely, (1) to identify feasible flowsheets and ranges of operating conditions capable of meeting process goals and, (2) to rank the alternatives according to economic measures to select candidates for more detailed study. These principles have been adopted in this study.

Before this can be done, in the development of any process, its various physical and chemical events need to be understood and accounted for. Crystallisation is a phase equilibria process, and hence the extent of the solid phase formation is mainly based on thermodynamic behaviour. The key concepts that play important role in the kinetics, throughput and yield of crystallisation processes include: solubility, degree of supersaturation and metastable zones,

In this chapter, an overview of pharmaceutical crystallisation process design will be provided, with a focus on three fundamental aspects, namely, the mechanisms of solubility and crystallisation; the main approaches to its operation and control; and a review of the available tools for analysis and synthesis.

2.1 Solubility and Crystallisation

It is essential, for the design of crystallisation processes, to first understand solubility behaviour. Analysis of crystallisation processes typically require that the solubility of the solute in the solvent be known as a function of temperature or solvent composition (Tavare, 1995). Solubility is defined as the maximum amount of a solid that can be dissolved in a particular solvent (or solvent mixture) at a given temperature and pressure, also known as the equilibrium or saturation concentration. The solubility data can be either obtained through experimental measurements or predicted from solubility models.

The equilibria between the solid and liquid phases (solid–liquid equilibria (SLE)) of a compound, are the thermodynamic foundation of all crystallisation processes. Solubility is typically expressed as a solubility curve whereby the temperature dependency can be demonstrated. Examples of typical solubility curves for diverse APIs in various solvents are shown in Figure 2.1. Data, measured by Hahnenkamp et al. (2010), illustrate that solubility can show a high temperature dependency (aspirin-ethanol), or a negligible temperature dependency (paracetamol-water). They further illustrate that an API (paracetamol) may reveal different solubility curves in

various solvents; and that distinct APIs may have different solubility curves in the same solvent (paracetamol and aspirin in acetone).

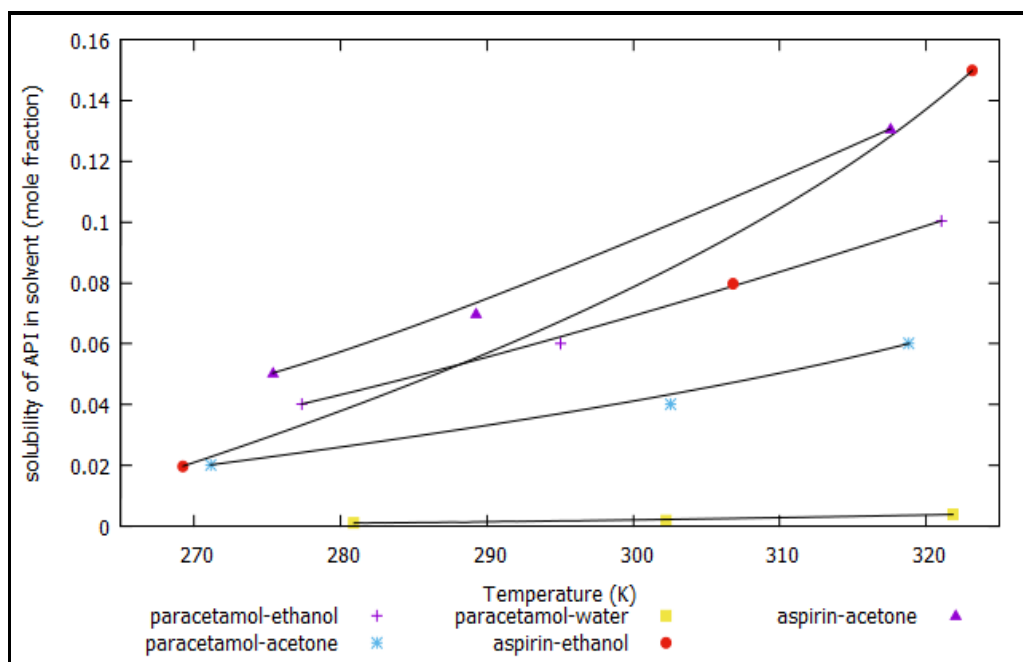


Figure 2.1. Examples of variation in API solubility curves.

A yield of crystals arises, primarily, from the difference in equilibrium solubility of a solute, in a solvent, at the initial and final operating conditions. The two key variables that affect solubility and control crystallisation processes are temperature and solvent composition (Mullin, 2001; Myerson, 2002). Solubility is also a function of pressure as shown in Equation 2.4, but the effect is generally negligible in the systems normally encountered in crystallization from solution (Mullin, 2001; Mersmann, 2001). The mechanism for crystallisation is illustrated in a typical solubility diagram, as shown in Figure 2.2. In this Figure, the temperature – composition space is divided into two regions by the solubility curve. The regions below and above the curve are termed under-saturated and supersaturated, respectively, and indicate the relative amount of dissolved solid, as compared to the saturated solution. In the supersaturated region, where the solute concentration exceeds the saturation concentration, the solution is unstable because the dissolved solid and solvent are not in equilibrium.

Like all non-equilibrium systems, the supersaturated solution inclines toward equilibrium by removing the solute from the solution in the form of nuclei, which then grow into crystals. Supersaturation is therefore a prerequisite step in crystallisation. This region above the solubility curve is further divided into two: a metastable region and a labile region. A labile, or unstable, area promotes the spontaneous and intense onset of nuclei. But, the uncontrolled nuclei formation

in this region leads to a decrease in solute available, and does not enhance crystal growth. On the other hand, the solution is supersaturated in the metastable zone, but the supersaturation is not sufficient for the crystals to appear spontaneously. Because a crystal can grow in this zone, the growth of seed crystals, of the same nature as the solute, can be promoted, limiting the appearance of additional nuclei. Hence the metastable zone is the desired operational zone in crystallisation processes.

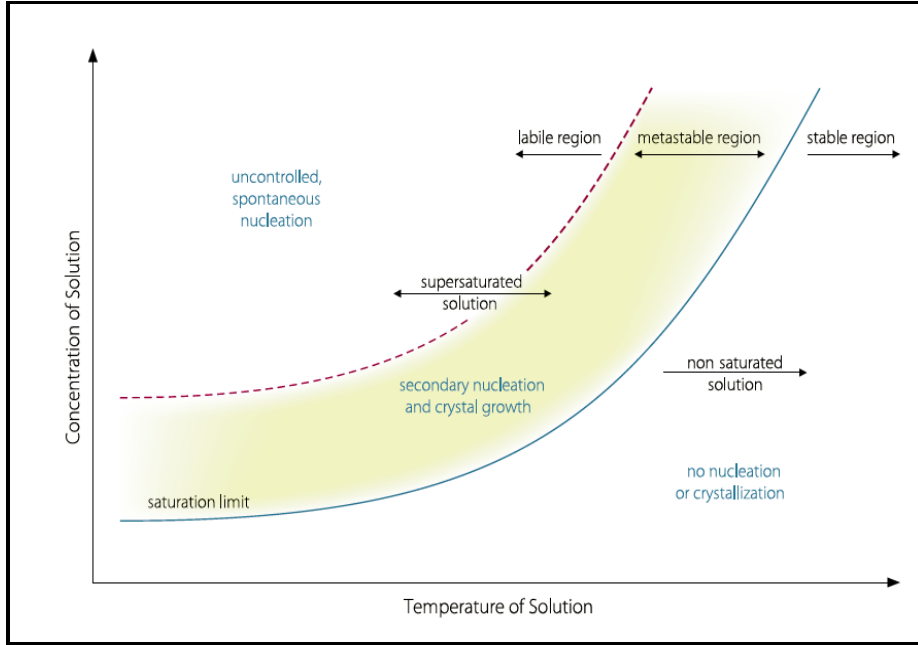


Figure 2.2. The various regions of interest in crystallisation (www.solvias.com).

Below the solubility curve, the stable area is a sub-area where the solution has a concentration less than the solubility limit. So there is no crystallisation, and the addition of crystals will lead to their dissolution.

The thermodynamic driving force for crystallisation is the difference in the chemical potential between the solute and the solution, and may be expressed as:

$$\Delta\mu_i = \mu_{i,solution} - \mu_i^* \quad 2.1$$

$$\Delta\mu_i = RT\ln(\mu_{i,solution}) - RT\ln(\mu_i^*) \quad 2.2$$

$$\exp\left(\frac{\Delta\mu_i}{RT}\right) = \frac{f_l}{f_s} = \frac{\gamma_i c}{\gamma^* c^*} = S \quad 2.3$$

Where S is defined as the relative supersaturation; $\Delta\mu_i$ is the difference in chemical potential; f_l and f_s are the fugacity of the liquid and solid phase respectively; c^* and γ^* are the concentration and activity coefficient at equilibrium; and c and γ are the actual concentration and activity coefficient of the solution.

Common expressions of supersaturation are shown by the equations presented below, where solubility and concentration units are adjusted for consistency:

$$\begin{aligned}\text{Supersaturation:} \quad \Delta c &= c - c^* \\ \text{Supersaturation ratio:} \quad S &= \frac{c}{c^*} \\ \text{Relative supersaturation:} \quad \sigma &= \frac{c - c^*}{c^*} = S - 1\end{aligned}$$

2.2 Solubility Models

Solubility data, as shown in Figure 2.1, can either be obtained through experimental measurements or through prediction by means of solubility models. Solubility models should accurately predict how the equilibrium solute concentration changes over the course of the crystallisation process. This understanding of the process of solubility is required for the development of a crystallisation model to, in turn, precisely predict crystal product properties, such as size, quality and yield. Solubility models can be based on either empirical or thermodynamic foundations. Empirical solubility models such as the Jouyban-Acree model (Jouyban and Acree, 2006) and Yalkowsky-Roseman model (Yalkowsky and Roseman, 1981) are equations fitted to experimental solubility data, and generally have no underlying theoretical foundation, while on the other hand, a thermodynamic solubility model both fits the data and is derived from thermodynamic principles.

The equation derived from the fundamental principles of thermodynamics which is used to determine the solubility of a solid solute (s) in a liquid solvent (derived in chapter 3 (Section 3.3)), is given by the following expression, where it is assumed that the solid phase is pure and that the triple point temperature can be replaced by the melting temperature:

$$-\ln x_i^l \gamma_i^l = \ln \left(\frac{f_i^l}{f_i^s} \right)_{\text{pure } i} = \frac{\Delta h_{m,i}}{RT} \left(1 - \frac{T}{T_{m,i}} \right) - \frac{\Delta c_{p,i}(T_{m,i} - T)}{RT} + \frac{\Delta c_{p,i}}{R} \ln \left(\frac{T_{m,i}}{T} \right) + \frac{\Delta V_i(P - P_t)}{RT}$$

2.4 Generally for systems where the pressure difference and change in heat capacities are small, the second, third and fourth terms on the right-hand-side are negligible in comparison to the first term, and the above expression is generally simplified to:

$$-\ln x_i^l \gamma_i^l = \frac{\Delta h_{m,i}}{RT} \left(1 - \frac{T}{T_{m,i}} \right) \quad \text{or} \quad x_i^l = \frac{1}{\gamma_i^l} \exp \left[\frac{-\Delta h_{m,i}}{RT} \left(1 - \frac{T}{T_{m,i}} \right) \right] \quad 2.5$$

Where x_i^l is the solubility; γ_i^l is the activity coefficient; $\Delta h_{m,i}$ is the heat of fusion; and $T_{m,i}$ is the melting point of component i .

In the above model, the values for the activity coefficient, heat of fusion and melting point can either be obtained experimentally or through predictive methods. It should be noted that the estimation of heat of fusion and melting point is generally not recommended.

Common types of thermodynamic models used to calculate the activity coefficient include those based on excess Gibbs energy, such as, Wilson, NRTL, or UNIQUAC (Walas, 1985). However, these correlative models require identification of binary interaction parameters from phase equilibrium data for each of the solvent–solvent, solvent–solute, and solute–solute binary mixtures. Unfortunately, the lack of experimental solubility data, on new APIs, limits the use of these thermodynamic models for process design and analysis in the pharmaceutical industry. (Chen, 2011).

There has been an increase in the use of predictive thermodynamic models for solubility calculations (as distinct from correlative models). The advantage of these predictive models is that no new experimental data is needed to calculate activity coefficients. The determination of the activity coefficient is essential, as it accounts for any non-ideal behaviour of the solution by accounting for the interactions that exist between the various molecules.

The various predictive methods that can be used to calculate the activity coefficient can be classified into 4 main groupings:

- Models based on group contribution methods, such as the UNIQUAC Functional Group Activity Coefficient (UNIFAC) and its many derivatives, such as modified UNIFAC (Grensemann and Gmehling, 2005), Pharma Mod-UNIFAC (Diedrichs and Gmehling, 2010) and UNISAC (Moodley et al. 2015)
- Models based on quantum theory: Conductor-like Screening Model, abbreviated as COSMO, and its derivatives the COSMO-RS (Klamt, 2005) and COSMO-SAC (Lin and Sandler, 2002);
- Models based on equation of state PC-SAFT (Kliener et al., 2009) and;
- Hybrid-data estimation methods, such as the Non-Random Two-Liquid - Segment Activity Coefficient (NRTL-SAC) (Chen and Song, 2004).

These predictive models and their ability to predict the solubility of APIs are fully described in chapter 3.

2.3 Physical processes during crystallisation

Whilst the particle size distribution (PSD) is not modelled or determined in this work, it is important to understand the variables that impact on the PSD, to ensure that any proposed operational strategy take these into account. Below, is an overview of the most important physical processes occurring during crystallisation. These concepts aid in an understanding of the recommended strategies for optimization studies in crystallisation. For a more in-depth description of these phenomena, the reader is referred to several of the texts available: Beckmann (2013), Tung et al. (2009), Jones (2002), Mersmann (2001), Myerson (2001) and Tarave (1995).

2.3.1 Nucleation

Nucleation is the starting point of crystal formation. This may occur through one of the following mechanisms

- *Primary nucleation* is the formation of a new solid phase from a clear liquid. This type of nucleation can be further subdivided into homogeneous and heterogeneous nucleation. In heterogeneous nucleation, nucleation starts on foreign substrates of mostly microscopic particles, dust or dirt particles. If such substrates are absent, new phase formation takes place by statistical fluctuations of solute entities clustering together, a mechanism referred to as homogeneous primary nucleation. It requires very high supersaturation conditions, as in the labile zone shown in Figure 2.2. (Jones, 2002 and Beckmann, 2013).
- *Secondary nucleation* is induced only when previously crystallized material is available. This nucleation mechanism generally occurs at much lower supersaturations than in primary nucleation. There are various types of secondary nucleation, but the most important source of secondary nuclei in crystallisation is contact nucleation, and occurs as a result of crystal collisions (Mullin, 2001).

2.3.2. Crystal growth and dissolution

Crystal growth is a desired phenomenon in crystallisation and it results from the addition of more solute molecules to the nucleation site or crystal lattice. Besides increasing crystal size, crystal growth also largely determines the key qualities of the crystal: crystal morphology, surface structure and purity of the crystal. Crystal growth is a three-step process, consisting of mass transfer, surface integration and heat transfer. Mass transfer and surface integration occur sequentially and in parallel with heat transfer. Mass transfer involves the diffusion of growth units (molecules, atoms or ions) to the crystal surface. Surface integration consists of surface diffusion, orientation and the actual incorporation into the lattice. The latent heat of crystallisation is released and transported to the crystal and solution.

2.3.3. Agglomeration and Breakage

An agglomerate is defined as the mass formed by the cementation of individual particles, probably by inter-particle forces during the collision of particles. Agglomerates are usually undesirable because they contain mother liquor between the primary crystals that form the agglomerate. This liquor is hard to remove during drying, and promotes caking of the product during storage. Breakage, as the fracture of a particle into one slightly smaller particle and many much smaller fragments, is defined as attrition. Breakage involves the fracture of a particle into two or more pieces.

Control strategies for crystallisation are primarily used to determine whether nucleation or growth should be the dominant process, depending on which of these process objectives is most critical for the desired overall outcome. The demand for increasing control of physical attributes, for final bulk pharmaceuticals, has necessitated a shift in emphasis from control of nucleation to control of growth (Tung , 2009). Both nucleation and growth are dependent on the degree of supersaturation, and hence, maintaining the degree of supersaturation within the metastable zone is crucial. This desired zone of operation is shown in Figure 2.3, which shows how the primary, secondary and growth rates vary with supersaturation. In the next section we will examine the various ways of generating and maintaining the required supersaturation.

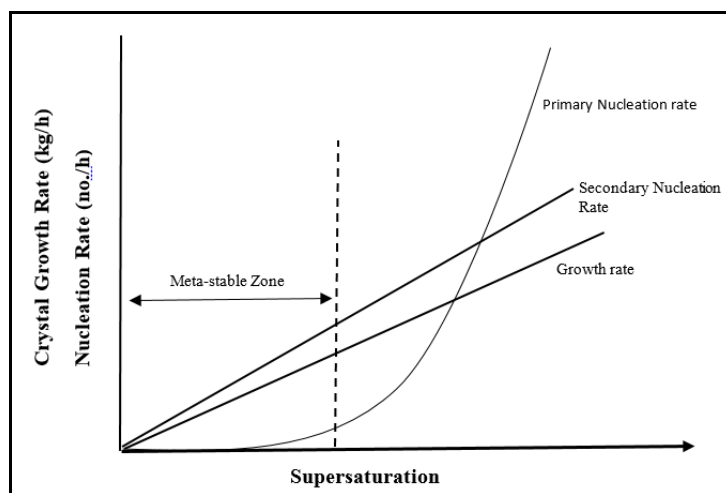


Figure 2.3. The influence of supersaturation on growth and nucleation rates (adapted from Moyers and Rousseau, 1987).

2.4. Modes of Crystallisation

The technique employed to generate supersaturation in a solution, for crystal formation, is referred to as the mode of operation. The mode chosen is dependent on the phase-equilibrium characteristics of the system. The usual techniques for generating supersaturation include: cooling, solvent evaporation, chemical reaction, anti-solvent addition, and common ion addition.

The choice of the mode of crystallisation is normally dependent on the system properties (e.g. solubility, heat of solvent vaporisation, feed stream composition, etc.); and sometimes, a combination of processes is employed to maximize the yield.

In this section, a qualitative discussion is presented on the processes that are used to create and maintain supersaturation conditions that promote crystallisation. These procedures are classified by the manner in which supersaturation is generated. They are briefly described in Table 2.1, and some are illustrated on a solubility curve, as shown in Figure 2.4.

Table 2.1. Methods of Supersaturation Generation.

Mode of Crystallisation	Description
Cooling	Crystallisation is achieved by cooling solvent from a high temperature to a low temperature at constant solvent composition. This mode is applicable to systems where the solubility has a strong temperature dependence. Since the solubility decreases with temperature, the solution becomes supersaturated.
Anti-solvent	Crystallisation is achieved by adding an anti-solvent to a solvent in which the solute is soluble. The anti-solvent is used to reduce solubility of the solute in the mixed solvent, and hence the mixed solvent becomes supersaturated.
Reactive	Crystallisation is achieved by changing the compound ionically or structurally through reaction. The reactants are often soluble with the product being insoluble. Reaction is used to change the concentration of the product above the solubility limit
Evaporative	Crystallisation is achieved by the evaporation of solvent that increases the solute concentration above the solubility limit, resulting in supersaturation. The evaporation of the solvent can be through flashing or heating.

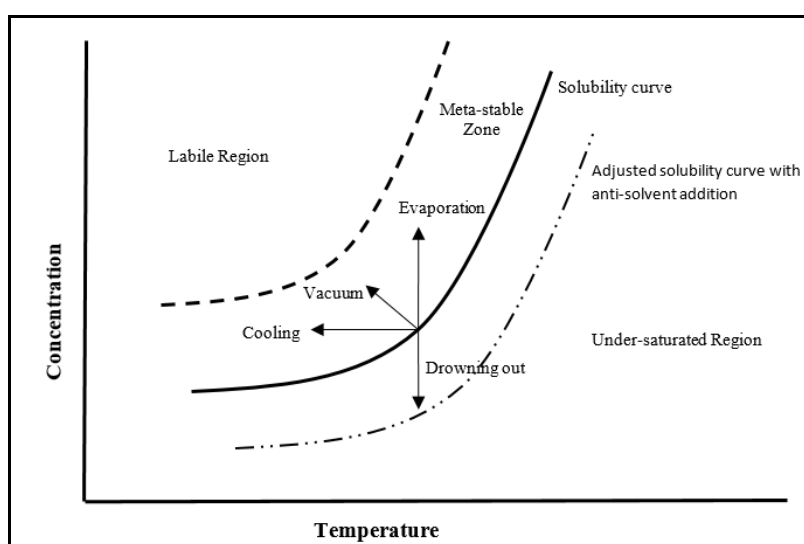


Figure 2.4. Solubility Diagram showing how the different modes of crystallisation influence supersaturation (adapted from Jones, 2002).

The processing implications of the various modes of supersaturation generation are briefly described below.

2.4.1. Cooling crystallisation

Cooling of a solution can be performed in a variety of ways, depending on the system and the desired result (yield, quality, crystal size, etc.). Natural cooling is the simplest method, but results in varying supersaturation as the cooling proceeds. A rapid decrease in the temperature has the potential of passing through the metastable region and reaching the uncontrolled nucleation region. Uncontrolled cooling can also lead to an accumulation of crystal scale on the cooling surface, caused by low temperatures at the wall. This crystal layer is triggered by spontaneous nucleation on the cold surface, followed by growth of the crystal layer. This encrustation acts as a fouling layer to heat transfer and can severely limit the cooling rate, leading to non-uniformity in the product.

When uncontrolled nucleation is not acceptable, cooling strategies can be utilized to match the cooling rate to the growing crystal surface area. Suggested cooling rates by Mullin and Nyvlt (1971) and Mullin (1993), can be used in the control of supersaturation. The suggested cooling rates are lower than natural cooling in order to maintain supersaturation in the growth region. The cooling rate can be increased with crystal growth. The suggested cooling rates also reduce scale layer formation by limiting temperature differences across the jacket. It has been suggested by Mersmann (2001) the formation of the scale layer can be avoided if the temperature difference between the crystallizing mixture and cooling fluid is less than the width of the metastable zone. Recommended cooling rates for organic compounds are in the order of 0.1–0.2 K/min (Beckmann, 2013).

2.4.2. Evaporation Crystallisation

The use of evaporation is widely applied to increase concentration by removing solvent. But, it gives rise to several nucleation and growth control problems. For drug substances (APIs) that require tight control of mean particle size and PSD, uncontrolled growth problems can make this method unsuitable.

The effect of the rate of evaporation rate is similar to the effect of the rate of cooling in creating supersaturation. Therefore similar methods of control are used to match the evaporation rate with the surface area available for growth. Seeding can also be useful if the saturation point can be accurately determined. The seed slurry can be added as the concentration reaches saturation point.

At the heating surface, local high temperatures and a high vaporization rate result in uncontrollable local supersaturation environments in which uncontrolled nucleation can be

excessive, particularly in those regions of poor bulk mixing (Tung , 2009). Crust formation above the heated surface can also lead to significant product quality issues. At the boiling surface, vapour disengagement can lead to very high local supersaturation. In addition, the vapour bubbles can cause local nucleation.

2.4.3. Anti-solvent Crystallisation

This procedure is widely used when cooling has limited effects on solubility. It is reported to have many inherent advantages over both batch cooling and concentration, in terms of crystallisation control. The control of supersaturation and crystal growth, is readily achievable by the rate anti-solvent addition. The obvious disadvantage of anti-solvent addition is the production of larger volumes of solvent mixtures which will also require a separation process for solvent recovery.

2.4.4. Reactive Crystallisation

Reactive crystallisation is defined as a process where the supersaturation of a crystallizing compound is the result of a chemical reaction. The reaction may occur between two complex organic compounds or can occur by means of acid or base neutralization, to form a salt of a complex compound. Reactions can be fast, compared to both the mass transfer rate and growth rate of crystals, thereby leading to high local supersaturation and nucleation (Tung , 2009). Unless the rate of the reaction that generates the supersaturation can be controlled, it is difficult to control crystal growth and particle size in reactive crystallisation. Reactive crystallisation operations are also commonly known as precipitation.

2.5. Synthesis and Analysis of Crystallisation Processes

As previously mentioned, the goal of conceptual design is to find the process flowsheets, equipment sizes and configurations, together with operating conditions that will be optimally viable in meeting the product specification after crystallisation. There are typically many alternatives, so some analysis for the ranking of alternatives is necessary. When the differences among alternatives are small, high accuracy is required to identify the true optimum. Conversely, larger differences among the alternatives mean that less accurate models will not lead to bad decisions. In this section, we review various approaches to conceptual design of crystallisation processes in order to develop a robust computational tool for solvent selection.

Douglas (1988) suggested that at the early stages of process development, only basic structures be considered with details being added at later stages. Thereby, minimizing experimental effort and use of models. But, as the conceptual design moves forward, accurate measurements and rigorous modelling are often necessary to support the decision to favour one alternative over another.

For process development, Ng and Wibowo (2003) proposed that, while a hierarchy of models of increasing complexity and accuracy is recommended, certain choices will require the determination of many experimental parameters, and the development and optimization of a detailed mathematical model. This is in contradiction to the fact that, typically, in the conceptual design phase of API development, important decisions have to be made at early stages, when information is still limited.

To address this need for detailed information at an early stage of crystallisation process design, short-cut methods (models), requiring limited information, are necessary. These shortcut models may not be accurate. However, by capturing the principal physical phenomena, they point to the correct trend under ‘what if’ scenarios, thus suggesting the right direction for problem solving. The hierarchy and complexity of models required over the development phase is illustrated in Figure 2.5.

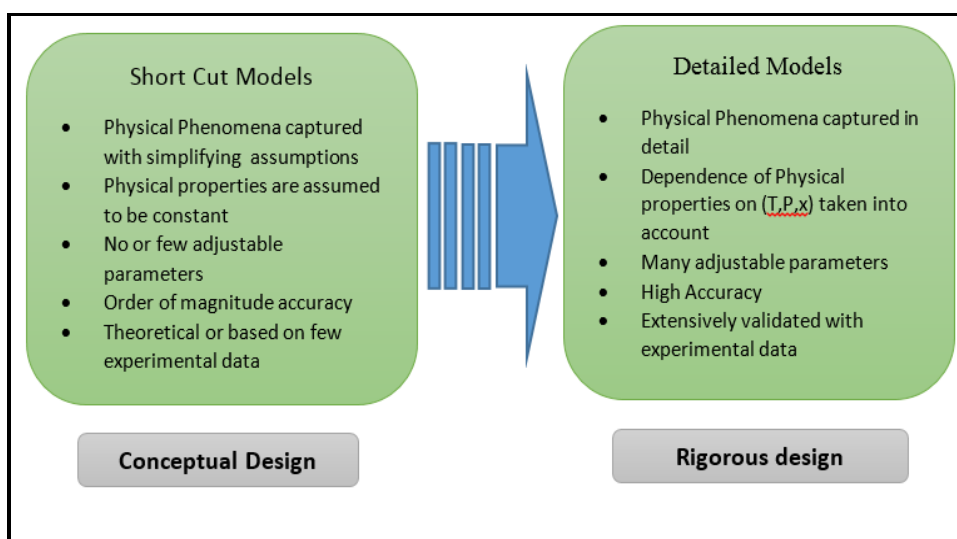


Figure 2.5. A hierarchy of models for process development (adapted from Ng and Wibowo, 2003).

In recent years, graphical methods have proven useful for the systematic generation of process flowsheets (Rajagopal et al., 1991; Wibowo et al., 2004). Graphical methods based on thermodynamic phase equilibria readily establish the thermodynamic limitations of a process. In distillation processes, such an analysis usually involves residue curves and distillation boundaries; while for extraction, the shape of the miscibility gap and the location of tie-lines are taken into account.

For crystallisation processes, the crystallisation path map is a useful tool for finding feasible flowsheets. These paths are trajectories of the liquid composition in a crystallizer as the solid is formed and removed from solution (Slaughter and Doherty, 1995). These trajectories are conveniently represented in the form of phase diagrams. Such phase diagrams can be predicted using a thermodynamic model, or can be obtained from experimental data (if available). For crystallisation, solid-liquid equilibrium (SLE) phase diagrams are used.

In order to evaluate the process design options for crystallisation processes, Moyers and Rousseau (1987) highlighted the need, in the earliest stages, for accurate solid-liquid equilibrium data. To this end, with the use of SLE phase diagrams the following can be established prior to any detailed investigations:

- The feasibility of the planned process; that is, determining if pure solute can in fact be crystallized from the feed solution.
- The feasibility of operating regions, process operations and pathways to generate the desired product.
- The thermodynamic limitations on crystallisation of a component. The feed composition and position of the eutectic fix the maximum attainable solute recovery.
- The identity of the maximum theoretical yield under any given condition.
- The effect of solvent / anti-solvent / co-solvent selection on the process.
- The temperature and/or pressure ranges of the crystallizer operation and the composition of the residue liquor exiting the crystallizer.

An example of a ternary solid-liquid equilibrium phase diagram of a system consisting of a solvent (S), product (P) and anti-solvent (A) is shown in Figure 2.6. This three-dimensional diagram is formed by putting together the T-x diagrams of three binary systems (P-A, P-S, and A-S). The surfaces bounded by the points $(T_{m,P} - BE_{S,P} - TE_{S,A,P} - BE_{A,P})$; $(T_{m,S} - BE_{S,P} - TE_{S,A,P} - BE_{S,A})$ and $(T_{m,A} - BE_{S,A} - TE_{S,A,P} - BE_{A,P})$ are the saturation surfaces (or solubility surfaces) for each of the three components P, S and A respectively. Two surfaces intersect at the eutectic trough, which represents the area where multiple components are saturated. Lines $(BE_{S,P} - TE_{S,A,P})$, $(BE_{A,P} - TE_{S,A,P})$, and $(BE_{S,A} - TE_{S,A,P})$ are the binary troughs for components (P and S), (A and P) and (A and S), respectively; and all three surfaces meet at the ternary eutectic point $(TE_{S,A,P})$, at which all three components are saturated at the same time. Because working with three-dimensional figures can be challenging, it is often desirable to reduce the dimensionality. Two key diagrams obtained from the phase diagram are the poly-thermal projection and isothermal cuts.

The poly-thermal projection is the top view one would get looking straight down, through the prism, from above. This projection can be seen at the base of Figure 2.6. The presence of eutectics establishes the boundaries that divide the projection into distinct crystallisation regions, which are non-overlapping and mutually exclusive. The aim is to divide the composition space into operating regions that are subspaces in which only a certain component can be crystallized. In each bounded region of the phase diagram only a single component can be crystallized in pure form.

Crystallisation of more than one component is possible for components sharing a eutectic manifold. The operating regions are bounded by various manifolds. The number of these manifolds is dependent on the number of components. Several components are crystallized at a eutectic manifold. The number of different types of eutectic manifolds depends on the number of components (n) involved. A system can have (n - binary eutectic manifolds in which 2 components are crystallized simultaneously, ($n-2$) ternary eutectic manifolds in which 3 components are crystallized simultaneously, etc.

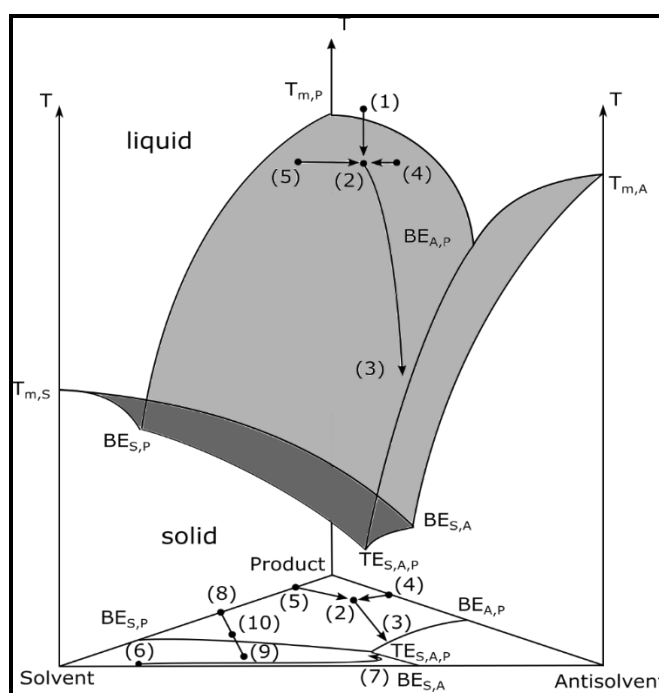


Figure 2.6. Solid-Liquid Equilibrium (SLE) phase diagram with the poly-thermal projection of a system consisting 3 components.

Note that every point on this projection has a different temperature, corresponding to the saturation surfaces in the original 3D diagram, thus the name, poly-thermal projection. Also, since only the saturation surfaces are displayed, this projection does not provide any information on the solid mixture regions that exist below the eutectic points.

On the other hand, an isothermal cut diagram represents a series of cuts taken at different temperatures. It is useful for evaluating systems at constant temperature, such as, evaporative or anti-solvent crystallisation. Whereas, a poly-thermal projection is more useful to identify the feed compositions from which a pure product can be obtained without being restricted to a certain temperature, and where the regions in which pure components can be crystallized out are clearly identified.

Equilibrium phase diagrams can be used to determine the effect of some basic operations in crystallisation and to visualise the resulting movements within the composition space. Using this technique, several researchers, (Dye and Ng, 1995; Wibowo and Ng, 2000, 2002; Schroer et al., 2001; Cisternas et al., 2006), have developed rules and guidelines for synthesising operational protocols for crystallisation-based separation processes, and some of the key elements are presented here.

The composition of the feed stream determines the location at which it should be introduced into the compositional space. The pure component, i , is represented by the melting point of component, i , resulting in an apex (end point) in the ternary poly-thermal projection. Our mixture shows three binary eutectic points, $BE_{i,j}$, and one ternary eutectic point, $TE_{i,j,k}$. The resulting eutectic curve between these points defines the operating region.

First, the attainable operating regions in the phase diagram need to be analyzed. The method to generate supersaturation in each region is closely related to the order in which the regions in the phase diagram will be visited. Within each operating region there are several tasks possible to induce crystallisation: temperature change; composition change by solvent removal; and composition change by anti-solvent addition or removal. The eutectic manifolds need to be crossed to move to a different operating region, which can be affected by the following tasks for composition manipulation: solvent addition or removal; anti-solvent addition or removal; and stream combination.

Some operations, and their resulting movements in the compositional space, are summarized in Table 2.2 and shown in Figure 2.6. Starting with a feed composition (1) one can cool down until the SLE surface is reached (2). For simplicity, we assume equilibrium operation without the need to sub-cool. Further cooling results in crystallisation of the product and movement along the SLE surface on a straight line away from the product apex (3). The addition of a solvent to a binary mixture containing product and anti-solvent (4) results in (2). Adding an anti-solvent to a binary mixture of the product dissolved in a solvent (5) also leads to a new composition (2). The evaporation of the solvent from (6) along the distillation curve results in the new composition (7). In order to change into a different operating region one can combine two different streams (8) and (9), resulting in (10).

Table 2.2. Impact of key basic operations in the crystallisation compositional space.

Operation	Resulting movement in compositional space
cooling	Create supersaturation within operating region only. Movement away from solids apex that is being crystallized. Isothermal cut lines can be used to quantify the changes.
evaporation	Create supersaturation within operating region or used to cross into different operating regions. Straight line movement from away from solvent apex through "feed" point towards desired solid/region to be crystallised. Lever arm rule can be used to quantify changes.
Solvent addition	Movement towards solvent apex. Used to cross into different operating region. Straight line movement away from last condition towards solvent apex. Lever arm rule can be used to quantify changes.
Anti-solvent addition	Create supersaturation within operating region or used to cross into different operating regions. Movement away from solids apex that is being crystallized. Anti-solvent composition cut lines can be used to quantify changes.
Stream combination	Changes composition of stream – used to cross operating regions. Movement towards component apex that is being enriched.

It is illustrated, in Figure 2.7, how the movements in the poly-thermal diagram were translated into a process flowsheet, using the rules of Table 2.2. The system consisted of two solutes, A and B, in solvent S. The feed composition was as indicated by point 1. The tie-line from point S to the feed point extended into compartment A, indicating that solute A should have been crystallised first. A bypass of the binary eutectic line, to reach compartment A, could be achieved by one of the following: evaporation, or anti-solvent addition, or stream combination.

Evaporation was chosen to bypass the binary eutectic line, until a tie-line radiating from point A was reached, which resulted in the maximum amount of A crystallising, as shown by the blue tie-line passing through point 2. Cooling crystallisation and solid recovery was then conducted until reaching a position close to the eutectic boundary, as indicated by point 3. Compartment B was entered by means of further evaporation of the remaining mother liquor. In order to maximize the amount of B crystals, evaporation was undertaken until the radiating tie-line from point B was reached (that results in maximum B), which was the line close to the axis AB.

Cooling was then undertaken, with solid recovery, until close to the AB eutectic line. A stream combination was implemented by adding solvent (recovered from the evaporation process) to the mother liquor, to cross over into compartment A. Solvent was added until a radiating tie-line was reached that resulted in maximum A. Once again, cooling with solid recovery was possible up to a point close to a eutectic line.

As can be seen, a number of strategies could be employed to recover more of both crystal A and crystal B. However, this may mean requiring more evaporators / cooling crystallisers and solid recovery units, and more utilities for heating and chilling. To determine the ratio of solvent to be

evaporated or the amount of crystals recovered in each operation, the “lever-arm” rule can be applied to the respective tie-lines. With these ratios, together with the initial feed rate and conditions, the material balance and energy balance can be completed.

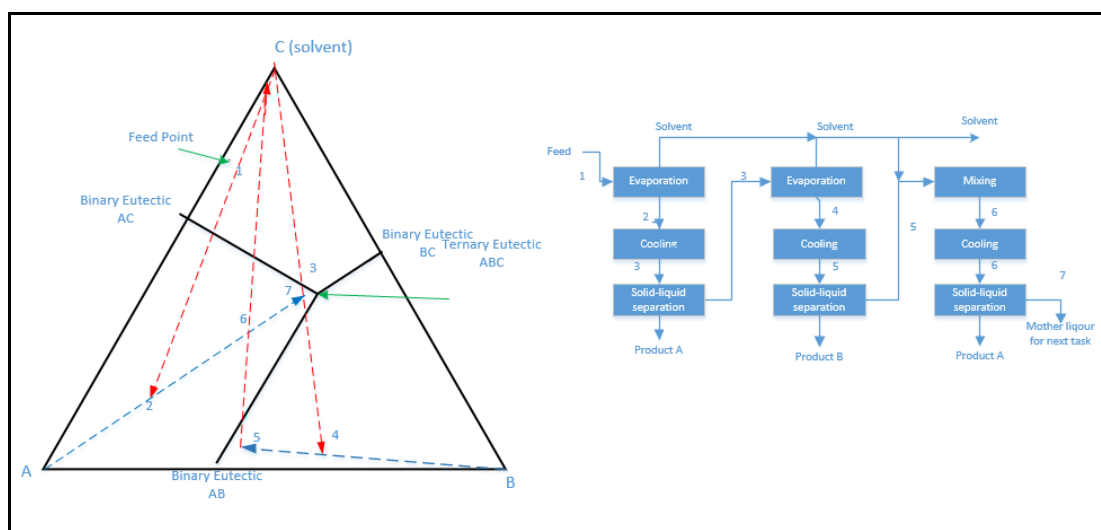


Figure 2.7. Examples of phase diagrams for ternary systems, the movements in the compositional space and the resulting operating protocol: 1-2 evaporation; 2-3 cooling and crystal recovery; 3-4 evaporation; 4-5 cooling and crystal recovery; 5-6 stream combination; 6-7 cooling and crystal recovery.

The example in Figure 2.7 reveals the usefulness of crystallisation maps in evaluating potential crystallisation schemes for the recovery of stipulated products. However, as the number of components increases, so does the complexity of the phase diagrams. To overcome this challenge, a computational tool can be developed to replicate the visual analysis.

This highlights the need for a computational framework: firstly, to generate the phase equilibria required for the analysis; secondly, to evaluate the various crystallisation modes of evaporation, cooling, and anti-solvent and; thirdly, to develop a process and economic evaluation of the suggested schemes in order to shortlist the flowsheets for further investigation. To this end, in the chapter 4, the details of a computational framework that has been developed in this research work, will be presented.

CHAPTER 3: THERMODYNAMIC FRAMEWORK FOR MULTICOMPONENT, MULTIPHASE EQUILIBRIA

The essential requirement in the development of any predictive process computational tool include: the understanding of the theoretical foundation of the key phenomena, and the capability, range of applicability, and accuracy of the various thermodynamic and process models that are used in the development of the process computational tool. In this work, solid-liquid phase equilibria are used as the theoretical foundation to understand crystallisation processes, and multicomponent multiphase flash calculations are used to establish the presence of solid-liquid equilibria and the subsequent changes to the phase equilibria under varying conditions of temperature and concentration.

In this chapter, the thermodynamic phenomena and models required to address the various aspects of multicomponent multi-phase equilibria are presented. But first, a brief outline is provided on the importance of phase equilibrium processes in solvent selection for crystallisation. The treatment of thermodynamics is then provided in three sections. The first section covers the preliminary thermodynamic concepts and the derivation of solid-liquid phase equilibrium (SLE) equations. In the second section an overview of correlative and predictive models currently available for phase equilibria calculations is presented, and finally, calculation procedures are presented for the determination the multicomponent multi-phase equilibria.

3.1 Use of Phase Diagrams to analyse crystallisation

Phase data is an essential component in the design, operation and optimization of mass transfer operations. Experimental measurements of phase data are generally limited to a few data points at specific conditions. These data points are then utilized to determine interaction parameters and model coefficients of appropriate thermodynamic models, to extend the use of the data measured beyond the experimental conditions. However, in solvent selection processes, making experimental measurements with various solvents, co-solvents and anti-solvent combinations at various conditions can be impractical in terms of cost and experimental time required.

The experimental effort required to rapidly determine a full phase diagram in mass transfer operations increases with the addition of components. Hence, reliable, predictive tools are necessary for the initial screening and selection processes. The predictions can guide the experimental work to confirm critical regions of interest.

Graphical representation of the phase data, in the form of phase diagrams, provide the thermodynamic information required for the design and synthesis of separation processes. Since the availability of accurate phase diagrams is important for successful application of analysis and

synthesis procedures, computational tools are essential methods developed to generate SLE phase diagrams.

The prediction of the solid-liquid-vapour equilibrium (SLVE) data for multicomponent systems, and the ability to generate phase diagrams and solubility curves, are requirements in the design and analysis of crystallisation processes. The operations that are associated with crystallisation, such as, cooling, heating, evaporation, and solvent addition, can be represented as specific types of movements on the phase diagrams, thus allowing the user to quickly evaluate process alternatives and develop feasible process schemes (Wibowo and Ng, 2000, 2002), (Wibowo et al., 2004).

Traditionally, SLE data, generated using the Schroder equation, is primarily used in the analysis of crystallisation processes. However, applications, with the combination of evaporation with co-solvent/anti-solvent systems, vacuum application and the sizing of downstream processes, require a more rigorous multiphase equilibria procedure. It requires a procedure that can calculate the amount and compositions of all three phases (liquid, solid and vapour) present during the process of crystallisation. In this work, multiphase equilibria are generated by means of algorithms based on multiphase flash calculations.

The phase equilibrium calculation is used to accurately predict the correct number of phases and their compositions, at equilibrium, in the system. Two kinds of approach are typically used to model multiphase flash calculations: the equation-solving approach (K-value method); and the minimization of the Gibbs free energy approach. Isofugacity conditions and mass balances form the set of equations in the equation-solving approach, and the stability test or the common tangent test forms the basis of Minimization of the global Gibbs free energy approach (Lucia, 2000), (Parekh and Mathias, 1998).

A robust solvent selection process must be capable of identifying and excluding solvent systems that exhibit immiscibility or azeotropic behaviour. Hence the phase behaviour of relevant systems (excluding azeotropic and immiscible systems) can be modelled as solid-liquid-vapour systems.

3.2. Development of thermodynamic framework

When selecting thermodynamic methods for a phase calculation, the first consideration is whether it is a pure component or a mixture of components system, and secondly, what type of behaviour is exhibited by the component or mixture. In thermodynamics, fluid behaviour is typically classified as ideal, regular, or non-ideal systems, and the behaviour is attributable to either physical or chemical intermolecular forces.

In fluid behaviour, physical forces would be due to collisions between molecules, affected primarily by the size and shape of the molecules. Chemical forces would be electromagnetic-type

forces at the molecular level, which tend to cause the molecules to group or associate in a non-random fashion.

Ideal fluid behaviour is experienced in systems where all the molecules are of virtually the same size, and no intermolecular forces of attraction or repulsion exist. An ideal / perfect gas consists of freely moving particles of negligible volume and intermolecular forces. Regular behaviour is experienced in systems where the non-idealities stem from moderate physical interactions, i.e., from differences in the size and shape of the molecules. Intermolecular associations are assumed to be minimal. The strict definition of a regular solution is one where the excess entropy of mixing is zero. This generally occurs in systems where the components are non-polar and do not differ appreciably in size, shape, and chemical behaviour. Non-ideal (polar) behaviour is experienced in systems where there are strong influences, predominantly from chemical or intermolecular forces of attraction or repulsion. The impact of the differences in the component or mixture behaviour will become explicit in the following sections.

3.2.1. Thermodynamic Preliminaries

The state of a thermodynamic system is defined by a set of variables associated with thermodynamic state, which include: pressure, temperature, volume, internal energy, enthalpy, and entropy. For a heterogeneous, closed system, containing m components, in π phases, the whole system reaches the equilibrium state if the intensive quantities (including chemical potential, μ) are identical for all π phases.

$$T^{(1)} = T^{(2)} = \dots = T^{(\pi)} \quad (3.1)$$

$$P^{(1)} = P^{(2)} = \dots = P^{(\pi)} \quad (3.2)$$

Walas (1985) proposed that, an equilibrium state is characterized as having a maximum entropy or a minimum energy function, at specified values of the two other properties of the particular fundamental equation. These possible extrema at equilibrium for different known combination of independent variables are identified in Table 3.1 below.

Table 3.1. Equilibrium Extrema for specified conditions (Walas, 1985).

Variable	Extrema Property and Energy function	
Independent variable	Maximum	Minimum
U, V	S	-
S, V	-	U , where $U = TS - PV + \sum x_i \mu_i$
P, H	S	-
P, S	-	H , where $H = TS + \sum x_i \mu_i$
T, V	-	A , where $A = -PV + \sum x_i \mu_i$
P, T	-	G , where $G = \sum x_i \mu_i$

Where the chemical potential, μ_i , is defined as the change in internal energy of the system per mole of substance i , and may be expressed in terms of any of the four fundamental groups of properties:

$$\mu_i = \left(\frac{\partial U}{\partial n_i}\right)_{SVn_j} = \left(\frac{\partial H}{\partial n_i}\right)_{SPn_j} = \left(\frac{\partial A}{\partial n_i}\right)_{TVn_j} = \left(\frac{\partial G}{\partial n_i}\right)_{PTn_j} \quad (3.3)$$

Because of the importance of temperature T and pressure P as independent properties, the chemical potential is generally taken to be the derivative of Gibbs energy with respect to the number of moles, referred to as partial molal Gibbs energy.

For the transfer of dn_i moles of a substance between two phases, at the same temperature T and pressure P , the change in Gibbs energy is

$$dG = (\mu_i^{(2)} - \mu_i^{(1)}) dn_i \quad (3.4)$$

Since G is a minimum at equilibrium, its derivative is zero:

$$\left(\frac{\partial G}{\partial n_i}\right)_{TPn_j} = 0, \text{ and therefore } \mu_i^{(1)} = \mu_i^{(2)} \quad (3.5)$$

When the transfer of more than one species between more than two phases occurs, equality of chemical potential extends to all phases and all species and can be expressed as:

$$\begin{aligned} \mu_1^{(1)} &= \mu_1^{(2)} = \dots = \mu_1^{(\pi)} \\ \mu_2^{(1)} &= \mu_2^{(2)} = \dots = \mu_2^{(\pi)} \\ &\dots \\ \mu_m^{(1)} &= \mu_m^{(2)} = \dots = \mu_m^{(\pi)} \end{aligned} \quad (3.6)$$

In order to relate the abstract chemical potential of a substance to physically measurable quantities, such as, temperature, pressure, and composition, we consider the isothermal change in the Gibbs energy of n moles of a perfect gas:

$$\left(\frac{\partial G}{\partial P}\right)_{T,n} = V = \frac{nRT}{P} \quad (3.7)$$

Integrating the above equation, we obtain:

$$G(T, P, n) - G(T, P_0, n) = nRT \ln \frac{P}{P_0} \quad (3.8)$$

Which is the difference of the Gibbs energy of a perfect gas at an arbitrary pressure P and a fixed reference pressure P_0 .

Since G for a pure component is proportional to n , equation (3.8), for one mole of pure component the equation can be rewritten as:

$$\mu(T, P, n) - \mu(T, P_0, n) = RT \ln \frac{P}{P_0} \quad (3.9)$$

Since the above equation (3.9) is limited to an ideal gas, Lewis and Randall (1923) introduced an auxiliary function concept of fugacity, which allows the above equation to be extended to real fluids. Lewis and Randall (1923) described fugacity as a measure of the tendency of a molecule to escape from the phase in which it is. Fugacity is considered the true (observable) system pressure, compensated for by molecular interactions. Therefore, for real fluids, pressure is replaced by fugacity in the above equation, and the change in chemical potential of a substance i , between a reference state, P_0 , T_0 , and the actual state, P , T , is given as:

$$\mu(T, P, n) - \mu(T, P_0, n) = \mu_2 - \mu_1^0 = RT \ln \left(\frac{f}{f_0} \right) \quad (3.10)$$

This represents an isothermal change in chemical potential when going from f^0 to f_i . Therefore the standard states for all components in all phases must be at the same temperature; the pressure and composition of the standard states, however, do not necessarily have to be the same. The phase equilibrium criterion given in terms of chemical potential can therefore be rewritten in terms of fugacity as:

$$\begin{aligned} f_1^{(1)} &= f_1^{(2)} = \dots = f_1^{(\pi)} \\ f_2^{(1)} &= f_2^{(2)} = \dots = f_2^{(\pi)} \\ f_m^{(1)} &= f_m^{(2)} = \dots = f_m^{(\pi)} \end{aligned} \quad (3.11)$$

For a pure, ideal gas, the fugacity is equal to the pressure, and for component i , in a mixture of an ideal gas, it is equal to its partial pressure, $y_i P$. Because all systems, either pure or mixed, approach ideal gas behaviour at very low pressures, the definition of fugacity is completed by the limit:

$$\frac{f_i}{y_i P} \rightarrow 1 \text{ as } P \rightarrow 0 \quad (3.13)$$

We further define the above dimensionless ratio as the fugacity coefficient, φ_i , and for a mixture of ideal gases, $\varphi_i = 1$.

It is clear that, if we want to take advantage of the fugacity criteria to perform equilibrium calculations, we need to have a means of calculation. With the definition of fugacity, in terms of chemical potential for a pure component and Maxwell's relationship being:

$$d\mu = RTd\ln f \quad \text{and} \quad \left(\frac{\partial \mu}{\partial P}\right)_T = \tilde{v} \quad (3.14)$$

$$RTd\ln f = \tilde{v}dP \quad (3.15)$$

The definition of the fugacity coefficient of $\varphi = \frac{f}{P}$ can be written as:

$$\ln \varphi = \ln f - \ln P \quad (3.16)$$

Therefore the equation can be written as:

$$RTd\ln \varphi = \tilde{v}dP - RTd\ln P \quad \text{or equivalently as} \quad RTd\ln \varphi = \tilde{v}dP - RT \frac{dP}{P} \quad (3.17)$$

Integrating:

$$\int_{\ln \varphi^0}^{\ln \varphi} d\ln \varphi = \int_{P^0}^P \left\{ \frac{\tilde{v}}{RT} - \frac{1}{P} \right\} dP$$

It is convenient to define the lower limit of integration as the ideal state, for which the values of fugacity coefficient, volume, and compressibility factor are known. At the ideal state, at the lower limit $P \rightarrow 0$, $\varphi \rightarrow 1$, $\ln \varphi \rightarrow 0$. Substituting this into the above equation, we have:

$$\ln \varphi = \int_0^P \left\{ \frac{\tilde{v}}{RT} - \frac{1}{P} \right\} dP \quad (3.18)$$

Expressed in terms of compressibility factor $Z = \frac{P\tilde{v}}{RT}$, we have:

$$\ln \varphi = \int_0^P \left\{ \frac{P\tilde{v}}{RT} - 1 \right\} \frac{dP}{P} = \int_0^P (Z - 1) \frac{dP}{P} \quad (3.19)$$

Or expressed in terms of fugacity:

$$\ln f = \ln P + \int_0^P (Z - 1) \frac{dP}{P} \quad (3.20)$$

Similarly, we can derive the expression for the fugacity coefficient of a component in a multicomponent mixture. Following a pattern similar to that which we have presented, beginning with the definition of fugacity for a component in terms of chemical potential, we can derive:

$$\ln \varphi_i = \int_0^P (\bar{Z}_i - 1) \frac{dP}{P} \quad \text{where} \quad \bar{Z}_i = \frac{P\bar{V}_i}{RT} \quad (3.21)$$

The above equations show that fugacity, or the fugacity coefficient, is a function of pressure, temperature and volume. To compute the fugacity coefficient with equations of state that are explicitly expressed in pressure, we can replace Vdp with its equivalent:

$$\int_{P_0}^P V dP = \int_{P_0}^P d(PV) - \int_{V_0}^V P dV = PV - P_0V_0 - \int_{V_0}^V P dV \quad (3.22)$$

Accordingly, the fugacity coefficient for a pure component may be evaluated as follows:

$$\ln \phi = Z - 1 - \ln Z - \frac{1}{RT} \int_{\infty}^V \left(P - \frac{RT}{V} \right) dV \quad (3.23)$$

And for mixtures, the partial fugacity is evaluated as follows:

$$\ln \phi_i = \frac{1}{RT} \int_{\infty}^V \left(\frac{\partial P}{\partial n_i} - \frac{RT}{V} \right) dV - \ln Z \quad (3.24)$$

Table 3-2 provides a list of partial fugacity coefficient expressions that result from the integrations of commonly used equations of state. A more comprehensive review of the various equations of state are provided in the texts by Walas (1984), Gmehling et al. (2012), and Kontogeorgis and Folas (2010).

To evaluate the fugacity of a liquid, at a pressure above the saturation pressure, a two-step approach is applied. First, at saturation the liquid fugacity equals the vapour fugacity:

Since $\phi_i^{sat} = \frac{f_i^{sat}}{P_i^{sat}}$, therefore $f_i^l = f_i^{sat} = \phi_i^{sat} P_i^{sat}$

In the second step, the change in fugacity from saturation pressure to system pressure (at constant T) is determined. This effect is generally small but is significant at high pressures. The fugacity is related to pressure at constant temperature by the equation:

$$dG_i = V_i dP - S_i dT = RT d \ln f_i \quad (3.33)$$

At constant T, we have

$$d \ln f_i = \frac{V_i^l}{RT} dP \quad (3.34)$$

Integrating from P_i^{sat} to P , we have

$$f_i^l = f_i^{sat} \exp \left\{ \frac{1}{RT} \int_{P_i^{sat}}^P V_i^l dP \right\} \quad (3.35)$$

The exponential term is referred to as the Poynting correction factor, which, to evaluate, requires liquid molar volume as a function of pressure.

Table 3.2. Partial Fugacity Coefficient Expressions for some Equations of State (Walas, 1985).

EoS Name	Equations
Viral	<p>Equation</p> $z = 1 + \frac{BP}{RT}$ <p>Fugacity Expression</p> $\ln \phi_i = \frac{P}{RT} \{ B_{ii} + 0.5 [\sum_i \sum_k y_j y_k (2\delta_{ji} - \delta_{jk})] \}$ <p>Where</p> $B = \sum \sum y_i y_j B_{ij} \quad \delta_{ji} = 2B_{ji} - B_{jj} - B_{ii} \quad \delta_{jk} = 2B_{jk} - B_{jj} - B_{kk}$
Van der Waals	<p>Equation</p> $P = \frac{RT}{V-b} - \frac{a}{V^2}$ <p>Fugacity Expression</p> $\ln \phi_i = \frac{b_i}{V-b} - \ln \left[\frac{V-b}{V} \right] - \frac{2\sqrt{a_i}}{RTV} \sum_{j=1}^n y_j \sqrt{a_{ij}} (1 - k_{ij}) - \ln Z$ <p>Where</p> $a = (\sum y_i \sqrt{a_i})^2 \text{ and } b = \sum y_i b_i$
Redlich -Kwong	<p>Equation</p> $P = \frac{RT}{V-b} - \frac{a}{\sqrt{T}V(V+b)}$ <p>Fugacity Expression</p> $\ln \phi_i = \frac{b_i}{b} (z-1) - \ln \left[Z \left(1 - \frac{b}{V} \right) \right] + \frac{1}{bRT^{1.5}} \left[\frac{ab_i}{b} - 2\sqrt{aa_i} \right] \ln \left(1 + \frac{b}{V} \right)$ <p>Where</p> $a = (\sum y_i \sqrt{a_i})^2 \text{ and } b = \sum y_i b_i$
Peng-Robinson	<p>Equation</p> $P = \frac{RT}{V-b} - \frac{a\alpha}{V(V+b)+b(V-b)}$ <p>Fugacity Expression</p> $\ln \phi_i = \frac{b_i}{b} (z-1) - \ln \left(Z - \frac{bP}{RT} \right) + \frac{a\alpha}{4.828 bRT} \left[\frac{b_i}{b} - \frac{2}{a\alpha} \sum_j y_j (a\alpha)_{ij} \right] \ln \left[\frac{Z+2.414b}{Z-0.414b} \right]$

When V_i^l is assumed independent of pressure (i.e incompressible fluid), the above equation becomes

$$f_i^l = \phi_i^{sat} P_i^{sat} \exp \left[\frac{V_i^l (P - P_i^{sat})}{RT} \right] \quad (3.36)$$

Note:

The above approach for computing fugacity is effective only if the EoS is adequate for the computation of the fugacity coefficients. This may not be true for two reasons: either the equation may not adequately represent the compound itself, or the mixing rules may not adequately represent, or quantify, what is happening to the molecules in a solution, i.e., fail to adequately model intermolecular forces between different molecules in a mixture.

For systems where the non-idealities derive from chemical or intermolecular forces of attraction or repulsion, the predictive methods described above are generally inadequate. For these

situations, it is necessary to use methods based upon the excess Gibbs free energy, that is, to use activity coefficient methods. To account for non-ideality of a substance, the quantity activity coefficient γ_i is introduced. The activity coefficient of a non-ideal solution is defined as the ratio of the activity a_i , and the mole fraction x_i , where Lewis defined the fugacity ratio as the activity a_i :

$$\gamma_i = \frac{a_i}{x_i} = \frac{f_i}{x_i f_i^0} \quad (3.37)$$

The relationship between the activity coefficient and Gibbs free energy can be obtained using the concept of excess properties. Excess functions are thermodynamic properties of solutions that are in excess of those of an ideal solution at the same conditions of temperature, pressure and composition. In general for property M:

the excess value of M = (the actual value of M)_(at T,P,x) – (the ideal value of M)_(at the same T,P,x)

$$M^E = M_{T,P,x} - M_{T,P,x}^{id} \quad (3.38)$$

The excess Gibbs energy is defined by

$$G^{ex} = G - G^{id} \quad (3.39)$$

And for mixtures

$$G_i^{ex} = G_i - G_i^{id} \quad (3.40)$$

$$G_i^{ex} = RT \ln \left(\frac{f_i}{x_i} \right) - RT \ln f_i^0 \quad (3.41)$$

$$G_i^{ex} = RT \ln \left(\frac{f_i}{x_i f_i^0} \right) = RT \ln \gamma_i \quad (3.42)$$

$$G^{ex} = \sum x_i G_i^{ex} = RT \sum x_i \ln \gamma_i \quad (3.43)$$

Several analytical expressions for the concentration dependence of the excess Gibbs Energy have been developed, and hence, it is a fairly simple process to find an analytical expression for the activity coefficient. These expressions/models account for the various interactions between the various components in the various interacting phases, with binary interaction parameters.

3.2. Solid-liquid Equilibria

The process of crystal formation is a solid-liquid equilibria phenomena, and the essential thermodynamic equations and relationships for Solid-Liquid-Vapour Equilibrium (SLVE) calculations presented here are analogous to those described by Lira-Galeana et al. (1996) for wax deposition in hydrocarbon streams. At a fixed temperature and pressure, a liquid phase (*l*) may coexist in equilibrium with a vapour phase (*v*) and a solid phase (*s*). At equilibrium, for every component *i* the following thermodynamic relationship applies:

$$f_i^v = f_i^l = f_i^s \quad i = 1, 2, \dots, N \quad (3.44)$$

Where f is the fugacity and N is the number of components.

The vapour phase can be described by an equation of state (EOS), the liquid phase by an activity-coefficient model or by an EOS, and the solid phase is generally described by an activity-coefficient model:

$$f_i^v = \phi_i^v y_i P; \quad f_i^l = \phi_i^l x_i^l P \text{ or } f_i^l = \gamma_i^l x_i^l f_{pure\ i}^l, \quad \text{and } f_i^s = \gamma_i^s x_i^s f_{pure\ i}^s \quad (3.45)$$

Where ϕ_i^v and ϕ_i^l are fugacity coefficients of component i in the vapour and liquid phases, respectively, and are computed from an EOS; and γ_i^l and γ_i^s are activity coefficients of component i in the liquid and the solid phases; respectively; and are computed from activity coefficient models.

Further, the use of equilibrium coefficients, K , generally used in VLE and LLE computations, are extended to describe the equilibrium relationships between the phases in a VLSE system. For the vapour-liquid phase, the commonly-used expression is:

$$K_i^{vl} = \frac{y_i}{x_i^l} = \frac{\phi_i^l}{\phi_i^v} \quad (3.46)$$

For the liquid-solid phase, fugacity can be described with the help of activity coefficients and the standard fugacities for the liquid and solid phases:

$$x_i^l \gamma_i^l f_{pure\ i}^l = x_i^s \gamma_i^s f_{pure\ i}^s \quad \text{or} \quad \frac{x_i^l \gamma_i^l}{x_i^s \gamma_i^s} = \left(\frac{f^s}{f^l} \right)_{pure\ i} \quad (3.47)$$

Where Lira-Galeana et al. (1996) proposed an analogous equilibrium constant:

$$K_i^{sl} = \frac{\gamma_i^l}{\gamma_i^s} \left(\frac{f^l}{f^s} \right)_{pure\ i} \quad \text{and therefore} \quad K_i^{sl} = \frac{x_i^s}{x_i^l} \quad (3.48)$$

The required ratio of the standard fugacities can be obtained by examining the thermodynamic cycle of the sublimation process of a solid. Because enthalpy, entropy and Gibbs energy are unique state functions, calculation of a difference in such a quantity is independent of the thermodynamic route of calculation. The described phase transition from liquid to solid evaluated at constant temperature T can be evaluated via an alternate path as shown in Figure 3.1. Since data for the triple point is not abundant, we can reasonably assume that the triple point temperature is well approximated by the melting temperature T_m , i.e. $T_r \approx T_m$.

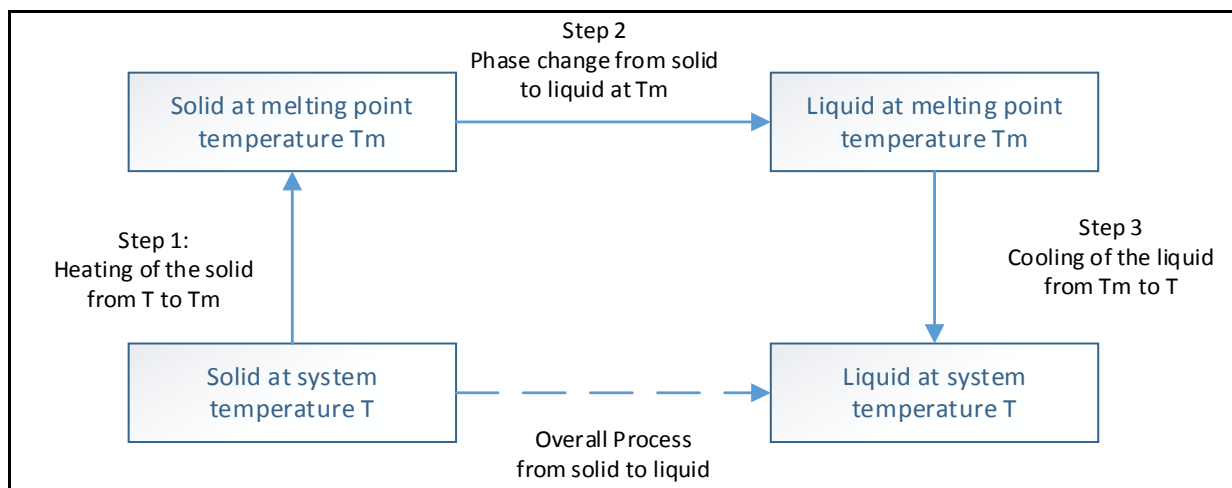


Figure 3.1. Thermodynamic cycle for the sublimation process of a solid.

First, the solid at state is heated from the system temperature T up to triple point temperature $T_r \approx T_m$ – step 1. Then, at the temperature $T_r \approx T_m$, the solid undergoes a phase change from solid to liquid – step 2. Finally, the liquid is cooled from the triple point temperature $T_r \approx T_m$ to the systems temperature T – step 3. The changes in the enthalpy and entropy, via the alternate hypothetical route, is shown in Figure 3.2.

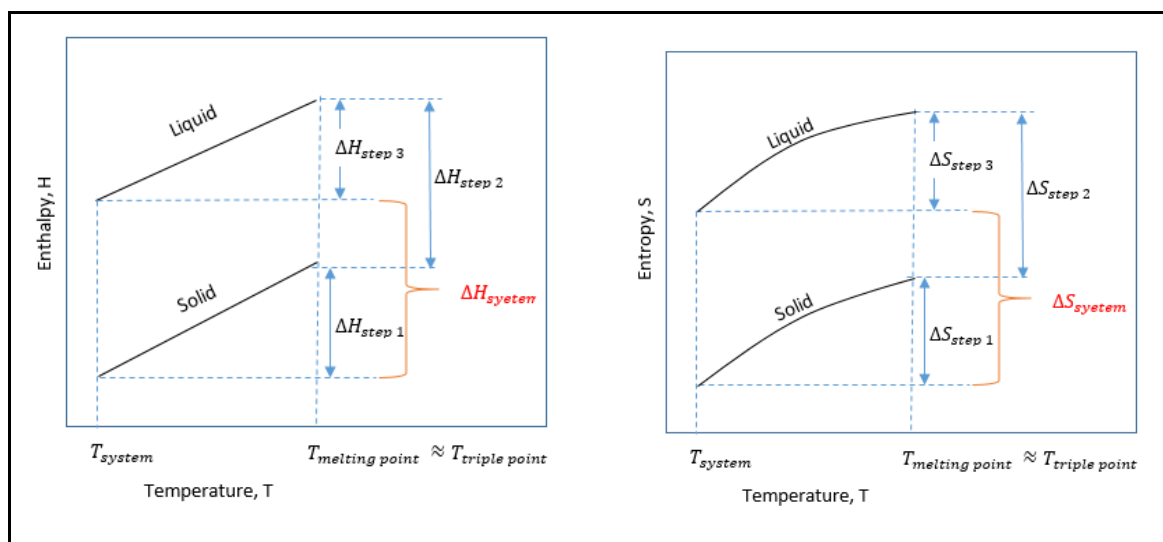


Figure 3.2. Enthalpy and Entropy changes during the sublimation process of a solid.

The Gibbs energy change for a transition from solid to liquid at system temperature is calculated as follows:

$$\Delta G_i = RT \ln \frac{f_i^l}{f_i^s} \quad (3.49)$$

The change in Gibbs energy can also be calculated with the fundamental equation for Gibbs energy. The Gibbs State function and the differential of the Gibbs energy, in terms of temperature and pressure, are:

$$G(T, P) = H(T, P) - TS(T, P) \quad (3.50)$$

$$dG(T, P) = dH(T, P) - TdS(T, P) \quad (3.51)$$

The differential of the entropy can be evaluated by considering the heat capacity and the Maxwell equation, which relate the derivative of the entropy to the pressure, and the derivative of the volume to temperature:

$$dS(T, P) = \left(\frac{\partial S}{\partial T}\right)_P dT + \left(\frac{\partial S}{\partial P}\right)_T dP = \frac{C_p}{T} dT - \left(\frac{\partial V}{\partial T}\right)_P dP = \frac{C_p}{T} dT - V\beta dP \quad (3.52)$$

Where, β is the volumetric thermal expansion coefficient, and $\beta = \frac{1}{V} \left(\frac{\partial V}{\partial T}\right)_P$; and the differential of the enthalpy is:

$$dH(T, P) = TdS(T, P) + VdP = C_p dT + V(1 - \beta T)dP \quad (3.53)$$

With reference to Figure 3.2, we have:

$$\Delta G_i = \Delta H_i - T\Delta S_i \quad (3.54)$$

Combining the two equations yields the ratio of the standard fugacity:

$$\ln \frac{f_i^l}{f_i^s} = \frac{\Delta G_i}{RT} = \frac{\Delta H_i}{RT} - \frac{\Delta S_i}{R} \quad (3.55)$$

The change in enthalpy and entropy can be determined from the thermodynamic cycle shown in Figure 3.2 and portrayed below.

Step 1	Step 2	Step 3
<i>Heating of solid from T to T_m</i>	<i>Phase change at T_m</i>	<i>Cooling of liquid from T_m to T</i>
$\Delta H_i = \int_T^{T_m} c_{p,i}^s dT + \int_P^{P_t} V^s(1 - T\beta^s)dP + \Delta H_i^{fusion} + \int_{T_m}^T c_{p,i}^l dT + \int_{P_t}^P V^l(1 - T\beta^l)dP$		
$\Delta S_i = \int_T^{T_m} \frac{c_{p,i}^s}{T} dT - \int_P^{P_t} V^s\beta^s dP + \frac{\Delta H_i^{fusion}}{T_m} + \int_{T_m}^T \frac{c_{p,i}^l}{T} dT - \int_{P_t}^P V^l\beta^l dP$		

Assuming that the difference in heat capacities and the volumes of the solid and liquid phases is constant for the temperature range:

$$\Delta c_{p,i} = c_{p,i}^l - c_{p,i}^s, \quad \text{and that} \quad V_i^l - V_i^s = \Delta V_i$$

Then:

$$\Delta H_i = \int_{T_m}^T \Delta C_{p,i} dT + \Delta H_i^{fusion} + \int_{P_t}^P \Delta V_i dP - \int_{P_t}^P T(V^l \beta^l - V^s \beta^s) dP \quad (3.56)$$

$$\Delta S_i = \int_{T_m}^T \frac{\Delta C_{p,i}}{T} dT + \frac{\Delta H_i^{fusion}}{T_m} - \int_{P_t}^P (V^l \beta^l - V^s \beta^s) dP \quad (3.57)$$

Therefore:

$$\begin{aligned} \Delta G_i = \int_{T_m}^T \Delta C_{p,i} dT + \Delta H_i^{fusion} + \int_{P_t}^P \Delta V_i dP - \int_{P_t}^P T(V^l \beta^l - V^s \beta^s) dP - T \int_{T_m}^T \frac{\Delta C_{p,i}}{T} dT - \\ \frac{T \Delta H_i^{fusion}}{T_m} + T \int_{P_t}^P (V^l \beta^l - V^s \beta^s) dP \end{aligned} \quad (3.58)$$

$$\Delta G_i = \int_{T_m}^T \Delta C_{p,i} dT + \Delta H_i^{fusion} \left(1 - \frac{T}{T_m}\right) + \int_{P_t}^P \Delta V_i dP - T \int_{T_m}^T \frac{\Delta C_{p,i}}{T} dT \quad (3.59)$$

$$\ln \frac{f_i^l}{f_i^s} = \frac{\Delta G_i}{RT} = \frac{\Delta C_{p,i}}{RT} (T - T_m) + \frac{\Delta H_i^{fusion}}{RT} \left(1 - \frac{T}{T_m}\right) - \frac{\Delta C_{p,i}}{R} \ln \frac{T}{T_m} + \frac{\Delta V_i (P - P_t)}{RT} \quad (3.60)$$

Which can be written as:

$$\ln \frac{f_i^l}{f_i^s} = \frac{\Delta H_i^{fusion}}{RT} \left(1 - \frac{T}{T_m}\right) - \frac{\Delta C_{p,i}}{R} \left[\frac{T_m}{T} - 1 - \ln \left(\frac{T}{T_m}\right) \right] + \frac{\Delta V_i (P - P_t)}{RT} \quad (3.61)$$

In the case of simple eutectic systems, the solid will crystallise out in pure form, hence equation (3.61) reduces to:

$$\begin{aligned} -\ln x_i^l \gamma_i^l &= \ln \left(\frac{f_i^l}{f_i^s} \right)_{pure\ i} \\ &= \frac{\Delta h_{m,i}}{RT} \left(1 - \frac{T}{T_{m,i}}\right) - \frac{\Delta c_{p,i}(T_{m,i}-T)}{RT} + \frac{\Delta c_{p,i}}{R} \ln \left(\frac{T_{m,i}}{T} \right) + \frac{\Delta V_i (P - P_t)}{RT} \end{aligned} \quad (3.62)$$

From equation (3.62), it can be seen that to calculate the ratio of the standard fugacity at a given temperature and pressure, only the melting temperature, the latent heat fusion and specific heat capacity of pure liquid (I) and pure solid (I) are required.

3.3. Activity Coefficient Models

The determination of the activity coefficient is essential, as it accounts for any non-ideal behaviour of the solution by accounting for the interactions that exist between the various molecules. Several analytical expressions for the concentration dependence of the excess Gibbs Energy have been developed, and hence, it is a fairly simple process to find an analytical expression for the activity coefficient. These expressions/models account for the various interactions between the various components in the various interacting phases, with binary interaction parameters. These models are generally classified into two broad techniques: correlative methods and predictive methods.

The correlative models require experimental data to regress and obtain model parameters. The binary interaction parameters for each of the solvent–solvent, solvent–solute, and solute–solute interactions are obtained from the regression binary phase equilibrium data. Two basic categories of models exist:

- a) Random mixing models, such as, Van Laar and Margules equations, and advanced, theoretically based models.
- b) Local composition models such as Wilsons, NRTL and UNIQUAC.

Unfortunately, the lack of experimental solubility data, on new APIs, limits the use of these thermodynamic models for conceptual process design and analysis in the pharmaceutical industry. (Chen, 2011).

There has been an increase in the use of predictive thermodynamic models for solubility calculations (as distinct from correlative models). The advantage of these predictive models is that no new experimental data is needed to calculate activity coefficients. The various predictive methods can be classified into four main groups:

- a) Models based on group contribution methods such as the Analytical-Solution-of-Groups (ASOG) (Kojima and Tochigi, 1979), the UNIQUAC Functional Group Activity Coefficient (UNIFAC) (Fredenslund et al., 1975), and its many derivatives, such as, modified UNIFAC (Grensemann and Gmehling, 2005) and Pharma Mod-UNIFAC (Diedrichs and Gmehling, 2010).
- b) Models based on quantum theory, such as, the Conductor-like Screening Model, abbreviated as COSMO, and its derivatives the COSMO-RS (Klamt, 2005), COSMO-SAC (Lin and Sandler, 2002) , and COSMO-RS(OL) (Grensemann and Gmehling, 2005).
- c) Models based on equation of state, such as, PC-SAFT (Kliener et al., 2009), Predictive Soave–Redlich–Kwong (PSRK) (Holderbaum and Gmehling, 1991), and the Volume-Translated Peng–Robinson (VTRP) (Ahlers and Gmehling, 2001)

- d) Segment Activity Coefficient Methods: Hansen Solubility Parameters, Modified Separation of Cohesive Energy Density (MOSCED) (Lazzaroni et al., 2005), Non-Random Two-Liquid - Segment Activity Coefficient (NRTL-SAC) (Chen and Song, 2004), Functional-segment activity coefficient (F-SAC) (Soares and Gerber, 2013); and UNiversal Segment Activity Coefficient model (UNISAC) (Moodley et al., 2015).

The accuracy of any model depends upon how closely the system of interest adheres to the model assumptions, and how accurately the required parameters are measured or predicted.

3.3.1. Correlative Models

If solubility data is available, it can be used to determine the parameters of correlative thermodynamic models. These predictive models can, in turn, calculate the activity coefficients required for phase equilibria calculations. Examples of the standard, local-composition correlative models applied for the estimation of pharmaceutical solubility data, are summarised here.

3.3.1.1. Wilson Model

Wilson (1964) developed a model for correlating phase data based on a “local-composition” concept. This theory implies that there is a distinct difference between the local and overall mixture composition, due to intermolecular forces and differences in molecular size, which result in a short-range order and non-random molecular orientation (Van Ness and Abbott, 1997). The Wilson model can be defined by means of the following excess Gibbs energy expression for a multicomponent system:

$$\frac{G^E}{RT} = \sum_{i=1}^m x_i \ln \left(\sum_{j=1}^m x_j \Lambda_{ij} \right) \quad (3.63)$$

Where the Wilson parameters are given by: $\Lambda_{ij} = \frac{v_j^l}{v_i^l} \exp \left(-\frac{\Delta\lambda_{ij}}{T} \right)$ and v_i^l and v_j^l are molar volumes of pure liquids at temperature T.

The activity coefficients for the Wilson model are given by:

$$\ln \gamma_i = 1 - \ln \left(\sum_j x_j \Lambda_{ij} \right) - \sum_k \left(\frac{x_k \Lambda_{ki}}{\sum_j x_j \Lambda_{kj}} \right) \quad (3.64)$$

3.2.1.2. Non-Random Two Liquid (NRTL) Model

Renon and Prausnitz (1968) proposed the NRTL model as an improvement on the Wilson model. The model is premised on the “local-composition” concept of Wilson (1964) and the two-liquid theory of Hildebrand and Scott (1964), and is based exclusively on molecular interactions. Three

adjustable parameters constitute the NRTL model, i.e., τ_{ij} , τ_{ji} and α_{ij} (symmetrical parameters). The NRTL model can be defined by the following excess Gibbs energy expression for a multicomponent system:

$$\frac{G^E}{RT} = \sum_{i=1}^m x_i \frac{\sum_{j=1}^m \tau_{ji} G_{ji} x_j}{\sum_{l=1}^m G_{li} x_l} \quad (3.65)$$

Where:

$$G_{ij} = \exp(-\alpha_{ij} \tau_{ij}) \quad \tau_{ij} = \frac{\Delta g_{ij}}{T} \quad \tau_{ii} = 0 \quad \text{and} \quad G_{ii} = 1$$

Δg_{ij} is the interaction parameter between components i and j , and α_{ij} is characterised as the non-randomness parameter, with values ranging from 0 to 1.

The activity coefficients for the NRTL model are given by:

$$\ln \gamma_i = \frac{\sum_j x_j \tau_{ji} G_{ji}}{\sum_k x_k G_{ki}} + \sum_j \frac{x_j G_{ij}}{\sum_k x_k G_{kj}} \left(\tau_{ij} - \frac{\sum_m x_m \tau_{mj} G_{mj}}{\sum_k x_k G_{kj}} \right) \quad (3.66)$$

3.2.1.3. Universal Quasi-Chemical (UNIQUAC) Model

The Universal Quasi-Chemical (UNIQUAC) excess Gibbs energy model was developed by Abrams and Prausnitz (1975). It combines characteristics of the Wilson model and the NRTL model, and can be viewed as an extension of the quasi-chemical lattice theory of Guggenheim (1952). This model is based on both an entropic contribution (attributed to the distinct composition, size and shape of molecules), and intermolecular interactions. Abrams and Prausnitz (1975) recommend the UNIQUAC model for the correlation of VLE and LLE data for both binary and multi-component systems. There are two components that constitute its defining equation, namely, the combinatorial and the residual components, which account for the entropic and enthalpic contributions respectively.

$$\frac{G^E}{RT} = \left(\frac{G^E}{RT} \right)_{\text{Combinatorial}} + \left(\frac{G^E}{RT} \right)_{\text{Residual}} \quad (3.67)$$

$$\text{Where} \quad \left(\frac{G^E}{RT} \right)_{\text{combinatorial}} = \left[\sum_{i=1}^m x_i \ln \frac{\phi_i}{x_i} + \frac{z}{2} \sum_{i=1}^m q_i x_i \ln \frac{\theta_i}{\phi_i} \right] \quad (3.68)$$

$$\left(\frac{G^E}{RT} \right)_{\text{Residual}} = - \left[\sum_{i=1}^m q_i x_i \ln \left(\sum_{j=1}^m \theta_j \tau_{ji} \right) \right] \quad (3.69)$$

Therefore, the UNIQUAC model can be written as:

$$\frac{G^E}{RT} = \left[\sum_{i=1}^m x_i \ln \frac{\phi_i}{x_i} + \frac{z}{2} \sum_{i=1}^m q_i x_i \ln \frac{\theta_i}{\phi_i} \right] - \left[\sum_{i=1}^m q_i x_i \ln \left(\sum_{j=1}^m \theta_j \tau_{ji} \right) \right] \quad (3.70)$$

Where the activity coefficient for the system can be expressed as:

$$\ln \gamma_i = (\ln \gamma_i)_{Combinatorial} + (\ln \gamma_i)_{Residual} \quad (3.71)$$

$$(\ln \gamma_i)_{comb} = \ln \left(\frac{\phi_i}{x_i} \right) + \left[1 - \frac{\phi_i}{x_i} \right] - 5q_i \left[\ln \frac{\phi_i}{\theta_i} + \left(1 - \frac{\phi_i}{\theta_i} \right) \right] \quad (3.72)$$

$$(\ln \gamma_i)_{resd} = q_i \left[1 - \ln \sum_k \theta_k \tau_{ki} - \sum_j \left(\frac{\theta_j \tau_{ji}}{\sum_k \theta_k \tau_{ki}} \right) \right] \quad (3.73)$$

$$\text{Where} \quad \theta_i = \frac{q_i x_i}{\sum_j q_j x_j} \quad \phi_i = \frac{r_i x_i}{\sum_j r_j x_j} \quad \tau_{ij} = \left[-\frac{(u_{ij} - u_{jj})}{RT} \right]$$

Where Φ_i is the average segment fraction of component i ; θ_i is the average surface area fraction of molecule i ; τ_{ij} is an adjustable parameter containing interaction terms, u_{ij} and u_{ji} ; r is a molecular size parameter; q is an external surface area parameter; z is the coordination number and is equal to 10; and u_{ij} and u_{ii} are adjustable energy parameters.

3.3.2. Predictive Methods:

The use of predictive thermodynamic models form an integral part in the development of a conceptual design computational tool, as calculations are generally required on systems for which no experimental measured data may be available. In this section, a brief description is presented on a selection of predictive activity coefficient models that have been applied to predict solubility and multiphase equilibria. Currently, the most successful predictive models for the generation of activity coefficients are the group contribution methods, such as, UNIFAC and modified UNIFAC (Grensemann and Gmehling, 2005).

UNIFAC models have proven to be reliable predictive models, computationally efficient, easy to program, and they have a wide application range whenever experimental data are not available. They are embedded in most commercial, chemical engineering simulators, such as, ASPEN, CHEMCAD, PROII, and ProSim (Xue et al., 2012).

Other predictive methods that are gaining popularity because they overcome some of the limitations of the UNIFAC models in certain fields of applications, are NRTL-SAC, COSMO-RS and its variations, and PC-SAFT. Some recently developed models, like the F-SAC and UNISAC have also shown some success.

3.3.2.1. Group Contribution Methods

UNIFAC, which is an acronym for UNIQUAC Functional Group Activity Coefficient, is classed as a group contribution (GC) method. The basic assumption behind this method is that a molecule is a construct of functional groups, and that a mixture does not consist of molecules but of functional groups. It is further assumed that each of the several functional groups that make up a

molecular entity, make a separate and additive contribution to a property of the mixture being considered.

The basic structure of the UNIFAC model is founded on four rules laid out by Wilson and Deal (1962), when they developed the first solution of the group method to determine activity coefficients, known as the Analytical Solution of Groups (ASOG). Wilson and Deal (1962) suggested four rules to describe a solution of groups approach. These are:

1. The partial molal excess free energy or, simply, the logarithm of the activity coefficient of a component, is assumed to be the sum of two contributions: one associated with differences in molecular size; and the other with interactions of structural groups. For molecular solute, j , in any solution:

$$\ln \gamma_i = \ln \gamma_i^S + \ln \gamma_i^G \quad (3.74)$$

2. The contribution associated with molecular size differences is assumed to be given by a Flory-Huggins relation, expressed in terms of the number of constituent atoms other than hydrogen. For solute, j , in solution of component i :

$$\ln \gamma_i^S = \ln \frac{n_i}{\sum_i x_i n_i} + 1 - \frac{n_i}{\sum_i x_i n_i} \quad (3.75)$$

Where

n_i = number of atoms (other than hydrogen) in molecular component i

x_i = molecular mole fraction of component i

And the summation is made over all components in solution.

3. The contribution from interactions of molecular “groups” is assumed to be the sum of the individual contributions of each solute “group” in the solution, less the sum of the individual contributions in the conventional standard state environment. For molecular solute, j , containing groups k :

$$\ln \gamma_i^G = \sum_k v_{kj} (\ln \Gamma_k - \ln \Gamma_k^*) \quad (3.76)$$

Where

v_{kj} = number of groups of type k in solute component j

Γ_k = activity coefficient of group k in the solution environment

Γ_k^* = the activity coefficient of group k in pure liquid j

4. Finally, the individual group contributions, Γ_k , in any environment, containing groups of given kinds, are assumed to be only a function of group concentrations.

$$\Gamma_k = f(X_1 \dots \dots X_k) \quad (3.78)$$

Where

$$X_k = \text{group fraction of group } k = \frac{\sum_j v_{kj} x_j}{\sum_i \sum_j v_{ij} x_j} \quad (3.79)$$

Using these principles, Derr and Deal (1969) developed their Analytical Solution of Groups (ASOG) method for correlating and predicting activity coefficients. Later, Kojima and Tochigi (1979) increased the range of application of the ASOG approach for a wider range of compounds, by adding more group parameters. The original UNIFAC (UNIQUAC Functional-group Activity Coefficients) group contribution method, developed by Fredenslund et al. (1977), combines the ASOG approach with a model for activity coefficients based on an extension of the quasi chemical theory of liquid mixtures (UNIQUAC).

Several modifications of the original UNIFAC model, such as modified UNIFAC (Lyngby) (Larsen et al., 1987), modified UNIFAC (Dortmund) (Weidlich and Gmehling, 1987), KT-UNIFAC (Kang et al., 2002), and Pharma-modified UNIFAC (Diedrichs and Gmehling, 2010), have been proposed to improve the performance of the model.

In essence, an activity coefficient group contribution method consists of the following key components:

- The definition of the functional groups used to “build” the molecules. These groups are generally any convenient structural unit that have been decided by the developers.
- Each functional group has area and volume structural parameters, and each functional group pair has two unique binary interaction parameters associated with that pair. The group interaction parameter-matrix database is developed from the regression of large sets of thermodynamic-consistent experimental data. Knowing the parameters of groups constituting the mixture is necessary for implementation.
- An equation to calculate the combinatorial contribution.
- An equation to calculate the residual contribution.

These components can then be used to predict the activity coefficients for other systems for which no experimentally obtained data is available, but which contain the same functional groups.

A brief description of some of the activity coefficient group contribution method equations are presented, and the main differences of the various versions are also described. Some of the essential differences in the various formulations of the UNIFAC Group Contribution Method are:

- a) The adoption of either a Flory-Huggins or the Staverman-Guggenheim type formulation for the combinatorial contribution term. The Staverman-Guggenheim potential consists of the original Flory-Huggins combinatorial along with a Staverman-Guggenheim

correction term to compensate for molecular ring formation as well as bending of flexible and branched molecules and crosslinks (Muzenda, 2013).

- b) The computation of the various contributing terms in the Flory-Huggins or the Staverman-Guggenheim type formulation may differ.
- c) Whilst the formulation of the residual contribution term is similar in all formulations, the computation of the temperature dependency of the group interaction parameters may differ.

In addition to the above differences, the definition of the type of groups may vary, and hence the values for the various parameters (obtained from the regression of experimental data) associated with the groups will differ.

3.3.2.1.1. Analytical Solution of Groups (ASOG):

The combinatorial term is estimated by using the athermal Flory- Huggins equation, which is expressed in terms of the number of constituent atoms in molecule i , other than hydrogen. This results in the combinatorial contribution with the summation being made over all the components j , in the solution:

$$\ln \gamma_i^C = \ln v_i^{FH} - \ln \sum_{j=1}^N x_j v_j^{FH} + \frac{1-v_i^{FH}}{\sum_{j=1}^N x_j v_j^{FH}} \quad (3.80)$$

Where v_i^{FH} is the measure of the size of molecule i

The group interactions contribution is:

$$\ln \gamma_i^G = \sum_k v_{kj} (\ln \Gamma_k - \ln \Gamma_k^*) \quad (3.81)$$

Where Γ_k (activity coefficient of group k) is given by Wilson's equation:

$$\ln \Gamma_k = 1 - \ln \left(\sum_{n=1}^{NG} X_n a_{kn} \right) - \sum_{n=1}^{NG} \left[\frac{X_n a_{nk}}{\sum_{m=1}^{NG} X_m a_{nm}} \right] \quad (3.82)$$

Where the group interaction parameter, a_{kn} , characteristic of groups k and n , is defined as:

$$a_{kn} = \exp \left(\frac{m_{kn} + n_{kn}}{T} \right) \quad (3.83)$$

And m_{kn} and n_{kn} are the group pair parameters, characteristic of groups k and n , which are independent of temperature, and:

$$X_k = \text{group fraction of group } k = \frac{\sum_{i=1}^{NC} x_i v_{ni}}{\sum_{i=1}^{NC} x_i \sum_{k=1}^{NG} v_{ki}} \quad (3.84)$$

3.3.2.1.2. Original UNIFAC

As with the ASOG, the UNIFAC model consists of a combinatorial contribution, which describes the excess Gibbs energy arising due to differences in molecular size and shape, and a residual term, which describes the excess Gibbs energy differences due to molecular interactions. In the UNIFAC method, the combinatorial contribution is estimated by using the UNIQUAC equation that contains differences in size and shape of the molecules in the mixture. In addition, functional group sizes and interaction surface areas are introduced from molecular structure data for pure compounds.

The UNIFAC equations for computing liquid phase activity coefficient is defined as:

$$\ln \gamma_i = \ln \gamma_i^C + \ln \gamma_i^R, \quad (3.85)$$

The combinatorial term is expressed as:

$$\ln \gamma_i^C = \ln \frac{\Phi_i}{x_i} + \frac{z}{2} q_i \ln \frac{\theta_i}{\Phi_i} + l_i - \frac{\Phi_i}{x_i} \sum_j x_j l_j \quad (3.86)$$

$$\text{Where:} \quad \theta_i = \frac{q_i x_i}{\sum_j q_j x_j} \quad \Phi_i = \frac{r_i x_i}{\sum_j r_j x_j} \quad l_i = \frac{z}{2} (r_i - q_i) - (r_i - 1)$$

Where θ_i is the molecular surface area fraction, and Φ_i the molecular volume fraction of component i . The component area parameter, q_i , is calculated as the sum of the group area parameters, Q_k , and the component volume parameter, r_i , is calculated as the sum of the group volume parameters, R_k . The area parameters and volume parameters are obtained from published tables, or calculated from van der Waals group surface areas, A_{wk} , and van der Waals group volume, V_{wk} , respectively, as defined by Bondi in 1968.

$$q_i = \sum_k v_k^{(i)} Q_k \quad r_i = \sum_k v_k^{(i)} R_k$$

Where v_k^i is the number of groups of kind k , in a molecule of component i , and:

$$Q_k = \frac{A_{wk}}{2.5 \times 10^9}, \text{ and } R_k = \frac{V_{wk}}{15.17}$$

The group residual activity coefficients are computed similarly to the residual part of the UNIQUAC equation but are adapted to conform to the solution-of-groups concept, and are expressed as:

$$\ln \gamma_i^R = \sum_k v_k^i \left(\ln \Gamma_k - \ln \Gamma_k^{(i)} \right) \quad (3.87)$$

The parameter Γ_k , is the residual activity coefficient of the group k , whereas parameter, $\Gamma_k^{(i)}$, represents the residual activity coefficient of group k , in a reference pure solution, which contains only molecules of type i . The group residual activity coefficients are calculated using:

$$\ln \Gamma_k = \ln \Gamma_k^{(i)} = Q_k \left[1 - \ln(\sum_m \Theta_m \Psi_{mk}) - \sum_m \frac{\Theta_m \Psi_{km}}{\sum_n \Theta_n \Psi_{nm}} \right] \quad (3.88)$$

Where Θ_m group area fraction and Ψ_{mk} are defined as the group interaction parameters which are expressed as:

$$\Theta_m = \frac{Q_m X_m}{\sum_n Q_n X_n} \quad \text{and} \quad \Psi_{mn} = \exp - \left[\frac{U_{mn} - U_{nn}}{T} \right] = \exp - \frac{a_{mn}}{T}$$

Where X_m is the mole fraction of group m , in the mixture, and a_{mn} is the interaction parameter between groups m and n :

$$X_m = \frac{\sum_j v_m^j X_j}{\sum_j \sum_n v_n^j X_j}$$

Limitations of the original UNIFAC have been highlighted by Fredenslund and Rasmussen (1985), Kontogeorgis and Folas (2010) and Gmehling et al. (2012):

- It cannot differentiate between isomers; because the assumptions of the solution-of-functional-groups concept uses only first order description of molecules.
- It cannot satisfactorily predict liquid-liquid equilibrium (LLE); mainly because the interaction parameters are characterized only by vapor liquid equilibrium (VLE) data.
- It poorly represents dilute systems and the prediction of complex systems, containing water and multifunctional chemicals; because of proximity effects (occurring when polar groups are close to each other).
- It is limited in application range to low pressures of between 10-15 atm, and to temperatures of 275-425 K, depending on the range of phase equilibria data temperatures used to regress the interaction parameters.
- It poorly predicts properties, such as, heat of mixing and infinite dilution activity coefficients; because of the weak temperature dependency of the interaction parameters.
- It is not applicable for non-condensable gases, electrolytes and polymers.

In order to improve on the accuracy and robustness of the original UNIFAC model, and to overcome the weaknesses listed above, different modifications to UNIFAC have been proposed by various developers. Changes to both combinatorial and residual terms, as well as the introduction of temperature-dependent group interaction parameters, are some of the modifications introduced. Some of these modifications are listed in the following sections.

3.3.2.1.3. Modified UNIFAC (Lyngby)

The Modified UNIFAC (Lyngby) Model by Larsen, Rasmussen and Fredenslund, published in 1987, incorporated the several changes into the combinatorial and residual terms.

In the combinatorial term, the Staverman-Guggenheim (S-G) correction to the Flory-Huggins combinatorial was dropped. The developers reported that the effects of the S-G corrections were often negligible and some excessively large S-G corrections resulted in negative excess entropy values, which were considered unrealistic by the developers. In addition, a modified group volume fraction parameter is introduced. The resulting combinatorial term is:

$$\ln \gamma_i^C = \ln \frac{\Phi_i}{x_i} + 1 - \frac{\Phi_i}{x_i} \quad (3.89)$$

Where the modified group volume fraction parameter Φ_i , is expressed as:

$$\Phi_i = \frac{x_i r_i^{2/3}}{\sum_j x_j r_j^{2/3}} \quad (3.90)$$

The residual term was left unchanged from the Original UNIFAC version. However a logarithmic temperature dependence was introduced to the interaction parameter:

$$a_{mn} = a_{mn,0} + a_{mn,1}(T - T_0) + a_{mn,2} \left(T \ln \frac{T_0}{T} + T - T_0 \right) \quad (3.91)$$

The developers reported an improved prediction of VLE and excess enthalpies. It is also capable of presenting liquid-liquid equilibria using the modified UNIFAC-VLE parameters.

3.3.2.1.4. Modified UNIFAC (Dortmund)

The Modified UNIFAC (Do) model, published by Gmehling and co-workers in 1987 (the same year as the Modified UNIFAC (Lyngby)), and incorporated the following changes to the Original UNIFAC:

- a) The Staverman- Guggenheim correction term as in the original UNIFAC was retained but the group volume parameter was modified.
- b) As was the case for Modified UNIFAC (Lyngby), temperature-dependent group interaction parameters were introduced
- c) Van der Waals volume and surface area parameters were introduced for cyclic alkanes, and alcohols were reclassified as primary, secondary and tertiary alcohols with their own van der Waals volume and surface area parameters.
- d) The fitting of Modified UNIFAC group interaction parameters was extended to include activity coefficients at infinite dilution, VLE and excess enthalpies in order to improve the accuracy of these parameters.

The resulting Modified UNIFAC (Do) model incorporated the following changes into the combinatorial and residual terms

$$\ln \gamma_i^C = \ln \frac{\varphi_i'}{x_i} + 1 - \frac{\varphi_i'}{x_i} - \frac{z}{2} q_i \left(\ln \frac{\Phi_i}{\theta_i} + 1 - \frac{\Phi_i}{\theta_i} \right) \quad (3.92)$$

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

In addition, the *van der Waals* volume and surface parameters R_k and Q_k of the structural UNIFAC groups were not calculated using the Bondi rules, but were optimized together with the group interaction parameters, using experimental data.

The residual term was left unchanged from the Original UNIFAC version. However a temperature dependence was introduced to the interaction parameter:

$$a_{mn} = a_{mn,0} + a_{mn,1}T + a_{mn,2}T^2 \quad (3.93)$$

Comparative studies by Voutsas and Tassios (1996) and Lohmann et al. (2001) on the performance of the various formulations of UNIFAC reported the Modified UNIFAC (Do) showed a greater range of applicability and reliability in most cases studied.

3.3.2.1.5. The Pharma-modified UNIFAC model

This model is a derivative of modified UNIFAC (Dortmund). It was established recently to overcome some limitations observed when the model is applied to API solutions. The equations of this new version of modified UNIFAC are the same as in the original model, excepting in the set of available functional groups, their volume and surface area parameters, and their interaction parameters. The Pharma-Mod model more accurately predicts complex compounds solubility data compared to the other UNIFAC models in most cases. However the parameter matrix requires further development to broaden its application.

3.3.2.1.6. KT UNIFAC Model

To overcome the limitation of distinguishing between isomers, and to handle systems with proximity effects, Kang et al., (2002) have proposed a model called the KT-UNIFAC model. In this formulation, the molecular structure of a compound is considered as a set of first-order and second-order groups. Estimation is performed at two levels: the basic level uses contributions from first-order simple groups, while the second level uses a small set of second-order groups, having the first-order groups as building blocks. The second-order groups provide more structural information on molecular fragments and take into account proximity effects, distinguishing among isomers.

The activity coefficient, in the KT-UNIFAC model, is calculated by addition of three contributions: a combinatorial term, a residual term, and a second-order residual term. The first one takes into account the molecular size and shape, the second, the molecular interactions, and the third, the second-order effects on molecular interactions:

$$\ln\gamma_i = \ln\gamma_i^C + \ln\gamma_i^R + w_{R2}\ln\gamma_i^{R2} \quad (3.94)$$

Where the term $w_{R2} = 0$ for the first order model, and $w_{R2} = 1$ for the second order model. The combinatorial and residual terms are the same as in the Modified UNIFAC (Lyngby), and group volume and group surface parameters are based on the methods of Bondi. The second order residual term calculates the activity coefficients due to the second order interactions. In the residual parts, the interaction parameters are expressed as linearly dependent on temperature:

$$a_{mn} = a_{mn,1} + a_{mn,2}(T - T_0) \quad (3.95)$$

The KT UNIFAC Model has its own set of first-order and second-order structural groups. A list of first-order and second-order groups, along with sample assignments, group occurrences, volume parameters, surface area parameters and main-group interaction parameters (first-order and second-order), can be found in Kang et al. (2002). Compared with some of the currently-used versions of UNIFAC, the KT-UNIFAC model makes significant improvements in accuracy, while providing a much wider range of applicability.

Other versions of the UNIFAC model have been developed to extend the application range of the original model. For various reasons, UNIFAC has been the preferred choice of foundation model, to be extended, either by adding new terms or by regressing new parameters, showing its usefulness and versatility.

Today, the Dortmund-modified UNIFAC has, probably, the most extensive parameters. It is widely used, and is continuously revised and extended in industrial/academic joint ventures. In recent years, numerous researchers have added new groups and subgroups, and have published new group parameters for this model (Abildskov et al., 2004).

3.3.2.2. NRTL-SAC

The NRTL-SAC model, first published by Chen and Song (2004), is derived from the polymer NRTL model of Chen (1993), which in turn is developed from the original NRTL model of Renon and Prausnitz (1968).

While the UNIFAC methods define a molecule by means of a set of predefined functional groups, the NRTL-SAC method defines the molecule in terms of a set of four predefined conceptual segments. As with Group contribution methods of UNIFAC and ASOG, the NRTL-SAC model also assumes that the activity coefficient consists of a combinatorial and residual part. The

combinatorial expression used is analogous to the Flory-Huggins expression, used in ASOG, whilst the residual part sums up the local interaction contributions of the predefined segments in a mixture.

The NRTL-SAC is considered a ‘hybrid’ model, because it initially requires selected experimental solubility data for the solute in a minimum of four reference solvents in order to determine its surface interaction parameters. Once the solute is characterised and its surface interaction parameters are known, the model can then be used as a predictive model.

The NRTL-SAC model describes a molecule by means of their potential, effective surface interactions, in terms of four types of segments: hydrophobic segment, electrostatic solvation segment, electrostatic polar segment, and hydrophilic segment. The four segment numbers, calculated for each molecule, are measures of the effective surface areas of the molecule that exhibit surface interaction characteristics of hydrophobicity (X), solvation (Y^-), polarity (Y^+), and hydrophilicity (Z). These four segments are used to describe the following types of inter-molecular interactions: the hydrophilic segment simulates polar molecular surfaces that are “hydrogen bond donor or acceptor”. The hydrophobic segment simulates molecular surfaces that show aversion to forming hydrogen bonds. The electrostatic segments (Y^- and Y^+) simulate molecular surfaces that are electron pair donor or acceptor. The electrostatic solvation segment is attractive to the hydrophilic segment, while the electrostatic polar segment is repulsive to the hydrophilic segment.

To determine the segment numbers of a solute molecule, solubility data in at least four reference solvents is needed. Experimental solubility data for the solute is measured in four to eight solvents with distinctive surface interaction characteristics, at or near room temperature. These solvents should include hydrophobic solvents, such as, hexane or heptane, hydrophilic solvents, such as, water and methanol, and polar solvents, such as, acetone, acetonitrile, DMSO, DMF, etc. The data obtained is regressed to determine the molecular parameters. The segment numbers obtained for a molecule defines the surface area that is accountable for that segment’s prescribed mode of interaction. The individual segments are assumed to interact with each other through binary interactions only, and they are calculated with binary interaction parameters that are constants of the model. Once the segment numbers of the solute molecule are determined from the data, the NRTL-SAC model can then be used to predict the solute solubility in other solvents or solvent mixtures.

A databank of molecular segment parameters for the common pure solvents and some compounds has been established, and is continually being developed by the authors. It is available with the commercial simulator ASPEN as an add-on module. The parameter tables are

likely to change as new equilibrium data and solvents are added to improve its accuracy and functionality.

An overview of the NRTL-SAC model will be given here with sufficient detail to understand its application. Full details of the model and its mathematical formulation are given in Chen and Song, (2004). The NRTL-SAC model computes the activity coefficient for component i from the expression:

$$\ln \gamma_i = \ln \gamma_i^C + \ln \gamma_i^R \quad (3.96)$$

Where the combinatorial part is calculated from:

$$\ln \gamma_i^C = \ln \frac{\phi_i}{x_i} + 1 - r_i \sum_j \frac{\phi_j}{x_j} \quad (3.97)$$

With the definitions:

$$\phi_i = \frac{r_i x_i}{\sum_j r_j x_j} \quad \text{and} \quad r_i = \sum_j r_{j,i}$$

Where x_i is the mole fraction of component i ; r_i is the total segment number of component i ; and ϕ_i is the segment mole fraction in the mixture. The residual term is defined as:

$$\ln \gamma_i^R = \sum_m r_{m,i} (\ln \Gamma_m^{lc} - \Gamma_m^{lc,i}) \quad (3.98)$$

Where the terms Γ_m^{lc} and $\Gamma_m^{lc,i}$ are the activity coefficients of segment m in solution, and in pure component i , respectively; and $r_{m,i}$ is the number of segment m in component i . These terms can be evaluated by:

$$\ln \Gamma_m^{lc} = \frac{\sum_j x_j G_{jm} \tau_{jm}}{\sum_k x_k G_{km}} + \sum_{m'} \frac{x_{m'} G_{mm'}}{\sum_k x_k G_{km'}} \left(\tau_{mm'} - \frac{\sum_j x_j G_{jm'} \tau_{jm'}}{\sum_k x_k G_{km'}} \right) \quad (3.99)$$

$$\ln \Gamma_m^{lc,i} = \frac{\sum_j x_{j,i} G_{jm} \tau_{jm}}{\sum_k x_{k,i} G_{km}} + \sum_{m'} \frac{x_{m',i} G_{mm'}}{\sum_k x_{k,i} G_{km'}} \left(\tau_{mm'} - \frac{\sum_j x_{j,i} G_{jm'} \tau_{jm'}}{\sum_k x_{k,i} G_{km'}} \right) \quad (3.100)$$

Where: $x_j = \frac{\sum_i x_i r_{j,i}}{\sum_i \sum_n x_i r_{i,n}}$ and $x_{j,i} = \frac{r_{j,i}}{\sum_n r_{n,i}}$

And n, j, k, m , and m' are the segment species index; x_i is the mole fraction of component i ; and x_j is the mole fraction of segment species j . G and τ are local binary quantities related to each other by the NRTL non-random factor parameter α : $G = \exp(-\alpha\tau)$.

Limitations of the model have identified by Kontogeorgis and Folas (2010) as follows: Firstly, the regression models are based on a simplified SLE equation, where the ΔC_p term is ignored. Whereas, recent investigations have shown that this term may be important in pharmaceutical

calculations. Secondly, all the interaction parameters are temperature independent, and their degree of accuracy varies, depending on the complexity of the molecule.

3.3.2.3. COSMO-RS and COSMO-SAC

Recent developments in computational chemistry yielded models based on the so-called continuum solvation family of models called COSMO (CONductor-like Screening MOdel). They use quantum mechanics calculations (for surface interaction energies), combined with statistical thermodynamics calculations, in order to predict the needed macroscopic thermodynamic properties. The key concepts have been extended to increase the applications of the model. COSMO-RS were pioneered by Klamt (1995), with further variations, like COSMO-SAC, proposed by Lin and Sandler (2002), and COSMO-RS(OL), proposed by Grensemann and Gmehling (2005).

These models are based on a three-step process: a molecule is deconstructed into a collection of very small surface elements, the charge density on each surface element is computed using a quantum electrostatic calculation. Along with a statistical mechanics analysis, this polarization charge density is used for quantification of the interaction energy of pairwise interacting surface segments. The unique characteristic of each molecule is its sigma (σ) profile that is a representation of charge density. The sigma profile for a pure component is calculated as:

$$p_i(\sigma) = \frac{A_i(\sigma)}{A_i} \quad (3.101)$$

In the COSMO-RS(OL) model, a further averaging is applied to obtain the final σ -profile. This additional averaging greatly increases the robustness of the method in the case of hydrogen bonding molecules (Constantinescu, 2009). The final σ -profile in COSMO-RS(OL) model is calculated as:

$$p_i(\sigma_n)_{\text{COSMO-RS(OL)}} = \frac{1}{3} \sum_{n=1}^{n+1} p_i(\sigma_n) \quad (3.102)$$

Where $A_i(\sigma)$ is the surface area of species i , with a screening charge density of σ ; and $A_i = \sum_{\sigma} A_i(\sigma)$ is the total surface area of species i . For a mixture, S , the σ -profile is determined as a mole fraction weighted average of the pure component contributions as:

$$p_s(\sigma) = \frac{\sum_{i=1}^c x_i A_i(\sigma)}{\sum_{i=1}^c x_i A_i} = \frac{\sum_{i=1}^c x_i A_i p_i(\sigma)}{\sum_{i=1}^c x_i A_i} \quad (3.103)$$

In COSMO-based models, in a very similar way to the UNIFAC models, the activity coefficient is computed as the result of two contributions:

$$\ln \gamma_{i/S} = \ln \gamma_{i/S}^{\text{Combinatorial}} + \ln \gamma_{i/S}^{\text{Residual}} \quad (3.104)$$

The following expressions are used by the COSMO-RS and COSMO-SAC models:

$$\ln\gamma_{i/S}^{COSMO-RS} = n_i \sum_{\sigma_m} p_i(\sigma_m) \exp \left[\frac{\mu_s(\sigma_m) - \mu_i(\sigma_m)}{RT} \right] - \lambda \ln \left(\frac{A_s}{A_i} \right) + \ln\gamma_{i/S}^{SG} \quad (3.105)$$

Where $\mu_s(\sigma_m)$ is the chemical potential of segment σ_m in the mixture; $\mu_i(\sigma_m)$ is the chemical potential of segment σ_m in pure component I ; A_s is the mole fraction weighted surface area of all the species in the mixture. Similarly, A_i is for pure component i and λ is a solvent specific adjustable parameter.

$$\ln\gamma_{i/S}^{COSMO-SAC} = n_i \sum_{\sigma_m} p_i(\sigma_m) [\ln\Gamma_s(\sigma_m) - \ln\Gamma_i(\sigma_m)] + \ln\gamma_{i/S}^{SG} \quad (3.106)$$

Where $\ln\Gamma_s(\sigma_m)$ is the segment activity coefficient of segment σ_m in the mixture; and $\ln\Gamma_i(\sigma_m)$ is the segment activity coefficient of segment σ_m in pure component I ; and $\ln\gamma_{i/S}^{SG}$ is the Staverman-Guggenheim (SG) combinatorial model, which accounts for the non-ideality that results from the size and shape difference between the species in a mixture. The volume and surface area parameters in the SG-combinatorial term are obtained from first principle COSMO calculations.

3.3.2.4. Functional-Segment Activity Coefficient Model (F-SAC)

This new model, developed by Soares and Gerber (2013), is based on a combination of concepts from two, well-developed methods: the group contribution models, such as UNIFAC, and the COSMO-SAC formulation, from which is drawn the interaction parameters between groups. The model assumes that each pure substance consists of several predefined functional groups (like the UNIFAC models), and that each functional group has its own apparent surface charges. Instead of using COSMO calculations to obtain the σ -profiles for the molecule, it is proposed that each predefined functional group has its own, empirically calibrated σ -profile, $p_k(\sigma)$, and that the σ -profile of a molecule i is given by the sum of the σ -profiles of the functional groups that make up the molecule.

In this model, each functional group is represented by three sets of empirical parameters: the functional group segment areas and the charge density of the positive, neutral and negative segments i.e.

$$p_k(\sigma)Q_k = \{(\sigma_k^+, Q_k^+); (0, Q_k^0); (\sigma_k^-, Q_k^-)\} \quad (3.107)$$

And that the σ -profile for the molecule i is the sum of σ -profiles of the functional groups that make-up molecule i :

$$p_i(\sigma)q_i = \sum_k v_k^i p_k(\sigma)Q_k \quad (3.108)$$

Where v_k^i is the number of functional groups of type k in molecule i .

F-SAC uses the Gibbs excess energy (GE) model, as the sum of two parts: a combinatorial contribution that accounts for differences in size and shape; and a residual contribution, which should include the differences in intermolecular forces between components.

$$\ln\gamma_{i/S} = \ln\gamma_{i/S}^{Combinatorial} + \ln\gamma_{i/S}^{Residual} \quad (3.109)$$

Where the F-SAC combinatorial contribution is similar to that of the modified UNIFAC (Do), as follows:

$$\ln\gamma_i^C = \ln\frac{\varphi_i'}{x_i} + 1 - \frac{\varphi_i'}{x_i} - \frac{z}{2}q_i\left(\ln\frac{\Phi_i}{\theta_i} + 1 - \frac{\Phi_i}{\theta_i}\right) \quad (3.110)$$

$$\text{Where } \varphi_i' = \frac{x_i r_i^{3/4}}{\sum_j x_j r_j^{3/4}} \quad \theta_i = \frac{q_i x_i}{\sum_j q_j x_j} \quad \Phi_i = \frac{r_i x_i}{\sum_j r_j x_j} \quad q_i = \sum_k v_k^{(i)} Q_k \quad r_i = \sum_k v_k^{(i)} R_k$$

The residual contribution is similar to that of the COSMO-SAC formulation, and is given by:

$$\ln\gamma_i^R = n_i \sum_{\sigma_m} p_i(\sigma_m) [\ln\Gamma_S(\sigma_m) - \ln\Gamma_i(\sigma_m)] \quad (3.111)$$

Where, n_i is the total number of segments in a molecule; $\ln\Gamma_S(\sigma_m)$ is the segment activity coefficient of segment σ_m in the mixture; and $\ln\Gamma_i(\sigma_m)$ is the segment activity coefficient of segment σ_m in pure component I , given by:

$$\ln\Gamma_S(\sigma_m) = -\ln\left\{\sum_{\sigma_n} p_s(\sigma_n) \Gamma_s(\sigma_n) \exp\left[\frac{-\Delta W(\sigma_m, \sigma_n)}{RT}\right]\right\} \quad (3.112)$$

$$\text{And} \quad \Delta W(\sigma_m, \sigma_n) = \left(\frac{\alpha'}{2}\right)(\sigma_m + \sigma_n)^2 + \frac{E^{HB}(\sigma_m, \sigma_n)}{2} \quad (3.113)$$

$$E^{HB}(\sigma_m, \sigma_n) = c_{hb} \max[0, \sigma_{acc} - \sigma_{hb}] \min[0, \sigma_{don} - \sigma_{hb}] \quad (3.114)$$

Where σ_{acc} and σ_{don} denote the larger and smaller values of σ_m and σ_n , respectively; σ_{hb} is the HB surface density cut-off; c_{hb} is a universal constant; and α' as the constant for the misfit energy is assigned a value of $8544.6 \text{ kcal } \text{\AA}^4 / (\text{mol. } e^2)$.

3.3.2.5. Universal Segment Activity Coefficient Model (UNISAC)

This new model, developed by Moodley et al. (2015), is based on a combination of concepts from two well developed methods: the original UNIFAC model and the NRTL-SAC model. It assumes that each pure substance consists of several predefined functional groups (as in the UNIFAC models); and then the “conceptual segment concept” of the NRTL-SAC model is applied to the functional groups. This results in each functional group being defined as a combination of four

basic segments that contribute to the different molecular interactions. In its formulation, the following groups were selected as the base groups to represent the four types of binary interaction base segments: Hydrophobic ($C - CH_3$), Hydrophilic (H_2O), Polar positive ($C - CN$) and Polar negative ($C - Cl$).

The UNISAC model also applies the Gibbs excess energy (GE) formulation, as the sum of two parts: a combinatorial contribution that accounts for differences in size and shape; and a residual contribution, which accounts for intermolecular forces between components:

$$\ln \gamma_i = \ln \gamma_i^{Combinatorial} + \ln \gamma_i^{Residual} \quad (3.115)$$

Where the UNISAC combinatorial contribution is similar to that of the modified UNIFAC (Do), as follows:

$$\ln \gamma_i^C = \ln \frac{\phi_i'}{x_i} + 1 - \frac{\phi_i'}{x_i} - \frac{z}{2} q_i \left(\ln \frac{\Phi_i}{\theta_i} + 1 - \frac{\Phi_i}{\theta_i} \right) \quad (3.116)$$

Where:

$$\phi_i' = \frac{x_i r_i^{3/4}}{\sum_j x_j r_j^{3/4}} \quad \theta_i = \frac{q_i x_i}{\sum_j q_j x_j} \quad \Phi_i = \frac{r_i x_i}{\sum_j r_j x_j} \quad q_i = \sum_k v_k^{(i)} Q_k \quad r_i = \sum_k v_k^{(i)} R_k$$

The residual contribution is similar to that of the COSMO-SAC formulation, and is given by:

$$\ln \gamma_i^R = \sum_k^N \Omega_{k,i} (\ln \Gamma_k - \ln \Gamma_{k,i}) \quad (3.117)$$

Where the terms, $\ln \Gamma_k$ and $\ln \Gamma_{k,i}$, are the activity coefficients of segment k in solution and in pure component i , respectively; and $\Omega_{k,i}$ is the total number of segment k in component i . These terms can be evaluated by:

$$\ln \Gamma_{k,i} = (1 - \ln(\sum_{m=1}^N \theta_{m,i} \psi_{m,k})) - \sum_{m=1}^N \frac{\theta_{m,i} \psi_{k,m}}{\sum_{n=1}^N \theta_{n,i} \psi_{n,m}} \quad (3.118)$$

$$\ln \Gamma_k = (1 - \ln(\sum_{m=1}^N \theta_m \psi_{m,k})) - \sum_{m=1}^N \frac{\theta_m \psi_{k,m}}{\sum_{n=1}^N \theta_n \psi_{n,m}} \quad (3.119)$$

Where:

$$\Omega_{k,i} = \sum_{l=1}^N v_{l,i} \zeta_{k,l} \quad \theta_{m,i} = \frac{\Omega_{m,i}}{\sum_{m=1}^N \Omega_{m,i}} \quad \theta_m = \frac{\sum_{i=1}^I \Omega_{m,i} x_i}{\sum_{m=1}^N \sum_{i=1}^I \Omega_{m,i} x_i}$$

Where $v_{l,i}$ is the number of groups of type l in a component I ; and $\zeta_{k,l}$ is the segment area of segment, k , in group l ; $\theta_{m,i}$ is the segment area fraction of segment m in the pure component i ; m and n are the segment based indices, N and I , are the total number of segments and components, respectively. While x_i is the mole fraction of component I , $\psi_{u,v}$ are the segment specific interactions, equivalent to the dimensionless group interactions between the base segments.

3.3.3 Performance of Predictive and hybrid models

In the foregoing sections, key thermodynamic aspects of describing real and non-ideal systems have been presented. In particular, the various key thermodynamics models to determine the activity coefficients have been described. The determination of the activity coefficient is essential, as it accounts for any non-ideal behaviour of the solution by accounting for the interactions that exist between the various molecules. The molecules of interest in this study are generally complex with multiple functional groups and hence models that can accurately account for the various molecular interactions is essential. The correlative models require identification of binary interaction parameters from phase equilibrium data for each of the solvent–solvent, solvent–solute, and solute–solute binary mixtures. Unfortunately, the lack of experimental solubility data, on new APIs, limits the use of these thermodynamic models for process design and analysis in the pharmaceutical industry. (Chen, 2011). Hence the increase in the need and use of predictive thermodynamic models in pharmaceutical studies. From the various models described, only the following models may be considered to be purely predictive: The UNIFAC models and its variations, COSMO and its variations, F-SAC and UNISAC. The advantage of these predictive models is that no new experimental data is needed to calculate activity coefficients. The other models like NRTL-SAC and PC-SAFT are considered to be hybrid models. They require some initial experimental data for the parametrisation of the models prior to its predictive capability.

Several studies have been performed on the measurement and prediction of the solubility of APIs, where the predictive performance of various models are compared to experimental measurements. Some of these comparative performance studies are summarised in Table 3.3. A general observation, reported in the various studies, is that the accuracy of the solubility obtained from any of the models, depends upon how much the system of interest deviates from the model assumptions, and how well the required parameters are known or can be predicted.

The hybrid models like the NRTL-SAC and PC-SAFT models tend to provide the best results of all activity coefficient models because these models uses experimental solubility data of the considered API for the initial parameterization of the model prior to their predictive capabilities. Hence, these models cannot be used in any process development and conceptual design work for systems of newly developed API's where no initial solubility data is available. For various applications, UNIFAC has been a preferred choice of model, extended either by adding new terms or by regressing new parameters, revealing the usefulness and versatility of the model. Today the Dortmund modified UNIFAC is probably the most extensive. It is widely used and is being revised and extended as part of an industrial/academic joint venture. In recent years, numerous researchers have added new groups and subgroups and have published new group parameters

(Abildskov et al., 2004). Kontogeorgis and Folas (2010) provide an extensive review of the various correlative and predictive activity coefficient models.

In the next section, the equations that are used to calculate solid-liquid phase equilibria is presented.

Table 3.3. Examples of Comparative Studies of Predictive Models for APIs Solubility.

APIs studied	Models Compared	Observations	Reference
Six different functionalized aromatic APIs	UNIFAC (Do) and COSMO-RS	The COSMO-RS method was capable of predicting the solubility of all 221 systems investigated. UNIFAC (Do) was capable of only predicting 80 % of the systems due to missing binary interaction parameters. Researchers reported comparable predictions against experimental data in literature.	Kolár et al. (2002)
Lovastatin, Simvastatin, Rofecoxib, and Etoricoxib	NRTL-SAC and COSMO-SAC	The semi-empirical NRTL-SAC model offered superior performance over the ab initio COSMO-SAC model. However limited solubility data in at least four solvents are needed for parameterization of the NRTL-SAC model prior to predictive capability.	Tung et. Al. (2007)
Paracetamol, Allopurinol, Furosemide And Budesonide	NRTL-SAC	The solubility data in pure organic solvents were used to regress the solute model parameters. The predicted solubility value varied in different solvent systems, ranging from an absolute deviation of 18% to 69%.	Mota et al. (2009)
Paracetamol, Ibuprofen, Sulfadiazine, P-Hydroxyphenylacetic Acid, And p-Aminophenylacetic Acid	PC-SAFT	The results obtained from PC-SAFT model were compared to the results from NRTL-SAC model and to experimental data. Both models give similar qualitative predictions and the performance. Although not in quantitative agreement, these results are accurate enough for solvent-screening purposes in the process development. The both models require some experimental data to regress and identify the model parameters prior to its predictive capability.	Ruether and Sadowski (2009)
Aspirin , Paracetamol and Ibuprofen	UNIFAC; Mod UNIFAC (Do), and COSMO-RS (OI)	The UNIFAC (Do) provides the lowest root mean square deviations for the temperature and the solubilities. The second best is achieved by the UNIFAC model, followed by COSMO-RS (OI). The UNIFAC (Do) is able to predict the solvent which shows the highest solubility for the two active pharmaceutical ingredients (aspirin, ibuprofen) investigated.	Hahnenkamp et al. (2010)
Paracetamol, Naproxen, Ibuprofen, Flurbiprofen, Ketoprofen, And Lovastatin	PC-SAFT	PC-SAFT predictions compare favourably with the experimental data of the various systems evaluated. Its predictive ability of the scheme was based on the appropriate parameterization of the pharmaceutical molecules. The regression of parameters was performed against the solubility of pharmaceuticals in three solvents, i.e., a hydrophilic solvent (water), a polar solvent, and a hydrophobic solvent.	Spyriouni et al.(2010)
Ibuprofen, Acetaminophen, Benzoic Acid, Salicylic Acid And 4-Aminobenzoic Acid, and Anthracene	UNIFAC; Mod UNIFAC (Do); COSMO-SAC and NRTL-SAC	It was found that UNIFAC models give the best order of magnitude results and could be useful method for rapid solubility estimations of an API in various solvents. COSMO-SAC needs more developments to increase its accuracy especially when hydrogen bonding is involved. NRTL-SAC model are also in good agreement with the experiments, but in that case the relevance of the results is strongly dependent on the required model parameters to be regressed from solubility data in single and mixed solvents prior to predictive application	Bouillot et al. (2011)
33 different APIs	COSMO-SAC	The COSMO-SAC model provides semi-quantitative accuracy for the activity coefficient of drug in solution. The predictions of solubility in mixture solvent is improved with the inclusion of an empirical Margules-type activity coefficient correction term to the result from the COSMO-SAC model. However this requires the determination the drug-solvent interaction parameters which are obtained using the experimental solubility data of the drug in pure solvent. This reduces the predictive nature of the model.	Shu and Lin (2011)

Table 3.3. Examples of Comparative Studies (Continued).

APIs studied	Models Compared	Observations	Reference
Between 12 and 70 different APIs with various models	Pharma Mod. UNIFAC; Hansen Solubility Model; UNIFAC; Mod. UNIFAC (Do); and NRTL-SAC	The correlative model NRTL-SAC provides the best results of all activity coefficient models because this model uses experimental solubility data of the considered API for parameterization of the model. The Pharma-Mod model more accurately predicted complex compounds solubility data compared to the other UNIFAC models in most cases. However the parameter matrix requires further development to broaden its application.	Diedrichs and Gmehling (2011)
3-Pentadecylphenol, Lovastatin, And Valsartan	NRTL-SAC and UNIFAC	The NRTL-SAC model showed relative advantage over the UNIFAC model in almost all cases, except for the systems containing light alcohols with water. However NRTL-SAC required some data for regression of parameters prior to use.	Sheikholeslamza deh and Rohani (2012)
18 APIs including Aspirin, Ibuprofen And Testosterone	UNISAC, UNIFAC and NRTL-SAC	The UNISAC model provides a superior prediction than the NRTL-SAC and UNIFAC models in over two-thirds of the systems investigated. The FIC and AIC scores recommended the UNISAC model in over 75% of the cases when compared to NRTL-SAC and UNIFAC for most systems tested. The UNISAC model is found to be competitive with Pharma-UNIFAC and modified UNIFAC (Do) models in the three solvent systems tested: alkanes, alcohols and water. The UNISAC model provides a means of performing qualitative predictions of solubility for complex pharmaceutical components.	Moodley et al. 2015
Acetylsalicylic Acid, Acetaminophen, Cimetidine and Famotidine	Pharma Mod. UNIFAC and Mod UNIFAC (Do)	The prediction results show that the Pharma Modified UNIFAC model gives a qualitative representation of the experimental solubilities, and give better predictions than the modified UNIFAC (Dortmund) model for the systems studied.	Matsuda et al. (2015)
Picric acid, Salicylic acid , 3-Nitrobenzoic acid, Biphenyl	Pharma-Mod UNIFAC and KT-UNIFAC	The Pharma-Mod model more accurately predicted complex compounds solubility data compared to the other UNIFAC models. The model parameter matrix needs further completion and extension.	Nouar et al. (2016)

3.4. Development of multiphase flash calculation equations

The generalised formulation of multiphase flash calculations that are used to calculate the various phase diagrams, associated with crystallisation, is presented. The equation-solving approach entails performing a stability test to know the number of phases, and in solving a set of non-linear algebraic equations deduced from mass balances and equilibrium relations. This method involves two sequential loops: first the determination of the number and type of coexisting phases, using a stability test like the tangent plane criterion; and second, the calculation of composition and ratio of the phases by solving the correct set of governing equations (Michelsen, 1982a, 1982b), (Michelsen and Mollerup, 2007).

If a system of n components is split into p phases, then the material balance for component I is given by:

$$z_i = \sum_{k=1}^{k=p} \theta_k x_i^k \quad (3.120)$$

where z_i is the mole fraction of component i in the feed; θ_k is the phase fraction of phase k ;

and x_i^k is the mole fraction of component i in phase k . If we choose one of the phases as the reference phase (such as the liquid phase), then the equilibrium constants can be defined as:

$$K_i^k = \frac{x_i^k}{x_i^l} \quad \text{or} \quad x_i^k = K_i^k x_i^l \quad (3.121)$$

And, therefore:

$$z_i = \sum_{k=1}^{k=p} \theta_k K_i^k x_i^l \quad (3.122)$$

$$x_i^l = \frac{z_i}{\sum_{k=1}^{k=p} \theta_k K_i^k} = \frac{z_i}{\theta_l K_i^l + \sum_{k=2}^{k=p} \theta_k K_i^k} = \frac{z_i}{\theta_l + \sum_{k=2}^{k=p} \theta_k K_i^k} \quad (3.123)$$

Since the sum of all mole fractions of each phase must be equal to 1, for each phase we have:

$$\sum_{i=1}^{i=n} x_i^k = 1 \quad \text{therefore we can define } (p-1) \text{ functions } f_k = \sum_{i=1}^{i=N} x_i^k - \sum_{i=1}^{i=N} x_i^l = 0$$

Therefore we have:

$$\sum_{i=1}^{i=N} K_i^k x_i^l - \sum_{i=1}^{i=N} x_i^l = x_i^l (\sum_{i=1}^{i=N} K_i^k - 1) = 0 \quad (3.124)$$

Therefore, from equation (3.124) we have:

$$\frac{\sum_{i=1}^{i=N} (K_i^k - 1) z_i}{\theta_l + \sum_{k=2}^{k=p} K_i^k \theta_k} = 0 \quad (3.125)$$

And the overall material balance corresponds to:

$$\sum_{k=1}^{k=p} \theta_k = 1 \quad \text{which can be written as } \theta_1 + \sum_{k=2}^{k=p} \theta_k = 1$$

Therefore equation (3.124) can be written as (p-1) equations:

$$\frac{\sum_{i=1}^{i=N} (K_i^k - 1) z_i}{1 + \sum_{k=2}^{k=p} (K_i^k - 1) \theta_k} = 0 \quad (3.126)$$

To satisfy positive phase concentrations for all phases, mass conservation constrains that the denominator term be greater than zero):

$$1 + \sum_{k=2}^{k=p} (K_i^k - 1) \theta_k > 0 \quad i = 1, 2, \dots, n \quad (3.127)$$

The incorporation of the following azeotropy and miscibility test, into the solvent selection tool, ensures correct prediction of the phases that exist during the multiphase flash calculations:

An azeotrope exist if:

$$\frac{\gamma_1^\infty P_1^{sat}}{P_2^{sat}} > 1 \text{ and } \frac{P_1^{sat}}{\gamma_2^\infty P_2^{sat}} < 1 \quad \text{OR} \quad \text{if } \frac{\gamma_1^\infty P_1^{sat}}{P_2^{sat}} < 1 \text{ and } \frac{P_1^{sat}}{\gamma_2^\infty P_2^{sat}} > 1 \quad (3.128)$$

Then at some composition x , $\alpha_{12} = 1$

Where γ_1^∞ represent the activity coefficient of component 1 at infinite dilution in component 2, and γ_1^∞ represent the activity coefficient of component 2 at infinite dilution in component 1.

The system exhibits complete miscibility if the following inequality is satisfied across the entire composition range:

$$\left[\frac{d^2 \Delta G}{d(x_i^l)^2} \right]_{T,P} > 0 \quad \text{OR} \quad \left[\frac{d \ln \gamma_i}{d x_i^l} \right]_{T,P} > -\frac{1}{x_i^l} \quad (3.129)$$

Since the Solvent Selection Tool eliminates the possibilities of immiscible systems, we will have at least three phases (vapour, liquid, and solid) with the possibility of multiple solid phases. From the stability criterion presented by Michelsen (1998b), component i may exist as a pure solid if:

$$f_i(P, T, z) - f_{pure\ i}^s(P, T) \geq 0 \quad (3.130)$$

Where $f_i(P, T, z)$ is the fugacity of component i with feed composition z . This stability analysis gives the number and identities of the crystallizing components.

For a vapour-liquid, multi-solid systems with N_s number of solid phases determined by the above stability criterion, we have the following unknowns: $(n-1)$ vapour phase compositions; $(n-1)$ liquid phase compositions; vapour phase fraction; and N_s solid phase fractions. With the set of $(p - 1)$ equations obtained from equation (3,125), and with the phase equilibria equations listed below, we are able to determine the phase fractions and compositions for equilibrium flash at a specific temperature and pressure:

N vapour-liquid isofugacity equations:

$$f_i^v(P, T, x_1^v, x_2^v, \dots, x_{n-1}^v) - f_i^l(P, T, x_1^l, x_2^l, \dots, x_{n-1}^l) = 0 \quad (i = 1, \dots, N) \quad (3.131)$$

And N_s liquid-solid isofugacity equations:

$$f_i^l(P, T, x_1^l, x_2^l, \dots, x_{n-1}^l) - f_{pure\ i}^s(P, T) = 0$$

$$(i = (N - N_s) + 1, (N - N_s) + 2, \dots, N) \quad (3.132)$$

As there are no analytical techniques for solving such sets of nonlinear coupled equations, iterative methods are used. There are several successful and proven iterative methods in phase calculation. The most established is the Successive Substitution Method (SSM) (Michelsen,

1982a, 1982b). SSM requires an initial guess for the K-factors, to calculate the molar phase fractions. At the outset, composition-independent correlations can be used to calculate the K-factors (Michelsen and Mollerup, 2007). This is followed by a calculation of phase compositions, fugacity coefficients, component fugacities, and activity coefficients in each fluid phase. Pure solid-phase fugacities can be calculated outside the iteration, as the solid phases are assumed to be pure and their fugacities, therefore, only depend on temperature and pressure. The phase equilibrium conditions are checked. If equilibrium conditions are not satisfied, the K-factors (using composition dependent correlations) are updated, after which the iteration procedure is repeated. The procedure known as the Boston-Britt algorithm, which is used by most commercial process simulators, solves the flash equations in an inner loop using simple models, and the simple model parameters are updated in an outer loop by calculating properties from rigorous models (Parekh and Mathias, 1998). The typical iteration scheme of the SSM method is shown in Figure 3.3.

Concluding remarks

In this work, solid-liquid phase equilibria is used as the theoretical foundation for the understanding crystallisation processes, and in this chapter, a brief outline has been provided on the importance of solid-liquid phase equilibrium processes in the selection of solvents for crystallisation. The thermodynamic foundations relating to the phenomena of phase equilibria, and the classical formulation of solid-liquid equilibria is presented. An overview of the thermodynamic models, including the equations of state, and the correlative and predictive activity coefficient models that account for the non-ideal behaviour in the computation of phase equilibria is presented. In addition, the calculation procedures for determining the multicomponent multi-phase equilibria using flash calculations is presented as these are used to establish the presence of solid-liquid equilibria and the subsequent changes to the phase equilibria under varying conditions of temperature and concentration.

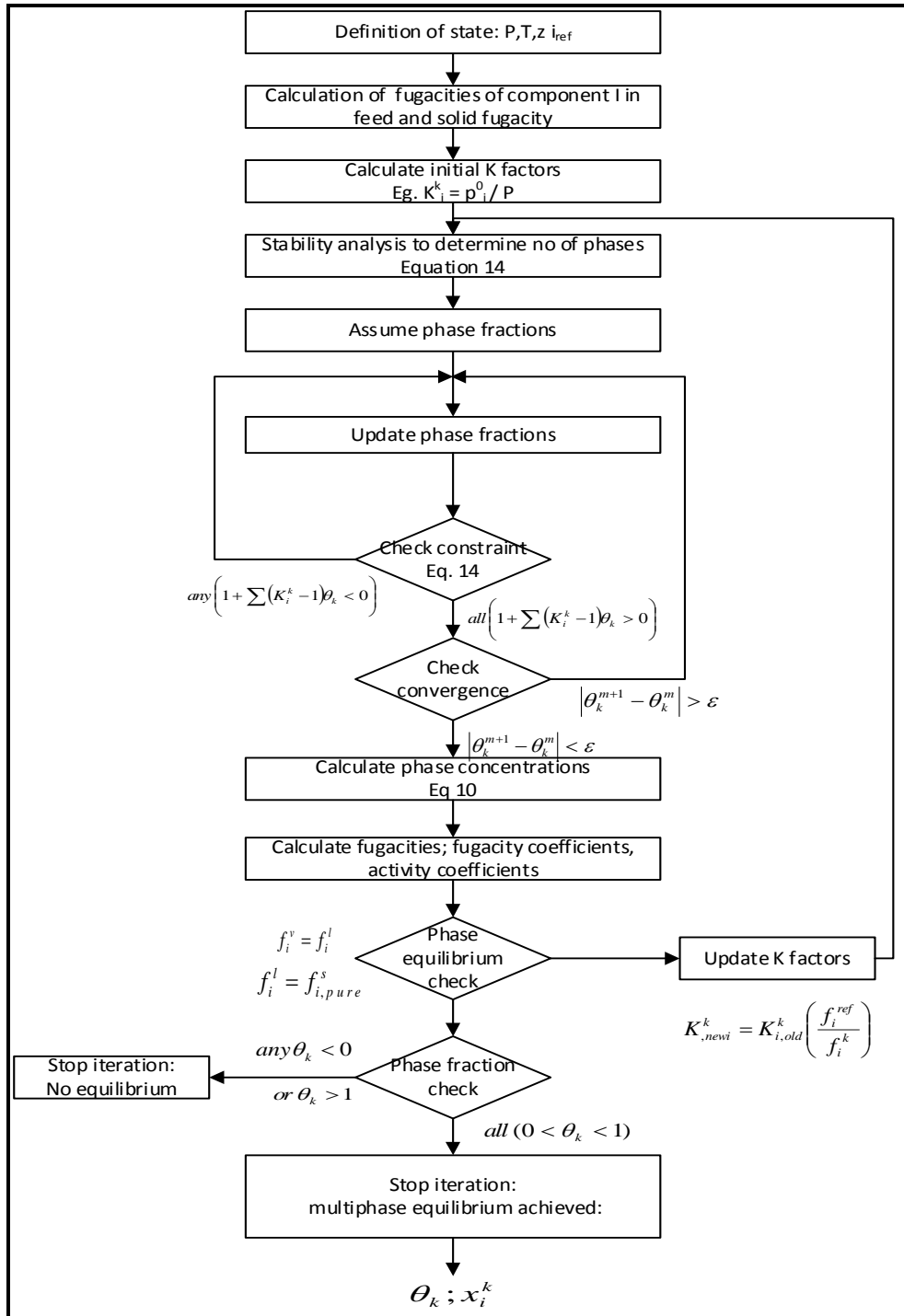


Figure 3.3: Typical iteration scheme of the Successive Substitution Method (SSM).

CHAPTER 4: COMPUTATIONAL METHODS

4.1. Introduction

The pharmaceutical industry is a highly competitive, regulated industry. The process of developing a new drug compound is not only lengthy and costly, it is also risky and complex. The average cost of developing a new drug is currently estimated to be in the region of \$800 million when capitalized, and the average development time, from patent (typically valid for twenty-two years) application to market entry, is seven to twelve years, leaving just ten to fifteen years remaining for sales, and good profitability (Crafts, 2007).

Considering the industry constraints, in designing a process for API production it is significant that there are several, necessary yet competing criteria to fulfil:

- The process must be able to consistently produce the required yield, purity and quality;
- The process should be uncomplicated, robust, and to be able to perform consistently;
- The required capital expenditure and operating costs should be minimal, and;
- The process should have minimal environmental impact.

As early as possible during the conceptual design phase it is critical to establish the optimal processing route and operating conditions that will result in maximum financial benefit with minimum environmental impact.

A key process parameter, in terms of impact, is the choice of solvent. Different solvent systems induce varying yields, which invariably impact directly on processing volumes, equipment size, and downstream processing requirements for solvent recovery and waste stream treatment (Buxton et al., 1999; Constable et al., 2007, Ruether and Sadowski, 2009). Ideally, solvents should be selected for their minimal operational and capital costs, while providing good yield of the API, good recovery and reuse, and staying within regulatory constraints.

Hence, in this chapter the development of a computational means to select the best solvent will be explored, within the overall process, in terms of operational, economic, and environmental performance requirements.

4.2 Development of a Computational Framework

4.2.1. Software selection

This section briefly describes the type of knowledge and some of the tools needed to develop a reliable simulation module. Some of the key requirements for the development of a computational tool to assess the overall performance of a solvent include:

1. A database of chemicals associated with the pharmaceutical industry including the various solvents, products, by-products, and impurities for the system under investigation.

2. Thermodynamic models that can predict the physical and chemical properties of all the components required in the various computations, including multiphase equilibria calculations.
3. Process models and algorithms that can perform the required solvent evaluations and process synthesis calculations for crystallisation.
4. An automated means to perform multiple calculations.
5. A means to record the results from multiple calculations as structured data sets for analysis.

The first two criteria already exist in many commercial software simulators, such as CHEMCAD, Aspen Plus, Hysys, Pro II, etc. These allow for user defined compounds as well as values of experimentally measured properties, such as binary interaction parameters, to be added into the component database. Chemical process simulators have become reliable tools that are widely used in process engineering. Process simulators contain strict models for most chemical process units, robust numerical algorithms for particular units, and large databases of physicochemical thermodynamic and transport properties. The capability of process simulators can make modelling and optimization of the process easier, while complex calculations can be carried out quickly by a process simulator. Nevertheless, resetting the input parameters to evaluate a range of scenarios can be time consuming, and storage of the multiple simulation results for comparative studies can be challenging.

The developers of commercial process simulation software have created various interfacing options to extend the capabilities and applications of the software. An interface serves as a technical and semantic bridge that allows reading from, and writing to the objects of another software package. But, depending on the purpose of the system, objects can be found on either side of the interface and the objects of a process simulator will differ in meaning to those of adjacent programs. Hence, an important requirement for the design of an interface lies in the standardization of meaning.

A technical bridge from the interface can be implemented by means of a file or data transfer, function calls or shared objects. The key benefit of interfacing is that it allows for the objects in the process simulator to be used beyond its boundaries. For example, CHEMCAD developers have created various interfacing options: between Microsoft Excel and CHEMCAD; between The Design Institute for Physical Properties (DIPPR) databank and CHEMCAD; and between the Dortmund Data Bank (DDBST) and CHEMCAD (Fricke and Schöneberger, 2015). In addition, CHEMCAD also allows a user to develop their own process models that can be integrated into any simulation.

The CHEMCAD user-defined process models are developed with VBA (Visual Basic for Application) programming language and are hence named VBA Unit Operations. These interfacing facilities now extend computational capabilities by allowing the resetting of input parameters, thus managing the operations of the process simulator. And, as the storage capacity of MS Excel is practically unlimited, it is an ideal, interfacing, complementary tool for the process simulator.

As a result, it was decided to exploit the interfacing capability of CHEMCAD, and to use it to construct a solvent selection tool. This was accomplished by developing an internal, interfacing VBA module embedded within the process simulator. The details of the VBA module is described in the following sections.

4.2.2. Computational Framework

A novel, modular, integrated computational framework is developed to simulate crystallisation process synthesis and optimization. The modular integrated framework integrates CHEMCAD, a VBA Module (with various crystallisation sub-routines) and Excel. The key criterion for the modular integrated framework is the accurate exchange of data between the VBA module algorithms, the CHEMCAD simulation algorithms and Excel. The specified and calculated variables extracted from the CHEMCAD simulation platform are dictated by, and used in the VBA platform, in this study. In the framework, the various calculations for the crystallisation processes are based on the CHEMCAD multiphase flash calculation simulation model. The extraction and processing of necessary data is then implemented by the VBA module. The overall framework, and the interactions of the VBA module with the various components of CHEMCAD and Excel, is shown in Figure 4.1.

This approach, of developing and integrating this VBA module within the commercial simulation programme CHEMCAD, has several advantages. The computations within the VBA modules have access to a full range of thermodynamic models and correlations, a comprehensive database of compounds and their pure and mixture properties, and rigorous, computational algorithms for process calculations and equipment design. In addition, software vendors regularly deliver updates for their programs to fix bugs and to deliver new functionality, including the updating and inclusion of new properties, thermodynamic models and methods. CHEMCAD also provides for user defined components to be added into its database. The VBA modules are integrated into CHEMCAD via a VBA unit operation; and an intuitive and user-friendly dialogue screen, built into a visual basic platform, allows users to send information into the VBA unit operation.

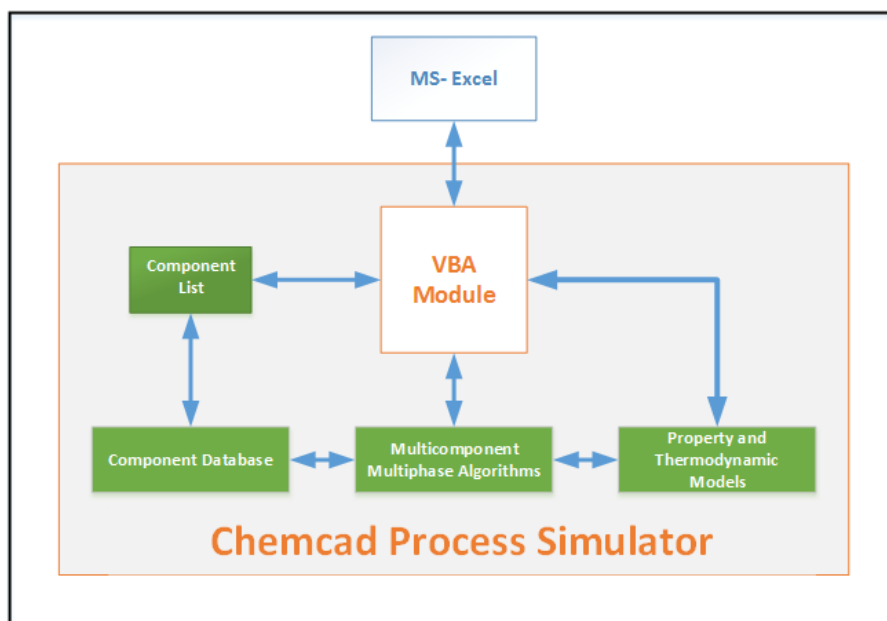


Figure 4.1. Components of the integrated computational framework.

The mutual connection between CHEMCAD and MS Excel is created by a VBA unit operation. A master/slave architecture is adopted, where a series of calculations, as dictated by the VBA module, are carried out systematically with the help of this software connection. The input, feed-stream parameters for every simulation are always set in MS Excel and transferred into CHEMCAD via the VBA Module. When the required computation is completed by the process simulator, the internal variables are further processed and updated within the VBA module sub-routines for the next calculation loop. The next calculation loop may be a change in the operation conditions, or a new component to be evaluated. The desired set of results, either for each calculation loop or at the end of a complete simulation, are transferred back to MS Excel. At the end of the calculations, the initial parameters and the calculation results are displayed in a structured form in the MS Excel file. Thus, the operational database of the crystallisation process is created. The resulting database enables us to define/analyse the optimum and highlight the unfavourable operating range of the system. This typical structure of the various sub-routines is illustrated in Figure 4.2.

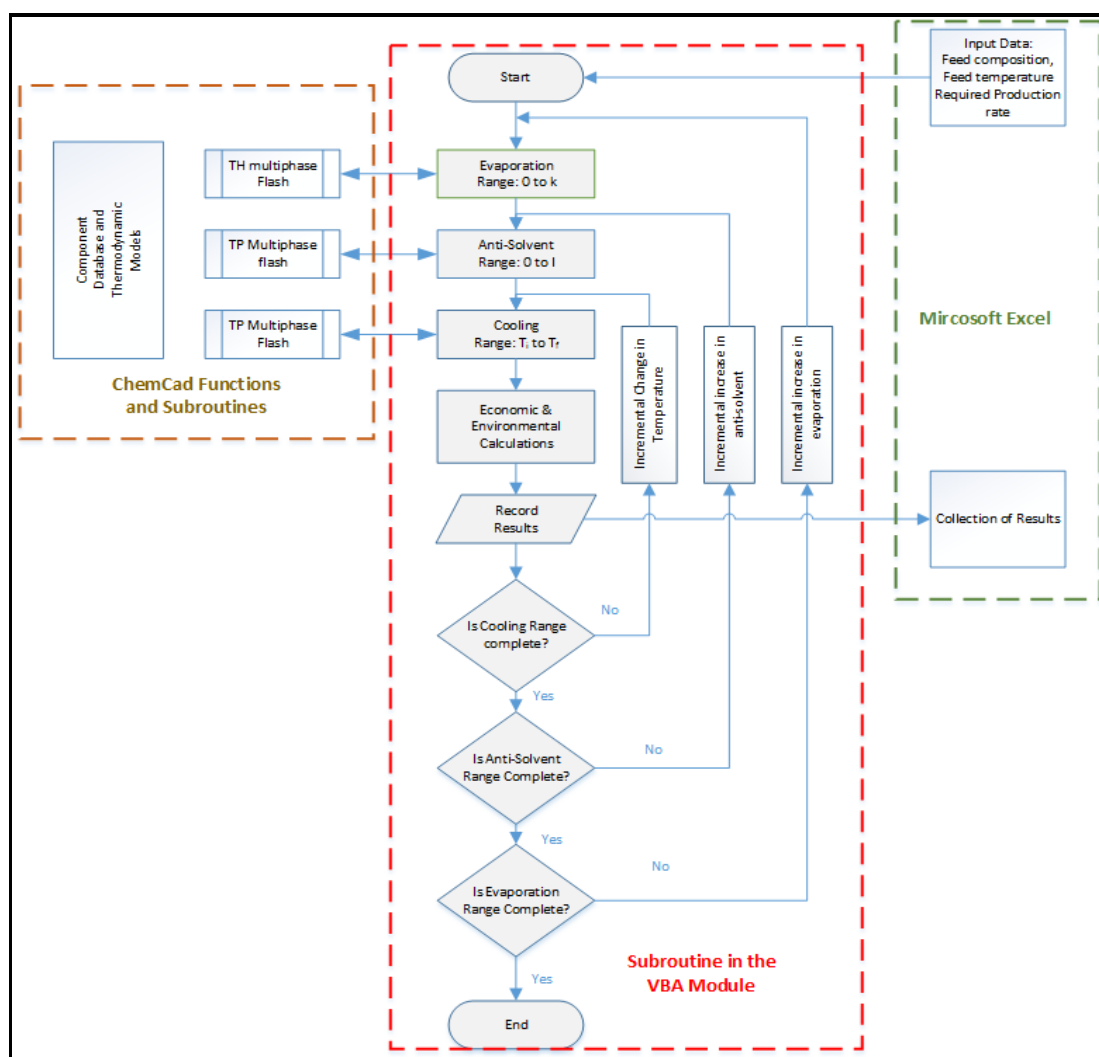


Figure 4.2. An Example of a Typical Sub-Routine and Its Communication Pathways.

To illustrate this interface and the two way communication between Excel – VBA Module and CHEMCAD, some examples of code statements and their functions in the various sub-routines of the VBA module) are presented in Table E.1. In addition, an example of the many sub-routines that make up the computational framework is presented in Appendix E.

4.4.3. The VBA Module

The heart of the computational solvent selection tool is the VBA module, which not only provides the necessary instructions to CHEMCAD, but has over 30 sub-routines that compute various calculations to fulfil its purpose as a complete conceptual design tool. The VBA module meets the following desired software development requirements: its models are appropriate for purpose, in terms of rigor, level of detail, accuracy, validity and generality; the algorithms are robust, generalizable, and efficient in terms of execution and storage; the software is easy to understand, maintain and modify, and it is transportable; and the user interface readily accepts input, and presents results in a useable form.

The Solvent Selection Tool, developed in this work, is capable of computing various outputs associated with crystallisation. Outputs include the prediction of the eutectic temperatures and compositions in the system. It can create various types of phase diagrams and solubility curves that allow the user to examine the overall composition space, to visualize crystallisation regions, to identify separation barriers, and to evaluate the various operations. The operations, such as heating, cooling, solvent addition, solvent removal, anti-solvent addition, and their combinations, can thereby be studied in order to systematically evaluate process alternatives.

The pivotal calculation tool is the multicomponent, multiphase, flash-calculation algorithm, which is presented in chapter 3. The operational protocols of the crystallisation process, to achieve the desired targets, are determined by conditions that lead to the formation of a solid phase, the extent to which a solid phase will continue, under varying temperatures and liquid phase compositions, to crystallise prior to the next solid phase forming. The methodology that is used to establish these protocols is presented in Figure 4. 3.

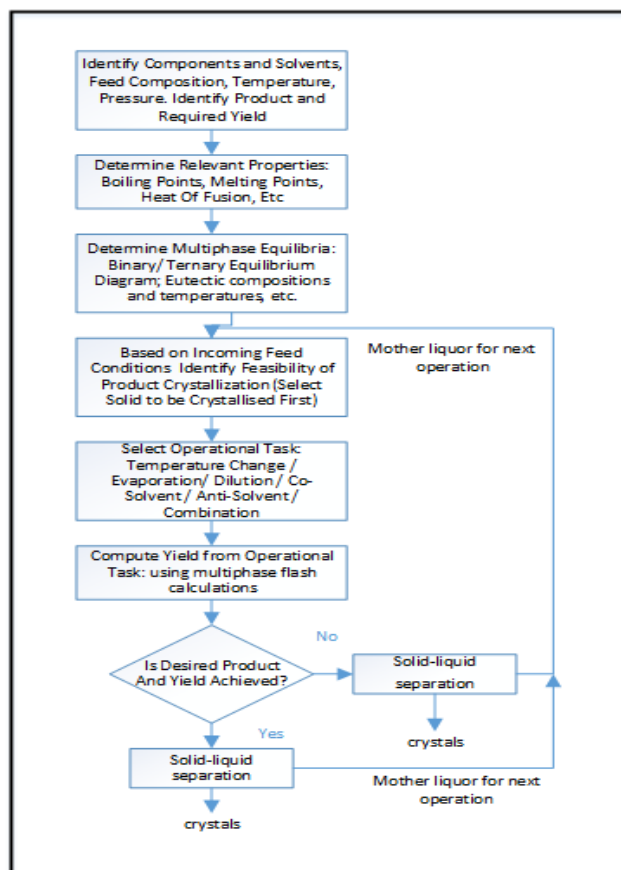


Figure 4. 3. Algorithm for the analysis of crystallisation to determine operating protocols.

The full computational capabilities of the VBA module are summarised in Figure 4.4. The VBA module consists of 32 sub-routines (150 pages of VBA code). Each of the calculations listed in Figure 4.4 may involve multiple sub-routines with multiple iterative CHEMCAD calculations.

Whilst some of the computational outputs present as a single answer after completing several iterations, e.g., run cool down will determine the yield at a specified temperature, others will output a matrix of results for a range of input variables for further analysis. A detailed description of the various menus and calculation capabilities of the developed computational tool is presented in Appendix B.

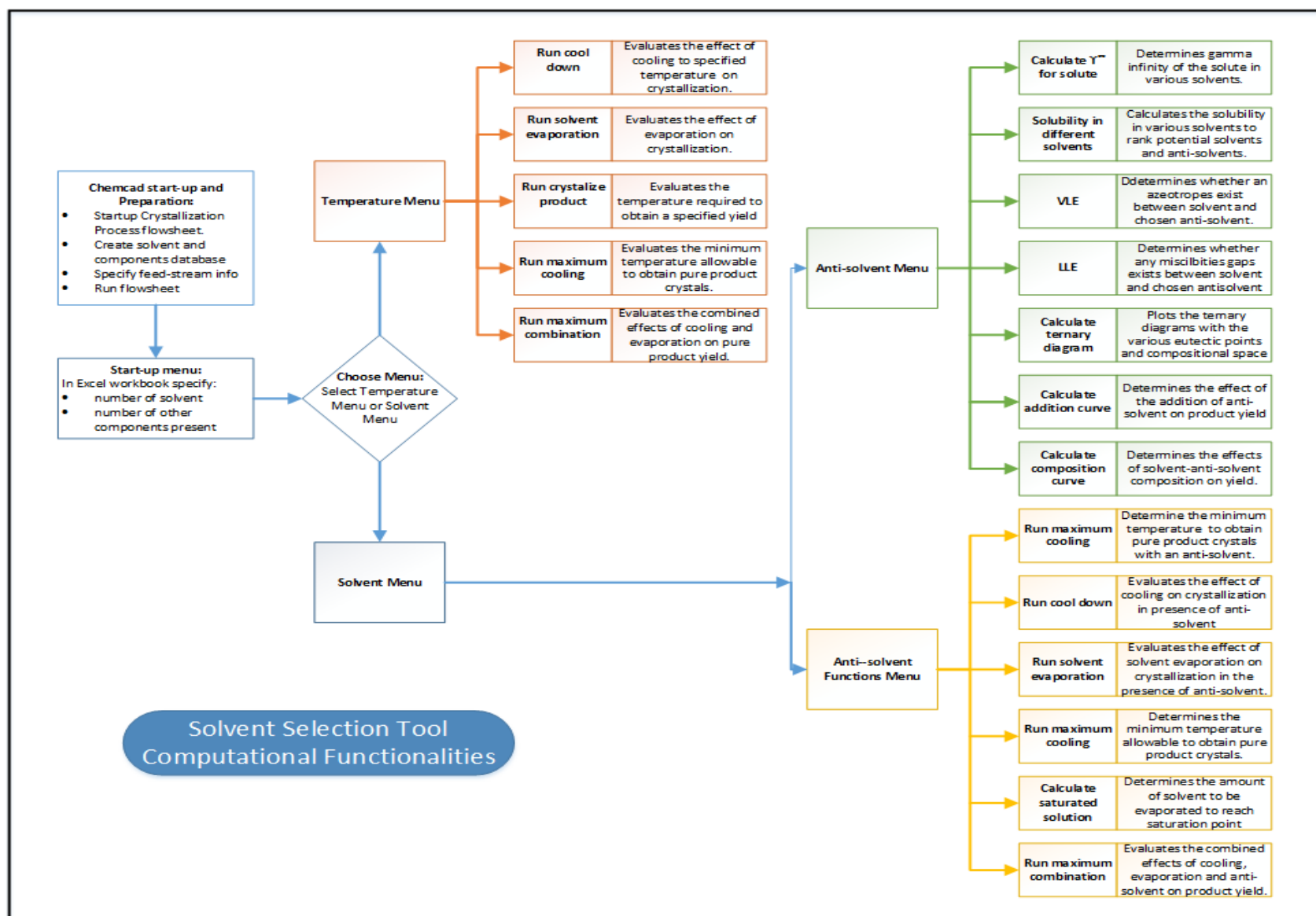


Figure 4.4. Computational Capabilities of the VBA Module developed.

4.3. Models and their Assumptions

4.3.1. Practices in the pharmaceutical industry

To develop a realistic computational tool for the pharmaceutical industry, it is essential to understand its manufacturing principles and procedures. General processing practices in the pharmaceutical industry, noted by Crafts (2007), Wieckhusen, (2006, 2011), and am Ende, (2011) include:

- Pharmaceutical manufacturers of non-generics almost exclusively use batch technologies. This minimizes the impact of off-specification products to discrete quantities that can be easily isolated for rework or disposal.
- Material recycling is generally avoided because it can introduce problems of quality control, particularly in respect to impurity of concentrations. Impurities, for the purposes of crystallisation, could be unreacted starting materials or byproducts. With recycling, the concentration of these impurities may increase to solubility, within the operating temperature ranges, and may crystallize simultaneously with the desired API, hence leading to contamination.
- Processes are designed to fit generic arrangements where feasible. Because of the small production volumes of many APIs it is typical for production to be done in multi-product batch plants that are furnished with generic equipment.
- The process of crystallisation is generally used to selectively separate out the API as a solid of very high purities. The key unit operation is the recovery of the desired API from the process stream exiting the reactor. The three main advantages of using crystallisation as a separation technique are: a high purity product in one process step; a comparatively low level of energy consumption; and relatively mild process conditions.

With these factors in mind, there are several competing measures of performance that can be used for ranking process schemes during the conceptual design phase. The general methodology selected to develop the selection tool is structured as follows:

1. Identify solvents that are acceptable for use in the pharmaceutical industry. Several Solvent selection guides have been developed by AstraZeneca, Pfizer, GlaxoSmithKline (GSK), and Sanofi to facilitate solvent selection (Prat, 2014), (Byrne et al., 2016). In this work, the GlaxoSmithKline (GSK) guide (Henderson, 2011) is used.
2. Evaluate the list of solvents to determine a list of candidate solvents/solvent blends that meet the *process specific performance criteria*: Crystal yields, physical property requirements (boiling point, flashpoint, miscibility etc.)
3. From (2), determine the optimal solvents/solvent blends/anti-solvents that meet *economic performance* expectations by considering the following:

- a. Operational costs (solvent usage and solvent recovery and reuse potential, energy usage, waste treatment, etc.). The methodology proposed by Ulrich and Vasudevan (2006)
 - b. Capital costs, which accounts for the processing equipment size and downstream processing complexity. The methodology presented in Turton (2008) is used.
4. Evaluation of the *environmental performance* of the selected solvent candidates using performance indices. Methodologies outlined in the field of Green Chemistry have been adopted (Sheldon, 2007) (Dunn, 2010) (Pistikopoulos & Stefams, 1998)

This stepwise procedure reduces the combinatorial complexity and avoids unnecessary computations. In addition, to reduce the computational complexity, the following assumptions are proposed:

- a) Crystallisation is the preferred unit operation for the recovery of the desired API. It will be assumed that the desired API with the desired properties (morphology and size distribution) are achieved in a well-controlled crystallisation process over a period of 20 hours. In addition, the discharge, cleaning and recharge cycle is 4 hours. Hence, the total batch time is 24 hours or 1 day. The batch size is limited to 1.25 cubic meter per crystallizer. The total volume of each crystallizer is 1.5 cubic meters (1.25 cubic meter working volume and 0.25 cubic meter of headspace.)
- b) The amount of solvent/s required for the process is only dependent on the equilibrium solubility relationship
- c) The size of storage (feed and waste streams) and processing vessels is dependent on the solvent requirements per batch or production output. It is assumed that the required solvent storage capacity is twice the daily solvent requirement plus a headspace (safety factor) of 25%.
- d) The plant operations that are considered to be directly affected by the choice of solvent are limited to: the reactor; the crystallizers and their required heat exchangers; solvent feed and waste storage tanks; and the solvent recovery system. Hence the capital costs are determined for: reactor; crystallizer; heat exchangers; solvent recovery evaporator; and solvent feed and waste storage tanks. The operational costs are determined for: crystallizer cooling / evaporation; solvent recovery heating; solvent cost and waste treatment. The operations not included in the financial models include: the centrifuge; washing station; any recrystallisation operations and drying.
- e) The waste treatment is only dependent on the volume of the unrecoverable solvent and is independent of the type of solvent and API. It will be assumed that there will be 80% recovery of solvent, hence 20 % of the daily solvent requirement plus the unrecovered API will require tertiary waste treatment or hazardous waste treatment methods
- f) The amount of desired API (kg) crystals will form the basis for all simulations. A production rate of 1000 kg of API per day will be the basis for all calculations.

- g) Evaporation is the preferred unit operation for solvent recovery of single solvent systems, and distillation may be required with co-solvent and anti-solvent applications.
- h) Steam generation in the pharmaceutical industry is typically a two stage process: steam that is generated in a standard boiler is used as a heat source to convert ultra-pure water (generated in a multi-effect evaporator) into steam which is then used for process heating. This dual loop system effectively increases the cost of heating.

4.3.2 Solvent Database

Solvents typically make up more than 80% of the material usage for active pharmaceutical ingredient (API) manufacture (Constable et al., 2007). Solvent use also consumes about 60% of the overall energy and accounts for 50% of the post-treatment greenhouse gas emissions (Jimenze-Gonzalez et al., 2005). Hence, solvent selection is a major consideration in the design of chemical and pharmaceutical processes.

Several initiatives within the pharmaceutical industry, to improve process efficiency and product quality through green chemistry and engineering, have resulted in the development of a valuable set of tools, including a solvent selection guide, Process Mass Intensity/LCA calculator, and a powerful reagent guide.

With the use of these solvent selection guides, appropriate solvents can be identified prior to screening experiments and less desirable solvents can be replaced in new and established processes. These solvent selection guides take into consideration the chemical functionality, physical properties, regulatory concerns, and safety/health/environment (SHE) impact. In this work, the GSK Solvent Guide is used to develop the database of solvents for this framework. The resulting solvent database used in this computational tool is presented in Appendix C.

4.3.3. Performance criteria and the Relevant Equations

4.3.3.1. Process Specific Performance Criteria: Yield, purity and quality

Yield, in the API crystallisation process, is an important target because it largely determines the profitability of the process. Both solvent and process should be chosen so that a high yield with the desired purity can be obtained. But, conditions leading to high yields may also favour crystallisation of by-products and undesirable polymorphs, which may result in contamination of the desired pure API crystals. Delivering the right polymorphic form or solvate is crucial because these forms usually differ in their solubility and dissolution rate and may even differ in physiological effects.

Particle size distribution (PSD) of the API is a key quality attribute. It affects the manufacturing process in terms of flow characteristics, the filtration process, isolation, and drying kinetics; interactions with excipients, and the proper delivery profiles of the drug (bioavailability). Ensuring consistent PSD with

each batch requires proper control of primary nucleation and crystal growth. The equation that is used to calculate the yield is derived from a material balance around a crystallizer (Jones, 2002; Mersmann, 2001; Myerson, 2001), and the resulting equation that is used to calculate the yield is:

$$f_1(x) = Y = \frac{WR[C_i - C_f(1-V)]}{1 - C_f(R-1)} \quad 4.1$$

Where:

C_i and C_f are the initial and final equilibrium solute content in the crystalliser (kg solute/kg solvent);

W is the initial mass of solvent (kg);

V is the solvent loss through evaporation (kg vapour/kg original mass of solvent) ;

R is the ratio of the molar mass of the hydrate and anhydrate solute;

Y is the yield of crystals (kg).

Note:

C_i and C_f are determined by the solubility equation: $\ln x_i^{sat} - \frac{\Delta H_{fus}}{T_m} \left(1 - \frac{T_m}{T}\right) + \ln \gamma_i^{sat} = 0$

It should be noted that the basis for all calculations is the desired production rate of API. Hence, from the feed composition and the initial and final operating conditions, we are able to calculate the required feed rate of all the feed components to meet the desired API production rate. From the completed material balance we are able to determine the sizes of the various equipment and utilities required.

4.3.3.2. Economic performance

In all stages of the design process, economic evaluation is crucial in the selection of process alternatives. Various methods are available in chemical engineering literature (Turton, 2008), (Peters, 1991), (Green, 2007), etc. for the economic evaluation of chemical processes. Some incorporate the concept of the ‘time value of money’, such as the net present value (NPV), and discounted cash flow methods. While these are measures of profitability over an extended time period, they require certain assumptions relating to interest rates and inflation. Alternatively, the total annualized cost (TAC) can also be used as an economic indicator/objective function for the evaluation of design alternatives and economic optimization. The economic method used in this work, which is developed in a VBA module, carries out standard cost calculations for fixed capital investment and operational cost, and computes total annualized cost.

The plant operations considered to be directly affected by the choice of solvent are limited to: the reactors; the crystallizers and their required heat exchangers; solvent feed and waste storage tanks; and the solvent recovery system. Hence, the capital costs are determined for: reactors; crystallizers; heat exchangers; solvent recovery evaporators; and solvent feed and waste storage tanks. The operational

costs are determined for: crystallizer cooling/heating requirements; solvent recovery heating requirements; solvent cost; and waste treatment. The operations not included in the financial models include: centrifuge; washing station; any recrystallization operations; and dryer.

The financial impact of a selected solvent to other available solvents is determined by comparing the TAC associated with each solvent, where the TAC is calculated by the expression:

$$\text{Total Annualised Cost (TAC)} = D \times \text{Fixed Capital Costs} + \text{Annual Operational Costs} \quad 4.2$$

Where D is the depreciation or capital recovery factor and is normally taken to be a value between 0.15 and 0.25, but can also be computed using depreciation calculation methods.

We can further rank the solvents using the concept of normalised total annualised cost (NTAC) which we define as:

$$\text{Normalised Total Annualised Cost (NTAC)} = \frac{(TAC)_{\text{solvent } i}}{(TAC)_{\text{lowest}}} \quad 4.3$$

The fixed capital costs include direct costs (equipment, installation, electrical, piping, instrumentation, etc), and indirect costs (contractor fees, construction expenses, contingencies, etc). It is estimated by multiplying the equipment cost by estimation factors. Peters and Timmerhaus (2003) present an estimation factor of 4.28 for solid-liquid processing plants. Hence the fixed capital cost can be estimated by:

$$\text{Fixed Capital Costs} = 4.28 \sum C_{eq} \quad 4.4$$

4.3.3.2.1. Equipment Costs

The equations and figures presented by Turton et al. (2009) are used to determine the capital equipment costs. These are based on a module factor approach to costing that was originally introduced by Guthrie (1974) and modified by Ulrich (1984). The general form of the equipment cost equation for equipment operation at ambient pressure using carbon steel construction is presented here:

$$\log_{10} C_p^0 = K_1 + K_2 \log_{10}(A) + K_3 (\log_{10}(A))^2 \quad 4.5$$

Where A is the capacity or size parameter; and K_1, K_2, K_3 are correlation parameters given in Table 4.1.

These cost equations are normalised to 2001 equipment costs. Further adjustments were made to account for different materials of construction and the current *CEPCI* index. The resulting expression used is:

$$C_{eq} = C_p^0 \times F_{BM} \times \frac{CEPCI_{2016}}{CEPCI_{reference}} \quad 4.6$$

Where F_{BM} is the adjustment factor for the material of construction for different equipment, and can be found in various tables presented in Turton et al. (2009), *CECPI* is the Chemical Engineering Plant Cost Index which is an inflation parameter for projects and is published monthly in “Chemical Engineering”.

The standard mass and energy balances, unit operation process models and heat transfer equations are used to determine the size of various storage vessels and process equipment in the process.

Table 4.1. Equipment Costing Parameters (Extracted from Turton et al., 2009).

Equipment	Cost Parameters			Size parameter	Material of Construction Factor
	K_1	K_2	K_3	A	F_{BM}
Crystalliser (batch)	4.5097	-0.8269	0.1344	Volume (m^3)	4.8
Shell and Tube heat exchanger (fixed tube)	4.3247	-0.3030	0.1634	Area (m^2)	5.95
Storage Tanks (shop fabricated)	3.4974	0.4485	0.1074	Volume (m^3)	8.07
Reactor (Jacketed-agitated)	4.1052	-0.4680	-0.0005	Volume (m^3)	4
Falling Film Evaporator (solvent recovery)	3.9119	0.8627	-0.0088	Area (m^2)	4

4.3.3.2.2. Operating Cost

The operating costs are based on the key utilities that are significantly influenced by the volume of selected solvent required to meet the desired production rate of API. These cost will include: the cost of solvent; cost of cooling (cooling crystallisation), cost of heating (evaporative crystallisation), cost of solvent recovery, and cost of tertiary waste treatment of unrecovered solvent, API and other

components. The various flowrates and heating and cooling are determined by material and energy balances setup in the VBA module. The utility costs is determined using the method proposed by Ulrich and Vasudevan (2006). The two factor utility cost equation is expressed as:

$$C_{Su} = a \times (CECPI) + b \times C_{Sf} \quad 4.7$$

Where C_{Su} is the price of the utility; a and b are utility cost coefficients; $CECPI$ is an inflation parameter for projects, and C_{Sf} is the price of fuel in \$/GJ. The $CECPI$ index used in this work is 556.8 (as at March 2016). The utility cost coefficients used in this work is presented in Table 4.2. The operating cost associated with a selected solvent is:

Where:

- UC_{rmi} is the unit cost of solvent/raw material i ;
- RM_i is the amount of solvent/raw material i required;
- UC_{ej} is the unit cost of cooling or heating;
- E_j is the amount of cooling and heating required;
- UC_{wst} is the unit cost of tertiary wastewater treatment;
- q is the amount of unrecoverable solvent and API.
- UC_{api} is the unit cost of unrecovered api;
- RM_{api} is the amount of unrecovered api.

Table 4.2. Utility Costing Parameters (Extracted from Ulrich and Vasudevan, 2006).

Utility	Cost unit	Cost equation coefficients	
		a	b
Electricity (purchased)	(\$/kWh)	1.3×10^{-4}	0.010
Steam	(\$/kg)	$2.7 \times 10^{-5} \times m_s^{-0.9}$	$0.0034 \times p^{0.05}$
Cooling Water	(\$/m ³)	$0.0001 + 3 \times 10^{-5} \times q^{-1}$	0.003
Refrigeration	(\$/kJ)	$0.6 \times Q_c^{-0.9} \times T^{-3}$	$1.1 \times 10^6 \times T^{-5}$
Tertiary Wastewater Treatment	(\$/m ³)	$0.001 + 2 \times 10^{-4} q^{-0.6}$	0.1
Solid/liquid disposal (Hazardous)	(\$/kg)	2.5×10^{-3}	-

Where:

- m_s = boiler steam capacity ($\frac{kg}{s}$); p = delivered pressure (bara);
- q = total water capacity (m^3/s) for cooling ; Q_c = total cooling capacity ($\frac{kJ}{s}$);

T = min cooling temperature K ; and q = plant capacity (m^3/s) for water treatment

4.3.3.2. Environmental Performance Indicators

Pollution prevention as a sustainable process solution is receiving growing interest in process industries, due to its considerable environmental and economic benefits. The two main process-related areas where this may be achieved are via reducing energy and mass usage in the process systems, and in toxic release into the environment. The incorporation of environmental ranking tools is particularly important, in preliminary process design, when a number of design alternatives need to be evaluated quickly at each level of process synthesis. The following index approaches are very simple and capable of giving a quick estimation of a design alternative.

4.3.3.2.1. The Environmental Factor

The field of Green Chemistry (Dunn et al., 2010) has proposed a practical, simple and flexible metric that can be applied to assess environmental performance. The Environmental (E) Factor which was proposed by Sheldon in 1992, shows how effective the overall production process is by comparing the amount of waste generated by the process against the amount of desirable product that gets created.

$$E = \frac{\text{Total waste produced}(kg)}{\text{Total product produced}(kg)} \quad 4.9$$

In this work, the yield of product and the amount of waste produced is directly dependent on the solvent selected and, hence, usage of the E -factor is a good environmental performance indicator. A lower E -Factor means better, environmentally, performing solvents.

4.3.3.2.2. The Energy Consumption Factor

An energy consumption (E_c) factor is also a metric that can be used to compare the energy efficiency of the selected solvent. The energy consumption factor refers to the total energy consumption per unit of product and is calculated as follows:

$$E_c = \frac{\sum \dot{H}}{\dot{M}_p} = \frac{\text{total energy consumption}}{\text{total product produced}} \text{ (kJ/kg product)} \quad 4.10$$

Where \dot{H} is the combination of heating, cooling and electrical energy required per kg of product.

In this work, since only the heating and cooling energy requirements are considered for each of the solvents evaluated during this conceptual design stage, the energy consumption factor is calculated using the heating/cooling load only.

4.4. Operating Conditions:

The operating conditions in any process are generally selected to promote the controlling phenomena in the desired direction to obtain the maximum yield. In addition, adverse conditions that may lead to safety issues or degradation of feed components or product are avoided. Typically, safety margins are incorporated into the operating conditions to ensure that adverse conditions seldom or never occur. In determining the operating temperatures for crystallisation with pure solvent, co-solvent or anti-solvent systems the following must be taken into account:

Operating temperature ranges:

The initial temperature (T_i) should be:

- Below the boiling point T_b of the solvent, to prevent uncontrolled loss of solvent.
- Below the flash point T_{flash} of the solvent, to prevent auto-ignition of air vapour mixture in the crystallizer.
- Below the degradation temperature T_{deg} of the API, to prevent loss of API.
- Below the vapourization temperature T_{vap} of the API, to prevent loss of API.

Applying a 10 degree safety margin below the lowest of the temperatures listed from a to d, then:

$$T_i = \min(T_b, T_{flash}, T_{deg}, T_{vap}) - 10K \quad 4.11$$

Also note that at T_i the feed solution should not be too close to the solubility point of the desired API, to prevent uncontrolled crystal formation. This may promote primary nucleation and limit crystal growth.

The final temperature (T_f) should be:

- Above melting point T_{msol} of solvent, to prevent freezing of solvent (unless Freeze crystallisation is the desired operation).
- Above melting point T_{mimp} of impurities, to prevent contamination of pure API, where the impurities could be unreacted starting materials or reaction byproducts (unless a fractional crystallisation operation is being implemented where an impurity is first removed).
- High enough to have low viscosity for the solvent, to facilitate movement of fluids.
- Above the approach temperature of 10°C of the cooling circuit. It is assumed that if cooling water is used as a coolant in the crystallizer, then T_{cw} is 20°C (site dependent). OR if a chilling circuit is used as a coolant, then T_{cw} could be as low as -20°C.

Applying a 10 degree safety margin above the lowest of the temperatures listed from a to c and considering the available coolant temperature, then:

$$\text{if } T_{msol} \text{ or } T_{mimp} \geq T_{cw} \text{ then } T_f = (T_{msol} \text{ or } T_{mimp}) + 10^{\circ}\text{C} \quad \text{else } T_f = T_{cw} + 10^{\circ}\text{C}$$

4.12

Other considerations:

- Solubility at the maximum feasible operating temperature should be sufficiently high to achieve the desired volume, and the yield of crystallisation must meet acceptable productivity targets.
- If the yield of crystallisation is too high (> ~90%) then co-crystallisation of impurities is more likely.
- The solvent should be easy to remove through drying and/or washing with a cleaning solvent.
- In evaporation applications, the evaporation temperature must be below the degradation temperature of the API.
- For systems where the solubility is not temperature sensitive, then anti-solvent or evaporation crystallisation should be considered.
- If the degradation temperature of the API is below the boiling point of the solvent in evaporation crystallisation, then the application of vacuum evaporation must be considered.

To assist in determining the type of crystallisation operation that is applicable for a particular application, a Decision Tree has been developed and is presented in Figure 4.5.

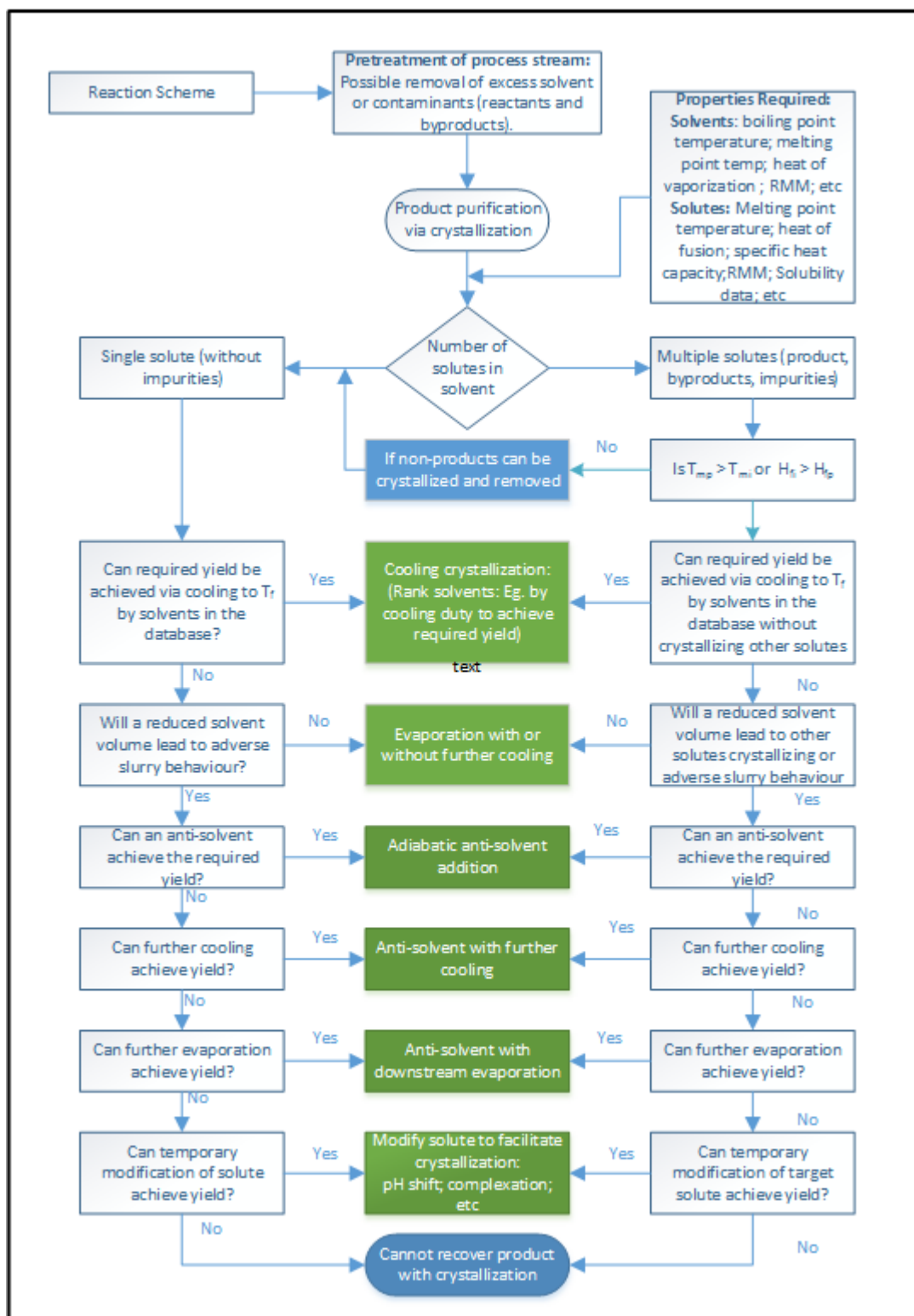


Figure 4.5. Decision Tree for the selection of crystallisation process.

Concluding Remarks

In this chapter, the various assumptions made during the development of the Solvent Selection Tool that can be applied to API crystallisation has been presented and justification provided. The various models and assumptions made for the computation of process, economic and environmental

performance are presented. The novel, modular, integrated computational framework based on a “master-slave” architecture with a commercial process simulator developed in this work to simulate crystallisation process synthesis and optimization is fully described.

In the next chapter, the Solvent Selection Tool is validated against plant and experimental data the computational capabilities is illustrated and discussed by means of a range of case studies.

CHAPTER 5: EXPERIMENTAL METHODS

Whilst verification and testing was performed consistently throughout the development process of the various subroutines and algorithms, against existing data in literature, some experimental work was undertaken to validate some of the simulation results with regards to co-solvent and anti-solvent behaviour. The solubility of APIs in binary solvent systems vary with the variation in composition of the binary solvent system. This variation in solubility forms the basis for the co-solvent and anti-solvent crystallisation. Co-solvent systems may show higher solubility in the binary solvent than in either of the pure solvents, whilst in anti-solvent systems the solubility in the binary solvent may be lower than in either of the pure solvents. In this work, we want to validate the co-solvent and anti-solvent effect by measurement of the solubility of selected APIs in selected binary systems and comparing the measured data to that predicted by the solvent selection tool.

In this section the experimental measurement of the solubility of above mentioned APIs in binary solvent systems is described.

5.1. Background

The equilibrium solubility or thermodynamic solubility of a compound is defined as the maximum quantity of that substance which can be completely dissolved at a given temperature and pressure in a given amount of solvent, and is thermodynamically valid as long as a solid phase exists which is in equilibrium with the solution phase.

The experimental methods used for the determination of solubility of solids can be classified as direct and indirect methods. In the direct methods, the solubility is measured by chemical analysis of the liquid and solid phases in equilibrium (analytical methods) or through the variation of a property of a saturated solution of known bulk composition (synthetic methods). In indirect methods, the solubility product (electrolytes only) is determined by the experimental measurements and the solubility is deduced from these measurements. (Cohen–Adad, 2003). The most suitable method for a given system depends on various factors such as available amount of substance (which is often a limiting factor during early development stages of pharmaceutical products), solvent properties like viscosity, boiling points, etc., required analytical techniques, and if additional solid-phase characterization is required (Beckmann, 2013). Some of these techniques have been comprehensively summarized in a special volume by IUPAC dedicated to the experimental measurement of solubility (Hefter, 2003).

Whatever the method finally adopted for a solubility determination, it must satisfy certain general requirements: the foremost of these is purity of the materials investigated; the second important consideration is the precise regulation and measurement of temperatures in the case of strong

temperature dependence of solubility and the third requirement to be met is the establishment of solubility equilibrium (Zimmerman, 1952). After the requirements described above have been fulfilled, a method of analysis or detection is chosen which is compatible with all other conditions in the system

In the direct methods, the procedure for measurement of solubility consists of two parts; firstly, the attainment of thermodynamic equilibrium and secondly, determination of the composition or solute concentration in the equilibrated solution. There are several methods mentioned in the literature for attainment of equilibrium (Jones, 2002), (Mullin, 1993). Generally the equilibrium solubility can be established in two ways: isothermal method and polythermal method. In the isothermal method, equilibrium solubility is achieved by either successively adding known amounts of solid to the solvent until saturation is achieved (addition method) or by equilibrating and analysing a solution containing an excess of solid (excess method). In the polythermal method, equilibrium solubility is achieved by controlled heating a solution of known composition with an initial excess of solid until last particles are dissolved. For detection, visual observation (e.g., under a microscope), turbidity measurements, particle-detecting inline probes or calorimetry may be used (Lorenz, 2013). These two techniques of establishing equilibrium i.e. isothermal (1) and polythermal (2) are shown in Figure 5.1

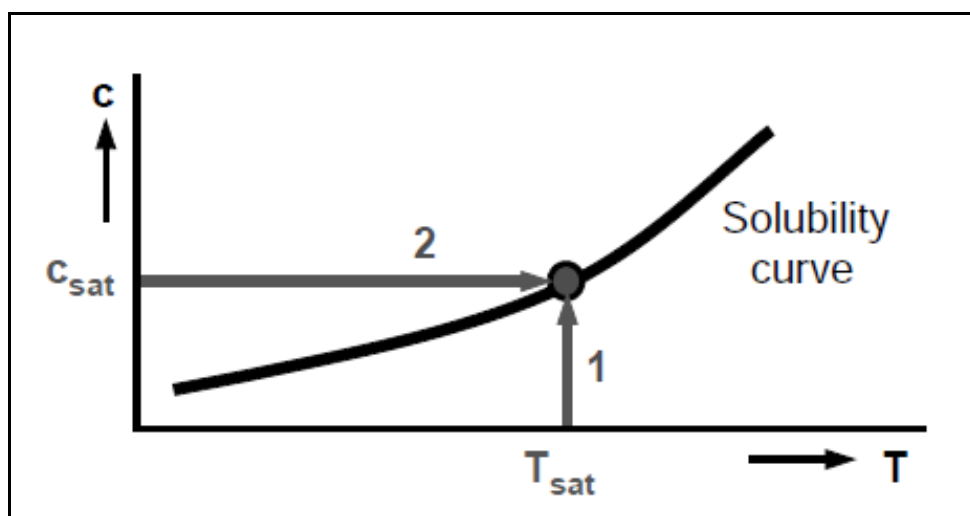


Figure 5.1. Determination of the saturation concentration by isothermal (1) and polythermal (2) methods. (Lorenz, 2013).

The principal device used for equilibrium solubility measurements by most researchers generally consists of: an equilibrium cell with a double-walled jacket that is maintained at constant temperature by circulation of thermostated fluid; an agitation device to ensure intimate mixing to eliminate temperature and composition gradients within the sample (magnetic, rotating, or vibrating stirrers are usually used); and the sampling pipet (also thermostated) equipped with a filter. Depending upon the specific conditions to be satisfied, and analysis or sampling technique numerous

modifications may be introduced into this basic design. Some typical examples of devices used for solubility measurements are shown in Figure 5.2 and Figure 5.3.

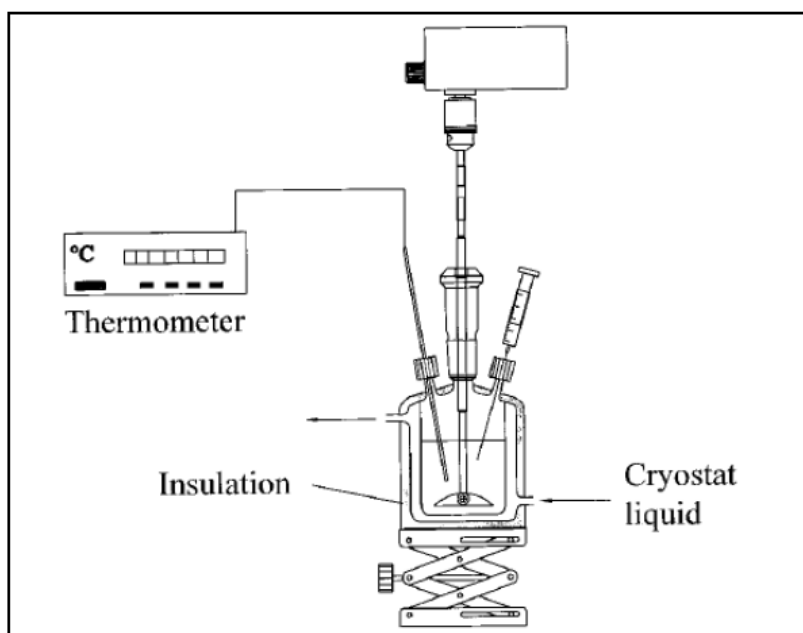


Figure 5.2. Apparatus used for analytical measurements of solid-liquid equilibria (Lohmann J, Joh R, and Gmehling J, 1997).

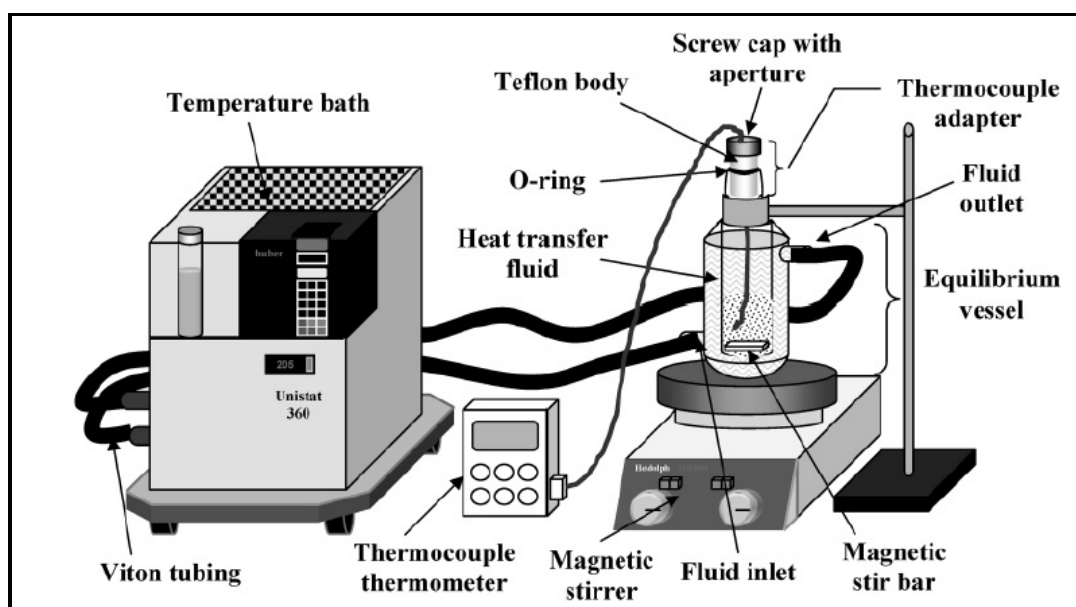


Figure 5.3. Experimental setup used for determining the solubility data by the polythermal solid-disappearance method. (Kwok et al., 2005).

Analytical and Synthetic Methods

In the analytical method, the solubility is measured under isothermal-isobaric conditions. A known amount of solvent together with solute that is in excess relative to its estimated solubility is charged into a closed vessel which is controlled to its desired temperature. The mixture is then thoroughly agitated through means of shaking or direct stirring for a sufficiently long time to ensure equilibrium. This may require several hours or days depending on the nature of the chemical system. After complete settling of the undissolved solute, samples of the saturated solution are taken and composition is obtained by chemical analysis or by comparison with a standard solution. Some techniques used to analyse the equilibrated solution include: liquid chromatography, spectroscopy (UV, IR, NMR and mass), differential scanning calorimetry, differential thermal analysis, thermogravimetric analysis, refractometry, polarimetry, etc. (Cohen-Adad, 2003). Examples of using this method for solubility determination include the works by Chiavone-Filho and Rasmussen (1993), Lohmann J et al., (1997), Kuramochi et al. (1996), Nordstrom and Rasmuson, 2006, Baka et al., 2008, Shalmashi and Eliassi 2008, Hsieng et al., 2009

The synthetic method adopted in solubility measurements exploits the principle that when a solute completely dissolves, there is an accompanying discontinuous change in physical properties. By detecting this discontinuity in the physical property (P) of a mixture of known composition, a point on the saturation boundary can be located. In principle any property (P) of the system can be used for the determination of solubility but the choice of a method will depend on the nature of the components, the range of temperature and of pressure, etc. Synthetic methods allow very precise measurements and they are time saving. Their application to binary systems is very easy but it requires a specific methodology for more complex systems. Some examples of the property (P) used in synthetic determination of solubility include: electrical conductivity or resistivity, refractive index, UV, visible or IR absorption, dielectric constant, temperature, composition, vapour pressure, pH, osmotic pressure, etc.[ref]. Examples of using this method for solubility determination include the works by Domanska et al., 1993, Chen and Ma, 2004, Domanska and Lachwa, 2005, Yongjin et al., 2005, Hahnenkamp et al., 2010 and Benziane et al., 2013.

5.2. Experimental procedure

To identify possible systems for validation, the solubility of three selected APIs in all possible combinations, of the 30 commonly used solvents in the pharmaceutical industry, were predicted with the solvent selection tool developed in this research. These predictions were based on the UNIFAC activity coefficient model. A strong variation of solubility, for the selected APIs, was observed in the ethanol and ethyl acetate binary systems. The selected APIs were Acetylsalicylic acid (aspirin), 4-Acetaminophenol (paracetamol), and 2-(4-Isobutylphenyl)-propanoic acid (ibuprofen).

The Thermogravimetric Analysis (TGA) method is used to measure the solubility of the APIs in the binary solvent mixture. Thermogravimetric analysis is an experimental method whereby changes in

mass are used to detect and measure the chemical or physical processes that occur upon heating a sample. The TGA instrument consists of a highly precise analytical balance to which the sample pan is attached. The sample pan is typically suspended within a controlled atmosphere that can be heated. With the use of isothermal or dynamic heating and cooling cycles, the TGA is programmed to hold a specific temperature once a change in mass is detected, then that temperature will be maintained until no further change is observed.

The accuracy of solubility measurements is dependent on many factors, including sample and instrument preparation, the accuracy of the balances used, as well as precise control of heating/cooling rates, and atmospheric conditions. In this study, the Differential Thermal Analysis (DTG 60AH) Shimadzu is used. This is connected to a Shimadzu Thermal Analyser (TA 60WS) and a Shimadzu Flow Controller (FC 60A). These are photographed in Figure 5.4. The APIs and solvents used in the experimental work, as shown in Table 5.1., were all obtained from Sigma-Aldrich.



Figure 5.4. The differential thermal analysis equipment used for solubility measurements.

Table 5.1. Chemicals used for ternary systems solubility measurements.

Compound	CAS No.	RMM	Purity %
Acetylsalicylic acid	50-78-2	180.16	99.0
Acetaminophen	103-90-2	151.16	99.0
Ibuprofen	15687-271	206.28	98.0
Ethanol	64-17-5	46.07	99.8
Ethyl Acetate	141-78-6	88.11	99.7

5.2.1. Samples Preparation

Sample preparation is undertaken as follows:

- 1 Firstly, the vials and pans were washed using soap and rinsed with distilled water. The vials were filled with acetone and were placed in the ultrasonic bath for 20 minutes at 303.15 K.
- 2 The pans and vials were emptied and dried in an oven at the temperature of 333.15 K for the period of 15 minutes. This was done to ensure that there was no acetone residual in the vials and they were ready for sample preparation.
- 3 The gravimetric method was used to prepare the solutions. The samples were prepared using an analytical balance, which has a manufacturer-stated-uncertainty of ± 0.0001 g in mass. Every measurement was repeated three times and the average of the three measurements was calculated.
- 4 The calculated/literature value amount was first added, plus 20% of API (to ensure saturated solution), to the vials.
- 5 The binary solvent systems of ethanol and ethyl acetate are varied from 1.0 to 0.0 mole fraction ethyl acetate. In addition to the pure solvent measurements, 8 intermediate concentrations were made. The calculated amount of the solvents were added and the weights noted, these later give the composition of the binary solvent system.
- 6 The samples were kept in the ultrasonic bath for 4 hours at the desired temperature for solubility measurements. Finally, the samples were kept at a single point in the water bath for 20 to 24 hours at the desired temperature for solubility measurements.

5.2.2. Solubility Measurement

Solubility measurements were undertaken:

- 1 Prior to solubility measurements, it is important to keep the room temperature, syringe and sampling pan at the desired temperature for solubility measurements. This was done to

minimize error. A small amount was drawn from the vial using a syringe.

- 2 The weight of the empty pan was measured and a few drops (3-5 drops) were added to the pan (it should be half full). The solvents will also start to evaporate immediately, so the first weight measurement was taken for the total amount.
- 3 The pans were placed at the TGA and the auto sampler was programmed. The temperature for the TGA has to be 10 degrees below the lower boiling point of the solvents. The residence time in the TGA is dependent on the volatility of the solvents.
- 4 The temperature of the chamber was set at 10 K below the boiling point of solvent. Then, the TA program was started, and the solvent was evaporated slowly until the mass was constant. The pan with the API was weighed using a mass balance, and the final value was recorded.
- 5 With the 3 weight values (the empty pan, the pan with the saturated solution and the final pan with the solid) the solubility of solute in the binary solvent system at a specific temperature can be calculated.
- 6 The solubility of the API in the solvent mixture is calculated by:

$$x_{API}^{eq} = \frac{M_{API}}{M_{API} + M_{ethyl\ acetate} + M_{ethanol}} \quad 5.1$$

where

$$M_i = \text{moles of component } i = \text{mass of component } i / \text{Relative Molecular Mass of } i$$

The procedure to determine the uncertainties associated with the solubility measurements by the method employed in this study is detailed by Moodley et al. (2017). The standard relative composition uncertainty of the measurements is due to the uncertainty in the experimental mass measurements.

$$u_r = \frac{f(u_{MM})}{x_i^{sat}} = \frac{u_c}{x_i^{exp}}$$

Where u_r is the standard relative uncertainty in composition, u_{MM} is the uncertainty in the measurement of mass, u_c is the standard uncertainty in composition, and x_i^{exp} is the measured composition. It is assumed that all the experimental uncertainties result from the uncertainty in the solute mass measurements and is justified with the small change of solubility within the standard experimental uncertainty of 0.1K. The standard uncertainty in composition also includes the uncertainty of repeatability of mass between each repeatability measurement.

CHAPTER 6: RESULTS AND DISCUSSION

The primary objective of this work is the development of a computational tool that can be used to select solvent systems and crystallisation methods for the production of pharmaceutical and natural products and their intermediates. Integral to the development of such computational tools are the means to verify their capability, range of applicability and degree of accuracy. These components are assessed and demonstrated in this chapter, presented in two distinct sections: in the first section the focus is on the validation of the process and the thermodynamic models and algorithms; and in the second section, the computational capability of the Solvent Selection Tool developed are demonstrated by means of a broad selection of applications.

6.1 Model Validation

The Solvent Selection Tool that is developed is based on a strong theoretical foundation, with a combination of predictive thermodynamic property models and process models. Whilst these models and algorithms may have inherent limitations in terms of range of applicability, the predictive tool is only as good as the accuracy of its predictions in relation to a working process. Hence, the reconciliation of available measurement data and the validation of process models are mutually connected.

Not only is process knowledge required to define a process model, but the process model itself is often required to reconcile measured variables with available measurements drawn from the fundamental material and energy balances of the industrial process. The process of determining the degree to which the Solvent Selection Tool and its associated systems are an accurate representation of the real world, from the perspective of the intended uses of the model, is presented in this section.

Whilst verification and testing was performed consistently throughout the development process of the various subroutines and algorithms, the validation process presented here sets out to verify the following three assumptions made in the development of the Solvent Selection Tool: firstly, that predictive activity coefficient models based on group contributions methods adequately predict the solubility and solid-liquid phase equilibria of complex molecules such as APIs; secondly, that multiphase equilibria flash calculations can be used to predict the crystallisation phenomena; and thirdly, that the various modes of crystallisation processes, such as cooling, evaporative and anti-solvent crystallisation, can be modelled by using multi-component, multiphase flash calculations.

This validation process was achieved by means of the following processes: A confirmation of the performance of the predictive activity coefficient models to reasonably predict the phase behaviour of complex molecule systems is achieved by comparing the outputs of the Solvent Selection Tool to measured experimental data obtained from the DDBST database, and also against some experimental measurements made in this work; and their ability to simulate real industrial crystallisation processes is

confirmed through a comparative analysis, using production plant data, for the crystallisation of a natural flavourant from a complex feed-stream.

6.1.1. Reliability of Predictive Activity Coefficient models:

Several studies have been undertaken to assess the implementation of API solubility prediction, where the predictive performance of various models is compared to experimental measurements. Some of these comparative performance studies are summarised in Table 3.3., and a few of these models are fully described in chapter 3.

A general observation, reported in the various studies, is that the accuracy of solubility prediction depends upon how much the system of interest conforms to, or deviates from, the model assumptions, and how accurately the required parameters are known or can be predicted. UNIFAC has been a preferred choice of model for various applications. It can be extended, either by adding new terms or by regressing new parameters, revealing its usefulness and versatility.

Today, the Dortmund modified UNIFAC is probably the most, well-developed model. It is widely used and is being revised and extended as part of various industrial/academic joint ventures. In recent years, numerous researchers have added new groups and subgroups and have published new group parameters (Abildskov et al., 2004).

Since the Solvent Selection Tool uses CHEMCAD as its computational engine, any predictive phase equilibria calculations performed are limited to the predictive models that are available in CHEMCAD. CHEMCAD's two predictive activity coefficient models, to which this work is limited, are the UNIFAC and Modified UNIFAC (Dortmund) models. The subgroup listing available on CHEMCAD and its UNIFAC Consortium supplement are presented in Appendix D.

Since the crystallisation process is predominantly dependent on solubility, the performance of significant models used to predict the solubility behaviour of APIs in complex solvent systems are presented here; and are compared to some of the experimental work undertaken. The measured data have standard uncertainties of $u(T) = 0.1K$ and $u(P) = 0.002MPa$ and a relative uncertainty of $u_r(x_i^{exp}) = 0.0003$. As outlined in section 5.1, the three selected APIs are Acetylsalicylic acid (Aspirin), 4-Acetaminophenol (Paracetamol) and 2-(4-Isobutylphenyl) propionic acid (Ibuprofen). The solubility is measured for a binary solvent system that showed both an increase and a decrease in API solubility over the solvent concentration range.

Whilst the accuracy of the predictions varied for the various APIs, the binary solvent systems, there was an evident ability of the solvent selection tool to predict the general solubility behaviour of co-solvency and anti-solvency over the binary solvent concentration range. The UNIFAC and Modified UNIFAC (Do) models' predictions for Acetylsalicylic acid are shown in Figure 6.1. As shown in Figure 6.2, the predictions for 2-(4-Isobutylphenyl) propionic acid generally revealed a reasonable and conservative

estimate of solubility, while the UNIFAC model revealed an over-prediction for 4-acetaminophenol, as shown in Figure 6.3. The UNIFAC model's over estimation of 4-acetaminophenol solubility, in a binary ethanol-ethyl acetate solvent, is predominantly due to the poor UNIFAC model prediction of the solubility of 4-acetaminophenol in ethyl acetate, as shown in Figure 6.4. However, the UNIFAC model shows a better prediction of the solubility of 4-acetaminophenol in ethanol, as shown in Figure 6.5. The predictive models solubility in pure solvents predictions are compared to data measured by Granberg and Rasmuson (1999).

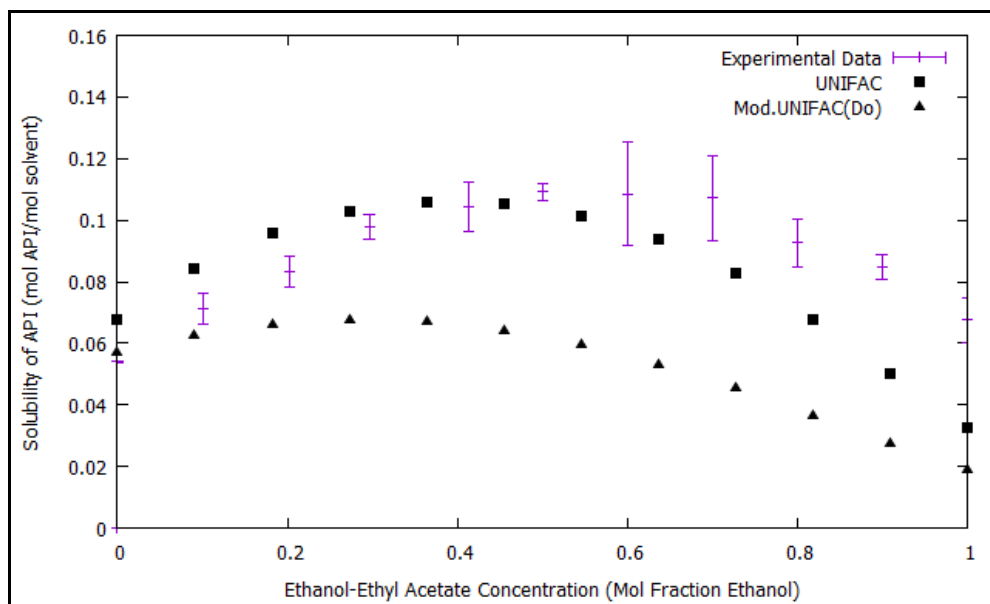


Figure 6.1. Comparison of Predicted and Measured Solubility of Acetylsalicylic Acid in a Binary Solvent at 25 °C.

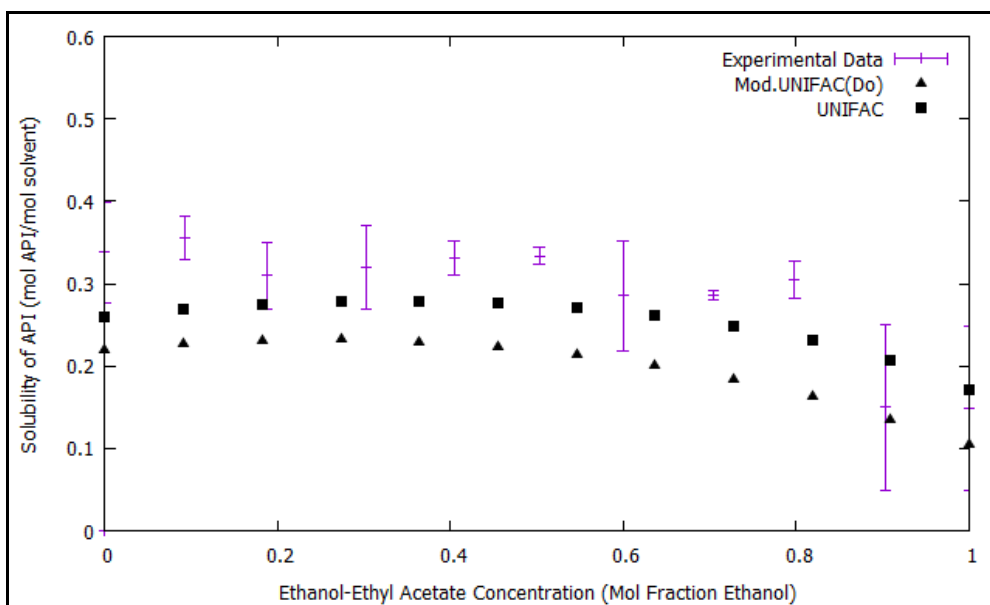


Figure 6.2. Comparison of Predicted and Measured Solubility of 2-(4-Isobutylphenyl) Propionic Acid in a Binary Solvent at 25 °C.

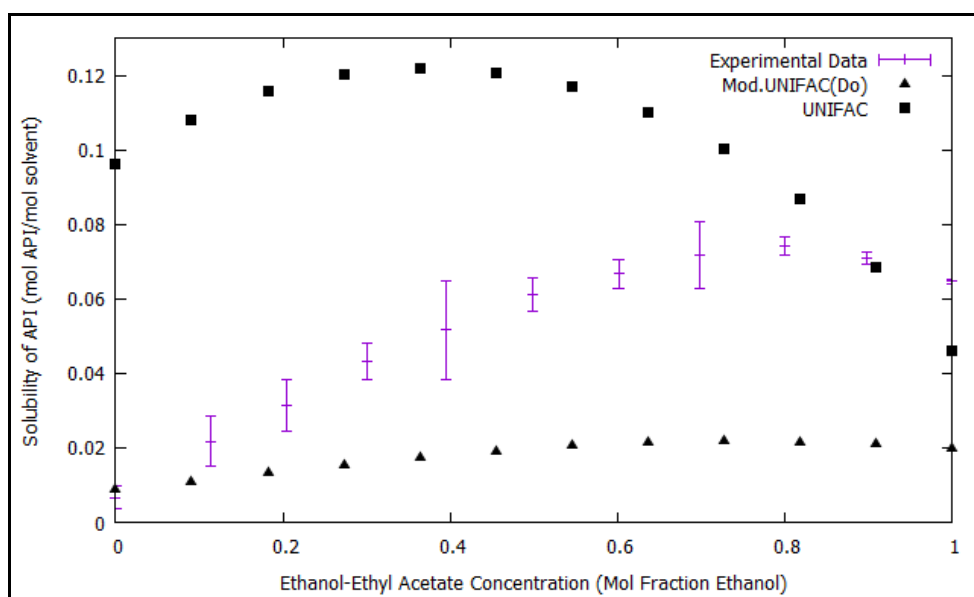


Figure 6.3. Comparison of Predicted and Measured Solubility of 4-Acetaminophenol in a Binary Solvent at 25 °C.

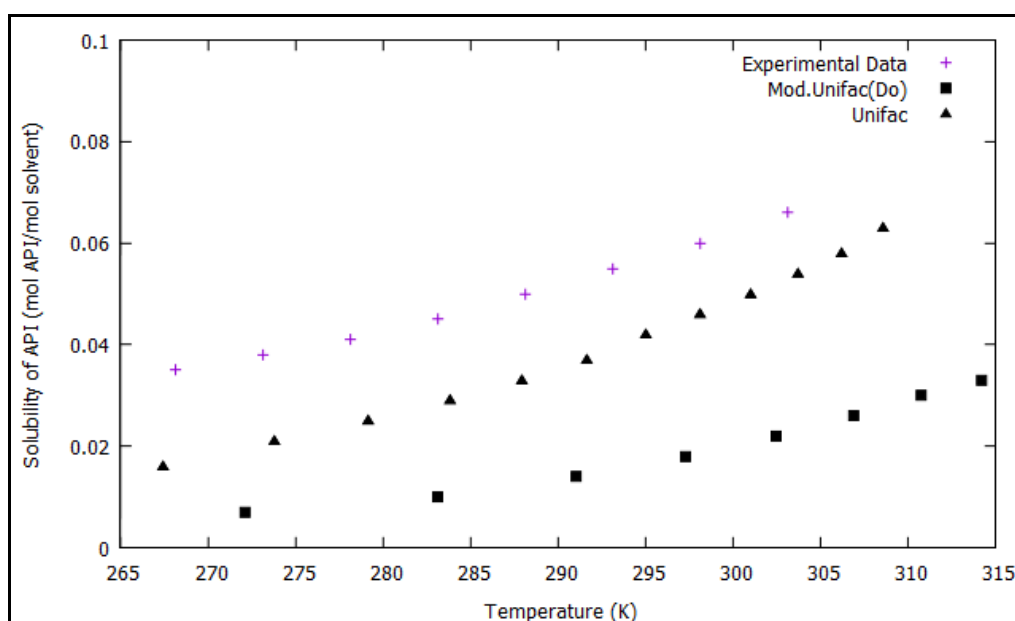


Figure 6.4. Comparison of Predicted and Measured Solubility of 4-Acetaminophenol in Ethanol Data of Granberg and Rasmuson (1999).

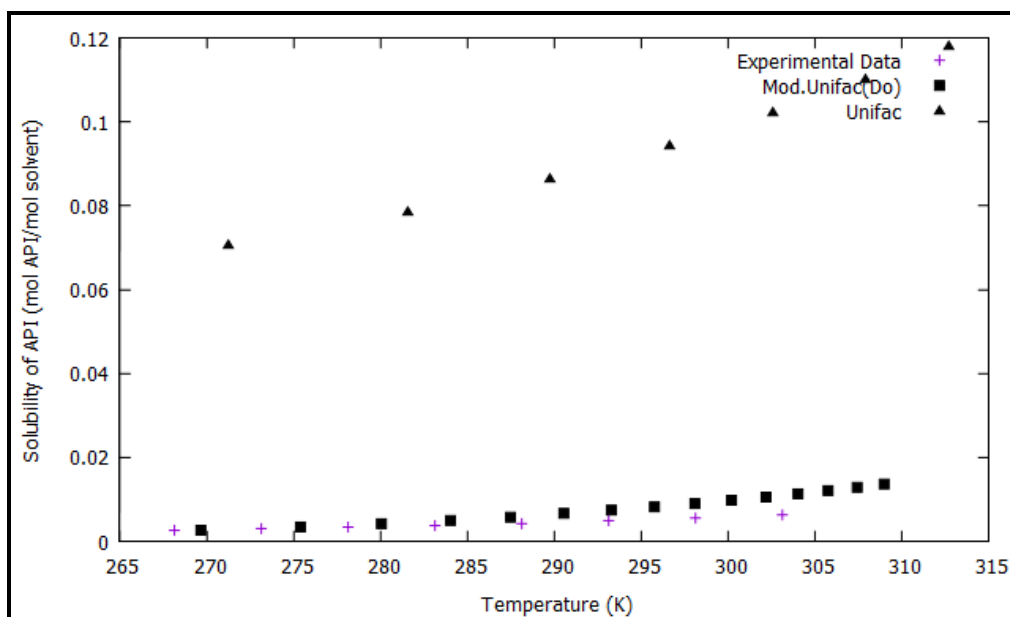


Figure 6.5. Comparison of Predicted and Measured Solubility of 4-Acetaminophenol in Ethyl Acetate Data of Granberg and Rasmuson (1999).

Both the UNIFAC and Modified UNIFAC (Do) models generally predict conservative values and are capable of predicting the general behaviour of complex systems. Improvement in capability and accuracy are the primary objectives of the ongoing development and refinement of the predictive and semi-predictive activity coefficient models. The current variance in the accuracy of the predictive and semi-predictive methods for complex multi-functional molecules, like APIs, warrants experimental work to confirm the results of predictive methods. Hence, as a conceptual design tool, predictive methods are invaluable in identifying the necessary experimental measurements required.

6.1.2. Reliability to Predict the Performance of Industrial Crystallisation Processes: Production of Natural Flavourant 2,3 Butanedione (Diacetyl).

6.1.2.1. Description of Production Process

Diacetyl is a naturally derived food flavourant that can be obtained as a by-product from several processes such as fermentation and biomass hydrolysis. A production facility downstream of a sugar mill process uses one of the waste streams of the sugar mill to produce a range of high value products, including the extraction of Diacetyl. After several separation and conversion processes, all the streams containing Diacetyl are combined for further processing. The key objective being the recovery of the Diacetyl.

The process consists of two sequential batch crystallisers designed to produce high purity Diacetyl crystals from a complex mixed-feed of several components. The feed stream to the primary crystalliser combines several process streams and the mother liquor from the secondary crystallizer. It contains the

following major components: acetone, acetaldehyde, ethanol, 2,3 pentanedione, 2,3 butanedione (Diacetyl) and water.

The recovery of Diacetyl through distillation is not an option because of the several complex azeotropes that exist in the feed stream. Fractional crystallisation is a viable process to recover the Diacetyl. The Diacetyl crystals obtained from the primary crystalliser are re-melted and combined with another water-Diacetyl process stream, and fed into the secondary crystalliser. The crystals obtained from the secondary crystalliser are purified to obtain a final purity of 98 % natural flavourant product. The crystallisers are cooled with an ethanol cooling circuit available at -35 °C.

The Solvent Selection Tool is used to simulate a cooling crystallisation process, and to compare the yields obtained to actual plant performance. The design feed specifications to the primary and secondary crystallisers are presented in Table 6.1, and the relevant physical properties of the feed components are presented in Table 6.2.

Table 6.1. Design Feed Composition to the Primary and Secondary Crystallisers.

Crystalliser Feed Components	Primary Crystalliser Design Feed Composition (%)	Secondary Crystalliser Design Feed Composition (%)
Acetone	6	2
Acetaldehyde	6	2
Ethanol	10	3.25
2,3 Pentanedione	6	2
2,3 Butanedione	63	84.25
Water	9	5.5

6.1.2.1. Performance of Primary Crystalliser

The primary crystalliser has been designed for a target production rate of 500 kg per batch, which represents a 50 % yield of Diacetyl as crystals per batch. Plant data for 1200 batches (which represent 1200 production days) have been sourced and analysed. The actual plant data show a broad spectrum of feed and operating conditions as compared to the design specification feed. This broad spectrum of feed and operating conditions is attributed to several causes which include: feed variation with seasonal biomass feedstock (sugar cane), the age of the feedstock, variation performance of upstream processes leading to variation in batch sizes, recycling of off-spec batches and plant operational problems.

Table 6.2. Major Feed Component Properties.

Crystalliser Feed Components	RMM (g/mol)	Boiling Point (°C)	Melting Point (°C)	Heat of Fusion (kJ/mol)	Specific Gravity
Acetone	58.08	56	-94	5.7	0.79
Acetaldehyde	44.05	20.8	-125	2.31	0.78
Ethanol	46.07	78.35	-114.15	4.64	0.78
2,3 Pentanedione	100.11	110	-52	7.84	0.99
2,3 Butanedione	86.09	87.5	-2.4	38.5	0.99
Water	18.01	100	0	6	1

In addition, the control of the crystallisation process is largely subjected to human decision, where the termination of crystallisation process is based purely on the physical appearance of the crystalliser content, and hence dependent on the experience of process operator. This broad spectrum of feed conditions, operating conditions and crystal yield is shown in Figure 6.6.

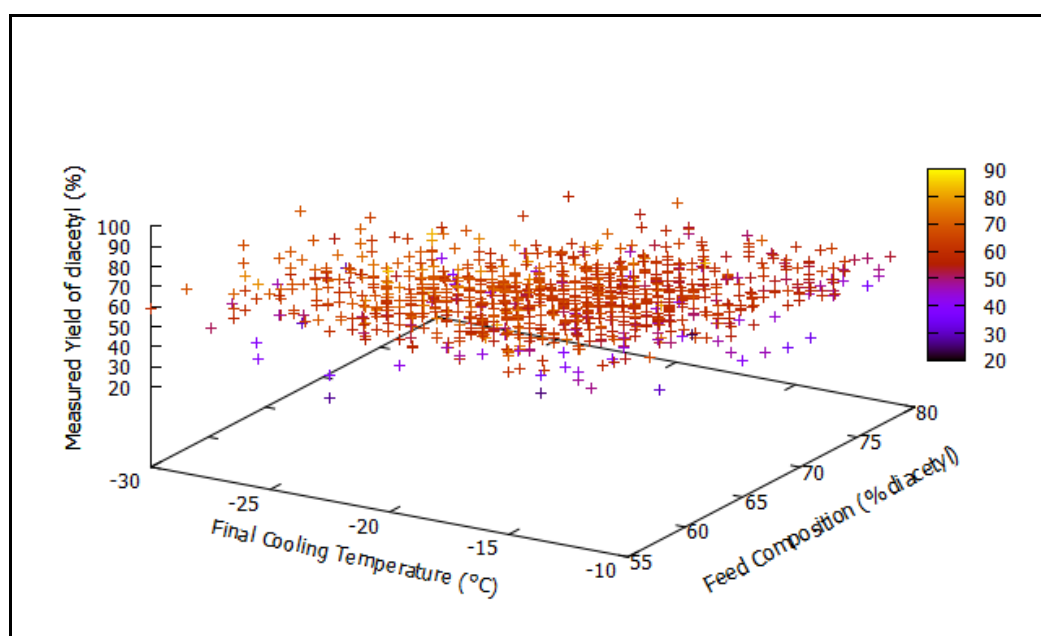


Figure 6.6. Production Plant Data for Primary Crystalliser for 1200 Production Days.

From 1200 data recordings, only 160 batches reflect the design feed composition of 63 % Diacetyl in the feed. The batch feed volume and proportion of the other major components are unknown. This sample of 160 data points was analysed and the results of a statistical analysis performed on the sample of data is presented in Table 6.3.

Table 6.3. Analysis of Primary Crystalliser Plant Performance Meeting the Design Feed Composition.

Statistical Value	Feed Composition (% Diacetyl)	Final Cooling Temperature (°C)	Diacetyl Crystal Yield (%)
Average Value	63	-19.4	50.6
Maximum Value	64	-11	85.2
Minimum Value	62	-31	1
Standard Deviation	0.6	3.7	19.6

This sample of data (which meets the feed composition of 63 % Diacetyl) is shown in Figure 6.7. The region enclosed with horizontal and vertical lines reflects those operations that were within the average operating conditions (mean value \pm standard deviation). The Solvent Selection Tool's crystal yield prediction for the complex feed, using three different predictive thermodynamic models for activity coefficient, is also shown in Figure 6.7. The three thermodynamic models used are: UNIFAC, Modified UNIFAC (Dortmund), and the Scatchard-Hildebrand model. At the average conditions of feed composition of 63 % Diacetyl and a final cooling temperature of -19.4 °C, a predicted yield of 45.3, 0 and 78 % are predicted by the Mod. UNIFAC (Do), UNIFAC, and Scatchard-Hildebrand models, respectively, compared to the plant average of 50.6 % at these conditions. It should be noted that the Scatchard-Hildebrand model was developed for non-polar solvents with positive deviations from ideality, and therefore may have limited application to this system.

From Figure 6.7, it can be seen that the Solvent Selection Tool, using the Modified UNIFAC (Dortmund) model for activity calculations, presents an acceptable prediction for the complex industrial system. The model under-predicts the average plant yield by approximately 10 %. As a conceptual design tool, this level of accuracy is considered acceptable.

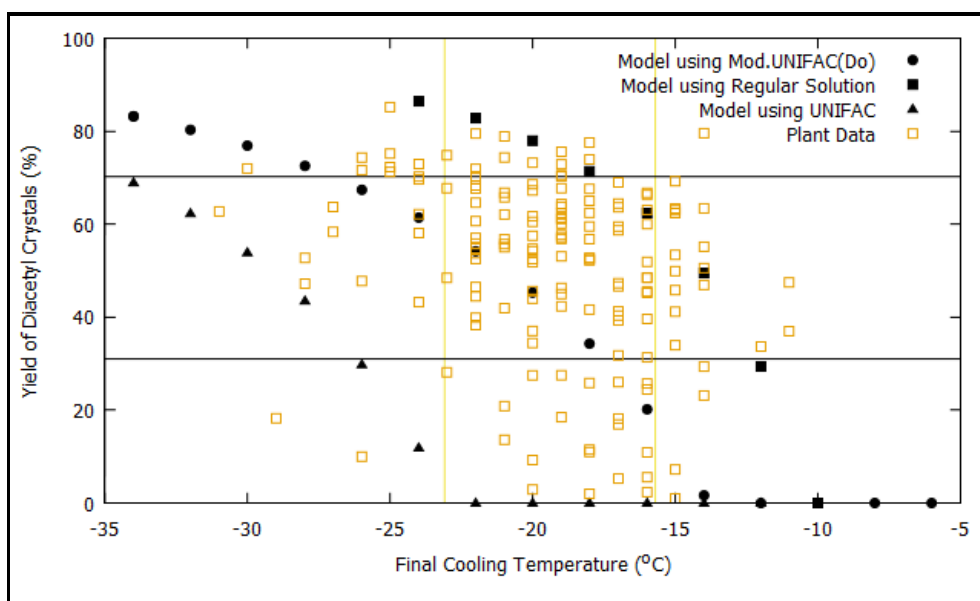


Figure 6.7. Comparison of Predicted and Actual Plant Performance of the Primary Crystalliser.

6.1.2.1. Sensitivity Analysis to Identify Key Variables Affecting Performance of Primary Crystalliser

A further preliminary evaluation was conducted to establish the factors that may contribute to the large variance of performance, as observed in the 1200 plant data. The presence of several “solvent” type components in the feed, like ethanol, water, acetone and acetaldehyde, warranted the evaluation of the impact of the feed composition on the performance of the system. The variation in composition of the various components in feed is largely attributed to the composition and age of the biomass (sugarcane) feed. Whilst the amount of Diacetyl in the feed stream will have an obvious direct impact on yield, where a higher initial concentration generally results in higher yields, the influence of the other components needs to be assessed.

The sensitivity on feed composition was evaluated using the Solvent Selection Tool, and it was observed that four of the main components: water, ethanol, acetone and acetaldehyde have varying effects on the yield. The influence of these four components on the Diacetyl crystal yield is shown in Figure 6.8.

It is observed that water and acetone have a co-solvency effect which result in the yield decreasing with increases in concentration in any of these two components in the feed-stream. A 50 % variation of water in design feed composition ($9\% \pm 4.5\%$) could result in the yield varying from 55 % to 40 % at $-20\text{ }^{\circ}\text{C}$. Similarly, a 50 % variation of acetone in design feed composition ($6\% \pm 3\%$) could result in the yield varying from 52 % to 44 % at $-20\text{ }^{\circ}\text{C}$. An increase in concentration of both components in the feed stream could result in even lower performance. A decrease in concentration of any of these components in the feed stream, from designed feed conditions, would lead to an increased yield.

The other two components in the feed-stream, ethanol and acetaldehyde, have an anti-solvent effect associated with increases in yield. A 50 % variation of ethanol in design feed composition ($10\% \pm 5\%$) could result in the yield varying from 41 % to 55 % at $-20\text{ }^{\circ}\text{C}$. Similarly, a 50 % variation of acetaldehyde in design feed composition ($6\% \pm 3\%$) could result in the yield varying from 45 % to 52 % at $-20\text{ }^{\circ}\text{C}$. An increase in concentration of both components in the feed stream could result in even higher yield in the primary crystalliser. A decrease in concentration in any of these components in the feed stream, resulting from the design feed conditions, would lead to lower yields.

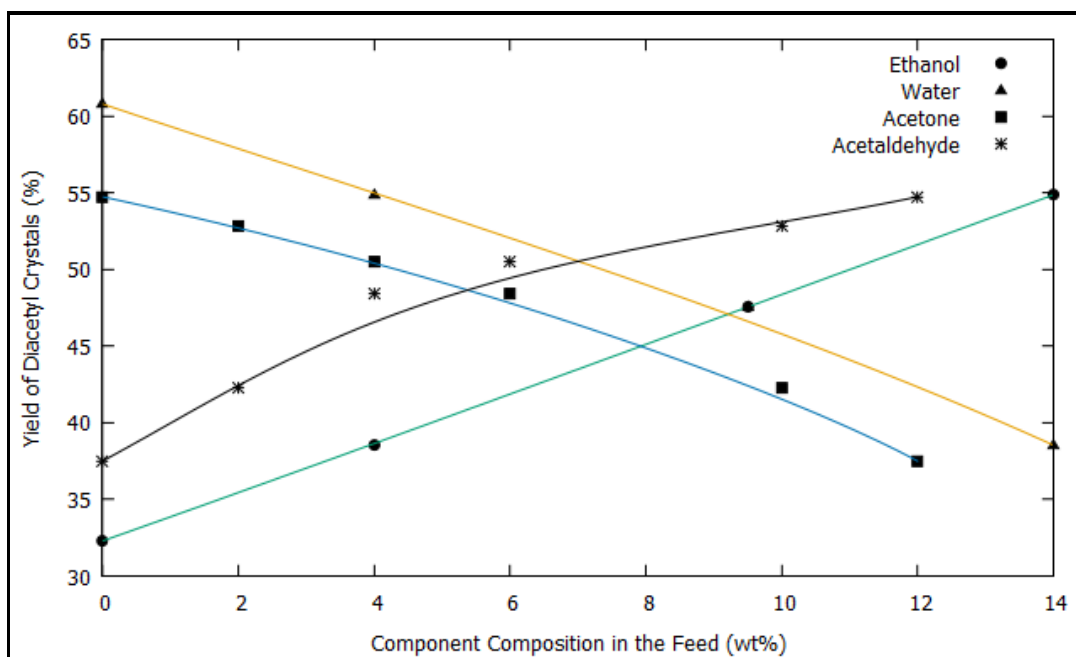


Figure 6.8. Effect of Feed Component Composition on the Primary Crystalliser Performance.

A sensitivity analysis confirms that variation in the feed composition of the four “solvent” components associated with the biomass feedstock and the various process streams may account for the large variance in yields, even when the Diacetyl concentration is within the design feed specification.

6.1.2.3. Performance of the Secondary Crystalliser

The feed to the secondary crystalliser consists of the Diacetyl crystals obtained from the primary crystalliser which are re-melted and combined with another water-Diacetyl process stream. The operation plant data is presented in Figure 6.9. From 1200 data recordings, approximately 50 % of the batches reflect compliance with the design feed composition of 85 % Diacetyl in the feed. The reasons for the large spectrum of feed specifications are the same as presented for the variation in feed to the primary crystalliser. The batch feed volume and the proportion of the other major components are unknown. This generated sample of 616 data points was analysed, and the results of the statistical analysis performed is presented in Table 6.4.

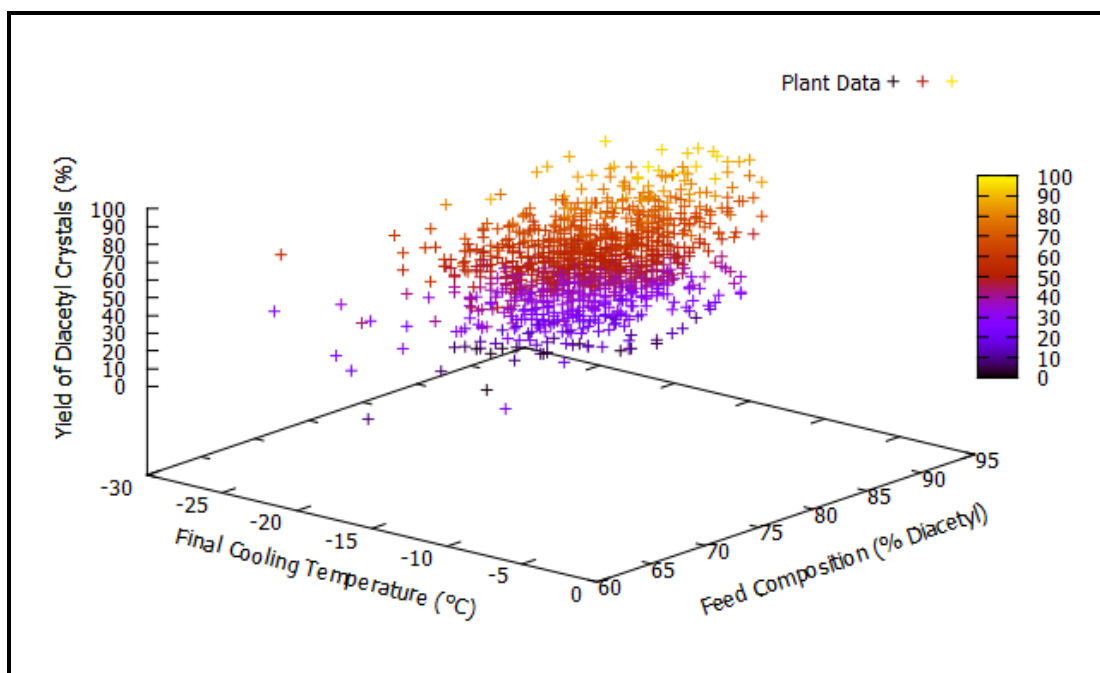


Figure 6.9. Production Plant Data for Primary Crystalliser for 1200 Production Days.

Table 6.4. Analysis of Secondary Crystalliser Plant Performance Meeting the Design Feed Composition.

Statistical Value	Feed Composition (% Diacetyl)	Final Cooling Temperature (°C)	Diacetyl Crystal Yield (%)
Average Value	84.1	-16.5	58.07
Maximum Value	90.0	-11.0	96.44
Minimum Value	80.1	-23.0	40.10
Standard Deviation	2.5	2.5	11.98

A sample of data that meets the feed composition of 85 % Diacetyl is shown in Figure 6.10. The region enclosed within the horizontal and vertical lines reflects those operations that were within the average operating conditions (mean value \pm standard deviation). The Solvent Selection Tool is used to predict the crystal yield of the complex feed into the secondary crystalliser using the Modified UNIFAC (Dortmund) model.

At the average conditions of feed composition of 84.1 % Diacetyl and cooling to -16.5 °C, a predicted yield of 73 % is obtained, compared to the plant average of 58.07 % at these conditions. From Figure 6.10, the theoretical model over-predicts the plant average yield by approximately 25 %. The observations on the sensitivity analysis performed for the primary crystalliser also apply to the secondary crystalliser. The variation in the feed composition to the second crystalliser may also contribute to the plant yields being lower than predicted yield values.

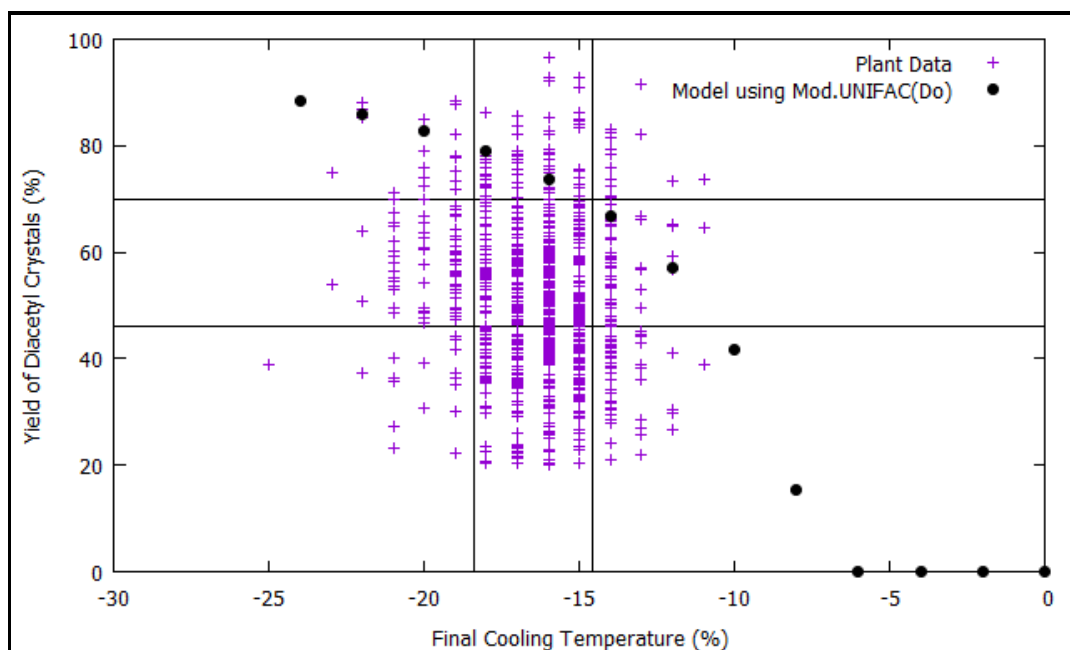


Figure 6.10. Comparison of Predicted and Actual Plant Performance of the Primary Crystalliser.

The fact that the process model closely describes the physical measurements from the real process is an adequate validation of the Solvent Selection Tool. The major differences between the calculated and reported results have been suitably explained with the sensitivity analysis performed, and the other reasons largely attributed to seasonal biomass feedstock and that the termination of the crystallisation process is dependent on operator observations.

In this section, we have provided adequate evidence that the thermodynamic and process models and the algorithms used in the Solvent Selection Tool are capable of being used as a conceptual design tool for decision making relating to crystalliser performance. The predictive activity coefficient models based on group contributions methods adequately predict the solubility and solid-liquid phase equilibria of complex molecules such as APIs and natural compounds. The multi-component, multiphase, flash calculations approach can be used to predict the crystallisation phenomena and the various modes of crystallisation processes such as cooling, evaporative and anti-solvent crystallisation.

6.2. Case Studies

The application and capacity of the developed solvent selection procedure will be illustrated through several case studies. The method explores the synergistic combination of multiphase, phase equilibria phenomena and process systems engineering methods to develop a decision tool for faster process design and process understanding during the conceptual phase of API development or the retrofit of an existing process.

The solvents considered will be restricted to a GSK list of solvents recommended for use in the pharmaceutical industry. The criteria used to compare the various solvents will include a process-related criterion such as yield, an economic criterion such as operating and fixed annualized costs, and environmental criteria such as waste generation and energy usage.

The case studies selected highlight the comprehensive computational capabilities of the Solvent Selection Tool and its invaluable role in decision making during conceptual and retrofit design of crystallisation processes. The key feature of the computational tool is that the only information required to perform all its calculations is the chemical structure of the components involved. The case studies considered are:

- Application One: *Identification of Potential Solvents for API crystallisation.*

For a newly developed API, identify a pool of suitable solvents that can be used for cooling crystallisation (desired operation in the pharmaceutical industry) that will give the required yield, quality and production rate. This case study illustrates the usefulness of the computational tool during the research and development phase of pharmaceutical products. It provides a decision making tool which is not limited to just the yield but also illustrates the estimated size of equipment required as well as the economic and environmental impact that is associated with all of the selected solvents. With the application of user-defined, decision making criteria the output of the computational tool will be a shortlist of solvents that fulfil the user-defined criteria.

- Application Two: *Identification of Modes of Crystallisation to fulfil production objectives.*

Whilst cooling crystallisation is the desired mode of crystallisation in the pharmaceutical industry, suitable alternative crystallisation processes need to be considered if the desired yield, quality and production rate cannot be achieved via cooling crystallisation. This will apply to systems that show little or no solubility variation in relation to temperature or production facilities with limited cooling capacity (eg. The Lowest achievable temperature in the plant may be limited to -5°C).

To overcome this production limitation, evaporative and anti-solvent crystallisation processes will be evaluated as alternatives to cooling crystallisation. In addition, a combination of systems (evaporation and cooling; evaporative and anti-solvent and cooling and anti-solvent systems) will be evaluated for an existing API and solvent system to identify the operating options that will give the required yield and production rate. Potential operating conditions will be shortlisted for scenarios with and without process and plant constraints. This case study illustrates the computational range and flexibility of the tool in process development as well in the re-evaluation and optimization of existing production systems.

Application Three: *Fractional crystallisation of multi-component API feed.* For a given API and solvent system in the presence of impurities (unreacted starting materials and byproducts), the operating options and conditions that will give the required yield and production rate will be identified. This case study illustrates the impact of the presence of impurities in a crystallisation feed-stream. It illustrates how the tool can be used for systems analysis to identify the phase equilibria limitations and then to synthesise an alternate processing route to obtain the maximum yield of pure product.

6.2.1. General Observations

A generally accepted phenomenon in chemical conversion processes is that for a given target production rate, the size of the processing facility, the amounts of raw materials required, as well as the utility requirements, are directly dependent on the yield/conversion of key operations. In API production, one such yield-dependent key operation is the crystallisation process.

For the target production rate of an API crystal, the attainable yield by crystallisation will impact on the upstream and downstream processing requirements. Low attainable yields invariably imply that larger volumes of materials must be processed to obtain the target production rate. This larger processing volume of materials will require larger equipment or greater numbers of modular units, and greater process utility requirements such as cooling, heating and waste treatment. Hence, the resulting primary objective during the selection or design of processes will be to select process options that render high yields.

The first step in the application of crystallisation as a unit operation is to examine equilibrium solubility data, as this is indicative of the options available to create the desired state of supersaturation, which is the driving force in crystallisation. The state of supersaturation may be created either by increasing the solute concentration or by decreasing the solute solubility. The solute concentration approach is generally achieved by evaporating the solvent, and the solute solubility approach is generally achieved by cooling, or anti-solvent addition. To assist with the screening of the type of crystallisation process to be selected, a decision making tool has been developed in this work and is shown in Figure 4.5.

From the decision-making flowsheet, the preferred choice for an API crystallisation process can be ranked as follows: the first choice is cooling crystallisation, followed by evaporative crystallisation, and last of all, anti-solvent crystallisation. This hierarchy of choice is because temperature and cooling rate are relatively easy to control. If the desired substance exhibits a strong temperature dependence for solubility, then the state of supersaturation can be achieved by cooling the mixture. If, on the other hand, the equilibrium line is relatively flat and shows little temperature dependence, then an evaporative process might become necessary.

If the yield from either of the processes is low, or there are process limitations such as thermal degradation of API at evaporation temperatures that prevents evaporation from being used, then a second solvent (anti-solvent) can be added to reduce the solubility. For a sparingly soluble solute, the addition of a second solvent to increase solubility could be selected. The addition of a co-solvent is an effective way of increasing productivity. The disadvantage of the use of additional solvents is the increased complexity of the process if solvent recovery and reuse is a priority. The addition of co-solvents and anti-solvents will also lead to an increase in processing volumes which will impact on equipment size and utility requirements.

In the applications that follow, reference will be made to various calculation procedures and algorithms of the Solvent Selection Tool. The full menu and calculations of the tool is shown in Figure 4.4, and a detailed description of computational capabilities is presented in Appendix B.

6.2.2 Applications

6.2.2.1 Application 1: Identification of a Pool of Potential Solvents for Cooling Crystallisation for a new API.

During the development phase of an API, or the retro-fit of an existing API manufacturing process, the use of computational process engineering tools to identify potential process options is essential, as it is unrestrictive in the range of options it can evaluate. In this application, the Solvent Selection Tool developed in this work will be employed to identify suitable solvents that can be used to ensure the desired yield and quality during the production of pure API crystals via crystallisation.

A generic and multi-tiered methodology is applied to identify the “best” solvent. Firstly, we evaluate the performance of all the solvents in the database for the specific task. Secondly, from this evaluation, we can select a pool of high-performing solvents based on a specific criterion, which in most cases will be the yield. Thirdly, we evaluate the operational implications of each of the solvents in the pool of selected solvents, and finally, we select the solvents for detailed evaluation, which may also include lab trials.

During this multi-tiered approach, the key questions that be will be used to evaluate the solvents are:

- a. Is the API fully soluble at the initial temperature T_i ? It is desired that at T_i the feed solution should be unsaturated to prevent uncontrolled crystal formation. This initial temperature is dependent on either:
 - i. The physical properties of the solvent or API (e.g., boiling point or degradation temperatures, etc.) as outlined in Section 4.4 i.e.

$$T_i = (T_b \text{ or } T_{flash} \text{ or } T_{deg} \text{ or } T_{vap}) - 10\text{ }^{\circ}\text{C}$$
 - ii. The upstream processing conditions (e.g., exit reaction temperatures, etc.)
- b. Is the required yield of pure API crystals achievable within the operating temperature range, where the final operating temperature as outlined in Section 4.4, is also dependent on either:

$$\text{if } T_{msol} \text{ or } T_{mimp} \geq T_{cw} \text{ then } T_f = (T_{msol} \text{ or } T_{mimp}) + 10\text{ }^{\circ}\text{C} \quad \text{or} \quad T_f = T_{cw} + 10\text{ }^{\circ}\text{C}?$$

- c. How much of the identified solvents will be required to ensure that the targeted production of API is achieved? Sufficient solvent is required to ensure that the API is completely dissolved at the start of the crystallisation process and to prevent uncontrolled/unfavourable crystallisation.
- d. What is the operating cost associated with a selected solvent to achieve the required production rate? The following operating cost will be determined for each of the solvents identified:
 - i. The cost of solvent required;
 - ii. The cost of cooling / or evaporation required;
 - iii. The cost of solvent recovery;
 - iv. The cost of treating the waste generated with the identified solvent; and
 - v. The loss associated with the unrecovered API.
- e. What is the estimated capital expenditure required for a selected solvent to achieve the required production rate? The following capital expenditure is considered:
 - i. the number of 1.5 m³ batch crystallisers are required to meet the production rate and the cost associated for the number of crystallisers required;
 - ii. the total heat transfer area required to achieve the cooling duty required to meet the production rate;
 - iii. The cost of the solvent storage vessel, and the cost of the waste solvent storage vessel; and
 - iv. The cost of solvent recovery process.
- f. What is the environmental impact of the selected solvent? The Green chemistry E-factor metric is used because of its flexibility and simplicity. In addition, the Energy Consumption Factor is also used as it is a good indicator of the carbon footprint associated with the selected solvent.

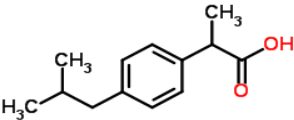
In summary, the computational algorithm must output the following information for decision making: the yield of pure API crystals; the operating costs; the capital costs; and the environmental impact associated with the resultant pool of solvents. A ranking methodology that takes into account the various performance criteria is applied to rank the potential solutions.

6.2.2.1.1: Application Description

For a given unsaturated/dilute feed, with a specified mole ratio of solvent to API and a targeted production rate, we want to identify a selection of solvents that can be used in the manufacture of pure API crystals using cooling crystallisation. The selected solvents must ensure that the feed is highly unsaturated at its entering conditions to prevent unfavourable crystallisation (as outlined in chapter 3) and then must use a controlled rate of cooling to achieve a specified, desired production rate.

The API considered is 2-(4-Isobutylphenyl) propanoic acid (Ibuprofen). It is an optically active compound with both *S* and *R*-isomers, of which the *S* (dextrorotatory) isomer is the more biologically active. Whilst Ibuprofen has two isomers, in this application it will be recovered as a racemate. Its structure and key physical properties are presented in Table 6.5.

Table 6.5. Structure and Properties of 2-(4-Isobutylphenyl) Propanoic Acid (Ibuprofen).

	Property	Structure
Name of API	2-(4-Isobutylphenyl)propionic acid	
Common Name	Ibuprofen	
Formula	C ₁₃ H ₁₈ O ₂	
Molar Mass	206.29 g/mol	
Density	1.03 g/ml	
Melting Point	75 to 78 °C	
Boiling Point	157 °C	
Heat of Fusion	128.9 ± 5.8 kJ/mol	

6.2.2.1.2: Process simulation

To determine the solvents that can be used in the recovery of the API, we determine the initial operating temperature, T_i , as discussed in section 4.4. The maximum operating temperature should be the lower of the following temperature limits: 10 degree safety margin below the boiling point of the solvent selected or 20 degree safety margin below the boiling point/degradation temperature of the API. It should be noted that the initial temperature of the crystallisation process may be dictated by the operating conditions of the upstream process. For example, let us say that water is the preferred solvent, since the boiling point of the solvent is much lower than the boiling point of the API, the maximum operating temperature will be 90 °C. However, if the exit temperature of the upstream operation is 60 °C, then the maximum initial temperature will be 60 °C, unless preheating is desired for other reasons (eg., lowering viscosity, etc.).

The final operating temperature, T_f , is dependent on the degree of cooling that can be achieved, and this is dependent on the utilities that are available in the plant. If cooling is to be achieved with a refrigerated cooling circuit, then subzero cooling temperatures can be achieved. The lower the cooling temperatures used, the higher will be the cooling costs per batch of API produced, however this may be offset by the higher yields and lower waste streams generated.

To illustrate the capacity of the developed computational tool, let's consider a feed-stream with a mole ratio of solvent to API of 5:2, and a constrained cooling capacity with the minimum attainable cooling temperature of 10 °C. The desired target production rate is 1000 kg of pure API crystals per day.

Using the solubility calculation facility on the anti-solvent menu, we can calculate the performance of all the solvents in the database. The algorithm will automatically exclude those solvents that have a melting point above the specified attainable cooling temperature, and will assume that the initial temperature of the solvent is $T_i = (T_b \text{ or } T_{flash} \text{ or } T_{deg} \text{ or } T_{vap}) - 10^\circ\text{C}$. The performance of the various solvents in the database at these conditions is shown in Figure 6.11. The list of solvents shown in Figure 6.11 excludes those solvents with a melting point of 10°C or higher. The performance of the solvents in the database varies from 0 % to 100 % yield.

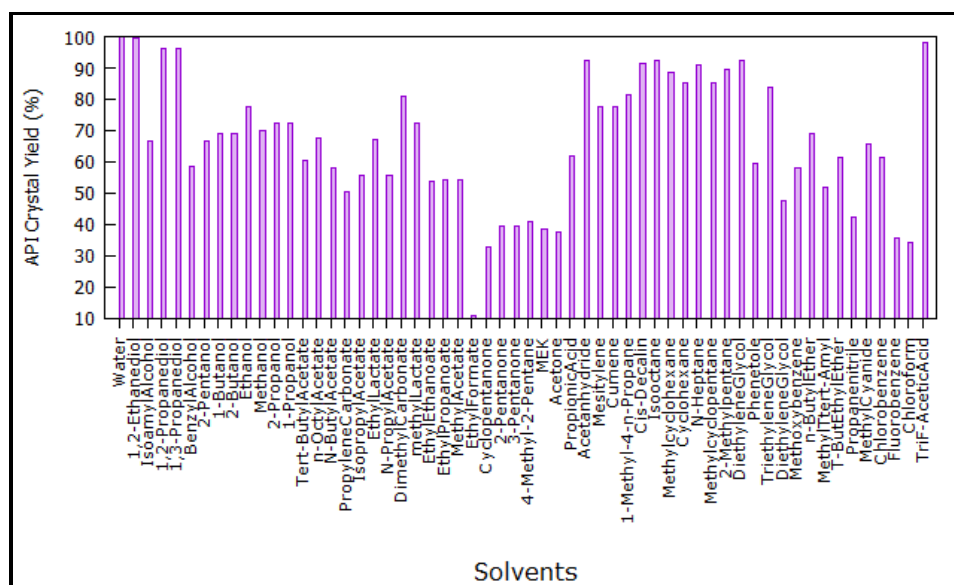


Figure 6.11. Solvent Performance When Cooled Down from NBP of Solvent to $T_f = 10^\circ\text{C}$.

We now further investigate the process engineering implications of the various solvents. Once the desired production rate is established, we can determine the effect each solvent will have on the size of plant required, the capital expenditure and the operational expenses associated with the selected solvent. As detailed in chapter 4, the plant operations that are considered to be directly affected by the choice of solvent are limited to the following: the crystallisers and their required heat exchangers; the solvent feed and waste storage tanks; and the solvent recovery system. Hence, the calculated capital costs are determined for these operations.

The operational costs are determined for: crystalliser cooling; solvent recovery heating; solvent cost; loss of unrecovered API; and waste treatment of unrecovered streams. The financial models exclude the following operations: the centrifuge; the washing station; any recrystallisation operations and the dryer. The results of the inclusion of the economic assessment for the various solvents in the database is shown in Figure 6.12.

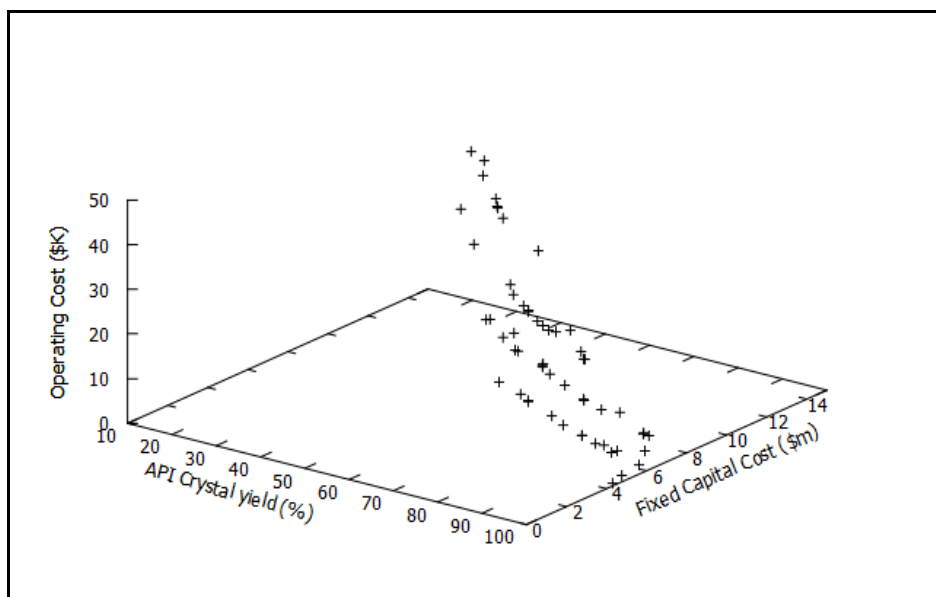


Figure 6.12. Operational and Economic Performance of the Solvents cooled to $T_f = 10\text{ }^{\circ}\text{C}$.

The general trend observed is that the fixed capital costs, as well as the operating costs, are dependent on the yield achieved in the process. Low yields require higher processing volumes to achieve the required production target, leading to higher capital and operating costs. The decreasing trend of fixed capital costs with increasing percentage yield indicates the decrease in size of the processing equipment required. Similarly, the decrease in operating costs with increasing percentage yield indicates the decrease in usage of solvent, waste-treatment, solvent recovery costs etc.

From the initial assessment, we can further expand the search to solvents that achieve a minimum specified yield. For a minimum desired yield of 85 %, the computational tool identifies at least 14 solvents that will fulfil this criterion when the feed is cooled from T_i to a final temperature of $10\text{ }^{\circ}\text{C}$. The resulting pool of potential solvents that can achieve a yield of 85 % pure API crystals for the initial feed ratio of 5:2 and cooling temperature of $10\text{ }^{\circ}\text{C}$ is shown in Figure 6.13.

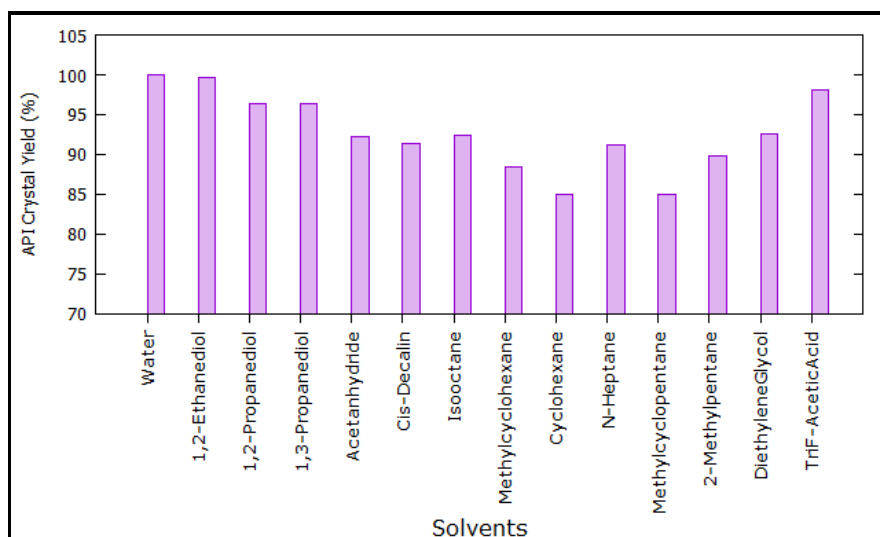


Figure 6.13. List of Solvents That Will Achieve at Least 85 % Yield of Pure API Crystals at $T_f = 10^\circ\text{C}$.

This pool of solvents can be further increased if the degree of cooling is increased. For example, if the cooling capacity of the plant is such that a final cooling temperature of -10°C is achievable, then the pool of solvents meeting the minimum desired yield of 85 % increases to 28 solvents. This automatically excludes those solvents which also solidify (crystallise) at temperatures greater than -10°C , for example water will crystallize at 0°C and so will be automatically excluded for this new operating condition.

If the mixture is cooled to 0°C , then the pool of potential solvents will be 17 solvents. Likewise, a change in feed composition will also impact on the number of potential solvents. For example, if the feed-stream has a solvent to API mole ratio of 10:2 (compared to 5:2 in the initial evaluation) and a final cooling temperature of 10°C , then the resulting pool of potential solvents is 6 solvents (compared to 14 for the 5:2 feed ratio).

The third variable that will affect the number and type of solvent in the selected pool will be the initial operating temperature. If T_i is dependent on upstream processing conditions, the resulting pool of solvents will exclude those solvents that experience supersaturation at temperatures of T_i and above. For example, if T_i is limited by the up-stream processing conditions to 40°C , then the resulting pool of potential solvents is 2 solvents compared to 14 for the 5:2 feed ratio and the final cooling temperature of 10°C .

These effects, of feed composition, feed temperature and degree of cooling available in the plant, on solvent selection is illustrated in Table 6.6. It demonstrates the computational capability of the Solvent Selection Tool, which will identify solvents for varying feed and process conditions that meet the specified criteria, for example, a minimum yield criteria of 85 %. This pool of solvents also excludes

those solvents which have been identified for legislative exclusion, or based on EHS criteria, as discussed in chapter 4.

Table 6.6. Number of Solvents that will Achieve 85 % API Crystal Yield Based on Feed Conditions and Available Cooling Capacity.

Feed Conditions		No of Solvents achieving 85 %		
Feed Temperature	Feed Composition	Yield For Plant Cooling Capacity		
	Solute/Solvent Ratio	–10 °C	0 °C	10 °C
<i>Solvent or API property limitation.</i>				
For example if the temperature is	2/5	28	17	14
based on solvent boiling point.				
$T_i = (T_{solventbp}) - 10\text{ °C}$	1/5	15	11	6
<i>Upstream Process limitation. For</i>				
example , $T_i = 40\text{ °C}$	2/5	21	5	2
	1/5	8	5	1

From the analysis provided in Table 6.6, it was found that the pool of potential solvents is dependent on three main variables: the feed temperature, T_i , the feed composition, and the final cooling temperature, T_f . The pool of potential solvents is reduced by a lower starting temperature, T_i . Only 2, or 5, or 21 solvents met the 85 % yield criteria when cooled from the initial feed temperature, $T_i = 40\text{ °C}$, to final crystallisation temperatures of $T_f = 10\text{ °C}$ or 0 °C or -10 °C , respectively.

Whilst the yield may be the main criterion for solvent selection, the influence of other key performance criteria, such as environmental impact, operating cost and capital cost of the shortlisted pool of solvents, must also be considered. These values are obtained for the following feed and operating conditions:

- Target production rate of 1000 kg per day of API crystals;
- Feed composition solute to solvent mole ratio of 2:5;
- Feed temperature is $T_i = (T_b \text{ or } T_{flash} \text{ or } T_{deg} \text{ or } T_{vap}) - 10\text{ °C}$;
- Final cooling temperature is 10 °C ;
- It is assumed that 80 % of the solvent is recovered and recycled and the cost of the replacement / make-up solvent is averaged at \$1000 per ton; and
- The unrecovered API is costed at \$20 per kg (average price of ibuprofen).

As has been observed earlier, with yield-dependent processes the plant size and utility requirements vary with yield. Using the additional, relevant performance criteria, such as operating costs, fixed annualised costs, environmental factors and energy consumption factors, the overall performance of a

selected solvent can be evaluated. The results obtained from the Solvent Selection Tool on the economic and environmental criteria for this pool of solvents are shown in Figure 6.14 to Figure 6.18.

The fixed annualised cost (FAC), which is a combination of the operating costs and a fraction of the fixed capital costs, is shown in Figure 6.14 for the 14 potential solvents, assuming 200 production days.

The operating cost is limited to: the cost of cooling from temperature T_i to temperature T_f ; the cost of recovering 80 % of the solvent for reuse; the cost of replacing unrecovered solvent that is sent to waste; the cost of tertiary treatment of the waste stream; and the cost of unrecovered product.

The fixed capital cost is limited to the equipment directly related to the solvent use; the size or number of crystallisers and their associated heat transfer area; solvent storage tanks; waste storage tanks, and the solvent recovery system. Figure 6.15 presents the operational cost per 1000kg of product for the potential solvents, and Figure 6.16 presents the fixed capital cost associated with each of the potential solvents.

The Environmental (E) Factor for the potential solvents, which is a measure of the amount of waste generated per kilogram of API crystallised is shown on Figure 6.17. It should be noted that the E Factor calculated here is based only the crystallisation process and excludes waste streams generated in the other associated operations such as the reactors, dryers etc. In addition, the calculations used in this work do not distinguish between organic and inorganic waste and the hazardous nature of the substance. Hence, if the hazardous nature of the solvents is included in the computation, then this may result in higher E-Factors for all the solvents in the database except for water.

The energy consumption (E_c) factor, which is a measure of the total energy required per kilogram of product crystallised for the potential solvents, is presented in Figure 6.18. In the application of cooling crystallisation, the E_c Factor accounts for the following: the heat load of the crystalliser to cool the content from the initial temperature to the final temperature of 10 °C, and the heat load required to recover 80 % of the solvent in the solvent recovery process. This “total energy” consumed per kilogram of API produced can be directly correlated to carbon footprint contribution of the selected solvent.

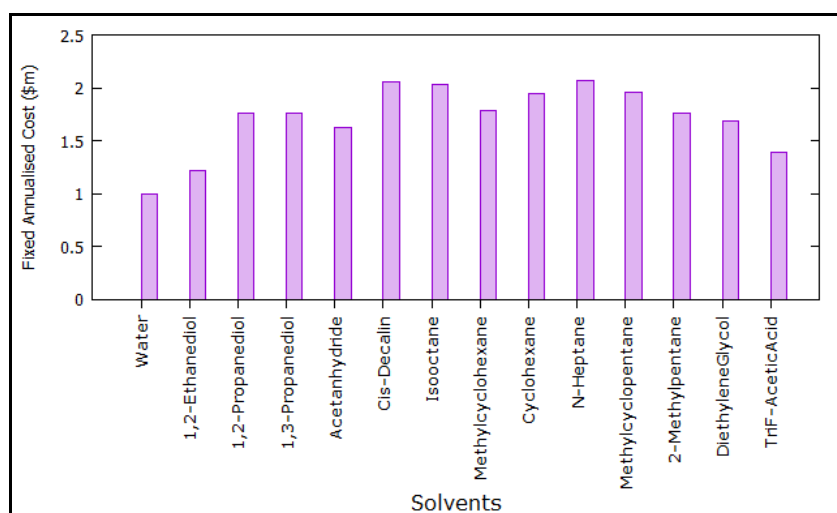


Figure 6.14. Fixed Annualised Cost of a 1000kg API/Day Production Plant With $T_f = 10^\circ\text{C}$.

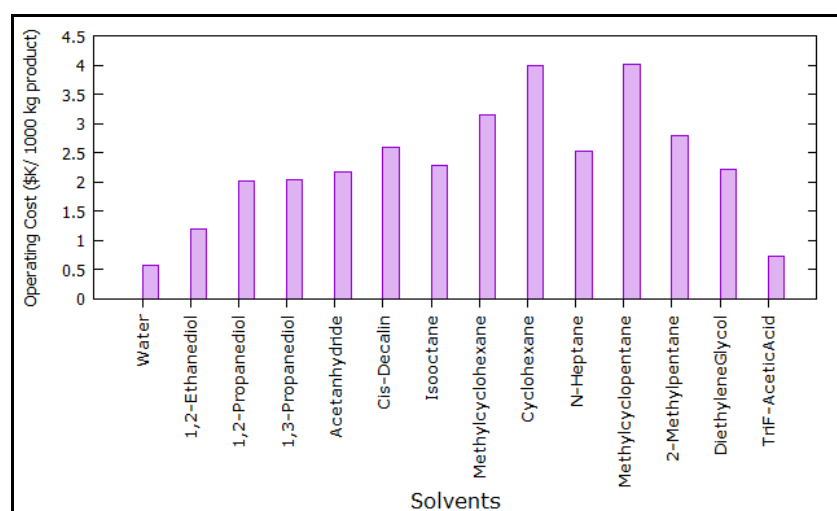


Figure 6.15. The Operational Cost of 1000kg API/Day Production Using Suggested Solvent for $T_f = 10^\circ\text{C}$.

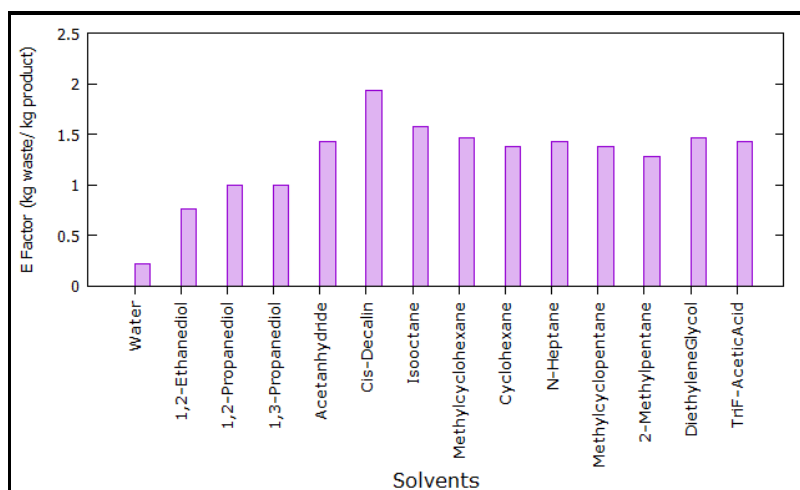


Figure 6.16. The Environmental (E) Factor of Suggested Pool of Solvents for $T_f = 10^\circ\text{C}$.

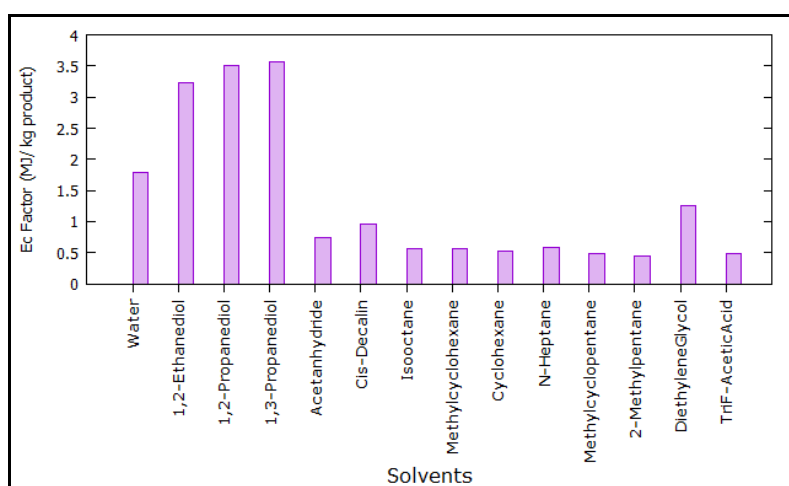


Figure 6.17. The Energy Consumption (Ec) Factor of Suggested Pool of Solvents for $T_f = 10^\circ\text{C}$.

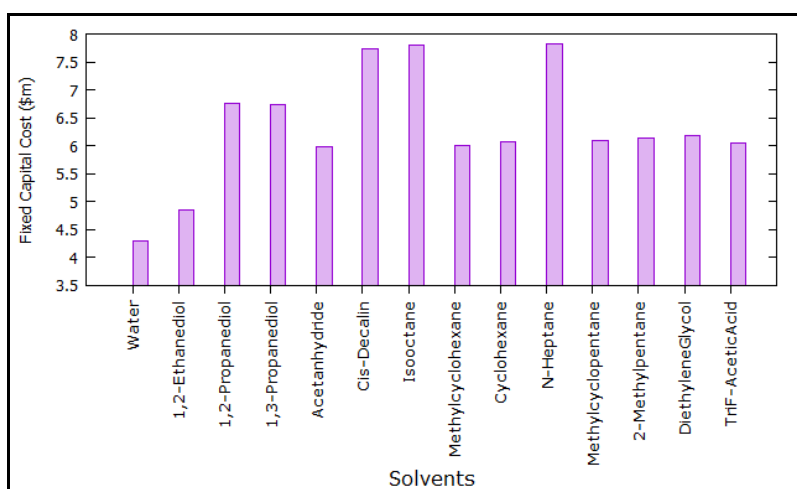


Figure 6.18. The Fixed Capital Cost of 200 Ton/Yr Production Facility Associated with Selected Solvent for $T_f = 10^\circ\text{C}$.

By applying user-defined weighting factors to various criteria, a stacked histogram can be developed to evaluate the cumulative effect of the selected criteria for each of the shortlisted solvents. For example, by applying equal weighting to the yield; FAC; E-Factor, and Ec-Factor, the potential solvents can be ranked as shown in Figure 6.19. The solvent with the lowest cumulative value will represent the best solvent. In this application, water is calculated to be the best solvent. It is important to note the feed temperature is assumed to be 90 °C (i.e. 10 °C below the boiling point of the selected solvent). This temperature is above the melting point of the API (78 °C), hence the API is completely soluble at the initial temperature, hence making water a high-ranking solvent.

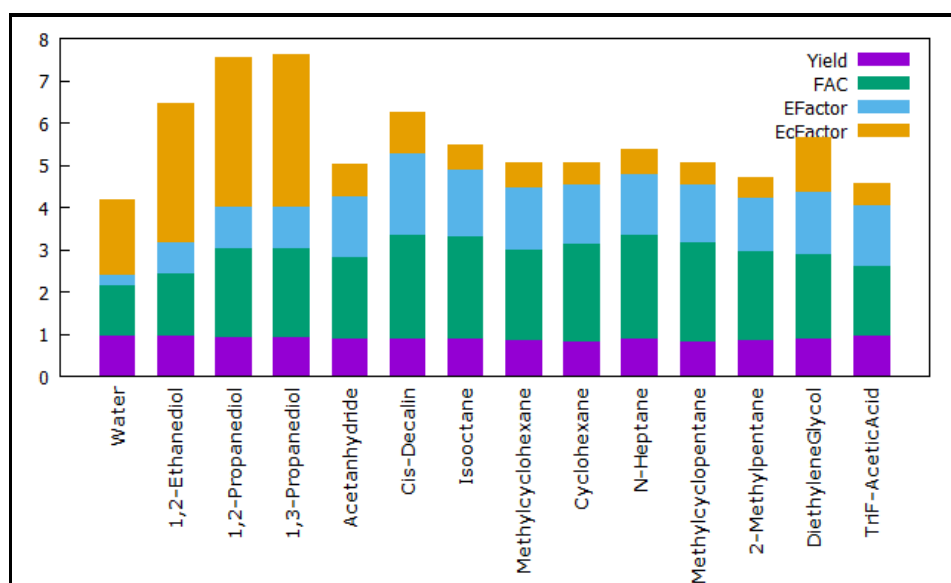


Figure 6.19. Ranking of Solvents based on equal weightings of yield, FAC, E-Factor and Ec-Factor for $T_f = 10\text{ }^{\circ}\text{C}$.

A more detailed assessment of the selected solvent is required. In chapter 2, the various desired operating conditions for crystallisation were outlined. The key aspects include: controlled cooling and operating within the “meta stable zone width” so as to promote crystal growth and avoid conditions that promote primary nucleation. Further evaluation of water as a solvent is conducted by calculating the yield profile from T_i to T_f . The prediction indicates that there is a 99.5 % yield achieved within a one degree drop in temperature at $63.5 \pm 1\text{ }^{\circ}\text{C}$.

Such rapid crystallisation is undesirable in API production as this results in a lack of control to achieve the required crystal morphology and crystal size distribution, and will lead to the inclusion of the solvent in the crystal matrix, giving rise to poor quality and purity. Recommended cooling rates for organic compounds are in the order of 0.1 – 0.2 K/min (Beckmann, 2013).

This second level assessment rules out water as a potential solvent. It shows that the solubility profile (or yield profile) over the possible operational range is very important in the selection or ranking of solvents. Whilst large variation of solubility with temperature is desirable in cooling crystallisation, systems that show large solubility variation with small temperature differences are undesirable, as it will result in a rapid increase in supersaturation with cooling, and will promote primary nucleation at the expense of crystal growth. A rapid increase in supersaturation may also result in the inclusion of mother liquor within the crystal structure, and hence contamination of the API crystals. Some typical profiles of variation in yield with cooling temperature are illustrated for the 14 short-listed solvents in Figure 6.20.

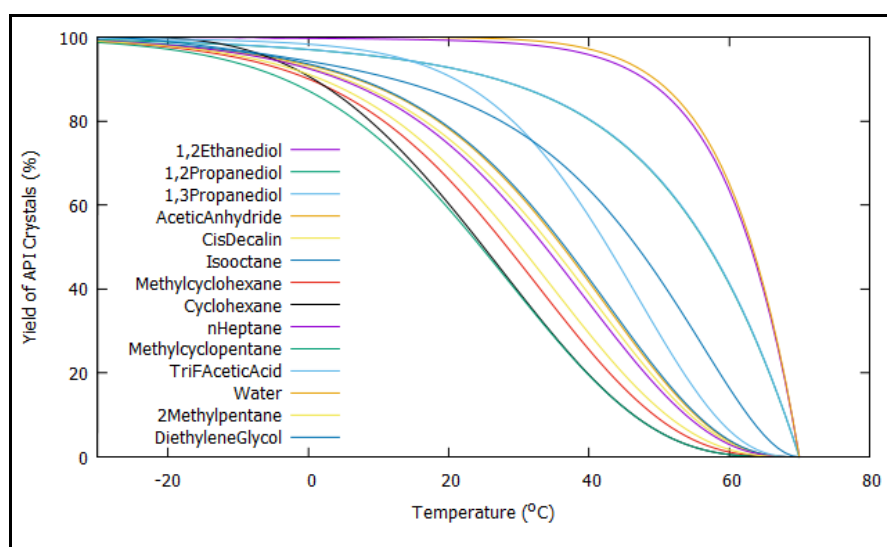


Figure 6.20. Yield – Cooling Temperature Profiles of the Shortlisted Solvents.

This information allows the user to quickly identify the potential operating temperature range that will fulfil production requirements, or, whether a particular user-selected solvent will result in the desired production and quality for a specified operating condition. This type of approach will eventually identify the most promising solvents that can be confirmed by means of experimental trials.

Since the operating temperatures affect not only the yield but also the operational and capital costs, the effect of the degree of cooling on a production plant should be evaluated. As an illustrative example, with the selection of n-heptane from the shortlist of solvents, the impact of varying degrees of cooling on the process, economic and environmental criteria is shown in Figure 6.21 to Figure 6.28. The data presented is obtained from the Solvent Selection Tool, where the major contributors to the operational costs are identified. Each graph also has a yield curve included to show how the various cost contributors vary with yield and cooling temperature.

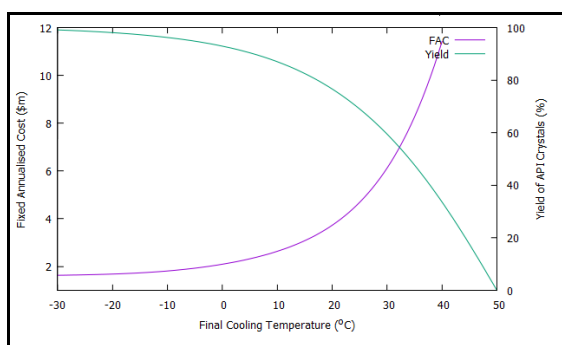


Figure 6.21. Effect of Cooling Capacity on the Fixed Annualised Cost and Yield.

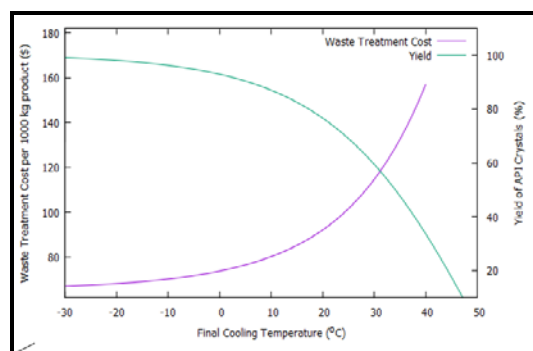


Figure 6.22. Effect of Yield on Waste Treatment Cost.

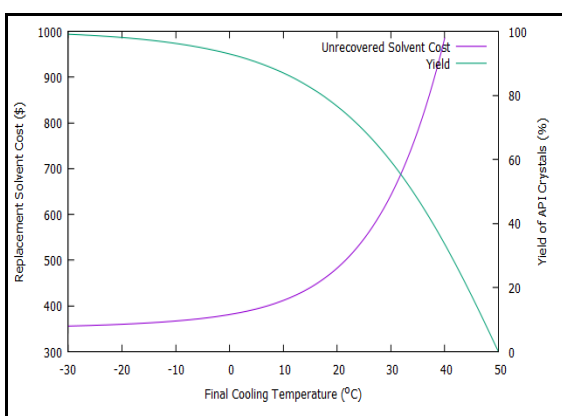


Figure 6.23. Effect of Yield on Solvent Replacement Cost.

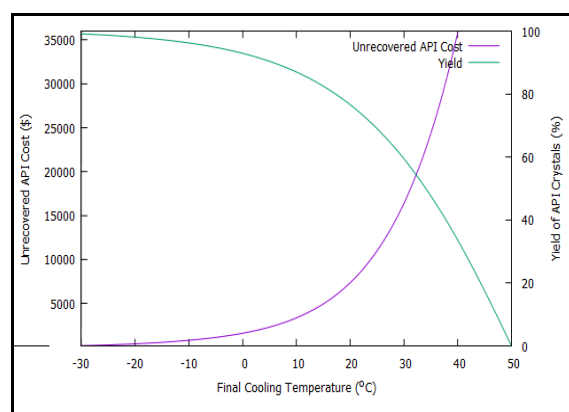


Figure 6.24. Effect of Yield on Unrecovered API.

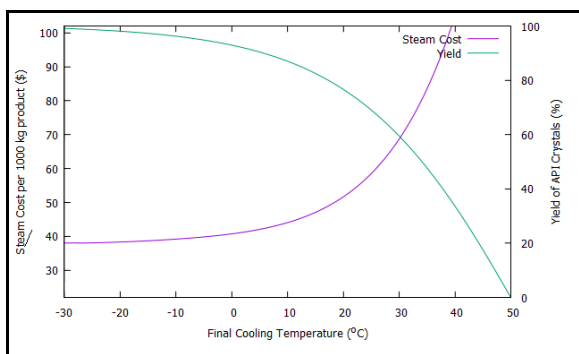


Figure 6.25. Effect of Yield on Solvent Recovery Steam Cost.

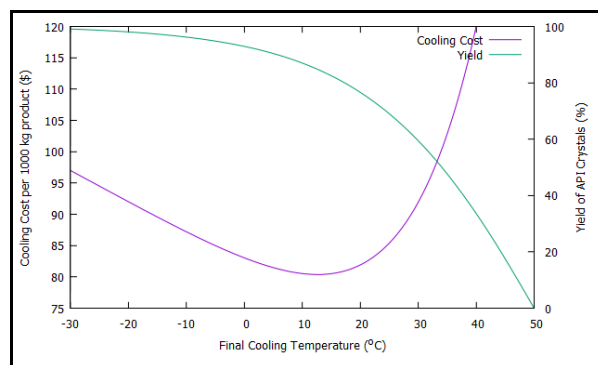


Figure 6.26. Effect of Yield on Cooling Cost.

The general observation is that lower final cooling temperatures promote higher yields which lead to lower capital and operation costs as well as lower environmental impacts. Whereas, lower yields lead to rapid increases in capital and operation costs as well as in a much higher impact on the environment. The yield obtained is dependent on the plant cooling capacity available and the final temperature to be achieved in the crystallisation process. The two major contributors to the operational costs are the solvent replacement cost and the cost associated with unrecovered API lost as waste.

The exponential increase in all the cost contributors with decreasing yield is primarily due to the larger processing volumes required to meet the production target of 1000 kg of API crystals per batch. It is evident that the size of a production facility for a required production rate will be dependent on the yields of the critical operations in the process. Low yields will require larger volumes of materials to be processed to ensure that the desired production rate is obtained. These larger processing volumes of materials will require larger equipment or a greater number of modular units, and greater process utility requirements such as cooling, heating and waste treatment.

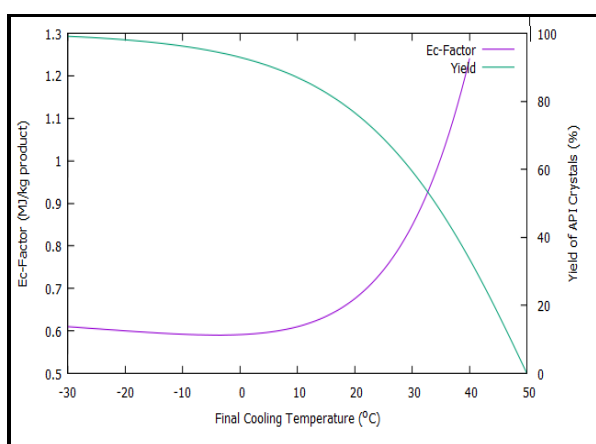


Figure 6.27. Effect of Yield on E_c Factor.

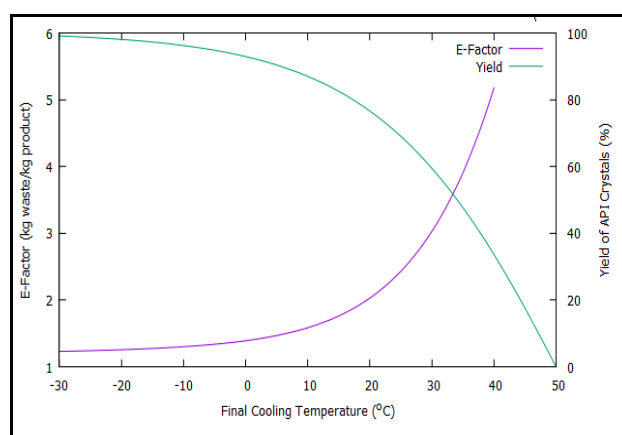


Figure 6.28. Effect of Yield on E Factor.

The objective of this case study was to identify and select potential solvents that can be used for the crystallisation of a newly developed API. With various computational options integrated into the Solvent Selection Tool, the potential solvents and operating conditions that meet the process objectives of high crystallisation yields can be easily identified. In addition, the operational costs, capital cost and the environmental impact is computed for each solvent in the pool of potential solvents, which facilitates decision making during the conceptual design stage.

6.2.2.2 Application 2: Evaluation of Cooling, Evaporative and Anti-solvent Crystallisation for a given application.

From the previous case study it is evident that the major contributors to the operational costs and the fixed annualized costs are the cost contributions associated with the loss of solvent and API in the waste stream. Lower crystallisation yields require larger feed flowrates (processing volumes) to achieve a required API production target. This leads to larger processing equipment (or increase in number of modular units), greater utility requirements and greater waste disposal, which lead to higher operational and fixed annualized costs. Hence, strategies to increase the recovery of API during the crystallisation operation are critical.

In this second case study we explore both how to obtain process-attainable maximum yield and the subsequent impacts on economic and environmental criteria.

Problem statement: For a given API and solvent system, identify the operational mode (cooling, evaporation, anti-solvent or combinations), and operating conditions that will give the required yield and production rate. The feed-stream, consisting of a molar ratio of 2 kmols API (2-(4-Isobutylphenyl)-propionic acid (Ibuprofen)) to 5 kmols solvent (ethanol) at $T_i = 50\text{ }^{\circ}\text{C}$. The production target is 1000kg per day with a minimum desired yield of 98 % of pure API crystals.

A generic and multi-tiered methodology is used to identify the best operating conditions. During this multi-tiered approach, the key questions that be will be used to evaluate the solvents are:

- a. Is cooling crystallisation adequate to achieve the required yield and production rate?
- b. Is evaporation crystallisation adequate to achieve the required yield and production rate
- c. Is anti-solvent addition adequate to achieve the required yield and production rate
- d. Is a combination of operations required?
- e. What is the operating cost associated with selected operations? The following operating cost will be determined for each of the operation option:
 - i. The cost of solvents (including anti-solvents) required for crystallisation;
 - ii. The cost of cooling / or evaporation required;
 - iii. The cost of treating the waste generated with the identified solvent/s;
 - iv. The cost of solvent recovery and;
 - v. The cost associated with unrecovered API.
- f. What is the estimated capital expenditure required for a selected operation to achieve the required production rate? The following capital expenditure is considered:
 - vi. the number of 1.25 m³ batch crystallisers required to meet the production rate, and their cost;
 - vii. the total heat transfer area required to achieve the cooling or evaporation duty to meet the production rate;
 - viii. The cost of the solvent storage vessel, and the cost of the waste solvent storage vessel and;
 - ix. The cost of the solvent recovery process.

Note: since the same equipment will be used to evaluate the various modes of crystallisation, the only changes to the capital cost will be associated with anti-solvent crystallisation because of the increased volumes of liquid and added sophistication required for solvent recovery.

Option 1: Cooling Crystallisation

To obtain a quick overview of potential options, the following sets of data were generated using two algorithms from the Solvent Selection Tool: the algorithm for cooling – evaporation studies and the algorithm for cooling – evaporation – anti-solvent studies. From the data generated by means of the computational tool, various graphs were generated to obtain a visualisation of the effects of cooling; evaporation and anti-solvent addition. The following graphs were generated to evaluate the full spectrum of options: effect of cooling on yield, shown in Figure 6.29; effect of evaporation and cooling on yield, shown in Figure 6.30, and effect of anti-solvent and cooling and yield, shown in Figure 6.31.

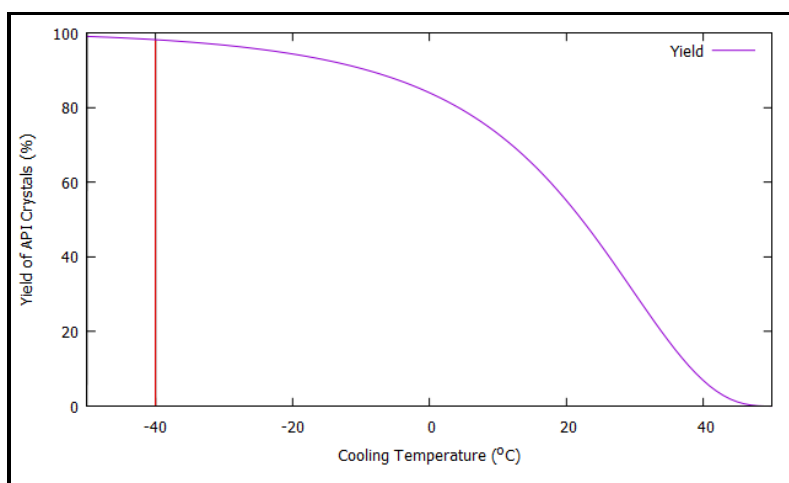


Figure 6.29. Effect of Cooling on Crystallisation Yield.

From Figure 6.29, it can be seen that the required yield of 98 % can be achieved using cooling crystallisation, if the final cooling temperature of -40 °C is achievable in the production facility. However, we may also achieve the required yield with a combination of evaporation and cooling. In Figure 6.30 it can be seen that with an evaporation stage preceding the cooling crystallisation stage, the degree of cooling required is reduced. For example, with 50 % evaporation, a cooling temperature of -14 °C is required to achieve the required 98 % yield. However, we need to be mindful that higher evaporations may lead to higher degrees of supersaturation and more rapid de-supersaturation upon cooling, which will promote rapid primary nucleation, leading to poor crystal growth.

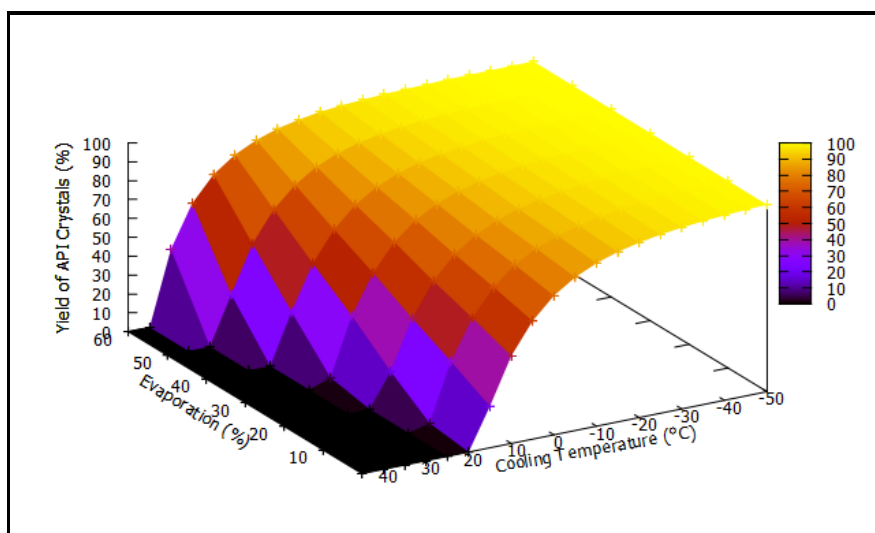


Figure 6.30. Effect of Evaporation and Cooling on Crystal Yield with no Anti-solvent.

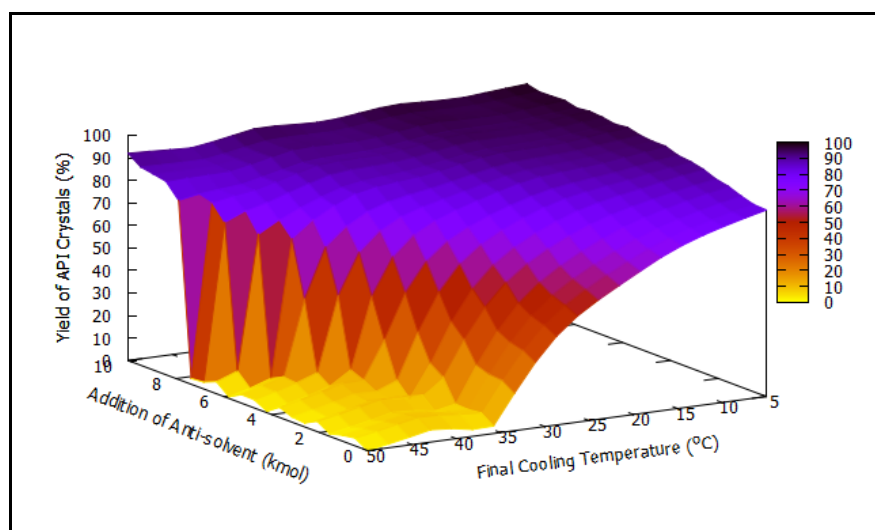


Figure 6.31. Effect of Anti-solvent Addition and Cooling on Crystal Yield with no Evaporation.

To explore the use of anti-solvents, it is first necessary to identify potential anti-solvents that can be used with the ethanol system. The “anti-solvent menus” as shown in Figure 4.4, can perform several calculations to identify potential anti-solvents. Whilst the detailed description of each menu is presented in Appendix B, in Table 6.7, we present aspects of the menu options that are used here to identify a suitable anti-solvent.

Table 6.7. Relevant Calculation Selection in the Anti-solvent Menu for Anti-Solvent Selection.

Menu Option	Computational Capability
Solubility	Calculates the attainable yield of API crystals with each solvent in the database at ambient conditions. It ranks the performance of each solvent classifies the solvents from good solvents to good anti-solvents.
VLE	It will perform the calculations to check whether any azeotropes exist in the solvent and selected anti-solvent.
LLE	It will perform the calculations to check whether any regions of immiscibility exist in the mixture of the solvent and selected anti-solvent.
Addition Curve	Calculates the changes in crystal yield with incremental increases in amount of anti-solvent.

Using the “solubility” subroutine, the Solvent Selection Tool can also be used to identify and rank the solvents from good solvents to poor solvents (good anti-solvents). In Figure 6.32, we illustrate the ranking of solvents from good solvents to good anti-solvents for the feed condition of 50 °C. All solvents with zero yield could be considered as good solvents because in these solvents the solute-solvent mixture at T_i will be under-saturated.

All solvents with high yields at temperature, T_i , could be considered as good anti-solvents because these solute-solvent mixtures result in highly super-saturated solutions. From Figure 6.32, the following solvents could be considered to be good anti-solvents: glycerol, ethylene glycol, water and 3-pentanone. Additional assessments are undertaken to evaluate the compatibility and performance of the shortlisted anti-solvents. These include: a miscibility test, an azeotrope test and a yield test. Using the subroutines called “VLE” and “LLE” we check whether these anti-solvents form azeotropes mixtures or miscibility gaps, respectively, with ethanol. The “Addition Curve” calculation is used to determine whether the required yield can be achieved by the anti-solvent, and the minimum amount of anti-solvent required to achieve the desired yield. Using the various subroutines outlined in Table 6.7, the following summary of results is presented in Table 6.8. and Figure 6.33:

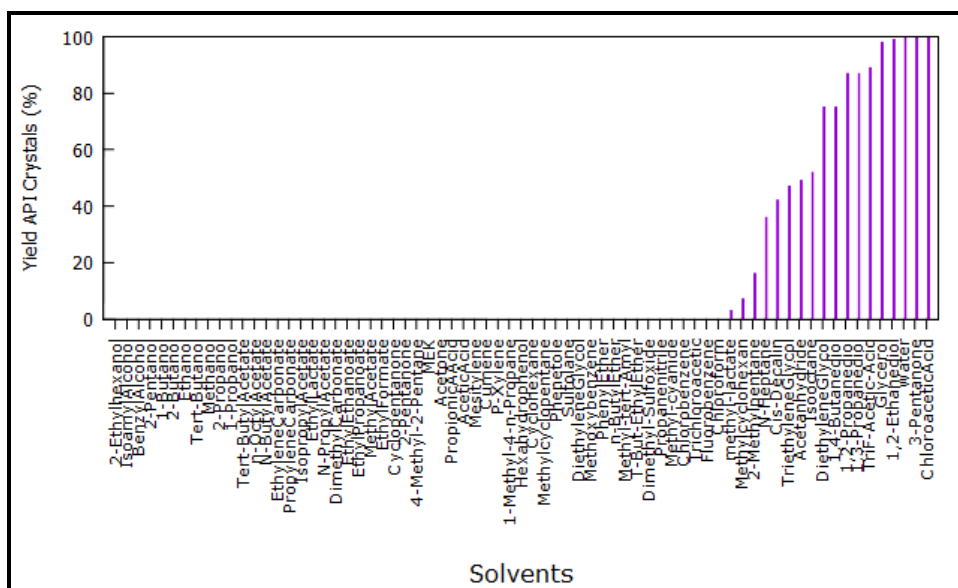


Figure 6.32. Identification of Good Solvents and Anti-Solvents at $T_i = 50\text{ }^{\circ}\text{C}$.

Table 6.8. Summary of Results on the Compatibility and Performance of Shortlisted Anti-Solvents at Ambient Conditions.

Anti-solvent	Forms	Forms	Maximum	Minimum Anti-
	Azeotrope	Immiscible	Attainable	solvent required
	with the solvent?	Regions with the solvent?	Yield? (%)	to achieve 98 % yield (kmol)
Ethylene Glycol	No	no	88.40	N/A
Glycerol	No	no	91.50	N/A
3-Pentanone	No	no	48.30	N/A
Water	Yes	no	99.50	12

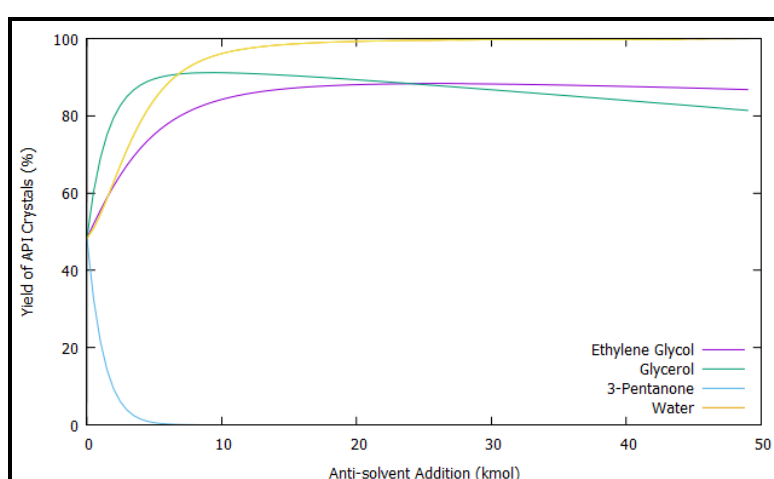


Figure 6.33. Addition Curves for Selected Anti-solvents at Ambient Conditions.

From the list of anti-solvent candidates, water is the only anti-solvent the meets the required yield criterion. Whilst fulfilling the miscibility criteria, water fails the azeotrope test. This implies that a

solvent - anti-solvent system will form a homogenous azeotrope, which will require a more complex solvent recovery system.

We can further explore the option of using water as an anti-solvent. The following ranges of operating conditions listed in Table 6.9 are used to investigate the process, economic and environmental implications of the various combinations of evaporation, cooling and anti-solvent crystallisation, to achieve a minimum yield of 98 % and a production target of 1000 kg pure API crystals per day.

Table 6.9. Range of Cooling, Evaporation and Anti-Solvent Addition Conditions Simulated.

Mode of Crystallisation	Lower limit	Upper limit	Required Yield and Production Rate
Cooling Temperature (°C)	-50	50	Minimum 98 % yield and 1000 kg API crystals per day
Evaporation (%)	0	50	
Anti-solvent Addition (kmol)	0	15	

Figure 6.34 illustrates the effect of anti-solvent addition and cooling on the yield of API crystals. It shows that various combinations of anti—solvent addition and degrees of cooling can readily achieve the required minimum yield of 98 %. The band of operating conditions can be further extended with the inclusion of evaporation as a potential mode.

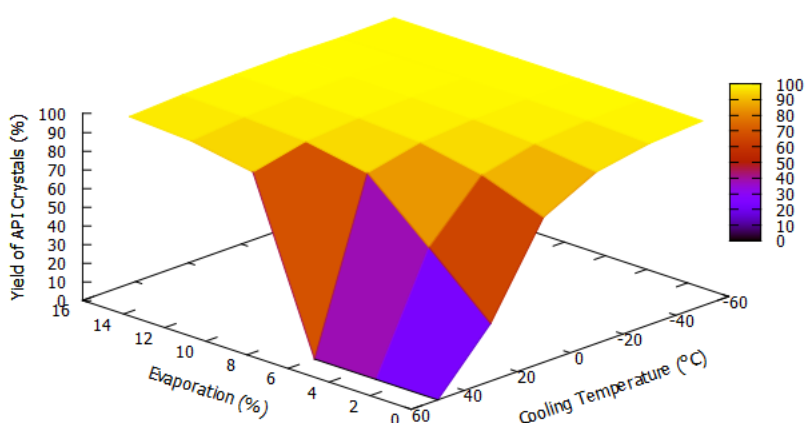


Figure 6.34. The Effect of Anti-Solvent Addition and Cooling on Yield.

Figure 6.35 shows the bands of potential operating conditions that fulfil the criteria of 98 % yield. The number of bands is merely dependent on the number of increments in the evaporation range. If a surface is constructed joining all the lowest points of each band and another surface is drawn joining all the top points of the bands, then the volume enclosed by these two surfaces represent all the possible

combinations of operating conditions that will fulfil the minimum required yield of 98 %. Hence these bands represent “slices” of the whole continuum of the possible operating conditions.

It is important to note that any operating condition outside the enclosed surface will not meet the desired yield or purity. Any point above the upper surface will result in lower yields, whilst any point below the lower surface will lead to contaminated crystals, where the anti-solvent water will also crystallise. This lower surface represents the eutectic surface.

The interpretation of this operational continuum is discussed with the aid of Figure 6.36 This diagram shows the range of operating conditions of cooling temperature and anti-solvent addition, when combined with 10 % evaporation, that will achieve a minimum yield of 98 %. For a given anti-solvent addition amount a vertical band of points exist. The top point in a vertical band indicates the temperature at which a minimum of 98 % yield of pure API crystals will be achieved, whilst the lowest point in the vertical band indicates the minimum cooling temperature allowable prior to contamination by solvent or anti-solvent crystallisation. This temperature is close to the eutectic point and, hence, this lower point represents the maximum yield attainable at this anti-solvent addition rate.

For example, if the conditions of 10 % evaporation and the addition of 8 kmols of anti-solvent are applied, then the operating cooling temperature of the crystallizer, to achieve 98 % yield of pure API crystals, is 15 °C, and the lowest permissible temperature is -15 °C, at which the yield obtained will be 99.6 %. Any further cooling will result in the crystal formation of the anti-solvent (water) resulting in a binary mixture of crystals and hence contamination of the API crystals.

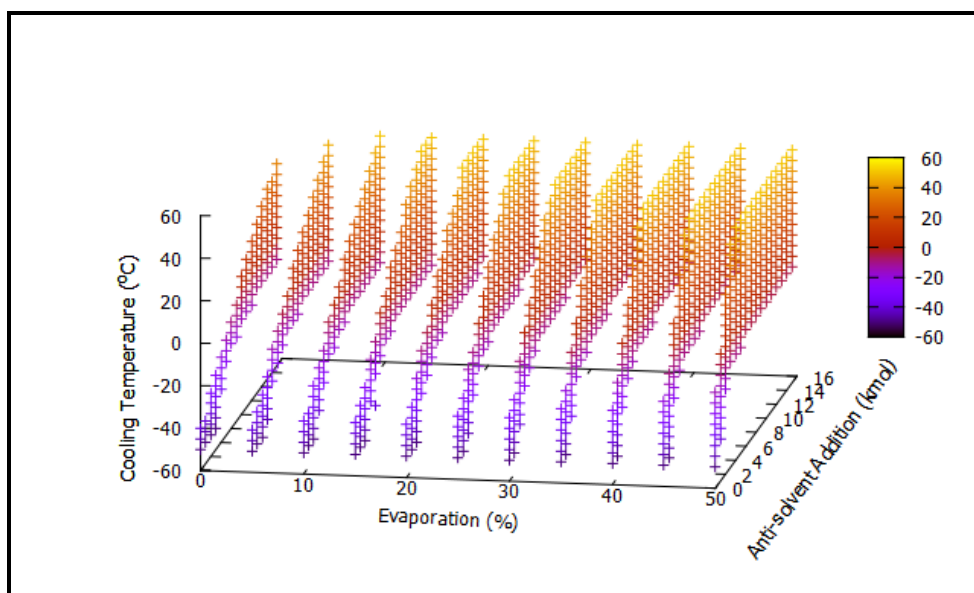


Figure 6.35. Potential Operating Conditions to Achieve a Minimum Yield of 98 % Pure API Crystals.

This ability of the Solvent Selection Tool to identify and exclude conditions that lead to eutectic mixtures is very useful when dealing with multi-component systems. The addition of co-solvents, or anti-solvents, or other by-products leads to an increase in the number of eutectic compositions and separation boundaries. As illustrated in this application, the addition of water, as an anti-solvent, limits the operational conditions when compared to a pure binary system. Conditions that lead to the crystallisation of the solvent, or anti-solvent, must be avoided to ensure the productions of pure API crystals.

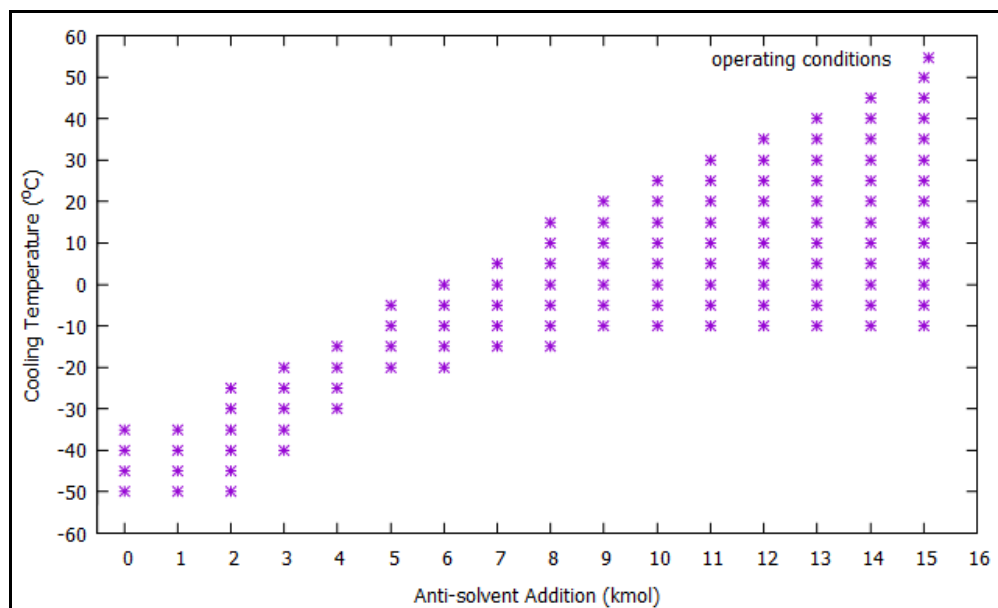


Figure 6.36. Band of Operating Conditions for the Combination of Cooling and Anti-solvent Addition to Achieve a Minimum Yield of 98 % Pure API Crystals with 10 % Evaporation.

An interesting phenomenon, illustrated by this eutectic point, is the depression of melting point of the anti-solvent. Whilst the melting point of water (anti-solvent) is 0,°C, in the given mixture of solvent (ethanol) and API, the temperature at which water will crystallise is lowered to -15 °C..

There is no limitation to the range of evaporation or cooling or anti-solvent addition conditions to be evaluated because the tool is designed to identify the operational conditions that will produce pure API crystals. Within this large pool of possible combinations of operating conditions that fulfil the yield and quality criteria, further criteria are required to screen and identify the best combination of crystallisation operating modes and conditions.

The two additional screening criteria are the economic criterion and the environmental criterion. For the economic criterion, use can be made of either the operating cost per batch of API produced, or use can be made of the fixed annualised cost, which is a combination of the annual operating cost and a ratio of the capital costs. For the environmental criterion, use can be made of the calculated environmental factor, or the energy consumption factor, or a combination of the two.

In addition to the criteria that can be used to rank the operating conditions, the final selection of will also be dependent on process and plant limitations. Such limitations could include: lowest attainable cooling temperature due to limitations of the refrigeration units, thermal degradation of the API at evaporation temperatures excluding evaporation as a potential mode, unless evaporation under vacuum is possible, etc.

In Table 6.10 some of the results that highlight the potential operating conditions, with or without process and plant constraints, are provided. These conditions are limited to the range of the four main variables used for this simulation: the required yield is specified at a minimum of 98 % pure crystals; the daily production target is 1000 kg of pure API crystals; the temperature range is -50 °C to 50 °C with anti-solvent addition of between 0 to 15 kmol; and the evaporation range is from 0 % to 50 %. In addition, the ranking of the operating conditions is based on an equal weighting of the performance criteria of yield, operating cost per batch, the waste generated per kg of API and the energy consumption per kg of API crystallised. The various performance criteria of operational cost, E-Factor and E_c-Factor are normalized by dividing the actual value obtained at specified conditions, divided by the lowest value obtained in the entire range evaluated. The normalised yield criteria is obtained by dividing the actual value obtained at specified conditions, divided by the minimum required yield. Since the proposed ranking method is based on the lowest cumulative value, the inverse of normalized yield value is used:

$$NCWS = \sum_i^n w_i NPC_i = w_y \times (Normalised Yield)^{-1} + w_{op} * Normalised Operating Cost + w_e \times E - Normalised E - Factor + w_{Ec} \times Normalised Ec - Factor$$

Where

$$w_i \text{ is the weighting and } NPC_i = \frac{PC_i \text{ for an operating condition}}{\text{lowest value for this PC}}$$

For the conditions described, and with no process and plant constraints, the best options for crystallisation is evaporation followed by cooling. From the samples of operating conditions shown in Table 6.10, this combination of evaporation and cooling is even more cost effective than just cooling crystallisation. The primary reason for this is that with evaporation, the processing volumes are reduced, which result in reduced downstream equipment sizes, reduced crystallisation cooling heat load, reduced solvent recovery heat load, and reduced waste treatment streams. Also evident from the sample of results presented in Table 6.10 and Figure 6.37, is the higher operating costs associated with anti-solvent addition. However, a combination of evaporation with anti-solvent addition will reduce the operating costs as shown in the last set of results in the table.

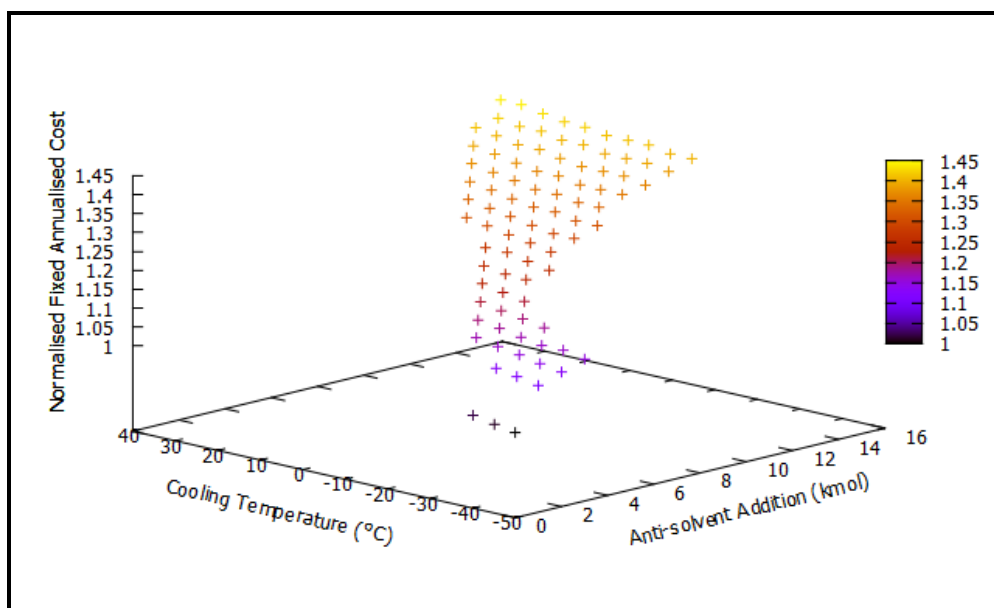


Figure 6.37. Effect of Cooling and Anti-Solvent Addition on the Fixed Annualised Cost.

Table 6.10. Examples of Operating Conditions for Operations With and Without Constraints.

Operating Conditions (OC)			Performance Criteria (PC)				Ranking (R)
Evaporation	Anti-solvent Addition	Cooling Temperature	Yield obtained	Operating Cost per Batch	E Factor	Ec Factor	Normalised cumulative weighted Score (NCWS)
%	(kmol)	(°C)	(%)	(\$)	$\frac{kg \text{ waste}}{kg \text{ API}}$	$\frac{MJ}{kg \text{ API}}$	
Without process or plant limitations:							
50.00	0.00	-50.00	99.72	379.59	0.11	5.05	0.996
45.00	0.00	-50.00	99.65	388.66	0.12	5.14	1.008
50.00	0.00	-45.00	99.62	399.60	0.12	5.05	1.012
40.00	0.00	-50.00	99.59	397.06	0.12	5.23	1.019
45.00	0.00	-45.00	99.53	413.68	0.12	5.14	1.028
35.00	0.00	-50.00	99.52	406.23	0.12	5.32	1.032
50.00	0.00	-40.00	99.50	425.37	0.12	5.05	1.032
With Cooling Only : no evaporation and no anti-solvent addition							
0.00	0.00	-50.00	99.05	472.80	0.12	5.95	1.119
With process and plant limitations eg. No evaporation due to API thermal degradation and lowest cooling temperature attainable in the plant is 0 °C							
0.00	8.00	0.00	98.39	1194.93	0.20	12.34	1.810
0.00	8.00	5.00	98.01	1276.43	0.21	12.38	1.865
0.00	9.00	0.00	98.71	1180.75	0.21	12.96	1.837
0.00	9.00	5.00	98.41	1245.00	0.21	12.98	1.880
0.00	9.00	10.00	98.05	1323.09	0.21	13.02	1.933
0.00	10.00	0.00	98.95	1183.69	0.21	13.57	1.876
0.00	10.00	5.00	98.71	1235.42	0.21	13.60	1.911
0.00	10.00	10.00	98.42	1297.91	0.22	13.62	1.953
0.00	11.00	0.00	99.13	1198.74	0.22	14.20	1.923
0.00	11.00	5.00	98.93	1241.15	0.22	14.21	1.952
With plant limitation: for example lowest attainable cooling temperature = 0 °C							
50.00	2.00	0.00	99.37	792.59	0.14	8.23	1.473
50.00	2.00	5.00	99.22	824.00	0.14	8.24	1.498
50.00	3.00	0.00	99.71	777.32	0.14	8.85	1.503
50.00	3.00	5.00	99.64	791.37	0.14	8.85	1.514
50.00	3.00	10.00	99.56	808.28	0.14	8.85	1.528
50.00	2.00	10.00	99.04	862.74	0.14	8.25	1.529
45.00	3.00	0.00	99.44	819.84	0.14	8.89	1.540
50.00	3.00	15.00	99.46	828.62	0.14	8.86	1.544

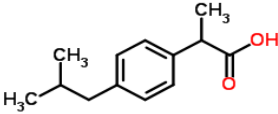
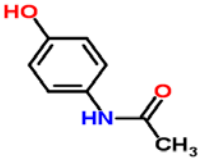
NB. The environmental factor only considers the waste due to the crystallisation process and does not account for waste generated in other processes, e.g., in the reaction process etc.

6.2.2.3 Application 3: Fractional Crystallisation

Multi-component API systems pose additional challenges when the application is further extended to determine modes of evaporation. We have seen that with an increase in the number of components present in the system, there is an increase in the number of eutectic compositions and hence separation boundaries. The ability to analyse multi-component systems and to identify separation limitations are essential requirements for a robust, crystallisation computational tool.

In this third application, we identify the operating options and conditions that will give the required yield and production rate for a system containing by-products/impurities. The system consists of two APIs in solution in a solvent, which need to be recovered as pure API crystals. The binary API systems consist of an unsaturated mixture of 2-(4-Isobutylphenyl) propanoic acid and acetaminophen in ethanol. The properties of the two APIs are presented in Table 6.11.

Table 6.11. Structure and Properties of APIs (NIST).

Property	Component 1	Component 2
Name of API	2-(4-Isobutylphenyl)propionic acid	4-Acetaminophenol
Common Name	Ibuprofen	paracetamol
Formula	$C_{13}H_{18}O_2$	$C_8H_9NO_2$
Molar Mass	206.29 g/mol	151.163 g/mol
Density	1.03 g/cm ³	1.293 g/cm ³
Melting Point	75 to 78 °C	168-172 °C
Boiling Point	157 °C	420 °C
Heat of Fusion	25.96 kJ/mol	27.51 kJ/mol
Structure		

Since there are two APIs that need to be recovered, a minimum of two sequential crystallisation processes are required. Either of the two APIs may be crystallised first, followed by the crystallisation of the other API. The cooling profile shown in Figure 6.38 is obtained for the multi-component system consisting of a ratio of 5 kmol of ethanol: 2 kmol 2-(4-Isobutylphenyl) propionic acid: 0.5 kmol of 4-Acetaminophenol.

From the yield profile it is evident that cooling crystallisation alone will not fulfill the requirements for the crystallisation of pure API crystals. In the cooling temperature range of 50 °C to -50 °C, both APIs will crystallise as the system is cooled. The same phenomena will apply if evaporation, followed by cooling, is implemented. Hence, to effectively recover both APIs in their pure state, alternative options need to be considered.

Using the anti-solvent identification process outlined in application 2, the following solvents, water and n-hexene were identified as potential anti-solvents for 2-(4-Isobutylphenyl) propionic acid and 4-Acetaminophenol, respectively.

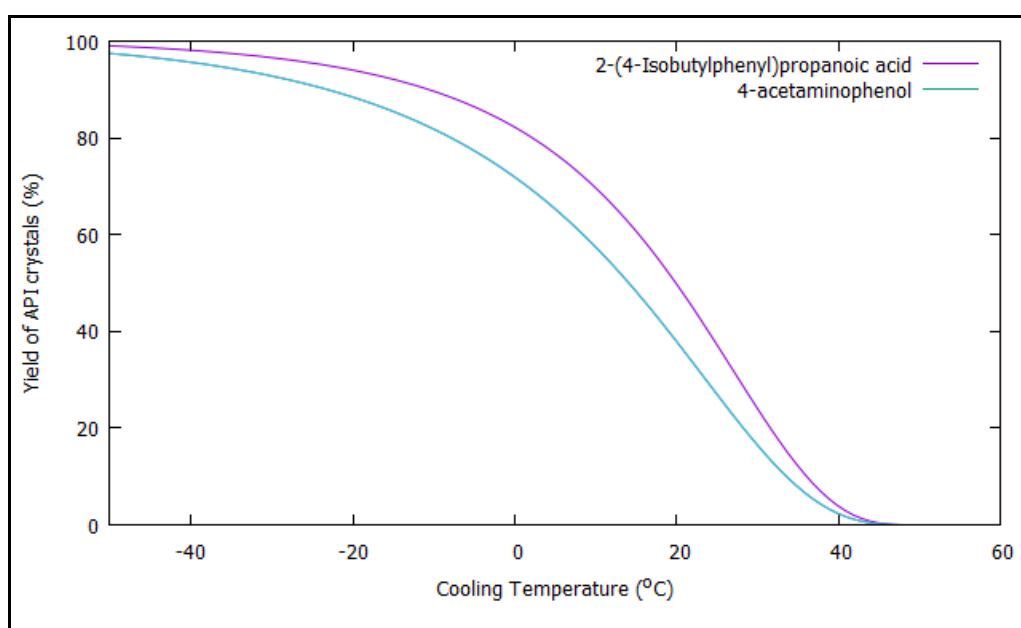


Figure 6.38. Effect of Cooling on the Crystallisation of a Binary API Mixture.

Following from application 2, where water is used as an anti-solvent for the recovery of 2-(4-Isobutylphenyl) propionic acid from ethanol, the presence of an additional component in the feed mixture will result in a different set of operating conditions, compared to the application where only one component is in solution. In Figure 6.39, conditions that will result in a 98 % recovery of pure 2-(4-Isobutylphenyl) propionic acid is presented. The difference in operating conditions is illustrated by comparing the operating conditions for the single API feed shown in Figure 6.34 and Figure 6.35 and the operating conditions obtained for the binary API feed shown in Figure 6.39 and Figure 6.40.

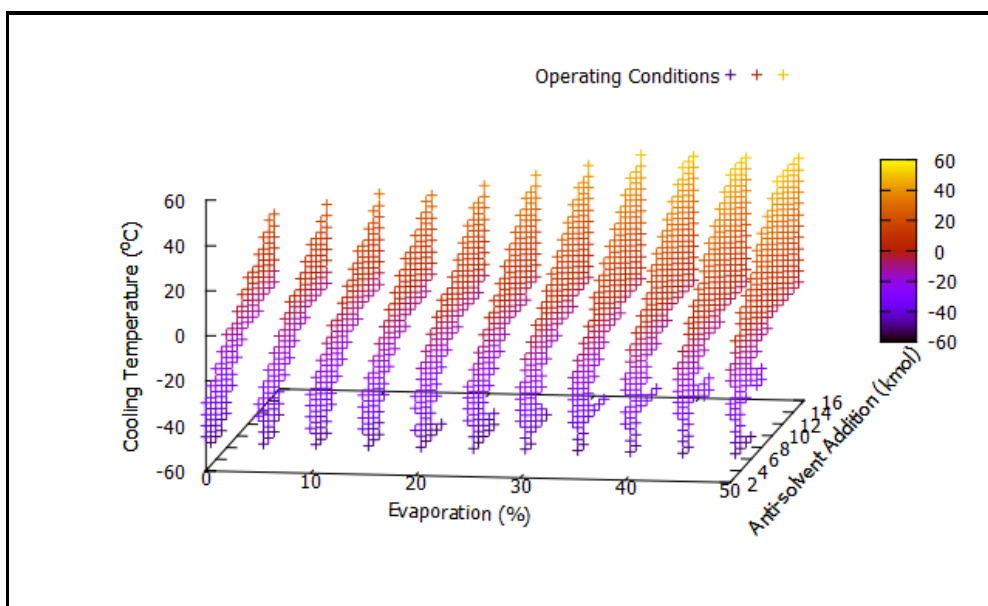


Figure 6.39. Potential Operating Conditions to Achieve a Minimum Yield of 98 % Pure API Crystals.

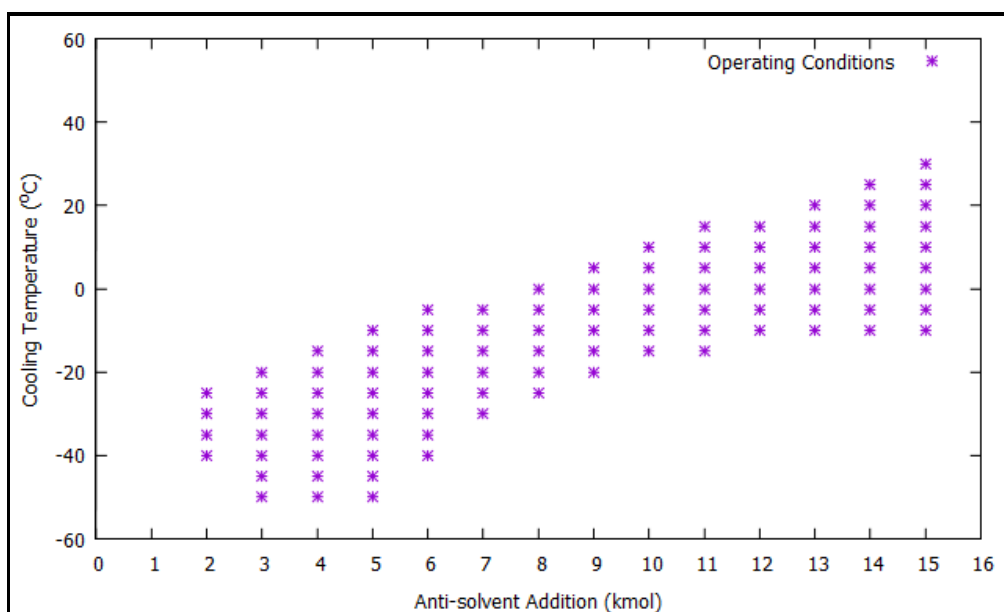


Figure 6.40. Band of Operating Conditions for the Combination of Cooling and Anti-solvent Addition to Achieve a Minimum Yield of 98 % Pure API Crystals With 10 % Evaporation.

Operating conditions based on the lowest operating cost for operations, with and without process and plant limitations, are presented in Table 6.12. In comparison to Table 6.10, it can be seen that the various performance criteria (operating cost, environmental factor and energy consumption factor) are higher than those for the same production rate of pure 2-(4-Isobutylphenyl) propionic acid in a binary feed.

**Table 6.12. Operating Conditions for the Recovery of First API From the Binary Mixture
Ranked on Lowest Operating Cost.**

Operating Conditions (OC)			Performance Criteria (PC)			
Evaporation	Anti-solvent Addition	Cooling Temperature	Yield obtained	Operating Cost per Batch	E Factor	Ec Factor
%	(kmol)	(°C)	(%)	(\$)	$\frac{kg\ waste}{kg\ API}$	$\frac{MJ}{kg\ API}$
Without process or plant limitations:						
0.00	3.00	-50.00	99.61	723.98	0.88	9.45
5.00	3.00	-50.00	99.64	730.98	0.88	9.46
0.00	2.00	-45.00	99.32	731.60	0.87	8.82
10.00	3.00	-50.00	99.67	738.22	0.88	9.47
15.00	3.00	-50.00	99.70	745.69	0.88	9.48
With process and plant limitations eg. no evaporation due to API thermal degradation and lowest cooling temperature attainable in the plant is 0 °C						
0.00	10.00	0.00	98.45	1346.84	0.96	14.04
0.00	9.00	0.00	98.18	1352.08	0.96	13.43
0.00	11.00	0.00	98.67	1352.79	0.97	14.66
0.00	12.00	0.00	98.86	1367.45	0.97	15.29
0.00	13.00	0.00	99.01	1388.98	0.98	15.91

Assuming 100 % crystal recovery of the pure 2-(4-Isobutylphenyl) propionic acid from the first crystallisation process, the remaining mother liquor is then subjected to a second crystallisation process. A combination of 80 % evaporation followed by cooling to -50 °C will result in a 91 % recovery of the second API. However, the purity obtained is 99.95 %, implying that there are traces of 2-(4-Isobutylphenyl) propionic acid (0.05 %) in the crystals obtained.

6.2. Concluding remarks

It is well established that the physical and chemical properties of chemicals in a system greatly influence the choice of operations, the size of the heat, mass and momentum transfer equipment, and the type and amount of utilities required. During the conceptual evaluation phase, the various factors that will impact on the selection of the final industrial-scale processing route must be evaluated.

In this chapter, the computational tool developed for the selection of solvents and anti-solvents for crystallisation processes has been demonstrated. The computational capability, range of applicability, and accuracy have been demonstrated by means of comparative studies using experimental and real plant data, and have been further demonstrated using several case studies. It has also been demonstrated that unlike the ternary diagram approach, which limits the analysis of systems to a maximum of 3

components (inclusive of solvent and anti-solvent), this computational tool can analyze multi-component systems containing more than three components.

It has been shown that using the multi-layer approach, potential solvents, mode of crystallisation and operating conditions can be identified, based on process performance, economic assessment and environmental assessment. It can be concluded that this computational tool can be used for decision making during the conceptual design phase of API development, as well as during retro-fit design or optimization studies of existing crystallisation processes.

Chapter 7: Conclusions and Recommendations

The main achievement of this work has been the development of a methodology that takes advantage of existing thermodynamic and process insights to provide feasible and near optimum solutions for the selection of solvents and anti-solvents for the synthesis and operational design of crystallisation processes. A robust and reliable, generic model-based crystallisation computational framework specifically targeting the pharmaceutical industry was developed that can predict and optimize the production of crystalline API materials, with the desired yield and purity based on solvent selection and selection of mode of crystallisation.

The computational framework explores the synergistic combination of multi-component multiphase flash calculations, phase equilibria phenomena, and process systems engineering methods, to establish the presence of solid-liquid equilibria of the components in a given feed, and the identification of the sequence of the precipitating solids under varying temperature and concentration changes. These computational capabilities allow the developed framework to provide the insight to exploit the change in the solubility boundaries and regions with temperature variation and concentration changes and to be used as a screening mechanism to quickly determine the most appropriate solvent(s) and type of crystallisation process/processes for a given application. It has been developed for faster process design and process understanding, that can be used in industry as a decision making tool during the conceptual design phase, and as a design or optimisation tool for retrofitting an existing process to maximize the overall process performance.

The successful embedding of the developed crystallisation computational framework within the commercial process simulation software CHEMCAD improves its robustness and extend its computational capabilities. The computations within the Solvent Selection module has access to a full range of thermodynamics models and correlations, a comprehensive database of compounds and their pure and mixture properties, and rigorous computational algorithms for process calculations and equipment design of the commercial simulator. In addition, the software vendors regularly deliver updates for their programs to fix bugs and to deliver new functionality including the updating and inclusion of new property and thermodynamic models and methods that enhance the accuracy of the predictive methods. Through several graphical user interfacing platforms, the user is able perform numerous calculations to comprehensively evaluate the process, economic and environmental impact of solvent selection.

The computational framework creates opportunities, not only for finding near optimal operating strategies, but also to investigate and develop a comprehensive understanding of the process, economic

and environmental impact of solvent and anti-solvent selection in crystallisation process. Specifically, the developed framework consists of several algorithms and subroutines that enables the following spectrum of computational capabilities:

- The optimum operability conditions for the crystallizer can be identified with minimal thermodynamic information on the system, by using multiphase flash calculations. Starting with just the chemical structure, and using the predictive thermodynamics property models like UNIFAC and other group contribution methods, the Solid-Liquid-Vapour equilibria (SLVE) phase behaviour can be calculated. The conditions at which there is a phase change from liquid to solid of a component of interest represents the onset of the crystallisation process for that component. By determining that change in the amount of solid formation by flash calculations due to temperature or composition change or both, the extend of crystal formation can be determined.
- The various eutectic temperatures and compositions that exist in the system can be predicted, and data can be generated for developing various types of phase diagrams and solubility curves. These allow for the overall composition space to be visualised and the separation barriers to be examined. In particular, the analysis of systems with multi-components and multiple saturation points is enabled.
- The operations such as heating, cooling, solvent addition, and solvent removal, can be simulated to systematically evaluate process alternatives. This enables the user to filter and screen solvents, and evaluate the effects of co-solvents, anti-solvents, other components, and impurities on the solute's solubility, in a specified temperature and composition range.
- The identification of the operating conditions that give maximum recovery of a desired compound, with a certain solvent or solvent mixture, and the calculation of the percent recoveries, and the total energy requirements (heating/cooling), under various operating conditions is enabled. The tool can be used to establish operating strategies, which may involve a combination of "cooling/heating", "co-solvent/anti-solvent addition", and "Evaporation" steps to meet the process objectives. All the important process alternatives can also be identified by this procedure, and can be systematically evaluated for quick screening purposes. The results obtained from this procedure will be mainly helpful in obtaining a quick, preliminary estimate of the best alternative.
- The framework can also be used to perform sensitivity analyses on the various input parameters to the process, and therefore identify important design variables in the process that will have the greatest impact on the overall performance.
- The inclusion of financial and environmental impact algorithms enhances and extend the applicability for significant and realistic comparative studies. The comparative investigation of the process engineering implications of the various solvents and modes of crystallisation can

be undertaken. Once the desired production rate is established, the effect each solvent will have on the size of various key equipment required, can be determined, along with the associated capital expenditure. In addition, the operational expenses and environmental impact associated with a selected solvent / anti-solvent and selected mode of crystallisation can be evaluated.

A series of validation processes and applications have shown that the thermodynamic and process insights embedded in this computation framework can be exploited to provide solutions for the synthesis and operational design of crystallisation processes, and in particular the impact a selected solvent, anti-solvent and mode of crystallisation may have on the overall performance of the process, as the goal of this thesis was originally set to be.

A key limitation of this methodology is that the accuracy of the predictions for complex molecules is dependent on the accuracy of predictive pure and mixture property models and predictive phase equilibria models that exist within the commercial simulator. Since the simulator allows for the inclusion of experimentally measured properties and user defined property models, this may for a specific purpose improve the accuracy for a particular application.

In summary, the innovative computational framework developed in this work and embedded into a commercial process simulator, provides a simple and fast way of conducting comparative studies to provide feasible, near optimum solutions that could be used for screening design and operational alternatives, and eventually be used for further rigorous optimisation studies. It presents a robust conceptual design tool for the rapid screening solvents and operational modes for crystallisation.

Chapter 8: Recommendations for Future work

Whilst the framework developed is intended to be used as a conceptual design tool to assist with decision making, several opportunities exist to enhance the capability, range of applicability and accuracy of the predictive design tool. These opportunities include:

1. Inclusion of a robust multi-objective optimisation algorithm into the Solvent Selection Framework may provide a more concise identification of the optimum operating conditions taking into account the process, economic and environmental requirements.
2. The inclusion of raw material and product cost, personnel costs and other operational cost not included into the existing framework may provide a relevant economic assessment of the process alternatives or be used as a production management tool to determine daily operational profitability requirements.
3. The models that have been used to calculate the equipment sizes and costs are generally simple and they can be improved to be more exact by including more details
4. The computation framework may be extended to include the other relevant API manufacture operations such as crystal washing, drying, etc. to fulfil the role as a process simulator for API manufacture.
5. The extension of the environmental performance index that accounts for plant and personnel safety, resource depletion, energy conservation and fugitive emission associated with the choice of solvent will improve the framework to present a holistic environmental assessment.
6. Whilst the existing framework uses the UNIFAC and Modified UNIFAC (Dortmund) predictive models for the phase equilibria computations, other emerging predictive models can be included to improve the accuracy of the computations with complex molecules.
7. The developed algorithms would also benefit from further validation from other case studies, experimental work and literature examples.

REFERENCES

- Abbas, A., Romagnoli, J. and Wideniski, D., 2013 Modeling of Crystallization Processes. *Process Systems Engineering: Dynamic Process Modeling, Volume 7*, pp.239-285.
- Abildskov, J. and Kontogeorgis, G.M., 2004. Chemical Product Design: A new challenge of applied thermodynamics. *Chemical Engineering Research and Design*, 82(11), pp.1505-1510.
- Abrams, D.S. and Prausnitz, J.M., 1975. Statistical thermodynamics of liquid mixtures: a new expression for the excess Gibbs energy of partly or completely miscible systems. *AIChE Journal*, 21(1), pp.116-128.
- Achenie LEK, Gani R, and Venkatasubramanian V. *Computer Aided Molecular Design: Theory and Practice*, Elsevier Science, Amsterdam (2003).
- Ahlers, J. and Gmehling, J., 2001. Development of an universal group contribution equation of state: I. Prediction of liquid densities for pure compounds with a volume translated Peng–Robinson equation of state. *Fluid Phase Equilibria*, 191(1), pp.177-188.
- Beckmann, W. ed., 2013. *Crystallisation: Basic concepts and industrial applications*. John Wiley & Sons.
- Berry, D.A. and Ng, K.M., 1997. Synthesis of reactive crystallisation processes. *AIChE Journal*, 43(7), pp.1737-1750.
- Berry, D.A., Dye, S.R. and Ng, K.M., 1997. Synthesis of drowning-out crystallisation-based separations. *AIChE Journal*, 43(1), pp.91-103.
- Bouillot, B., Teychené, S. and Biscans, B., 2011. An evaluation of thermodynamic models for the prediction of drug and drug-like molecule solubility in organic solvents. *Fluid Phase Equilibria*, 309(1), pp.36-52.
- Buxton, A., Livingston, A.G. and Pistikopoulos, E.N., 1999. Optimal design of solvent blends for environmental impact minimization. *AIChE Journal*, 45(4), pp.817-843.
- Byrne, F.P., Jin, S., Paggiola, G., Petchey, T.H., Clark, J.H., Farmer, T.J., Hunt, A.J., McElroy, C.R. and Sherwood, J., 2016. Tools and techniques for solvent selection: green solvent selection guides. *Sustainable Chemical Processes*, 4(1),

- Cabezas, H., Bare, J.C. and Mallick, S.K., 1999. Pollution prevention with chemical process simulators: the generalized waste reduction (WAR) algorithm—full version. *Computers & Chemical Engineering*, 23(4), pp.623-634.
- Chang, W.C. and Ng, K.M., 1998. Synthesis of processing system around a crystallizer. *AIChE Journal*, 44(10), pp.2240-2251.
- Chemstations, C. and en CHEMCAD, B., 2001. Process Flowsheet Simulator. *Chemstations Inc., Houston, TX, USA*.
- Chen, C.C. and Crafts, P.A., 2006. Correlation and prediction of drug molecule solubility in mixed solvent systems with the nonrandom two-liquid segment activity coefficient (NRTL-SAC) model. *Industrial & Engineering Chemistry Research*, 45(13), pp.4816-4824.
- Chen, C.C. and Mathias, P.M., 2002. Applied thermodynamics for process modeling. *AIChE Journal*, 48(2), pp.194-200.
- Chen, C.C. and Song, Y., 2004. Generalized electrolyte-NRTL model for mixed-solvent electrolyte systems. *AIChE Journal*, 50(8), pp.1928-1941.
- Chen, C.C. and Song, Y., 2004. Solubility modeling with a nonrandom two-liquid segment activity coefficient model. *Industrial & Engineering Chemistry Research*, 43(26), pp.8354-8362
- Chen, C.C., 1993. A segment-based local composition model for the Gibbs energy of polymer solutions. *Fluid Phase Equilibria*, 83, pp.301-312.
- Chen, C.C., 2010. Molecular thermodynamics for pharmaceutical process modeling and simulation. *Chemical Engineering in the Pharmaceutical Industry: R&D to Manufacturing*, pp.505-519.
- Cisternas, L., Cueto, J. and Swaney, R., 2004. Flowsheet synthesis of fractional crystallisation processes with cake washing. *Computers & Chemical Engineering*, 28(5), pp.613-623.
- Cisternas, L.A. and Swaney, R.E., 1998. Separation system synthesis for fractional crystallisation from solution using a network flow model. *Industrial & Engineering Chemistry Research*, 37(7), pp.2761-2769.
- Cisternas, L.A., 1999. Optimal design of crystallisation-based separation schemes. *AIChE Journal*, 45(7), pp.1477-1487
- Cisternas, L.A., Vásquez, C.M. and Swaney, R.E., 2006. On the design of crystallisation-based separation processes: Review and extension. *AIChE Journal*, 52(5), pp.1754-1769.

Constable, D.J., Jimenez-Gonzalez, C. and Henderson, R.K., 2007. Perspective on solvent use in the pharmaceutical industry. *Organic Process Research & Development*, 11(1), pp.133-137.

Constantinescu D. and Gmehling J., 2016, Further Development of Modified UNIFAC (Dortmund): Revision and Extension 6,” *Journal of Chemical Engineering Data*, pp. 2738-2748.

Crafts, P., 2007. The Role of Solubility Modeling and Crystallisation in the Design of Active Pharmaceutical Ingredients. In: KM Ng, R Gani, K Dam-Johansen (eds), *Chemical Product Design: Toward a Perspective through Case Studies* Elsevier B.V.

David J. am Ende. 2010, *Chemical Engineering in the Pharmaceutical industry: R&D to Manufacturing*. Wiley.

Deal, C.H. and Derr, E.L., 1968. Group contributions in mixtures. *Industrial & Engineering Chemistry*, 60(4), pp.28-38.

Dey, S.R. and Ng, K.M., 1995. Fractional crystallisation: Design alternatives and tradeoffs. *AIChE Journal*, 41(11), pp.2427-2438.

Diedrichs, A. and Gmehling, J., 2010. Solubility calculation of active pharmaceutical ingredients in alkanes, alcohols, water and their mixtures using various activity coefficient models. *Industrial & Engineering Chemistry Research*, 50(3), pp.1757-1769.

Doherty, M.F. and Malone, M.F., 2001. *Conceptual Design of Distillation Systems*, McGraw Hill, New York.

Douglas, J., 1988, *Conceptual Design of Chemical Process*, McGraw-Hill.

Dunn, P.J., Wells, A. and Williams, M.T. eds., 2010. *Green Chemistry in the Pharmaceutical Industry*. John Wiley & Sons.

Fitch, B., 1970. How to design fractional crystallisation processes. *Industrial & Engineering Chemistry*, 62(12), pp.6-33.

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

Giovanoglou, A., Barlatier, J., Adjiman, C.S., Pistikopoulos, E.N. and Cordiner, J.L., 2003. Optimal solvent design for batch separation based on economic performance. *AIChE Journal*, 49(12), pp.3095-3109.

Gmehling J, Dortmund Data Bank, DDBST GmbH, [Online]. Available: <http://www.ddbst.com>.

- Gmehling J, Kolbe B, Kleiber M, and Rarey J, 2012. *Chemical Thermodynamics for Process Simulation*, Wiley, 2012.
- Granberg, R.A. and Rasmuson, Å.C., 1999. Solubility of paracetamol in pure solvents. *Journal of Chemical & Engineering Data*, 44(6), pp.1391-1395.
- Grensemann, H. and Gmehling, J., 2005. Performance of a conductor-like screening model for real solvents model in comparison to classical group contribution methods. *Industrial & Engineering Chemistry Research*, 44(5), pp.1610-1624.
- Grodowska, K. and Parczewski, A., 2010. Organic solvents in the pharmaceutical industry. *Acta Pol Pharm*, 67(1), pp.3-12.
- Guggenheim, E.A., 1952. *Mixtures: the theory of the equilibrium properties of some simple classes of mixtures solutions and alloys*. Clarendon Press.
- Guthrie, K.M., 1974. *Process plant estimating, evaluation, and control*. Craftsman Book Company of America.
- Hahnenkamp, I., Graubner, G. and Gmehling, J., 2010. Measurement and prediction of solubilities of active pharmaceutical ingredients. *International Journal of Pharmaceutics*, 388(1), pp.73-81.
- Heidemann, R.A., 1983. Computation of high pressure phase equilibria. *Fluid phase equilibria*, 14, pp.55-78.
- Henderson, R.K., Jiménez-González, C., Constable, D.J., Alston, S.R., Inglis, G.G., Fisher, G., Sherwood, J., Binks, S.P. and Curzons, A.D., 2011. Expanding GSK's solvent selection guide—embedding sustainability into solvent selection starting at medicinal chemistry. *Green Chemistry*, 13(4), pp.854-862.
- Hildebrand J.H., Scott R.L., 1964. *Solubility of Non-Electrolytes*. 3rd ed., Reinhold, New York, Dover, New York
- Holderbaum, T. and Gmehling, J., 1991. PSRK: A group contribution equation of state based on UNIFAC. *Fluid Phase Equilibria*, 70(2), pp.251-265
- Horstmann S., 2005. PSRK Group Contribution Equation of State: Comprehensive Revision and Extension IV, Including Critical Constants and a-Function Parameters for 1000 Components, *Fluid Phase Equilibria*, pp. 157-164.
- Hsieh, C.M., Lin, S.T. and Vrabec, J., 2014. Considering the dispersive interactions in the COSMO-SAC model for more accurate predictions of fluid phase behavior. *Fluid Phase Equilibria*, 367, pp.109-116.

Jones, A.G., 2002, *Crystallisation Process Systems*. Butterworth-Heinemann.

Jouyban, A. and Acree, W.E., 2006. Solubility prediction in non-aqueous binary solvents using a combination of Jouyban-Acree and Abraham models. *Fluid Phase Equilibria*, 249(1), pp.24-32.

Kang, J.W., Abildskov, J., Gani, R. and Cobas, J., 2002. Estimation of mixture properties from first-and second-order group contributions with the UNIFAC model. *Industrial & Engineering Chemistry Research*, 41(13), pp.3260-3273.

Karunanithi, A.T., Achenie, L.E. and Gani, R., 2006. A computer-aided molecular design framework for crystallisation solvent design. *Chemical Engineering Science*, 61(4), pp.1247-1260.

Karunanithi, A.T., Acquah, C. and Achenie, L.E., 2008. Tuning the morphology of pharmaceutical compounds via model based solvent selection. *Chinese Journal of Chemical Engineering*, 16(3), pp.465-473.

Klamt, A., 2005. *COSMO-RS: From Quantum Chemistry to Fluid Phase Thermodynamics and Drug Design*. Elsevier.

Klamt, A., Jonas, V., Bürger, T. and Lohrenz, J.C., 1998. Refinement and parametrization of COSMO-RS. *The Journal of Physical Chemistry A*, 102(26), pp.5074-5085.

Kleiner, M., Tumakaka, F. and Sadowski, G., 2009. Thermodynamic Modeling of Complex Systems. In *Molecular Thermodynamics of Complex Systems* (pp. 75-108). Springer.

Kojima, K. and Tochigi, K., 1979. Prediction of Vapor-Liquid Equilibria Using ASOG. *Elsevier, Amsterdam*.

Kolář, P., Shen, J.W., Tsuboi, A. and Ishikawa, T., 2002. Solvent selection for pharmaceuticals. *Fluid Phase Equilibria*, 194, pp.771-782.

Kontogeorgis, G.M., Folas, G.K., 2010, *Thermodynamic Models for Industrial Applications: From Classical and Advanced Mixing Rules to Association Theories*. Wiley .

Kroenlein K, C. D. Muzny, CD, Kazakov, AF, Chirico RD, Magee JW, Abdulagatov I, and Frenke M, NIST Web Thermo Tables,[Online]. Available: <http://www.nist.gov/srd/nistwebsub3.cfm>.

Kwok, K.S., Chan, H.C., Chan, C.K. and Ng, K.M., 2005. Experimental Determination of Solid–Liquid Equilibrium Phase Diagrams for Crystallization-Based Process Synthesis. *Industrial & Engineering Chemistry Research*, 44(10), pp.3788-3798.

- Larsen, B.L., Rasmussen, P. and Fredenslund, A., 1987. A modified UNIFAC group-contribution model for prediction of phase equilibria and heats of mixing. *Industrial & Engineering Chemistry Research*, 26(11), pp.2274-2286
- Lazzaroni, M.J., Bush, D., Eckert, C.A., Frank, T.C., Gupta, S. and Olson, J.D., 2005. Revision of MOSCED parameters and extension to solid solubility calculations. *Industrial & Engineering Chemistry Research*, 44(11), pp.4075-4083.
- Leibovici, C.F. and Neoschil, J., 1995. A solution of Rachford-Rice equations for multiphase systems. *Fluid Phase Equilibria*, 112(2), pp.217-221.
- Lewis, G.N. and Randall, M., 1923. *Thermodynamics/Thermodynamics and the free energy of chemical substances*. McGraw-Hill.
- Lin, S.T. and Sandler, S.I., 2002. A priori phase equilibrium prediction from a segment contribution solvation model. *Industrial & Engineering Chemistry Research*, 41(5), pp.899-913.
- Lira-Galeana, C., Firoozabadi, A. and Prausnitz, J.M., 1996. Thermodynamics of wax precipitation in petroleum mixtures. *AIChE Journal*, 42(1), pp.239-248.
- Lucia, A., Padmanabhan, L. and Venkataraman, S., 2000. Multiphase equilibrium flash calculations. *Computers & Chemical Engineering*, 24(12), pp.2557-2569.
- Matsuda, H., Mori, K., Tomioka, M., Kariyasu, N., Fukami, T., Kurihara, K., Tochigi, K. and Tomono, K., 2015. Determination and prediction of solubilities of active pharmaceutical ingredients in selected organic solvents. *Fluid Phase Equilibria*, 406, pp.116-123.
- Mersmann, A., 2001, *Crystallisation Technology Handbook*, Marcel Dekker Inc., New York.
- Michelsen M and Mollerup J, 2007 *Thermodynamic Models: Fundamentals and Computational Aspects*, Holte: Tie-Line Publications.
- Michelsen, M.L., 1982. The isothermal flash problem. Part I. Stability. *Fluid Phase Equilibria*, 9(1), pp.1-19.
- Michelsen, M.L., 1982. The isothermal flash problem. Part II. Phase-split calculation. *Fluid Phase Equilibria*, 9(1), pp.21-40.
- Moodley, K., Rarey, J. and Ramjugernath, D., 2015. A universal segment approach for the prediction of the activity coefficient of complex pharmaceuticals in non-electrolyte solvents. *Fluid Phase Equilibria*, 396, pp.98-110.

Moodley, K., Rarey, J. and Ramjugernath, D., 2017. Experimental solubility data for prednisolone and hydrocortisone in various solvents between (293.2 and 328.2) K by employing combined DTA/TGA. *Journal of Molecular Liquids*, 240, pp.303-312.

Mota, F.L., Carneiro, A.P., Queimada, A.J., Pinho, S.P. and Macedo, E.A., 2009. Temperature and solvent effects in the solubility of some pharmaceutical compounds: Measurements and modeling. *European Journal of Pharmaceutical Sciences*, 37(3), pp.499-507.

Moyers Jr, C.G. and Rousseau, R.W., 1987. Crystallisation operations. *Handbook of Separation Process Technology*, Wiley, New York.

Mullin J ,1993. *Crystallisation*, Oxford: Butterworth-Heinemann Ltd.

Mullin, J.W. and Nývlt, J., 1971. Programmed cooling of batch crystallizers. *Chemical Engineering Science*, 26(3), pp.369-377.

Myerson, A., 2002. *Handbook of Industrial Crystallisation*. Butterworth-Heinemann.

Nagy, Z.K., Chew, J.W., Fujiwara, M. and Braatz, R.D., 2008. Comparative performance of concentration and temperature controlled batch crystallisations. *Journal of Process Control*, 18(3), pp.399-407.

Nass K. K., 1994. Rational Solvent Selection for Cooling Crystallisations. *Industrial & Engineering Chemistry Research*, 1994,33.

Ng, K.M. and Wibowo, C., 2003. Beyond process design: The emergence of a process development focus. *Korean Journal of Chemical Engineering*, 20(5), pp.791-798.

Ng, K.M., 2002. Design and development of solids processes—a process systems engineering perspective. *Powder Technology*, 126(3), pp.205-210.

Nouar, A., Benmessaoud, I., Koutchoukali, O. and Koutchoukali, M.S., 2016. Solubility Prediction of Active Pharmaceutical Compounds with the UNIFAC Model. *International Journal of Thermophysics*, 37(3), pp.1-15.

Nowee, S.M., Abbas, A. and Romagnoli, J.A., 2007. Optimization in seeded cooling crystallisation: A parameter estimation and dynamic optimization study. *Chemical Engineering and Processing: Process Intensification*, 46(11), pp.1096-1110

Nowee, S.M., Abbas, A. and Romagnoli, J.A., 2008. Anti-solvent crystallisation: Model identification, experimental validation and dynamic simulation. *Chemical Engineering Science*, 63(22), pp.5457-5467.

- Parekh, V.S. and Mathias, P.M., 1998. Efficient flash calculations for chemical process design—extension of the Boston–Britt “Inside–out” flash algorithm to extreme conditions and new flash types. *Computers & Chemical Engineering*, 22(10), pp.1371-1380.
- Pistikopoulos, E.N. and Stefanis, S.K., 1998. Optimal solvent design for environmental impact minimization. *Computers and Chemical Engineering*, 6(22), pp.717-733.
- Peters, M.S., Timmerhaus, K.D., West, R.E., Timmerhaus, K. and West, R., 2003. *Plant design and economics for chemical engineers* (5th Edition). McGraw-Hill.
- Rajagopal, S., Ng, K.M. and Douglas, J.M., 1991. Design and economic trade-offs of extractive crystallisation processes. *AIChE Journal*, 37(3), pp.437-447.
- Rajagopal, S., Ng, K.M. and Douglas, J.M., 1992. A hierarchical procedure for the conceptual design of solids processes. *Computers & Chemical Engineering*, 16(7), pp.675-689.
- Renon, H. and Prausnitz, J.M., 1968. Local compositions in thermodynamic excess functions for liquid mixtures. *AIChE Journal*, 14(1), pp.135-144.
- Ruether, F. and Sadowski, G., 2009. Modeling the solubility of pharmaceuticals in pure solvents and solvent mixtures for drug process design. *Journal of Pharmaceutical Sciences*, 98(11), pp.4205-4215.
- Schmid, B., Schedemann, A., and Gmehling, J., 2014. Extension of the VTPR Group Contribution Equation of State: Group Interaction Parameters for Additional 192 Group Combinations and Typical Results, *Industrial & Engineering Chemistry Research*, pp. 3393-3405.
- Schroer, J.W., Wibowo, C. and Ng, K.M., 2001. Synthesis of chiral crystallisation processes. *AIChE Journal*, 47(2), pp.369-387.
- Sheikholeslamzadeh, E. and Rohani, S., 2011. Solubility prediction of pharmaceutical and chemical compounds in pure and mixed solvents using predictive models. *Industrial & Engineering Chemistry Research*, 51(1), pp.464-473.
- Sheldon, R.A., 1992. Organic synthesis-past, present and future. *Chemistry and Industry*, (23), pp.903-6.
- Sheldon, R.A., 2005. Green solvents for sustainable organic synthesis: state of the art. *Green Chemistry*, 7(5), pp.267-278.
- Sheldon, R.A., 2007. The E factor: fifteen years on. *Green Chemistry*, 9(12), pp.1273-1283.

Shu, C.C. and Lin, S.T., 2010. Prediction of drug solubility in mixed solvent systems using the COSMO-SAC activity coefficient model. *Industrial & Engineering Chemistry Research*, 50(1), pp.142-147.

Slaughter, D.W. and Doherty, M.F., 1995. Calculation of solid-liquid equilibrium and crystallisation paths for melt crystallisation processes. *Chemical Engineering Science*, 50(11), pp.1679-1694.

Soares, R.D.P. and Gerber, R.P., 2013. Functional-segment activity coefficient model. 1. Model formulation. *Industrial & Engineering Chemistry Research*, 52(32), pp.11159-11171.

Spyriouni, T., Krokidis, X. and Economou, I.G., 2011. Thermodynamics of pharmaceuticals: Prediction of solubility in pure and mixed solvents with PC-SAFT. *Fluid Phase Equilibria*, 302(1), pp.331-337.

Talaviya S and Majmudar F., 2012. Green chemistry: A tool in Pharmaceutical Chemistry. *NHL Journal of Medical Sciences* July 2012, Vol. 1, Issue 1.

Tavare, N.S., 1995. Batch Crystallizer. In *Industrial Crystallisation* (pp. 93-139). Springer US.

The UNIFAC Consortium, [Online]. Available: <http://unifac.ddbst.de/>.

Treble, M.A., 1989. A preliminary evaluation of two and three phase flash initiation procedures. *Fluid Phase Equilibria*, 53, pp.113-122.

Tung, H.H., Tabora, J., Variankaval, N., Bakken, D. and Chen, C.C., 2008. Prediction of pharmaceutical solubility Via NRTL-SAC and COSMO-SAC. *Journal of Pharmaceutical Sciences*, 97(5), pp.1813-1820.

Tung Hsien-Hsin, 2009 *Crystallisation of Organic Compounds: An Industrial Perspective*. Wiley,

Turton, R., Bailie, R.C., Whiting, W.B. and Shaeiwitz, J.A., 2008. *Analysis, Synthesis and Design of Chemical Processes*. Pearson Education.

Ulrich, G.D. and Vasudevan, P.T., 2006. How to estimate utility costs. *Chemical Engineering*, 113(4), p.66.

Ulrich, G.D., 1984. *A Guide To Chemical Engineering Process Design and Economics* (p. 295). New York: Wiley.

Valvani, S.C., Yalkowsky, S.H. and Roseman, T.J., 1981. Solubility and partitioning IV: Aqueous solubility and octanol-water partition coefficients of liquid nonelectrolytes. *Journal of Pharmaceutical Sciences*, 70(5), pp.502-507.

Van Ness, H.C. and Abbott, M., 1997. Thermodynamics, *Perry's Chemical Engineers' Handbook*, chapter 4 (Thermodynamics). McGraw-Hill.

Variankaval, N., Cote, A.S. and Doherty, M.F., 2008. From form to function: crystallisation of active pharmaceutical ingredients. *AIChE Journal*, 54(7), pp.1682-1688.

Walas, S.M., 1985. *Phase Equilibria in Chemical Engineering*. Butterworth-Heinemann.

Weidlich, U. and Gmehling, J., 1987. A modified UNIFAC model. I: Prediction of VLE, hE, and γ^∞ . *Industrial & engineering chemistry research*, 26(7), pp.1372-1381.

Wibowo, C. and Ng, K.M., 2000. Unified approach for synthesizing crystallisation-based separation processes. *AIChE Journal*, 46(7), pp.1400-1421.

Wibowo, C. and Ng, K.M., 2001. Operational issues in solids processing plants: Systems view. *AIChE Journal*, 47(1), pp.107-125.

Wibowo, C. and Ng, K.M., 2002. Workflow for process synthesis and development: crystallisation and solids processing. *Industrial & Engineering Chemistry Research*, 41(16), pp.3839-3848.

Wibowo, C., O'Young, L. and Ng, K.M., 2004. Streamlining crystallisation process design. *Chemical Engineering Progress*, 100(1), pp.30-39.

Widenski, D.J., 2012. *A Thermodynamic Framework for the Modeling and Optimization of Crystallisation Processes*, Doctoral dissertation, Louisiana State University.

Widenski, D.J., Abbas, A. and Romagnoli, J.A., 2010. Comparison of different solubility equations for modeling in cooling crystallisation. *Chemical Engineering and Processing: Process Intensification*, 49(12), pp.1284-1297.

Wieckhusen, D., 2013. Development of Batch Crystallisations. In: Beckmannn, W., Editor: *Crystallisation: Basic Concepts and Industrial Applications*, Wiley, pp.187-202

Wilson, G.M., 1964. Vapor-liquid equilibrium. XI. A new expression for the excess free energy of mixing. *Journal of the American Chemical Society*, 86(2), pp.127-130.

Xue, Z., Mu, T. and Gmehling, J., 2012. Comparison of the a Priori COSMO-RS models and group contribution methods: Original UNIFAC, modified UNIFAC (Do), and modified UNIFAC (Do) consortium. *Industrial & Engineering Chemistry Research*, 51(36), pp.11809-11817.

APPENDIX A: Alternate SLE Derivation

Solid – Liquid Equilibria Equation from Classical Thermodynamic Relation

The classical approach used for relating the solid and liquid fugacities in a pure compound system can be obtained from classical thermodynamic relations. In the framework, the solid and liquid fugacities at melting are evaluated starting from the Fundamental Property Relation (FPR) of a pure compound system. It consists in obtaining the liquid and solid fugacities starting from the triple point conditions, and the pressure and temperature variations are evaluated in terms of fugacity.

The FPR of a pure component in a generic phase α in terms of volume and enthalpy can be expressed as

$$d\left(\frac{G^\alpha}{RT}\right) = \frac{V^\alpha}{RT} dP - \frac{H^\alpha}{RT^2} dT = d\left(\frac{\mu^\alpha}{RT}\right) = d\left(\frac{RT \ln f^{\alpha,0}}{RT}\right) = d(\ln f^{\alpha,0})$$

The variation of the total Gibbs energy in an isothermal and isobaric step can be represented by the first and second identity:

$$\left(\frac{d \ln f^{\alpha,0}}{dP}\right)_T = \frac{V^\alpha}{RT} \quad , \quad \left(\frac{d \ln f^{\alpha,0}}{dT}\right)_P = -\frac{H^\alpha}{RT^2}$$

It follows that the variation of the fugacity in the phase α from condition P_1, T_1 to condition P_2, T_2 can be expressed as:

$$\iint_{P_1 T_1}^{P_2 T_2} d \ln f^{\alpha,0} dT dP = \int_{P_1}^{P_2} \frac{V^\alpha}{RT} dP + \int_{T_1}^{T_2} -\frac{H^\alpha}{RT^2} dT$$

Assuming the following relationships

$$V_{P_1, T_1}^\alpha = V_{P_2, T_2}^\alpha \quad ; \quad H^\alpha = H_{P_1, T_1}^\alpha + Cp^\alpha(T - T_1) \quad ; \quad \text{and} \quad Cp_{P_1, T_1}^\alpha = Cp_{P_2, T_2}^\alpha$$

$$\begin{aligned} \ln \frac{f^{\alpha,0}(P_2, T_2)}{f^{\alpha,0}(P_1, T_1)} &= \frac{V_{P_1, T_1}^\alpha}{RT} \int_{P_1}^{P_2} dP - \frac{1}{R} \int_{T_2}^{T_1} \frac{H_{P_1, T_1}^\alpha}{T^2} dT - \frac{1}{R} \int_{T_1}^{T_2} \frac{Cp_{P_1, T_1}^\alpha}{T^2} (T - T_1) dT \\ &= \frac{V_{P_1, T_1}^\alpha}{RT} (P_2 - P_1) - \frac{H_{P_1, T_1}^\alpha}{R} \left(-\frac{1}{T_2} + \frac{1}{T_1}\right) - \frac{Cp_{P_1, T_1}^\alpha}{R} \left[\frac{T_1}{T_2} - 1 - \ln\left(\frac{T_1}{T_2}\right)\right] \\ &= \frac{V_{P_1, T_1}^\alpha}{RT} (P_2 - P_1) + \frac{H_{P_1, T_1}^\alpha}{RT_1} \left(\frac{T_1}{T_2} - 1\right) - \frac{Cp_{P_1, T_1}^\alpha}{R} \left[\frac{T_1}{T_2} - 1 - \ln\left(\frac{T_1}{T_2}\right)\right] \end{aligned}$$

This equation can be applied independently to the solid and liquid fugacities from the triple point temperature and pressure T_t, P_t up to the melting temperature and pressure T and P , resulting in the following equations for the liquid and solid phases respectively:

$$\ln \frac{f^{l,0}(P, T)}{f^{l,0}(P_t, T_t)} = \frac{V_{P_t, T_t}^l}{RT} (P - P_t) + \frac{H_{P_t, T_t}^l}{RT_t} \left(\frac{T_t}{T} - 1 \right) - \frac{Cp_{P_t, T_t}^l}{R} \left[\frac{T_t}{T} - 1 - \ln \left(\frac{T_t}{T} \right) \right]$$

$$\ln \frac{f^{s,0}(P, T)}{f^{s,0}(P_t, T_t)} = \frac{V_{P_t, T_t}^s}{RT} (P - P_t) + \frac{H_{P_t, T_t}^s}{RT_t} \left(\frac{T_t}{T} - 1 \right) - \frac{Cp_{P_t, T_t}^s}{R} \left[\frac{T_t}{T} - 1 - \ln \left(\frac{T_t}{T} \right) \right]$$

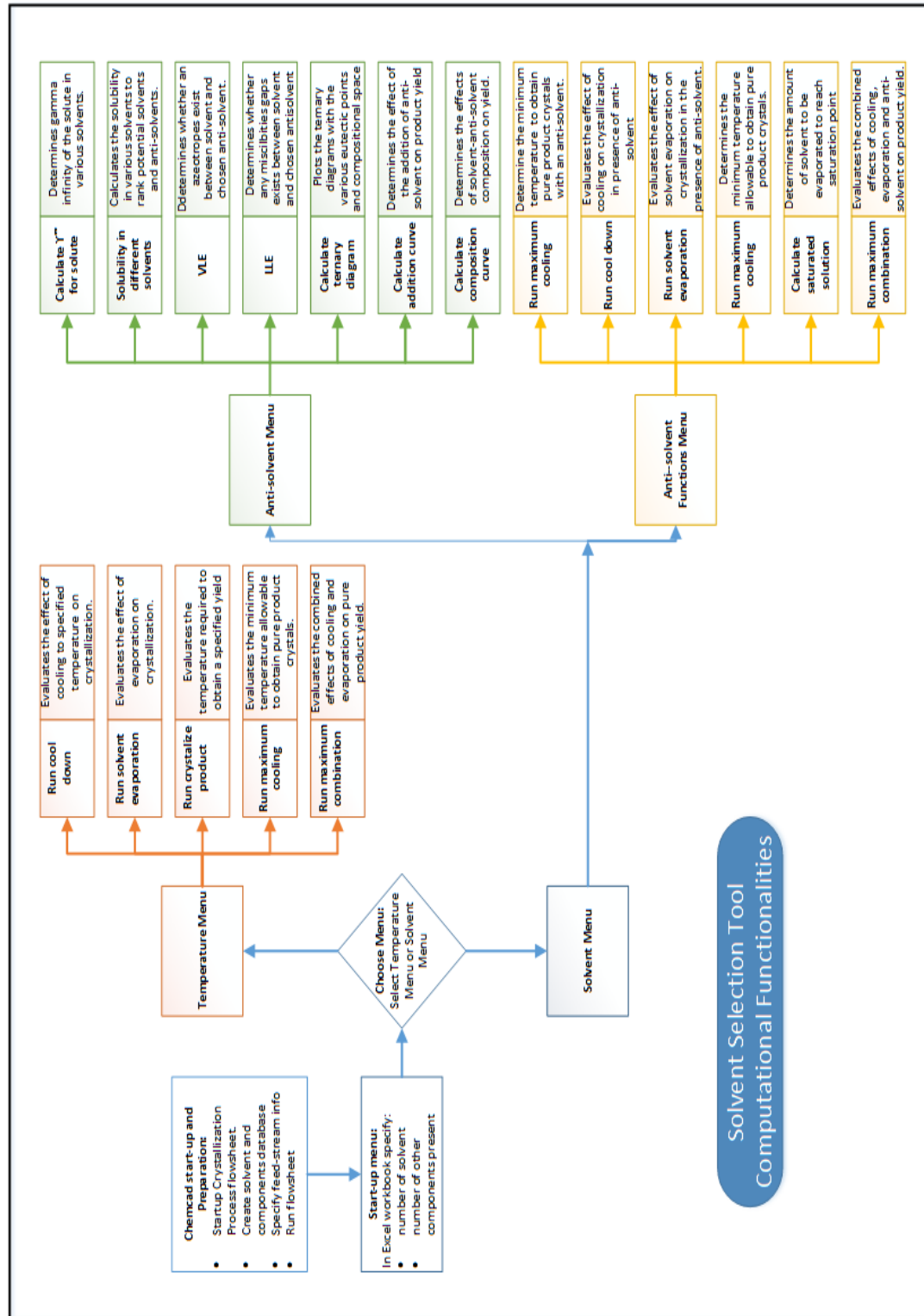
Applying the isofugacity condition between the solid and liquid phases at the triple point i.e $f^{s,0}(P_t, T_t) = f^{l,0}(P_t, T_t)$, the above equations can be combined to yield:

$$\ln \frac{f^{l,0}(P, T)}{f^{s,0}(P, T)} = \frac{(V_{P_t, T_t}^l - V_{P_t, T_t}^s)}{RT} (P - P_t) + \frac{(H_{P_t, T_t}^l - H_{P_t, T_t}^s)}{RT_t} \left(\frac{T_t}{T} - 1 \right) - \frac{(Cp_{P_t, T_t}^l - Cp_{P_t, T_t}^s)}{R} \left[\frac{T_t}{T} - 1 - \ln \left(\frac{T_t}{T} \right) \right]$$

$$\ln \frac{f^{l,0}(P, T)}{f^{s,0}(P, T)} = \frac{\Delta V_{P_t, T_t}^{sle}}{RT} (P - P_t) + \frac{\Delta H_{P_t, T_t}^{sle}}{RT_t} \left(\frac{T_t}{T} - 1 \right) - \frac{\Delta Cp_{P_t, T_t}^{sle}}{R} \left[\frac{T_t}{T} - 1 - \ln \left(\frac{T_t}{T} \right) \right]$$

APPENDIX B: Operating Manual

The Solvent Selection Tool



B.1. Introduction

This tool is helpful to optimize a crystallisation process with the simulation software CHEMCAD. To use it, you need CHEMCAD Version 6.4 or updated and a CHEMCAD-File called "Crystallisationexpert.cc6". This file includes the VBA Tool and allows you to optimize your crystallisation process fast and efficiently.

Whilst the VBA unit operation is embedded within CHEMCAD, the VBA unit operation controls CHEMCAD and has many functions to help you to synthesize and optimize your crystallisation process. The tool requires some set-up arrangements that will be described step by step later. The different menus with all their functions are explained in later sections.

All calculations are based on the Multiphase-Flash algorithm. This allows a simultaneous calculation of the vapor, liquid and solid phase behavior. This is important for a calculation close to the reality. Furthermore, the Multiphase- Flash algorithm controlled by the VBA Tool works very fast and allows a quick overview of your crystallisation process.

In general, the Solvent Selection Tool enables calculations in cooling crystallisation, evaporative crystallisation, anti-solvent or co-solvent and the combination of these crystallisation methods. Furthermore, it is possible to optimize the selection of an efficient solvent for the crystallisation process. The tool is directly connected to Microsoft Excel and exported all results directly to different sheets. This Excel interface allows for the results to be further analysed using the full spectrum tools and applications available in Excel.

In the tool are already 30 common solvents that are used in the pharmaceutical industry as possible solvent, anti-solvent or co-solvent preloaded, further solvents, by-products and impurities can be added to the crystallisation task. The tool deals in every function with problems for crystallizing impurities like by-products or solvents.

B.2. Simulation Preparation

In this section, we will explain step by step all arrangements that must be done prior to the use the Tool for the optimization of a crystallisation process. First, copy the original file "Crystallisationexpert.cc6" and rename it. Work with the copy! Open the file and you will see the flow sheet with just one VBA-Unit. This Unit controls is the whole Solvent Selection Tool and includes all VBA-Codes in one module. The first view of this file in CHEMCAD is shown in Figure B.1.

Before you can start with the optimization of the crystallisation you need to establish the component list for your simulation. Since in the crystallisation process, any component (including the solvents) in the feed-stream may crystallise depending on the conditions, we must ensure that for each component in the simulation component list, each component's solid clone must be created and included in the component list. Currently CHEMCAD allows for maximum of 50 liquid-solid pairs to be evaluated in a simulation and hence the maximum number of feed components for the crystallisation simulation is limited to 50. Whilst the programme has a preloaded solvent database of 30 solvents. These can be changed. It is important to note that developed algorithms make certain decisions/assumptions based on the position of the compound in the component list eg. It will automatically assume that the last two components in the database are the solute and crystal (solid) form of the desired product.

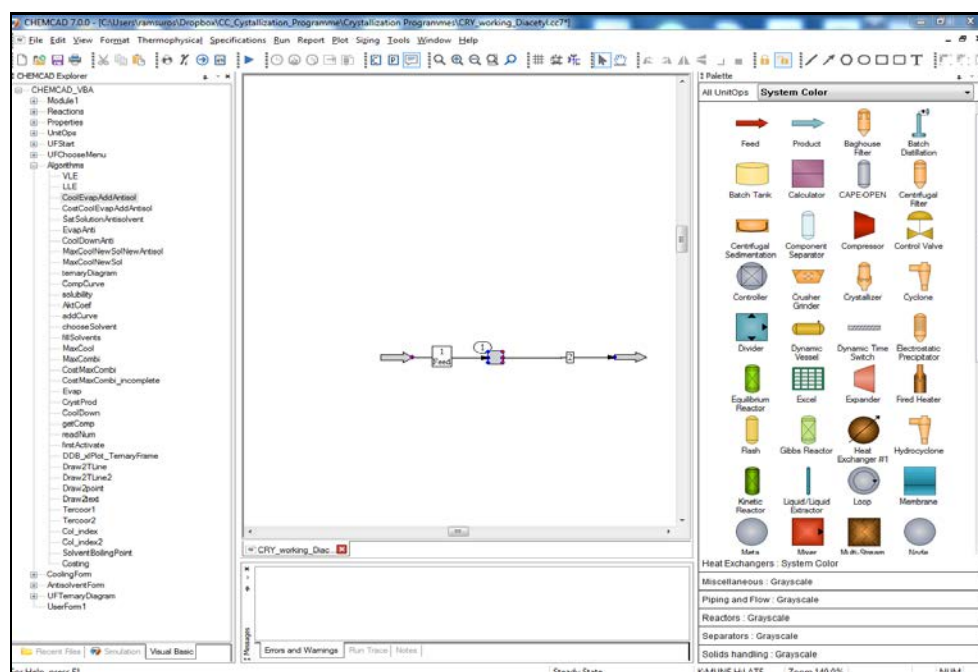


Figure B.1. View of CHEMCAD File with list of sub-routines.

Hence it important to ensure that that various components are arranged in the order reflected in Table B1. An illustrated example of a simulation component list is shown in Figure B.2. The procedure to create a component list for your simulation is as follows:

1. *First add solvents out of the CHEMCAD component database (all liquid) – must be 30*
2. *Then clone these solvents and create the solid phase. They have to be cloned in the same order you have added the liquid solvents.*
3. *Then you add additional solvents (if any) to the simulation and clone them to define the solids (also in the order you have added the liquids)*
4. *Then you add by-products (in any) to the simulation and clone them to define the solids (also in the order you have added the liquids)*
5. *Then you add the product and clone this one to define that as a solid.*

User defined components can be added into the CHEMCAD database. The procedure is outlined in the CHEMCAD “Help” menu. Large complex molecules may be fragmented into the various UNIFAC subgroups using “Artist” from the DDBST suite of programmes or equivalent tools. The fragmentation data is required by CHEMCAD to determine the various properties required in the simulation. Available experimentally measured data can also be uploaded into CHEMCAD

The calculations are based on the selected thermodynamic models. If you have good experimental data, you can regress binary interaction parameters (BIP's) and use NRTL (*non-random two liquids*) as the thermodynamic model. If you have no experimental data, you have to use a group contribution method like UNIFAC. From previous calculations, we recommend Modified UNIFAC (Dortmund), because it is closer to the real behavior. In Appendix D, is the list of subgroups available for the UNIFAC and Modified UNIFAC (Dortmund) models in CHEMCAD.

Table B.1: Scheme to Create Component List for Simulation.

Component number in simulation	Class	Notation in VBA Code
1 : : 30 (ncomp/2)	Preloaded Solvent – liquid phase	1 to ncomp/2
31 : : 60 (ncomp)	Preloaded Solvent – solid phase (crystals)	(ncomp/2 +1) to ncomp
61 : : 60+ (NofaS)	Further added Solvents – liquid phase	(ncomp +1) to (ncomp + numAddsol)
60 + 1 + NofaS : : 60 + 2*NofaS	Further added Solvents – solid phase (crystals)	(ncomp + 1 + numAddsol) to (ncomp + 2*NumAddsol)
60+ 2*(NofaS+1) : : 60 + 2*NofaS + NoB	By-products – liquid phase	ncomp + 2*(numAddsol +1) to (ncomp + 2*numAddsol + numAddothers)
60 + 2*NofaS + NoB +1 : : 60 + 2*NofaS + 2*NoB	By-products – solid phase (crystals)	(ncomp + 2*numAddsol+ numAddothers +1) to (ncomp + 2*numAddsol + 2*numaddothers)
60 + 2*NofaS + 2*NoB +1	Product – liquid phase (solute)	nges
60 + 2*NofaS + 2*NoB +2	Product –solid phase (crystals)	Nges + 1

ID	SolventName	RMM	Bpt(R)	Bpt(C)	SG					
1	Ethylene Glycol	62.068	846.81	197.3	1.11925					
2	Water	18.015	671.67	99.99999	1					
3	Acetic Acid	60.053	703.89	117.9	1.054249					
4	Ethanol	46.069	632.592	78.28999	0.796303					
5	Methanol	32.042	608.13	64.7	0.800598					
6	2-Propanol	60.096	639.738	82.25999	0.793767					
7	2-Ethoxyethanol	90.123	734.67	135	0.936549					
8	Monophenol	94.113	818.982	181.84	1.080172					
9	1-Butanol	74.123	703.458	117.66	0.815512					
10	Ethyl Acetate	88.106	630.378	77.05999	0.906758					
11	Pyridine	79.101	699.138	115.26	0.9875					
12	Acetone	58.08	592.92	56.24999	0.798561					
13	Acetonitrile	41.053	638.55	81.59999	0.7866					
14	N-Methyl-2-Pyrro	99.133	859.356	204.27	1.0357					
15	Diethyl Ether	74.123	553.644	34.42999	0.720751					
16	MEK	72.107	635.022	79.63998	0.81133					
17	Tetrahydrofuran	72.107	608.4	64.85001	0.892476					
18	Dichloromethane	84.932	563.22	39.74998	1.339496					
19	1,4-Dioxane	88.106	674.046	101.32	1.041598					
20	Cyclohexanone	98.145	772.02	155.75	0.953027					
21	Cyclohexanol	100.1611	781.2	160.85	0.972158					
22	Chloroform	119.377	601.794	61.18	1.504048					
23	P-Xylene	106.167	740.718	138.36	0.8657					
24	Cyclohexane	84.161	636.966	80.72	0.7834					
25	N-Hexane	86.177	615.384	68.73002	0.6633					
26	N-Pentane	72.15	556.596	36.07	0.63073					
27	3-Pentanone	86.134	675.252	101.99	0.820775					
28	Toluene	92.141	690.804	110.63	0.8718					
29	Glycerol	92.095	1009.8	287.85	1.26533					
30	Carbon Tetra-Cl	153.822	629.622	76.64001	1.604781					
31	Ethylene Glycol	62.068	846.81	197.3	1.11925					
32	Water (s)	18.015	671.67	99.99999	1					
33	Acetic Acid (s)	60.053	703.89	117.9	1.054249					
34	Ethanol (s)	46.069	632.592	78.28999	0.796303					
35	Methanol (s)	32.042	608.13	64.7	0.800598					
36	2-Propanol (s)	60.096	639.738	82.25999	0.793767					
37	2-Ethoxyethanol	90.123	734.67	135	0.936549					
38	Phenol (s)	94.113	818.982	181.84	1.080172					
39	1-Butanol (s)	74.123	703.458	117.66	0.815512					
40	Ethyl Acetate (s)	88.106	630.378	77.05999	0.906758					
41	Pyridine (s)	79.101	699.138	115.26	0.9875					
42	Acetone (s)	58.08	592.92	56.24999	0.798561					
43	Acetonitrile (s)	41.053	638.55	81.59999	0.7866					
44	N-Methyl-2-Pyrro	99.133	859.356	204.27	1.0357					
45	Diethyl Ether (s)	74.123	553.644	34.42999	0.720751					
46	2-Butanone (s)	72.107	635.022	79.63998	0.81133					
47	Tetrahydrofurane	72.107	608.4	64.85001	0.892476					
48	Dichloromethane	84.932	563.22	39.74998	1.339496					
49	1,4-Dioxane (s)	88.106	674.046	101.32	1.041598					
50	Cyclohexanone (s)	98.145	772.02	155.75	0.953027					
51	Cyclohexanol (s)	100.1611	781.2	160.85	0.972158					
52	Chloroform (s)	119.377	601.794	61.18	1.504048					
53	P-Xylene (s)	106.167	740.718	138.36	0.8657					
54	Cyclohexane (s)	84.161	636.966	80.72	0.7834					
55	N-Hexane (s)	86.177	615.384	68.73002	0.6633					
56	N-Pentane (s)	72.15	556.596	36.07	0.63073					
57	Clone Diethyl Ke	86.134	675.252	101.99	0.820775					
58	Toluene (s)	92.141	690.804	110.63	0.8718					
59	Clone Glycerol	92.095	1009.8	287.85	1.26533					
60	Clone Carbon Tet	153.822	629.622	76.64001	1.604781					
61	Benzene	78.114	635.832	80.08998	0.8844					
62	Benzene (s)	78.114	635.832	80.08998	0.8844					
63	acetaminophen	151.1658	954	256.85	0.757233					
64	Clone acetaminop	151.1658	954	256.85	0.757233					
65	ibuprofen	206.2859	1058.4	314.85	0.860222					
66	Clone ibuprofen	206.2859	1058.4	314.85	0.860222					
						Solvents - liquid phase				
						Solvents - clone solid phase				
						Additional Solvent liquid phase				
						Additional Solvent clone solid phase				
						By-product in liquid phase				
						By-product clone solid phase				
						Product in liquid phase (solute)				
						Product clone solid phase				

FIGURE B.2. Example of Simulation Component List (Output from CHEMCAD into Excel).

Now that all component setups are done and you can start the tool by running the simulation. The programme is driven by easy to use Graphical User Interfaces (GUIs) and several Excel worksheets that will automatically open for loading of input data and storage of simulation results.

On running the CHEMCAD flowsheet, the start-menu will pop up and an excel workbook will open.

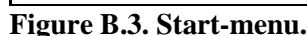


Figure B.4. Number of components.

Figure B.5. Feed composition InputRequirements.

1. Fill in number of additional solvents in excel workbook – sheet1 – Figure B.4
2. Fill in the additional number of components in the systems eg by-products and unreacted starting materials.- Figure B.4
3. Activate **Read number of components** Figure B.3
4. You will be required to Enter feed-stream composition and the target production rate per day into sheet 1- Figure B5. Nb the number of entries required is determined by the information entered in steps 1 and 2.
5. Activate **Read Composition** Figure B.3

A GUI will pop up containing two buttons for the two menus. Here we select the type of crystallisation process, we would like to evaluate:

- A. **Temperature Menu** evaluates Crystallisation by cooling and/or evaporation only
- B. **Anti-solvent Menu** evaluates Crystallisation in the presence of an anti-solvent

B.3.3. Menu 3: Temperature Menu

The Temperature Menu is programed for a selected solvent and feed composition (as defined in menu 1). This menu has many functions to optimize the crystallisation process via cooling and evaporative crystallisation and a combination of both crystallisation methods. After a calculation is finished all results will be exported to Excel and a “Calculation Finished” message-box will appear. All functions will be started via the related button on the right side of the menu shown in Figure B.6, and will be calculated at 1 atmosphere. Higher pressure can be done and will require the change to be made in the code. The computational capability of each function is briefly described below and typical results obtained are presented in Figure B.7, Figure B.8, and Figure B.9.

Figure B.6. Temperature Menu.

Note:

Number of steps for cooling:

The lower and upper bond describe the range in which you want to have a look for the crystallisation behaviour. The number of steps define how many points you want to calculate in this range. For example upper bond 25 degrees, lower bond -50 degrees, number of steps 15 means you will calculate at -50, -45, -40, ..., 15, 20, 25 Degrees Celsius.

Number of steps for evaporation:

For the solvent evaporation it is the same! Upper bond means how much solvent you want to evaporate in maximum for example 40 %, lower bond defines the minimum solvent evaporation for example 0 % . Number of steps defines how many points you want to calculate in this range for example 5, then you would calculate for a solvent evaporation of 0, 10, 20, 30, and 40 %.

Temperature Menu

Run cool down

Evaluates the effect of cooling on crystallisation. Insert final cooling temperature to determine effect on crystallisation process.

Run solvent evaporation

Evaluates the effect of evaporation on crystallisation. Insert desired % of solvent to be removed to determine effect on crystallisation process.

Run crystallize product

Evaluates the temperature required to obtain a specified yield of product. Note: At this temperature, other components in the feed may also crystallize, which will also be identified in the result sheet.

Run maximum cooling

This determines the minimum temperature allowable to obtain pure product crystals. This can also be used to determine the approximate eutectic temperature.

Run maximum combination Evaluates the combined effects of cooling and evaporation on the crystallisation process. It will also calculate the financial and environmental impacts associated with all combinations.

Results in Excel worksheet						Function	
1	A	B	C	D	E	F	Run cool down Results of the function cool down to -10 °C.
2	Temperature Menu results						
3	Results from calculation: Cool down to -10°C						
4	Componentname	Liquid Stream	Solid Stream	Vapour Stream	amount cristallized (%)		
5	ibuprofen	0,695059478	1,304940581	0	65,2470245		
6	Tetrahydrofuran	5	0	0	0		
7	Temperature	-9,99999254					
8							
9							
1	A	B	C	D	E	F	Run solvent evaporation Results of the function for evaporation of 25 %.
2	Temperature Menu results						
3	Results from calculation: Evaporate amount of 25% solvent						
4	Componentname product	Liquid Stream	Solid Stream	Vapour Stream	amount cristallized (%)		
5	ibuprofen	1,99998498	0	1,50445E-05	0		
6	Tetrahydrofuran	3,739936829	0	1,260063171	0		
7							
8	Temperature	82,35615099	Enthalpy needed [MJ]	923,4881429			
9							
10							
1	A	B	C	D	E		Run crystalize product Results of the function for crystallisation an amount of 90 % product.
2	Temperature Menu results						
3	Results from calculation: Crystallize amount of 90% product						
4	Componentname product	Liquid Stream	Solid Stream	Vapour Stream	amount cristallized (%)		
5	ibuprofen	0,193774715	1,806225419	0	90,31126404		
6	Tetrahydrofuran	5	0	0	0		
7	Temperature [°C]	-44,99953478					
8							
9							
1	A	B	C	D	E		Run maximum cooling The results show that for the given feed, the eutectic temperature is -108.99 °C. A maximum yield of pure component is 99.64% when the feed is cooled to -107.99 °C.
2	Temperature Menu results						
3	Results from calculation: Crystallize as much product as possible before any other component crystallizes						
4	First calculation with impurities			Temperature in °C	-108,9991007		
5	Componentname product	Liquid Stream	Solid Stream	Vapour Stream	amount cristallized (%)		
6	ibuprofen	0	2	0	100		
7	Tetrahydrofuran	0	5	0	100		
8							
9							
10	Last calculation without impurities			Temperature in °C	-107,9991075		
11	Componentname product	Liquid Stream	Solid Stream	Vapour Stream	amount cristallized (%)		
12	ibuprofen	0,007045247	1,992954969	0	99,6477356		
13	Tetrahydrofuran	5	0	0	0		
14							
15							

Figure B.7: Typical results obtained from various functions of the Temperature Menu.

		% Solvent evaporated										
		0	5	10	15	20	25	30	35	40	45	50
Cooling	-10.00	94.77	95.00	95.29	95.58	95.80	96.09	96.36	96.59	96.86	97.13	97.39
Temperat	-7.50	94.22	94.47	94.79	95.11	95.36	95.67	95.98	96.22	96.52	96.82	97.11
	-5.00	93.59	93.87	94.23	94.58	94.85	95.20	95.54	95.81	96.15	96.48	96.80
	-2.50	92.88	93.20	93.59	93.98	94.29	94.67	95.05	95.35	95.72	96.09	96.44
	0.00	92.08	92.43	92.87	93.30	93.64	94.07	94.49	94.83	95.24	95.65	96.05
	2.50	91.17	91.56	92.04	92.53	92.91	93.39	93.86	94.23	94.69	95.15	95.59
	5.00	90.12	90.56	91.10	91.64	92.07	92.60	93.13	93.55	94.06	94.57	95.07
	7.50	88.91	89.40	90.01	90.62	91.10	91.70	92.29	92.76	93.34	93.91	94.46
	10.00	87.50	88.06	88.74	89.43	89.97	90.64	91.31	91.84	92.49	93.13	93.76
	12.50	85.85	86.48	87.26	88.03	88.64	89.41	90.16	90.76	91.50	92.23	92.94
	15.00	83.90	84.61	85.49	86.37	87.07	87.94	88.80	89.49	90.33	91.15	91.96
	17.50	81.56	82.37	83.39	84.39	85.20	86.19	87.18	87.96	88.92	89.87	90.79
	20.00	78.73	79.67	80.84	82.00	82.92	84.07	85.21	86.11	87.22	88.31	89.38
	22.50	75.26	76.35	77.71	79.07	80.14	81.48	82.80	83.85	85.14	86.41	87.65
	25.00	70.97	72.25	73.85	75.43	76.70	78.26	79.81	81.04	82.56	84.05	85.51
	27.50	65.58	67.10	68.99	70.88	72.37	74.23	76.07	77.53	79.33	81.09	82.83
	30.00	58.77	60.59	62.86	65.11	66.91	69.13	71.34	73.08	75.24	77.35	79.43
	32.50	50.10	52.31	55.05	57.78	59.95	62.65	65.31	67.43	70.03	72.60	75.10
	35.00	39.09	41.79	45.14	48.47	51.12	54.40	57.66	60.24	63.42	66.55	69.61
	37.50	25.22	28.52	32.63	36.72	39.98	44.01	48.01	51.18	55.09	58.93	62.69
	40.00	7.95	12.01	17.08	22.11	26.12	31.09	36.01	39.90	44.71	49.44	54.07

Figure B.8. Matrix of Yield results for “Run Maximum Combination” function.

Total Operating Cost (\$)		0	5	10	15	20	25	30	35	40	45	50
-10.00	1358.90	1306.75	1242.19	1178.39	1127.95	1065.66	1004.35	956.06	896.79	838.86	782.50	
-7.50	1485.67	1427.31	1355.10	1283.78	1227.42	1157.88	1089.46	1035.61	969.54	905.00	842.25	
-5.00	1629.90	1564.38	1483.37	1403.42	1340.30	1262.45	1185.90	1125.70	1051.87	979.81	909.78	
-2.50	1794.96	1721.15	1629.96	1540.05	1469.09	1381.68	1295.79	1228.28	1145.56	1064.87	986.51	
0.00	1985.18	1901.68	1798.61	1697.08	1617.03	1518.49	1421.77	1345.81	1252.80	1162.15	1074.19	
2.50	2206.09	2111.16	1994.07	1878.88	1788.16	1676.59	1567.20	1481.37	1376.37	1274.16	1175.05	
5.00	2464.96	2356.36	2222.60	2091.17	1987.79	1860.79	1736.44	1638.98	1519.89	1404.07	1291.92	
7.50	2771.43	2646.30	2492.41	2341.44	2222.86	2077.39	1935.16	1823.84	1687.99	1556.06	1428.47	
10.00	3138.57	2993.14	2814.59	2639.77	2502.68	2334.79	2170.94	2042.89	1886.87	1735.60	1589.53	
12.50	3584.58	3413.73	3204.43	3000.00	2840.00	2644.46	2454.02	2305.49	2124.84	1950.04	1781.56	
15.00	4135.30	3931.97	3683.53	3441.54	3252.65	3022.34	2798.66	2624.58	2413.35	2209.44	2013.38	
17.50	4828.71	4582.72	4283.16	3992.42	3766.19	3491.20	3224.98	3018.42	2768.45	2527.87	2297.21	
20.00	5722.51	5418.64	5050.21	4694.29	4418.44	4084.39	3762.40	3513.46	3213.26	2925.42	2650.43	
22.50	6907.93	6522.42	6057.58	5611.17	5267.06	4852.46	4454.95	4149.09	3781.89	3431.50	3098.28	
25.00	8537.19	8030.27	7423.73	6846.02	6403.80	5874.65	5370.98	4985.78	4526.10	4090.27	3678.25	
27.50	10880.19	10180.30	9352.04	8572.23	7981.23	7280.63	6620.36	6119.74	5527.11	4969.92	4447.42	
30.00	14458.56	13422.77	12217.16	11101.48	10268.22	9293.74	8388.40	7710.21	6916.37	6178.89	5494.78	
32.50	20413.03	18710.85	16782.19	15045.74	13778.06	12326.42	11006.42	10035.19	8917.07	7895.77	6963.22	
35.00	31780.73	28448.91	24855.95	21775.47	19612.71	17220.76	15120.08	13617.69	11931.42	10430.36	9091.17	
37.50	60244.14	50955.46	42019.24	35118.34	30639.22	25999.32	22171.86	19565.21	16760.59	14365.18	12303.83	
40.00	234378.80	148298.19	98430.62	71486.03	57477.58	45105.42	36215.66	30731.01	25278.95	20951.53	17447.31	

Figure B.9. Matrix of Total Operating Cost (\$) results for “Run Maximum Combination” function.

Figure B.8 and Figure B.9 are examples of data that is obtainable from the Solvent Selection Tool. Here is shown a matrix of results for a combination of evaporation and cooling crystallisation processes. This is the result obtained for a system consisting of 2 kmols of ibuprofen in 5 kmols of ethanol. The effect of cooling from 40 °C to -10 °C, and also evaporating from 0% evaporation to 50% evaporation. In Figure B.8, we see that a 54% yield of pure ibuprofen crystals is obtainable if we evaporate 50% of ethanol without any further cooling. However, a yield of 97.39% is obtainable if the evaporated mixture is further cooled to -10 °C. From Figure B.9, we see that this is a very cost

effective as this represents the lowest operating cost from all possible combinations. The individual cost contributors such as cooling costs, waste treatment costs, etc are also simultaneously outputted into Excel. A graphical package such as GNUPLLOT can be used to transform the matrix of results into a 3D graph.

B.3.4. Menu 4: The Solvent/Anti-solvent Menu

The second and more complicated menu is the Solvent/Anti-solvent Menu. In this menu, just the feed composition is fixed and used. The menu contains two pages: The first called “Choose Anti-solvent” offers functions for the solvent and anti-solvent/co-solvent selection. The drop down menus allow the user to select: the desired product, the solvent to use, and the anti-solvent to use. The list of components in these dropdown menus are the same entered into CHEMCAD as the feed components.

B.3.4.1. Menu 4.1: Choose Anti-Solvent

The first page of the Solvent/Anti-solvent Menu contains computational tools for the selection of an efficient solvent / co-solvent / anti-solvent for the crystallisation process. Furthermore, it contains functions to select a potential anti-solvent or co-solvent. It can be used to classify the preloaded 30 solvents and further added solvents.

Note:

1. For the ternary diagram sub menu

The accuracy is an option for the algorithm to reach the eutectic point. For example: The algorithm starts with the feed composition, cools until the component starts to crystallize and then cools further down and is looking for a second component starting to crystallize. For this we programmed a temperature interval of 1 °C. When the second component crystallizes for the first time the algorithm will increase the temperature by 3 °C and cool down again to find the exact eutectic point. The temperature interval for this is the accuracy!!! I would recommend to test first a higher accuracy like 0.01 °C. If you couldn't get good exact eutectic points with this accuracy you must decrease the value.

Solvent Menu: Choose anti-solvent

Calculate γ^∞ for solute:

Determines gamma infinity of the solute in various solvents. It will calculate the activity coefficients for the selected solute in all solvents at infinite dilution.

Solubility in different solvents:

Calculates the solubility of the solute in various solvents at 25 °C, to identify potential solvents and anti-solvents.

VLE:

Determines whether an azeotropes exist between solvent and chosen anti-solvent.

LLE:

Determines whether any immiscibilities gaps exists between solvent and chosen anti-solvent

This function determines the ternary diagrams with the various eutectic points and displays the compositional space for the system.

This determines the effect of the addition of anti-solvent on the crystallisation process.

Solvent-anti-solvent composition
Determines the effects of solvent-anti-solvent composition on yield. The solvent-anti-solvent composition is varied from 0 to 1 mole fraction

[illegible]

Figure B.10. Results for the calculation of the *solubility in different solvents* function as an example.

Solvent	yield(%)	no cry	CapcostCry(\$)	CapcostHEX(\$)	Capcostsoltank(\$)	Capcostwsoltank(\$)	Capcostsevtap(m\$)	Chilling costs(\$)	wastetreatcost(\$)	steamcost(\$)	solvent costs	operatingcost(\$)	fixedCapcost(m\$)	E factor	Ec	FAC (m\$)	
Water	1.00	2.00	508628.16	327750.50	20695.31	13789.10	129951.93	448.53	28.58	49.48	43.67	570.31	4.30	0.22	1784.42	1.19	
1,2-Ethanediol	1.00	2.00	508628.16	411597.28	31767.02	18017.92	157394.73	889.40	45.42	61.92	134.88	1200.53	4.85	0.76	3239.98	1.45	
1,2-Propanediol	0.96	3.00	762942.25	551134.75	36477.98	21012.86	199364.45	923.67	55.46	81.70	183.77	2016.69	6.75	1.00	3508.18	2.09	
1,3-Propanediol	0.96	3.00	762942.25	564009.25	36202.98	20915.95	184699.28	959.25	55.15	74.69	180.78	2038.92	6.74	1.00	3559.61	2.09	
Acetanhydride	0.92	3.00	762942.25	432922.44	41940.12	24267.30	134292.69	103.33	65.38	51.42	246.51	2155.92	5.98	1.43	740.95	1.93	
Cis-Decalin	0.91	4.00	1017256.31	571082.56	54226.54	28409.54	134708.41	171.43	76.91	51.61	407.07	2597.32	7.74	1.93	951.47	2.45	
Isocotane	0.92	4.00	1017256.31	614679.63	55899.34	28638.97	106500.53	80.20	77.52	39.22	430.71	2288.04	7.81	1.58	569.41	2.41	
Methylcyclohexane	0.88	3.00	762942.25	447853.63	49930.01	28013.31	111501.06	75.02	75.85	41.38	348.19	3141.50	6.00	1.47	571.28	2.13	
Cyclohexane	0.85	3.00	762942.25	467061.31	46731.62	28268.33	111138.30	57.54	76.54	41.22	306.19	3999.23	6.07	1.38	516.35	2.32	
N-Heptane	0.91	4.00	1017256.31	617449.13	52792.20	28022.25	111459.38	77.93	75.88	41.36	387.12	2520.41	7.82	1.43	580.06	2.46	
Methylcyclopentane	0.85	3.00	762942.25	476586.38	47678.46	28553.77	108855.32	52.03	77.30	40.23	318.45	4013.96	6.10	1.38	491.30	2.33	
2-Methylpentane	0.90	3.00	762942.25	490174.41	50342.26	27716.83	100268.26	45.97	75.06	36.56	353.72	2793.51	6.13	1.28	442.41	2.09	
DiethyleneGlycol	0.93	3.00	762942.25	436767.09	42015.86	24182.03	178783.80	216.02	65.13	71.89	247.42	2217.84	6.19	1.47	1255.51	1.99	
TriF-AceticAcid	0.98	3.00	762942.25	482653.66	36790.88	20423.15	110080.19	48.49	53.57	40.76	187.20	715.77	6.05	1.43	484.79	1.66	

Figure B.11. Results showing Process, Economic and Environmental Performance for shortlisted solvents also obtained from the *solubility in different solvents* function.

In Figures B.10 and B.11, we see the full capability of the Solvent selection Tool. Besides calculating the solubility of a solute in a solvent, underlying algorithms also rank the solvents from good solvents to good anti-solvents. This is useful to identify potential solvents that can be used for a particular application. In addition, a full economic and environmental assessment can be done as shown in Figure B.11. Here is shown a shortlist of solvents that can achieve a minimum yield of 85 % for a particular application, and the capital cost, operating cost and environmental indicators are calculated for each solvent. Hence, sufficient data is presented to assist with decision making in the conceptual design or retro-fit process.

After you selected a solvent and an anti-solvent from the drop down menus you have first to check for miscibility via the LLE-Button. Behind the LLE-Button is a code based on a thermodynamic Gibbs energy model. We calculate for 100 points the Gibbs energy via the activity coefficients and look if there exist just one minimum (no miscibility gap) or there exist a second minimum in the Gibbs energy (miscibility gap). The function just returns via a message if there is a miscibility gap or not. Solvent systems that experience miscibility gaps must be avoided.

For a downstream solvent recovery processing, it could be interesting to check if an azeotrope exists between the selected solvent and anti-solvent, as this will determine the complexity of the recovery process. Because of this, an algorithm is implemented to check if an azeotrope exists between the selected solvent and anti-solvent. A simple calculation is performed by calculating saturated vapor pressures via the Antoine equation and the activity coefficient at infinitive dilution. The following conditions are checked:

$$\frac{\gamma_1^\infty P_1^{sat}}{P_2^{sat}} > 1 \text{ and } \frac{P_1^{sat}}{\gamma_2^\infty P_2^{sat}} < 1 \text{ OR if } \frac{\gamma_1^\infty P_1^{sat}}{P_2^{sat}} < 1 \text{ and } \frac{P_1^{sat}}{\gamma_2^\infty P_2^{sat}} > 1$$

then at some composition x , $\alpha_{12} = 1$

where γ_1^∞ represents the activity coefficient of component 1

at infinite dilution in component 2

where γ_2^∞ represents the activity coefficient of component 2

at infinite dilution in component 1

One of the most important things to develop all different process alternatives is to know the SLE behavior. Out of this reason, an algorithm was developed to predict a ternary SLE. This is very important in cases of impurities or the employment of an anti-solvent or co-solvent. The function "calculate ternary diagram" requires the selection of the three components via the drop down menus "choose solute", "choose solvent" and "choose anti-solvent". After the selection of 3 components, click on the button "calculate ternary diagram". This will open a new GUI as shown in Figure B.12.

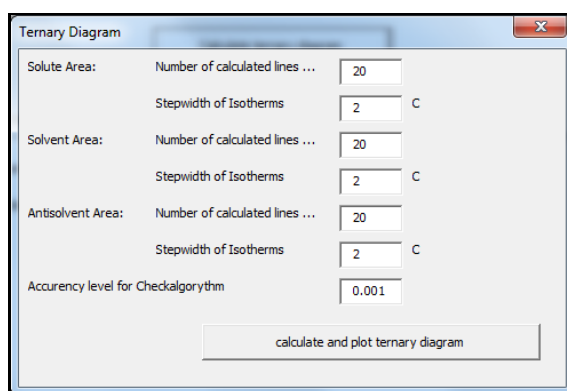


Figure B.12. – GUI to calculate ternary diagram.

In Figure B.13 we show the resulting already plotted ternary diagram for the components Benzene, Cyclohexane and 2-Butanone as an example with 3 binary eutectic points and one ternary eutectic point. The eutectic lines are plotted in black.

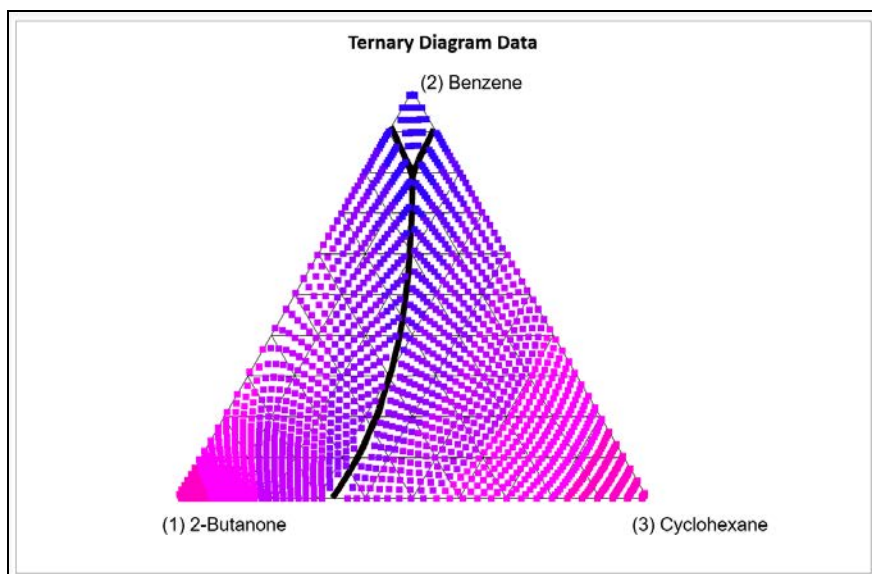


Figure B.13. Resulting diagram in Excel for Benzene, Cyclohexane and 2-Butanone as an example.

The function "Calculate addition curve" was implemented to calculate the effect of anti-solvent addition. It shows how the crystallisation is affected as the anti-solvent is gradually added into the process. From this computation, the dosage of anti-solvent can be determined to achieve a specified yield. The function will calculate for the inserted dosage and number of steps the crystallized amount of product via a TPFlash at 25 °C. The results are exported to Excel and can be plotted. The typical results for this function is shown in Figure B.14. The results are exported to Excel and can be plotted directly in Excel or with any other program like Origin or GNUPLLOT.

Antisolvent Menu Results					
Results from calculation: Add the antisolvent Water in 11 steps in increments of 1 kmol to the solution					
Addition Curve at 25°C					
Solvent	Antisolvent	ibuprofen, liq	ibuprofen, solid	ibuprofen, cryst.	
Ethanol	Water				
5	0.00	0.59	1.41	70.75	0
5	1.00	0.30	1.70	84.78	0
5	2.00	0.15	1.85	92.32	0
5	3.00	0.09	1.91	95.73	0
5	4.00	0.05	1.95	97.40	0
5	5.00	0.03	1.97	98.30	0
5	6.00	0.02	1.98	98.83	0
5	7.00	0.02	1.98	99.16	0
5	8.00	0.01	1.99	99.37	0
5	9.00	0.01	1.99	99.52	0
5	10.00	0.01	1.99	99.62	0

Figure B.14. Results of an example calculation for the Addition Curve function.

The next function "Calculate composition curve" is to check the efficiency of the selected solvent/anti-solvent system. This function calculates the solubility of the product over a composition range from pure solvent to pure anti-solvent. The number of calculated points must be inserted over the GUI. The function calculates for the different composition the crystallized amount of product via a TPFlash at any specified temperature. The typical results for this function is shown in Figure B.15. The results are exported to Excel and can be plotted directly in Excel or with any other program like Origin or GNUPLOT.

Antisolvent Menu Results					
Composition Curve at 25°C					
Solvent	Antisolvent	ibuprofen, liq	ibuprofen, solid	ibuprofen, cryst.	
Ethanol	Water				
0.714	0.000	0.084	0.202	70.748	
0.649	0.065	0.056	0.230	80.384	
0.584	0.130	0.033	0.253	88.498	
0.519	0.195	0.017	0.268	93.950	
0.455	0.260	0.008	0.277	97.061	
0.390	0.325	0.004	0.282	98.683	
0.325	0.390	0.002	0.284	99.468	
0.260	0.455	0.001	0.285	99.814	
0.195	0.519	0.000	0.286	99.946	
0.130	0.584	0.000	0.286	99.988	
0.065	0.649	0.000	0.286	99.998	
0.000	0.714	0.000	0.286	100.000	

Figure B.15. Results of an example calculation for the Composition Curve function.

B.3.3.2. Menu 4.2: Anti-Solvent Functions

The second page of the Solvent/Anti-solvent Menu contains various functions to have a closer look at the crystallisation behavior with the new selected solvent and anti-solvent from the first page. This menu has many functions to optimize the crystallisation process via cooling and evaporative crystallisation in the presence of an anti-solvent. All functions will be started by the related button on the right side of the menu shown and described below, and will be calculated at 1 atmosphere. Different pressure can be done and will require the change to be made in the code. The computational capability of each function is briefly described below and typical results obtained are presented in Figure B.16.

The first function "run maximum cooling" is implemented to check the crystallisation potential with the new solvent just via cooling down. This function works in the same way as the function "run maximum cooling" from the Temperature Menu only with the new solvent.

The next three functions "run cool down", "run solvent evaporation" and "run maximum cooling" are implemented for the calculation of a specific binary solvent crystallisation behavior. All these functions change the amount of solvent from the feed stream to an inserted new composition. In these next example calculation, we will have a closer look at the binary solvent system containing ibuprofen in a 50 % of solvent (ethanol) and 50 % of the selected anti-solvent (water).

Solvent Menu

Choose Antisolvent Antisolvent functions

Crystallize as much product as possible before any other component crystallize just via cooling down depending on new chosen solvent

run maximum cooling

For all calculations: amount of Solvent 50 % amount of Antisolvent 50 %

Cool down to ... -10 °C run cool down

Evaporate amount of solvent ... 10 % run solvent evaporation

run maximum cooling

Crystallize as much product as possible before any other component crystallize just via previous solvent evaporation, cooling down and adding an Antisolvent

calculate saturated solution

Lower and Upper bond for adding antisolvent in kmol

Crystallize as much product as possible before any other component crystallize via previous solvent evaporation, cooling down and adding an Antisolvent

Number of steps for cooling: Lower and Upper bond for temperature range for cooling in °C

Number of steps for evaporation: Lower and Upper amount of previous evaporated solvent in %

Number of steps for antisolvent adding: Lower and Upper added Antisolvent in kmol

Yield run maximum combination

Note:

Number of steps for cooling:

The lower and upper bond describe the range in which you want to have a look for the crystallisation behaviour. The number of steps define how many points you want to calculate in this range. For example upper bond 25 degrees, lower bond -50 degrees, number of steps 15 means you will calculate at -50,-45,-40,.....,15, 20, 25 Degrees Celsius.

Number of steps for evaporation:

The upper bond means how much solvent you want to evaporate in maximum for example 40 %, lower bond defines the minimum solvent evaporation for example 0 %. Number of steps defines how many points you want to calculate in this range for example 5, then you would calculate for a solvent evaporation of 0, 10, 20, 30, and 40 %.

Number of steps for anti-solvent addition:

The same with the Anti-solvent addition: the lower bond defines the minimum of added Anti-solvent for example 0 kmol/h, the upper bond describes the maximum for the Anti-solvent addition for example 25 kmol/h, number of steps define how many points you want to calculate in this range for example 6 this means you will calculate for an Anti-solvent addition of 0, 5, 10, 15, 20 and 25 kmol/h.

Anti-solvent Functions menu

Run maximum cooling:

This determines the minimum temperature allowable to obtain pure product crystals in the presence of an anti-solvent.

Run cool down

Evaluates the effect of cooling on crystallisation in presence of anti-solvent, for a given solvent – anti-solvent mixture.

Run solvent evaporation

Evaluates the effect of solvent evaporation on crystallisation, for a given solvent – anti-solvent mixture. Insert desired % of solvent to be removed to determine yield.

Calculate saturated solution:

Determines the amount of solvent to be evaporated to reach saturation point prior to anti-solvent addition.

Run maximum combination

evaluates the combined effects of cooling, evaporation and anti-solvent addition on pure product yield.

1	Solvent Menu results					
2	Results from calculation: Crystallize as much product as possible before any other component crystallizes					
3						
4	First calculation with impurities			Temperature in °C	-28,99964328	
5	Componentname	Liquid Stream	Solid Stream	Vapour Stream	amount crystallized (%)	
6	ibuprofen	0,009089603	1,990910411	0	99,54551697	
7	Water	2,469812155	0,030187923	0	1,207516909	
8	Ethanol	2,500000238	0	0	0	
9						
10						
11	Last calculation without impurities			Temperature in °C	-27,99965007	
12	Componentname	Liquid Stream	Solid Stream	Vapour Stream	amount crystallized (%)	
13	ibuprofen	0,009409579	1,990590453	0	99,52951813	
14	Water	2,499999762	0	0	0	
15	Ethanol	2,499999762	0	0	0	
16						

Temperature Menu results						
Results from calculation: Cool down to						
First calculation with impurities				Temperature in °C	-9,99999254	
Componentname product	Liquid Stream	Solid Stream	Vapour Stream	amount crystallized (%)		
ibuprofen	0,003047345	1,996952772	0	99,84763336		
Water	2,5	0	0	0	0	
Ethanol	2,5	0	0	0	0	

Temperature Menu results						
Results from calculation: Cool down to						
First calculation with impurities				Temperature in °C	39,99999051	
Componentname product	Liquid Stream	Solid Stream	Vapour Stream	amount crystallized (%)		
ibuprofen	0,043277714	1,95669651	0	97,83609009		
Water	1,562204599	0	0	0	0	
Ethanol	1,896890998	0	0	0	0	

Results for an example calculation of the *maximum cooling function* for a binary solvent system.

Results for an example calculation of the *run cool down function* for a given solvent – anti-solvent mixture.

Results for an example calculation of the *Run solvent evaporation function* for a given solvent – anti-solvent mixture.

Figure B.16. Examples of typical results for Anti-solvent Functions menu.

The last function is implemented to enable a quick overview about possible combinations from solvent evaporation, adding an anti-solvent and cooling down to a specific temperature. In addition, the economic and environmental performance is evaluated. For all three crystallisation methods, one has to define the upper and lower bond and also the number of steps, which should be calculated within this range. First the lowest amount of solvent will be evaporated, then the first amount of anti-solvent will be added and after that cooled to the lowest temperature. Then the amount of evaporated solvent and the added anti-solvent will be hold constant and the temperature will be changed in the selected range. After this the temperature change is repeated for the next value of added anti-solvent.

This procedure will be continued, until all possible combinations of added amount of anti-solvent in the whole temperature range are finished for the first fixed amount of previous evaporated solvent. Then this whole procedure is repeated for all selected amounts of evaporation. The evaporation function and the temperature function work in the same way like the evaporation function and the cool down function of the Temperature Menu. The addition of an anti-solvent changes the resulting composition after the solvent evaporation leading to a new composition, which is selected for the temperature behavior with a TPFash. We recommend to do first a calculation with just 10 steps for each crystallisation method to get an overview and then have a closer look to interesting regions. Otherwise the calculation of, for example 100 points of each crystallisation will lead to 1000000

combinations and this will take a long time to calculate.

This function produces 3 set of results: the results of the calculation of all possible combinations as shown in Figure B.17; the economic and environmental results as shown in Figure B.18, and then the function also filters and reports on all combinations leading to the user defined yield inserted on the GUI and where no impurities or solvent crystallize as shown in Figure B.19. The results are exported to Excel and can be plotted directly in Excel or with any other program like Origin or GNUPLLOT.

Results for combination of previous solvent evaporation, adding amount of antisolvent and cooling down							
choosen solvent			Ethanol				
choosen antisolvent			Water				
Amount of evaporation (%)	Amount of added antisolvent(kmol/hr)	temperature (C)	liquid product (kmol/hr)	solid product (kmol/hr)	amount of cristallization (%)	solvent crystallized (kmol/hr)	by-product crystallized (kmol/hr)
0	0	-49.99999254	0.01896108	1.981034766	99.0514255	0	0
0	0	44.99999254	0.025614511	1.974385142	98.71927643	0	0
0	0	-39.99999254	0.034158424	1.965841293	98.29207611	0	0
0	0	-34.99999254	0.045054846	1.954944068	97.74726105	0	0
0	0	-29.99999254	0.058878891	1.941120505	97.05605316	0	0
0	0	-24.99999254	0.076364078	1.923635483	96.18179321	0	0
0	0	-19.99999254	0.098466344	1.901533365	95.07668304	0	0
0	0	-14.99999254	0.126453948	1.873543839	93.67720032	0	0
0	0	-9.99999254	0.162058115	1.837911408	91.89709473	0	0
0	0	-4.99999254	0.207670167	1.792329788	89.61649323	0	0
0	0	7.45985E 06	0.266688108	1.733311534	86.66558838	0	0
0	0	5.00000746	0.344002813	1.655996919	82.79985809	0	0
0	0	10.00000746	0.446715772	1.553284168	77.66420746	0	0
0	0	14.99999051	0.585081875	1.414917827	70.74590302	0	0
0	0	19.99999051	0.773522556	1.226477385	61.32387161	0	0
0	0	24.99999051	1.031373501	0.96862644	48.43132401	0	0
0	0	29.99999051	1.38335526	0.618645012	30.83222961	0	0

Figure B.17. Extract of the results of the calculation of all possible combinations.

Total Operating Cost (\$)	Waste Treatment Cost (\$)	Cooling Cost (\$)	Replacement Solvent Cost (\$)	Cost of unrecovered api	Steam cost for evaporation (\$)	Solvent Recovery Cost (\$)	Scale-up Ratio	Cooling Required (k/hour)	Heating Required (k/hour)	Total feed (kg)	EFactor	EcFactor	FAC
472.80	48.13	43.78	141.57	191.47	0.00	47.85	2.45	2865568.50	2690504.00	1573.23	0.12	5.95	1672620.00
542.05	48.69	43.83	142.05	259.47	0.00	48.01	2.46	2869082.75	2699565.50	1578.53	0.13	5.96	1687794.25
631.73	49.39	43.93	142.66	347.52	0.00	48.22	2.47	2875317.50	2711298.25	1585.39	0.13	5.98	1707584.88
747.24	50.28	44.08	143.46	460.93	0.00	48.49	2.48	2884991.00	2726410.25	1594.23	0.14	6.01	1733213.88
895.64	51.38	44.29	144.48	606.65	0.00	48.84	2.50	2899037.75	2745827.50	1605.58	0.15	6.05	1766279.38
1086.38	52.75	44.59	145.79	793.96	0.00	49.28	2.52	2918706.75	2770786.00	1620.18	0.16	6.10	1808914.88
1332.44	54.44	45.00	147.49	1035.66	0.00	49.85	2.55	2945719.25	2802991.75	1639.01	0.17	6.16	1864065.50
1652.30	56.53	45.57	149.69	1349.91	0.00	50.60	2.59	2982519.00	2844866.25	1663.49	0.19	6.24	1935916.88
2073.10	59.11	46.33	152.59	1763.48	0.00	51.58	2.64	3032700.50	2899974.00	1695.72	0.21	6.36	2030631.00
2636.41	62.33	47.39	156.47	2317.32	0.00	52.89	2.70	3101768.25	2973773.00	1738.87	0.24	6.51	2157661.75
3408.95	66.38	48.87	161.80	3077.21	0.00	54.69	2.80	3198558.00	3075028.25	1798.08	0.28	6.72	2332217.00
4503.78	71.55	51.00	169.36	4154.63	0.00	57.25	2.93	3338039.50	3218594.25	1882.03	0.34	7.03	2580139.75
6125.99	78.32	54.20	180.56	5751.89	0.00	61.03	3.12	3547368.75	3431427.50	2006.48	0.43	7.48	2948512.25
8682.19	87.51	59.28	198.21	8270.19	0.00	67.00	3.43	3880378.00	3766990.75	2202.69	0.57	8.20	3531244.75
13088.43	100.63	68.11	228.67	12613.73	0.00	77.29	3.95	4458330.00	4345764.50	2541.12	0.81	9.44	4541889.50
21889.98	121.15	85.83	289.54	21295.59	0.00	97.87	5.00	5618157.50	5502619.00	3217.58	1.30	11.93	7034782.50
45770.02	160.23	134.05	454.81	44867.20	0.00	153.73	7.86	8774281.00	8643525.00	5054.18	2.61	18.68	13715854.00
271917.28	324.91	591.57	2021.34	268296.22	0.00	683.25	34.94	38721088.00	38415344.00	22462.81	15.02	82.77	90073328.00

Figure B.18. Extract of the results of the Economic and Environmental Performance calculations of all possible combinations.

results with required yield and without impurities									
Amount of evaporation (%)	Amount of added antisolvent (kmol/hr)	temperature (C)	E-factor (kgwaste/kgapi)	Ec-factor (MJ/kgapi)	amount of crystallization (%)	solvent crystallized (kmol/hr)	by-product crystallized (kmol/hr)	OP Cost	F A C
0.000	0.000	-50.000	0.122	5.950	99.052	0.000		472.803	1672620.000
0.000	0.000	-45.000	0.126	5.964	98.719	0.000		542.053	1687794.250
0.000	0.000	-40.000	0.131	5.984	98.292	0.000		631.729	1707584.875
0.000	1.000	-50.000	0.130	7.780	99.137	0.000		653.330	1861727.125
0.000	1.000	-45.000	0.134	7.797	98.839	0.000		715.887	1875484.375
0.000	1.000	-40.000	0.138	7.820	98.457	0.000		796.430	1893340.750
0.000	2.000	-50.000	0.137	8.414	99.351	0.000		663.137	1893522.750
0.000	2.000	-45.000	0.139	8.427	99.125	0.000		710.371	1903851.750
0.000	2.000	-40.000	0.142	8.444	98.836	0.000		771.143	1917290.750
0.000	2.000	-35.000	0.147	8.468	98.470	0.000		848.793	1934607.125
0.000	2.000	-30.000	0.152	8.501	98.007	0.000		947.602	1956786.625
0.000	3.000	-50.000	0.143	9.050	99.517	0.000		683.058	1925994.875
0.000	3.000	-45.000	0.145	9.058	99.348	0.000		718.237	1933592.375
0.000	3.000	-40.000	0.148	9.071	99.133	0.000		763.420	1943507.000
0.000	3.000	-35.000	0.151	9.088	98.860	0.000		820.973	1956284.750
0.000	3.000	-30.000	0.155	9.112	98.516	0.000		893.893	1972620.500
0.000	3.000	-25.000	0.160	9.145	98.085	0.000		985.988	1993395.125
0.000	4.000	-35.000	0.156	9.713	99.144	0.000		815.559	1981868.375
0.000	4.000	-30.000	0.159	9.731	98.887	0.000		869.816	1993952.750
0.000	4.000	-25.000	0.163	9.754	98.566	0.000		937.949	2009272.750
0.000	4.000	-20.000	0.168	9.786	98.167	0.000		1023.320	2028612.875
0.000	5.000	-25.000	0.168	10.372	98.913	0.000		918.992	2030822.250
0.000	5.000	-20.000	0.172	10.395	98.613	0.000		982.814	2045219.375
0.000	5.000	-15.000	0.176	10.426	98.242	0.000		1062.169	2063263.625
0.000	6.000	-20.000	0.177	11.013	98.934	0.000		969.134	2067177.625
0.000	6.000	-15.000	0.180	11.035	98.651	0.000		1029.273	2080783.750
0.000	6.000	-10.000	0.184	11.065	98.305	0.000		1103.508	2097721.750
0.000	7.000	-20.000	0.183	11.636	99.167	0.000		974.032	2092754.875
0.000	7.000	-15.000	0.185	11.653	98.948	0.000		1020.553	2103192.250
0.000	7.000	-10.000	0.188	11.675	98.680	0.000		1077.679	2116152.500
0.000	7.000	-5.000	0.192	11.704	98.354	0.000		1147.739	2132190.000
0.000	8.000	-15.000	0.191	12.276	99.167	0.000		1028.541	2128887.500
0.000	8.000	-10.000	0.194	12.293	98.956	0.000		1073.419	2138980.750
0.000	8.000	-5.000	0.197	12.315	98.700	0.000		1128.150	2151433.000
0.000	8.000	0.000	0.201	12.344	98.389	0.000		1194.927	2166768.250
0.000	8.000	5.000	0.205	12.381	98.013	0.000		1276.428	2185630.000
0.000	9.000	-15.000	0.198	12.904	99.330	0.000		1048.229	2156798.000
0.000	9.000	-10.000	0.200	12.916	99.161	0.000		1084.139	2164776.000
0.000	9.000	-5.000	0.203	12.933	98.956	0.000		1127.783	2174617.250
0.000	9.000	0.000	0.206	12.955	98.709	0.000		1180.745	2186703.000
0.000	9.000	5.000	0.209	12.984	98.411	0.000		1245.000	2201507.750
0.000	9.000	10.000	0.214	13.021	98.051	0.000		1323.088	2219645.500
0.000	10.000	-15.000	0.206	13.536	99.454	0.000		1076.227	2186198.000
0.000	10.000	-10.000	0.207	13.544	99.316	0.000		1105.443	2192586.750
0.000	10.000	-5.000	0.209	13.556	99.150	0.000		1140.848	2200474.500
0.000	10.000	0.000	0.212	13.573	98.949	0.000		1183.687	2210162.750

Figure B.19. Extract of the filtered results for a minimum yield of 98% of pure API crystals.

The End

APPENDIX C: Solvent Database

GlaxoSmithKline List of Solvents in the Pharmaceutical Industry					
Classification	Solvent	Cas Number	DDBST ID	Melting Point	Boiling Point
Water	Water	7732-18-5	174	0	100
Alcohol	2-Ethyl hexanol	104-76-7	379	-76	185
	Glycerol	56-81-5	230	18	290
	Cyclohexanol	108-93-0	252	25	161
	Ethylene glycol	107-21-1	8	-13	197
	1,4-butanediol	110-63-4	614	20	235
	Isoamyl alcohol	123-51-3	266	-117	131
	1,2-propanediol	57-55-6	282	-60	188
	1,3-propanediol	504-63-2	730	-27	214
	Benzyl alcohol	100-51-6	24	-15	205
	2-Pentanol	6032-29-7	766	-50	119
	1-Butanol	71-36-3	39	-89	118
	2-Butanol	78-92-2	22	-115	100
	Ethanol IMS	64-17-5	11	-114	78
	t-Butanol	75-65-0	153	25	82
	Methanol	67-56-1	110	-98	65
	2-Propanol	67-63-0	95	-88	82
	1-Propanol	71-23-8	140	-127	97
	2-Methoxyethanol	109-86-4	113	-85	124
Ester	t-Butyl acetate	540-88-5	1099	-78	95
	n-octyl acetate	112-14-1	612	-39	210
	Butyl acetate	123-86-4	80	-77	126
	Ethylene carbonate	96-49-1	1713	36	248
	Propylene carbonate	108-32-7	728	-55	242
	Isopropyl acetate	108-21-4	380	-73	89
	Ethyl lactate	97-64-3	2291	-23	154
	Propyl acetate	109-60-4	238	-92	102
	Dimethyl carbonate	616-38-6	451	-1	91
	methyl lactate	547-64-8	2290	-66	144
	Ethyl acetate	141-78-6	21	-84	77
	Ethyl propionate	105-37-3	205	-74	99
	Methyl acetate	79-20-9	82	-98	57
	Ethyl formate	109-94-4	16	-80	54
Ketone	Cyclohexanone	108-94-1	250	-32	155
	Cyclopentanone	120-92-3	241	-51	131
	2-Pentanone	107-87-9	137	-78	102
	3-Pentanone	96-22-0	285	-42	102
	Methylisobutyl ketone	108-10-1	117	-84	117
	Acetone	67-64-1	4	-95	56
	Methylethyl ketone	78-93-3	40	-87	80
Acid	Propionic acid	1979-09-04	141	-21	141
	acetic anhydride	108-24-7	233	-73	140
	Acetic acid	64-19-7	84	17	118

Classification	Solvent	Cas Number	DDBST ID	Melting Point	Boiling Point
Aromatic	Mesitylene	108-67-8	487	-45	165
	Cumene	98-82-8	351	-96	152
	p-Xylene	106-42-3	176	-13	138
	Toluene	108-88-3	161	-95	111
	Benzene	71-43-2	31	6	80
Hydrocarbon	cis-Decalin	493-01-6	315	-43	196
	ISOPAR G	64742-48-9	???	-60	163
	Isooctane	540-84-1	97	-107	99
	Methyl cyclohexane	108-87-2	1540	-127	101
	Cyclohexane	110-82-7	50	7	81
	Heptane	142-82-5	91	-91	98
	Pentane	109-66-0	134	-130	36
	Methylcyclopentane	96-37-7	55	-142	72
	2-Methylpentane	107-83-5	111	-153	60
	Hexane	110-54-3	89	-95	69
	Petroleum spirit	8032-32-4	???	-73	55
Ether	Di(ethylene glycol)	111-46-6	463	-10	246
	Ethoxybenzene	103-73-1	609	-29	170
	Tri(ethylene glycol)	112-27-6	443	-7	285
	Sulfolane	126-33-0	542	28	282
	DEG monobutyl ether	112-34-5	404	-68	231
	Anisole	100-66-3	18	-38	154
	Diphenyl ether	101-84-8	505	27	258
	Dibutyl ether	142-96-1	57	-95	140
	t-Amyl methyl ether	994-05-8	876	-80	86
	t-Butylmethyl ether	1634-04-4	822	-109	55
	Cyclopentyl methyl ether	5614-37-9	8047	-140	106
	t-Butyl ethyl ether	637-92-3	1409	-74	70
	2-Methyltetrahydrofuran	96-47-9	294	-137	78
	Diethyl ether	60-29-7	12	-116	35
	Bis(2-methoxyethyl) ether	111-96-6	835	-68	162
	Dimethyl ether	115-10-6	580	-141	-25
	1,4-Dioxane	123-91-1	75	12	102
	Tetrahydrofuran	109-99-9	159	-108	65
	1,2-Dimethoxyethane	110-71-4	213	-58	85
	Diisopropyl ether	108-20-3	96	-86	68
Dipolar Aprotic	Dimethylpropylene urea	7226-23-5	2191	-23	247
	Dimethyl sulphoxide	67-68-5	151	19	189
	Formamide	1975-12-02	701	3	220
	Dimethyl formamide	1968-12-02	72	-61	153
	N-Methylformamide	123-39-7	226	-4	200
	N-Methyl pyrrolidone	872-50-4	284	-24	202
	Propanenitrile	107-12-0	326	-93	97
	Dimethyl acetamide	127-19-5	227	-20	165
	Acetonitrile	1975-05-08	3	-45	82

Classification	Solvent	Cas Number	DDBST ID	Melting Point	Boiling Point
Halogenated	1,2-Dichlorobenzene	95-50-1	802	-17	180
	1,2,4-Trichlorobenzene	120-82-1	467	17	214
	Chlorobenzene	108-90-7	27	-45	132
	trichloroacetonitrile	545-06-2	3146	-42	83
	Chloroacetic acid	1979-11-08	295	61	189
	trichloroacetic acid	76-03-9	775	58	197
	Perfluorocyclohexane	355-68-0	715	51	53
	Carbon tetrachloride	56-23-5	157	-23	77
	Dichloromethane	1975-09-02	70	-95	40
	Perfluorohexane	355-42-0	466	-86	57
	Fluorobenzene	462-06-6	183	-42	85
	Perfluorotoluene	434-64-0	4681	-66	104
	Chloroform	67-66-3	47	-64	61
	Perfluorocyclic ether	335-36-4		-88	103
	Trifluoroacetic acid	27881		-15	72
	Trifluorotoluene	36015	169	-29	102
	1,2-Dichloroethane	107-06-2	68	-36	84
	2,2,2-Trifluoroethanol	75-89-8	1086	-43	74
Base	N,N-Dimethylaniline	121-69-7	63	3	194
	Triethylamine	121-44-8	162	-115	89
	Pyridine	110-86-1	144	-42	115
Other	Nitromethane	75-52-5	125	-29	101
	carbon disulfide	75-15-0	149	-111	46

APPENDIX D: CHEMCAD Unifac Groups

UNIFAC Group Specifications Subgroup Listing for CHEMCAD UNIFAC Models

The **VLE**, **LLE**, and **Do** columns represent **UNIFAC VLE**, **UNIFAC LLE**, and **Modified (Dortmund) UNIFAC**, respectively.

The *Subgroup number* is the number to assign to a component for the given subgroup.

Main Group	Subgroup	Subgroup number	VLE	LLE	Do	Example component	Groups for Example Component
CH2	CH3	1	X	X	X	butane	2 CH3, 2 CH2
	CH2	2	X	X	X	butane	2 CH3, 2 CH2
	CH	3	X	X	X	i-butane	3 CH3, 1 CH
	C	4	X	X	X	2,2-dimethylpropane	4 CH3, 1 C
	c-CH2	3095	*	-	X	cyclohexane	6 c-CH2
	c-CH	3100	*	-	X	methylcyclohexane	1 CH3, 5 c-CH2, 1 c-CH
	c-C	3105	*	-	X	1,1-dimethylcyclohexane	2 CH3, 5 c-CH2, 1 c-C
C=C	CH2=CH	5	X	X	X	1-hexene	1 CH3, 3 CH2, 1 CH2=CH
	CH=CH	6	X	X	X	2-hexene	2 CH3, 2 CH2, 1 CH=CH
	CH2=C	7	X	X	X	2-methyl-1- butene	2 CH3, 1 CH2, 1 CH2=C
	CH=C	8	X	X	X	2-methyl-2- butene	3 CH3, 1 CH=C
	C=C	9	X	*	X	2,3-dimethylbutene	4 CH3, 1 C=C
	C=C=C	3295	-	-	\$		
	=CHCH=	3300	-	-	\$		
	=CCH=	3305	-	-	\$		
ACH	ACH	10	X	X	X	benzene	6 ACH
	AC	11	X	X	X	styrene	1 CH2=CH, 5 ACH, 1 AC
ACCH2	ACCH3	12	X	X	X	toluene	5 ACH, 1 ACCH3
	ACCH2	13	X	X	X	ethylbenzene	1 CH3, 5 ACH, 1 ACCH2
	ACCH	14	X	X	X	cumene	2 CH3, 5 ACH, 1 ACCH
OH	OH	15	X	X	X	1-propanol	1 CH3, 2 CH2, 1 OH(p)
	OH(s)	3000	*	*	X	2-propanol	2 CH3, 1 CH, 1 OH(s)
	OH(t)	3005	*	*	X	tert-butanol	3 CH3, 1 C , 1 OH (t)
CH3OH	CH3OH	16	X	*	X	methanol	1 CH3OH
H2O	H2O	17	X	X	X	water	H2O
ACOH	ACOH	18	X	X	X	phenol	5 ACH, 1 ACOH
Ketone							
CH2CO	CH3CO	19	X	X	X	2-butanone	1 CH3, 1 CH2, 1 CH3CO
	CH2CO	20	X	X	X	3-pentanone	2 CH3, 1 CH2, 1 CH2CO
Aldehyde							

CHO	CHO	21	X	X	X	acetaldehyde	1 CH ₃ , 1 CHO
-----	-----	----	---	---	---	--------------	---------------------------

Esters							
CCOO	CH ₃ COO	22	X	X	X	butyl acetate	1 CH ₃ , 3 CH ₂ , 1 CH ₃ COO
	CH ₂ COO	23	X	X	X	butyl propanoate	2 CH ₃ , 3 CH ₂ , 1 CH ₂ COO
HCOO	HCOO	24	X	*	X	ethyl formate	1 CH ₃ , 1 CH ₂ , 1 HCOO
	CHCOO	144	\$	*	\$		
	CCOO	145	\$	*	\$		
Ether							
CH ₂ O	CH ₃ O	25	X	X	X	dimethyl ether	1 CH ₃ , 1 CH ₃ O
	CH ₂ O	26	X	X	X	diethyl ether	2 CH ₃ , 1 CH ₂ , 1 CH ₂ O
	CH-O	27	X	X	X	diisopropyl ether	4 CH ₃ , 1 CH, 1 CH-O
	fCH ₂ O	28	X	X	-	tetrahydrofuran	3 CH ₂ , 1 fCH ₂ O
Amine							
CNH ₂	CH ₃ NH ₂	29	X	*	X	methylamine	1 CH ₃ NH ₂
	CH ₂ NH ₂	30	X	*	X	n-propylamine	1CH ₃ , 1 CH ₂ , 1 CH ₂ NH ₂
	CHNH ₂	31	X	*	X	isopropylamine	2 CH ₃ , 1 CHNH ₂
	CNH ₂	3090	-	-	X	tert-butylamine	3 CH ₃ , 1 CNH ₂
CNH	CH ₃ NH	32	X	*	X	dimethylamine	1 CH ₃ , 1CH ₃ NH
	CH ₂ NH	33	X	*	X	diethylamine	2 CH ₃ , 1 CH ₂ , 1 CH ₂ NH
	CHNH	34	X	*	X	diisopropylamine	4 CH ₃ , 1 CH, 1CHNH
(C ₃)N	CH ₃ N	35	X	*	X	trimethylamine	2 CH ₃ , 1 CH ₃ N
	CH ₂ N	36	X	*	X	triethylamine	3 CH ₃ , 2 CH ₂ , 1 CH ₂ N
Tert-N	TERT-N	85	X	*	-	triethylamine	3 CH ₃ , 3 CH ₂ , 1 >N-
ACNH ₂	ACNH ₂	37	X	X	X	aniline	5 ACH, 1 ACNH ₂
(Pyridines)	C ₅ H ₅ N	38	X	X	-	pyridine	1 C ₅ H ₅ N
C ₅ H _n N	C ₅ H ₄ N	39	X	X	-	2-methylpyridine	1 CH ₃ , 1 C ₅ H ₄ N
	C ₅ H ₃ N	40	X	X	-	2,3-dimethylpyridine	2 CH ₃ , 1 C ₅ H ₃ N
Pyridine	AC ₂ H ₂ N	3010	-	-	X	pyridine	1 AC ₂ H ₂ N, 3 ACH
	AC ₂ HN	3015	-	-	X	2-methylpyridine	1 AC ₂ HN, 3 ACH, 1 CH ₃
	AC ₂ N	3020	-	-	X	2,5-dimethylpyridine	1 AC ₂ N, 3 ACH, 2 CH ₃
CCN	CH ₃ CN	41	X	X	X	acetonitrile	1 CH ₃ CN
	CH ₂ CN	42	X	X	X	propionitrile	1 CH ₃ , 1 CH ₂ CN
COOH	COOH	43	X	X	X	acetic acid	1 CH ₃ , 1 COOH
	HCOOH	44	X	X	X	formic acid	1 HCOOH
CCl	CH ₂ Cl	45	X	X	X	1-chlorobutane	1 CH ₃ , 2CH ₂ , 1CH ₂ Cl
	CHCl	46	X	X	X	2-chloro-propane	2 CH ₃ , 1 CHCl
	CCl	47	X	X	X	2-chloro-2-methyl propane	2 CH ₃ , 1 CCl
CCl ₂	CH ₂ Cl ₂	48	X	X	X	dichloromethane	1 CH ₂ Cl ₂
	CHCl ₂	49	X	X	X	1,1-dichloroethane	1 CH ₃ , 1 CCl ₂
	CCl ₂	50	X	X	X	2,2-dichloropropane	2 CH ₃ , 1 CCl ₂
CCl ₃	CHCl ₃	51	X	X	X	chloroform	1 CHCl ₃
	CCl ₃	52	X	X	X	1,1,1-trichloroethane	1 CH ₃ , 1 CCl ₃
CCl ₄	CCl ₄	53	X	X	X	carbon tetrachloride	1 CCl ₄

ACCl	ACCl	54	X	X	X	chlorobenzene	5 ACH, 1 ACCl
Cl(C=C)	Cl(C=C)	70	X	*	X	trichloroethene	1 CH=C, 3 Cl-(C=C)
CNO2	CH3NO2	55	X	X	X	nitromethane	1 CH3NO2
	CH2NO2	56	X	X	X	1-nitropropane	1 CH3, 1 CH2, 1 CH2NO2
	CHNO2	57	X	X	X	2-nitropropane	2 CH3, 1 CHNO2
ACNO2	ACNO2	58	X	X	X	nitro-benzene	5 ACH, 1 ACNO2
CS2	CS2	59	X	*	X	carbon disulfide	1 CS2
CH3SH	CH3SH	60	X	*	X	methanethiol	1 CH3SH
	CH2SH	61	X	*	X	ethanethiol	1 CH3, 1 CH2SH
furfural	furfural	62	X	X	X	furfural	1 furfural
DOH	(CH2OH)2	63	X	X	X	1,2-ethanediol	1 (CH2OH)2
I	I	64	X	*	X	iodoethane	1 CH3, 1 CH2, 1 I
Br	Br	65	X	*	X	bromomethane	1 CH3, 1 Br
C<->C	CH<->C	66	X	*	X	1-hexyne	1 CH3, 3 CH2, 1 CH<->C
	C<->C	67	X	*	X	2-hexyne	2 CH3, 2 CH2, 1 C<->C
DMSO	(CH3)2SO	68	X	X	X	Dimethyl sulfoxide	1 (CH3)2SO
Acrylonitrile	acrylonitrile	69	X	*	X	acrylonitrile	1 acrylonitrile
ACF	ACF	71	X	*	X	hexafluorobenzene	6 ACF
DMF	DMF	72	X	X	X	N,N-dimethylformamide	1 DMF
	HCON(CH2)2	73	X	*	X	N,N-diethylformamide	2 CH3, 1 HCON(CH2)2
CF2	CF3	74	X	*	X	perfluorohexane	2 CF3, 4 CF2
	CF2	75	X	*	X	perfluorohexane	2 CF3, 4 CF2
	CF	76	X	*	X	perfluoromethylcyclohexane	1 CF3, 5 CF2, 1 CF
COO	COO	77	X	*	X	methyl acrylate	1 CH3, 1CH2=CH, 1 COO
c-CH2O	c-CH2O[CH2]½	3075	-	-	X	1,3-dioxane	1 c-CH2, 2 c-CH2O[CH2](1/2)
	c-[CH2]½O[CH2]½	3080	-	-	X	1,3,5-trioxane	3 c-[CH2]1/2O[CH2]1/2
	c-CH2OCH2	3085	-	-	X	tetrahydrofuran	2 c-CH2, 1 c-CH2OCH2
SiH2	SiH3	78	X	*	-	methylsilane	1 CH3, 1 SiH3
	SiH2	79	X	*	-	diethylsilane	2 CH3, 2 CH2, 1 SiH2
	SiH	80	X	*	-	heptamethyltrisiloxane	7 CH3, 2 SiO, 1 SiH
	Si	81	X	*	-	hexamethyldisiloxane	6 CH3, 1 SiO, 1 Si
SiO	SiH2O	82	X	*	-	1,3-dimethyldisiloxane	2 CH3, 1 SiH2O, 1 SiH2
	SiHO	83	X	*	-	1,1,3,3-tetramethyldisiloxane	4 CH3, 1 SiHO, 1 SiH
	SiO	84	X	*	-	octamethylcyclotetrasiloxane	8 CH3, 4 SiO
Chlorofluorocarbons							
	CCl3F	86	X	*	\$	trichlorofluoromethane	1 CCl3F
	CCl2F	87	X	*	\$	tetrachloro-1,2-difluoroethane	2 CCl2F
	HCCl2F	88	X	*	\$	dichlorofluoromethane	1 HCCl2F
	HCClF	89	X	*	\$	1-chloro-1,2,2,2-tetrafluoroethane	1 CF3, 1 HCClF

	CCIF2	90	X	*	\$	1,2-dichlorotetrafluoroethane	2 CCIF2
	HCCIF2	91	X	*	\$	chlorodifluoromethane	1 HCCIF2
	CCIF3	92	X	*	\$	chlorotrifluoromethane	1 CCIF3
	CCI2F2	93	X	*	\$	dichlorodifluoromethane	1 CCI2F2
Amide	CONH2	94	X	*	\$	acetamide	1 CH3, 1 CONH2
CONMeCH2	CONHCH3	95	X	*	X	N-methylacetamide	1 CH3, 1 CONHCH3
	CONHCH2	96	X	*	X	N-ethylacetamide	2 CH3, 1 CONHCH2
	CONHC	3183	-	-	\$	N-tert-Butyl-Acetamide	4 CH3, 1 CONHC
CONR2	CON(CH3)2	97	X	*	X	N,N-dimethylacetamide	1 CH3, 1 CON(CH3)2
	CONCH3CH2	98	X	*	X	N,N-methylethylacetamide	2 CH3, 1 CONCH3CH2
	CON(CH2)2	99	X	*	X	N,N-diethylacetamide	3 CH3, 1 CON(CH2)2
NMP	NMP	109	X	*	-	N-methylpyrrolidone	1 NMP
Pyrrolidone	cy-CON-CH3	3055	-	-	X	N-methylpyrrolidone	1 cy-CON-CH3, 3 cy-CH2
cy-CONC	cy-CON-CH2	3060	-	-	X	N-ethylpyrrolidone	1 cy-CON-CH2, 3 cy-CH2, 1 CH3
	cy-CON-CH	3065	-	-	X	N-isopropylpyrrolidone	1 cy-CON-CH, 3 cy-CH2, 2 CH3
	cy-CON-C	3070	-	-	X	N-tert-butylpyrrolidone	1 cy-CON-C, 3 cy-CH2, 3 CH3
Ethoxy	C2H5O2	100	X	*	\$	2-ethoxyethanol	1 CH3, 1 CH2, 1 C2H5O2
OCCOH	C2H4O2	101	X	*	\$	2-ethoxy-1-propanol	2 CH3, 1 CH2, 1 C2H4O2
CH2S	CH3S	102	X	*	\$	dimethylsulfide	1 CH3, 1 CH3S
	CH2S	103	X	*	\$	diethylsulfide	2 CH3, 1 CH2, 1 CH2S
	CHS	104	X	*	\$	diisopropylsulfide	4 CH3, 1 CH, 1 CHS
Morpholine	MORPH	105	X	*	-	morpholine	1 Morph
Thiophene	C4H4S	106	X	*	-	thiophene	1 C4H4S
(CS)	C4H3S	107	X	*	-	2-methylthiophene	1 CH3, 1 C4H3S
	C4H2S	108	X	*	-	2,3-dimethylthiophene	2 CH3, 1 C4H2S
NCO	NCO	1109	\$	-	\$	Butylisocyanate	1 CH3, 2 CH2, 1 NCO
Epoxide	H2COCH	110	X	*	X	propylene oxide	1 H2COCH, 1 CH3
	H2COC	131	X	-	-	2-methyl propylene oxide	1 H2COC, 2 CH3
	HCOCH	111	X	*	X	2,3-epoxybutane	1 HCOCH, 2 CH3
	HCOC	112	X	*	\$	2-methyl, 2,3-butylene oxide	1 HCOC, 3 CH3
	H2COCH2	113	X	*	\$	ethylene oxide	1 H2COCH2
Thiophene	AC2H2S	3040	-	-	X	thiophene	2 ACH, 1 AC2H2S
(ACS)	AC2HS	3045	-	-	X	2-methylthiophene	1 CH3, 2 ACH, 1 AC2HS
	AC2S	3050	-	-	X	2,5-dimethylthiophene	2 CH3, 2 ACH, 1 AC2S
Anhydrides	OCOCO	114	\$	-	\$	acetic anhydride	1 OCOCO, 2 CH3
Carbonates	(CH3O2)2CO	3025	\$	-	X	dimethylcarbonate	(CH3O)2CO
	(CH2O2)2CO	3030	\$	-	X	diethylcarbonate	1 (CH2O)2CO, 2 CH3
	CH2OCH3OCO	3035	\$	-	X	methyl-ethyl-carbonate	1 CH2OCH3OCO, 1 CH3
	CHOCH2OCO	120	\$	-	\$	Ethyl-Isopropyl-Carbonate	1 CHOCH2OCO, 3 CH3
Sulfones	(CH2)2Su	118	\$	-	\$	sulfolane	1 (CH2)2SU, 2 CH2
	CH2SuCH	119	\$	-	\$	2,4-dimethylsulfolane	1 CH2SuCH, 2 CH3,
							1 CH2, 1 CH

HCONR	HCONHCH3	121	\$	-	\$	N-Methyl-formamide	1 HCONHCH3
	HCONHCH2	122	\$	-	\$	N-Ethyl-formamide	1 CH3, 1 HCONHCH2
ACCN	ACCN	123	\$	-	\$	Benzonitrile	5 ACH, 1 ACCN
cy-CONH	cy-CONH	124	\$	-	\$	e-Caprolactam	4 CH2, 1 cy-CONH
Lactone	cy-COO-C	125	\$	-	\$	g-Butyrolactone	2 CH2, 1 cy-COO-C
peroxide	-O-O-	126	\$	-	\$	Di-Tert-Butylperoxide	6 CH2, 2 C, 1 -O-O-
	-O-OH	127	\$	-	\$	Tert-Butylhydroperoxide	3 CH2, 1 C, 1 -O-OH
Acetals	O-CH2-O	128	\$	-	\$	Dimethoxymethane	2 CH3, 1 O-CH2-O
	O-CH-O	129	\$	-	\$	1,1-Dimethoxyethane	3 CH3, 1 O-CH-O
	O-C-O	130	\$	-	\$	2,2-Dimethoxypropane	4 CH3, 1 O-C-O
Aniline	ACN(CH3)2	132	\$	-	\$	N,N-Dimethylaniline	5 ACH, 1 ACN(CH3)2
	ACN(CH2)2	133	\$	-	\$	N,N-Diethylaniline	2 CH3, 5 ACH,
							1 ACN(CH2)2
	ACNCH3CH2	134	\$	-	\$	N-Ethyl-N-methylaniline	1 CH3, 5 ACH,
							1 ACNCH3CH2
	ACNHCH3	135	\$	-	\$	N-Methylanilin	5 ACH, 1 ACNHCH3
	ACNHCH2	136	\$	-	\$	N-Ethylanilin	1 CH3, 5 ACH,
							1 ACNHCH2
ACBr	ACBr	137	\$	-	-	Brombenzene	5 ACH, 1 ACBr
Oxime	HCNOH	138	\$	-	\$	Propionaldehydoxime	1 HCNOH, 1 CH3, 1 CH2
	CNOH	139	\$	-	\$	Acetoneoxime	1 CNOH, 2 CH3
ACCHO	ACCHO	3200	-	-	\$	Benzaldehyde	1 ACCHO, 5 ACH
ACCOOH	ACCOOH	3205	-	-	\$	Benzoic Acid	1 ACCOOH, 5 ACH
ACCOO	ACCOO	3210	-	-	\$	Benzylbenzoate	1 ACCOO, 1 ACCH2, 10 ACH
ACCO	ACCOCH3	3285	-	-	\$		
	ACCOCH2	3290	-	-	\$		
CFH	CFH3	3215	-	-	\$	R 41	1 CFH3
	CFH2	3220	-	-	\$	R 161	1 CH3, 1 CFH2
	CFH	3225	-	-	\$	R225BB	1 CFCl, 1 CF2H, 1 CF2Cl
	CFCIH2	3230	-	-	\$	R 31	1 CFCIH2
	CFCI	3235	-	-	\$	R225BB	1 CFCl, 1 CF2H, 1 CF2Cl
	CF2H2	3240	-	-	\$	R 32	1 CF2H2
	CF2H	3245	-	-	\$	R 134	2 CF2H
CF3H	CF3H	3250	-	-	\$	R 23	1 CF3H
	(CH3)-CF3	3260	-	-	\$	R143A	1 CH3, 1 (CH3)-CF3
CF4	CF4	3255	-	-	\$	R 14	CF4
CF3Cl	CF2ClBr	3265	-	-	\$	R13B1	1 CF3Br
	CF2ClBr	3270	-	-	\$	R12B1	1 CF2ClBr
Furane	AC2H2O	140	\$	-	\$	Furan	2 ACH, 1 AC2H2O
	AC2HO	141	\$	-	\$	2-Methylfuran	1 CH3, 2 ACH, 1 AC2HO
	AC2O	142	\$	-	\$	2,5-Dimethylfuran	2 CH3, 2 ACH, 1 AC2O
c- Amine	c-CH2NH	162	-	-	\$	Pyrrolidine	1 c-CH2NH, 3 c-CH2
	c-CHNH	163	-	-	\$	2-Methylpiperidine	1 c-CHNH, 1 CH3, 4 c-CH2
	c-CN	164	-	-	\$	2,2-Dimethylpiperidine	1 c-CN, 2 CH3, 4 c-CH2
	c-CNCH3	165	-	-	\$	N-Methylpyrrolidine	1 c-CNCH3, 3 c-CH2

	c-CNCH ₂	166	-	-	\$	N-Ethylpiperidine	c-CNCH ₂ , 1 CH ₃ , 5 c-CH ₂
	c-CNCH	167	-	-	\$	N-Isopropylpiperidine	1 c-CNCH, 2 CH ₃ , 5 c-CH ₂
	c-CNC	168	-	-	\$	N-tert-Butylpyrrolidine	1 c-CNC, 3 CH ₃ , 5 c-CH ₂

Additional Subgroups for UNIFAC LLE Model							
Main Group	Subgroup		VLE	LLE	Do	Example component	
P1	P1 (1-propanol)	501	-	X	-	1-propanol	1 P1
P2	P2 (2-propanol)	502	-	X	-	2-propanol	2 P2
DEOH	(HOCH ₂ CH ₂) ₂ O	503	-	X	-	diethylene glycol	1 (HOCH ₂ CH ₂) ₂ O
TCE	CCl ₂ =CHCl	504	-	X	-	trichloroethylene	1 CCl ₂ =CHCl
MFA	HCONHCH ₃	505	-	X	-	methylformamide	1 HCONHCH ₃
TMS	1 (CH ₂) ₂ SO	506	-	X	-	tetramethylenesulfone	1 (CH ₂) ₂ SO

Additional Subgroups for PSRK Model							
Main Group	Subgroup	Subgroup number	VLE	LLE	Do	Example component	Groups for Example Component
CO ₂	CO ₂	1001	X	-	-	carbon dioxide	1 CO ₂
CH ₄	CH ₄	1002	X	-	-	methane	1 CH ₄
N ₂	N ₂	1003	X	-	-	nitrogen	1 N ₂
H ₂ S	H ₂ S	1004	X	-	-	hydrogen sulfide	1 H ₂ S
H ₂	H ₂	1005	X	-	-	hydrogen	1 H ₂
CO	CO	1006	X	-	-	carbon monoxide	1 CO
H ₂ C=CH ₂	H ₂ C=CH ₂	1007	X	-	-	ethene	1 CH=CH
CH ^o CH	CH ^o CH	1008	X	-	-	ethyne	1 CH<->CH
NH ₃	NH ₃	1009	X	-	-	ammonia	1 NH ₃
Ar	Ar	1010	X	-	-	argon	1 Ar
O ₂	O ₂	1011	X	-	-	oxygen	1 O ₂
SO ₂	SO ₂	1012	X	-	-	sulfur dioxide	1 SO ₂
NO	NO	1013	X	-	-	nitric oxide	1 NO
N ₂ O	N ₂ O	1014	X	-	-	dinitrogen monoxide	1 N ₂ O
SF ₆	SF ₆	1015	X	-	-	sulfur hexafluoride	1 SF ₆
He	He	1016	X	-	-	helium	1 He
Ne	Ne	1017	X	-	-	neon	1 Ne
Kr	Kr	1018	X	-	-	krypton	1 Kr
Xe	Xe	1019	X	-	-	xenon	1 Xe
HCl	HCl	1020	X	-	-	hydrogen chloride	1 HCl
HBr	HBr	1021	X	-	-	hydrogen bromide	1 HBr
CHSH	CHSH	1022	X	-	-	iso-propyl mercaptan	1 CHSH, 2 CH ₃
CSH	CSH	1023	X	-	-	tert-butyl mercaptan	1 CSH, 3 CH ₃
COC	COC	1025	X	-	-	2,3-dimethyl 2,3 butylene oxide	1 COC, 4 CH ₃

HF	HF	1026	X	-	-	hydrogen fluoride	1 HF
HI	HI	1027	X	-	-	hydrogen iodide	1 HI
COS	COS	1028	X	-	-	carbonyl sulfide	1 COS

Notes

- UNIFAC / UNIFAC LLE subgroups 3000 3005 are identical to OH.
- 3095, 3100, and 3105 are identical to 2, 3, and 4, respectively CH₂, CH, C

Legend

X	CHEMCAD has data for this subgroup.
\$	The subgroup is available to UNIFAC Consortium members in a supplement to CHEMCAD.
*	The optimized subgroup is not specified for this model. UNIFAC subgroup will be used as a default.
-	The subgroup does not exist in the model.
cy	Denotes a cyclic hydrocarbon.
<->	Denotes a triple bond.
A	Indicates an aromatic ring.
R	Indicates a hydrocarbon branch.
Me	Indicates a methyl group (-CH ₃).

APPENDIX E: Sample Code

Example of Sub-Routine Code of the Solvent Selection Tool

The Solvent Selection Tool is developed as a VBA module and embedded into the commercial process simulation software CHEMCAD. The Solvent selection Tool consist of 36 subroutines made up of 150 pages of VBA code.

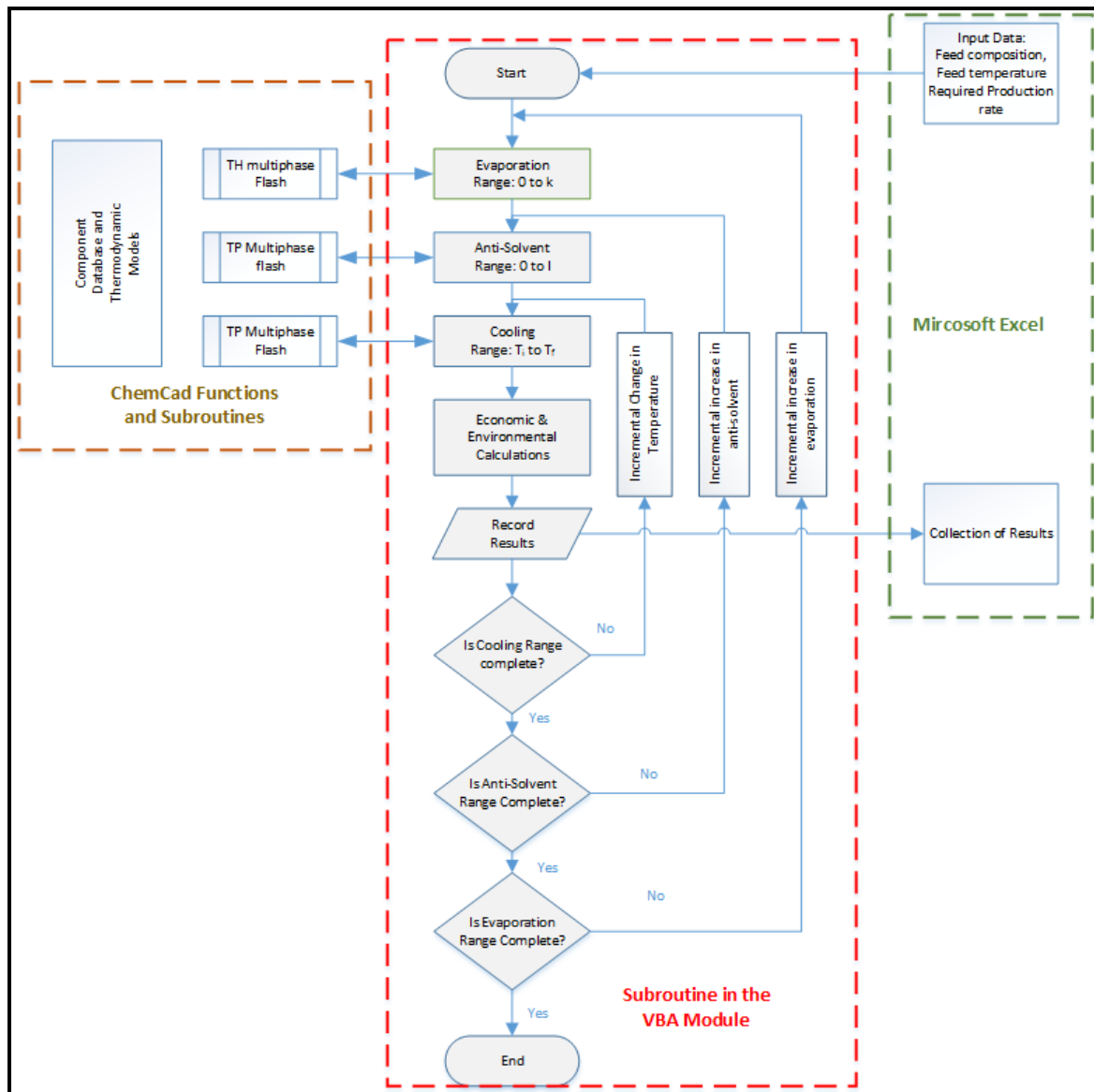
To illustrate the interface and the two way communication between Excel – VBA Module and CHEMCAD, some examples of code statements and their functions in the various sub-routines of the VBA module) are presented in Table E.1. These sample coding statements (note that this is not a full subroutine.)

Table E.1. VBA code showing a technical and semantic bridge between Excel, the VBA module and CHEMCAD.

Function	Typical Statements in the VBA module
<u>Activating CHEMCAD from VBA Module:</u> Requesting access to functions from CHEMCAD, for example: Stream information, pure component data, flash calculations, and thermodynamic models for phase calculations respectively.	Set strinfo = ccentry.GetStreamInfo Set objpp = ccentry.GetCompPPData Set fl = ccentry.GetFlash Set kval = ccentry.GetkValues
<u>Activating MS Excel from VBA Module:</u> Activates an Excel workbook and label a sheet. Also used for creating input data menus in Excel	Set wb = xls.workbooks.add Set ws = wb.Sheets("Sheet1") Xls.Visible = true
<u>Input of relevant data from Excel into VBA module:</u> Read the feed conditions from the Excel sheet. For eg. the flowrate and Temperature from cell (2,9) and cell (4,9) respectively	xCC(sol) = ws.cells(2,9) tempcel = ws.cells(4,9)
<u>Accessing CHEMCAD Algorithms:</u> <u>Handover Input file:</u> Define the feed stream for CHEMCAD from the values obtained from Excel sheet1. <u>Execute Calculation:</u> Request CHEMCAD to perform a TP flash calculation on the feed-stream. <u>Handover output file:</u> Request the following results on the liquid stream out from the flash calculation performed: the temperature, the enthalpy and the total mole flowrate.	Calcfl = fl.defineFeedStream(Temprank, pressin, 0, xCC) Calcfl = fl.calculateTPFlash(temprank, pressin) Istream = fl.getliquidStream(tempoutLiq, enthoutLiq, totmolrateoutliq)
<u>Output of VBA Module results into Excel:</u> Write results obtained from CHEMCAD flash calculations and VBA module calculations into the Excel worksheet. The data obtained from the simulation is further ranked to consider only those solutions which meet the following criteria: The yield is \geq the minimum yield specified by the user, and conditions where no by-products or solvent have also crystallised.	For i = 1 To maxrow If (wb.ActiveSheet.Cells(6 + i, 6) >= yield And wb.ActiveSheet.Cells(6 + i, 7) = 0 And wb.ActiveSheet.Cells(6 + i, 8) = 0) Then ws.Cells(n + 6, 10) = wb.ActiveSheet.Cells(6 + i, 1) ws.Cells(n + 6, 11) = wb.ActiveSheet.Cells(6 + i, 2)

An example of the sub-routine called: Sub CoolEvapAddAntisol()

This subroutine is activated by the "Run Maximum Combination" button in the anti-solvent functions menu, and it evaluates the effects of anti-solvent addition, evaporation of solvent and cooling combinations. The architecture of this sub-routine is presented in the diagram below.



Note: That 3 other subroutines that first activates CHEMCAD, creates an excel spreadsheet for input of feed conditions, and inputs data into CHEMCAD are not shown here.

Sub CoolEvapAddAntisol()

Dim calcfl As Integer

Dim lstream As Integer

Dim xCCOutAnti() As Single

Dim xCCOutEvap() As Single

Dim xCCAntisolCombi() As Single

Dim pressIn As Single

Dim tempOutLiq As Single

Dim tempOutvap As Single

Dim tempRoom As Single

Dim pressOutLiq As Single

Dim pressOutVap As Single

Dim enthoutliq As Single

Dim enthoutvap As Single

Dim totMolRateOutLiq As Single

Dim totMolRateOutVap As Single

Dim flowrateOutLiq() As Single

Dim flowrateOutVap() As Single

Dim incrementEvap As Single

Dim incrementCool As Single

Dim incrementAddAntisol As Single

Dim amountEvap As Single

Dim tempCool As Single

Dim bestResult As Single

Dim bestI As Integer

Dim bestK As Integer

Dim tempDiffCool As Single

Dim enthIn As Single

Dim enthStart As Single

Dim sumFlowrateLiq As Single

Dim sumFlowrateVap As Single

Dim flowrateRatio As Single

Dim amountAntiSol As Single


```

Dim i As Integer
Dim p As Integer
Dim k As Integer
Dim j As Integer
Dim h As Integer
Dim l As Integer
Dim m As Integer
Dim n As Integer
Dim checkSolvent As Single
Dim checkByProducts As Single
Dim there As Boolean
there = False

```

'Open a worksheet in excel to record results of simulation

```

For Each ws In Worksheets
    If ws.Name = "CoolEvapAnti-solvent" Then
        there = True
    End If
Next ws
If there = True Then
    Worksheets("CoolEvapAnti-solvent").Activate
Else
    ActiveWorkbook.Worksheets.Add.Name = "CoolEvapAnti-solvent"
    Worksheets("CoolEvapAnti-solvent").Activate
End If
Set ws = ActiveSheet

```

'Define inserted properties: Range and intervals for all calculations

```

Dim upperLimitEvap As Single
upperLimitEvap = 1 - (Anti-solventForm.TBUpperEvap.Value / 100)
Dim lowerLimitEvap As Single
lowerLimitEvap = 1 - (Anti-solventForm.TBLowerEvap.Value / 100)

Dim upperLimitTempCool As Single

```

```

upperLimitTempCool = (Anti-solventForm.TBUpperCool.Value * 1.8) + 491.67
Dim lowerLimitTempCool As Single
lowerLimitTempCool = (Anti-solventForm.TBLowerCool.Value * 1.8) + 491.67
Dim upperLimitAddAntisol As Single
upperLimitAddAntisol = Anti-solventForm.TBUpperAntisol.Value
Dim lowerLimitAddAntisol As Single
lowerLimitAddAntisol = Anti-solventForm.TBLowerAntisol.Value
Dim numPointsCool As Integer
numPointsCool = Anti-solventForm.TBNStepCool.Value - 1
Dim numPointsEvap As Integer
numPointsEvap = Anti-solventForm.TBNStepEvap.Value - 1
Dim numPointsAntisol As Integer
numPointsAntisol = Anti-solventForm.TBNStepsAntisol.Value - 1
incrementEvap = (upperLimitEvap - lowerLimitEvap) / numPointsEvap
incrementCool = (upperLimitTempCool - lowerLimitTempCool) / numPointsCool
incrementAddAntisol = (upperLimitAddAntisol - lowerLimitAddAntisol) / numPointsAntisol

```

Call Subroutine to select: Solute, Solvent and Anti-solvent from drop down menus

```

Call chooseSolvent
ReDim xCCOutEvap(nges)
ReDim xCCOutAnti(nges)
ReDim xCCAntisolCombi(nges)
ReDim flowrateOutLiq(nges)
ReDim flowrateOutVap(nges)

```

Defining Normal pressure (1 bar) in psia and temperature in Rankine

```

pressIn = 14.5047
tempRoom = 536.67

```

Defining initial temperature, amount of solvent evaporated and amount of anti-solvent added

```

tempCool = lowerLimitTempCool
amountEvap = lowerLimitEvap
amountAntiSol = lowerLimitAddAntisol
m = 1

```

'read composition of product and byproducts

```
For p = ncomp + 1 To nges
    xCCAntisolCombi(p) = xCC(p)
Next p
```

'define solvent and anti-solvent

```
xCCAntisolCombi(possol) = wb.Sheets("sheet1").Cells(4, 9)
```

'label excel results

```
ws.Cells(1, 1) = "Results for combination of previous solvent evaporation, adding amount of anti-  
solvent and cooling down"  
ws.Cells(3, 1) = "chosen solvent"  
ws.Cells(3, 4) = compname(possol)  
ws.Cells(4, 1) = "chosen anti-solvent"  
ws.Cells(4, 4) = compname(posAntisol)  
ws.Cells(6, 1) = "Amount of evaporation (%)"  
ws.Cells(6, 2) = "Amount of added anti-solvent(kmol/hr)"  
ws.Cells(6, 3) = "temperature (C)"  
ws.Cells(6, 4) = "liquid product (kmol/hr)"  
ws.Cells(6, 5) = "solid product (kmol/hr)"  
ws.Cells(6, 6) = "amount of crystallization (%)"  
ws.Cells(6, 7) = "solvent crystallized (kmol/hr)"  
ws.Cells(6, 8) = "by-product crystallized (kmol/hr)"
```

'Set start enthalpy for solvent evaporation calculations

'Instruction to CHEMCAD to perform flash calculations on defined streams

```
calcfl = fl.DefineFeedStream(tempRank, pressIn, 0, xCCAntisolCombi)  
calcfl = fl.CalculateTPFlash(tempRank, pressIn)
```

'Requesting the following calculated values from CHEMCAD calculation

```
Istream = fl.GetLiquidStream(tempOutLiq, pressOutLiq, enthoutliq, totMolRateOutLiq,  
flowrateOutLiq)  
Istream = fl.GetVaporStream(tempOutvap, pressOutVap, enthoutvap, totMolRateOutVap,  
flowrateOutVap)  
enthStart = enthoutliq + enthoutvap  
enthIn = enthStart  
flowrateRatio = 1
```

'Calculate datapoints and start evaporation calculations

```
For k = 0 To numPointsEvap
```

```

ReDim xCCOutEvap(nges)
ReDim xCCAntisolCombi(nges)
ReDim xCCOutAnti(nges)
' read composition of product and byproducts
For p = ncomp + 1 To nges
    xCCAntisolCombi(p) = xCC(p)
Next p
'define solvent and anti-solvent
xCCAntisolCombi(possol) = wb.Sheets("sheet1").Cells(4, 9)
amountEvap = lowerLimitEvap + k * incrementEvap
enthIn = enthStart
xCCAntisolCombi(posAntisol) = 0
While amountEvap <= flowrateRatio
    ' Increase of added enthalpy nearly 1 MJ/h
    sumFlowrateLiq = 0
    sumFlowrateVap = 0
    enthIn = enthIn + 947.8171 ' nearly 1 MJ/h
    calcfl = fl.DefineFeedStream(tempRoom, pressIn, 0, xCCAntisolCombi)
    calcfl = fl.CalculateHPFlash(enthIn, pressIn)
    'Check if Flash has not converged
    If calcfl = 1 Then
        MsgBox "not converged"
    End If
    Istream = fl.GetLiquidStream(tempOutLiq, pressOutLiq, enthoutliq, totMolRateOutLiq,
    flowrateOutLiq)
    Istream = fl.GetVaporStream(tempOutvap, pressOutVap, enthoutvap, totMolRateOutVap,
    flowrateOutVap)
'Insert second loop for any additional solvents
    For h = 1 To ncomp / 2
        sumFlowrateLiq = sumFlowrateLiq + flowrateOutLiq(h)
        sumFlowrateVap = sumFlowrateVap + flowrateOutVap(h)
    Next h

    For h = ncomp + 1 To ncomp + numAddSol
        sumFlowrateLiq = sumFlowrateLiq + flowrateOutLiq(h)

```

```

        sumFlowrateVap = sumFlowrateVap + flowrateOutVap(h)
    Next h
    If Not (sumFlowrateLiq = 0 And sumFlowrateVap = 0) Then
        flowrateRatio = sumFlowrateLiq / (sumFlowrateLiq + sumFlowrateVap)
    End If
Wend

```

'Anti-Solvent Addition

```

For j = 1 To nges
    xCCOutEvap(j) = flowrateOutLiq(j)
Next i
For l = 0 To numPointsAntisol
    amountAntiSol = lowerLimitAddAntisol + l * incrementAddAntisol
    xCCOutEvap(posAntisol) = amountAntiSol
    For j = 1 To nges
        xCCOutAnti(j) = flowrateOutLiq(j)
    Next j

```

'Cooling Calculations loop

```

tempCool = lowerLimitTempCool
For i = 0 To numPointsCool
    calcfl = fl.DefineFeedStream(tempCool, pressIn, 0, xCCOutAnti)
    calcfl = fl.CalculateTPFlash(tempCool, pressIn)

    'Check if Flash has not converged
    If calcfl = 1 Then
        MsgBox "not converged"
    End If

    lstream = fl.GetLiquidStream(tempOutLiq, pressOutLiq, enthoutliq, totMolRateOutLiq, flowrateOutLiq)
    lstream = fl.GetVaporStream(tempOutvap, pressOutVap, enthoutvap, totMolRateOutVap, flowrateOutVap)

    ws.Cells(6 + m, 1) = (1 - amountEvap) * 100
    ws.Cells(6 + m, 2) = amountAntiSol
    ws.Cells(6 + m, 3) = (tempCool - 491.67) / 1.8
    ws.Cells(6 + m, 4) = flowrateOutLiq(nges - 1)
    ws.Cells(6 + m, 5) = flowrateOutLiq(nges)

```

ws.Cells(6 + m, 6) = (flowrateOutLiq(nges) / (flowrateOutLiq(nges) + flowrateOutLiq(nges - 1))) * 100

“Mass balance check and recording of final results”

checkSolvent = 0

If numAddSol <> 0 Then

 If flowrateOutLiq(possol + ncomp / 2) <> 0 Then

 checkSolvent = checkSolvent + flowrateOutLiq(possol + ncomp / 2)

 End If

Else

 checkSolvent = checkSolvent + flowrateOutLiq(possol + numAddSol)

End If

If numAddSol <> 0 Then

 If flowrateOutLiq(posAntisol + ncomp / 2) <> 0 Then

 checkSolvent = checkSolvent + flowrateOutLiq(posAntisol + ncomp / 2)

 End If

Else

 checkSolvent = checkSolvent + flowrateOutLiq(posAntisol + numAddSol)

End If

If checkSolvent = 0 Then

 ws.Cells(6 + m, 7) = 0

Else

 ws.Cells(6 + m, 7) = checkSolvent

End If

checkByProducts = 0

For h = ncomp + 2 * numAddSol + numAddOthers + 1 To ncomp + 2 * numAddSol + 2 * numAddOthers

 If flowrateOutLiq(h) <> 0 Then

 checkByProducts = checkByProducts + flowrateOutLiq(h)

 ws.Cells(6 + m, 8) = checkByProducts

Else

 ws.Cells(6 + m, 8) = 0

End If

```

Next h
m = m + 1
tempCool = tempCool + incrementCool
Next i
Next l
Next k

```

```

Dim maxrow As Integer
maxrow = m
'fit columns
Columns("A:A").ColumnWidth = 25
Columns("B:B").EntireColumn.AutoFit
Columns("C:C").EntireColumn.AutoFit
Columns("D:D").EntireColumn.AutoFit
Columns("E:E").EntireColumn.AutoFit
Columns("F:F").EntireColumn.AutoFit
Columns("G:G").EntireColumn.AutoFit
Columns("H:H").EntireColumn.AutoFit

```

'search for results that meet user defined criterion

```

Dim yield As Single
yield = Anti-solventForm.TBYield.Value
n = 1
ws.Cells(4, 10) = "results with required yield and without impurities"
ws.Cells(6, 10) = "Amount of evaporation (%)"
ws.Cells(6, 11) = "Amount of added anti-solvent (kmol/hr)"
ws.Cells(6, 12) = "temperature (C)"
ws.Cells(6, 13) = "liquid product (kmol/hr)"
ws.Cells(6, 14) = "solid product (kmol/hr)"
ws.Cells(6, 15) = "amount of cristallization (%)"
ws.Cells(6, 16) = "solvent crystallized (kmol/hr)"
ws.Cells(6, 17) = "by-product crystallized (kmol/hr)"

i = 1

```

For i = 1 To maxrow

**If (wb.ActiveSheet.Cells(6 + i, 6) >= yield And wb.ActiveSheet.Cells(6 + i, 7) = 0 And
wb.ActiveSheet.Cells(6 + i, 8) = 0) Then**

ws.Cells(n + 6, 10) = wb.ActiveSheet.Cells(6 + i, 1)

ws.Cells(n + 6, 11) = wb.ActiveSheet.Cells(6 + i, 2)

ws.Cells(n + 6, 12) = wb.ActiveSheet.Cells(6 + i, 3)

ws.Cells(n + 6, 13) = wb.ActiveSheet.Cells(6 + i, 4)

ws.Cells(n + 6, 14) = wb.ActiveSheet.Cells(6 + i, 5)

ws.Cells(n + 6, 15) = wb.ActiveSheet.Cells(6 + i, 6)

ws.Cells(n + 6, 16) = wb.ActiveSheet.Cells(6 + i, 7)

ws.Cells(n + 6, 17) = wb.ActiveSheet.Cells(6 + i, 8)

n = n + 1

End If

Next i

'fit columns

Columns("J:J").ColumnWidth = 21.86

Columns("K:K").EntireColumn.AutoFit

Columns("L:L").EntireColumn.AutoFit

Columns("M:M").EntireColumn.AutoFit

Columns("N:N").EntireColumn.AutoFit

Columns("O:O").EntireColumn.AutoFit

Columns("P:P").EntireColumn.AutoFit

Columns("Q:Q").EntireColumn.AutoFit

MsgBox ("Calculation finished")

End Sub

Appendix F: Publication (submitted)

Development of a Computational Tool for the Analysis and Synthesis of Crystallization Processes.

Pascal Böwer¹; Thomas Teusch¹; Suresh Ramsuroop²; Jürgen Rarey^{1,3,4}; Deresh Ramjugernath³

¹Carl von Ossietzky University of Oldenburg, Germany.

²Durban University of Technology, South Africa.

³University of KwaZulu Natal, South Africa.

⁴DDBST GmbH, Oldenburg, Germany

Abstract

A computational tool integrated into a commercial simulation software (CHEMCAD) is developed for the analysis and synthesis of crystallization processes. The tool utilizes the comprehensive thermodynamic models and the rigorous computational algorithms available. The crystallization calculations are formulated as multicomponent multiphase equilibrium phase calculations and support the analysis of various modes of crystallization such as cooling, evaporative and the use of mass separating agents such as anti-solvents and cosolvents. The tool is demonstrated with applications related to the crystallization of API's and LCD crystals.

Introduction

With the general availability of computational software for the design and simulation of chemical processes, rapid evaluation of process alternatives and the influence of different parameters on performance and feasibility is readily available, especially for distillation, gas stripping and scrubbing, extraction, etc. in large continuous processes. These calculations heavily rely on the availability of reliable thermodynamic models and their parameters for the description of the pure component and mixture behaviours of the components involved. Parameters are usually regressed to experimental data found in large electronic databases (e.g. DDB [1], NIST [2]). Due to the large number of possible binary combinations of compounds, the required binary interaction parameters (BIP) are especially difficult to obtain. In order to apply these methods to varying or new processes, therefore predictive methods like UNIFAC [3], mod. UNIFAC [4], PSRK [5], VTPR [6] and COSMO-RS [7] or COSMO-SAC [8] [9] have been developed and are mostly available in process simulation software. Following the success of these methods, new variations were recently developed with a special view on pharmaceuticals and pharmaceutical intermediates (NRTL-SAC [10], Pharma mod. UNIFAC [11], etc.)

In contrast to the rigorous simulation of a different unit operation or a complete chemical plant, conceptual design relies heavily on simplified concepts (shortcut methods, infinite-infinite analysis), which employ a variety of graphical representations (e.g. residual curve plots) []. While these methods and tools are nowadays state-of-the-art for design and optimization of continuous processes, they are increasingly applied to smaller scale batch processes.

Compared to distillation and extraction, which both involve fluid phases and are characterized by very high scale-up factors, crystallization design is both more governed by kinetic effects and involves numerous important factors like nucleation, different crystal morphologies, etc. .

Simulation results for organic crystallization processes can therefore most often not be directly applied to the final process. Nevertheless, a number of researchers [12] [13] [14] have developed computerized strategies to generate and evaluate crystallization strategies.

These methods are of special importance for the purification of products from reaction mixtures in pharmaceutical synthesis due to the fact that the final product approval after expensive clinical tests not only covers the molecular structure but also each individual step in synthesis and purification of intermediates and the final product.

In process simulation, crystallization of organic non-electrolyte components is usually restricted to eutectic systems, in which the components crystallize as pure separate solid phases. While the thermodynamic models describe phase equilibrium, their results can also be used to judge the degree of supersaturation as a driving force for nucleation and crystal growth. In addition to the required parameters for pure component and liquid mixture behaviour, the melting points and heats of fusion of the solids are required and usually nowadays easily accessible by differential scanning calorimetry. The thermodynamic formalism is derived in detail below.

For the description of the real liquid phase behaviour, interaction parameters (BIP) are required between the complex product, by-products and remaining starting materials (solutes) in the mixture and a limited number of common used solvents [15]. At not too high concentrations of the solutes, solute-solute interaction can be ignored. The solvent-solute BIPs can be estimated by group contribution (UNIFAC, mod. UNIFAC, PSRK, VTPR, etc.), simplified quantum chemical methods (COSMO-RS, COSMO-SAC) or from the solution behaviour in a limited number of mixtures (NRTL-SAC).

This work is aimed at developing and implementing shortcut procedures and calculation code to assist the chemist and engineer in the interactive development of product recovery and purification strategies via crystallization including evaluation of different solvents and anti-solvents (drowning-out).

Whenever possible, functionalities available in a commercial process simulator (CHEMCAD®) were used as this guarantees the regular future update and extension with respect to models and parameters. In addition, the simulator provides interfaces and load-procedures for third party products like NIST REFPROP and the UNIFAC Consortium parameters [16]. CHEMCAD was chosen as it provides consistent vapour-liquid-solid equilibrium calculations in all streams and unit operations.

Thermodynamic framework for solid-liquid-vapor equilibria

The objective of the phase equilibrium calculation is to predict the correct number of phases at equilibrium present in the system and their respective compositions. Two kinds of approaches are usually used to model multiphase flash calculations: the equation-solving approach (K-value method) and minimization of the Gibbs free energy [17]. Isotugacity conditions and mass balances form the set of equations in the equation-solving approach, and the stability test or the common tangent test forms the basis of minimization of the global Gibbs free energy approach [18] [19] [20].

Since the developed solvent selection process eliminates solvent systems that exhibit immiscibility as potential solvents, the mixtures are modeled as solid – liquid – vapour (SLV) systems.

The essential thermodynamic equations and relationships for SLV equilibrium calculations presented here are analogous to those described by *Lira-Galeana et al.* [21] for wax deposition in hydrocarbon streams. At a fixed temperature and pressure, a liquid phase (*l*) may coexist in equilibrium with a vapour phase (*v*) and a solid phase (*s*). At equilibrium, for every component *i* the following thermodynamic relationship applies:

$$f_i^v = f_i^l = f_i^s \quad i = 1, 2, \dots, N \quad (1)$$

Where *f* is the fugacity and *N* is the number of components. The vapour phase can be described by an equation of state (EOS), the liquid phase by an activity-coefficient model or by an EOS, and the solid phase is generally described by an activity-coefficient model, i.e.

$$f_i^v = y_i \phi_i^v P; \quad f_i^l = x_i^l \phi_i^l P \text{ or } f_i^l = x_i^l \gamma_i^l f_{pure\ i}^l, \quad \text{and } f_i^s = x_i^s \gamma_i^s f_{pure\ i}^s \quad (2)$$

Where ϕ_i^v and ϕ_i^l are fugacity coefficients of component i in the vapour and liquid phases respectively and are computed from an EOS, and γ_i^l and γ_i^s are activity coefficients of component i in the liquid and the solid phases respectively and are computed from activity coefficient models.

Further, the use of distribution coefficients K which are generally used in VLE and LLE computations are extended to describe the equilibrium relationships between the phases in a SLV system. For the vapour – liquid phase, the commonly used expression is

$$K_i^{vl} = \frac{y_i}{x_i^l} = \frac{\phi_i^l}{\phi_i^v} \quad (3)$$

For the liquid-solid phase, the fugacity can be described with the help of activity coefficients and the standard fugacities for the liquid and solid phases, i.e

$$x_i^l \gamma_i^l f_{pure\ i}^l = x_i^s \gamma_i^s f_{pure\ i}^s \quad \text{or} \quad \frac{x_i^l \gamma_i^l}{x_i^s \gamma_i^s} = \left(\frac{f^s}{f^l} \right)_{pure\ i} \quad (4)$$

Where *Lira-Galeana et al.* [21] proposed an analogous equilibrium constant

$$K_i^{sl} = \frac{\gamma_i^l}{\gamma_i^s} \left(\frac{f^l}{f^s} \right)_{pure\ i} \text{ and therefore } K_i^{sl} = \frac{x_i^s}{x_i^l} \quad (5)$$

The required ratio of the standard fugacities of the pure components can be obtained by examining the thermodynamic cycle of the sublimation process of a solid. From the cycle shown in Figure 2, it can be deduced that the ratio of the standard fugacities can be expressed by the change in Gibbs energy [17].

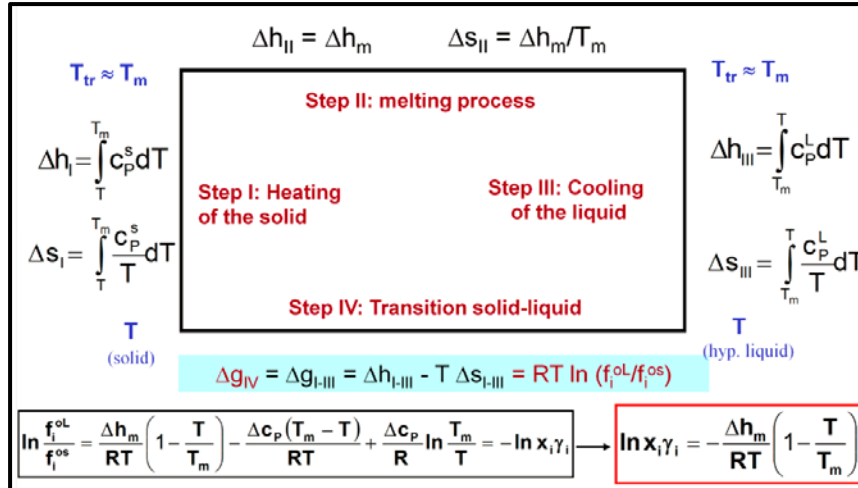


Figure 2: Thermodynamic Cycle for the derivation of an expression for the ratio of standard fugacities. ([17], page 409).

With the simplifications that the melting temperature is nearly identical to the triple point temperature, that heat of fusion is approximately identical to the change in enthalpy of the solid-liquid phase transition at triple point, and that heat capacity difference is negligible, *Gmehling et al.* [17] have shown that:

$$\ln \left(\frac{f^l}{f^s} \right)_{pure\ i} = -\ln \left(\frac{x_i^l \gamma_i^l}{x_i^s \gamma_i^s} \right) = \frac{\Delta h_{m,i}}{RT} \left(1 - \frac{T}{T_{m,i}} \right) \quad (6)$$

In the case of simple eutectic systems, the solid will crystallize in pure form, hence equation 6 reduces to

$$-\ln x_i^l \gamma_i^l = \ln \left(\frac{f^l}{f^s} \right)_{pure\ i} = \frac{\Delta h_{m,i}}{RT} \left(1 - \frac{T}{T_{m,i}} \right) \quad (7)$$

From equation 7, it can be seen that to calculate the ratio of the standard fugacity at a given temperature and pressure, only the melting temperature, the latent heat of fusion and specific heat capacity of pure liquid *i* and pure solid *i* are required.

Conceptual Design of crystallization processes

“The goal of a conceptual design is to find the best process flowsheet and estimate the optimum design conditions. The problem is difficult because very many process alternatives could be considered.” [22] Conceptual design means the handling of chemical processes from first principles

(like thermodynamic model etc.). The aim of conceptual design is to find the “best” flowsheet alternative via optimization of variables like costs, efficiency, etc.. Systematic procedures for synthesis of separation flowsheets via distillation and extraction etc. were already developed by different researchers (Dogerthy, etc.). A hierarchically procedure for the development of a separation process via crystallization is given by Douglas et al. [22]. The procedure is based on an input-output structure containing a separation train with a recycle structure. During the last decades this procedure was improved, i. e. by Ng [23] [24] [25]. Separation of non-electrolytes via crystallization is generally based on two possible approaches; temperature change and composition change. When lowering the temperature, crystallization of a component simultaneously leads to a composition change along the solid – liquid equilibrium curve. Composition change other than that obtained from crystallization of one or several components may also be achieved by: evaporating a solvent or solvent mixture from the solution, or by adding a further solvent as in anti-solvent (drowning out) crystallization. In case of adiabatic evaporation, both temperature and the amount of solvent are changed. In order to achieve a sufficient rate of crystallization, the system needs to be supersaturated by a certain degree that can be judged from the knowledge of the solid-liquid equilibrium. Therefore, visualization of the solid-liquid equilibrium in form of various different diagrams is of great importance for the conceptual design of crystallization processes. Examples will be presented for the use of graphical representations in binary, ternary and higher systems. Approaches for the development of flowsheet alternatives based on ternary SLE diagrams are given by Ng [23] [24] [25] in detail.

Basic operation steps

In general, crystallization via temperature or composition change can for example be achieved using basic operations like e. g. cooling, evaporating, adding of an anti-solvent or co-solvent, stream splitting (with concentration change) or combination, extraction, pH-shift, salting-out and reactive crystallization. The resulting movements in a ternary diagram caused by some basic operation steps are shown in Figure 3. The pure component i represented by the melting point of component i , results in an apex (end point) in the ternary polythermal projection. Our mixture shows three binary eutectic points ($BE_{i,j}$) and one ternary eutectic point ($TE_{i,j,k}$). The resulting eutectic curves between these points define the operating regions. Various manifolds bound the operating regions. The number of these manifolds is equal to the number of components in the feed. In each bounded region, only a single component can be crystallized in pure form. Simultaneous crystallization of more than one component is possible for components sharing a eutectic manifold. Movement across a manifold into another operation region is achieved by either evaporation, addition of a solvent or anti-solvent, and combination or splitting streams.

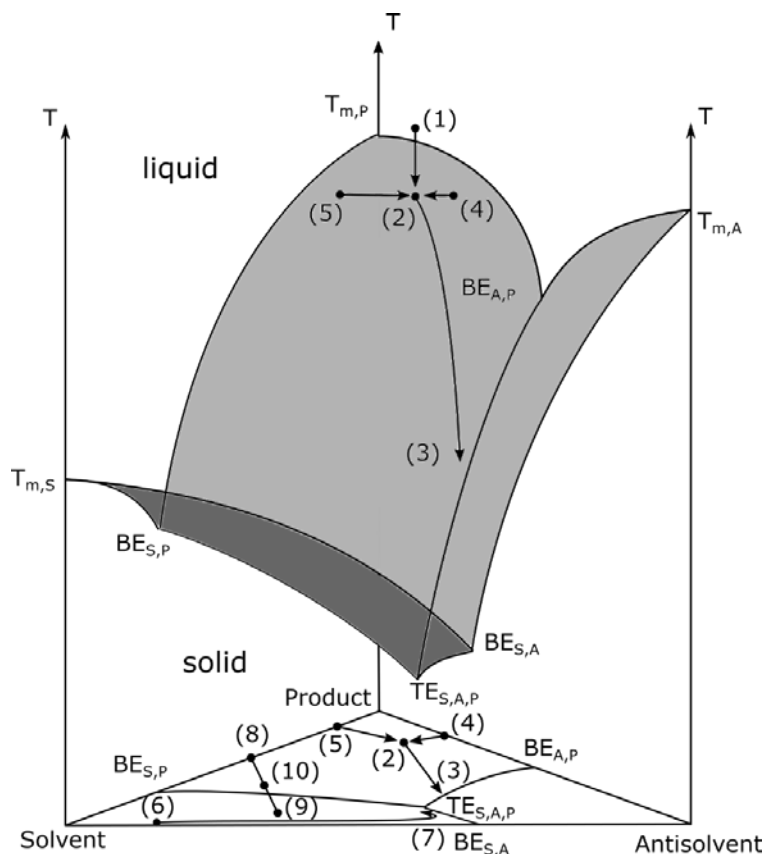


Figure 3: Isobaric SLE phase diagram for a system containing a solvent (S), an anti-solvent (A) and a product (P) with a ternary polythermal projection.

Starting with a feed composition (1) one can cool down until the SLE surface is reached (2). For simplicity we assume equilibrium operation without the need to subcool. Further cooling results in crystallization of the product and movement along the SLE surface on a straight line away from the product apex (3). The addition of a solvent to a binary mixture containing the product and anti-solvent (4) results in (2). Adding an anti-solvent to a binary mixture of the product dissolved in a solvent (5) also leads to a new composition (2). The evaporation of the solvent from (6) along the distillation curve results in the new composition (7). In order to change into a different operating region one can combine two different streams (8) and (9) resulting in (10).

Whilst cooling crystallization is the preferred mode of crystallization, recovery of the product is limited by its solubility at the lowest feasible temperature for the available equipment or the temperature, at which another component in the mixture would crystallize. The slope of the SLE curve should be flat enough for an efficient crystallization (sufficient change of solubility with temperature). In addition to the costs involved in low temperature cooling, crystallization at very low temperatures may be very slow. Some discussion of the effect of the real liquid mixture behaviour on solvent selection is given below.

If an efficient crystallization via cooling is not feasible, one should consider other crystallization methods like solvent evaporation or adding a co-solvent or anti-solvent. An example of a decision tree that can assist in determining the crystallization trajectory in temperature and composition space is shown in Figure 4.

The possible temperature range is in any case between the achievable lowest temperature in the equipment (lower limit) and the decomposition temperature of the components in the mixture or the maximum vapour pressure of the mixture that can be handled by the equipment.

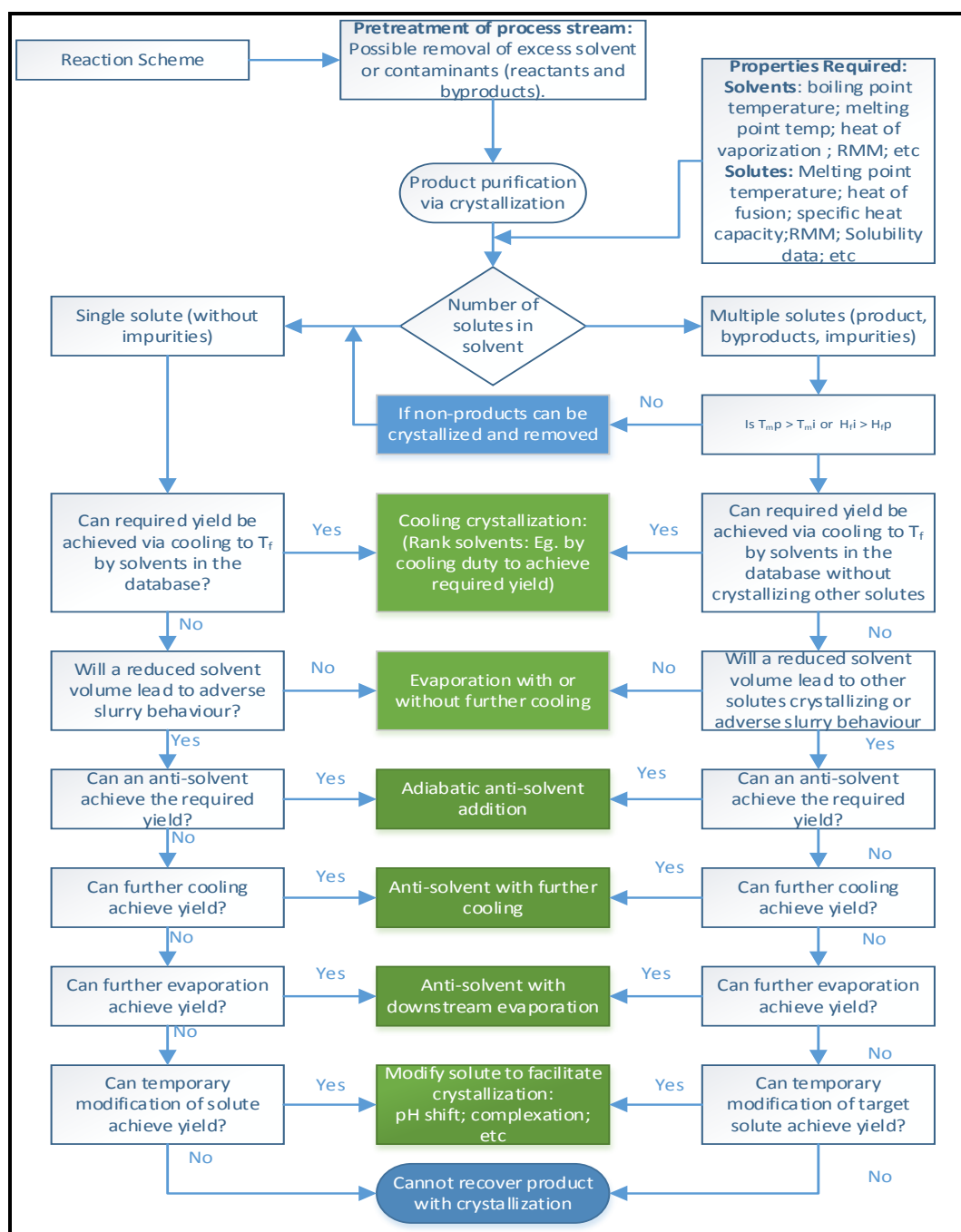


Figure 4: Example of Decision Tree to determine operating protocol for crystallization.

The scheme shown in Figure 5 shows one possible sequencing of these individual steps. The stream from the reactor R contains the target product (P), the by-products (BP) and the starting materials (SM). The first option is to recover the product via cooling down or evaporate the solvent. In this case one gets an amount of the pure product and in case of evaporation the amount of solvent is reduced too and the evaporated pure solvent can be recycled. After this step, adding an anti-solvent is preferred. This will lead to a binary solvent system where the solubility of the product should decrease and the solubility for the impurities (SM, BP) should increase. Again an amount of product can be recovered. After adding the anti-solvent, one can evaporate from the binary solvent/anti-solvent the solvent or an azeotrop, which should lead preferably to a higher amount of anti-solvent and a change

of the solubility so one can recover further product. The anti-solvent should therefore have a lower volatility than the solvent. The stream should now contain by-product and starting materials and a only small amounts of the product.

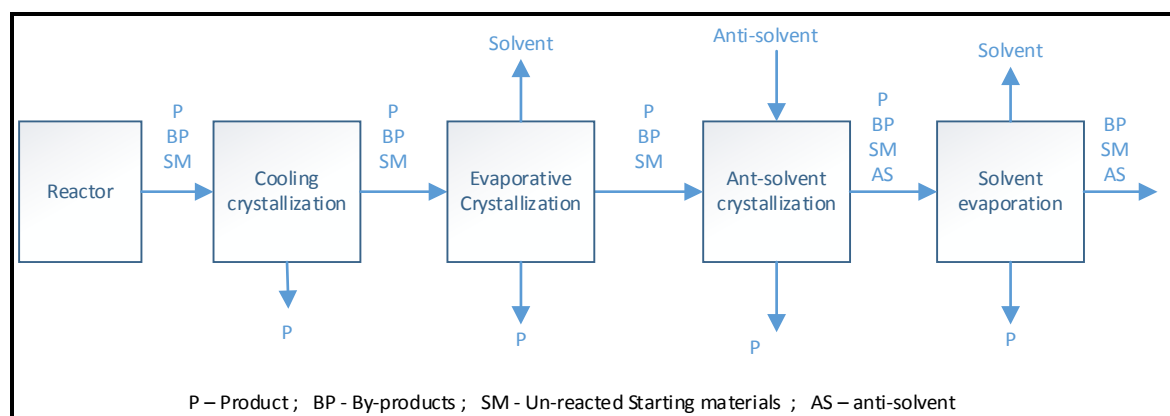


Figure 5: Example of sequencing possibilities for maximum recovery of product.

Criteria for solvent and anti-solvent/co-solvent selection for crystallization processes

During the development phase of an API, it is prudent to determine the best solvent or solvents to be used in the manufacturing process, because once the clinical trial phase has been conducted, the legislative and regulatory process prevents changes to the production without further clinical trials. To determine possible process alternatives, the knowledge of ternary SLEs including the solvent, the product and probably an appropriate anti-solvent or an existing by-product is unavoidable. From these SLEs one can also define the performance of a possible solvent. The measurement of the required phase equilibrium data points is often difficult and lengthy. These data are often available in commercial data banks like DDB [1] and NIST [2]. If the required data points are not available, prediction is recommended. The presented computational software tool is able to predict and plot ternary SLE diagrams using all predictive models available in the simulator. The multiphase equilibria are generated by algorithms that are based on multiphase flash calculations.

Equilibrium phase diagrams can be effectively used to visualize the movements in composition space associated with the different basic operations. Several researchers [24] [13] [14] [26] have developed rules and guidelines for synthesising operational protocols for crystallization based separation processes, and some of the key elements are presented here for a eutectic mixture without adduct formation (mixed crystals with defined stoichiometry).

The selection of the “best” solvents is based on some variables like the required yield, temperature range (e.g. cooling below 0 °C requires a refrigerant, decomposition temperatures of involved components), costs and feasibility. Figure 6 shows the predicted binary SLEs for Ibuprofen in various solvents (UNIFAC). Ibuprofen is best soluble in THF, good soluble in Toluene and poorly soluble in *n*-Hexane. As can be seen from Figure 6, *n*-Hexane would be the preferable solvent in case of crystallization at higher temperatures, because much less product is lost in the mother liquor than in case of the “better” solvents. If a crystallization is to be performed at lower temperatures, THF provides both a higher solubility and a much higher concentration change with temperature.

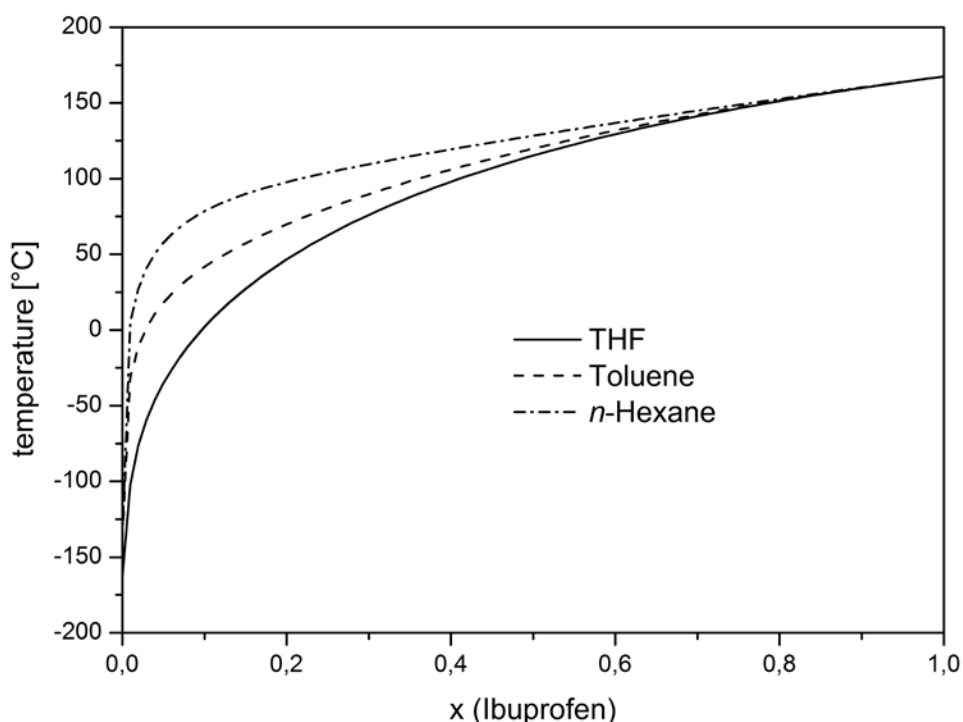


Figure 6: SLEs of Ibuprofen in three different solvents calculated via UNIFAC.

For a crystallization, the addition of an anti-solvent can be very efficient. Starting the crystallization with a stream with a high product content in a good solvent and adding a miscible good anti-solvent leads to a high crystallization yield via low cooling requirement. An equivalent yield without anti-solvent addition would require a very drastic cooling.

Development of a computational tool

In order to assist the engineer in evaluating different crystallization strategies, a VBA Tool in CHEMCAD was developed. All calculations are based on the Multiphase-Flash Algorithm. This allows a simultaneous calculation of the vapour, liquid and solid phase behaviour. It allows to model cooling and evaporative crystallization, adding an anti-solvent or cosolvent and the combination of these methods. Furthermore, it is possible to identify an optimal solvent.

With yield-dependent processes, the plant size and utility requirements vary with yield. Low yields require higher processing volumes to achieve the required production target, leading to higher capital and operating costs. To facilitate decision making during the conceptual design stage, further criteria are required to screen and identify the best solvent and combination of crystallisation operating modes and conditions. In addition to the crystal yield criterion, the two additional screening criteria are included into the Solvent Selection Tool: the economic criterion and the environmental criterion. For the economic criterion, use is made of either the operating cost per batch of API produced or the Fixed Annualised Cost which is a combination of the annual operating cost and a ratio of the capital costs. For the environmental criterion, indices commonly used in green chemistry [27] – the Environmental Factor and the Energy Consumption Factor is used.

Once the desired production rate is established, the effect each solvent will have on: the size of plant required, the capital expenditure and the operational expenses can be determined. The plant operations that are considered to be directly affected by the choice of solvent is limited to the following: the crystalliser size or number of modular units and its required heat exchanger area; the solvent feed and waste storage tanks; refrigeration unit and the solvent recovery system. The capital cost calculations is based on a module factor approach presented by Turton et al. [28]. The operating costs is based on the key utilities that are significantly influenced by the selected solvent required to meet the desired production rate of API. These costs include the cost of solvent; cost of cooling (cooling crystallization), cost of heating (evaporative crystallization), cost of solvent recovery, and cost of tertiary waste treatment of unrecovered solvent, API and other components. The various flowrates and heating and cooling are determined by material and energy balances setup in the VBA module. The utility costs is determined using the method proposed by Ulrich and Vasudevan [29] .

The Environmental (E) Factor for the potential solvents is a measure of the amount of waste generated per kilogram of API crystallised, and the Energy Consumption (E_c) Factor is a measure of the total energy required per kilogram of product crystallised. This energy factor accounts for the following: the heat load of the crystalliser for cool or evaporation, and the heat load required to recover 80 % of the solvent in the solvent recovery process. This “total energy” consumed per kilogram of API produced can be directly correlated to carbon footprint contribution of the selected solvent.

The ranking of the operating conditions can be based on a user defined weighting of the performance criteria of yield, operating cost or fixed annualised cost per batch, the waste generated per kg of API crystallised and the energy consumption per kg of API crystallised. A cumulative effect of the selected criteria for each of the solvent or operating condition is determined, and the solvent and operating condition with the lowest Net Cumulative Weighted Score (NCWS) will represent the best solvent and operating condition.

The various performance criteria of operational cost, E-Factor and E_c -Factor are normalized by dividing the actual value obtained at specified conditions divided by the lowest value obtained in the entire range evaluated. The normalized yield criteria is obtained by dividing the actual value obtained at specified conditions divided by the minimum required yield. Since the proposed ranking method is based on the lowest cumulative value, the inverse of normalized yield value is used. The Net Cumulative Weighted Score (NCWS) is calculated by:

$$NCWS = \sum_i^n w_i NPC_i = w_y \times (Normalised Yield)^{-1} + w_{op} * Normalised Operating Cost + w_e \times Normalised E - Factor + w_{Ec} \times Normalised Ec - Factor \quad (8)$$

where

$$w_i \text{ is the user defined weighting and } Normalised Performance Criteria (NPC_i) = \frac{\text{value of } PC_i \text{ for an operating condition}}{\text{lowest value for this PC}}$$

The tool is directly connected to Microsoft Excel® and all results are exported directly to different sheets. Besides product, solvent and anti-solvent also the solubility of by-products and remaining starting materials is taken into account.

Workflow description

The typical input requirements, calculation and display options and result format choices are shown in Figure 7.

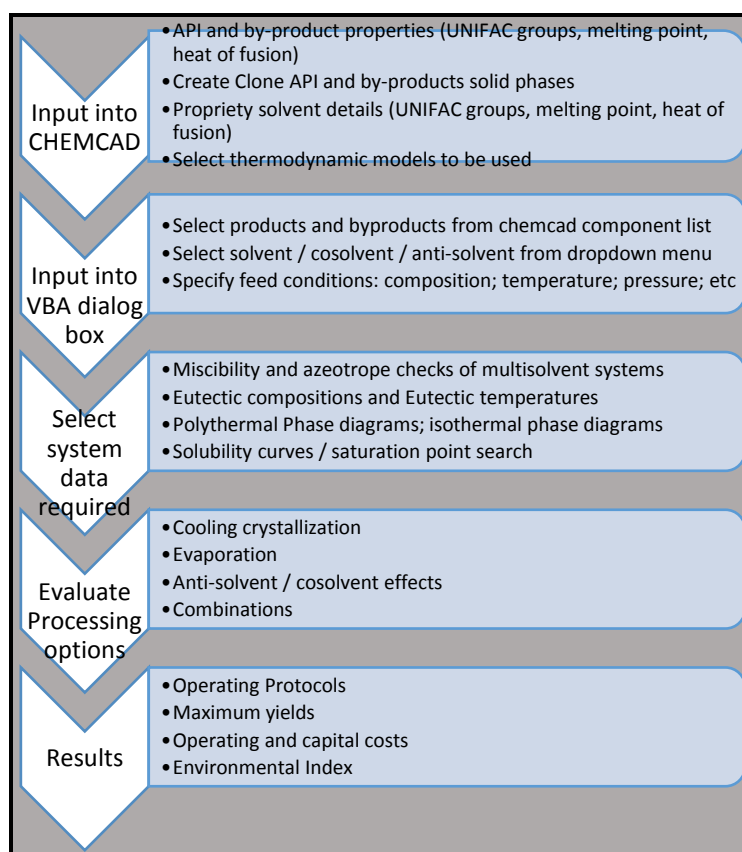


Figure 7: Outline of the computer-assisted workflow.

Crystallization dialog

The functions of the program are available from two different main menus (Temperature Menu and Solvent/Anti-solvent Menu).

In the Temperature Menu of the VBA tool is able to deal with a crystallization via cooling down or evaporation of the solvent. If this leads to required yield, adding an anti-solvent is not necessary. Furthermore, the Temperature Menu offers several computational options: determination of the eutectic temperatures and compositions; determination of maximum evaporation or maximum cooling that can be done to achieve a pure solid i.e. prior to a second solute crystallizing out. In the Temperature Menu were the selected Solvent and feed composition is specified, there are many functions to optimize the crystallization process via cooling and evaporative crystallization and a combination of both crystallization methods.

The Solvent/Anti-solvent Menu is a tool to find the best solvent and if necessary the best anti-solvent for crystallization. By creating a database of solvents that are used in pharmaceutical processes, the computational algorithm can determine best combination of the solvent and anti-solvent by

calculating the solubility of all compounds in all of the solvents and arrays them to good solvents and good anti-solvents. The database of solvents built into the selection tool is based on the GSK Solvent Selection Guide [15]. Ideally, the anti-solvent should decrease the solubility of the product and should increase the solubility of the impurities. Furthermore, it has to be miscible with the solvent and in case a recovery of the solvents is considered it should not form an azeotrope or adduct with the solvent. Also for the possibility of further solvent evaporation the boiling point of the solvent should be below the boiling point of the anti-solvent. Because of the increasing volume during the addition of an anti-solvent, it is recommended to add the anti-solvent to a saturated solution. This will also lead to a higher amount of the anti-solvent in the binary solvent system, leading to higher crystallization yields.

Applications examples

The application of the tool is illustrated by three examples, i. e. the selection of solvents and anti-solvents in the absence of any by-products or impurities, solvent selection in the presence of impurities, and application to a product used in liquid crystal displays.

Ibuprofen

Ibuprofen has to be recovered from a solution containing 1 kmol Ibuprofen in 2.5 kmol Tetrahydrofuran. Cooling to -50 °C leads to 0.92 recovery, that can be increased to 0.96 kmol upon evaporation of 1.25 kmol of solvent.

Through the solvent selection tool, Ethanol was proposed as a better solvent. With this solvent it can be shown, that with the same composition (2 kmol Ibuprofen dissolved in 5 kmol Ethanol) one can achieve 99 % pure Ibuprofen via cooling down to -50 °C. For a combination of a previous solvent evaporation of 50 % and then cooling down to -50 °C the amount of pure Ibuprofen can be increased to 99.5 %. It must be highlighted that these conditions are impractical because of the extremely low temperature required to achieve the desired yields, which will lead to high energy consumption. Instead of this drastic cooling, an anti-solvent could be used. In this case, 3 % of the Ethanol are first removed by evaporation to yield a saturated solution. From the solvents considered here, water may act as a good anti-solvent. The effect of the addition of water to the saturated Ibuprofen solution in Ethanol with and without further cooling is shown in **Error! Reference source not found.** The left and right faces of the graph show the yield obtained for pure cooling crystallisation and pure anti-solvent crystallisation respectively. Whilst high yields may be obtained at low temperatures for cooling crystallisation only or high anti-solvent addition for anti-solvent crystallisation only, high yields can also be obtained at milder conditions for the combination of cooling and anti-solvent crystallisation. The process implications of the various possible combinations of anti-solvent solvent addition and cooling is determined by calculating its associated capital and operating cost. Assuming an equal weighting of the four performance criteria of yield, fixed annualised costs, waste generated (E factor) and energy used (E_c Factor), Figure 8 illustrates how the Net Cumulative Weighted Score (NCWS) can be used to determine the best operating protocols that will result in a minimum yield of 98%. The addition of anti-solvents leads to higher processing volumes which impact on the capital costs (larger equipment size) and operating costs (cooling/heating, solvent recovery, etc) hence leading to higher NCWS values. Similarly, very low operating temperatures require larger refrigeration and utility requirements also leading to higher NCWS values. It should be noted that a different weighting of the four performance criteria (as guided by the company policy and objectives) will result in a different profile to that shown in figure 8 i.e. the user may have a lower weighting to the environmental criteria and higher weightings for the yield and economic criteria. Using this method, further alternatives may also be evaluated and verified by experiments.

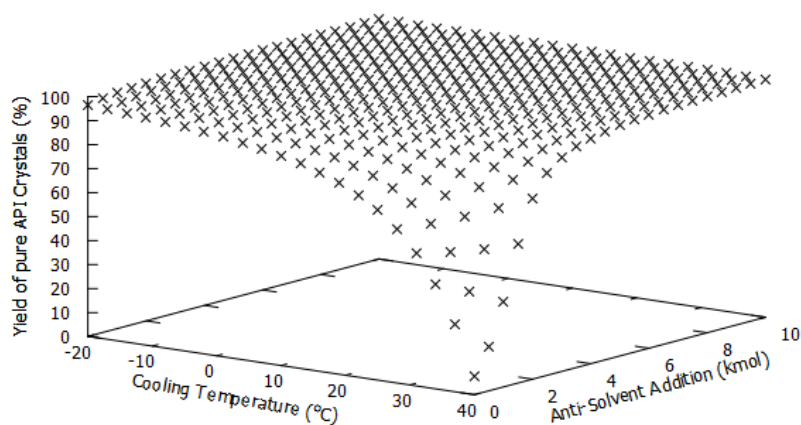


Figure 7 Effect of cooling and anti-solvent addition on Ibuprofen crystal yield

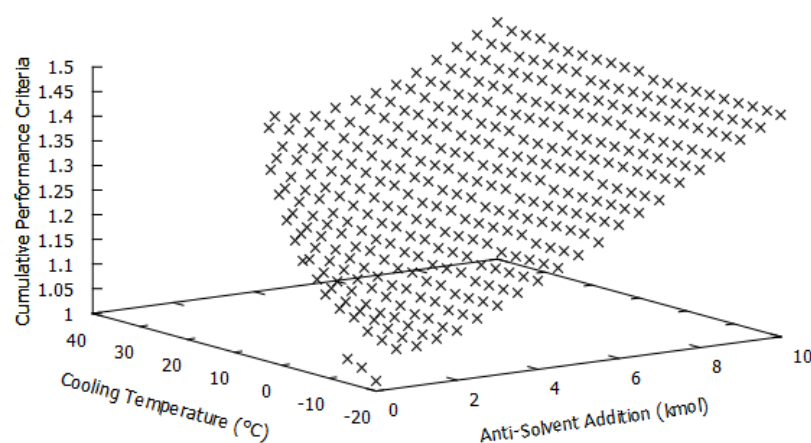


Figure 8 Cumulative performance taking into account yield, fixed annualised cost, waste generated and energy consumption.

Ibuprofen/Paracetamol

In this example, a mixture containing Ibuprofen as the product and Paracetamol as a by-product in Acetonitrile is considered. The reaction mixture contains 2 kmol Ibuprofen and 0.5 kmol Paracetamol dissolved in 10 kmol Acetonitrile. Due to the low solubility of Paracetamol it is not possible to crystallize pure Ibuprofen from the reaction mixture via cooling or solvent evaporation. It is therefore considered to add an anti-solvent to the reaction mixture in order to either crystallize and thereby remove the impurity (Paracetamol) and then crystallize the pure Ibuprofen or crystallize Ibuprofen while keeping Paracetamol in solution by an anti-solvent selective for Ibuprofen. The first option can be realized by using *n*-Hexane as an anti-solvent. The removal of Paracetamol from the mixture by

anti-solvent crystallization is shown in 9 for various temperatures. From this figure the minimum temperature can also be obtained below which *n*-Hexane addition would also lead to Ibuprofen precipitation.

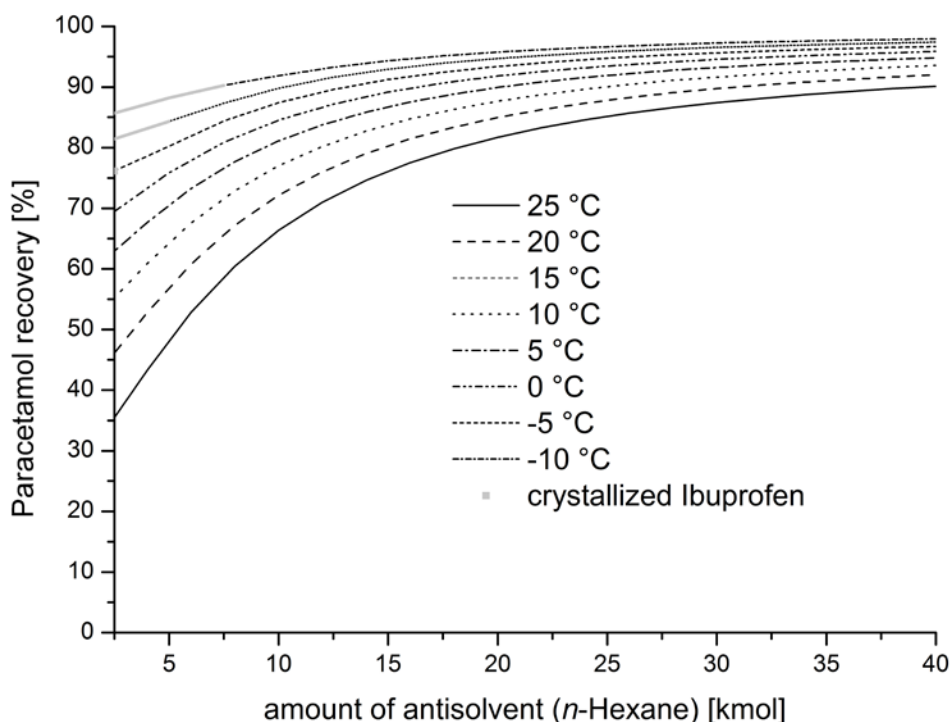


Figure 9: Recovery of Paracetamol from Paracetamol/Ibuprofen in Acetonitrile via anti-solvent addition (*n*-Hexane) at different temperatures.

91.8 % of the paracetamol can be removed via adding 10 kmol of *n*-Hexane and cooling down to -10 °C. Further addition of the anti-solvent has only a minor effect on the Paracetamol yield. After removal of most of the Paracetamol it is possible to crystallize 88 % of the Ibuprofen via evaporating 95 % of the solvent mixture at 15 °C.

An alternate option is to use an anti-solvent to yield Ibuprofen while keeping Paracetamol in solution. Using the crystallization tool we identified water as a possible anti-solvent combined with Ethanol as solvent. In this case, it is possible to crystallize 92.72 % of Ibuprofen via adding 20 kmol of water at 25 °C. Figure 10 shows the amount of crystallized Ibuprofen as function of the amount of anti-solvent (water) at 25 °C

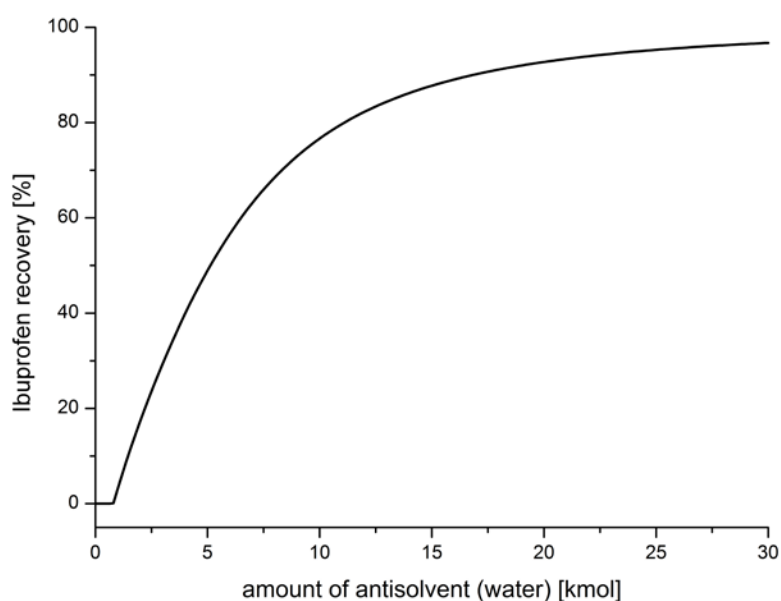


Figure 10: Recovery of Ibuprofen from an ethanolic solution of Ibuprofen and Paracetamol by anti-solvent (water) addition at 25 °C.

Liquid crystal component

The last example deals with an unknown component where only UNIFAC-groups and compositions are available from a real crystallization process. The high value product is a propriety LCD crystal and the composition and properties of the process stream is shown in Table 2 below.

Table 2. Composition and properties of LCD crystal process stream

	Product	By-product 1	By-product 2	Solvent THF
Composition [mol%]	12.73	0.0095	0.5315	86.729
Melting point [°C]	51.00	171.71	-44.58	-108.39
Heat of Fusion [J/g]	78.77	157.96	74.01	86.73

The only known properties of the products and by-products are the heat of fusion and the melting point as well as the UNIFAC-groups of the molecules. The primary objective is to crystallize the maximum amount of pure LCD crystals. Because of the higher melting temperature and the higher heat of fusion, the difficulty is the separation of by-product 1 from and the product.

The available data are sufficient to simulate the crystallization behaviour. We found that the maximum cooling to -13 °C without any by-product crystallization leads to 36 % crystallized product. However, with solvent evaporation of 95 % Tetrahydrofuran and then cooling down to 18.5 °C it is possible to crystallize 76 % of the product without precipitating impurities. Evaluation of different solvents pointed to Acetonitrile as a promising alternative. Using this solvent, it is possible to crystallize 86 % of the product just via cooling to 1 °C. If 74.25 % of the Acetonitrile are evaporated before cooling to 13.2 °C, 91.5 % of the product can be recovered. When using water as an anti-solvent and Acetonitrile as the solvent, adding 10 kmol of water to the solution and then cooling to 5 °C will yield to 97.8 % of the product. It should be noted that maintaining a higher temperature during crystallization will quite

definitely also reduce the required crystallization time and thus shorten the use of the crystallizer. A summary of the results for the various process options are presented in Table 3. It should be noted that the option with the maximum yield is not necessarily the optimised solution since the operation and capital cost associated with each option has to be evaluated.

Table 3. Summary of results for some processing options for the LCD crystal stream

	Scenario 1	Scenario 2	Scenario 3	Scenario 4	Scenario 5
Solvent	Tetrahydrofuran	Tetrahydrofuran	Using Acetonitrile as solvent	Using Acetonitrile as solvent	Using Acetonitrile as solvent
Crystallization protocol	Cooling to -13 °C	95% Solvent evaporation then cooling to 18.5 °C	Cooling to 1 °C	74.25% Solvent evaporation then cooling to 13.2 °C	Anti-solvent addition then cooling to 5 °C
Yield of pure crystals	36%	76%	86%	91.5%	97.8%

Conclusion

It has been shown that using the computational tool developed in this work it is easy to evaluate process alternatives for crystallization during the development stage of any project. We focus on the solvent selection for a specific crystallization problem. Furthermore, we implemented three crystallization methods: Cooling crystallization, evaporative crystallization, adding an anti-solvent or cosolvent and combination of all these methods. The results obtained can be effectively used for decision making prior to incurring of substantial product development costs, and can also be used to decide on the direction for further research and experimental work. The ability to develop a platform which can integrate existing commercial simulation software into company specific operations allows the users to have access to a full range of thermodynamics models and correlations, the comprehensive database of compounds and their pure and mixture properties, and the rigorous computational algorithms for process calculations and equipment design. The software code is freely available from the authors upon request.

Acknowledgement

Please insert project fundings etc. here

References

- [1] J. R. J. M. J. Gmehling, "Dortmund Data Bank," DDBST GmbH, [Online]. Available: <http://www.ddbst.com>.
- [2] K. Kroenlein, C. D. Muzny, A. F. Kazakov, D. V., R. D. Chirico, J. W. Magee, I. Abdulagatov and M. Frenkel, "NIST Web Thermo Tables," [Online]. Available: <http://www.nist.gov/srd/nistwebsub3.cfm>.
- [3] L. J. G. J. Wittig R., *Ind. Eng. Chem. Res.*, pp. 183-188, 2003.
- [4] D. Constantinescu and J. Gmehling, "Further Development of Modified UNIFAC (Dortmund): Revision and Extension 6," *J. Chem. Eng. Data*, pp. 2738-2748, 2016.
- [5] J. A. K. J. F. K. G. J. 2. Horstmann S., "PSRK Group Contribution Equation of State: Comprehensive Revision and Extension IV, Including Critical Constants and a-Function Parameters for 1000 Components," *Fluid Phase Equilib.*, pp. 157-164, 2005.
- [6] B. Schmid, A. Schedemann and J. Gmehling, "Extension of the VTPR Group Contribution Equation of State: Group Interaction Parameters for Additional 192 Group Combinations and Typical Results," *Ind. Eng. Chem. Res.*, pp. 3393-3405, 2014.
- [7] A. Klamt, V. Jonas, T. Bürger and J. C. W. Lohrenz, "Refinement and Parametrization of COSMO-RS," *J. Phys. Chem.*, pp. 5074-5085, 1998.
- [8] C.-M. Hsieh, S.-T. Lin and J. Vrabec, "Considering the dispersive interactions in the COSMO-SAC model for more accurate predictions of fluid phase behavior," *Fluid Phase Equilibria*, pp. 109-116, 2014.
- [9] C.-M. Hsieh, S.-T. Lin and J. Vrabec, "Corrigendum to: Considering the dispersive interactions in the COSMO-SAC model for more accurate predictions of fluid phase behavior," *Fluid Phase Equilibria*, pp. 14-15, 2014.
- [10] H. H. Tung, J. Tabora, N. Variankaval, D. Bakken and C. C. Chen, "Prediction of pharmaceutical solubility Via NRTL-SAC and COSMO-SAC," *J. Pharm. Sci.*, pp. 1813-1820, 2008.
- [11] A. Diedrichs and J. Gmehling, "Solubility Calculation of Active Pharmaceutical Ingredients in Alkanes, Alcohols, Water and their Mixtures Using Various Activity Coefficient Models," *Ind. Eng. Chem. Res.*, pp. 1757-1769, 2011.
- [12] L. Cisternas, "Optimal design of crystallization-based separation schemes.," *AIChE Journal*, vol. 45, no. 7, pp. 1477-1487, 1999.
- [13] D. Berry , S. Dye and K. Ng, "Synthesis of drowning-out crystallization- based separation," *AIChE*, vol. 43, 1997.

- [14] L. Cisternas , C. Va´squez and R. Swaney, "On the Design of Crystallization-Based Separation Processes: Review and Extension May 2006 Vol. 52, No. 5 *AIChE Journal*," *AIChE*, vol. 52, no. 5, 2006.
- [15] K. Grodowska and A. Parczewski, "Organic solvents in the pharmaceutical industry," *Acta Pol. Pharm.*, pp. 3-12, 2010.
- [16] "The UNIFAC Comsortium," [Online]. Available: <http://unifac.ddbst.de/>.
- [17] J. Gmehling, B. Kolbe, M. Kleiber and J. Rarey, *Chemical Thermodynamics for Process Simulation*, Wiley, 2012.
- [18] A. Lucia, L. Padmanabhan and S. Venkataraman, "Multiphase equilibrium flash calculations," *Computers and Chemical Engineering*, vol. 24, 2000.
- [19] V. Parekh and P. Mathias, "Efficient flash calculations for chemical process design – extension of the Boston-Britt "inside-out" flash algorithm to extreme conditions and new flash types," *Computers Chem Engng*, vol. 22, no. 10, 1998.
- [20] K. Moodley, J. Rarey and D. Ramjugernath, "Application of the bio-inspired Krill Herd optimization technique to phase equilibrium calculations," *Computers & Chemical Engineering*, 2015.
- [21] C. Lira-Galeana, A. Firoozabadi and J. Prausnitz, " Thermodynamics of Wax Precipitation in Petroleum Mixtures," *AIChE*, vol. 42, no. 1, 1996.
- [22] J. Douglas, *Conceptual Design of Chemical Process*, McGraw-Hill , 1988.
- [23] C. Wibowo and K. M. Ng, "Unified Approach for Synthesizing Crystallization-Based Separation Processes.," *AIChE*, vol. 46, no. 7, 2000.
- [24] C. Wibowo and K. Ng, "Unified Approach for Synthesizing Crystallization-Based Separation Processes," *AIChE*, vol. 46, no. 7, 2000.
- [25] C. Wibowo and K. Ng, "Workflow for Process Synthesis and Development: Crystallization and Solids Processing," *Ind. Eng. Chem. Res*, vol. 41, 2002.
- [26] B. Fitch, " How to design fractional crystallization processes," *Industrial and Engineering Chemistry*, vol. 62, no. 12, 1970.
- [27] P. J. Dunn, "The importance of green chemistry in process research and development.," *Chemical Society Reviews* , vol. 41, no. 4, pp. 1452-1461., 2012.
- [28] R. B. R. W. W. a. S. J. Turton, *Analysis, synthesis and design of chemical processes. .* , Pearson Education, 2008.
- [29] G. a. V. P. Ulrich, "How to estimate utility costs.," *Chemical Engineering*, vol. 113, no. 4, p. p66, 2006.
- [30] M. Trebble, "A preliminary evaluation of two and three phase flash initiation procedures," *Fluid Phase Equilibria*, vol. 53, 1989.

- [31] C. Wilbowo, L. O'Young and K. Ng, "Streamlining Crystallization Process Design," *Chemical Engineering Progress*, no. Jan, Jan 2004.
- [32] A. Buxton, A. Livingston and E. Pistikopoulos, "Optimal Design of Solvent Blends for Environmental Impact Minimization," *AIChE*, vol. 45, no. 4, 1999.
- [33] D. Constable, C. Jimenez-Gonzalez and R. Henderson, "Perspective on Solvent Use in the Pharmaceutical Industry," *Organic Process Research & Development*, vol. 11, 2007.
- [34] A. Jones, *Crystallization Process Systems.*, Butterworth-Heinemann, 2002.
- [35] A. Mersmann, *Crystallization Technology Handbook*, New York : Marcel Dekker Inc, 2001.
- [36] J. Mullin, *Crystallization*, Oxford: Butterworth-Heinemann Ltd., 1993.
- [37] H.-H. Tung, *Crystallization of Organic Compounds: An Industrial Perspective*, Wiley, 2009.
- [38] C. Leibovici and J. Neoschil, "A solution of Rachford-Rice equations for multiphase systems," *Fluid Phase Equilibria*, vol. 112, 1995.
- [39] M. Michelsen, "The isothermal flash problem part i: Stability," *Fluid Phase Equilibria*, vol. 9, 1982.
- [40] M. Michelsen, "The isothermal flash problem part ii: phase split calculation," *Fluid Phase Equilibria*, vol. 9, 1982.
- [41] R. Heldemann, "Computation of High Pressure Phase Equilibria," *Fluid Phase Equilibria*, vol. 14, 1983.
- [42] M. Michelsen and J. Mollerup, *Thermodynamic Models: Fundamentals and Computational Aspects*, Holte: Tie-Line Publications, 2007.
- [43] R. Henderson, C. Jimenez-Gonzalez, D. Constable, S. Alston, G. Inglis, G. Fisher, J. Sherwood, S. Binks and A. Curzons, "Expanding GSK's solvent selection guide – embedding sustainability into solvent selection starting at medicinal chemistry," *Green Chemistry*, vol. March, 2011.

Appendix G: Experimental Solubility Data

Experimental and Predicted Solubility of 4-Acetaminophenol in Ethanol and Ethyl Acetate at 298.15 ± 0.1 K								
Experimental Solubility Data					Predicted Solubility Values			
$x_{ethanol}$	$x_{ethyl\ acetate}$	x_i^{exp}			$x_{ethanol}$	$x_{ethyl\ acetate}$	$x_i^{predict}$ UNIFAC	$x_i^{predict}$ Mod UNIFAC
1.0000	0.0000	0.0647	±	0.0007	1.0000	0.0000	0.0459	0.0198
0.8982	0.1018	0.0710	±	0.0017	0.9091	0.0909	0.0686	0.0211
0.8007	0.1993	0.0741	±	0.0024	0.8182	0.1818	0.0867	0.0218
0.6983	0.3017	0.0718	±	0.0090	0.7273	0.2727	0.1003	0.0219
0.6015	0.3985	0.0667	±	0.0040	0.6364	0.3636	0.1102	0.0215
0.4986	0.5014	0.0612	±	0.0044	0.5455	0.4545	0.1168	0.0206
0.3947	0.6053	0.0516	±	0.0131	0.4545	0.5455	0.1206	0.0192
0.3008	0.6992	0.0433	±	0.0048	0.3636	0.6364	0.1217	0.0174
0.2040	0.7960	0.0314	±	0.0071	0.2727	0.7273	0.1201	0.0154
0.1135	0.8865	0.0218	±	0.0069	0.1818	0.8182	0.1156	0.0133
0.0000	1.0000	0.0067	±	0.0029	0.0909	0.9091	0.1079	0.0111
					0.0000	1.0000	0.0962	0.0090

Standard uncertainties u are $u(T) = 0.1K$, $u(P) = 0.002MPa$ and the standard relative uncertainties are $u_r(x_i^{exp}) = 0.0003$

Experimental and Predicted Solubility of 2-(4-Isobutylphenyl) Propionic Acid in Ethanol and Ethyl Acetate at 298.15 ± 0.1 K								
Experimental Solubility Data					Predicted Solubility Values			
$x_{ethanol}$	$x_{ethyl\ acetate}$	x_i^{exp}			$x_{ethanol}$	$x_{ethyl\ acetate}$	$x_i^{predict}$ UNIFAC	$x_i^{predict}$ Mod UNIFAC
1.0000	0.0000	0.1489	±	0.0106	1.0000	0.0000	0.1713	0.1047
0.9034	0.0966	0.1497	±	0.0118	0.9091	0.0909	0.2062	0.1361
0.7981	0.2019	0.3042	±	0.0224	0.8182	0.1818	0.2310	0.1631
0.7040	0.2960	0.2858	±	0.0050	0.7273	0.2727	0.2488	0.1849
0.5995	0.4005	0.2854	±	0.0671	0.6364	0.3636	0.2617	0.2021
0.5025	0.4975	0.3339	±	0.0112	0.5455	0.4545	0.2705	0.2151
0.4039	0.5961	0.3316	±	0.0210	0.4545	0.5455	0.2761	0.2243
0.3026	0.6974	0.3202	±	0.0512	0.3636	0.6364	0.2787	0.2302
0.1871	0.8129	0.3096	±	0.0400	0.2727	0.7273	0.2784	0.2327
0.0918	0.9082	0.3553	±	0.0257	0.1818	0.8182	0.2752	0.2319
0.0000	1.0000	0.3385	±	0.0612	0.0909	0.9091	0.2689	0.2276
					0.0000	1.0000	0.2589	0.2192

Standard uncertainties u are $u(T) = 0.1K$, $u(P) = 0.002MPa$ and the standard relative uncertainties are $u_r(x_i^{exp}) = 0.0003$

**Experimental and Predicted Solubility of Acetylsalicylic Acid
in Ethanol and Ethyl Acetate at 298.15 ± 0.1 K**

Experimental Solubility Data					Predicted Solubility Values			
$x_{ethanol}$	$x_{ethyl\ acetate}$	x_i^{exp}			$x_{ethanol}$	$x_{ethyl\ acetate}$	$x_i^{predict}$ UNIFAC	$x_i^{predict}$ Mod UNIFAC
1.000	0.000	0.068	±	0.007	1.0000	0.0000	0.0324	0.0193
0.899	0.101	0.085	±	0.004	0.9091	0.0909	0.0503	0.0276
0.800	0.200	0.093	±	0.008	0.8182	0.1818	0.0679	0.0366
0.700	0.300	0.107	±	0.014	0.7273	0.2727	0.0826	0.0454
0.600	0.400	0.108	±	0.017	0.6364	0.3636	0.0938	0.0533
0.501	0.499	0.109	±	0.003	0.5455	0.4545	0.1013	0.0598
0.413	0.587	0.104	±	0.008	0.4545	0.5455	0.1053	0.0644
0.296	0.704	0.098	±	0.004	0.3636	0.6364	0.1058	0.0671
0.202	0.798	0.083	±	0.005	0.2727	0.7273	0.1027	0.0678
0.101	0.899	0.071	±	0.005	0.1818	0.8182	0.0958	0.0663
0.000	1.000	0.054	±	0.000	0.0909	0.9091	0.0843	0.0628
					0.0000	1.0000	0.0675	0.0572

Standard uncertainties u are $u(T) = 0.1K$, $u(P) = 0.002MPa$ and the standard relative uncertainties are $u_r(x_i^{exp}) = 0.0003$