September 2013

Pharmaceutical Process Modeling, Optimization, and Control

Ehsan Sheikholeslamzadeh  
The University of Western Ontario

Supervisor
Sohrab Rohani  
The University of Western Ontario

Graduate Program in Chemical and Biochemical Engineering

A thesis submitted in partial fulfillment of the requirements for the degree in Doctor of Philosophy

© Ehsan Sheikholeslamzadeh 2013

Follow this and additional works at: https://ir.lib.uwo.ca/etd

Part of the Process Control and Systems Commons, Thermodynamics Commons, and the Transport Phenomena Commons

Recommended Citation
https://ir.lib.uwo.ca/etd/1622

This Dissertation/Thesis is brought to you for free and open access by Scholarship@Western. It has been accepted for inclusion in Electronic Thesis and Dissertation Repository by an authorized administrator of Scholarship@Western. For more information, please contact tadam@uwo.ca, wlswadmin@uwo.ca.
Pharmaceutical Process Modeling, Optimization, and Control

(Thesis format: Integrated-Article)

By

Ehsan Sheikholeslamzadeh

Graduate Program in Engineering Science
Department of Chemical and Biochemical Engineering
A thesis submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

The School of Graduate and Postdoctoral Studies
The University of Western Ontario
London, Ontario, Canada
© Ehsan Sheikholeslamzadeh 2013
Abstract

In this project, the aim was to achieve two important objectives and solve some challenges that the pharmaceutical industry is facing. It will be shown that the NRTL-SAC (Non-random Two Liquid Segment Activity Coefficient) model can best predict the solubility of different pharmaceutical and chemical components in pure and mixed solvents by comparing the results with the well-known model of the UNIFAC. The four parameters that are used in the NRTL-SAC model will be found through nonlinear parameter estimation technique. This project also covers the VLE, LLE, and VLLE phase behaviour calculations using the mentioned models to verify their applicability in industries that use solvents as their main process materials (such as pharmaceutical processes). It will be explained that the NRTL-SAC model is efficient and less complex than the UNIFAC model when dealing with multi-component systems of solvents. The solvent screening process is then modeled using a novel method of modeling and optimization which resulted in a significant change in the objective functions from single to binary solvent combinations. The proposed method shows the efficient selection of single, binary, and ternary solvent systems with the optimal crystallization operating conditions to achieve the desired objectives. However, the change from binary to ternary system of solvents did not have a significant effect on the performance functions. The study on the crystallization process of a polymorphic transformation phenomenon is another part of the project which was modeled and optimized. The novel method of modeling for polymorphic transformation of L-glutamic acid enabled us to develop an optimal control strategy of the system consisting of a variety of process conditions (such as seeded and un-seeded crystallization). The outcome of this part of the project gives a detailed understanding of polymorphic transformation systems with optimal conditions that can be implemented for such processes. Finally some useful experimental work that has been done in the area of nucleation and polymorphic transformation of L-glutamic acid using a powerful spectroscopic probe (Lasentec FBRM) will be explained. The nucleation detection and the change from metastable polymorph to the stable one can be performed using the in-situ FBRM which was used in this project.
Keywords:

Mathematical modeling, Solubility measurement and prediction, Activity coefficient models, NRTL-SAC and UNIFAC, VLE and VLLE systems of solvents, Solvent screening, Nonlinear programming (NLP), Process optimization, Polymorphic transformation, Optimal control, Nucleation detection, FBRM and online measurement, XRPD
Co-Authorship Statement

Chapter 02: Writing the manuscript, experimental work, data gathering, and parameter estimation were performed by the author. This work was supervised by Dr. Sohrab Rohani, whose input was very important for the improvement of its content. His recommendations were incorporated in this chapter. A version of this chapter has been published in the Industrial & Engineering Chemistry Research Journal:


Chapter 03: The original draft of this chapter was prepared by the author. All the required data was gathered and analysed by the written codes in Matlab environment. The implementation of the models were done by the author. The valuable recommendations by Dr. Sohrab Rohani strengthened the content of the paper. A version of this chapter has been published in Fluid Phase Equilibria Journal:


Chapter 04: The original draft of this chapter was prepared by the author. Further review was performed by Dr. Sohrab Rohani, whose recommendations were very helpful in improving the contents of the manuscript. Dr. Chau-Chyun Chen from Aspen Technology gave some valuable helps in performing a few revisions on the manuscript. A version of this chapter has been published in Industrial & Engineering Chemistry Research Journal:

Chapter 05: The original draft of this chapter was prepared by the author. Further review was performed by Dr. Sohrab Rohani, whose recommendations were very helpful in improving the contents of the manuscript. A version of this chapter has been published in Industrial & Engineering Chemistry Research Journal:


Chapter 06: This chapter contains a full experimental work which its draft will be prepared and reviewed by Dr. Sohrab Rohani for submission to a suitable scientific journal.
Acknowledgments

I would like to take this opportunity to acknowledge and thank those who made this work possible. First, I would like to express my appreciation to my supervisor, Dr. Sohrab Rohani, for his warm support and invaluable feedback throughout the progress of this project. He was always open to novel ideas and also provided a good atmosphere that helped me to reach my full potential. I am always grateful for that.

I deeply appreciate the recommendations and invaluable feedback from Dr. Chau-Chyun Chen from Aspen Technology regarding the solvent screening project. His advice was beneficial in revising the manuscript related to the solvent screening work.

My several thanks to all of the people in our research group whom their cooperation in doing my project lead to successful outcomes in this project.

I would like to give special thanks to my mother, father, brother, and my sisters, who never gave up to support and always, motivated me to pursue my research during my PhD at Western University. My brother helped me a lot in all aspects of my life during his stay here at Western while doing his masters in electrical engineering. I hope I have made them proud. They deserve it.

Finally, I would like to thank all of my friends here in Canada and overseas for their support. Listing all of them would fill a book in itself, so I merely will have to limit myself to a few words: Thank you all.
Table of Contents

Title page ....................................................................................................................................... i
Abstract ......................................................................................................................................... ii
Co-Authorship Statement ............................................................................................................ iv
Acknowledgments ...................................................................................................................... vi
Table of Contents ........................................................................................................................ vii
List of Tables ................................................................................................................................ xii
List of Figures ............................................................................................................................. xv
Nomenclature .............................................................................................................................. xx

Chapter 01: Introduction ............................................................................................................ 1
1.1. Project outline ........................................................................................................................ 8
1.2. References ............................................................................................................................ 11

Chapter 02: Solubility Measurement and Prediction of Pharmaceutical and Chemical Compounds in Pure and Mixed Solvents ................................................................................. 12
2.1. Introduction .......................................................................................................................... 13
2.2. Thermodynamic theory and modeling ................................................................................. 16
   2.2.1. Universal Functional Activity Coefficient (UNIFAC) model ...................................... 18
   2.2.2. Non-random Two-liquid Segment Activity Coefficient (NRTL-SAC) ........................ 22
Chapter 03: Vapour-Liquid and Vapour-Liquid-Liquid Equilibrium Modeling for Binary, Ternary, and Quaternary Systems of Solvents

3.1. Introduction

3.2. Theoretical background

3.2.1. Thermodynamic models

3.2.1.1. Phase behaviour calculations

3.2.1.2. Vapour-liquid equilibrium

3.2.1.3. Liquid-liquid equilibrium
3.2.1.4. Vapour-liquid-liquid equilibrium ................................................................. 53

3.3. Test cases .............................................................................................................. 53

3.3.1. VLE system of Toluene-Ethanol and Toluene-Isopropanol at different pressures ..... 53

3.3.2. VLE data for binary and ternary systems of Ethanol-Water-Ethylene Glycol ............ 57

3.3.3. Binary and ternary system of Diethyl Ether-Methanol-Butanol .............................. 60

3.3.4. VLE and VLLE equilibrium of Water-Ethanol-Hexane-Toluene and Water-Ethanol-
Cyclohexane-Isooctane ........................................................................................................ 62

3.4. Conclusion .............................................................................................................. 66

3.5. References .............................................................................................................. 67

Chapter 04: Optimal Solvent Screening for the Crystallization of Pharmaceutical
Compounds from Multi-Solvent Systems ......................................................................... 70

4.1. Introduction ............................................................................................................ 71

4.2. Thermodynamic background .................................................................................. 72

4.2.1. VLE, LLE, and SLE prediction ............................................................................ 73

4.2.2. Optimization method and procedure .................................................................... 74

4.2.3. Mathematical formulation of the model ............................................................... 76

4.3. Results and discussion ............................................................................................ 83

4.3.1. Single solvent screening ...................................................................................... 84

4.3.2. Binary solvent screening ..................................................................................... 87

4.3.3. Ternary solvent screening .................................................................................... 92

4.4. Conclusion .............................................................................................................. 96
Chapter 05: Modeling and Optimal Control of Solution Mediated Polymorphic Transformation of L-Glutamic Acid  ............................................................................................................................ 99

5.1. Introduction ........................................................................................................................................................................ 100

5.2. Process modeling and simulation .............................................................................................................................................. 102

5.2.1. Numerical solution of the model ........................................................................................................................................... 103

5.2.2. Objective functions for the optimal control of the process ........................................................................................................ 110

5.3. The model compound and the simulation using the conventional cooling policies .............................................................. 111

Case 1. Supersaturation with respect to both forms (Region A in Figure 5-1) ................................................. 112

Case 2. Supersaturation with respect to the stable form and undersaturation with respect to the metastable form ........................................................................................................................................... 115

Case 2.1. Un-seeded condition (Region B in Figure 5-1) .......................................................................................... 115

Case 2.2. Seeded only with the stable form (Region B in Figure 5-1) .......................................................... 118

Case 2.3. Seeded with both polymorphs (Region B in Figure 5-1) ............................................................................. 121

5.4. Optimal control for polymorphic transformation in the presence of dissolution process ........................................................................................................................................................................... 124

5.5. Conclusion ........................................................................................................................................................................................................ 128

5.6. References ........................................................................................................................................................................................... 128

Chapter 06: Experimental Studies on Crystallization of Polymorphic Transformation of L-Glutamic Acid .............................................................................................................................................. 131
6.1. Introduction .......................................................... 132

6.2. Nucleation detection and kinetic rate .......................................................... 134

   6.2.1. Experimental procedure ......................................................... 134

   6.2.2. Results and discussion ......................................................... 135

6.3. Polymorphic transformation process ....................................................... 140

6.4. Conclusion ............................................................................. 143

6.5. References ............................................................................ 144

Chapter 07: Conclusions and Future Works .............................................. 146

7.1. Solid solubility prediction and measurement ......................................... 147

7.2. VLE and VLLE prediction for industrial solvent systems .................... 148

7.3. Solvent screening process .................................................................. 148

7.4. Polymorphic transformation process modeling and optimal control .......... 148

7.5. Experimental results for crystallization of L-glutamic acid ................... 149

7.6. Future Works ....................................................................... 149

Appendix: ..................................................................................... 152
List of Tables

Chapter 02

Table 2-1. NRTL binary interaction constants for conceptual segments ........................................... 25

Table 2-2. Physical properties of the model compounds used for thermodynamic modeling and prediction ................................................................................................................................................. 29

Table 2-3. Functional groups that are defined in UNIFAC and their replication in two different structures of 3-pentadecylphenol ........................................................................................................................................ 31

Table 2-4. Optimized segment numbers for the model compounds using the NRTL-SAC method .................................................................................................................................................................. 32

Table 2-5. Experimental solubility of 3-pentadecylphenol in pure IPA and its mixtures with water ......................................................................................................................................................... 33

Table 2-6. Functional groups that have been used in modeling with the UNIFAC model ............................................................................................................................................................................. 36

Table 2-7. Average relative deviation for the UNIFAC and NRTL-SAC in three methods ...... 40

Table 2-8. Optimized segment numbers for each method used for NRTL-SAC model .............. 40

Chapter 03

Table 3-1. Deviation of predictive model results from corresponding experimental values of Chen et al. [17] ............................................................................................................................................................. 55

Table 3-2. Absolute relative deviation (%ARD) of predicted and experimental vapour mole fraction of components in each binary system ........................................................................................................ 58
Table 3-3. Binary solution deviation of model vs. experimental data for three pairs of Diethyl ether-Methanol-Butanol (%ARD) ................................................................. 61

Table 3-4. Average absolute deviation (%AAD) for predictive models and two correlative models for a ternary system of Diethyl Ether-Methanol-Butanol .............................................. 62

Table 3-5. %ARD for the two quaternary systems of VLLE .......................................................... 64

Chapter 04

Table 4-1. Segment numbers, Antoine coefficients of vapour pressure, and normal boiling points of solvents studied in the current work [19] with the class number assigned to each solvent by Food and Drug Administration (FDA) [18] ................................................................................ 75

Table 4-1. Continued .................................................................................................................. 76

Table 4-2. Segment numbers and physical properties of selected components from the literature ........................................................................................................................................ 83

Table 4-3. The optimal single solvent for each of model components with their starting and ending operating temperature at 40°C and 20°C, respectively ................................................. 85

Table 4-4. Optimal conditions of operation for cooling and antisolvent crystallization for the first objective function ($J_1$) ........................................................................................................... 91

Table 4-5. Optimum yield of first objective function ($J_1$) for model components with its corresponding yield based on second objective and the CPU time allocated for calculation for binary solvent optimization ........................................................................................................... 92

Table 4-6. Optimal solvent mole ratios, volume ratios, and temperature for the start and final part of crystallization .............................................................................................................. 95

Table 4-7. Optimal values of yield of crystallization for ternary mixture of solvents based on the first objective function and the second objective function ................................................. 96
Table 4-8. Optimal yield values for single, binary, and ternary solvent combinations .......... 97

Chapter 05

Table 5-1. Kinetic parameters with their values and units ......................................................... 112

Table 5-2. Seed size distribution parameters for stable (α) and metastable (β) forms used in equation (5-7) ........................................................................................................................................... 113

Table 5-3. Optimal values of mass of stable, metastable, and the size of stable form with their values of conventional cooling policies .......................................................................................... 125

Table 5-4. Important results of the crystallization process using the two methods of MOC and MOMC ........................................................................................................................................ 127

Chapter 06

Table 6-1. Kinetic data on the nucleation of L-glutamic acid from FBRM measurement ........ 138
List of Figures

Chapter 01

Figure 1-1. Schematic diagram of typical processes that involved in the pharmaceutical manufacturing ............................................................................................................................... 3

Figure 1-2. Project summary with the brief title of each section ................................................................. 10

Chapter 02

Figure 2-1. The algorithm of converging to the solubility of a ternary system using UNIFAC model ........................................................................................................................................... 17

Figure 2-2. Algorithm flowchart for parameter estimation using NRTL-SAC model ............... 22

Figure 2-3. The chemical structure of 3-pentadecylphenol in compact form (right) and the length of bonds in Angstrom (Å) (left) ................................................................................................... 24

Figure 2-4. Chemical structure of lovastatin ................................................................. 26

Figure 2-5. Chemical structure of valsartan ................................................................. 27

Figure 2-6. Characterization of re-crystallized 3-pentadecylphenol with DSC (left) and XRPD (right) ......................................................... ................................................................................ 28

Figure 2-7. Experimental points of the solubility of 3-pentadecylphenol in two solvents and their curve of estimation using NRTL-SAC model. The curves of UNFAC are also shown here for comparison ................................................................. 32

Figure 2-8. Experimental and predicted solubilities for 3-pentadecylphenol in solvents using UNIFAC and NRTL-SAC models ................................................................. 34
Figure 2-9. Experimental and predicted values of the solubility data of 3-pentadecylphenol in IPA-water mixture at different temperatures ................................................................. 35

Figure 2-10. Schematic view of functional groups presented in lovastatin ........................................... 36

Figure 2-11. Experimental data with the predictions of solubility in four selected solvents using three methods of the NRTL-SAC and the UNIFAC method ....................................................... 37

Figure 2-12. Predicted and experimental values of valsartan in two solvents: left: acetonitrile, right: butyl acetate ....................................................................................................................... 39

Figure 2-13. Solubility prediction of valsartan in three ethyl acetate + hexane mixtures for different mole percent of ethyl acetate as a solvent ............................................................... 42

Figure 2-14. Schematic diagram for finding the fugacity change from solid to liquid state of a pure substance ............................................................................................................................. 43

Chapter 03

Figure 3-1. Predicted and experimental results for systems (A) ethanol-toluene at 101.3kPa, (B) ethanol-toluene at 201.3kPa, (c) isopropanol-toluene at 101.3kPa, and (D) isopropanol-toluene at 201.3kPa ...................................................................................................................................... 56

Figure 3-2. Txy diagram of VLE systems; (A) ethanol-water, (B) ethanol-ethylene glycol, and (C) water-ethylene glycol all at 101.3kPa ....................................................................................................................... 57

Figure 3-3. Vapour phase mole fractions of ethanol (top-left), water (top-right), ethylene glycol (bottom-left), and saturated temperature of the solution (bottom-right) from experimental data and theoretical models for the ternary system of ethanol-water-ethylene glycol .................... 59

Figure 3-4. Three binary phase diagrams of (A) diethyl ether-methanol, (B) diethyl ether-butanol, and (C) methanol-butanol at 101.3kPa from experiment and models ......................... 60

Figure 3-5. The results of prediction vs. experimental runs for VLE system of water-hexane-toluene; (A) bubble point temperature, (B) water, (C) hexane, and (D) toluene mole fraction in vapour phase ............................................................................................................................. 63
Figure 3-6. Predicted and experimental phase mole fractions of three species in a quaternary system of VLLE, 1) system of Water-Ethanol-Hexane-Toluene, prediction based on (A) organic phase and (B) aqueous phase, 2) system of Water-Ethanol-Cyclohexane-Isooctane, prediction based on (C) organic phase and (D) aqueous phase .......................................................... 65

Chapter 04

Figure 4-1. Chemical structure of model molecules used in current study, a: lovastatin, b: valsartan, c: paracetamol, d: budesonide, e: allopurinol, f: furosemide, g: sulfadiazine .......... 84

Figure 4-2. The obtained solubility curves for valsartan in pyridine and water from NRTL-SAC model ........................................................................................................................................... 86

Figure 4-3. Comparative illustration of yields for two objective functions for sulfadiazine ..... 87

Figure 4-4. Solubility of lovastatin in the mixture of cumene and acetonitrile in the temperature range of 300-340K ...................................................................................................................... 89

Chapter 05

Figure 5-1. The schematic representation of the polymorphs solubility and three distinct regions in between the curves for an arbitrary component; (Region A) nucleation and growth of stable and metastable forms, (Region B) nucleation and growth of stable form and dissolution of the metastable form, (Region C) dissolution of both forms .......................................................... 102

Figure 5-2. A sample class (j\textsuperscript{th} class) and its neighbour segments; the arrows represent (1) the growth of particles from a class of lower average size, (2) crystal growth from the main class, (3) dissolution from a class higher in average size, (4) crystal dissolution from the main class .... 107

Figure 5-3. Algorithm for solving polymorphic transformation using MOMC method ........ 109
Figure 5-4. (a) Supersaturation profile, (b) Three cooling policies implemented on the crystallization system, (c) the ratio of the stable to metastable masses, and (d) mass of metastable form of L-glutamic acid over time (for the case 1) .......................................................................................... 114

Figure 5-5. Time evolution of supersaturation with respect to the solubility of the stable form (top) and average size of the stable form based on number-weighted definition (bottom) in case of un-seeded operation .......................................................................................................................... 116

Figure 5-6. Difference of the mass of stable form produced from the metastable in a wide range of cooling policies for three different initial concentrations .................................................................................. 118

Figure 5-7. Concentration evolution (top) and supersaturation (bottom) over time for three different seeding average sizes (µm) ............................................................................................................................................... 119

Figure 5-8. Supersaturation profile (top) and average size of particles (bottom) for three different masses of stable form seeds ............................................................................................................................................ 120

Figure 5-9. Polymorphic transformation in the presence of stable and metastable form seeds using three cooling policies, (a) average size of metastable, (b) average size of stable form, (c) supersaturation with respect to stable solubility, (d) mass of stable and metastable form ...... 123

Figure 5-10. Optimal trajectory of (a) temperature for different objective functions, (b) average size of stable form ($J_2$), (c) mass of stable form ($J_1$), and (d) Mass of metastable form ($J_3$) .... 126

Chapter 06

Figure 6-1. The solubility and nucleation curves for an arbitrary component ......................... 133

Figure 6-2. FBRM user interface program ...................................................................................... 133

Figure 6-3. (a) Measured saturated and nucleation points of solutions of different with the (b) solubility and nucleation curves for each cooling rate ............................................................... 136

Figure 6-4. A schematic diagram of cyclic automated operation for the saturation temperature of 40°C ......................................................................................................................................................... 137
Figure 6-5. Diagrams of the plot of logarithm of cooling rate versus logarithm of MSZW for four different saturation temperatures at constant stirring rate (250rpm) ........................................ 139

Figure 6-6. Particle size distribution for five different samples of alpha to beta L-glutamic acid ..................................................................................................................................................... 140

Figure 6-7. Four distinct regions of crystallization of L-glutamic acid: (a) complete dissolution of stable form, (b) Addition and dissolution of metastable form, (c) Polymorphic transformation, dissolution of metastable form and formation of stable form, and (d) complete dissolution of metastable form and formation of stable form ................................................................. 141

Figure 6-8. X-ray powder diffraction of the polymorphs of L-glutamic acid ................................. 142
Nomenclature

Chapter 02

$\gamma_i$ Activity Coefficient

$x_j$ Mole Fraction of Species $j$ in solution

$\Lambda_{zi}$ Binary Interaction Coefficient

$V_i^L$ Molar Volume of Saturated Liquid of Species $i$

$R$ Universal Gas Constant

$T$ Temperature

$f_i(T,P)$ Fugacity of Species $i$

$\Delta H_{\text{fus}}$ Enthalpy Change of Fusion

$\Delta C_p$ Change in the Heat Capacity at Constant Pressure from Solid to Liquid State

$T_t$ Triple Point Temperature

$T_m$ Melting Point Temperature

$x_s$ Mole Fraction of a Solute in Solution

Chapter 03

$x_i^\alpha$ Mole Fraction of Species $i$ in Phase $\alpha$

$P_i^{\text{sat}}$ Saturated Pressure of Species $i$

$P_{\text{tot}}$ Total Pressure of the System

$A_i, B_i, \text{and } C_i$ Coefficients Used in Antoine Equation

Chapter 04

$S_{\text{initial}}$ Initial Solute Concentration per Mass of Solvent
\( S_{\text{final}} \) Final Solute Concentration per Mass of Solvent

\( M_s \) Mass of Solvent

\( J_1 \) and \( J_2 \) Objective Functions Used in Optimization

\( T_{\text{initial}} \) Initial Operating Temperature of Crystallization

\( T_{\text{final}} \) Final Operating Temperature of Crystallization

\( T_{\text{melting}} \) Melting Point Temperature of the Solid

\( T_{\text{solvent sat}} \) Solvent Bubble Point Temperature

\( R_i \) Solvent Ratio of \( i^{th} \) to \( i+1^{th} \) Solvent

\( x_{i,0}^s \) Initial Mole Fraction of Species \( i \) in solution

\( x_{i,f}^s \) Final Mole Fraction of Species \( i \) in solution

**Chapter 05**

\( f_i \) Number Density Function for Particle \( i \)

\( G_i \) Growth Rate of Species \( i \)

\( D_i \) Dissolution Rate of Species \( i \)

\( B_{\text{nucleation},i} \) Birth Rate of Species \( i \)

\( \delta \) Delta Dirac Function

\( C(t) \) Solute Concentration at time \( t \)

\( C_s \) Heat Capacity of Solid \( s \)

\( \rho_s \) Density of Solid \( s \)

\( \Delta H_{\text{crystallization}} \) Enthalpy Change of Crystallization

\( \Delta H_{\text{dissolution}} \) Enthalpy Change of Dissolution
UA  Overall Heat Transfer Coefficient of the System

$\Delta T_{lm}$  Log Mean Temperature

$\alpha_{v,i}$  Volume Shape Factor of Solid $i$

$\rho_{c,i}$  Density of Species $i$ in the solution

$\mu_{i}^{j}$  $j^{th}$ Moment of $i^{th}$ Polymorph

$\varepsilon_{i}, \sigma_{seed,i}, I_{seed,i}$  Parameters Used in Solid Distribution of the Seeds

$V_j$  Size of the $j^{th}$ Class

$V(t)$  Total Volume of the Solution

$J_1, J_2, and J_3$  Objective Functions Used in Dynamic Optimization Procedure

Chapter 06

$S$  Supersaturation

$CR$  Cooling Rate

$J_n$  Nucleation Rate

$MSZW$  Metastable Zone Width

$m, k_n$  Nucleation Parameters
Chapter 1

Introduction
This chapter provides a brief overview of the materials that will be discussed in detail in the next chapters of this thesis. In this chapter, the project subject with its explanation is first demonstrated briefly and then, the main challenges that the industry is dealing with will be mentioned. The questions that were raised in this field directed us in performing different sections of the project. The objective(s) that will be addressed in each chapter with a brief description of the methods and procedures are discussed in subsequent paragraphs.

One of the main unit operations in the chemical industry, especially in the pharmaceutical industry is the crystallization. In the pharmaceutical and chemical industry, the development of a drug or chemical compound with desirable characteristics is of prime significance. In achieving this objective, the choice of a proper solvent for crystallization, operating condition, and design parameters is very important to lead to better separation and purification, lower cost, and better product quality. The solids that are produced in a crystallization unit will affect the downstream unit processes, such as filtration and drying [1]. Solution crystallization is widely used in chemical and pharmaceutical industry during final and intermediate parts of purification and separation processes [2]. The chemical synthesis of a drug molecule often involves several synthetic steps in series and each step contains many unit operations. Each unit which is involved in the production of a specific pharmaceutical product has some challenges regarding the method of operation to reach to the optimal performance in production. A schematic diagram of a general pharmaceutical process for a solid active pharmaceutical ingredient (API) synthesis and development to the final product is shown in Figure 1-1.

In the current project, we will focus on the product development and the crystallization process. These two stages of drug production have significant effect on the performance of other unit operations downstream of the crystallization process. In developing a specific pharmaceutical product, there are many issues that the researchers have to deal with. Some examples include the physico-chemical properties of the potential component under study, such as potency, solubility, pK_a, lipophilicity, metabolic stability, absorption, etc [2]. The drug candidate should have the best characteristic in terms of the mentioned properties. Traditionally, the task of selection of the potential component that has all the desirable properties was time-intensive and very expensive.
Figure 1-1. Schematic diagram of typical processes that involved in the pharmaceutical manufacturing

Since the amount of the component under study at early stages of the development is low and often expensive, the researchers had to manage the experimental work so as to minimize the loss, and therefore, a few data could be found. This problem is more significant when the solubility of the candidate drug has to be examined in several solvents and their mixtures [3]. Several pharmaceutical companies have implemented high-throughput screening to profile compounds in terms of properties such as solubility [4]. Although this method is helpful in identification of the early failures in development without spending resources on clinical trials, it needs proper amount of solvent(s) and pharmaceutical components, which is often expensive. The mentioned reasons directed us to build a comprehensive algorithm which can overcome those problems in an efficient and accurate way. The following are the main questions that need to be addressed in the synthesis and development of a pharmaceutical product:

1. How to predict the solubility of a chemical or pharmaceutical in pure and mixed solvent(s) at different operating temperatures and solvent compositions?
2. What solvent or mixture of solvents will result in a favorable operation of crystallization of a given chemical component?
3. How to comprehensively model a solvent screening process for a specific pharmaceutical component?
4. Which conditions (temperature and solvent composition) can optimize the production of an API?
5. What thermodynamic model(s) to pick to investigate all the possible behaviors affecting the process, such as the vapor-liquid (VLE), liquid-liquid (LLE), and solid-liquid equilibrium (SLE)?

Chapters 2-4 cover all the details about the above questions.

As it is mentioned before, measuring physical properties of the new discovered drugs is time-consuming and expensive. Therefore, the need to have a reliable and accurate method of finding the solubility of pharmaceutical compounds in solvents or solvent mixtures is beneficial. This subject is completely discussed and addressed in chapter 2. The method that is picked for modeling the solubility should have the least dependency on experimental data and provide accurate predictions. As a result, applying different predictive thermodynamic models and comparing them to find which method can result in accurate solubility prediction is desirable. The outcome of this part of the project can help the researchers in the pharmaceutical industry and more generally, in the particulate and fine chemical industry, to understand the behaviour of the solids in suspensions and their solubility in different pure and mixed solvents. The bioavailability of a pharmaceutical compound depends heavily on its solubility and dissolution rate in the blood stream. Having an accurate predictive model describing these properties, results in less dependency on the process measurement, saves time, and cost. In order to address this objective, we applied the predictive methods of the UNIFAC and a recent developed model, the NRTL-SAC, to predict the phase behaviour of solid-liquid, vapor-liquid, liquid-liquid systems, and solubility prediction of different pharmaceuticals. Three model compounds were selected to study and experiments were performed to characterize and measure the solubility of the compounds in different pure and mixed solvents. The measured data and predicted values are compared using two predictive models.
The process of selecting the right solvent combination for a crystallization process is referred to as solvent screening. Normally, the solvent screening is done with high-throughput experiments at different operating conditions. This process is very time-intensive and costly. Having the knowledge of the phase behaviour of different solvent compositions helps one to select the proper operating conditions for the best production performance. If the operating temperature for the crystallization process is out of the safe range (i.e., the starting crystallization temperature exceeds the solute’s melting point or solvent mixture’s bubble point temperature), there will be a failure in the process operation that leads to loss of valuable products and time. As a result, prior to designing an optimized process for production of a pharmaceutical, we have to ensure that the operating conditions will meet the safe and reliable behaviour of the production unit. Chapter 3 discusses the phase behaviour of different solvents that are used in pharmaceutical industries. The equilibrium condition of a variety of binary, ternary, and quaternary combinations of solvents at various operating conditions (pressure and temperature) will be modeled and verified. The use of a powerful method to predict different solvent compositions within a desirable accuracy is not only helpful to the pharmaceutical industry, but also to other common chemical processes. We developed a preliminary software program to model the phase behaviour of the solvent combinations using the new robust model of NRTL-SAC.

After finding the best operating conditions for each combination of solvents, the production rate of the crystallization process has to be calculated. By comparing the production rates at different combinations of solvents, one can pick a solvent or a solvent mixture which results in the maximum production rate. It should be noted that in addition to the production rate, other important objective functions can be studied (e.g., the minimum mass of solvent usage). At this stage of process design, a non-linear programming should be used in which the linear and non-linear constraints are included. The nonlinear constraint originates from the phase equilibrium calculations. It is worth noting that the crystallization process in general, can be run by three methods: cooling, antisolvent addition, and evaporation, or a combination thereof. For the current study, we examined the case of cooling and antisolvent crystallization process, simultaneously. An optimization procedure was developed to examine all the possible combinations of cooling and antisolvent crystallization based on different objective functions. The procedure used to obtain the optimal process conditions is discussed in chapter 4. The
outcome of the proposed method in chapter 4 is helpful for preliminary studies of solvent screening in picking the optimum solvent combination and also, finding the best operating conditions of the process. It should be noted that the proposed algorithm is capable of handling other related processes that are involved in the production of a chemical compound from solvent mixture (such as reactive or extractive distillation).

After the drug is selected as a potential candidate for production in large scale, the necessary processes for manufacturing the component have to be designed. The competition in the pharmaceutical industry and on the other hand, the increase in demand by the society, has made the industry focusing more on the efficiency of the manufacturing units [5]. The process modeling and simulation is a powerful means to facilitate the task of moving from synthesizing to large-scale manufacturing of a component. The solid form of a drug can have different physical properties in terms of its crystalline shape. The most preferred crystalline shape is the one which is thermodynamically stable. However, the stable form of the component may show poor solubility and dissolution rate in the human body. The solubility and absorbance of a drug in the body are the main parameters for the selection a suitable form (polymorph) of a pharmaceutical component. In this case, other forms of that component (less stable forms) may be selected for synthesis and development [6]. The separation of the crystals of different polymorphic forms of a drug is a great challenge in industry [7]. Most of the studies in the field of polymorphic crystallization are qualitative descriptions. The complexity of the quantitative study of this process comes from the presence of different phenomena such as thermodynamics, kinetics, reactor dynamics, and population balances [7-11]. This issue is more difficult to study when other related phenomena such as agglomeration and breakage of the particles are also present. The study on the dynamics and process modeling of polymorphic crystallization systems has recently started and thus, there is a wide range of unsolved problems in this field. Therefore, we investigated the polymorphic processes, their comprehensive modeling, and dynamic optimization of such processes. The followings are the main questions that can be raised to be addressed in this area of study:

1. Which method can be used to model a polymorphic crystallization process?
2. How to reduce the time of computation for modeling polymorphic transformation processes?
3. How to minimize or maximize the production of a specific polymorph?
4. How to generate the particle size distribution of different polymorphs during and at the end of a batch crystallization process?
5. How to implement the dynamic optimization procedure to model a polymorphic transformation process in different conditions (comprehensive modeling and optimization)?

To answer all the above concerns, one needs to know the fundamentals of crystallization and different effective parameters that influence the process performance. For example, the method of implementing the cooling rate has to be studied to achieve the desirable product with optimum crystal size distribution. Most of the newly discovered pharmaceuticals show different crystalline forms in solid state that is called polymorphism. Depending on a method which is implemented on the process, the product can have a mixture of polymorphs or just one form. From all of the possible crystalline forms of a pharmaceutical component, a very few (in most cases one) of those are bioavailable. Different polymorphs of a pharmaceutical component have different properties in dissolving and absorbing by the body. It is important for the pharmaceutical industry to produce as much desirable and bioavailable drug as possible to gain the highest benefit. Chapter 5 discusses the conventional and optimal cooling profiles which can maximize some objective functions that are currently used in industry for achieving the desirable product. The polymorphic transformation process is studied in this chapter. The model component which we studied was L-glutamic acid that has stable and metastable forms. A comprehensive modeling of the process can be used effectively by the particulate industry that are facing complex polymorphic systems and want to find the dynamic operating conditions that lead to their desirable objectives [8, 11]. It will be shown that a method of moments (MoM) that is used for performing the population and moments of particles cannot be applied for generation of particle size distribution (PSD). In addition to the production rate of each polymorph and final average size of the product, the size distribution of the particles that are produced at the end of the batch is important. This will be more significant when designing filtration and particle transportation systems since they are dependent on the size distribution of the solids that are processed. The conventional method of generating PSD is the method of classes (MoC) which has been used to model crystallization systems containing only one form of crystal. The main disadvantage of this method is its time-demanding calculation. In our study, we have combined the method of classes
(MoC) with the method of moments (MoM) to: 1) reduce the CPU time and 2) generate PSD of both stable and metastable polymorphs during the crystallization process. Although, there are a few other methods for calculating PSD from the population balance equation (such as finite element method), they are difficult to model. Our new proposed method (MoMC) generates the PSD in a few seconds with the same accuracy of other methods. The mentioned method of MoMC and all the related materials will be discussed in chapter 5. The outcome of chapter 5 can be used effectively by the research centers of industries to study the effect of different parameters in the size distribution of the polymorphs. The proposed MoMC method can also be effectively used in real-time control of polymorphic transformation systems where the size distribution has to be adjusted within a specified range.

We also have done several experiments to monitor the nucleation and polymorphic transformation behaviour of L-glutamic acid with different cooling rates. The nucleation behaviour of L-glutamic acid with different cooling and agitation rates will be discussed in chapter 6. The development of the new methods of monitoring the crystal size distribution such as Focused Beam Reflectance Method (FBRM) has made the online control of crystallization systems more easily than before. At first, the use of FBRM probe in monitoring the onset of nucleation is shown. Then, it is used for detecting the cloud and saturated points corresponding to different solution concentrations. The automated method for doing such experiment will be explained in chapter 6. Next, the FBRM technology will be incorporated in detecting the polymorphic transformation of L-glutamic acid from its metastable form to stable during the course of crystallization. The FBRM can qualitatively show the time and the size distribution of polymorphic transformation of the particles. Finally, the use of XRPD as another tool for polymorph detection in offline mode will be presented.

1.1. Project Outline

The project outline in schematic flowchart is shown in Figure 1-2. Based on this flowchart, the whole project is divided into two phases: 1) development phase, in which the pharmaceutical is discovered and developed and 2) production phase, in which the pharmaceutical component is manufactured using different chemical and physical processes. In order to answer the previous mentioned questions, we picked predictive thermodynamic models to verify their capability in generating accurate and reliable data to be used for drug development. The most important part
of the pharmaceutical development is the solvent screening process. This is best studied and modeled in our work. In order to get the comprehensive view of the solvent screening process, we needed to ensure the solvent(s) are operated within the safe and reliable process conditions. Therefore, different phase behaviors of the solvent(s) with the solids, such as VLE and VLLE were studied in the development phase.

In the development phase, the polymorphic transformation of complex components using the current and the novel method (which was developed in our work), were done. As most of the pharmaceutical processes are performed in batch operation, the dynamic optimization procedure was used to achieve two different important objectives.

The conclusion and future works of the stated projects will be discussed in chapter 7. The program codes that were developed in Matlab environment are listed in the Appendix.
Figure 1-2. Project summary with the brief title of each section
1.2. References


Chapter 2

Solubility Measurement and Prediction of Pharmaceutical and Chemical Compounds in Pure and Mixed Solvents

A version of this chapter has been published as:

2.1. Introduction

One of the main properties of the polymorphic compounds is their solubility difference in solvents and solvent mixtures. With this important property, different polymorphs of a drug can be separated by crystallization. In the synthesis of pharmaceuticals, there are many unit operations and factors which affect the overall performance of the unit, such as the solvent selection, operating temperature and supersaturation [1]. Empirical selection of solvents requires extensive experimentation and high cost [2]. The predictive thermodynamic models can be a good choice in estimation of the phase behaviour and solubility of drugs in different solvents and solvent mixtures [3]. There have been many thermodynamic models for the prediction of the phase behaviour of vapour-liquid and liquid-liquid systems, such as the Margules equation [4], the Wilson equation [5], the Van Laar equation [6], the NRTL equation [7], and the UNIQUAC equation [8]. These models can also be used for solid-liquid equilibrium behaviour. Generally, the equations of activity coefficient for solubility prediction can be divided into the two categories:

1. The correlative models, which require many experimental data at different conditions. In some cases the data on ternary mixtures are also needed, such as for Wilson’s equation [5].
2. The group-contribution and predictive models, which only require the chemical structure of the molecule and/or a few experimental data points to predict the phase behavior of the solid in different solvents, such as the UNIFAC and NRTL-SAC models.

From the two above categories, the first one is not very useful for solubility prediction and solvent screening purposes [1]. The main reason for this is the lack of experimental data for the binary interaction parameters of the solute-solvent, solute-antisolvent, and solvent-antisolvent systems. As an example, the activity coefficient from the Wilson’s equation of state is found from [5]:

\[ \ln \gamma_i = -\ln \left[ \sum_{j=1}^{n} x_j \Lambda_{ij} \right] + 1 - \sum_{z=1}^{n} \frac{x_z \Lambda_{zi}}{\sum_{k=1}^{n} x_k \Lambda_{zk}} \quad (2 - 1) \]

The binary interaction parameter is:
\[ \Lambda_{ij} = \frac{V_i^L}{V_i^L \exp \left( -\frac{\lambda_{ij} - \lambda_{ii}}{RT} \right) } \]  \hspace{1cm} (2-2)

Where \( i \) and \( j \) refer to the compounds present in the solution. From the equation (2-1) it can be seen that for obtaining the activity coefficient of a component 1 in a pure solvent 2, we need four interaction parameters (\( \Lambda_{12}, \Lambda_{21}, \Lambda_{11} \) and \( \Lambda_{22} \)) which are temperature dependent. In addition, from equation (2-2) it is evident that for calculating the value of the binary interaction parameters, additional experimental data, such as molar volume is needed. Other models which belong to the first category have the same limitations as Wilson’s method. Matsuda et al. [9] used the Wilson’s model to predict the solubility of Salicylic acid, Benzocaine, Acetanilide, Phenacetin in water mixed with different co-solvents. They referred to several literatures in order to get the parameters for Wilson’s equation. For pure parameters they used the Tassios method [10] followed by using DECHEMA VLE collection [11]. In addition, they had to consider some assumptions which led to some errors in prediction (i.e., they used the simplified method to find the interaction parameters, or the molar volumes were estimated using the group contribution method) which led to some errors in prediction.

From the second category, the universal functional activity coefficient (UNIFAC) model is one of the well-developed ones. This model is basically used for prediction of the phase behaviour of non-electrolytic and non-ideal systems. This model was first introduced by Fredenslund et al. in 1975 [12]. They used the UNIFAC model to predict the solubility of Naphthalene, Anthracene and Phenanthrene in various solvents and their mixtures. They showed the applicability of the UNIFAC model in prediction of the phase behaviour of solutes in solvents. After that some modifications have been made on the original form of the UNIFAC model in order to make the prediction better [13]. Sheikhzadeh et al. [14] used the UNIFAC and its modified form in predicting the solubility of the two forms of Buspirone-hydrochloride in isopropanol and its mixture with water. They concluded that the modified UNIFAC forms are not suitable for prediction of highly soluble drugs. Gracin and coworkers [15] studied the solubility of a wide variety of chemical compounds in water and some organic solvents. Although they obtained some reasonable predictions for a few chemicals, but the model was not proper for solvent screening purposes. They used some known group parameters in UNIFAC method for complex and unknown groups. In addition, because of the dependency of the property of some functional
groups on the rest of the molecule, they concluded that the UNIFAC was not a suitable method for crystallization process design and solvent screening. Kan et al. [16] used the UNIFAC method for solubility prediction of some compounds from alkanes, alkyl benzenes, alkenes, and benzenes. The solubility of 11 out of 14 compounds was best predicted by UNIFAC model in their study. The chlorinated alkenes, phthalates, and long-chain alkanes were not predicted well.

The NRTL-SAC model was first introduced by Chen et al. [17] in 2004. This model was proposed in order to compensate the weakness of the UNIFAC in predicting the solubility of complex chemical molecules with functional groups that had not been studied for the UNIFAC parameters. Also, in some cases the UNIFAC group addition rule becomes invalid [18]. One of the main advantages of NRTL-SAC model in comparison to the other predictive methods is its ability to predict organic electrolyte systems [17]. The UNIFAC method identifies the molecule in terms of its functional groups, while the NRTL-SAC model divides the whole surface of the molecule to four segments. Each compound can have three conceptual segments [17], hydrophobic, hydrophilic, and polar with four segment numbers. Hydrophobic segment (X) accounts for the molecular surfaces that do not tend to form a hydrogen bond, such as hexane. The polar segment (Y- and Y+) does not belong to the hydrophobic nor hydrophilic segment. The polar attractive segment (Y-) shows attractive interaction with hydrophilic segment, while the polar repulsive segment (Y+) has repulsive characteristic with hydrophilic segment. Hydrophilic segment (Z) contributes to the part of the molecule which tends to form a hydrogen bond, such as water.

The four segments are identified in terms of the interactions between the molecules in the solution which is expressed in the phase equilibrium experimental data. Chen et al. [17] selected water as a reference for hydrophilic segment, acetonitrile as a reference for polar segment, and hexane as a reference for hydrophobic segment.

In order to predict the solubility of a chemical in a solvent or a mixture of solvents, one should know the segment numbers for solute and solvent(s). Chen et al. has studied extensive data of liquid-liquid equilibrium (LLE) and vapor-liquid equilibrium (VLE) for many solvents. Based on the reference solvents (water, acetonitrile, and hexane), they found the optimized values of segment numbers for other common solvents. Many solvents have only one or two segment numbers that are significant compared to the other segment numbers. Up to now, there are 62
common solvents which their segment numbers have been adjusted and tabulated for calculation of the phase behaviour [17].

2.2. Thermodynamic theory and modeling

For a solid in equilibrium with itself in a solution, there is a thermodynamic relation [18]:

\[ \hat{f}_i^S(T, P) = \hat{f}_i^L(T, P, x_i) \]  \hspace{1cm} (2 - 3)

Where \( \hat{f}_i^S(T, P) \) denotes fugacity of component \( i \) in solid phase and \( \hat{f}_i^L(T, P, x_i) \) represents the fugacity in the liquid phase. The procedure to get the equation (2-5) from fundamental relations of thermodynamics is explained.

With plugging in the proper definitions in equation (2-3), we get:

\[ f_i^S = x_i \gamma_i f_i^L \]  \hspace{1cm} (2 - 4)

Where \( f_i^S \) and \( f_i^L \) are the fugacity of pure component \( i \) in solid and liquid states, respectively. In order to find the ratio of the two fugacities, we need to include all the thermodynamic processes that are needed to start from solid and reach to the liquid state. The three processes are shown in Figure (2-14). Path A in this Figure shows the changing from process conditions to the state where the solid starts melting. Path B shows the melting process at constant temperature and pressure. Path C indicates the change of conditions from melting to the process state. The sum of the three paths will give us the whole change from solid to liquid state of a pure component.

From the fundamental rules in thermodynamics, we have:

\[ \Delta G_{S\rightarrow L} = RT \ln \frac{f_i^L}{f_i^S} \]  \hspace{1cm} (2 - 5)

And the Gibbs energy change is related to change of enthalpy and entropy:

\[ \Delta G_{S\rightarrow L} = \Delta H_{S\rightarrow L} - T \Delta S_{S\rightarrow L} \]  \hspace{1cm} (2 - 6)
From Figure (2-1), the whole change in enthalpy is found from:

\[
\Delta H_{S \rightarrow L} = \Delta H_A + \Delta H_B + \Delta H_C = \int_T^{T_f} C_{p, \text{solid}}dT + \Delta H_{\text{fusion}} + \int_T^{T_f} C_{p, \text{liquid}}dT
\]

\[
= \int_{T_f}^{T} \Delta C_p dT + \Delta H_{\text{fusion}} \quad \text{(2 - 7)}
\]

In which \( \Delta C_p = C_{p, \text{liquid}} - C_{p, \text{solid}} \). In the same manner, the entropy change from solid to liquid state can be found from:

\[
\Delta S_{S \rightarrow L} = \int_{T_f}^{T} \frac{\Delta C_p}{T} dT + \Delta S_{\text{fusion}} \quad \text{(2 - 8)}
\]

Also, from thermodynamic rules, \( \Delta S_{\text{fusion}} = \frac{\Delta H_{\text{fusion}}}{T_{\text{fusion}}} \). If we substitute equations (2-7 and 2-8) into equation (2-6), then:

\[
RT \ln \frac{f_L}{f_S} = \int_{T_f}^{T} \Delta C_p dT + \Delta H_{\text{fusion}} + \int_{T_f}^{T} \frac{\Delta C_p}{T} dT + \frac{\Delta H_{\text{fusion}}}{T_{\text{fusion}}} \quad \text{(2 - 9)}
\]
If we neglect the terms including change in the heat capacity (because of the large value of heat of fusion compared to the heat capacities), then we will get to the equation (2-5).

After some mathematical calculations we get to the following equation:

\[
\ln \frac{f_i^L}{f_i^s} = -\frac{\Delta H_{\text{fus}}}{R} \left( \frac{1}{T_t} - \frac{1}{T} \right) + \frac{\Delta C_p}{R} \left[ \frac{T_t}{T} - 1 - \ln \left( \frac{T_t}{T} \right) \right]
\]  
(2 - 10)

Where \( f_i^L \) and \( f_i^s \) are the pure fugacities of the liquid and solid state of the component \( i \) and \( \Delta C_p \) is the heat capacity change during the phase transition from solid to liquid. With applying the two assumptions the equation (2-4) can be further simplified: (1) the second term in the right-side of equation (2-4) can be neglected, because of the magnitude of \( \Delta C_p \) compared to \( \Delta H_{\text{fus}} \), and (2) The triple-point temperature, \( T_t \), can be substituted by melting-point temperature, \( T_m \), at regular pressures. After considering the two assumptions, the equation (2-4) will be simplified:

\[
\ln x_s = -\frac{\Delta H_{\text{fus}}}{R} \left( \frac{1}{T_m} - \frac{1}{T} \right) - \ln \gamma_s
\]  
(2 - 11)

Where the index \( s \) refers to the solid phase, \( \Delta H_{\text{fus}} \) is the heat of fusion, and \( T_m \) is the melting point. Equation (2-11) is used in prediction of the solubility of a solid in a solvent. For an ideal solution, \( \gamma_s = 1 \) and therefore, the Van’t-Hoff equation will be derived. To predict the solubility of a real solution, we need the physical properties of the solid (\( \Delta H_{\text{fus}} \) and \( T_m \) that can be obtained using thermal methods, such as DSC and TGA) and a proper model describing \( \gamma_s \).

2.2.1. Universal Functional Activity Coefficient model (UNIFAC)

The group-contribution models divide the contribution of the activity coefficient to two parts:

- Combinatorial part. This part includes the contribution of the chemical structure and the size (volume and surface of the molecule) of the compound.
- Residual part. This part includes the contribution of the group size and binary interaction between pairs of the functional groups.

With the above definition, the total activity of a component in the solution is the sum of the two parts:
\[ \ln \gamma_i = \ln \gamma_i^C + \ln \gamma_i^R \quad (2 - 12) \]

In which \( \gamma_i \) is the activity coefficient of component \( i \) in the solution, \( \gamma_i^C \) is the combinatorial part and \( \gamma_i^R \) is the residual part. Up to this point all of the group contribution and activity coefficient methods (i.e. NRTL-SAC) are the same, but the method in which the activities are calculated is different. In the UNIFAC model the combinatorial part for the component \( i \) is found from the following equation [12]:

\[ \ln \gamma_i^c = \ln \left[ \frac{\phi_i}{x_i} \right] + \frac{z}{2} q_i \ln \left[ \frac{\theta_i}{\phi_i} \right] + L_i - \frac{\phi_i}{x_i} \sum_{j=1}^{n} x_j L_j \quad (2 - 13) \]

Where:

\[ L_i = \frac{z}{2} (r_i - q_i) - (r_i - 1) \quad (2 - 14) \]

\( z \) is the coordination number and is taken to be 10. In equation (2-13), \( \phi_i \) is the segment fraction and \( \theta_i \) is the area fraction of component \( i \) and is related to the mole fraction of the species \( i \) in the mixture:

\[ \phi_i = \frac{r_i x_i}{\sum_{j=1}^{n} r_j x_j} \quad (2 - 15) \]

\[ \theta_i = \frac{q_i x_i}{\sum_{j=1}^{n} q_j x_j} \quad (2 - 16) \]

\( q_i \) and \( r_i \) are the pure component surface area and volumes (Van der Waals), respectively. These parameters are not temperature dependent and are only functions of chemical structure of a functional group. In the UNIFAC model for every functional group there is a unique value for surface area and volume that can be found in common texts and handbooks [18]. The first step in modeling the UNIFAC for a specific binary or ternary system is to break down the chemical structure of a molecule into the basic functional groups. As it is suggested in thermodynamic textbooks [19], the optimum way of breaking down is the one which results in the minimum number of sub-groups with each sub-group having the maximum replicates. The \( q_i \) and \( r_i \) can be found from equations (2-17 and 2-18):
\[ r_i = \sum_k v_k^i R_k \quad (2 - 17) \]

\[ q_i = \sum_k v_k^i Q_k \quad (2 - 18) \]

\( v_k^i \) is the number of sub-group \( k \) in component \( i \). The residual part of the UNIFAC is found from the below equation:

\[ \ln \gamma_i^R = \sum_{k=1}^{n_k} v_k^i \left[ \ln \Gamma_k - \ln \Gamma_k^1 \right] \quad (2 - 19) \]

In equation (2-19) \( \Gamma_k \) is the residual activity coefficient of sub-group \( k \) in the mixture and \( \Gamma_k^1 \) is that value in a pure solution of the component \( i \). This term is added so when the mole fraction approaches unity, the term \( \ln \gamma_i^R \) tends to zero (\( \gamma_i^R \to 1 \)). The residual activity coefficient of sub-group \( k \) in a solution is given by:

\[ \ln \Gamma_k = Q_k \left[ 1 - \ln \sum_m \theta_m \psi_{mk} - \sum_m \left[ \frac{\theta_m \psi_{mk}}{\sum_n \theta_n \psi_{nm}} \right] \right] \quad (2 - 20) \]

Equation (2-19) is also applicable to the case of \( \Gamma_k^1 \), in which the parameters of the right-hand side of the equation are written based on the pure component \( i \). \( \theta_m \) is the area fraction of the functional group \( m \) in the mixture:

\[ \theta_m = \frac{Q_m X_m}{\sum_n Q_n X_n} \quad (2 - 21) \]

\( X_m \) is the mole fraction of sub-group \( m \) in the mixture. \( \psi_{nm} \) is the group interaction parameter between groups \( n \) and \( m \) and is dependent on the temperature:

\[ \psi_{nm} = \exp \left[ - \frac{u_{nm} - u_{mm}}{RT} \right] = \exp \left[ - \frac{a_{nm}}{T} \right] \quad (2 - 22) \]

The group interaction parameter \( a_{nm} \) is found from the large sets of VLE and LLE data in the literature which are tabulated for many sub-groups. It is worth noting that \( a_{nm} \neq a_{mn} \). There are
some modifications to the original UNIFAC equation in order to make the model robust for some complex systems. In the UNIFAC-DM method, the modification is made on the combinatorial part:

\[
\ln \gamma_i^c = \ln \left[ \frac{\Phi_i}{x_i} \right] + \frac{Z}{2} q_i \ln \left[ \frac{\theta_i}{\Phi_i} \right] + L_i - \frac{\Phi_i}{x_i} \sum_{j=1}^{n} x_j L_j
\]  

(2 - 23)

In which the term \( \frac{\Phi_i}{x_i} \) is defined as:

\[
\frac{\Phi_i}{x_i} = \frac{r^{3/4}}{\sum_j x_j r_j^{3/4}}
\]  

(2 - 24)

The algorithm for finding the solute solubility in the mixture of ternary solution (solvent/co-solvent/solute) is shown in Figure 2-2. The algorithm starts with known values, such as the physical properties of the solute. After making an initial guess for the solubility, the program obtains the activity coefficients and the new solubility is found and is compared with the old value and the calculations are repeated to converge to a unique value for solubility. This procedure is done for all of the experimental data points.
Figure 2-2. The algorithm of converging to the solubility of a ternary system using UNIFAC model

2.2.2. Non-random Two-liquid Segment Activity Coefficient (NRTL-SAC)

According to Chen et al. [17], the NRTL-SAC model is based on the derivation of the original NRTL model for polymers. From equation (2-6) the activity coefficient is made up of two terms, combinatorial and residual. Like the UNIFAC model, the activity coefficients must be generated in order to obtain the solubility. In the NRTL-SAC model, the combinatorial part is calculated by equation (2-19):
\[ \ln \gamma_i^c = \ln \frac{\phi_i}{x_i} + 1 - r_i \sum_j \frac{\phi_j}{x_j} \]  \hspace{1cm} (2 - 25)

With the definitions:

\[ r_i = \sum_j r_{j,i} \]  \hspace{1cm} (2 - 26)

\[ \phi_i = \frac{r_i x_i}{\sum_j r_j x_j} \]  \hspace{1cm} (2 - 27)

Where \( x_i \) is the mole fraction of component \( i \), \( r_{m,i} \) is the number of segment \( m \), \( r_i \) is the total segment number in component \( i \) (\( r_1 \) refers to the value of \( X \), \( r_2 \) refers to \( Y^- \), \( r_3 \) refers to \( Y^+ \), and \( r_4 \) refers to \( Z \)), and \( \phi_i \) is the segment mole fraction in the mixture. The residual term is defined as:

\[ \ln \gamma_i^R = \ln \gamma_i^lc = \sum_m r_{m,i} (\ln \Gamma_m^lc - \ln \Gamma_m^lc,i) \]  \hspace{1cm} (2 - 28)

In equation (2-28) there are two terms, \( \ln \Gamma_m^lc \) and \( \ln \Gamma_m^lc,i \) which are the activity coefficients of segment \( m \) in solution and component \( i \), respectively.

The two mentioned terms are found using the equations (2-29 and 2-30):

\[ \ln \Gamma_m^lc = \frac{\sum_j x_j G_{j,m} \tau_{j,m}}{\sum_k x_k G_{k,m}} + \sum_m \frac{x_m G_{m,m'}}{\sum_k x_k G_{k,m'}} \left( \tau_{m,m'} - \frac{\sum_j x_j G_{j,m} \tau_{j,m'}}{\sum_k x_k G_{k,m'}} \right) \]  \hspace{1cm} (2 - 29)

\[ \ln \Gamma_m^lc,i = \frac{\sum_j x_j G_{j,m} \tau_{j,m}}{\sum_k x_k G_{k,m}} + \sum_{m'} \frac{x_{m',i} G_{m,m'}}{\sum_k x_k G_{k,m'}} \left( \tau_{m,m'} - \frac{\sum_j x_j G_{j,m} \tau_{j,m'}}{\sum_k x_k G_{k,m'}} \right) \]  \hspace{1cm} (2 - 30)

In the two above equations \( l \) refers to component and \( j, k, m \), and \( m' \) refer to the segments in each component. \( x_{j,i} \) is the segment-based mole fraction of segment species \( j \) in component \( l \) only. The mole fractions of segments in the whole solution and in components are defined as below:

\[ x_j = \frac{\sum_l x_{j,l} r_{j,l}}{\sum_z \sum_l x_{z,l} r_{j,z}} \]  \hspace{1cm} (2 - 31)

\[ x_{j,l} = \frac{r_{j,l}}{\sum_l r_{j,l}} \]  \hspace{1cm} (2 - 32)
$G_{ij}$ and $\tau_{ij}$ are the local binary values which can be related to each other based on NRTL non-random parameter $\alpha_{ij}$, and are shown by their values in Table 2-1. $G_{ij}$ and $\tau_{ij}$ have the following relation:

$$G_{ij} = e^{-\alpha_{ij} \tau_{ij}} \quad (2-33)$$

Therefore, from fixed values of $\tau_{ij}$ and $\alpha_{ij}$ one can find $G_{ij}$. The segment numbers for the common solvents can be found from the literature [20]. After putting the values of segments for solvents and initial guess values for the solute segments, the written code for NRTL-SAC starts solving for the mole fractions at saturation for all of the species in the solution (see Figure 2-3).

Figure 2-3. Algorithm flowchart for parameter estimation using NRTL-SAC model
The only thing here which is different from the UNIFAC model is the parameter estimation loop. This part of the model (which has a separate Matlab code) uses a few experimental data in order to fit the model output to the experimental data. After optimizing the process and finding the four segment numbers for the solute (which were unknown initially), the values can be set for that compound and be used for further predictions. For the NRTL-SAC models, we used the \textit{lsqnonlin} routine of the Maltab software which is suitable for solving the nonlinear least-squares problems. The algorithm for solving the least-squares is based on interior-reflective Newton method [21]. After finding the adjustable parameters for the selected solutes, we used the model equation to predict the solubility of the compound in pure/mixed solvents. For UNIFAC model there is no need to find the adjustable parameters. The average relative deviation (ARD) for the whole data is calculated for each case in order to find which model best describes the system:

$$\text{ARD} = \frac{1}{N} \sum_{i} \left| \frac{x_{i}^{\text{model}} - x_{i}^{\exp}}{x_{i}^{\exp}} \right|$$

(2 - 34)

Where $N$ is the total number of data points. The dependency of the binary parameters in the NRTL-SAC model for the segments to temperature is ignored. The values of the binary interactions for the segments are shown in Table 2-1.

<table>
<thead>
<tr>
<th>Table 2-1. NRTL binary interaction constants for conceptual segments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Segment (1)</td>
</tr>
<tr>
<td>Segment (2)</td>
</tr>
<tr>
<td>$\tau_{12}$</td>
</tr>
<tr>
<td>$\tau_{21}$</td>
</tr>
<tr>
<td>$\alpha_{12} = \alpha_{21}$</td>
</tr>
</tbody>
</table>

For binary parameters between the two segments, Chen et al. [17] did the following:

1. For binary interaction between hydrophobic (X) and hydrophilic (Z), they used the experimental data of LLE for hexane-water system. The values of both parameters are high, stating the strong repulsive nature of the two segments.
2. The values for binary segment interaction between hydrophobic (X) and polar (Y) segments were determined from LLE data of the hexane-acetonitrile mixture.
3. The binary interaction between polar (Y) and hydrophilic (Z) segments were found from the vapour-liquid equilibrium (VLE) data available in the literature for acetonitrile-water system.
4. The value of $\alpha$ for the liquid-liquid mixtures was set to 0.2 and for vapour-liquid mixtures was set to 0.3.
5. The interaction between polar attractive (Y-) and polar repulsive (Y+) segments in the mixture was assumed to be ideal and the binary interactions were set to zero.

2.3. Materials and methods

2.3.1. Model compounds

The model compounds for our study are: 3-pentadecylphenol [22], lovastatin [23, 24] and valsartan [25, 26]. 3-Pentadecylphenol with the chemical formula of $C_{21}H_{36}O$, and molecular weight of 304.51 is the product of catalytic hydrogenation of cardanol. Its main usage is in agriculture industry as an emulsifier and coating material. The chemical structure of this compound is shown in Figure 2-4. It has a phenolic head and a linear alkyne group. Because of having a long hydrocarbon chain, and near straight shape of the molecule, one can consider the molecule as if it does not have a benzene ring. This assumption can be reasonable because the benzene ring has a non-polar characteristic and has a minor effect on non-ideality of the solution. The dimensions are shown in Figure 2-4. The sizes of bonds are calculated form ChemDoodle software, version 3.3.1. From this Figure, the length of the hydrocarbon chain is about 29.53 Å and that of the benzene-ring is 4.73 Å. This means that about 16% of the volume of the molecule is occupied by the benzene-ring, which is non-polar and the other part of molecule is mostly the alkyl chain.

![Figure 2-4. The chemical structure of 3-pentadecylphenol in compact form (right) and the length of bonds in Angstrom (Å) (left)](image)

The solubility of this compound was studied in six pure solvents, ethanol, 1-butanol, toluene, acetone, tetrachloromethane, and ethyl acetate [22]. The authors used a laser detector in order to
find the point of saturation. In the study by Mao et al. [22] they used the Wilson’s equation to get the adjustable parameters for every solute-solvent system. They did not do any prediction of the solubility, but only fitted the experimental data to the Wilson’s model and derived the values of binary interactions for the system. We also conducted solubility measurements of this compound in pure and mixed solvents. The material and the experimental procedure will be discussed in the next section. In our study we used the UNIFAC and the NRTL-SAC model to predict the solubility of the compound and compared it with the experimental data.

Lovastatin is a drug that belongs to a group of compounds that lower the lipids of the body, called statins [23]. There are other statins approved in many countries, like simvastatin and fluvastatin. The chemical name of this drug is (butanoic acid 2-methyl-1,2,3,7,8,8a-hexahydro-3,7-dimethyl-8-[2-(tetrahydro-4-hydroxy-6-oxo-2H-pyran-2-yl) ethyl]-1-naphalenyl ester) with the chemical formula of C_{24}H_{36}O_{5} and molecular weight of 404.54. The chemical structure of this drug is represented in Figure 2-5. There are two main studies that have been reported on the solubility of this compound [23, 24].

![Chemical structure of lovastatin](image)

**Figure 2-5.** Chemical structure of lovastatin

In the study by Gyabaah et al. [24] the solubility of lovastatin in a group of acetate- solvents was measured. The solvents were ethyl acetate, propyl acetate, isopropyl acetate, butyl acetate, sec-buty acetate, isobutyl acetate, tert-buty acetate, and 2-butanone. All of the solubilities were measured in the range of 285-313K. In their study they used the NRTL model to find the binary interaction parameters for solute-solvent mixtures. In another study the solubility of lovastatin in six alcohols was measured [23]. It was found that in alcohol groups the solubility of lovastatin
increases from ethanol to 1-butanol and then decreases as the carbon chain length increases. This can be rationalized by the solute-solvent interactions.

Valsartan with the chemical formula of C_{24}H_{29}N_{5}O_{3} and chemical name (S)-N-(1-carboxy-2-methyl-1-yl)-N-pentnoyl-N-[2’-(1H-tetrazol-5-yl) biphenyl-4-yl methyl]-amine, is used orally for the treatment of hypertension [37]. The chemical structure of valsartan is illustrated in Figure 2-6.

![Chemical structure of valsartan](image)

**Figure 2-6.** Chemical structure of valsartan

This drug is extracted from the product mixture by ethyl acetate at the end of the production process, because of its high solubility [25]. However, some problems such as poor yield, slow filtration, and long-time drying are the results of recrystallization with ethyl acetate. For solving this problem hexane has been used as an antisolvent in order to overcome those problems. Liu et al. [26] found the solubility of the mixture of ethyl acetate + hexane at different molar ratios over the temperature range of 278.15 to 313.15K. The synthetic method was used to obtain the solubility of valsartan in the solvent mixture. In another study the solubility of valsartan in six different solvents: methyl acetate, n-butyl acetate, acetonitrile, N,N-dimethylformamide, dichloromethane, and chloroform over the temperature range of 278.15 to 313.15K, was measured [25]. The melting point temperature, heat of fusion, and entropy change of melting of the three mentioned model compounds are given in Table 2-2.
### Table 2-2. Physical properties of the model compounds used for thermodynamic modeling and prediction

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>3-pentadecylphenol, [22]</td>
<td>322.35</td>
<td>38092</td>
<td>118.14</td>
</tr>
<tr>
<td>Lovastatin, [23, 24]</td>
<td>445.50</td>
<td>43136</td>
<td>96.86</td>
</tr>
<tr>
<td>Valsartan, [25, 26]</td>
<td>380.65</td>
<td>31647</td>
<td>84.97</td>
</tr>
</tbody>
</table>

#### 2.3.2. Experimental setup and procedure

The solubility of 3-pentadecylphenol in isopropyl alcohol and the mixture of isopropyl alcohol and water were measured. 3-Pentadecylphenol (90% purity) was purchased from Sigma Aldrich. The compound was recrystallized for purification with ethanol [22]. A specified amount of 3-pentadecylphenol was dissolved in ethanol in 50°C. After reaching to the desired temperature the solution was filtered by filter paper (VWR, Grade 410, qualitative). The clear solution was again maintained at that temperature for 1 hour. Then the solution was cooled with cooling rate of 0.2°C/min until reaching to 20°C. The solution was kept at that temperature for 1 hour. The precipitated solids were filtered and dried. In order to identify its purity we used the differential scanning calorimetry (DSC) analysis (DSC, Mettler Toledo, Chicago, IL). The sample was put in a 40µl aluminum crucible with a hole in the cap to allow venting. The heating rate was 0.5°C/min in the temperature range of 20°C to 120°C. The DSC curve and the obtained melting temperature were compared to those of the purified compound. The results showed good purity of the sample and effective experimental procedure.

After getting the purified compound, it was used for determining the solubility in IPA (isopropyl alcohol) solvent. The gravitational method was used to obtain the solubility data. At first, a known amount of 3-pentadecylphenol was dissolved in 50ml of IPA solvent at 40°C. The
solution was prepared in 250ml Bellco jacketed vessel (Vinelean, NJ). A bath circulator (Julabo, Germany) was used for heating and cooling. A Teflon-coated thermocouple was used for reading the temperature in the flask. For mixing a top-mounted electromagnetically stirrer was employed. After complete dissolution of all of the particles in the solvent, the temperature was cooled down to a specified temperature to get the data of solubility. The solution was kept at constant temperature for 1 hour. A membrane disk filter (VWR, Mississauga, ON) with 0.45µm was used in order to filter the impurities from the sample solution. The clear saturated solution was kept in closed weighted vials. The vial with solution was weighed and transferred to the oven (at 60°C) for 1 day. The weight of solute and solvent were recorded for every sample. The mentioned experimental procedure was repeated three times for each temperature to increase the reliability. The replicates at every temperature were averaged and the standard deviation was calculated for data analysis. In addition to the single solvent experimentation, we used solvent mixtures. Water was selected as an antisolvent because 3-pentadecylphenol is sparingly soluble in water, but fairly soluble in IPA. The solubility of 3-pentadecylphenol was measured at different solvent volume fractions.

2.3.2.1. Characterization methods

2.3.2.1.1. X-ray powder diffraction (XRPD)

X-ray powder diffraction was conducted by XRPD equipment (Rigaku, Miniflex) with CuKα radiation. The scan angle was between 5°-40° with the step angle of 0.05°.

2.3.2.1.2. Differential scanning calorimetry (DSC)

Thermal analysis was conducted by differential scanning calorimetry (DSC, Mettler Toledo, Chicago, IL). The sample of 3-5mg was prepared in a covered 40µl aluminum crucible with a hole in the lid to allow venting. The heating rate of 1°C/min was employed and the temperature range was set from 20°C to 120°C. The N₂ flow was used on the crucible with a rate of 100ml/min. The calibration was performed using indium.
2.4. Results and discussion

2.4.1. Solubility prediction of 3-pentadecylphenol.

In order to model the solubility of 3-pentadecylphenol using the UNIFAC model, the chemical structure of the solute was broken down to functional groups. For 3-pentadecylphenol, the alkyl chain length is much longer than the benzene-ring. The prediction was based on the two functional group arrangements: (1) structure with benzene ring and (2) structure without benzene ring. We found some functional groups that each have one or more replicates in the structure. The functional groups for two conditions are shown in Table 2-3. The results of characterization using XRPD and DSC instruments are shown in Figure 2-7.

**Table 2-3.** Functional groups that are defined in UNIFAC and their replication in two different structures of 3-pentadecylphenol

<table>
<thead>
<tr>
<th>Sub-group</th>
<th>Condition I (with benzene ring)</th>
<th>Condition II (without benzene ring)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACOH</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>ACCH₂</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>ACCH</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>CH₃</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>OH</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>CH₂</td>
<td>13</td>
<td>14</td>
</tr>
</tbody>
</table>

The solubility of 3-pentadecylphenol in the solvents was calculated using the algorithm which is shown in Figure 2-2. The results of the prediction using UNIFAC are shown in Figures 2-8 and 2-8. For the NRTL-SAC model the estimation using two solvents are shown in Figure 2-8 and the prediction is shown for other solvents in Figures (2-9 and 2-10). Except for butanol and ethanol, for which the two conditions (with and without benzene ring) have nearly the same prediction, for other four solvents, there is a big difference between the solubility predictions for the two conditions. For acetone, with the inclusion of the benzene ring, the prediction is very well. For ethyl acetate, tetrachloromethane and to some extent toluene, the assumption of no benzene ring shows better prediction. The reason for the different behaviour in the two proposed
structure, is mainly due to the electron-rich nature of the benzene ring. When the benzene ring is present, the electron cloud on the ring helps to interact better with some polar solvents, such as acetone. On the other hand, when we assume no benzene ring in the structure, the molecule will be an alcohol with long alkyl chain, which has OH group at the terminal of the structure and makes it easier to interact with non-polar or less polar solvents, such as tetrachloromethane.

![Figure 2-7. Characterization of re-crystallized 3-pentadecylphenol with DSC (left) and XRPD (right)](image)

For the NRTL-SAC model prediction, we need to find the four segment numbers for the chemical compound, and then use the segment values to predict the solubility of the compound in other solvents. We selected 1-butanol and acetone for parameter estimation and optimized the values of segment numbers which are shown in Table 2-4.

<table>
<thead>
<tr>
<th>Component</th>
<th>X (Hydrophobic)</th>
<th>Y- (Polar attractive)</th>
<th>Y+ (Polar repulsive)</th>
<th>Z (Hydrophilic)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-pentadecylphenol</td>
<td>0.674</td>
<td>0.000</td>
<td>0.571</td>
<td>0.398</td>
</tr>
<tr>
<td>Lovastatin</td>
<td>1.175</td>
<td>0.000</td>
<td>0.548</td>
<td>0.882</td>
</tr>
<tr>
<td>Valsartan</td>
<td>0.000</td>
<td>0.946</td>
<td>0.000</td>
<td>0.539</td>
</tr>
</tbody>
</table>
Using the obtained segment numbers for the compound, the NRTL-SAC model predicted solubility in other solvents (Figures 2-9 and 2-10). It can be seen from the results that except for toluene, the solubility in other three solvents were predicted satisfactorily. For all of the four solvents, the NRTL-SAC model resulted in better prediction compared to the UNIFAC. As it is evident from Table 2-5 the polar attractive segment is zero, which implies the molecule has a repulsive nature against polar solvent, due to the electron-rich part of benzene ring. As it was mentioned previously, about 84% of the total volume of the molecule is occupied by the alkyl group, and the remaining is the phenolic part. This makes the molecule to have a good interaction with other polar solvents. For non-polar solvents (such as toluene) the interaction is not the same. That is why the prediction for toluene has some deviation from the experimental results.

**Table 2-5.** Experimental solubility of 3-pentadecylphenol in pure IPA and its mixtures with water (the mean values with their standard deviation)

<table>
<thead>
<tr>
<th>Temperature [°C]</th>
<th>15</th>
<th>20</th>
<th>25</th>
<th>30</th>
<th>35</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pure IPA</td>
<td>0.7072 ± 0.0631</td>
<td>0.9436 ± 0.0804</td>
<td>1.4109 ± 0.1233</td>
<td>2.2355 ± 0.1011</td>
<td>3.8502 ± 0.4150</td>
</tr>
<tr>
<td>80 vol.% IPA</td>
<td>0.7635 ± 0.0659</td>
<td>0.9755 ± 0.1213</td>
<td>1.6026 ± 0.1519</td>
<td>2.3653 ± 0.2608</td>
<td>4.4036 ± 0.2307</td>
</tr>
<tr>
<td>50 vol.% IPA</td>
<td>0.2302 ± 0.0286</td>
<td>0.3980 ± 0.0419</td>
<td>0.6704 ± 0.0681</td>
<td>1.2739 ± 0.1396</td>
<td>3.1857 ± 0.2867</td>
</tr>
<tr>
<td>20 vol.% IPA</td>
<td>0.0152 ± 0.0040</td>
<td>0.0188 ± 0.0031</td>
<td>0.0235 ± 0.0045</td>
<td>0.0273 ± 0.0028</td>
<td>0.0329 ± 0.0037</td>
</tr>
</tbody>
</table>

For a system of IPA-water we see that with the addition of water to the pure IPA, the solubility increases to a certain value at constant temperature. With more addition of water to the mixture the solubility decreases and approaches zero as the volume fraction of water goes to unity in the mixture. The maximum value of solubility is predicted by the NRTL-SAC (Figure 2-10).
Figure 2-8. Experimental points of the solubility of 3-pentadecylphenol in two solvents and their curve of estimation using NRTL-SAC model. The curves of UNFAC are also shown here for comparison.
Figure 2-9. Experimental and predicted solubilities for 3-pentadecylphenol in solvents using UNIFAC and NRTL-SAC models
2.4.2. Solubility prediction of lovastatin. From the chemical structure of lovastatin, the functional groups in the UNIFAC model are listed in Table 2-6. The experimental data were obtained from literature [23]. The functional groups are shown schematically in Figure 2-11. Using the obtained group functions and their replicates in the chemical structure, we modeled the thermodynamic equilibrium of lovastatin in alcohol and acetate group solvents.

**Table 2-6.** Functional groups that have been used in modeling with the UNIFAC model

<table>
<thead>
<tr>
<th>Sub-group</th>
<th>Number of replicates</th>
</tr>
</thead>
<tbody>
<tr>
<td>CH₃</td>
<td>4</td>
</tr>
<tr>
<td>CH₂</td>
<td>6</td>
</tr>
<tr>
<td>CH</td>
<td>6</td>
</tr>
<tr>
<td>C=C</td>
<td>2</td>
</tr>
<tr>
<td>CH-O</td>
<td>1</td>
</tr>
<tr>
<td>C-O-C</td>
<td>2</td>
</tr>
<tr>
<td>COCH₃</td>
<td>2</td>
</tr>
</tbody>
</table>
From the acetate group we selected methyl acetate and ethyl acetate for estimation of the NRTL-SAC segment numbers. We made three ways of modeling the solubility of lovastatin in solvents:

- Parameter estimation with two solvents from alcohol group only and solubility prediction for all the remaining solvents
- Parameter estimation with two solvents from acetate group only and prediction of solubility for the remaining solvents
- Parameter estimation with two solvents from alcohol and two solvents from acetate group and solubility prediction for other remaining solvents.

Also we selected ethanol and 1-pentanol from alcohol group for alcohol estimation. After finding the segment numbers from a few set of solvents, we predicted the solubility of lovastatin in other remaining solvents. Finally, all alcohol and acetate solvents were put in one group and again the estimation was repeated with the same compounds as before. After optimizing the adjustable parameters using four solvents, the solubility for the remaining eight solvents were predicted. The summary of the methods are shown in Table 2-7. For the first estimation (method I) we selected two alcohols from the set of 5 alcohols in ref. [23]. Our selection was based on the alkyl chain in order to consider the effect of its length on the parameter estimation. Therefore, we used ethanol and 1-pentanol for parameter estimation. With the adjusted value for segment numbers that were generated for alcohol group, we then predicted the solubility of lovastatin in remaining three alcohols.

![Figure 2-11. Schematic view of functional groups presented in lovastatin](image-url)
In Figure 2-12 the experimental data for four selected solvents with the predicted curves for the NRTL-SAC and UNIFAC are shown. From those four curves, one is for the UNIFAC prediction method, and the rest are for the NRTL-SAC model. The curve of “NRTL-SAC” results in the best prediction of the data. Other methods are also shown for comparison. The next possible way of estimation (method II) can be on the acetate group. In this group there are six acetate solvents plus acetone as a basic solvent [24]. Two solvents were selected from the group (methyl- and ethyl- acetate). By having their experimental values for these solvents the error between the experimental and estimated values were minimized. The segment numbers that are generated here were used to predict the solubility of lovastatin in the remaining 5 solvents. The final part for estimation (method III) deals with four solvents from both the alcohols and acetate groups. The solvents that were selected are the ones mentioned in method I and II. In this case, the estimation is made by minimizing the error between the experimental and estimated solubility of 48 data point. For all of the three cases the ARD (average relative deviation) was calculated in order to compare the three methods, qualitatively. The results and summary of the ARD are shown in Table 2-7. The values of optimized segment numbers for each method are shown in Table 2-8.
Figure 2-12. Experimental data with the predictions of solubility in four selected solvents using three methods of the NRTL-SAC and the UNIFAC method, experimental data from ref. [24]
Table 2-7. Average relative deviation for the UNIFAC and NRTL-SAC in three methods

<table>
<thead>
<tr>
<th>Method</th>
<th>%ARD based on alcohol solvents</th>
<th>%ARD based on acetate solvents</th>
<th>%ARD considering both alcohol and acetate solvents</th>
</tr>
</thead>
<tbody>
<tr>
<td>NRTL-SAC (method III)</td>
<td>0.1946</td>
<td>0.1658</td>
<td>0.2401</td>
</tr>
<tr>
<td>NRTL-SAC (method II)</td>
<td>0.4980</td>
<td>0.1305</td>
<td>0.3306</td>
</tr>
<tr>
<td>NRTL-SAC (method I)</td>
<td>0.4735</td>
<td>0.4841</td>
<td>0.3499</td>
</tr>
<tr>
<td>UNIFAC</td>
<td>0.4795</td>
<td>0.7064</td>
<td>0.6153</td>
</tr>
</tbody>
</table>

Table 2-8. Optimized segment numbers for each method used for NRTL-SAC model

<table>
<thead>
<tr>
<th>Method</th>
<th>X (Hydrophobic)</th>
<th>Y- (Polar attractive)</th>
<th>Y+ (Polar repulsive)</th>
<th>Z (Hydrophilic)</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>0.820</td>
<td>0.366</td>
<td>0.484</td>
<td>0.277</td>
</tr>
<tr>
<td>II</td>
<td>1.154</td>
<td>0.001</td>
<td>0.831</td>
<td>1.393</td>
</tr>
<tr>
<td>III</td>
<td>1.175</td>
<td>0.000</td>
<td>0.548</td>
<td>0.882</td>
</tr>
</tbody>
</table>

Referring to Table 2-7 it can be seen that from the four methods of predicting the solubility, method III is the best one which describes and predicts the solubility of lovastatin. The next is NRTL-SAC which was based on estimation of the two acetate solvents (method II). The UNIFAC method has the highest average deviation from the experimental values. In Table 2-7 in the second column, the ARD is based on only the deviation of predicted from experimental solubility of alcohols. The prediction by the UNIFAC model and the method I (which was based on parameter estimation of alcohols only) are nearly the same. However the minimum average deviation in this column is for the method III. In column three the ARD is based on the acetate solvent data. The method II has the lowest deviation which shows the selection of methyl- and ethyl- acetate from the acetate solvents was satisfactory. From Table 2-8 it is evident that the values of hydrophobic and polar attractive segments are very near to those of the method II.
(acetate estimation). However, this behaviour is different for polar repulsive and hydrophilic segments, which are nearly the arithmetic average of the methods I and II. The method III was the optimum one to minimize the total average deviation and hence best describes the solubility of lovastatin in alcohols and acetate solvents. Finally it can be understood from the results that the UNIFAC method was not able to predict the solubility data as well as the NTRL-SAC.

2.4.3. Solubility prediction of valsartan

For the solubility prediction of valsartan, we used the experimental data of Liu et al. [25, 26] to estimate the adjustable parameters for NRTL-SAC. The chemical structure of valsartan is shown in Figure 2-6. One group contains four nitrogen atoms in a pentagonal ring. To the best of our knowledge and from the thermodynamic references and newly published articles, this group does not have any UNIFAC parameter. This is one of the weaknesses that the UNIFAC model has in face of newly complex chemical and pharmaceutical materials. In order to predict the pure solvent solubility and solvent + antisolvent solubility, three solubility data sets were selected for estimation and the remaining data were predicted using the adjusted segment numbers.

The solubility data in methyl acetate, DMF, and dichloromethane were used for parameter estimation. After optimizing the model and obtaining the segment numbers, the solubility of n-butyl acetate and acetonitrile were predicted. The ARD for this prediction is 0.3843 which is very good for prediction purposes. In the two graphs in Figure 2-13 the square points are the coordination of the point that shows the experimental solubility (x-axis) and the prediction of the solubility (y-axis). For the case of solvent mixture [26], the optimized segment numbers (Table 2-4) was used to predict the solubility. Here the ethyl acetate and hexane is used as solvent and antisolvent, respectively. It is worth noting that if \( x_1 \) is the mole fraction of the solute in the solution, then with having the molar ratio (or mass ratio) of antisolvent to solvent, one can find the mole fractions of the two solvents in the system. With the assumption that the molar ratio of antisolvent to the solute-free solution is \( M_w \), then the mole fractions of each species in the mixture will be:

\[
\begin{align*}
    x_1 &= \text{mole fraction of solute} \\
    x_2 &= (1 - x_1)(1 - M_w) \\
    x_3 &= \frac{M_w}{(1 - M_w)} x_2
\end{align*}
\]


Figure 2-13. Predicted and experimental values of valsartan in two solvents: left: acetonitrile, right: butyl acetate

Therefore, the mole fractions of all of the three compounds are functions of the solute mole fraction, which is a function of the activity coefficient. As a result, for this type of calculation the trial and error method have to be used to converge to a unique value of $\Delta u_1$. Having this included in the modeling code, we started with the trial and error procedure to attain the solubility for each antisolvent to solvent ratio over a range of temperature. Curves of solubility prediction for pure ethyl acetate to the solution of up to the 49.44 mol%. ethyl acetate are illustrated in Figure 2-14.
From the optimized segment numbers for valsartan we see that this drug does not have the hydrophobic and polar-repulsive parts in its chemical structure. On the other hand, it has polar attractive and hydrophilic segments, which can be rationalized by the carboxy group (COOH) and also hydrogen head. Both Figures 2-13 and 2-14 show satisfactory prediction of the model at different mixture ratios and temperatures. This shows the robustness of the NRTL-SAC model to predict the ternary systems of solute-solvent-antisolvent.
2.5. Conclusion

In this study we used two predictive methods for the phase behaviour of solid liquid equilibrium (SLE) systems. There are many thermodynamic models which can describe the solubility of a pharmaceutical or a chemical compound in a solvent or mixtures of solvents. However, the correlative models have some restrictions such as the need to have a wide range of experimental data in order to find the binary interaction parameters (i.e., Wilson). The UNIFAC model was able to describe the solubility of the three selected solutes in the solvents with a reasonable accuracy. We used some modifications in optimization of the adjustable parameters for NRTL-SAC model. The NRTL-SAC showed a very good accuracy for pure solvent systems. Also, for the mixture of solvents (in section 2.4.3) we were not able to use the UNIFAC model. Using a few experimental data points for segment number estimation of the NRTL-SAC method allowed us to predict the solubility in pure and also mixed solvents, successfully. For the case of solvent-antisolvent solubility prediction, 3-pentadecylphenol phase behaviour in the mixture of IPA-water was best predicted by the NRTL-SAC model. The maximum point of solubility was predicted and verified by experimental data points. The NRTL-SAC predicted the solubility of valsartan in a mixture of hexane and ethyl acetate, successfully.

2.6. References


21- Matlab software help documents, version 7.6.0.324 (R2008a).

Chapter 3

Vapour-Liquid and Vapour-Liquid-Liquid Equilibrium Modeling for Binary, Ternary, and Quaternary Systems of Solvents

A version of this chapter has been published as:

3.1. Introduction

Thermodynamic models have been used to aid chemical engineers in designing various chemical and pharmaceutical processes. Although high-throughput solubility measurement techniques are developing, they are time-demanding and expensive in finding the best operating conditions for a specific process. This issue is more important when one deals with the crystallization of a chemical compound in various solvents and their possible combinations. In this case, depending on a solvent or solvent mixture, so many possible cases have to be tested in order to reach the maximum solubility. In many applications where there are at least two phases in equilibrium (such as crystallization process), the need to have an accurate thermodynamic model with less dependency on experimental data is significant. Many thermodynamic models, such as Wilson [1], Van Laar [2], NRTL [3], and UNIQUAC [4] have been used to help researchers design chemical processes. The main disadvantage of these equations of state (EOS) is their dependency on experimental data. For example, to predict the phase behaviour of a ternary vapour-liquid equilibrium (VLE) system, one should have three sets of binary experimental data for all the constituents of the system. Another problem dealing with the models based on the equation of state is that the adjusted parameters are valid at a given pressure or temperature of the system. As the total pressure and temperature of the system change, the fugacities in the vapour and liquid phases can be changed, which may affect the mole fractions of species and other dependent properties, such as bubble and dew point temperatures. One of the main advantages of predictive models is that they do not require any experimental data on two, three, or more component mixtures.

Since the universal functional activity coefficient (UNIFAC) model was proposed by Fredenslund et al. [5], there have been much effort to find a model which can describe the phase behaviour (i.e., vapour-liquid equilibrium) with no or limited experimentation. The UNIFAC model is powerful in predicting vapour-liquid and liquid-liquid equilibrium systems with average molecular weights of 100g/mol. However, when the molecular weight of a chemical species is in the range of 200-600g/mol with complex chemical structure, the UNIFAC model and its modifications fail to generate reasonable predictions [6].

Chen and Crafts [7] proposed a new predictive model, Non-random Two Liquid Segment Activity Coefficient (NRTL-SAC) which is based on polymer NRTL model [8]. According to
This model, four individual conceptual segments are assigned to each component, regardless of its nature or chemical structure complexity. The segment numbers are unique for each compound. According to the NRTL-SAC model, the molecule surface can have at least one or more of the following segments:

- Hydrophobic segment (x),
- Polar-attractive segment (y-),
- Polar-repulsive segment (y+) and
- Hydrophilic segment (z)

A hydrophobic segment (X) does not form a hydrogen bond, such as heptane. A hydrophilic segment (Z) contributes to the part of the molecule which tends to form a hydrogen bond, such as methanol. The polar segments (Y- and Y+) do not belong to the hydrophobic nor hydrophilic segment. The polar attractive segment (Y-) shows attractive interaction with hydrophilic segment, while the polar repulsive segment (Y+) has repulsive characteristic with hydrophilic segment, such as acetone. Depending on the nature and chemical structure of a molecule, the contribution of the above four mentioned segments varies. Chen and Song have extracted the numerical attributes of segments for common solvents. Therefore, for a given set of components, one can calculate the phase behaviour of the mixture with using four parameters which are constant and unique for each constituent of the solution. The segment numbers of most common industrial solvents have been estimated and tabulated in the literature [6].

Recent works have been conducted on modeling phase equilibrium systems and finding segment numbers for new components. Chen and Crafts [7] found the segment numbers for acetaminophen, sulfadiazine, cimetidine, and sulfamerazine. Mota et al. [9] used this model to adjust the segment numbers for allopurinol, furosemide, and budesonide. Sheikholeslamzadeh and Rohani [10] used the NRTL-SAC and UNIFAC model to predict the solubility of 3-pentadecylphenol, lovastatin, and valsartan in various solvents and their mixtures. They concluded that the NRTL-SAC model can be used in many industrial applications and pharmaceutical processes. The NRTL-SAC model has shown promising predictions and applicability in many areas of interest. Ferreira and Pinho [11] have investigated the solubility of flavonoids in pure solvents, experimentally and theoretically using the NRTL-SAC method.
In order to ensure the predictive capability of a model in vapour-liquid equilibrium calculations (specifically bubble point temperature prediction), different solvent mixtures should be employed and the model predictions should be compared with the corresponding values from other correlative models and experimental data.

Equations (3-1) and (3-2) are used to quantify the deviations of the suggested model predictions from the experimental data:

- Average Relative Deviation, \[ \text{ARD} = \frac{1}{N} \sum_{i=1}^{N} \left| \frac{x_i^{\text{exp}} - x_i^{\text{calc}}}{x_i^{\text{exp}}} \right| \] (3 - 1)
- Average Absolute Deviation, \[ \text{AAD} = \frac{1}{N} \sum_{i=1}^{N} |x_i^{\text{exp}} - x_i^{\text{calc}}| \] (3 - 2)

Where \( x_i^{\text{exp}} \) and \( x_i^{\text{calc}} \) are experimental and theoretical mole fractions of a specific compound in a solution, respectively. In a multiphase system at equilibrium, the fugacity of every compound in each phase is equal to the corresponding value in other phases [3]:

\[ f_1^a = f_1^b = \ldots = f_i^y \quad i = 1, 2, \ldots, N_c \] (3 - 3)

\( N_c \) is the total number of compounds present in the system. For a vapour-liquid-liquid equilibrium system, equation (3-3) can be written in the following form:

\[ x_i^a \gamma_i^{up} p_i^{sat} = x_i^b \gamma_i^{bp} p_i^{sat} = y_i p_{tot} \varphi_i \] (3 - 4)

where \( x_i^a \), \( x_i^b \), and \( y_i \) are the mole fractions of component \( i \) in two liquid and vapour phases, respectively. \( p_i^{sat} \) is the vapor pressure of component \( i \) at an arbitrary temperature and total pressure \( p_{tot} \). \( \varphi_i \) is the fugacity coefficient of component \( i \) in the vapour phase, which can be equal to unity for low to moderate pressures (the studied cases in this work were at low pressures). For VLLE systems, one can select an arbitrary phase as a reference for bubble point or dew point calculations. For our study, we selected organic and aqueous liquid phases as reference. The data on saturated pressure of the solvents have been found from the NIST database [12]. The mathematical form of vapor pressure for a pure compound (Antoine equation) is:

\[ \log_{10}(p_i^{sat}) = A_i - \frac{B_i}{T + C_i} \quad i = 1, 2, \ldots, N_c \quad (3 - 5) \quad \text{P [bar], T [K]} \]
The flash calculations were written in separate codes for every thermodynamic model. For the UNIFAC model, the activity coefficient is temperature dependent, thus we had to build a separate code for trial and error procedure. All the codes were written in the Matlab environment and further details can be provided to the interested reader. The details of the procedure for finding the bubble, dew, and flash calculations can be found elsewhere [13].

3.2. Theoretical background

3.2.1. Thermodynamic models

The mathematical description of the two thermodynamic models, UNIFAC and NRTL-SAC, have been introduced in chapter 2. Therefore, we don't mention the details of how the models predict the phase behaviour of the systems.

3.2.1.1. Phase behaviour calculations

3.2.1.2. Vapour-liquid equilibrium

In order to find the flash calculation of the systems containing vapour and liquid, we must start from the equilibrium equation for the VLE system:

\[ y_i \phi_i P_{tot} = y_i x_i P_{sat} \]  \hspace{1cm} (3 – 6)

Which \( y_i \) and \( x_i \) are the mole fractions of component \( i \) in vapour and liquid phases, respectively. \( \phi_i \) and \( y_i \) denote the deviation from ideal conditions for vapour and liquid phases, respectively. Equation (3-6) is the basis for determining the bubble, dew, and flash calculations. The bubble point pressure for a given temperature is found from:

\[ P_{tot}^{\text{bubble}} = \sum_{i=1}^{N_c} y_i x_i P_{sat} \]  \hspace{1cm} (3 – 6)

The procedure to find the bubble point temperature of a liquid solution at a given pressure is as follows:

1. Estimate bubble temperature: \( T = \sum_i x_i T_i^{sat} \)  \hspace{1cm} (3 – 7)
2. Evaluation of \( P_i^{sat} \) and \( y_i \) at the given temperature
3. Calculation of $P_j^{\text{sat}}$ by:

$$P_j^{\text{sat}} = \frac{1}{\sum \varphi_i \left( \frac{P_i^{\text{sat}}}{P_j^{\text{sat}}} \right)}$$

(3 - 8)

4. Calculation of new bubble temperature by:

$$T = \frac{B_j}{A_j - \ln P_j^{\text{sat}}} - C_j$$

(3 - 9)

5. Evaluation of the new $P_i^{\text{sat}}$, calculation of $y_i$, $\varphi_i$, $\gamma_i$, $P_j^{\text{sat}}$, and finally $T$

6. If the calculated $T$ from step 5 and 4 are different the trial and error from step 1 should be repeated.

### 3.2.1.3. Liquid-liquid equilibrium

For the systems exhibiting the liquid-liquid equilibrium behaviour, we use the fundamental relation of equation (3-4). Thus:

$$x_i^a \gamma_i^a = x_i^b \gamma_i^b$$

(3 - 10)

$\gamma_i^a$ and $\gamma_i^b$ are generally the functions of mole fraction of constituents, temperature, and pressure. Equation (3-10) is written for each of the species in the solution and solved for the system at specific pressure and temperature of interest.

### 3.2.1.4. Vapour-liquid-liquid equilibrium

In order to investigate the existence of three phases in a system, equation (3-4) should be written for each pair of phases and the results from each calculation should be re-corrected by the results of other phases. In this study, we have selected both aqueous and organic phases for model validation.

### 3.3. Test cases

#### 3.3.1. VLE system of Toluene-Ethanol and Toluene-Isopropanol at different pressures

Ethanol, toluene, and isopropanol are common solvents used in large volumes in various chemical and pharmaceutical industries. Also, they are used as an additive to other solvents to promote their physical and/or chemical characteristics. Because of the formation of azeotrope for toluene-ethanol and toluene-isopropanol systems, they are of great interest in azeotropic distillation process design. According to a study by Stichlmair et al. [16], pressure swing distillation process can be employed to overcome the azeotropic nature of the mixture. In order
to have an accurate knowledge of vapour-liquid behaviour of these binary systems at different temperatures and pressures and to simulate and optimize the unit operations, a reliable and flexible thermodynamic model has to be used. In the work by Chen et al. [17], a recirculating apparatus was used for VLE measurements of toluene-ethanol and toluene-isopropanol systems. They found that the system was highly non-ideal and exhibited azeotropic behaviour. The VLE data were extracted at four different pressures (101.3, 121.3, 161.3, and 201.3 kPa) assuming that the vapour phase can be treated as an ideal mixture. Chen et al. [17] also made use of Wilson, NRTL, and UNIQUAC equations of state to fit the measured data with the experimental values and found the interaction parameters, but did not predict the experimental data with the models.

In our study, we used the NRTL-SAC and UNIFAC models to predict these highly non-ideal systems in liquid phase. Figure 3-1 compares the model prediction and the experimental data for the systems. The NRTL-SAC model predicts the experimental data of bubble and dew point with a reasonable accuracy. It is shown that UNIFAC model could anticipate the azeotrope formation for all of the cases with respect to the mole fractions of ethanol, with close agreement with the literature data. Because of this ability, the UNIFAC model can be used in simulation and optimization of azeotropic distillation of these systems, effectively. On the other hand, the NRTL-SAC model could not predict the azeotrope point for either of two systems.

The average relative deviation of two models for two solvents from the experimental data is shown in Table 3-1. From this table, the average deviations of prediction from the experimental value for vapour phase mole fractions with the use of the UNIFAC model are about one third of the corresponding values from the NRTL-SAC model. Also, for bubble temperature deviations, the ARD from the NRTL-SAC are three times larger than the UNIFAC model. It is worth noting that the average relative deviations for both binary systems using NRTL-SAC model are nearly the same. However, this is not the same when using the UNIFAC model. The deviation for the system of toluene-ethanol is less than that of the toluene-isopropanol. It shows that the UNIFAC can better predict the light alcohol systems than the heavier ones.

From the results that are shown in Table 3-1, two more conclusions can be drawn:

- The ARD of the bubble point temperature for both models is small, which shows the applicability of two methods for accurate bubble temperature calculations.
At higher pressures, the ARD is generally increased. This is because of neglecting the vapour phase fugacity coefficients which has been made in calculation. In spite of the mentioned fact, the deviation is not that much to include a model to predict the vapour phase fugacities for pressures up to 2 atm. For higher operating pressures (i.e., 5 atm and above), the vapour phase should also be considered in evaluation of species’ fugacity coefficients.

**Table 3-1. Deviation of predictive model results from corresponding experimental values of Chen et al. [17]**

<table>
<thead>
<tr>
<th>Thermodynamic model</th>
<th>Pressure, [kPa]</th>
<th>%ARD (equilibrium temperature and vapour-phase mole fractions)</th>
<th>System 1</th>
<th>System 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Temperature, [K]</td>
<td>Ethanol</td>
<td>Toluene</td>
</tr>
<tr>
<td>NRTL-SAC</td>
<td>101.3</td>
<td>0.46</td>
<td>5.14</td>
<td>9.81</td>
</tr>
<tr>
<td></td>
<td>201.3</td>
<td>0.64</td>
<td>7.63</td>
<td>9.72</td>
</tr>
<tr>
<td>UNIFAC</td>
<td>101.3</td>
<td>0.10</td>
<td>1.08</td>
<td>3.15</td>
</tr>
<tr>
<td></td>
<td>201.3</td>
<td>0.27</td>
<td>2.03</td>
<td>3.47</td>
</tr>
</tbody>
</table>
Figure 3-1. Predicted and experimental results for systems (A) ethanol-toluene at 101.3kPa, (B) ethanol-toluene at 201.3kPa, (c) isopropanol-toluene at 101.3kPa, and (D) isopropanol-toluene at 201.3kPa.
3.3.2. VLE data for binary and ternary systems of Ethanol-Water-Ethylene Glycol

Nowadays with increasing fossil fuel prices, the need to find alternative energy is so significant. Ethanol is a promising candidate in blending with the fuel to increase its performance in burning and decreasing air pollution by transportation vehicles. In addition to the environmental and efficiency characteristics, the cost saving of blending ethanol is important. The performance of the final fuel blend increases with an increase in the ethanol purity. It is shown that using glycols can enhance the separation of ethanol from water (dehydration process) with consuming less energy than conventional distillation processes [18]. In the work by Kamihama et al. [19], the effect of adding glycol to the mixture of ethanol and water (hydrated alcohol) on separation performance has been investigated.

![Figure 3-2. Txy diagram of VLE systems; (A) ethanol-water, (B) ethanol-ethylene glycol, and (C) water-ethylene glycol all at 101.3kPa](image-url)
They conducted three main experiments on every pair of solvents from the ternary system of ethanol-water-ethylene glycol. They concluded that ethylene glycol has an effect on separation of ethanol and water beyond the azeotrope point. We have made use of two predictive models to predict the experimental data of this system.

According to Figure 3-2 the UNIFAC model can predict the phase behaviour of ethanol-water system well. The other two systems, specifically the case of water-ethylene glycol, were not predicted satisfactorily. The NRTL-SAC, however, can give good results for all the binary systems, especially for water-ethylene glycol. It should be noted that in the cases which there is not an azeotrope point, the NRTL-SAC has the ability to predict the experimental points with the same or even lower deviation from experimental points than the UNIFAC method (Figures 3-2B and 3-2C). However, in the presence of the azeotrope point, the UNIFAC yields a better prediction (Figure 3-2A). In order to have a quantitative assessment of the predictive ability of two methods, we calculated the average relative deviation for three binary systems. As it is seen from Table 3-2, the NRTL-SAC model can best predict the systems of water-ethylene glycol and ethanol-ethylene glycol, while UNIFAC resulted in lower ARD for ethanol-water system.

**Table 3-2.** Absolute relative deviation (%ARD) of predicted and experimental vapour mole fraction of components in each binary system

<table>
<thead>
<tr>
<th>Model</th>
<th>System 1</th>
<th>System 2</th>
<th>System 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ethanol</td>
<td>Water</td>
<td>Water</td>
</tr>
<tr>
<td>NRTL-SAC</td>
<td>3.68</td>
<td>7.23</td>
<td>1.42</td>
</tr>
<tr>
<td>UNIFAC</td>
<td>1.22</td>
<td>2.19</td>
<td>3.58</td>
</tr>
</tbody>
</table>

Kamihama et al. [19] also performed the VLE measurements for ternary system of ethanol-water-ethylene glycol at 101.3kPa. They used the correlative models (such as UNIQUAC) to find the binary parameters using three sets of binary data, and then predicted the ternary system. For the NRTL-SAC model, we just used four conceptual segments of each solvent, which are already found in the literature [6]. The NRTL-SAC model predicted the equilibrium of the
ternary solvent system with a high degree of accuracy. This model could predict the vapour mole fractions, as well as the equilibrium temperature. From Figure 3-3, except for ethylene glycol, the mole fractions of ethanol and water in the vapour phase were predicted with nearly zero error by the NRTL-SAC model. However, there were a few deviations for the results from the UNIFAC model. The ethylene glycol vapour phase mole fraction for all of the experimental runs were nearly zero and if we consider the contribution of inherent errors of experiments, the NRLT-SAC model also described the behaviour of ethylene glycol in the vapour phase, well. As it is evident from this figure, the NRLT-SAC could successfully anticipate the bubble and dew points of the system. The results of the modeling showed that the NRLT-SAC predicts the data with one third average relative deviation compared to the UNIFAC method.

Figure 3-3. Vapour phase mole fractions of ethanol (top-left), water (top-right), ethylene glycol (bottom-left), and saturated temperature of the solution (bottom-right) from experimental data and theoretical models for the ternary system of ethanol-water-ethylene glycol
3.3.3. Binary and ternary system of Diethyl Ether-Methanol-Butanol

In another study by Gao et al. [20], binary and ternary combinations of diethyl ether with methanol and 1-butanol were examined in vapour and liquid phase equilibrium. This study is mainly due to the importance of the presence of by-products (i.e., diethyl ether) during syngas conversion to methanol. By adjusting binary interaction parameters, they could anticipate the VLE behaviour of the three-component system.

![Figure 3-4. Three binary phase diagrams of (A) diethyl ether-methanol, (B) diethyl ether-butanol, and (C) methanol-butanol at 101.3kPa from experiment and models.](image)

Having a precise knowledge of the phase equilibria inside the syngas reactor will help the engineer in efficient designing and optimizing the process. Even more, accurate design of the downstream processes (such as distillation column) depends on thermodynamic equilibrium in such systems. We examined the predictability of two models for the binary and ternary data [20].
First, the binary data on every pair of the ternary system was modeled with both predictive methods. The results of binary simulation are shown in Table 3-3 and Figure 3-4. Two main conclusions which can be drawn from binary results are:

- For system 1 which contains diethyl ether as one of the constituents, the NRTL-SAC shows better performance in prediction, specifically saturation temperature. For this system there is an azeotrope point near to diethyl ether unit mole fraction, which is also predicted by the NRTL-SAC model.
- For the second and third systems, the UNIFAC and NRTL-SAC show nearly the same prediction.

**Table 3-3.** Binary solution deviation of model vs. experimental data for three pairs of Diethyl ether-Methanol-Butanol (%ARD)

<table>
<thead>
<tr>
<th>Model</th>
<th>System 1</th>
<th>System 2</th>
<th>System 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diethyl ether</td>
<td>Methanol</td>
<td>Saturated temperature</td>
<td>Diethyl ether</td>
</tr>
<tr>
<td>NRTL-SAC</td>
<td>2.23</td>
<td>6.28</td>
<td>0.15</td>
</tr>
<tr>
<td>UNIFAC</td>
<td>4.28</td>
<td>9.54</td>
<td>0.42</td>
</tr>
</tbody>
</table>

In addition to the binary calculation, the ternary system was also studied. The numerical values of absolute average deviation are shown in Table 3-4. The reason for adopting the AAD rather than ARD for this system is to compare the results of Gao et al. [20] with our work. It is seen from the table that the NRTL-SAC has a better accuracy in prediction in comparison with the UNIFAC method. If the AAD of mole fractions of methanol and butanol are compared with the UNIQUAC model, the ability of prediction of the NRTL-SAC model (with only four parameters) will be obvious. The UNIFAC did not yield satisfactory results for this ternary system.
Table 3-4. Average absolute deviation (%AAD) for predictive models and two correlative models for a ternary system of Diethyl Ether-Methanol-Butanol

<table>
<thead>
<tr>
<th>Model</th>
<th>Diethyl ether</th>
<th>Methanol</th>
<th>n-Butanol</th>
</tr>
</thead>
<tbody>
<tr>
<td>NRTL-SAC</td>
<td>2.28</td>
<td>1.77</td>
<td>1.29</td>
</tr>
<tr>
<td>UNIFAC</td>
<td>5.33</td>
<td>4.13</td>
<td>1.21</td>
</tr>
<tr>
<td>UNIQUAC [20]</td>
<td>1.71</td>
<td>1.97</td>
<td>2.62</td>
</tr>
<tr>
<td>Wilson [20]</td>
<td>1.68</td>
<td>1.36</td>
<td>2.04</td>
</tr>
</tbody>
</table>

3.3.4. VLE and VLLE equilibrium of Water-Ethanol-Hexane-Toluene and Water-Ethanol-Cyclohexane-Isooctane.

Generally the system of water-hydrocarbons is of importance, because they are mainly produced and consumed in most oil and chemical industries. Pequenin et al. [21] used ethanol plus hydrocarbon mixture (i.e., n-hexane and toluene) as a simulated gasoline. In the first study they worked with the system of water-ethanol-cyclohexane-isooctane [21]. In the second work, they performed experiments on ternary and quaternary systems of water-hexane-toluene and water-ethanol-hexane-toluene, respectively [22]. Both studies were conducted under 101.3kPa total pressure. For emulsification and preventing temperature fluctuations inside the solution, they employed an ultrasound homogenizer. For each data point, they measured the boiling temperature and mole fractions of the liquid phases and vapour phase of the system.

VLE data of Water-Hexane-Toluene. The study on the system of water-hexane-toluene showed that in aqueous phase the mole fraction of water is nearly unit, while in the organic phase hexane and toluene were the dominant parts of the phase. Since the aqueous phase was almost pure water, this phase was not considered for equilibrium calculations. Therefore, we selected the organic and vapour phase for our study. The mole fractions of the vapour phase and the corresponding equilibrium solution temperatures with predicted values are shown in Figure 3-5. As it is evident from the results, the bubble temperature of the four runs from the NRTL-SAC is well predicted compared to the UNIFAC model. For water and hexane mole fraction in vapour phase, the NRTL-SAC model could anticipate the phase behaviour with a higher degree of
accuracy compared to the UNIFAC model. For toluene at low mole fractions, the NRTL-SAC results are closer to their corresponding experimental values, while for higher mole fractions, the UNIFAC was the superior model in prediction.

**Figure 3-5.** The results of prediction vs. experimental runs for VLE system of water-hexane-toluene; (A) bubble point temperature, (B) water, (C) hexane, and (D) toluene mole fraction in vapour phase

**VLLE system of Water-Ethanol-Hexane-Toluene (system 1) and Water-Ethanol-Cyclohexane-Isooctane (system 2).** For the VLLE system of quaternary compounds, we made the calculation based on two methods:

- VLLE prediction based on aqueous phase
- VLLE prediction based on organic phase

As long as the equilibrium condition is maintained, the fugacity of the compounds in each pair of phases should be equal (Equation (3-4)). Therefore, we first used aqueous phase as a base phase to predict the data. After that, the organic phase was chosen as the reference phase for
calculation. For quantitative analysis of the modelling procedure, the average relative deviation for each of four mole fractions for the two systems is shown in Table 3-5. From this table, the total average deviation for system 1 using NRTL-SAC model is 0.1003 and 0.1269 for organic and aqueous phases, respectively. In case of the UNIFAC model, the average deviation for the whole system 1 is 0.1269 and 0.1057 for organic and aqueous phase, respectively. According to this result, the prediction performance for the two models is nearly the same. When the organic phase is used for calculation, the NRTL-SAC gives better results.

### Table 3-5. %ARD for the two quaternary systems of VLLE

<table>
<thead>
<tr>
<th>Thermodynamic model</th>
<th>%ARD (vapour-phase mole fractions) based on aqueous phase</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>System 1</td>
<td>System 2</td>
</tr>
<tr>
<td></td>
<td>Water Ethanol Hexane Toluene</td>
<td>Water Ethanol Cyclohexane Isooctane</td>
</tr>
<tr>
<td>NRTL-SAC</td>
<td>6.07 3.94 8.93 5.95</td>
<td>9.53 8.29 10.19 10.31</td>
</tr>
<tr>
<td>UNIFAC</td>
<td>4.29 5.04 12.48 8.71</td>
<td>10.10 10.23 9.59 11.60</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>%ARD (vapour-phase mole fractions) based on organic phase</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>System 1</td>
<td>System 2</td>
</tr>
<tr>
<td></td>
<td>Water Ethanol Hexane Toluene</td>
<td>Water Ethanol Cyclohexane Isooctane</td>
</tr>
<tr>
<td>NRTL-SAC</td>
<td>3.40 3.19 4.67 1.60</td>
<td>9.50 9.32 6.67 12.69</td>
</tr>
<tr>
<td>UNIFAC</td>
<td>12.04 4.94 6.72 1.09</td>
<td>6.32 4.76 2.44 10.44</td>
</tr>
</tbody>
</table>

According to the results depicted in Table 3-5, the NRTL-SAC model could predict the vapour phase mole fractions of water, ethanol, and hexane with lower error compared to the UNIFAC model. For the toluene prediction, the result depended on the method which the phase calculation was based. If the aqueous phase was selected, then the error from NRTL-SAC model would be lower, and for the organic phase, the UNIFAC would have the lower prediction error. The results
Figure 3-6. Predicted and experimental phase mole fractions of three species in a quaternary system of VLLE, 1) system of Water-Ethanol-Hexane-Toluene, prediction based on (A) organic phase and (B) aqueous phase, 2) system of Water-Ethanol-Cyclohexane-Isooctane, prediction based on (C) organic phase and (D) aqueous phase.
of prediction are shown in Figure 3-6(A-B). The experimental and theoretical points are depicted in triangles to show the mole fractions of the quaternary system. According to the Figure 3-6 the two models could predict the quaternary system, except UNIFAC for the case of organic phase base calculation. There is some deviation from experimental results for the UNIFAC when the organic phase is selected as a base phase. We conclude that the NRTL-SAC can be a better choice in VLLE calculation of this system, rather than the UNIFAC. From Table 3-5, we conclude that the NRTL-SAC can be a better choice in VLLE calculation of this system compared to the UNIFAC. For system 2 the total average deviation from NRTL-SAC model is 0.1056 and 0.1078 for organic and aqueous based calculation. The two deviations are nearly the same, which confirms the liquid-liquid-vapour equilibrium condition in the system from the NRTL-SAC model. The deviations obtained by the UNIFAC model were 0.0805 and 0.1141 for organic and aqueous phase base, respectively. The two deviation values differ from each other, illustrating that the UNIFAC model cannot verify the three phase formation, successfully. The system of water-hexane-toluene and the quaternary system of it having methanol as the added component were modeled by the two methods. For the ternary system, the bubble point was perfectly verified using NRTL-SAC model, while the mole fractions of three compounds have a few deviations from experimental values. For the two systems of quaternary solvents, depending on selection of the aqueous or organic phase as a reference for calculation, the results were different. If the aqueous phase was considered as a reference phase, the results from two models showed large deviations, especially for UNIFAC model. However, when the organic phase was selected as a reference phase, the results were much better than the former case. The main reason for this behaviour can be the slight solubility of hydrocarbons in aqueous phase, while the most portion of the hydrocarbons are in organic phase.

3.4. Conclusion

In this work, the applicability of two predictive activity coefficient based models was examined. The experimental data of five different types of VLE and VLLE systems that are common in industry were used for the evaluation. The NRTL-SAC and UNIFAC models were selected for modeling the systems. In general, the NRTL-SAC model could predict all the binary systems, except light alcohols in water and in aromatics. The accuracy of the NRTL-SAC model was higher in ethylene glycol with alcohols and dimethyl ether with water. The NRTL-SAC model,
also could best predict the ternary system of solvents. The UNIFAC model could accurately predict the binary system of light alcohols with water and aromatic hydrocarbons. The ternary systems, however, were modeled by the UNIFAC model with a less accuracy than the NRTL-SAC model. For the quaternary systems, if the organic phase was selected as a reference phase, the NRTL-SAC could give good results for water-ethanol-hexane-toluene, while the UNIFAC gave better results for water-ethanol-cyclohexane-isooctane system. It can be concluded that because of fewer parameters and ease of computation, the NRTL-SAC is a more powerful method compared to the UNIFAC.

3.5. References


Chapter 4

Optimal Solvent Screening for the Crystallization of Pharmaceutical Compounds from Multi-Solvent Systems

A version of this chapter has been published as:

4.1. Introduction

In the production of many pharmaceuticals and chemicals, the solvent selection for the crystallization process is very important [1]. The solubility of an active pharmaceutical ingredient (API) in a solvent or solvent mixture can affect the drug release and absorption in the body and its transport in a living organism. The drug availability for experimental solubility determination during the early stages of a new drug development is very restrictive [2]. In addition, the difficulties in experimentation and inherent experimental errors are other complexities. On the other hand, accurate thermodynamic models can provide guidelines for the optimal selection of solvent combinations. Some thermodynamic methods, such as NRTL [3], Wilson [4], UNIQUAC [5], and Van Laar [6] rely on huge experimental data. For systems of ternary and quaternary solvents, the use of these models requires vast number of experimental runs to determine the binary interaction parameters [7]. The predictive thermodynamic models such as the UNIFAC [8], NRTL-SAC [2], and COSMO-SAC [9-10], require information on the molecular structure of the solute and solvents. The UNIFAC model has shown a good predictive ability for a wide range of VLE and LLE systems. However, it fails to deal with the systems in which the molecules have complex and unknown functional groups [2]. The NRTL-SAC model has been proposed by Chen and Song [2]. This model is derived from the polymer non-random two-liquid model [11]. In this model the molecules of solute and solvent are characterized by three segments; hydrophobic (X), hydrophilic (Z), and polar (Y). The polar segment itself is divided to polar attractive (Y-) and polar repulsive (Y+). These conceptual segments are found using some experimental data on VLE and LLE systems. The four conceptual segments assigned to each molecule (solute or solvent) can be used as a descriptor of that molecule when used in VLE, LLE, and SLE calculations. This model is simple and flexible in modeling the thermodynamic phase equilibrium of binary and multi-component systems. The advantage of this model, compared to other predictive methods such as UNIFAC, is its simplicity in defining a molecule with only four segment numbers, no matter how complex the chemical structure of a molecule is.

The VLE and SLE predictive ability of the NRTL-SAC model has been verified in some recent studies. Chen and Crafts used the NRLT-SAC model to find the segment numbers of paracetamol and sulfadiazine [12]. In another study they showed this model can be used in VLE
calculation of common solvents [2]. Sheikholeslamzadeh and Rohani found the conceptual segment numbers of 3-pentadecylphenol, lovastatin, and valsartan in single and binary solvents [13]. Sheikholeslamzadeh and Rohani have also studied different systems of solvents in binary, ternary, and quaternary mixtures to verify the applicability of the NRTL-SAC model in predicting the VLE and VLLE behaviour of those systems [14]. They showed the NRTL-SAC model best describes such a complex system of solvents in good agreement with the experimental data. The results showed the NRTL-SAC model could give better results in ternary and quaternary mixtures than the UNIFAC model. Mota et al. used the NRTL-SAC model to estimate the segment numbers of budesonide, allopurinol, and furosemide in different solvents [15].

In this work the NRTL-SAC model is written in separate programs for VLE, LLE, and SLE prediction of multi-solvent systems with seven model pharmaceutical molecules. A new optimization procedure is developed to calculate the best solvent or mixture of solvents to maximize the yield of crystallization in a cooling-antisolvent mode. It should be noted that in the current study, the process modeling and optimization are based on initial and final points of the operation only. The path of cooling or addition of antisolvent is not considered here and in addition, the initial and final operating points of crystallization are at equilibrium. It is worth noting that throughout this work, the component refers to a solid, solute, or solvent under study. A model molecule (an active pharmaceutical ingredient, API, or any chemical compound) can be in the form of a solute (dissolved in a solvent or a solvent mixture) or in the form of a solid (a crystallized or amorphous form of a solute).

4.2. Thermodynamic background

As it was mentioned in chapter 2, the group-contribution models divide the contribution of the activity coefficient to two parts, the combinatorial and residual parts. The combinatorial part includes the contribution of the chemical structure and the size (volume and surface of the molecule) of the component. The residual part includes the contribution of the group size and binary interaction between pairs of the functional groups. The thermodynamic model of NRTL-SAC has been reviewed in previous chapters.
4.2.1. VLE, LLE, and SLE prediction

For any component in equilibrium with itself in a multi-phase system, its fugacity in each phase is equal to the corresponding value in other phases [16]:

\[ f_i^α = f_i^β = \cdots = f_i^γ \quad i = 1, 2, \ldots, N \quad (4 - 1) \]

\( N \) is the total number of solvents present in the system. For a vapour-liquid and solid-liquid equilibrium system, equation 4-1 can be written in the following forms [18]:

SLE: \( x_i y_i p_i^{sat} = f_i^s \) \quad (4 - 2)

VLE: \( x_i y_i p_i^{sat} = y_i p_{tot} \varphi_i \) \quad (4 - 3)

LLE: \( (x_i y_i)^α = (x_i y_i)^β \) \quad (4 - 4)

Where \( x_i \) and \( y_i \) are the mole fractions of component \( i \) in liquid and vapour phases, respectively. \( p_i^{sat} \) is the saturation pressure of component \( i \) inside the system at any arbitrary temperature and total pressure \( p_{tot} \) and \( f_i^s \) is the fugacity of component \( i \) in the system. \( \varphi_i \) is the fugacity coefficient of component \( i \) in the vapour phase, which can be equal to unity for low to moderate pressures. The gaseous mixture is assumed ideal (Lewis-Randall rule) that is possible for low and moderate pressures. According to equation (4-4), the \( \gamma_i^α \) and \( \gamma_i^β \) are activity coefficients of component \( i \) in two liquid phases of \( α \) and \( β \), respectively. Each of equations 4-2 to 4-4 is written for each phase equilibrium calculation in conjunction with the mass balance for the whole system. The resulting model equations will enable us to solve the mole and phase fractions in the system under study.

In order to predict the solid-liquid equilibrium in multi-solvent systems, the activity coefficient of solute in solution has to be known. The following equation gives the mole fraction of solute at equilibrium condition [16]:

\[ \ln x_s = - \frac{\Delta H_{fus}}{R} \left( \frac{1}{T_m} - \frac{1}{T} \right) - \ln \gamma_s \quad (4 - 5) \]

Where \( x_s \) refers to the solute mole fraction, \( \Delta H_{fus} \) is the heat of fusion, and \( T_m \) is the melting temperature of the solid at normal conditions, which can be measured using characterization
methods (DSC and TGA) [13]. The $\gamma_s$ is evaluated by activity coefficient thermodynamic models such as NRTL-SAC model.

### 4.2.2. Optimization method and procedure

In order to examine the best solvent or mixture of solvents for the crystallization process of a certain component, a new optimization algorithm was developed with the following objective functions:

$$J_1(T, y) = S_{\text{initial}} - S_{\text{final}} = \frac{M_{\text{initial}} - M_{\text{final}}}{M_s} \quad (4 - 6)$$

$$J_2(T, y) = \frac{S_{\text{initial}} - S_{\text{final}}}{S_{\text{initial}}} \quad (4 - 7)$$

Where $J_1(T, y)$ is the yield of the solute precipitation based on the mass of solvent (or mixtures of solvents). The solubility of the model molecules ($S_{\text{initial}}$ or $S_{\text{final}}$) are in mass of solute per mass of solute-free solvent. This objective is a function of solubility at both initial and final states of crystallization and the mass of solvent mixture, $M_s$. In equation (4-6) the $M_{\text{initial}}$ and $M_{\text{final}}$ denote the initial and final mass of dissolved solute in the solvent of mass $M_s$. $J_2(T, y)$ is the crystallization yield based on the initial mass of solute. The main difference between the two objective functions is their dependency on solvent mass which is used in crystallization. For each of the above objective functions, there are many constraints which should be met in order to run the crystallization process safely and reliably. The optimization procedure with detail formulations in single, binary, and ternary systems are demonstrated in appendix A.

The solvents which were used in the current study with their segment numbers, Antoine coefficients, and normal boiling points are listed in Table 4-1. There is also a column which shows the class number of each solvent used in our study, based on the recommendation of U.S. Food and Drug Administration (FDA) for industries [18]. The solvents are divided to three classes, with the less recommended solvent for crystallization as class 1 and the favourable solvent as class 3. As an example, benzene is listed as class 1 due to its carcinogenic nature. There are many solvents in classes 2 and 3, which can be used in the pharmaceutical industries, but caution should be taken in separation and purification processes in the downstream of crystallization process.
Table 4-1. Segment numbers, Antoine coefficients of vapour pressure, and normal boiling points of solvents studied in the current work [19] with the class number assigned to each solvent by Food and Drug Administration (FDA) [18]

<table>
<thead>
<tr>
<th>Solvent</th>
<th>X</th>
<th>Y-</th>
<th>Y+</th>
<th>Z</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>Tₙ [K]</th>
<th>Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>acetic acid</td>
<td>0.048</td>
<td>0.222</td>
<td>0.195</td>
<td>0.026</td>
<td>4.682</td>
<td>1642.540</td>
<td>-39.764</td>
<td>390.58</td>
<td>3</td>
</tr>
<tr>
<td>acetone</td>
<td>0.131</td>
<td>0.109</td>
<td>0.513</td>
<td>0.000</td>
<td>4.424</td>
<td>1312.253</td>
<td>-32.445</td>
<td>329.03</td>
<td>3</td>
</tr>
<tr>
<td>acetonitrile</td>
<td>0.018</td>
<td>0.131</td>
<td>0.883</td>
<td>0.000</td>
<td>4.279</td>
<td>1355.374</td>
<td>-37.853</td>
<td>354.62</td>
<td>2</td>
</tr>
<tr>
<td>anisole</td>
<td>0.536</td>
<td>0.010</td>
<td>0.653</td>
<td>0.000</td>
<td>4.177</td>
<td>1489.756</td>
<td>-69.607</td>
<td>426.24</td>
<td>3</td>
</tr>
<tr>
<td>benzene</td>
<td>0.615</td>
<td>0.000</td>
<td>0.281</td>
<td>0.000</td>
<td>4.018</td>
<td>1203.835</td>
<td>-53.226</td>
<td>352.83</td>
<td>1</td>
</tr>
<tr>
<td>1-butanol</td>
<td>0.425</td>
<td>0.004</td>
<td>0.000</td>
<td>0.490</td>
<td>4.546</td>
<td>1440.231</td>
<td>-61.362</td>
<td>398.81</td>
<td>3</td>
</tr>
<tr>
<td>2-butanol</td>
<td>0.343</td>
<td>0.069</td>
<td>0.000</td>
<td>0.393</td>
<td>4.329</td>
<td>1158.672</td>
<td>-104.493</td>
<td>372.31</td>
<td>3</td>
</tr>
<tr>
<td>n-butyl acetate</td>
<td>0.317</td>
<td>0.030</td>
<td>0.330</td>
<td>0.000</td>
<td>4.268</td>
<td>1440.231</td>
<td>-61.362</td>
<td>398.81</td>
<td>3</td>
</tr>
<tr>
<td>MTBE</td>
<td>0.483</td>
<td>0.105</td>
<td>0.142</td>
<td>0.000</td>
<td>4.039</td>
<td>1149.261</td>
<td>-43.150</td>
<td>327.71</td>
<td>3</td>
</tr>
<tr>
<td>carbon tetrachloride</td>
<td>0.739</td>
<td>0.027</td>
<td>0.142</td>
<td>0.000</td>
<td>4.023</td>
<td>1221.781</td>
<td>-45.739</td>
<td>349.44</td>
<td>1</td>
</tr>
<tr>
<td>chlorobenzene</td>
<td>0.727</td>
<td>0.024</td>
<td>0.484</td>
<td>0.000</td>
<td>4.111</td>
<td>1435.675</td>
<td>-55.124</td>
<td>404.37</td>
<td>2</td>
</tr>
<tr>
<td>chloroform</td>
<td>0.393</td>
<td>0.000</td>
<td>0.167</td>
<td>0.000</td>
<td>4.208</td>
<td>1233.129</td>
<td>-42.550</td>
<td>353.20</td>
<td>1</td>
</tr>
<tr>
<td>cumene</td>
<td>1.161</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>4.054</td>
<td>1455.811</td>
<td>-65.981</td>
<td>425.04</td>
<td>3</td>
</tr>
<tr>
<td>cyclohexane</td>
<td>0.892</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>4.140</td>
<td>1316.554</td>
<td>-35.581</td>
<td>353.60</td>
<td>2</td>
</tr>
<tr>
<td>1,2-dichloroethane</td>
<td>0.394</td>
<td>0.000</td>
<td>0.691</td>
<td>0.000</td>
<td>4.585</td>
<td>1521.789</td>
<td>-24.670</td>
<td>356.56</td>
<td>1</td>
</tr>
<tr>
<td>1,1-dichloroethylene</td>
<td>0.529</td>
<td>0.000</td>
<td>0.208</td>
<td>0.000</td>
<td>4.097</td>
<td>1099.400</td>
<td>-35.950</td>
<td>304.29</td>
<td>1</td>
</tr>
<tr>
<td>1,2-dichloroethylene</td>
<td>0.188</td>
<td>0.000</td>
<td>0.832</td>
<td>0.000</td>
<td>4.147</td>
<td>1205.400</td>
<td>-42.550</td>
<td>333.20</td>
<td>2</td>
</tr>
<tr>
<td>dichloromethane</td>
<td>0.459</td>
<td>0.000</td>
<td>0.427</td>
<td>0.038</td>
<td>4.537</td>
<td>1327.016</td>
<td>-20.474</td>
<td>312.97</td>
<td>2</td>
</tr>
<tr>
<td>1,2-dimethoxymethane</td>
<td>0.277</td>
<td>0.030</td>
<td>0.077</td>
<td>0.057</td>
<td>4.037</td>
<td>1068.350</td>
<td>-50.409</td>
<td>315.06</td>
<td>2</td>
</tr>
<tr>
<td>N,N-dimethylacetamide</td>
<td>0.160</td>
<td>0.778</td>
<td>0.193</td>
<td>0.000</td>
<td>6.095</td>
<td>2725.960</td>
<td>28.209</td>
<td>419.07</td>
<td>2</td>
</tr>
<tr>
<td>N,N-dimethylformamide</td>
<td>0.180</td>
<td>0.752</td>
<td>0.254</td>
<td>0.000</td>
<td>3.931</td>
<td>1337.716</td>
<td>-82.648</td>
<td>422.97</td>
<td>2</td>
</tr>
<tr>
<td>dimethyl sulfoxide</td>
<td>0.000</td>
<td>1.114</td>
<td>0.000</td>
<td>0.000</td>
<td>4.491</td>
<td>1807.002</td>
<td>-60.995</td>
<td>463.35</td>
<td>3</td>
</tr>
<tr>
<td>1,4-dioxane</td>
<td>0.154</td>
<td>0.086</td>
<td>0.401</td>
<td>0.000</td>
<td>4.581</td>
<td>1570.093</td>
<td>-31.297</td>
<td>374.01</td>
<td>2</td>
</tr>
<tr>
<td>ethanol</td>
<td>0.251</td>
<td>0.030</td>
<td>0.000</td>
<td>0.630</td>
<td>5.247</td>
<td>1598.673</td>
<td>-46.424</td>
<td>351.12</td>
<td>3</td>
</tr>
<tr>
<td>ethyl acetate</td>
<td>0.339</td>
<td>0.058</td>
<td>0.441</td>
<td>0.000</td>
<td>4.228</td>
<td>1245.702</td>
<td>-55.189</td>
<td>349.81</td>
<td>3</td>
</tr>
<tr>
<td>ethylene glycol</td>
<td>0.000</td>
<td>0.043</td>
<td>0.000</td>
<td>0.852</td>
<td>5.337</td>
<td>2161.910</td>
<td>-64.720</td>
<td>469.80</td>
<td>2</td>
</tr>
<tr>
<td>diethyl ether</td>
<td>0.387</td>
<td>0.028</td>
<td>0.177</td>
<td>0.000</td>
<td>4.022</td>
<td>1062.640</td>
<td>-44.930</td>
<td>309.14</td>
<td>3</td>
</tr>
<tr>
<td>ethyl formate</td>
<td>0.256</td>
<td>0.305</td>
<td>0.000</td>
<td>0.000</td>
<td>4.133</td>
<td>1123.943</td>
<td>54.903</td>
<td>326.86</td>
<td>3</td>
</tr>
<tr>
<td>formamide</td>
<td>0.000</td>
<td>0.089</td>
<td>0.341</td>
<td>0.252</td>
<td>7.585</td>
<td>3881.305</td>
<td>27.655</td>
<td>484.04</td>
<td>2</td>
</tr>
<tr>
<td>formic acid</td>
<td>0.000</td>
<td>0.090</td>
<td>0.000</td>
<td>0.420</td>
<td>2.001</td>
<td>515.000</td>
<td>-139.408</td>
<td>396.75</td>
<td>3</td>
</tr>
<tr>
<td>n-heptane</td>
<td>1.152</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>4.028</td>
<td>1268.636</td>
<td>-56.199</td>
<td>371.13</td>
<td>3</td>
</tr>
<tr>
<td>n-hexane</td>
<td>1.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>4.003</td>
<td>1171.530</td>
<td>-48.784</td>
<td>341.47</td>
<td>2</td>
</tr>
<tr>
<td>isobutyl acetate</td>
<td>0.620</td>
<td>0.183</td>
<td>0.541</td>
<td>0.000</td>
<td>4.537</td>
<td>1625.875</td>
<td>-32.494</td>
<td>390.87</td>
<td>3</td>
</tr>
<tr>
<td>isopropyl acetate</td>
<td>0.552</td>
<td>0.154</td>
<td>0.498</td>
<td>0.000</td>
<td>4.552</td>
<td>1490.877</td>
<td>-34.098</td>
<td>361.64</td>
<td>3</td>
</tr>
<tr>
<td>methanol</td>
<td>0.090</td>
<td>0.139</td>
<td>0.000</td>
<td>0.594</td>
<td>5.204</td>
<td>1581.341</td>
<td>-33.500</td>
<td>337.37</td>
<td>2</td>
</tr>
<tr>
<td>2-methoxyethanol</td>
<td>0.082</td>
<td>0.095</td>
<td>0.180</td>
<td>0.361</td>
<td>5.064</td>
<td>1853.556</td>
<td>-30.838</td>
<td>396.87</td>
<td>2</td>
</tr>
<tr>
<td>Solvent</td>
<td>X</td>
<td>Y^-</td>
<td>Y^+</td>
<td>Z</td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>T_b [K]</td>
<td>Class</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>-----</td>
<td>------</td>
<td>------</td>
<td>-----</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>---------</td>
<td>-------</td>
</tr>
<tr>
<td>methyl acetate</td>
<td>0.239</td>
<td>0.000</td>
<td>0.338</td>
<td>0.000</td>
<td>4.204</td>
<td>1164.426</td>
<td>-52.690</td>
<td>329.69</td>
<td>3</td>
</tr>
<tr>
<td>3-methyl 1-butanol</td>
<td>0.419</td>
<td>0.000</td>
<td>0.538</td>
<td>0.314</td>
<td>5.080</td>
<td>1932.043</td>
<td>-28.698</td>
<td>408.99</td>
<td>3</td>
</tr>
<tr>
<td>methyl butyl ketone</td>
<td>0.673</td>
<td>0.224</td>
<td>0.469</td>
<td>0.000</td>
<td>5.667</td>
<td>2011.668</td>
<td>-45.364</td>
<td>400.33</td>
<td>2</td>
</tr>
<tr>
<td>methyl cyclohexane</td>
<td>1.053</td>
<td>0.000</td>
<td>0.246</td>
<td>0.000</td>
<td>3.948</td>
<td>1270.763</td>
<td>-51.734</td>
<td>373.62</td>
<td>2</td>
</tr>
<tr>
<td>methyl ethyl ketone</td>
<td>0.261</td>
<td>0.095</td>
<td>0.463</td>
<td>0.000</td>
<td>3.989</td>
<td>1150.207</td>
<td>-63.904</td>
<td>352.22</td>
<td>3</td>
</tr>
<tr>
<td>methyl isobutyl ketone</td>
<td>0.673</td>
<td>0.224</td>
<td>0.469</td>
<td>0.000</td>
<td>3.953</td>
<td>1254.095</td>
<td>-71.537</td>
<td>388.79</td>
<td>3</td>
</tr>
<tr>
<td>isobutanol</td>
<td>0.566</td>
<td>0.000</td>
<td>0.067</td>
<td>0.485</td>
<td>4.431</td>
<td>1236.991</td>
<td>-101.528</td>
<td>380.68</td>
<td>NA</td>
</tr>
<tr>
<td>N-methyl-2-pyrrolidone</td>
<td>0.252</td>
<td>0.790</td>
<td>0.281</td>
<td>0.000</td>
<td>12.657</td>
<td>4112.280</td>
<td>-66.866</td>
<td>391.76</td>
<td>2</td>
</tr>
<tr>
<td>nitomethane</td>
<td>0.122</td>
<td>0.000</td>
<td>1.032</td>
<td>0.051</td>
<td>4.405</td>
<td>1446.196</td>
<td>-45.633</td>
<td>373.91</td>
<td>2</td>
</tr>
<tr>
<td>n-pentane</td>
<td>0.898</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>3.989</td>
<td>1070.617</td>
<td>-40.454</td>
<td>308.83</td>
<td>3</td>
</tr>
<tr>
<td>1-pentanol</td>
<td>0.458</td>
<td>0.024</td>
<td>0.000</td>
<td>0.491</td>
<td>4.324</td>
<td>1297.689</td>
<td>-110.669</td>
<td>410.77</td>
<td>3</td>
</tr>
<tr>
<td>1-propanol</td>
<td>0.374</td>
<td>0.013</td>
<td>0.000</td>
<td>0.530</td>
<td>5.314</td>
<td>1690.864</td>
<td>-51.804</td>
<td>370.00</td>
<td>3</td>
</tr>
<tr>
<td>isopropyl alcohol (IPA)</td>
<td>0.332</td>
<td>0.000</td>
<td>0.000</td>
<td>0.636</td>
<td>4.861</td>
<td>1357.427</td>
<td>-75.814</td>
<td>355.06</td>
<td>3</td>
</tr>
<tr>
<td>n-propyl acetate</td>
<td>0.514</td>
<td>0.134</td>
<td>0.587</td>
<td>0.000</td>
<td>4.144</td>
<td>1283.861</td>
<td>-64.378</td>
<td>374.20</td>
<td>3</td>
</tr>
<tr>
<td>pyridine</td>
<td>0.135</td>
<td>0.000</td>
<td>0.305</td>
<td>0.000</td>
<td>4.163</td>
<td>1371.358</td>
<td>-58.496</td>
<td>387.93</td>
<td>2</td>
</tr>
<tr>
<td>sulfolane</td>
<td>0.209</td>
<td>0.089</td>
<td>0.000</td>
<td>0.249</td>
<td>4.533</td>
<td>2255.469</td>
<td>-61.757</td>
<td>559.33</td>
<td>2</td>
</tr>
<tr>
<td>tetrahydrofurane</td>
<td>0.235</td>
<td>0.040</td>
<td>0.320</td>
<td>0.708</td>
<td>4.121</td>
<td>1202.942</td>
<td>-46.818</td>
<td>338.71</td>
<td>2</td>
</tr>
<tr>
<td>tetrahydroxynaphthalene</td>
<td>0.924</td>
<td>0.000</td>
<td>0.865</td>
<td>0.000</td>
<td>4.127</td>
<td>1690.912</td>
<td>-70.229</td>
<td>479.98</td>
<td>2</td>
</tr>
<tr>
<td>toluene</td>
<td>0.604</td>
<td>0.000</td>
<td>0.304</td>
<td>0.000</td>
<td>4.142</td>
<td>1377.578</td>
<td>-50.507</td>
<td>383.13</td>
<td>2</td>
</tr>
<tr>
<td>1,1,1-trichloroethane</td>
<td>0.548</td>
<td>0.000</td>
<td>0.287</td>
<td>0.000</td>
<td>5.886</td>
<td>2210.179</td>
<td>34.902</td>
<td>340.59</td>
<td>1</td>
</tr>
<tr>
<td>trichloroethylene</td>
<td>0.552</td>
<td>0.000</td>
<td>0.262</td>
<td>0.000</td>
<td>3.553</td>
<td>974.538</td>
<td>-85.811</td>
<td>360.06</td>
<td>2</td>
</tr>
<tr>
<td>m-xylene</td>
<td>0.758</td>
<td>0.021</td>
<td>0.316</td>
<td>0.000</td>
<td>4.134</td>
<td>1462.266</td>
<td>-58.045</td>
<td>411.76</td>
<td>2</td>
</tr>
<tr>
<td>water</td>
<td>0.000</td>
<td>0.000</td>
<td>1.000</td>
<td>0.000</td>
<td>6.210</td>
<td>2354.731</td>
<td>7.559</td>
<td>371.65</td>
<td>NA</td>
</tr>
<tr>
<td>trimethylamine</td>
<td>0.403</td>
<td>0.030</td>
<td>0.000</td>
<td>0.000</td>
<td>4.016</td>
<td>970.297</td>
<td>-34.060</td>
<td>275.66</td>
<td>NA</td>
</tr>
<tr>
<td>1-octanol</td>
<td>0.867</td>
<td>0.000</td>
<td>0.000</td>
<td>0.534</td>
<td>6.477</td>
<td>2603.359</td>
<td>-48.799</td>
<td>450.75</td>
<td>NA</td>
</tr>
<tr>
<td>n-octane</td>
<td>1.253</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>4.049</td>
<td>1355.126</td>
<td>-63.633</td>
<td>398.34</td>
<td>NA</td>
</tr>
</tbody>
</table>

### 4.2.3. Mathematical formulation of the model

In this section, the mathematical formulation of the optimization procedure and the algorithm to find the constituents mole fractions for single, binary, and ternary solvent mixture is demonstrated.

**Single solvent**
For safety reasons, we enforce the program to choose the upper limit of crystallization temperature to be less than the melting point of the solid. Therefore, the mathematical representation of the problem for single solvent optimization is:

\[
\min J_u(T) \quad u = 1, 2 \quad (4-8)
\]

\[T_{\text{initial}}, T_{\text{final}}\]

Subject to

\[
\{T_{\text{initial}}, T_{\text{final}}\} - 5 < T_{\text{melting}}
\]

\[
\{T_{\text{initial}}, T_{\text{final}}\} - 10 < T_{\text{solvent}^\text{sat}}
\]

\[T_{\text{initial}} > T_{\text{final}}\]

\[T_{\text{initial}}\] and \[T_{\text{final}}\] are the initial and final temperatures of the crystallization operation. The vapour pressure of a pure component (Antoine equation) is found from:

\[
\log_{10}(p_i^\text{sat}) = A_i - \frac{B_i}{T_{\text{sat}} + C_i} \quad i = 1, 2, \ldots, N \quad \text{P [bar]}, T [K] \quad (4 - 9)
\]

The values of constants in equation (4-2A) for all the solvents were found from National Institute of Standards and Technology source [18] and are listed in Table 4-1.

**Binary solvents**

For the case of binary solvent calculation, in addition to the constraints that are implemented in the single solvent case, there are additional nonlinear constraints. The initial and final operation temperatures are calculated based on the bubble point and freezing point of the solvent mixture. According to the results by Sheikholeslamzadeh and Rohani the NRTL-SAC model can be best used to predict the VLE behaviour of multicomponent systems [14]. Therefore, we adopted the NRTL-SAC model to predict the bubble point temperature of the solvent mixture besides the SLE calculations. The VLE calculations for binary and multi-solvent systems were coded using methods described elsewhere [16]. In addition to the initial and final temperatures as optimization variables, initial and final mole fractions of solvents (or their volume ratio) are also included in the optimization procedure. Therefore, we need to have a framework to relate the mole fractions of solvents and solute together. For a system of N solvents and a solute, the procedure for finding the mole fractions of the constituents of the solution is a trial-and-error
method. If we assume the mole fraction of a solute and solvent \( z = 1, \ldots, N \) in the solution is \( x_{\text{solute}} \) and \( x_z^s \), respectively, then the ratio of mole fractions of two solvents is:

\[
R_1 = \frac{x_1^s}{x_2^s}, \quad R_2 = \frac{x_2^s}{x_3^s}, \ldots, \quad R_{N-1} = \frac{x_{N-1}^s}{x_N^s}
\]  

(4-10)

And the following relations are obtained for each solvent’s mole fraction:

\[
x_1^s = R_1 x_2^s = R_1 R_2 R_3 \ldots R_{N-1} x_N^s
\]  

(4-11)

\[
x_2^s = R_2 x_3^s = R_2 R_3 \ldots R_{N-1} x_N^s
\]  

(4-12)

\[x_N^s = R_{N-1} x_N^s
\]

(4-13)

The sum of the mole fractions of all the components is unity, therefore:

\[
x_1^s + x_2^s + x_3^s + \ldots + x_{N-1}^s + x_N^s + x_{\text{solute}} = 1
\]  

(4-14)

With the substitution of equations 4-11 to 4-13 in equation 4-14, we get the following:

\[
x_N^s (R_1 R_2 \ldots R_{N-1} + R_2 R_3 \ldots R_{N-1} + \ldots + R_{N-1} + 1) + x_{\text{solute}} = 1
\]  

(4-15)

Or in compact form:

\[
x_N^s \sum_{i=1}^{N-1} \prod_{j=i}^{N-1} R_j + x_N^s + x_{\text{solute}} = 1
\]  

(4-16)

First the solute mole fraction is guessed and with the known values of solvent molar ratios (or volume ratios) the mole fraction of the last solvent is found. The other solvent mole fractions will be found using equations 4-11 to 4-13. The trial and error calculation is repeated until the error between two consecutive mole fractions for each component in the solution is minimized. The procedure to find the solubility of a solute in a system of \( N \) solvents is described elsewhere [15]. This procedure is inside the optimization loop used in the present work. The optimization problem for binary solvent case is:
\[
\begin{align*}
\min J_u(T, R) & \quad u = 1, 2 \\
T_{\text{initial}}, T_{\text{final}}, R_{\text{initial}}, R_{\text{final}}
\end{align*}
\]

Subject to
\[
\begin{align*}
30^\circ C & < T_{\text{initial}} < \min\{\left(T_{\text{melting}} - 5\right), 65^\circ C\} \\
15^\circ C & < T_{\text{final}} < \min\{\left(T_{\text{melting}} - 5\right), 20^\circ C\} \\
0 & \leq \left\{\frac{x_{1,0}^s}{x_{2,0}^s}, \frac{x_{1,f}^s}{x_{2,f}^s}\right\} \leq 1000
\end{align*}
\]

\[
T_{\text{final}} < T_{\text{initial}}
\]

\[
T_{\text{initial}} < T^{\text{sat}}_i (P_{\text{tot}}, x_{1,0}^s, x_{2,0}^s) - 10
\]

\[
T_{\text{final}} < T^{\text{sat}}_i (P_{\text{tot}}, x_{1,f}^s, x_{2,f}^s) - 10
\]

Where \(x_{1,0}^s\) and \(x_{2,0}^s\) are initial solvent mole fractions which are functions of temperature. \(x_{1,f}^s\) and \(x_{2,f}^s\) are the final solvent mole fractions, which are functions of final temperature of the crystallizer. \(T^{\text{sat}}\) is the bubble temperature of the system of solvents at total pressure of \(P_{\text{tot}}\) and mole fractions of components. Therefore it can be seen that another trail-and-error procedure should be performed with the optimization for each combination of solvents, simultaneously. This procedure will be repeated for each possible combination of solvents and finally the values of optimization will be compared and the maximum yield will be picked as the optimal solvent combination. If each combination of solvents is denoted by \(S_{m,n}\) in which \(m\) and \(n\) are solvent IDs (which was defined for all the solvents in the program environment), then the optimized value of operating conditions for each combination will be contained in each element of the following matrix (matrix in matrix):

\[
\text{OM}_u = \begin{bmatrix}
0 & \begin{bmatrix} T_{\text{initial}} & T_{\text{final}} \\ R_{\text{initial}} & R_{\text{final}} \end{bmatrix}_{1,2} & \cdots & \begin{bmatrix} T_{\text{initial}} & T_{\text{final}} \\ R_{\text{initial}} & R_{\text{final}} \end{bmatrix}_{1,n} \\
0 & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \begin{bmatrix} T_{\text{initial}} & T_{\text{final}} \\ R_{\text{initial}} & R_{\text{final}} \end{bmatrix}_{m,n}
\end{bmatrix}
\]

\[(4-18)\]
The OM is the optimization matrix that is made by optimizing combination of solvents. \(OY_u\) is the matrix of objective functions for \(u = 1,2\). It should be noted that both matrices are upper diagonal and the elements of the same solvent are zero. In this study the total number of non-zero elements of the matrix for the binary solvent was 1891 (which is calculated from \(\frac{62 \times 61}{2}\)).

**Ternary and quaternary solvents**

When the number of solvents in a crystallization process gets larger, the number of combinations will grow largely. As an example, if one has to examine all the possible combinations of a ternary solvent mixture from a list of 62 solvents, the total number of 37820 (\(\frac{62 \times 61 \times 60}{3!}\)) cases should be considered for optimization. For each case the optimization procedure consists of linear and nonlinear constraints in addition to VLE and SLE calculations. This procedure is computationally very intensive. Therefore, we selected a subset from all 62 solvents for each model molecule to reduce the CPU time. The selection of solvents for ternary calculation was based on the optimal group of solvents which were found for binary combinations. Additional solvents which have different segment numbers from the selected binary solvents were also used for ternary calculation. The additional solvents were selected from the solvents with similar segment numbers and a less hazardous class of solvents.

For the quaternary solvent optimization, the computational time for all the possible combinations is much higher (there is a total number of 557845 cases for quaternary system). In the case of three or four solvent mixtures, the mathematical formulation for optimization is almost the same as previously described. The difference is in the number of mole fractions present in calculation of VLE and SLE. Also the nonlinear constraints of the framework will be changed according to the number of components present in the system. Therefore, the optimization formulation for a system of \(N\) solvents and a solute can be given by:

\[
\min J_u(T, R) \quad u = 1, 2 \quad (4 - 20)
\]

\(T_{\text{initial}}, T_{\text{final}}, R\)
Subject to

\[ 30^\circ \text{C} < T_{\text{initial}} < \min\{ (T_{\text{melting}} - 5), 65^\circ \text{C} \} \]
\[ 15^\circ \text{C} < T_{\text{final}} < \min\{ (T_{\text{melting}} - 5), 20^\circ \text{C} \} \]

\[ 0 \leq \left\{ \frac{x_{1,0}^s}{x_{1,f}^s}, \frac{x_{1,0}^s}{x_{2,0}^s}, \ldots, \frac{x_{N-1,0}^s}{x_{N-1,f}^s}, \frac{x_{N-1,0}^s}{x_{N,f}^s} \right\} \leq 1000 \]

\[ T_{\text{final}} < T_{\text{initial}} \]

\[ T_{\text{initial}} < T_{\text{sat}}(P_{\text{tot}}, x_{1,0}^s, x_{2,0}^s, \ldots, x_{N,0}^s) - 10 \]

\[ T_{\text{final}} < T_{\text{sat}}(P_{\text{tot}}, x_{1,f}^s, x_{2,f}^s, \ldots, x_{N,f}^s) - 10 \]

In which \( R = (R_1, R_2, \ldots, R_N) \) and accounts for initial and final ratio of solvents. The above operation will be performed for all of the combinations of solvents of interest. The mole fractions of solvents in the nonlinear constraint are found from the following:

\[ x_1^s = \left[ \frac{R_1 R_2 \ldots R_{N-1}}{R_1 R_2 \ldots R_{N-1} + \ldots + R_{N-1} + 1} \right] \quad (4 - 21) \]

\[ x_2^s = \left[ \frac{R_2 R_3 \ldots R_{N-1}}{R_2 R_3 \ldots R_{N-1} + \ldots + R_{N-1} + 1} \right] \]

\[ \vdots \]

\[ x_{N-1}^s = \left[ \frac{R_{N-1}}{R_2 R_3 \ldots R_{N-1} + \ldots + R_{N-1} + 1} \right] \]

For each optimization, the above mole fractions will be initialized and a trial-and-error procedure is performed to stabilize the final results for each multi-solvent system and its corresponding SLE containing the solute. If we assume the case of ternary solvent mixture, then we will have a three-dimensional upper diagonal matrix which contains all the operating points of each ternary solvent combination at optimal condition. Each non-zero element represents a \( 2 \times 3 \) matrix that consists of the operating points. As it is shown in equation 4-15A, \( \mathbf{OM}_k \) is the matrix which corresponds to all binary combinations with the \( k^{th} \) solvent (the third solvent in a ternary solvent system):
\[
\begin{align*}
OM_k &= \begin{bmatrix}
0 & \frac{T_{\text{initial}}}{R_{\text{initial}}} & \frac{T_{\text{final}}}{R_{\text{final}}} & \cdots & \frac{T_{\text{initial}}}{R_{\text{initial}}} & \frac{T_{\text{final}}}{R_{\text{final}}} \\
0 & \frac{R_{\text{initial}}}{R_{\text{final}}} & \frac{R_{\text{final}}}{R_{\text{final}}} & \cdots & \frac{R_{\text{initial}}}{R_{\text{final}}} & \frac{R_{\text{final}}}{R_{\text{final}}} \\
0 & 1 & 1 & \cdots & 1 & 1 \\
0 & 0 & 0 & \cdots & 0 & 0
\end{bmatrix}
\end{align*}
\]

\[k = 1, 2, \ldots, N \quad (4-22)\]

\[
OY_{k,u} = \begin{bmatrix}
0 & S_{1,2,k} & \cdots & S_{1,n,k} \\
0 & \ddots & \vdots & \ddots \\
0 & \cdots & S_{m,n,k}
\end{bmatrix}
\]

\[u = 1, 2 \quad (4-23)\]

\(N\) is the total number of solvents that are added to the binary combination. \(R_{\text{initial}}^i\) and \(R_{\text{final}}^i\) are the initial and final ratios of the \(i\)th solvent ratio \(\left(\frac{x_i}{x_{i+1}}\right)\). For the quaternary system of solvents the condition will be more complex, as the matrix of operating conditions will be four dimensional. For this case, each element inside \(OM_{k,\theta}\) is a three-dimensional sub-matrix which contains just the values of three-solvent-combinations for the first solvent picked in four-dimensional matrix. The matrix is shown in equation 4-24.

\[
\begin{align*}
OM_{k,\theta} &= \begin{bmatrix}
0 & \frac{T_{\text{initial}}}{R_{\text{initial}}} & \frac{T_{\text{final}}}{R_{\text{final}}} & \cdots & \frac{T_{\text{initial}}}{R_{\text{initial}}} & \frac{T_{\text{final}}}{R_{\text{final}}} \\
0 & \frac{R_{\text{initial}}}{R_{\text{final}}} & \frac{R_{\text{final}}}{R_{\text{final}}} & \cdots & \frac{R_{\text{initial}}}{R_{\text{final}}} & \frac{R_{\text{final}}}{R_{\text{final}}} \\
0 & 1 & 1 & \cdots & 1 & 1 \\
0 & 0 & 0 & \cdots & 0 & 0
\end{bmatrix}
\end{align*}
\]

\[(4 - 24)\]
4.3. Results and discussion

The selected model molecules are listed in Table 4-2 [12, 13-15]. The physical properties of model molecules with their four conceptual segment numbers which were used in our study are also shown in Table 4-2. First, the result of optimization for single solvent case will be discussed. Then, binary and ternary systems of solvents will be presented. As it was mentioned earlier, the computation time for finding the best solvent mixture in binary and higher number of solvent mixtures is prohibitively high. Therefore, the quaternary case is not presented in the current study since the marginal improvement in the objective functions was found to be negligible despite the enormous increase in the computational time. The chemical structures of the model molecules used in this study are shown in Figure 4-2. For all of the computation works, a computer with 2.67GHz Core (2) Duo processor was used. The computations were primarily conducted in Matlab environment. It should be noted that the molar ratios which are mentioned later are the division of the two solvents’ molar contents in a mixture (equation 5-3A). This has not to be treated as the mole fractions of solvents in the mixture.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Lovastatin [13]</td>
<td>445.50</td>
<td>43.14</td>
<td>96.84</td>
<td>1.175</td>
<td>0.000</td>
<td>0.548</td>
<td>0.882</td>
</tr>
<tr>
<td>Valsartan [13]</td>
<td>380.65</td>
<td>31.65</td>
<td>83.15</td>
<td>0.000</td>
<td>0.946</td>
<td>0.000</td>
<td>0.539</td>
</tr>
<tr>
<td>Paracetamol [12]</td>
<td>443.20</td>
<td>27.60</td>
<td>62.28</td>
<td>0.416</td>
<td>0.016</td>
<td>0.168</td>
<td>1.861</td>
</tr>
<tr>
<td>Budesonide [15]</td>
<td>534.00</td>
<td>34.70</td>
<td>64.98</td>
<td>1.000</td>
<td>0.178</td>
<td>0.005</td>
<td>1.079</td>
</tr>
<tr>
<td>Allopurinol [15]</td>
<td>653.50</td>
<td>38.50</td>
<td>58.91</td>
<td>0.016</td>
<td>0.002</td>
<td>1.169</td>
<td>0.000</td>
</tr>
<tr>
<td>Furosemide [15]</td>
<td>534.30</td>
<td>48.70</td>
<td>91.15</td>
<td>0.600</td>
<td>0.127</td>
<td>0.010</td>
<td>1.620</td>
</tr>
<tr>
<td>Sulfadiazine [12]</td>
<td>538.80</td>
<td>31.20</td>
<td>57.91</td>
<td>0.757</td>
<td>0.000</td>
<td>0.000</td>
<td>1.940</td>
</tr>
</tbody>
</table>
Figure 4-1. Chemical structure of model molecules used in current study, a: lovastatin, b: valsartan, c: paracetamol, d: budesonide, e: allopurinol, f: furosemide, g: sulfadiazine

4.3.1. Single solvent screening

According to the description in previous parts, the study on solvent screening was conducted using single, binary and ternary solvent mixtures. The first category deals with solvent selection from single solvents. The selections were made among 62 solvents commonly used in research centers and industries [2]. For each of the objective functions that was defined in previous
sections, the optimization was performed to find the suitable solvent for crystallization process. In Table 4-3 the optimal solvent for the single solvent crystallization process of each component is given. It should be noted that the initial and final temperatures for all of the simulations in single solvent case were chosen as 40°C and 20°C, respectively. The operating temperatures for the process were selected to satisfy the safety of the crystallization process.

Table 4-3. The optimal single solvent for each of model components with their starting and ending operating temperature at 40°C and 20°C, respectively

<table>
<thead>
<tr>
<th>Model molecule</th>
<th>Based on $J_1$</th>
<th>Based on $J_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Solvent</td>
<td>Yield ($\frac{g}{g\text{ solvent}}$)</td>
</tr>
<tr>
<td>Lovastatin</td>
<td>DMF$^a$</td>
<td>0.14</td>
</tr>
<tr>
<td>Valsartan</td>
<td>Water</td>
<td>3.01</td>
</tr>
<tr>
<td>Paracetamol</td>
<td>Methanol</td>
<td>0.23</td>
</tr>
<tr>
<td>Budesonide</td>
<td>Ethyl formate</td>
<td>0.13</td>
</tr>
<tr>
<td>Allopurinol</td>
<td>Acetonitrile</td>
<td>$9.60 \times 10^{-4}$</td>
</tr>
<tr>
<td>Furosemide</td>
<td>Ethyl formate</td>
<td>0.03</td>
</tr>
<tr>
<td>Sulfadiazine</td>
<td>DMSO$^b$</td>
<td>0.16</td>
</tr>
</tbody>
</table>

$a$: N,N-dimethyl formamide, $b$: Dimethyl sulfoxide

It can be seen from Table 4-3 that for each pharmaceutical model molecule, there is a solvent which can produce the maximum yield based on one of the objective functions. It is worth noting that for all of 7 solutes, the solvent which yielded the highest value for objective $J_1$ was different from that of the other objective, $J_2$, which is because of the definition of two objective functions. For the first objective, in addition to maximizing the yield (amount of crystallized solid), the minimum consumption of solvent is also included. While the second objective function aims to maximize the product yield with respect to the initial solute used for crystallization process. This is more clearly illustrated in Figure 4-2. This figure shows the solubility curves of valsartan in pyridine and water. From Table 4-3 it is seen that water will maximize the first objective function, while pyridine optimizes the second objective function. In order to maximize the
second objective, the final solubility should be as low as possible, while the initial solubility has minor effect in optimizing this objective function. This shows that one may find some solvents with near-to-zero solubility at final stage of crystallization and hence, optimize the yield based on second objective function. However, for the first objective function, in addition of the final solubility, the initial solubility and solvent usage (mass) also play an important role. The final selection of the solvent can be better done by considering other parameters such as environmental and hazardous conditions made by using specific solvent. For example, the solvent cost is one of the important issues when dealing with the large scale production of pharmaceuticals or specialty chemicals. If this is the case, the first objective can help an engineer decide the favourable solvent for the process. On the other hand, if the price of solvent is not crucial and the chemical is much more expensive, the yield based on the second objective can help finding the best solvent.

Figure 4-2. The obtained solubility curves for valsartan in pyridine and water from NRTL-SAC model

In order to have a comparative study on the values of two objective functions for the crystallization of the model molecules in 62 solvents, we used a scatter chart as is shown in Figure 4-3, which is the single solvent solubility for sulfadiazine. The x-axis represents the solvent identification number (i.e., the ID for water is 59) and y-axes on the right and left
illustrate the yield based on the second and first objective functions, respectively. As it is evident from the figure, for all of the solvents, the second objective function is in the range of 45-62%, while the first objective function has a value of nearly zero for about 45 solvents. It means that despite having a reasonable value for the second objective function, many of the solvents could not give favourable results for the first objective function. Therefore, three solvent candidates with the most favourable yields for both objective functions can be selected as N,N-dimethylacetamide, DMF, and DMSO. From the three solvents, the DMSO can be the best solvent, with the yield for the first and second objective functions being equal to $\frac{0.16}{g \text{ solvent}}$ and 47.55%, respectively. However, if one based the selection on the first objective function only, the DMF would be considered as the optimal single solvent.

![Figure 4-3. Comparative illustration of yields for two objective functions for sulfadiazine](image)

4.3.2. Binary solvent screening

In the case of binary solvent screening, the first objective function was used for all possible combinations of solvents (which is 1891 binary combinations). For the binary solvent case,
always one can find a second solvent which leads to complete crystallization, which maximizes
the second objective function. The results of optimal process conditions at initial and final part of
the process when using the first objective function are given in Table 4-4. For each component
there are two solvents which will result in the maximum crystallization of solute per mass of
solvent. In the first column the two selected optimal solvents are shown with their molar and
volume ratios given in the third and fourth columns. The solvent selections and conditions for
each model molecule is discussed. The maximum yield of crystallization for all of the
components based on the first objective function and the corresponding value in terms of second
objective function is shown in Table 4-4. As it was mentioned, the second objective function can
be maximized with some different combinations of solvents, while the first objective is much
restrictive to solvent selections. It should be noted that the value of zero for molar or volume
ratio of solvents in Table 4-4 means that part of crystallization runs with the pure second solvent.
The same thing applies when the ratio is a large number that corresponds to negligible second
solvent. The optimal solvents with the operating conditions for crystallization of each model
component are discussed below.

Lovastatin. In this case, the initial and final temperatures exactly lie on the limits of constraints
for temperature. After optimizing all the possible solvent combinations, dimethyl sulfoxide
(DMSO) and ethyl formate were found as optimal solvent mixture. In Figure 4-4, the solubility
of lovastatin in the mixture of cumene and acetonitrile is demonstrated. As it is evident from the
results, for all the temperatures, the molar ratio of cumene:acetonitrile with the value of 0.35 will
yield the maximum solubility. For the case of crystallization with these solvents, the initial and
final state and temperature can be found from Figure 4-4. This procedure was conducted for all
the binary combinations of solvents to maximize the crystallization yield of lovastatin. From the
results shown in Table 4-4, the volume ratio of DMSO:ethyl formate at the beginning of the
crystallization is 0.47 and the final value is 32.43. This means the crystallization process should
be started with nearly half of DMSO with one unit of ethyl formate at 65°C to dissolve as much
lovastatin as possible. However, at the end of the process at 15°C, the DMSO has to be added to
the vessel to achieve the final ratio of solvents and as a result, maximizing the crystallization
yield. The temperature ranges are exactly the values which were initially set for the
crystallization without any change. This shows that the bubble point of the solvent mixture from
the start to final part of the process and melting point of solute are higher than the operating temperatures. The optimum value of yield for lovastatin is $1.08 \frac{g}{g \text{ solvent}}$.

**Figure 4-4.** Solubility of lovastatin in the mixture of cumene and acetonitrile in the temperature range of 300-340K.

**Valsartan.** The two optimal solvents for this component are n-heptane and water with the molar ratio of 0.00 and 28.57 at the beginning and end of crystallization, respectively. From Table 4-4, it can be seen that the volume ratio of n-heptane to water at both conditions shows that one of the solvents has a high solubility (water) and the other has a very poor solubility (an anti-solvent).

**Paracetamol.** For this model molecule, the initial and final states of solvent mole and volume fractions are at the extreme values, like the valsartan case. The optimum solvents for paracetamol are cumene and methanol. As the solubility of paracetamol in methanol is very high, specifically at higher temperatures, the initial condition of the crystallizer should be with pure methanol at the highest possible temperature. The final state is reached by adding cumene to achieve the volume ratio of 83.31 (cumene:methanol) to crystallize the solute. This high volume ratio of cumene to methanol shows the huge addition of cumene as an antisolvent. Also, the temperature should be selected as low as possible at the final stage.
**Budesonide.** The solvents suitable for the highest yield for this component are cyclohexane and ethyl formate. The procedure of the crystallization is the same as paracetamol, but with different solvent ratios at the initial and final stages of operation.

**Allopurinol.** This component generally has a very low solubility in most solvents. Because of having N-H bonds in its chemical structure, two solvents having the amine and nitrile group were found as optimum solvents for the crystallization process. Trimethylamine is the best one to dissolve allopurinol, while the acetonitrile at high concentration is a poor solvent. Therefore, at the initial stage of the process, a combination of both solvents was identified for the maximum dissolution. The mixture of two solvents has a high concentration of both nitrile and amine groups, which may be of help in attracting and dissolving the solid molecule. Other solvents were not able to dissolve this component compared with the mentioned two solvents. It is worth noting that the boiling point of trimethylamine and acetonitrile are very low and as a result, the initial temperature was adopted 39°C for the initial operation.

**Furosemide.** The optimal binary solvents for this component were found to be ethyl formate and formic acid as solvent and antisolvent, respectively. The results show that at the end of crystallization process, the volume fraction of solvent should reach to zero, as the antisolvent will crystallize almost all of the furosemide dissolved in the solution. The yield of this process is $\Delta_\text{u} \frac{g}{g_{\text{solvent}}}$ which is very high. Trying other combinations of solvents resulted in lower yields. On the other hand, as the first objective function was defined based on the moles of solute per mole of solvent mixture, the lower molecular weight of mixture of solvents corresponds to the less mass of solvents consumed. The molecular weight of formic acid and ethyl formate are low and hence, this makes crystallization operation economical.
Table 4-4. Optimal conditions of operation for cooling and antisolvent crystallization for the first objective function ($J_1$)

<table>
<thead>
<tr>
<th>Model component</th>
<th>Initial condition</th>
<th>Final condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lovastatin</td>
<td>Solvents (first: second)</td>
<td>Molar ratio ($r_1$)</td>
</tr>
<tr>
<td></td>
<td>DMSO</td>
<td>0.54</td>
</tr>
<tr>
<td></td>
<td>Ethyl formate</td>
<td></td>
</tr>
<tr>
<td>Valsartan</td>
<td>n-heptane</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>Water</td>
<td></td>
</tr>
<tr>
<td>Paracetamol</td>
<td>Cumene</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>Methanol</td>
<td></td>
</tr>
<tr>
<td>Budesonide</td>
<td>Cyclohexane</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>Ethyl formate</td>
<td></td>
</tr>
<tr>
<td>Allopurinol</td>
<td>Acetonitrile</td>
<td>4.52</td>
</tr>
<tr>
<td></td>
<td>Trimethyl amine</td>
<td></td>
</tr>
<tr>
<td>Furosemide</td>
<td>Ethyl formate</td>
<td>64.74</td>
</tr>
<tr>
<td></td>
<td>Formic acid</td>
<td></td>
</tr>
<tr>
<td>Sulfadiazine</td>
<td>Cyclohexane</td>
<td>0.08</td>
</tr>
<tr>
<td></td>
<td>DMSO</td>
<td></td>
</tr>
</tbody>
</table>
**Sulfadiazine.** In this case, DMSO is found to be the best solvent and cyclohexane as an antisolvent. The reason why DMSO was picked for the crystallization can be the sulphide groups which are similar in both solute and solvent. The cyclohexane is a relatively non-polar solvent which is a poor solvent for this component.

<table>
<thead>
<tr>
<th>Component</th>
<th>Yield based on $J_1$ (g solvent)</th>
<th>Yield based on $J_2$ ($\frac{g}{g \text{ initial solute}} \times 100$)</th>
<th>CPU time (hr)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lovastatin</td>
<td>1.08</td>
<td>96.58</td>
<td>14.79</td>
</tr>
<tr>
<td>Valsartan</td>
<td>14.24</td>
<td>100.00</td>
<td>2.94</td>
</tr>
<tr>
<td>Paracetamol</td>
<td>0.97</td>
<td>100.00</td>
<td>15.41</td>
</tr>
<tr>
<td>Budesonide</td>
<td>0.48</td>
<td>100.00</td>
<td>3.31</td>
</tr>
<tr>
<td>Allopurinol</td>
<td>$9.67 \times 10^{-4}$</td>
<td>95.17</td>
<td>0.42</td>
</tr>
<tr>
<td>Furosemide</td>
<td>0.13</td>
<td>62.14</td>
<td>1.00</td>
</tr>
<tr>
<td>Sulfadiazine</td>
<td>0.64</td>
<td>99.42</td>
<td>6.78</td>
</tr>
</tbody>
</table>

**Table 4-5.** Optimum yield of first objective function ($J_1$) for model components with its corresponding yield based on second objective and the CPU time allocated for calculation for binary solvent optimization.

### 4.3.3. Ternary solvent screening

As it was mentioned in previous sections, the computation time will increase rapidly as the number of solvents under study increases. From Table 4-5, it can be seen that the required time to optimize the 62 binary combinations of solvents for the model molecules vary from 1-25hr CPU time. On the other hand, careful study of all common 62 solvents reveals that around half of them have similar segment numbers. Therefore, 30 solvents from the pool of 62 solvents were chosen for ternary solvent optimization. The optimum solvents in binary case are also included in this subset. The optimum solvent mole and volume ratios are listed in Table 4-6 for each model molecule. It should be noted that $r_i$ and $r_i'$ refer to the ratios of mole or volume of solvent $i^{th}$ to $(i+1)^{th}$ in each part of crystallization. The results of optimization for the solvents in ternary
case for the first objective function are shown in Table 4-6. This table shows three solvent combinations for each model molecule with their initial mole and volume ratios at the beginning and end of crystallization.

**Lovastatin.** DMSO is selected as the first solvent, which is similar in the binary case as the initial solvent for crystallization. The other two solvents screened here are different from that of the binary case. However, the combination of the two other solvents (cyclohexane and trimethylamine) is in such a way that the yield is nearly the same in both conditions, which shows that DMSO has a major effect in crystallization of this component.

**Valsartan.** For this model molecule, n-heptane and water solvents are the same as in the binary case. Water is selected as the initial solvent for crystallization. The molar and volume ratios of two other solvents to water are zero. Ethyl formate is in very small amount with respect to n-heptane at the end of the process, which shows the ternary optimization of solvents is nearly the same as binary case.

**Paracetamol.** For this component, methanol is the major solvent initially (the same as binary case), however methanol content is nearly zero at the final part of crystallization. The final part of crystallization runs with the cumene as the antisolvent with a small amount of water.

**Budesonide.** As it is evident from Table 4-6, and using equation 4-4A to 4-6A (in appendix A), the mole fraction of ethyl formate and formic acid are 0.96 and 0.04, respectively. This shows that a minute amount of formic acid is added to ethyl format at initial part of crystallization. To increase the yield, formic acid should be added to the system in large amount as is shown in Table 4-6. Here, the temperature range of crystallization is also bounded nearly to the two limits. The only difference between this case and the binary case is that formic acid is used as an antisolvent. The reason for this is the presence of DMSO in the system, which interacts with other solvents to behave like the cyclohexane interaction in the binary case.

**Allopurinol.** The solvent combination for ternary system has a higher yield in first objective function (nearly 15% higher than the binary case). Acetonitrile and trimethylamine are selected in both binary and ternary systems, however cyclohexane was added at the end of the crystallization process to crystallize as much solute as possible in the ternary solvent case.
**Furosemide.** Ethyl formate is selected in ternary system in initial part of crystallization. While the final part of crystallization runs with 1,4-dioxane instead of formic acid in the binary solvent combination. It seems that the yield of the first objective function has changed slightly, while the second objective for the ternary optimization changed significantly.

**Sulfadiazine.** DMSO has poor dissolving power for sulfadiazine and that’s why it has been selected in both binary and ternary systems as a crystallizing anti-solvent agent at the final part of crystallization process. However, dichloromethane has also been selected in combination with DMSO in ternary system at the final part of the process. The initial part of crystallization runs with trimethylamine for the ternary case which has an excellent dissolving characteristic for sulfadiazine.

The results of yield optimization based on both objective functions are summarized in Table 4-7. There are some additional points which must be mentioned based on the calculations of the ternary solvent case:

- From Table 4-8, the yield based on optimization of the first objective function does not change significantly compared to the binary case. It is seen that the maximum yield for a specific component for binary and ternary and also higher combinations is nearly the same.
- For nearly all the seven model molecules, the optimum solvents from binary and ternary mixtures were nearly the same. The solvents which were picked up in the ternary case were almost from the list of solvents in binary case (Table 4-4).
- In order to find an optimal solvent or a solvent mixture for a model molecule, the selection should be made from the pool of binary solvents. As the cost of solvent recovery and filtration increases with the presence of more solvents. In addition, the yield change from the binary to ternary solvent systems is not significant.
<table>
<thead>
<tr>
<th>Model molecule</th>
<th>Solvents (first:second:third)</th>
<th>Initial condition</th>
<th>Final condition</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Molar ratio ($r_1,r_2$)</td>
<td>Volume ratio ($r_1',r_2'$)</td>
<td>Temperature (°C)</td>
</tr>
<tr>
<td>Lovastatin</td>
<td>DMSO</td>
<td>2.28, 3.29</td>
<td>1.50, 4.03</td>
</tr>
<tr>
<td>Cyclohexane</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trimethylamine</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Valsartan</td>
<td>Ethyl formate</td>
<td>0.00, 0.00</td>
<td>0.00, 0.00</td>
</tr>
<tr>
<td>n-heptane</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Water</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paracetamol</td>
<td>Water</td>
<td>7.66, 0.00</td>
<td>1.00, 0.00</td>
</tr>
<tr>
<td>Cumene</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Methanol</td>
<td>DMSO</td>
<td>0.00, 28.00</td>
<td>0.00, 59.93</td>
</tr>
<tr>
<td>Budesonide</td>
<td>Ethyl formate</td>
<td>0.00, 28.00</td>
<td>0.00, 59.93</td>
</tr>
<tr>
<td></td>
<td>Formic acid</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Allopurinol</td>
<td>Cyclohexane</td>
<td>0.00, 17.55</td>
<td>0.00, 10.38</td>
</tr>
<tr>
<td></td>
<td>Acetonitrile</td>
<td>0.00, 17.55</td>
<td>0.00, 10.38</td>
</tr>
<tr>
<td></td>
<td>Trimethylamine</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DMSO</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Furosemide</td>
<td>Ethyl formate</td>
<td>0.16, 101.30</td>
<td>0.12, 95.94</td>
</tr>
<tr>
<td></td>
<td>1,4-dioxane</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DMSO</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sulfadiazine</td>
<td>Trimethylamine</td>
<td>9.83, 344.48</td>
<td>7.91, 475.93</td>
</tr>
<tr>
<td></td>
<td>DMSO</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4-6. Optimal solvent mole ratios, volume ratios, and temperature for the start and final part of crystallization.
Table 4-7. Optimal values of yield of crystallization for ternary mixture of solvents based on the first objective function and the second objective function

<table>
<thead>
<tr>
<th>Component</th>
<th>Yield based on $J_1$ ($\frac{g}{g_{\text{solvent}}}$)</th>
<th>Yield based on $J_2$ ($\frac{g}{g_{\text{initial solute}}}$ $\times$ 100)</th>
<th>CPU time (hr)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lovastatin</td>
<td>1.10</td>
<td>99.96</td>
<td>46.92</td>
</tr>
<tr>
<td>Valsartan</td>
<td>14.25</td>
<td>100.00</td>
<td>12.34</td>
</tr>
<tr>
<td>Paracetamol</td>
<td>0.97</td>
<td>100.00</td>
<td>76.20</td>
</tr>
<tr>
<td>Budesonide</td>
<td>0.49</td>
<td>99.84</td>
<td>13.29</td>
</tr>
<tr>
<td>Allopurinol</td>
<td>0.0014</td>
<td>96.42</td>
<td>1.31</td>
</tr>
<tr>
<td>Furosemide</td>
<td>0.14</td>
<td>99.88</td>
<td>4.00</td>
</tr>
<tr>
<td>Sulfadiazine</td>
<td>0.66</td>
<td>99.98</td>
<td>42.78</td>
</tr>
</tbody>
</table>

4.4. Conclusion

In this work the optimal solvent screening for the crystallization of seven pharmaceutical molecules based on a modified optimization algorithm and the NRTL-SAC thermodynamic modeling of VLE and SLE of 62 common industrial solvents was successfully accomplished. The algorithm showed good predictive capability in VLE, LLE, and SLE systems. Two different objective functions were defined for the optimization. The results of single, binary, and ternary solvent screening calculations are shown in Table 8. It is evident that, in general, the best choice to maximize the crystallization yield of a batch cooling-antisolvent process is a binary solvent system. As an example, the crystallization yield of valsartan and paracetamol increased by 4 and 3 times with respect to the single optimal solvent, respectively. The additional computational and complexities created do not justify ternary solvent systems. The ternary solvent systems only affect the second objective function, which for some cases such as furosemide results in an improved yield. It should be noted that for some optimization cases, the high dilution condition was happened at either or both ends of crystallization, which shows the possibility of applying other process methods, such as evaporative crystallization. This work shows the capability of the thermodynamic model in conjunction with the proposed optimization procedure to yield the best combination of solvents for optimal performance of a crystallization process, specifically in a pharmaceutical industry.
Table 4-8. Optimal yield values for single, binary, and ternary solvent combinations

<table>
<thead>
<tr>
<th>Component</th>
<th>Single</th>
<th></th>
<th>Binary</th>
<th></th>
<th></th>
<th>Ternary</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>J_1</td>
<td>J_2</td>
<td>J_1</td>
<td>J_2</td>
<td>J_1</td>
<td>J_2</td>
</tr>
<tr>
<td>Lovastatin</td>
<td>0.14</td>
<td>70.51</td>
<td>1.08</td>
<td>96.58</td>
<td>1.10</td>
<td>99.96</td>
<td></td>
</tr>
<tr>
<td>Valsartan</td>
<td>3.01</td>
<td>83.64</td>
<td>14.24</td>
<td>100.00</td>
<td>14.25</td>
<td>100.00</td>
<td></td>
</tr>
<tr>
<td>Paracetamol</td>
<td>0.23</td>
<td>61.88</td>
<td>0.97</td>
<td>100.00</td>
<td>0.97</td>
<td>100.00</td>
<td></td>
</tr>
<tr>
<td>Budesonide</td>
<td>0.13</td>
<td>61.00</td>
<td>0.48</td>
<td>100.00</td>
<td>0.49</td>
<td>99.84</td>
<td></td>
</tr>
<tr>
<td>Allopurinol</td>
<td>9.60 × 10^{-4}</td>
<td>63.55</td>
<td>9.67 × 10^{-4}</td>
<td>95.17</td>
<td>0.0014</td>
<td>96.42</td>
<td></td>
</tr>
<tr>
<td>Furosemide</td>
<td>0.03</td>
<td>72.29</td>
<td>0.13</td>
<td>62.14</td>
<td>0.14</td>
<td>99.88</td>
<td></td>
</tr>
<tr>
<td>Sulfadiazine</td>
<td>0.16</td>
<td>61.86</td>
<td>0.64</td>
<td>99.42</td>
<td>0.66</td>
<td>99.98</td>
<td></td>
</tr>
</tbody>
</table>

4.5. References

A version of this chapter has been published as:

5.1. Introduction

Crystallization process is an important unit operation in the chemical and pharmaceutical industries. This process involves transport phenomena in which the driving force is supersaturation that causes the diffusion of mass from the solution to the solid phase [1]. The difference between the activity of a component in the bulk of a solution and the surface (saturation point) is the driving force for transport of material. The supersaturation is mostly achieved by cooling, evaporation, or anti-solvent addition [2]. Advances in technologies of manufacturing pharmaceutical drugs have had much effect on reducing the time of production and increasing the safety and profitability of the drugs [3]. Many chemicals and pharmaceuticals have the ability to crystallize in different lattice shapes that is called polymorphism. The polymorphs of a substance have the same properties in the solution and gaseous phase, but they have different behaviour in solid state. There are some properties which make the polymorphs differ from each other. These include solubility, dissolution rate, bioavailability, colour, viscosity, density, crystal shape, and mechanical properties [4]. For a pharmaceutical product, the desired polymorph, the size and the shape of the crystals are important [5]. The efficiency and performance of many downstream unit operations (such as filtration, drying and milling) in the particulate industries depend on the quality of the solids produced in the crystallization operation. Most crystallization processes (especially for the pharmaceuticals) are performed in batch units. Although the industrial batch processes for crystallization of components are well-understood, there are still some factors to study which can affect the size distribution of particles and the polymorphic distribution of the product [6].

The modeling theory for crystallization processes and generally, particulate systems, was first proposed by Hulburt and Katz [7]. The methodology pursued and developed by Randolph and Larson [8]. The theory of particulate processes and population balance equations with their analytical solutions was extensively studied by Ramkrishna [9]. Analytical solutions of population balance equation (PBE) are limited to simple problems. For more complex systems where there are more than one form of crystal (such as polymorphic systems) and in the presence of dissolution process, the analytical solutions do not exist. There are some numerical methods of PBE solutions, namely method of moments (Hulburt and Katz [7], Randolph and Larson [8], Ramkrishna [9], and Marchisio et al. [10]), method of classes (Marshal et al. [11] and Hounslow
et al. [12]), orthogonal collocation [13], Monte Carlo simulation [14], and finite element method [15]. From the mentioned methods, method of moments (MoM) and method of classes (MoC) are easier and faster than other methods which can be used to optimize and control the crystallization systems. The ease of using the MoM is of importance when dealing with polymorphic transformation systems.

The method of moments reduces the PBE to a set of ordinary differential equations in terms of moments. The main drawback of this method is that it cannot generate the crystal size distribution [1]. Another important note about this method is its inability in solving crystallization systems in which the kinetic rates are non-linear and size-dependent. Some of the moments of a population density can be approximated experimentally through the use of sensors such as focused beam reflectance measurement, FBRM [16]. In the current work we used the method of moments with the dissolution term added to incorporate the transformation of the metastable to stable form. The entire concentration-temperature (C-T) surface was divided to sub-regions and a slightly different model was applied to each region. Because of the rapid calculation of the proposed method, the optimal control and real-time optimization of such systems can be very efficient.

In the method of classes, the PBE is transformed into a finite set of ODEs which are produced by discretization sizing technique [17]. The main difference of this method with the method of characteristic is the basis of formulation. In method of classes, the ODEs which are generated from the PBE are in terms of number of particles in each class over the size domain, while in the method of characteristic the ODEs are based on number density function. The advantage of using number of particles instead of number density function is the faster numerical processing for the class method [17].

In the current work, we also have used both methods of moment and classes to get particle moments, CSD, and concentration evolution over time. In order to speed up the calculation of ODE sets for the class method, we used MOM and MOC in successive procedure. The output of the MOM method which consists of birth, growth, and dissolution rates is used as input for the MOC method. In this way all the variables which make the conventional MOC a very time-intensive method, will be firstly calculated using method of moments, coupled with the mass balance for the whole system.
5.2. Process modeling and simulation

The mathematical framework of a crystallizer consists of the population balance equation, mass, and energy balances [7]. The general one-dimensional form of the population balance equation for a batch and perfectly mixed system that includes the dissolution term is:

\[
\frac{\partial f_i(t,L)}{\partial t} + \frac{\partial [G_i(t,L)f_i(t,L)]}{\partial L} - \frac{\partial [D_i(t,L)f_i(t,L)]}{\partial L} = B_{\text{nucleation},i}(t,L)\delta(L-L^*) \tag{5-1}
\]

where \(i\) refers to each polymorph in the system. \(f_i(t,L)\), \(G_i(t,L)\), \(D_i(t,L)\), \(B_{\text{nucleation},i}\), and \(\delta\) are referred to number density function, growth rate, dissolution rate, nucleation rate of polymorph \(i\), and Dirac delta function, respectively. For the current study we have made some assumptions in order to simplify the modeling. The breakage and agglomeration of particles were not taken into account in the system and the volume of the batch is kept constant. The birth term is multiplied by the Dirac delta function \((\delta)\) to impose the nucleation process for particles of size \((L^*)\). The term for dissolution rate \(D(t,L)\) is added to equation 5-1, as the process under study represents a polymorphic transformation system. It should be noted that for a polymorphic system, depending on where the process concentration is located on the concentration-temperature (C-T) diagram (Figure 5-1), the dissolution term has to be added or removed from the equation 5-1.

![Figure 5-1. The schematic representation of the polymorphs solubility and three distinct regions in between the curves for an arbitrary component; (Region A) nucleation and growth of stable and metastable forms, (Region B) nucleation and growth of stable form and dissolution of the metastable form, (Region C) dissolution of both forms.](image)
In addition to the PBE we need to have mass and energy balance in order to complete the model description. The mass balance for the batch crystallization system will be:

\[
\frac{dC(t)}{dt} = -3 \sum_{i} \rho_{c_i} \alpha_{v_i} \int_{0}^{\infty} G_{i}(t, L)L^{2}f_{i}(t, L)dL + 3 \sum_{i} \rho_{c_i} \alpha_{v_i} \int_{0}^{\infty} D_{i}(t, L)L^{2}f_{i}(t, L)dL \tag{5-2}
\]

\(C, \rho_{c_i}, \) and \(\alpha_{v_i}\) refer to solute concentration, particle density, and volume shape factor of polymorph \(i\), respectively. We added the second term on the right side of equation 5-2 to include the effect of dissolution of polymorphs. The energy balance for a jacketed vessel is:

\[
\rho_{s}c_{s} \frac{dT(t)}{dt} = -3 \sum_{i} \Delta H_{\text{crystallization},i} \rho_{c_i} \alpha_{v_i} \int_{0}^{\infty} G_{i}(t, L)L^{2}f_{i}(t, L)dL
\]

\[+ 3 \sum_{i} \Delta H_{\text{dissolution},i} \rho_{c_i} \alpha_{v_i} \int_{0}^{\infty} D_{i}(t, L)L^{2}f_{i}(t, L)dL - UA\Delta T_{\text{im}} \tag{5-3}\]

\(\rho_{s}, c_{s}, \Delta H_{\text{crystallization},i}, \Delta H_{\text{dissolution},i}, UA,\) and \(T_{\text{im}}\) are the solution density, specific heat of the solution, heat of crystallization, heat of dissolution of polymorph \(i\) in solution, overall heat transfer coefficient, and log-mean temperature of the crystallizer, respectively. In the current study the temperature of the system can be controlled using a cascade controller. Therefore, the system temperature is adjusted to the proposed trajectory and the equation 5-3 need not be solved.

5.2.1. Numerical solution of the model

In order to distinguish the seeded and newly nucleated crystals, we divide the particles into two parts:

- Seeded particles with a growth governing process (although in rare cases the nucleation can also occur on the surface of particles [18]).
- Newly nucleated particles at high supersaturation levels and then grown in size.

With multiplication of the equation 5-1 by the \(L^{j}\) and after rearrangements, the following formulation will be derived for each polymorph in the system:
\[
\frac{d\mu_i^0}{dt} = B_{\text{nucleation},i} \quad (5-4)
\]

\[
\frac{d\mu_i^j}{dt} = jG_i\mu_i^{j-1} - jD_i\mu_i^{j-1} \quad j = 1, 2, ..., 5 \quad (5-5)
\]

Equation 5-4 and 5-5 are written for \( j = 0 \) to 5 to cover the zeroth to the fifth moment of each polymorph \( i \) present in the system. \( \mu_i^j \) is the \( j \)th moment of polymorph \( i \). Based on the moments of polymorphs, the mean size of particles of each polymorph can be estimated [1]:

\[
[L_{m,n}]_i = \frac{\int_0^\infty L^m f_i(t,L) dL}{\int_0^\infty L^n f_i(t,L) dL} \quad (5-6)
\]

\([L_{1,0}]\) is the number-averaged and \([L_{4,3}]\) is the volume-averaged mean size for each polymorph.

Having the initial conditions, such as the seed size distribution, concentration, and temperature, equation 5-4 to 5-5 with mass balance will be solved for each polymorph over time. The seed size distribution in our study is assumed as [19]:

\[
f_i(0,L) = f_{\text{seed},i}(L) = \frac{\epsilon_i}{\sqrt{2\pi}\sigma_{\text{seed},i}} \exp \left[ -\frac{(L - \mu_{\text{seed},i})^2}{2\sigma_{\text{seed},i}^2} \right] \quad (5-7)
\]

In which \( \sigma_{\text{seed},i} \) and \( \mu_{\text{seed},i} \) are the standard deviation and average size of the seed of polymorph \( i \) and \( \epsilon_i \) is a constant which is used to set the mass of each polymorph seed. Also the mass of each polymorph at the start of the process can be evaluated from:

\[
m_{\text{seed},i} = \rho_{\text{seed},i} \alpha \int_0^\infty f_{\text{seed},i}(t,L)L^3 dL \quad (5-8)
\]

\( \rho_{\text{seed},i} \) and \( f_{\text{seed},i} \) are the density and number density of polymorph \( i \) in the seed. The Dormand-Prince method (Matlab, 2008) was used to solve the ODE set. The moments for each polymorph are written for the nucleated particles and seeds, separately. For a system of two polymorphs (which is the dominant case in crystallization systems), there would be a total of 25 ODEs that need to be solved, simultaneously. The initial conditions for each set of equations are the moments of particles for each polymorph (from \( 0^{th} \) to \( 5^{th} \) moments of seeds), initial temperature, and concentration.
Two important evaluated parameters from solving the equations of moment and mass balance for the polymorphic system are birth and growth rates throughout the process. In order to get the crystal size distribution of the particles in the slurry for each time step, we used the method of classes for each polymorph, including the effect of growth and dissolution on the whole model framework. The conventional method of classes (MOC) can be used for any type of crystallization system, no matter what type of growth function is applied to the model. Although we can apply the MOC for any type of system, but the time-demanding nature of the method is the main drawback. This will be a significant issue when this method has to be used as a model for control and optimization purposes (i.e., optimal control). To overcome this issue, we combined the method of moments for the polymorphic system to the method of classes. In this way the results of MOM will be successively used by MOC to generate CSD of polymorphs throughout the process time, which we named it as the method of moment and classes (MOMC).

The upcoming formulas were written for all of the polymorphs present in the system. Let us divide the whole length domain (x-dimension) by \( m \) classes and time domain by \( n \). The smallest value for the length domain is \( L^* \) (that is the size of particles generated through birth process), which we assume zero for numerical calculations. The increment of length for each class of particles is:

\[
\Delta u_j = L_j - L_{j-1} \quad (5 - 9)
\]

Each class of particles has the average size of:

\[
\overline{L}_j = \frac{L_j + L_{j-1}}{2} \quad (5 - 10)
\]

In order to discretize the PBE through the length domain, the equation based on number density function has to be changed to the total number per class version. The number of particles at each class of length \( \Delta u_j \) is:

\[
N_{ii} = \int_{j-1}^{j} f_i(t, L) dL \quad (5 - 11)
\]
By substitution of equation (5-11) into equation (5-1), the following will be generated for each class:

\[
\frac{1}{V(t)} \frac{dN_i(t,L_j) V(t)}{dt} + \int_{j-1}^{j} \partial \left[ G_i(t,L_j) f_i(t,L_j) \right] \frac{dL}{\partial L} - \int_{j-1}^{j} \partial \left[ D_i(t,L_j) f_i(t,L_j) \right] \frac{dL}{\partial L} = (B_{\text{nucleation}} \text{ or } 0) \quad (5 - 12)
\]

The nucleation term is non-zero for the first class. If we assume the value of number density at each node (such as \( j \)) is the arithmetic average of the average number density of its neighbours, then:

\[
f_i(t,L_j) = \frac{f_i(t,L_{j-1}) + f_i(t,L_j)}{2} \quad (5 - 13)
\]

And each term on the right side of equation (5-12) can be substituted with the average value of number of particles in the class:

\[
f_i(t,L_j) = \frac{N_{j-1,i} + N_{j,i}}{2} \quad (5 - 14)
\]

Which gives the general discretized form of PBE equation for each class of \( j \):

\[
\frac{1}{V(t)} \frac{dN_i(t,L_j) V(t)}{dt} = \frac{G_{j-1,i}}{2 \Box_{j-1}} N_{j-1,i} \frac{G_{ji}}{2 \Box_j} N_{ji} + \frac{D_{j+1,i}}{2 \Box_{j+1}} N_{j+1,i} - \frac{D_{ji}}{2 \Box_j} N_{ji} \quad (5 - 15)
\]

The above formula is written for \( m-2 \) classes, except the first and last classes. For the two remaining classes (boundaries) we have the following notes:

- First class; There is no flow of particle from the previous class, and in addition, the nucleation will only occur here, thus:

\[
\frac{1}{V(t)} \frac{dN_{i1} V(t)}{dt} = -\frac{G_{i1}}{2 \Box_1} N_{i1} + \frac{D_{2i}}{2 \Box_2} N_{2i} - \frac{D_{i1}}{2 \Box_1} N_{i1} + B_{\text{nucleation}} \quad (5 - 16)
\]

- Final class; There is no flow out of this class, and therefore:

\[
\frac{1}{V(t)} \frac{dN_{mi} V(t)}{dt} = -\frac{G_{mi}}{2 \Box_m} N_{mi} + \frac{G_{m-1i}}{2 \Box_{m-1}} N_{m-1i} - \frac{D_{mi}}{2 \Box_m} N_{mi} \quad (5 - 17)
\]
Or in matrix notation:

\[
\frac{1}{V(t)} \frac{d[V \mathbf{N}(t)]}{dt} = \omega \mathbf{N}_{1} + \tau \mathbf{N}_{1} + \mathbf{B} \delta (L - L')
\]

(5 - 18) \( I = \) for all polymorphs

In which:

\[
\omega = \frac{G}{2\theta} \begin{bmatrix}
-1 & 1 & 0 & 0 & 0 & \cdots & 0 \\
0 & -1 & 1 & 0 & 0 & \cdots & 0 \\
0 & 0 & -1 & 1 & 0 & \cdots & 0 \\
0 & 0 & 0 & -1 & 1 & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & 0 & \cdots & 0 & -1
\end{bmatrix}
\]

(5 - 19)

\[
\tau = \frac{D}{2\theta} \begin{bmatrix}
-1 & 1 & 0 & 0 & 0 & \cdots & 0 \\
0 & -1 & 1 & 0 & 0 & \cdots & 0 \\
0 & 0 & -1 & 1 & 0 & \cdots & 0 \\
0 & 0 & 0 & -1 & 1 & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & 0 & \cdots & 0 & -1
\end{bmatrix}
\]

(5 - 20)

**Figure 5-2.** A sample class (j\textsuperscript{th} class) and its neighbour segments; the arrows represent (1) the growth of particles from a class of lower average size, (2) crystal growth from the main class, (3) dissolution from a class higher in average size, (4) crystal dissolution from the main class.
In above notations the class size is set constant throughout the length domain. As it was shown earlier, the birth and growth rates for the whole dynamic process can be found from the MOM outputs. Then the values for each time step are used to accelerate the MOC method to generate the CSD of particles. The schematic diagram of the MOC method for polymorphic system is shown in Figure 5-2. Based on the initial, boundary conditions, the kinetic parameters of growth, nucleation, dissolution, and simulation parameters (i.e., time and length domain step changes), the model will evaluate all the moments, concentration, birth, growth rates, and CSD over time. In order to solve the ODE set, which contains nucleation, growth, and dissolution parameters the method of Runge-Kutta with the order of (4,5) has been used as the starting solution. In order to prevent failure of the mentioned method in some points, the backward numerical differentiation (Gear’s method) has been added to the program to resolve those points. The output of the MOM method is the initial points for the MOC method of the polymorphic crystallization. As the method was clarified before, based on $m$ segments in size domain and $n$ time increments, there are $m \times n$ ODE sets that have to be solved. Using MOM as a preliminary method for MOC can speed up the CPU time, significantly. The flowchart for the algorithm of MOMC method is shown in Figure 5-3.
Figure 5-3. Algorithm for solving polymorphic transformation using MOMC method
5.2.2. Objective functions for the optimal control of the process

For the optimal control of a polymorphic transformation crystallization system, we defined three objective functions which are of importance in industry:

- \( J_1 \), maximize the mass of stable polymorph at the final batch time
- \( J_2 \), maximize the average size of particles for the stable form at final batch time
- \( J_3 \), maximize the mass of metastable form at the final batch time

The optimization procedure involves the discretization of the time domain to a finite time steps \((t \in [0, t_f])\) and then, implementing the constrained NLP to optimize each segment of the control vector. In order to perform the dynamic optimization, we used the Karush-Kuhn-Tucker (KKT) method followed by conversion of constrained optimization function to a set of unconstrained non-linear problems. This algorithm is efficient for the polymorphic transformation system, in which there are highly non-linear functions of growth, dissolution, and nucleation with mass balance equation. The mathematical formulation of the optimal control for the process is:

Maximize \( J_1(x(t), u(t), \tau) \) \hspace{1cm} (5 – 21)

Subject to:

\[
\begin{align*}
\frac{dx}{dt} &= f_1(x(t), u(t), \pi) \hspace{1cm} (5 – 22) \\
y(t) &= g(x(t), u(t), \pi) \hspace{1cm} (5 – 23) \\
R_{\text{min}} &\leq \left| \frac{du(t)}{dt} \right| \leq R_{\text{max}} \hspace{1cm} (5 – 24) \\
U_{\text{min}} &\leq u(t) \leq U_{\text{max}} \hspace{1cm} (5 – 25) \\
u_{\text{final}} &= \text{Fixed value} \hspace{1cm} (5 – 26)
\end{align*}
\]

where \( 1 \) and \( i \) denote the objective function number and the polymorph, respectively. \( u(t) \) is the control vector (such as operating temperature or anti-solvent flow rate) and \( x(t) \) is the state variables, which consists of moments, solute concentration, etc. \( y(t) \) is any measurement vector such as mass or volume of polymorphs during the process time. \( \pi \) denotes all the parameters that are used in kinetic rates. \( R_{\text{min}} \) and \( R_{\text{max}} \) are the minimum and maximum cooling rates. \( u_{\text{final}} \) is the final value of the control vector, which enforces the optimization procedure to bound the control
at the final batch time. The initial guess for temperature profile is a linear cooling, which for most cases provides a good start for dynamic optimization procedure.

5.3. The model compound and the simulation using the conventional cooling policies

One of the interesting model compounds to study in the field of polymorphic transformation is L-glutamic acid, which has a stable ($\alpha$) and a metastable ($\beta$) form [20]. Scholl et al. conducted extensive experiments on polymorphic transformation of L-glutamic acid and found the kinetic parameters representing the important behaviours during the process [21]. Cornel et al. studied the effect of process parameters (such as stirring rate and impurity) on the transformation rate of metastable to stable at 45°C [22]. Ono et al. used the Raman spectroscopy to quantify the fraction of stable to metastable during crystallization of L-glutamic acid at different temperatures [18]. They concluded that the temperature has a strong effect on transformation process. The kinetic parameters which are used in our study have been found by Scholl et al [21]. Also the data on the solubility of two forms were found from literature and fitted [20]. All possible cases of crystallization scenarios based on the initial concentration of solute and its position with respect to the solubility of two polymorphs, have been studied. For all the cases the volume of system is kept constant and the temperature profile is changed. The following are the nucleation, growth, and dissolution kinetic equations implemented in our model [21]:

\[
B_\alpha = k_{n\alpha} S_\alpha^7 \exp \left[ -\frac{K_{n\alpha}}{\ln^2 S_\alpha} \right] \tag{5 - 27}
\]

\[
G_\alpha = k_{g\alpha} (S_\alpha - 1)^5 \exp \left[ -\frac{K_{g\alpha}}{S_\alpha - 1} \right] \tag{5 - 28}
\]

\[
D_\alpha = k_{d\alpha} (1 - S_\alpha) \tag{5 - 29}
\]

\[
B_\beta = k_{n\beta} S_\beta^7 \exp \left[ -\frac{K_{n\beta}}{\ln^2 S_\beta} \right] + k_{s\beta} \mu_a^2 \exp \left[ -\frac{K_{s\beta}}{\ln S_\beta} \right] \tag{5 - 30}
\]

\[
G_\beta = k_{g\beta} (S_\beta - 1)^5 \exp \left[ -\frac{K_{g\beta}}{S_\beta - 1} \right] \tag{5 - 31}
\]

$B$, $G$, and $D$ are birth, growth, and dissolution rates, respectively. In the above equations, supersaturation is defined as the ratio of solute concentration to the solubility concentration at a given temperature. The kinetic parameters which were used in this study are shown in Table 5-1 [21, 23].
Table 5-1. Kinetic parameters with their values and units

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Value</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k_{n\alpha}$</td>
<td>$8.0 \times 10^5$</td>
<td>$\frac{#}{m^3 s}$</td>
</tr>
<tr>
<td>$K_{n\alpha}$</td>
<td>$1.0 \times 10^{-1}$</td>
<td>-</td>
</tr>
<tr>
<td>$k_{g\alpha}$</td>
<td>$2.5 \times 10^{-7}$</td>
<td>$\frac{m}{s}$</td>
</tr>
<tr>
<td>$K_{g\alpha}$</td>
<td>$9.0 \times 10^{-2}$</td>
<td>-</td>
</tr>
<tr>
<td>$k_{d\alpha}$</td>
<td>$3.5 \times 10^{-5}$</td>
<td>$\frac{m}{s}$</td>
</tr>
<tr>
<td>$k_{n\beta}$</td>
<td>$5.4 \times 10^4$</td>
<td>$\frac{#}{m^3 s}$</td>
</tr>
<tr>
<td>$K_{n\beta}$</td>
<td>$1.5 \times 10^1$</td>
<td>-</td>
</tr>
<tr>
<td>$k_{s\beta}$</td>
<td>$6.0 \times 10^4$</td>
<td>$\frac{#}{m^2 s}$</td>
</tr>
<tr>
<td>$K_{s\beta}$</td>
<td>$1.0 \times 10^{-3}$</td>
<td>-</td>
</tr>
<tr>
<td>$k_{g\beta}$</td>
<td>$6.5 \times 10^{-8}$</td>
<td>$\frac{m}{s}$</td>
</tr>
<tr>
<td>$K_{g\beta}$</td>
<td>$1.6 \times 10^{-1}$</td>
<td>-</td>
</tr>
</tbody>
</table>

Case 1. Supersaturation with respect to both forms (Region A in Figure 5-1)

This case has been studied by Cornel et al., experimentally [22]. In their study they performed different cases of polymorphic transformation with initial concentration high enough to start nucleation and growth of both polymorphs. All experiments were conducted at constant temperature of 45°C. In order to examine the behaviour of the crystallization process, we assumed that both stable and metastable forms were present, initially. The seed size distribution of both polymorphs is given by equation 5-7, with parameters given in Table 5-2. The product quality indices are of high importance in the particulate industry, such as the ratio of the mass of
the two forms and the ratio of the seeded volume to the newly nucleated volume for each polymorph.

**Table 5-2.** Seed size distribution parameters for stable (α) and metastable (β) forms used in equation 5-7

<table>
<thead>
<tr>
<th>Polymorph</th>
<th>ε</th>
<th>μ (m)</th>
<th>σ (m)</th>
<th>αν</th>
<th>ρc (kg/m³)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stable</td>
<td>1 x 10¹⁰</td>
<td>50 x 10⁻⁶</td>
<td>2 x 10⁻⁶</td>
<td>0.031</td>
<td>1540</td>
</tr>
<tr>
<td>Metastable</td>
<td>1 x 10¹⁰</td>
<td>30 x 10⁻⁶</td>
<td>2 x 10⁻⁶</td>
<td>0.48</td>
<td>1540</td>
</tr>
</tbody>
</table>

In this case study we examined the effect of cooling profile on the supersaturation profile. The cooling profile was implemented from the following relation [24, 25]:

\[ T(t) = T_o - (T_o - T_f) \left( \frac{t}{t_{batch}} \right)^P \]  \( (5 - 32) \)

where \( T_o, T_f, t_{batch}, \) and \( P \) are initial temperature, final temperature, batch time, and a power number of cooling policy. The value of \( P \) can be changed according to the desired policy which will be implemented on the system. It should be noted that the cooling policy from equation 5-20 is not necessarily give the optimal trajectory. The three following values of \( P \) were selected to cover three common cooling conditions:

- \( P = 0.1 \), corresponds to an approximation of natural cooling,
- \( P = 1.0 \), which represents the linear cooling, and
- \( P = 3.0 \), that can be considered as a non-linear approximate optimal cooling.

The initial solute concentration and temperature are 43g solute/kg solvent and 45°C, respectively. According to the results shown in Figure 5-4a, the solute concentration drops rapidly using the natural cooling. Therefore most of the supersaturation is consumed initially, yielding a rich content of fine particles. For the higher values of \( P \), the concentration curve shows a sudden decrease at around 0.25 h, and then, there is a moderate supersaturation until nearly the end of the process. Therefore, the profile with higher value of \( P \) will result in a better quality of product. It should be noted that for all three conditions we assumed simultaneous nucleation and growth of both polymorphs (region A in Figure 5-1).
Figure 5-4. (a) Supersaturation profile, (b) Three cooling policies implemented on the crystallization system, (c) the ratio of the stable to metastable masses, and (d) mass of metastable form of L-glutamic acid over time (for the case 1)
Figure 5-4c shows that the ratio of the stable to the metastable form for all of the three cooling policies drops in the first 0.1 h of the process and then increases until the end of the batch time. However, the values of this ratio are lower than unit for the whole process and for all of the three cooling policies. The decrease in the ratio in the first part of the process is due to the rapid growth and nucleation of metastable form in comparison to the stable form. The cooling policy with $P = 3.0$ has the highest value of ratios and favours the production of the stable form. From Figure 5-2d it can be seen that the mass of the metastable form is increased for all of the three policies, however, for $P = 0.1$ the increase in metastable production is the highest. Under these conditions, if the objective is to produce the maximum amount of metastable form, the cooling policy of $P = 0.1$ should be selected to favour the formation of that form.

**Case 2. Supersaturation with respect to the stable form and undersaturation with respect to the metastable form**

If the initial solute concentration lies in between the solubility curves of the stable and metastable forms, then the metastable form dissolves. This increases the solute concentration and thus, the supersaturation with respect to the stable form is raised. The metastable form dissolves and the stable form is produced. The transformation from the metastable to the stable form will last until all of the metastable form is dissolved. In this case, we considered three scenarios:

- Un-seeded condition,
- Seeded with only the stable form, and
- Seeded with both polymorphs.

**Case 2.1. Un-seeded condition (Region B in Figure 5-1)**

At $45^\circ$C the solubility of the stable and metastable forms are 16.46 and 23.01 g solute/kg solvent, respectively. In order to start the crystallization process with an un-seeded condition, there should be a large supersaturation initially and maintained during the process. Therefore we selected the initial supersaturation with respect to the stable form as much as possible to induce the nucleation of this form. It should be noted that the process operating point at start is in region B of the Figure 5-1. Because of the small distance of the starting point of the process to the metastable solubility curve, it is more likely that the C-T operating curve crosses the metastable solubility curve and produces the metastable form to some extent. The batch time for the un-
seeded case was set to 15 h. This is because of the kinetic behaviour of nucleation for L-glutamic acid which enforces the nucleation to start with much delay. Based on the Figure 5-5, the supersaturation is a strong function of the cooling rate. For the sudden drop in temperature at the early stages of the process, there is a sudden increase in supersaturation and then a decrease, instantaneously. However, the supersaturation profiles for other cooling rates are different. As the power number for cooling policy (P, equation 5-20) increases, the average supersaturation is higher and thus, the product quantity and quality will be changed.

![Figure 5-5](image)

**Figure 5-5.** Time evolution of supersaturation with respect to the solubility of the stable form (top) and average size of the stable form based on number-weighted definition (bottom) in case of un-seeded operation.
For the natural cooling profile, the average size of the stable form increases until nearly 4 h of the process, as the supersaturation is high in the early stages of the operation. The average size (number-weight) remains constant from the 5 h of the process until the end of the batch. This behaviour is also the same for the case of linear cooling, however in the first 4 h of the process, the average size is nearly zero. For the case of $P = 3.0$, the size remains nearly zero for the first 7.5 h of the operation. The second part of the process (for $P = 3.0$) shows the growth in size. During this second period, the temperature is suddenly reduced to 25°C. Because of this sudden change, the nucleation occurs rapidly followed by growth and an increase in the size of the stable form. It is worth noting that the average size of the particles is based on the number of particles present in the system, which is different from the volume average size. This type of size was considered here, as the dominant process is nucleation, instead of the growth. All of the three curves of average sizes in Figure 5-5 show a maximum point during the process. This happens because of the rate of change of the first moment of particles is faster with respect to the zeroth moment of particles at a specific time of the process. However, after that time, the size of particles tends to be constant. The mass of the stable form produced minus the mass of the metastable form for a range of cooling policies from a near-to-flat cooling policy during the considered batch time ($P = 4.0$) to a sudden drop at initial ($P = 0.01$) is illustrated in Figure 5-6. The batch time for all of the power numbers is 2 h.

Based on the results shown in Figure 5-6, two conclusions can be drawn:

- For small $P$ values up to 0.2, the stable form production is inversely related to the initial concentration of batch system and this behavior will be reversed for power numbers higher than 0.2.
- For each initial concentration, there is an optimum point at which the difference between the stable and metastable form masses is maximized. For three different initial concentrations (while all other process and seeding parameters are kept constant), the optimum power number is related to the initial concentration and its value ranges from 0.17 to 0.47 for initial concentration from 17 to 21 g solute/kg solvent.

It can be concluded that in order to achieve as much stable form as possible for the cooling policies given by equation 5-20, one has to select a near-to-natural cooling policy. However this
case is only valid when the process is conducted in an un-seeded manner and the batch time is short.

![Graph](image)

**Figure 5-6.** Difference of the mass of stable form produced from the metastable in a wide range of cooling policies for three different initial concentrations

**Case 2.2. Seeded only with the stable form (Region B in Figure 5-1)**

Here, we seed the crystallizer with the stable polymorph in the absence of the metastable form. The initial solute concentration for all conditions was set to 20 g solute/kg solvent, 45°C as the initial temperature, and 20°C as the final temperature. The seed parameters were taken from Table 5-1, except the average size (µ) and ε which were varied over three values. At first, the ε was adjusted so as to maintain the same mass of 0.6 g seeds/kg solvent for all three conditions and µ was changed from 50 to 150µm. The operation time for this case was set to 5 h and the linear cooling policy was adopted for all of the three seed conditions. We studied the effect of changing the seed size on the process performance. From Figure 5-7 it is found that no metastable form is produced during the process because of undersaturation with respect to metastable form, while the mass of the stable form increases during the run. It should be noted that the mass of the stable form at the end of the batch for all three seed sizes was almost the same (12.48 g/kg solvent), which implies the independency of product mass with seed size. This fact is in agreement with the findings by Cornel et al [22]. For the seeds of higher mean size, the
precipitation process will be more rapid than the seeds with smaller mean size. As a result, the supersaturation will be higher for the case of smaller seed size.

The effect of changing the seed mass of the stable form on the process output is shown in Figure 5-8. The seed parameters are the same as the ones in Table 5-1 except for the $\varepsilon$, which defines the mass of seed. The average size of the seed in three conditions was 50µm. The initial concentration, initial and final temperature, and other conditions were the same as above.

**Figure 5-7.** Concentration evolution (top) and supersaturation (bottom) over time for three different seeding average sizes (µm)
Figure 5-8. Supersaturation profile (top) and average size of particles (bottom) for three different masses of stable form seeds

As the mass of seeds increases, the concentration of solution is dropped more rapidly. This is because of the presence of more particles that are used as precipitation sites for crystallization from the solution. Therefore, as it is depicted in Figure 5-8, the supersaturation for the case of seed mass of 0.6 g/kg solvent has the highest value, while the curve shows the lower values for the higher mass of seeds. Because finally the concentration is the same for three cases, the same mass of crystals will be precipitated on the seed surfaces. As a result, at lower seed loading (or the lower numbers of particles) results in larger average size (Figure 5-8).
Case 2.3. Seeded with both polymorphs (Region B in Figure 5-1)

In the final case we studied the effect of the presence of stable and metastable form seeds together. In order to see the effect of dissolution and cooling policy on the product quality, three cooling scenarios were selected. The seeding for all three cooling policies was the same, with the average seed size of 40µm and standard deviation of 2µm. Initial solute concentration for all conditions was set to 20 g solute/kg solvent, 45°C as initial temperature, and 20°C as final temperature. The seeding mass of metastable and stable forms was 2.39 and 0.03 g seed/kg solvent, respectively. The reason for setting much higher value for metastable seeding mass compared with the stable form is the dissolution process of the metastable form, which will be transferred to the stable form (depending on cooling policy adopted). The process variables are shown in Figure 5-9.

Figure 5-9a illustrates the average size of the metastable form (L_{1,0}) for the cooling scenarios. The natural cooling policy (P = 0.1) results in the largest mean size over the entire process (this is due to the short batch time employed). The sharp increase in the average size is a result of fast cooling process, which enforces the C-T curve to cross the metastable solubility curve, and results in the formation of metastable polymorph. The next highest average size is for linear and non-linear (P = 3.0) scenarios. It is seen that the size of metastable form for these two procedures lies on each other until nearly 0.5 h after the start of the process. The reason for this behaviour is that the metastable solubility curve has an approximately equal distance with the C-T curve for both forms and thus, the same undersaturation. However, after this time, the size of the metastable form differs for the two policies, as the cooling path shows its effect on the process output.

Figure 5-9b shows the average size of the stable form over process time. Due to the reduction in size of metastable form and its dissolution, the stable form will nucleate and grow. It is interesting to see that the rapid cooling makes the stable form grow in size for nearly 2.5 h of the process and then increase very slightly. For the case of linear cooling, the size of stable form rapidly grows (even higher than P = 3.0) and then, approaching a constant value. The last cooling policy (P = 3.0) shows a near-to-linear growth of stable form, and decreases in slope in the last 1 h of the process. This behaviour is a result of keeping a high and nearly constant supersaturation throughout the run.
In Figure 5-9c the supersaturation curves for three methods are displayed. As it is expected for the rapid cooling policy, the supersaturation sharply increases, and then decreases. The linear cooling has a lower slope in the C-T curve, but still has high supersaturation in the first 2.5 h of the process. However, for P = 3.0, because of (1) lower slope in temperature profile and (2) dissolution of the metastable form at early stage of the process, the supersaturation is maintained at a nearly constant level. After the whole dissolution of the metastable form, and on the other hand, cooling of the process (the decrease in temperature starts nearly 2 h after the start of the process), again the supersaturation is increased. The supersaturation level is nearly constant and high with an average value of 1.3.

Finally, in Figure 5-9d the masses of the stable and metastable forms over time for three cooling policies are illustrated. The rapid cooling leads to a high production of metastable form compared with the other two cooling methods. The final mass of the stable form for the linear and non-linear cooling (P = 3.0) is almost the same, while the rapid cooling gives the lowest mass of the stable form and results in highest mass of the metastable form. The initial mass of the stable seed form is nearly zero, while the mass of the metastable form seed is around 2.4 g seed/kg solvent. At the end of the batch process, the rapid cooling gives nearly three times of the metastable product mass as it was seeded, initially. For the other two procedures, almost all of the metastable form is transformed to the stable form.
Figure 5-9. Polymorphic transformation in the presence of stable and metastable form seeds using three cooling policies, (a) average size of metastable, (b) average size of stable form, (c) supersaturation with respect to stable solubility, (d) mass of stable and metastable form.
5.4. Optimal control for polymorphic transformation in the presence of dissolution process

With the conventional cooling discussed above, some cooling policies resulted in higher masses of the stable or metastable form, while the average size of products was not favourable. On the other hand, some policies achieved a low mass of product with a desirable average size. However, using an optimal control policy with the objectives defined in section 1.3, it is possible to compromise between the yield of the desirable form and its quality in terms of the average crystal size. All cases are for the seeded crystallization with both the stable and metastable forms with the initial concentration between the two solubility curves. The initial solute concentration, initial, and final temperature were selected as 20 g solute/kg solvent, 45°C, and 20°C respectively. The batch time was selected as 4 h for all of the conditions. The \( R_{\text{min}} \) and \( R_{\text{max}} \) were taken as 0.0 and 2.0°C/min, respectively. The initial mass of the stable and metastable forms were 0.03 and 2.39 g seed/kg solvent and the average size of the stable and metastable form seeds was 40µm with the standard deviation of 2µm.

Figure 5-10 shows the optimal cooling curves in addition of the three conventional cooling policies. Figure 5-10a shows the six cooling policies for the crystallization process. For the first objective function (\( J_1 \), maximizing the mass of the stable form at the end of the batch), the curve of cooling is nearly a combination of the two non-linear and linear curves. The optimal cooling curve for the second objective function (\( J_2 \), maximizing the average size of the stable form at the end of the process) lies nearly on the cooling policy of \( P = 3.0 \) until 2 h of the process start. The slope will be higher as the process continues to the end of the batch time. The optimal curve for the third objective function (\( J_3 \), maximizing the mass of metastable form at the end of the process) is very different from other cooling policies. Until 1 h into the process, the curve is nearly the same as the non-linear mode (\( P=3.0 \)). From the 1 h to the 2 h of the process, the temperature suddenly drops to the value of 20°C and is kept constant at this temperature until the end of the process.

The average size of the stable form is illustrated in Figure 5-10b. This shows that the optimal cooling policy will reach even larger final product size of particles in comparison to the non-linear cooling. It is worth noting that other values of \( P \) higher and lower that \( P = 3.0 \) were tried to confirm that the optimal policy has the highest value of mean size. Figures 5-10c and 5-10d show the mass of the stable and metastable forms, respectively. It can be seen from Figure 5-10d
that the optimal cooling policy for the third objective function, $J_3$, produces the highest mass of the metastable product, even higher than the rapid cooling. The numerical values of the objective functions for the optimal and regular policies are shown in Table 5-3.

### Table 5-3. Optimal values of mass of stable, metastable, and the size of stable form with their values of conventional cooling policies

<table>
<thead>
<tr>
<th>Objective function</th>
<th>Natural cooling (P = 0.1)</th>
<th>Linear cooling (P = 1.0)</th>
<th>Non-linear cooling (P = 3.0)</th>
<th>Optimal cooling</th>
</tr>
</thead>
<tbody>
<tr>
<td>$J_1 \left( \frac{g}{kg\text{ solvent}} \right)$</td>
<td>8.71</td>
<td>13.02</td>
<td>13.41</td>
<td>13.85</td>
</tr>
<tr>
<td>$J_2 (\mu m)$</td>
<td>65.54</td>
<td>98.60</td>
<td>148.50</td>
<td>153.83</td>
</tr>
<tr>
<td>$J_3 \left( \frac{g}{kg\text{ solvent}} \right)$</td>
<td>5.87</td>
<td>1.62</td>
<td>0.67</td>
<td>7.80</td>
</tr>
</tbody>
</table>

Table 5-4 shows the results of simulation for each of the previous cases when the methods of MOC and MOMC are applied. From the results that are depicted in this table, the MOMC could best simulate the polymorphic transformation process within the less CPU time than the method of classes and also, gives the same results.
Figure 5-10. Optimal trajectory of (a) temperature for different objective functions, (b) average size of stable form ($J_2$), (c) mass of stable form ($J_1$), and (d) Mass of metastable form ($J_3$)
Table 5-4. Important results of the crystallization process using the two methods of MOC and MOMC

<table>
<thead>
<tr>
<th></th>
<th>MOC method (case 1)</th>
<th>MOMC method (case I)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of size intervals</td>
<td>100</td>
<td>300</td>
</tr>
<tr>
<td></td>
<td>[-]</td>
<td>300</td>
</tr>
<tr>
<td>Number of time intervals</td>
<td>100</td>
<td>400</td>
</tr>
<tr>
<td></td>
<td>[-]</td>
<td>500</td>
</tr>
<tr>
<td>CPU time</td>
<td>290.84 Second</td>
<td>7.18 Second</td>
</tr>
<tr>
<td></td>
<td></td>
<td>7.24 Second</td>
</tr>
<tr>
<td>Mass balance error</td>
<td>1.38e-11 gr/kg</td>
<td>3.55e-14 gr/kg</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.03e-13 gr/kg</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>MOC method (case 2.1)</th>
<th>MOMC method (case 2.1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of size intervals</td>
<td>100</td>
<td>300</td>
</tr>
<tr>
<td></td>
<td>[-]</td>
<td>300</td>
</tr>
<tr>
<td>Number of time intervals</td>
<td>100</td>
<td>400</td>
</tr>
<tr>
<td></td>
<td>[-]</td>
<td>500</td>
</tr>
<tr>
<td>CPU time</td>
<td>237.73 Second</td>
<td>12.16 Second</td>
</tr>
<tr>
<td></td>
<td></td>
<td>12.89 Second</td>
</tr>
<tr>
<td>Mass balance error</td>
<td>1.46e-12 gr/kg</td>
<td>8.01e-013 gr/kg</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.46e-12 gr/kg</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>MOC method (case 2.2)</th>
<th>MOMC method (case 2.2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of size intervals</td>
<td>100</td>
<td>300</td>
</tr>
<tr>
<td></td>
<td>[-]</td>
<td>300</td>
</tr>
<tr>
<td>Number of time intervals</td>
<td>100</td>
<td>400</td>
</tr>
<tr>
<td></td>
<td>[-]</td>
<td>500</td>
</tr>
<tr>
<td>CPU time</td>
<td>290.84 Second</td>
<td>13.48 Second</td>
</tr>
<tr>
<td></td>
<td></td>
<td>14.14 Second</td>
</tr>
<tr>
<td>Mass balance error</td>
<td>1.38e-11 gr/kg</td>
<td>2.65e-14 gr/kg</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.03e-13 gr/kg</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>MOC method (case 2.3)</th>
<th>MOMC method (case 2.3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of size intervals</td>
<td>100</td>
<td>300</td>
</tr>
<tr>
<td></td>
<td>[-]</td>
<td>300</td>
</tr>
<tr>
<td>Number of time intervals</td>
<td>100</td>
<td>400</td>
</tr>
<tr>
<td></td>
<td>[-]</td>
<td>500</td>
</tr>
<tr>
<td>CPU time</td>
<td>237.73 Second</td>
<td>13.76 Second</td>
</tr>
<tr>
<td></td>
<td></td>
<td>14.42 Second</td>
</tr>
<tr>
<td>Mass balance error</td>
<td>1.46e-12 gr/kg</td>
<td>6.11e-13 gr/kg</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.46e-12 gr/kg</td>
</tr>
</tbody>
</table>
5.5. Conclusion

In this study, modeling of a batch polymorphic crystallization process of L-glutamic acid with the dissolution term of the metastable form was developed. The model was solved using the method of moments. The polymorphic transformation process was studied for solutions supersaturated with respect to both forms and undersaturation with the metastable form. The effect of cooling policies and seed average size and loadings were studied. The method is so flexible that can handle any kind of process conditions and seeding parameters. The optimal control policy was used with three single-objective functions. The optimal profiles show a better performance in comparison with the conventional cooling policies. Because of changing the stability of the polymorphic forms beyond a specific temperature in enantiotropic systems (transition temperature), the modeling and optimal control of such processes is of importance and interest. Finally, the novel method of MOMC was developed with the combination of two methods of MOC and MOM. The resulted method showed efficiency in calculation time required for process modeling. The MOMC method could generate the crystal size distribution of the crystallization system during the progress of time.

5.6. References


Chapter 6

Experimental Studies on Crystallization of Polymorphic Transformation of L-Glutamic Acid
6.1. Introduction

In the previous chapters, the main focus was on the solubility of different pharmaceuticals in pure and mixed solvents and also, the crystallization process in solvent screening and polymorphic transformation systems. In chapter 2, the solubility of pharmaceutical component in pure and mixed solvents using thermo-gravimetric method was shown. In this chapter, the experiments on monitoring of the nucleation and polymorphic transformation of L-glutamic acid are discussed.

It is known that the nucleation in a solution can occur only if there is a high supersaturation. As it is mentioned previously, the solubility of a solid in a solvent can be measured by different methods, such as the thermo-gravimetric method which we used in this project. The metastable zone width however, can't be detected by such methods. At each concentration, there is a region between the saturated and supersolubility curve known as metastable zone width (MSZW). The schematic representation of the MSZW is shown in Figure 6-1. The MSZW is dependent on various process parameters, such as concentration, presence of impurities, cooling rate, and agitation rate [1]. There are some techniques that have been employed to detect the onset of nucleation, such as electrical conductivity [2], turbidity [3], ultrasound [4], and focus beam reflectance measurement (FBRM) [1, 6].

Since a few years ago, the use of in-situ spectroscopic measurements has been started in many applications of crystallization [6-9]. From the above mentioned methods, the FBRM is an accurate method in detecting the nucleation process. Thus, we used Lasentec FBRM for our studies in nucleation and also, polymorphic transformation. The FBRM shows the counts of the particles dispersed in a solvent as a unit of no. / volume (Figure 6-2). The real time measurement data is sent to the central computer (HEL pc) for recording and further processing.

In addition to the experiments for the nucleation detection and kinetics, the polymorphic transformation study was performed with the help of FBRM. For L-glutamic acid, the two polymorphs have different shapes: stable form is needle-like and metastable form is prismatic. Because of the difference between the particle shapes, the FBRM can qualitatively detect any significant change in the particle shape distribution and thus, detect polymorphic transformation process. It should be noted that the detection depends on the crystalline shape of the polymorphs.
of a compound. This gives us the ability to detect the polymorphic transformation using FBRM, however, there are many other chemicals which have polymorphs that can't be detected using FBRM technology.

![Solubility and nucleation curves](image)

**Figure 6-1.** The solubility and nucleation curves for an arbitrary component

Figure 6-2 demonstrates the user interface program for the FBRM probe.

![FBRM user interface program](image)

**Figure 6-2.** FBRM user interface program
6.2. Nucleation onset and rate detection

The approach to estimate the nucleation kinetics of a solid in a solution from MSZW experiments was introduced by Nyvlt [10]. He assumed that the nucleation rate at the onset of nucleation process corresponds to the supersaturation rate. Based on this assumption, the rate of supersaturation generation can be formulated as a function of the cooling rate:

\[ \frac{dS}{dt} = CR \frac{dC_{sat}}{dT} \]  

(6 – 1)

Where S, CR, and C_{sat} are referred to supersaturation, cooling rate, and saturation concentration at a given temperature. Based on the works by Nyvlt, the rate of nucleation can be shown as:

\[ J_n = k_n\Delta C^m \]  

(6 – 2)

Where \( J_n \) is the nucleation rate in mass of solid / mass of solvent. He showed that by plotting the log(MSZW) versus log(CR), one can find the two important parameters of nucleation equation (m and \( k_n \)):

\[ \log CR = m\log(MSZW) + \log k_n + (m - 1)\log \left( \frac{dC_{sat}}{dT} \right) \]  

(6 – 3)

The two parameters of m and \( k_n \) that are found from the above equation will be used in the nucleation rate equation for modeling and prediction purposes.

6.2.1. Experimental procedure

In order to find the nucleation rate of L-glutamic acid at different temperatures, we prepared five different sets of experiments which were run by an automated program that was built in Matlab environment. The saturation temperatures chosen for the experiments performed were: 55, 50, 45, and 40°C. Four cooling rates of 0.5, 0.4, 0.3, and 0.2°C/min were selected for nucleation experiments. In all of the experiments, the stirring rate was adopted as 250 rpm. The stirrer was a
45° pitch blade made of teflon coated material. The stirring rate was controlled by WinISO software which is the interface between the operator and the equipment.

First, the required amount of solvent and solid are added together at a temperature higher than the saturation temperature in the 1-L crystallization vessel which was equipped with a cooling/heating jacket. The automated process is programmed in the following manner:

1. After the required amount of solid and solvent are added to the vessel, the program starts to heat the solution 5°C higher than the corresponding saturation temperature of the solution.
2. After reaching the specified temperature, the FBRM is used to ensure the complete dissolution of the solids. The temperature will be cooled down with the specified rates (such as 0.5°C per min) until the nucleation is detected.
3. After the nucleation (or cloud point) is detected by the FBRM, the temperature is kept constant for 1 hour while the nucleation and growth proceed.
4. The temperature is raised with the same rate of the cooling process until it reaches to the point mentioned in step 1.
5. The process for other cooling rates will be done (steps 1-4) until all the four cooling rates are tested for the solution.
6. At this point, the program calls the pump to add required solvent to the vessel to dilute the solution to the lower saturated temperature.
7. Steps 1-4 are repeated for the new saturation temperature.
8. Step 6 is performed for all of the four saturated temperatures.

It should be noted that all of the above steps were performed three times to get the average values of saturation and nucleation points.

6.2.2. Results and discussion

In all of the experiments, the beta form (stable form) of the L-glutamic acid were used. The results of the saturation and cloud points for different experiments are shown in Figures 6-3. As it can be seen in this figure, the higher cooling rate results in a wider MSZW. For the lower
saturation temperature, the MSZW is wider and the slope of the nucleation curve is decreased. Using the results of Figure 6-3, the MSZW as a function of the cooling rate can be found. As it can be seen from Figure 6-3a, the solubility curve is measured for four different solution concentrations. The solubility points were found using the heating of the nucleated solution until

**Figure 6-3.** (a) Measured saturated and nucleation points of solutions of different with the (b) solubility and nucleation curves for each cooling rate
the point where all of the nucleated particles disappear. The metastable points for each solution concentration is found from cooling the saturated solution with different cooling rates. One example of this procedure is shown in Figure 6-4. As it is shown in this figure, the process starts from the highest temperature (normally 5°C above the saturation point) and maintained at that temperature for an hour or more to ensure total dissolution. After reaching to the complete dissolution, the temperature is decreased with a specified rate until the cloud point is found. Thereafter, the temperature is maintained constant for about 1 hour while the nucleation is happening. The temperature again is raised with the same rate to the previous starting value for dissolution of the nucleated particles. This cycle is repeated three times for a specific solution concentration and cooling rate. The arithmetic average of the resulted values are shown in Figure 6-3b.

![Figure 6-4](image)

**Figure 6-4.** A schematic diagram of cyclic automated operation for the saturation temperature of 40°C

It can be seen from Figure 6-3b that as the cooling rate for a given solution concentration decreases, the metastable zone width is decreased. As the cooling rate decreases, the time required for the saturated solution to start nucleating is decreased. As the cooling rate goes up, the solution takes some time to respond to the changes in temperature and thus, the MSZW is wider for higher cooling rates. It is worth noting that for some cooling rates that are very high,
the time required for the onset of nucleation is enough to cross the melting point of the solvent. That's why we didn't cool the solution beyond the 0.5°C/min.

For each of the saturation points, we examined four cooling rates and therefore, the plot of logarithm of MSZW versus logarithm of cooling rate can be constructed (Figure 6-5). The four plots in Figure 6-5 show the data points with their regressed curves. If we compare equation (6-2) with the fitted curves of the diagrams in Figure 6-5, we can find the values of parameters m and \( k_n \). Table 6-1 contains the calculated values of m and \( k_n \) for different saturated temperatures.

Table 6-1. Kinetic data on the nucleation of L-glutamic acid from FBRM measurement

<table>
<thead>
<tr>
<th>Saturation temperature [°C]</th>
<th>m</th>
<th>( \frac{dC^{\text{sat}}}{dT} )</th>
<th>( k_n )</th>
</tr>
</thead>
<tbody>
<tr>
<td>55</td>
<td>2.955</td>
<td>0.79</td>
<td>( 2.78 \times 10^{-4} )</td>
</tr>
<tr>
<td>50</td>
<td>2.820</td>
<td>0.71</td>
<td>( 3.84 \times 10^{-4} )</td>
</tr>
<tr>
<td>45</td>
<td>2.624</td>
<td>0.58</td>
<td>( 4.32 \times 10^{-4} )</td>
</tr>
<tr>
<td>40</td>
<td>2.468</td>
<td>0.50</td>
<td>( 5.84 \times 10^{-4} )</td>
</tr>
</tbody>
</table>

It can be seen from Table 6-1 that the parameter m (the power in equation (6-2)) corresponds to the saturation temperature. It is important to note that the two parameters of nucleation have strong dependency on temperature. Therefore, the two following equations are derived from Table 6-1:

\[
m = 0.033T + 1.142 \quad (6 - 4)
\]

\[
k_n = 0.001 - 2 \times 10^{-5}T \quad (6 - 5)
\]

If one has to find the overall nucleation kinetic equation for the process, the parameters from equations (6-4) and (6-5) can be put into equation (6-2) to get:

\[
J_n = (0.001 - 2 \times 10^{-5}T)\Delta C^{(0.033T + 1.142)} \quad (6 - 6)
\]
Figure 6-5. Diagrams of the plot of logarithm of cooling rate versus logarithm of MSZW for four different saturation temperatures at constant stirring rate (250rpm)
6.3. Polymorphic transformation process

Pure L-glutamic acid (beta form, 99% purity) and monosodium salt of L-glutamic acid (higher than 99%) were purchased from Sigma Aldrich and used as received without further purification. The other form of L-glutamic acid (alpha form) is not common in commercial usage and therefore, we produced it with the method reported in the literature [11, 12]. In order to produce the alpha form, we employed the pH-shift reaction to precipitate the crystals. The monosodium salt was dissolved in a specified volume of ultrapure water and after complete dissolution at around 5°C, the fuming hydrochloric acid (37-38%) was added to the solution drop wise. After complete addition, the precipitated crystals were filtered and dried. To ensure the production of the alpha form, XRPD was employed.

In order to check if the FBRM can clearly detect the shape of the two forms of L-glutamic acid, we made 11 samples of different mass ratios of alpha to beta forms of the component. The samples contained the pure forms of the polymorphs, as well. The mass of the samples were chosen to be the same. All of the samples were added to the saturated solutions of L-glutamic acid to prevent further dissolution. The results of the FBRM measurement after stabilizing, are shown in Figure 6-6.

![Figure 6-6. Particle size distribution for five different samples of alpha to beta L-glutamic acid](image-url)
As it is evident from the Figure 6-6, the samples with the less amount of beta form exhibit the higher counts in larger class chord length (100-250µm). On the other hand, the samples with higher amount of beta form compared to the alpha form show finer chord lengths of particles, specifically in the range of 1-100µm. This confirms the fact that the metastable form has the needle-like shape and therefore, the higher chord length. With this ability of qualitatively detecting the polymorphic forms of L-glutamic acid, we performed the polymorphic transformation of the alpha to beta form. The polymorphic transformation experiments were done at constant temperature of 45°C. At first the saturated solution of beta L-glutamic acid at 45°C was prepared and kept constant at that temperature for 2 hours to ensure complete dissolution. We added the pure alpha form into the solution around 90 min into the run. This is to ensure no solid is dispersed in the solution. Once the solid is added to the vessel, as expected, the number of the counts in the class of 100-250µm started rising and maintained in the range of 3000-4000 counts/s.

![Figure 6-7](image)

**Figure 6-7.** Four distinct regions of crystallization of L-glutamic acid: (a) complete dissolution of stable form, (b) Addition and dissolution of metastable form, (c) Polymorphic transformation, dissolution of metastable form and formation of stable form, and (d) complete dissolution of metastable form and formation of stable form
This region is denoted by B in Figure 6-7. After nearly 300 min from the start of the process, the counts for the range of 100-250 µm fell gradually until the complete dissolution within 100 min.

It is seen that while the number of particles in the range of 100-250µm started to decrease, the counts in the range of 1-22µm (mostly the beta form) started increasing. The polymorphic transformation process happens in region C of the Figure 6-7. Once the whole metastable form is
disappeared, there will be just the stable form and thus, we see the crystallization (growth) of the stable form and increasing the counts of the particles (region D in Figure 6-7).

In another experiment, we did a rapid cooling of the solution of alpha form to transform to beta form. The polymorphic transformation of L-glutamic acid was performed starting at 80°C. First we dissolved 2.5gr of solid in 200ml of ultrapure water and we let the solids to dissolve completely in the solvent. The complete dissolution was verified by FBRM probe. After reaching to this point, the cooler started cooling down to 45°C with a ramp of 1.5°C/min. Soon after reaching to 45°C, a sample from the solution was filtered and the dried by vacuum drying at 60°C. Other samples of the solution were filtered and dried in 1h, 2.5h, and 5h of the process. The dried solids were characterized by x-ray powder diffraction (XRPD) and the results show the polymorphic transformation during the course of the process (Figure 6-8).

The XRPD results in the range of 2-theta between 5-30° is shown in Figure 6-8a. The characteristic peak is between 2-theta of 10-11°. Figure 6-8b magnifies the mentioned range to show the change in the intensities for the two forms in more detail.

6.4. Conclusion

In this chapter, we examined two important aspects of crystallization, nucleation and polymorphic transformation. In both experiments the FBRM spectroscopy was employed to detect the onset of the nucleation and particle chord length in different classes. In the nucleation experiments, the solutions at different saturation temperatures were subjected to different cooling rates to determine the cloud points and then, finding the kinetic parameters of nucleation. We could find the overall nucleation rate equation for the range of temperature that the experiments were done. The polymorphic transformation of the L-glutamic acid at 45°C was performed and qualitatively detected using FBRM probe. The use of FBRM for detection of the polymorphic transformation was validated using 11 samples of different metastable to stable forms. This study proved the ability of FBRM probe in detection of nucleation, polymorphic crystallization processes, and other related applications.
6.5. References


Chapter 7

Conclusions and Future Works
This project successfully addressed the two main challenges that a pharmaceutical and in general, chemical industry is facing. The first challenge is to have a powerful, accurate, and also flexible thermodynamic model in predicting the phase behaviours of systems exhibiting equilibrium between phases. The main focus was to model the solid-liquid equilibrium condition in systems containing a pharmaceutical component in a pure or mixed solvent. Another issue that this project aimed to solve was to model the crystallization systems consisting of polymorphic transformation process. This type of phenomenon has been discussed in the industrial crystallization audience since a few years ago and thus, finding a reliable and fast method of solving the process variables in dynamic mode will help the researchers and industry in predicting the behaviour of such systems and optimizing the operating conditions. The main contribution of this project can be divided into two streams: 1) Applying a new thermodynamic model to a wide variety of industrial applications and verifying the different phase behaviors of the systems involved in pharmaceutical industry and 2) introducing a new combined numerical method of MoMC which could result in a faster and more accurate process simulation of polymorphic transformation systems and then be used in dynamic optimization of the process. In this chapter the brief overview of the important findings in each of the previous chapters is discussed and finally, the future possible works in the two main streams of research in this project will be demonstrated.

7.1. Solid solubility prediction and measurement

The two models of NRTL-SAC and UNIFAC were selected for predicting the solubility of three chemical components in pure and mixed solvents. In order for us to use the NRTL-SAC model, we had to find the four characteristic parameters of each of those components to be used for prediction purposes. Therefore, we selected a portion of the experimental data on different pure and mixed solvents for nonlinear parameter estimation. Then, the NRTL-SAC model was used to predict the rest of the experimental data and thus, validate the method. The common and powerful model of UNIFAC was also employed to compare the results with the NRTL-SAC model. It is worth noting that for the valsartan, the UNIFAC model could not be used, as the functional groups in the chemical structure of the component have not been defined yet. The overall ability of the NRTL-SAC model in prediction of the solubility in different conditions,
directed us to further study the method in modeling other phase behaviour systems which was the subject of chapter 3.

7.2. VLE, and VLLE prediction for industrial solvent systems

Using the two models of activity coefficient prediction mentioned in section 7.1, we could successfully implement the NRTL-SAC model for six different binary, ternary, and quaternary systems of industrial solvents. The results from the model prediction when compared to the UNIFAC model and experimental data, showed a satisfactory application of the proposed model. The results of this chapter, especially the bubble point calculation were used as a constraint nonlinear function in chapter 4.

7.3. Solvent screening process

We developed a novel algorithm of finding the best combination of solvents in binary and ternary mixture of solvents for a given crystallization process. The single solvent investigation was also performed using the NRTL-SAC method as a thermodynamic framework for prediction purposes. The resulted package showed its strength in modeling complex and time-intensive process of solvent screening for seven selected model components. It should be noted that the outcomes from chapters 2 and 3 were all used to support the calculation procedure in this chapter.

7.4. Polymorphic transformation process modeling and optimal control

In this chapter we developed a novel method of combining methods of moment and classes together to take advantage of both methods. The resulting method could be used to generate CSD in addition to the moments of particles and other necessary variables in the system. The MoMC method was used in order to model the batch cooling crystallization of a polymorphic transformation system in different conditions and with a variety of seed sizes and masses. The method was also applied to an optimal control procedure to get the optimal trajectory in the course of crystallization for three important objective functions used in the pharmaceutical industry. The outcome of this chapter was to build a comprehensive model to include all possible scenarios that happen in the crystallization of a polymorphic transformation system. The optimal
control method that applied to the model, showed an improvement in the performance of the crystallizer for different objective functions.

### 7.5. Experimental results for crystallization of L-glutamic acid

In chapter 5, we used the dynamical model of the polymorphic crystallization system to predict the behaviour of the system in different conditions. The experimental part of the crystallization process, including the nucleation detection, nucleation kinetic study, and polymorphic transformation process were all performed in chapter 6. The use of particle chord length measurement probe (FBRM) in detecting the onset of nucleation and disappearance (saturation point) was shown. The FBRM also showed its capability in detecting the qualitative transformation from metastable to stable form of L-glutamic acid at constant temperature. The polymorphic transformation of this model compound was also detected by offline method of XRPD using different samples from the system at various times of the process.

### 7.6. Future Works

The results from this project showed that there are more works that are needed to be performed in the mentioned areas of solution thermodynamics and polymorphic crystallization. For the first part of study, the use of the NRTL-SAC model was comprehensively studied. However, there are some points which need to be addressed in solubility prediction of pharmaceutical components:

- The presence of impurities in the solvent or solid, which may affect solid solubility
- The presence of two liquid phases with a solid dissolved in the liquid phase

In chapter 3, six important solvent systems that are common in industry were modeled using the NRTL-SAC model. There are some works that can be done in this field, such as:

- The study on other systems that exhibit azeotropic behaviors using this model and other activity coefficient based methods
In the solvent screening process, there are many factors that can affect the objective functions defined in chapter 4. Therefore, the study of those factors can be beneficial to improve the accuracy of the model:

- Some solvents and solids that are used in our study may react with each other at some conditions (e.g., temperature). Therefore, having a precise knowledge of the reaction between the combinations of solvents can be included in the program and may significantly help the researchers in doing the experimental work safely and without loss of materials.
- In the current study, the path of crystallization from the initial to final conditions is not studied. As it was stated in chapter 4, the assumption was made to consider just two points of operation (initial and final state). However, in order to have a deeper understanding of the process, one can include the population balance equations with required nucleation and growth kinetics in the calculation. It should be noted that in order to achieve this, a wide range of experimental data on kinetics of nucleation and growth are needed for each system of solvents.
- In addition to the two objective functions that are used in our project, one can define other cost functions to be optimized.

In chapter 5, we studied the polymorphic transformation of L-glutamic acid using method of moments and also the novel method of MoMC. A few assumptions were made for making the calculation process easier and faster. In some crystallization processes, the agglomeration and breakage are common and thus, neglecting their effect on process modeling can result in large deviations from reality. Therefore, including the effect of other processes, such as breakage and agglomeration in the model, can give a more accurate result for some systems.

The objective functions that were defined for optimal control procedure come from industrial point of view for crystallization processes. There are other objective functions, such as PSD distribution, which can be studied using the novel method of MoMC, as this is very fast in calculation compared to the other common methods, such as MoC.

In the area of experimentation, there are some works that need to be done in the field of polymorphic crystallization process and optimal control. Some examples include the use of
Raman to detect online composition of polymorphs during the process or application of FBRM combined with Raman to detect the polymorphic transformation with more accuracy.
Appendixes
In the appendix, the codes that were used in all parts of the project are included. It should be noted that depending on the objective in the project, a few parts of the codes needs to be changed. These changes are not including the main program, but rather contain the parameter values, the number of components in the system, and other minor changes. Therefore, the mentioned codes can be successfully used for the purposes of fluid phase behaviour calculation and crystallization of different polymorphic systems in dynamic mode. **A. Matlab codes for thermodynamic models**

**A.1. NRTL-SAC model**

```matlab
function Solubility = Sol(T,y,r,delta_S_R,T_melting,MW)
N = numel(r(1,:));
Tau = [0,1.643,1.643,6.547;1.834,0,0,-2;1.834,0,0,2;10.949,1.787,1.787,0];  %---- Tau (NRTL binary parameters for conceptual segments)
alpha = [0,.2,.2,.2;.2,0,0,.3;.2,0,0,.3;.2,.3,.3,0];
x(1) = .5;
if (N>=3)
    F(1:N-2) = 1;
    for i=1:N-2
        for j=i:N-2
            F(i) = F(i)*y(j);
        end
    end
else
    F = 0;
end
SUM_SOLVENT = sum(F) + 1;
x(N) = (1-x(1))/SUM_SOLVENT;
if (N>=3)
    for i = N-1:-1:2
        x(i) = x(i+1)*y(i-1);
    end
end
Error = 1;
counter = 0;
while (Error>0.000000001)
    R(1:N) = 0;
    for j=1:N
        for i=1:4
            R(j) = R(j) + r(i,j);
        end
    end
    phi(1:N) = 0;
    for j=1:N
        for i=1:N
            phi(j) = phi(j) + (R(i)*x(i));
        end
        phi(j) = R(j)*x(j)/phi(j);
    end
    L_Gamma_C(1:N)=0;
end
```

153
for i=1:N
    for j=1:N
        L_Gamma_C(i) = L_Gamma_C(i) - R(i)*(phi(j)/R(j));
    end
    L_Gamma_C(i) = L_Gamma_C(i) + log(phi(i)/x(i)) + 1;
end

%---------------- Residual part
%---------------- local binary quantities ---------------
for i=1:4
    for j=1:4
        G(i,j) = exp(-alpha(i,j)*Tau(i,j));
    end
end

%----------- Activity coefficient of segment species m (Gamma(m)) -------
Sum = 0;
for i=1:4
    for j=1:N
        Sum = Sum + x(j)*r(i,j);
    end
end
X(1:4) = 0;
for j=1:4
    for z=1:N
        X(j) = X(j) + (x(z)*r(j,z))/Sum;
    end
end
for j=1:N
    for i=1:4
        XX(i,j) = r(i,j)/R(j);
    end
end
for j=1:N
    for i=1:4
        Sum_1(j) = Sum_1(j) + X(i)*G(i,j)*Tau(i,j);
        Sum_2(j) = Sum_2(j) + X(i)*G(i,j);
    end
end
SUM(1:4,1:4) = 0;
SUM_1(1:4,1:4) = 0;
for j=1:4
    SUM_1(j) = X(1)*G(1,j) + X(2)*G(2,j) + X(3)*G(3,j) + X(4)*G(4,j);
    SUM_1(j) = X(1)*G(1,j)*Tau(1,j) + X(2)*G(2,j)*Tau(2,j) + X(3)*G(3,j)*Tau(3,j) +
              X(4)*G(4,j)*Tau(4,j);
end
for k=1:N
    L_Gamma_m_lc = (Sum_1./Sum_2 + Sum_3);
end
for j=1:4
    for i=1:4
        Sum_1_I(j,k) = Sum_1_I(j,k) + XX(i,k)*G(i,j)*Tau(i,j);
        Sum_2_I(j,k) = Sum_2_I(j,k) + XX(i,k)*G(i,j);
    end
end
end
for i=1:N
    for j=1:4
        SUM_I(j,i) = XX(1,i)*G(1,j) + XX(2,i)*G(2,j) + XX(3,i)*G(3,j) + XX(4,i)*G(4,j);
    end
end
for k=1:N
    for i=1:4
        for j=1:4
            Sum_3_I(j,k) = Sum_3_I(j,k) + XX(i,k)*G(j,i)/SUM_I(i,k)*(Tau(j,i) - SUM_I_1(i,k)/SUM_I(i,k));
        end
    end
end
L_Gamma_m_lc_I = (Sum_1_I./Sum_2_I + Sum_3_I);
L_Gamma_R(1:N) = 0;
for j = 1:N
    for i = 1:4
        L_Gamma_R(j) = L_Gamma_R(j) + r(i,j)*(L_Gamma_m_lc(i) - L_Gamma_m_lc_I(i,j));
    end
end
Gamma = exp(L_Gamma_C + L_Gamma_R);
Ln_x = delta_S_R.*(1 - T_melting./T) - log(Gamma(1));
x_new(1) = exp(Ln_x);
x_new(N) = (1 - x_new(1))/SUM_SOLVENT;
if (N>=3)
    for i = N-1:-1:2
        x_new(i) = x_new(i+1)*y(i-1);
    end
end
Error = abs(x(1) - x_new(1));
x = x_new;
counter = counter + 1;
if counter> 2000
    x = 0;
    break;
end
M_solution = 0;
for i=1:(numel(x)-1)
    M_solution = M_solution + x(i+1)/sum(x(2:end))*MW(i+1);
end
Solubility = x(1)/(1-x(1))*MW(1)/M_solution;
if (Solubility<0)
    Solubility = NaN;
end
A.2. UNIFAC model

function L_Gamma = UNIFAC_new(T,x,a,R,Q,v,K)

NC = nnz(x);

r(1:NC) = 0;
q(1:NC) = 0;
for i=1:NC
    for j=1:K
        r(i) = r(i) + R(j)*v(j,i);
        q(i) = q(i) + Q(j)*v(j,i);
    end
end

L(1:NC) = 0;
Phi(1:NC) = 0;
Teta(1:NC) = 0;
SUM_Phi = 0;
SUM_Teta = 0;
SUM_L = 0;
for i=1:NC
    L(i) = 5*(r(i) - q(i)) - (r(i) - 1);
    SUM_Phi = SUM_Phi + x(i)*r(i);
    SUM_Teta = SUM_Teta + x(i)*q(i);
    SUM_L = SUM_L + x(i)*L(i);
end
for i=1:NC
    Phi(i) = x(i)*r(i)/SUM_Phi;
    Teta(i) = x(i)*q(i)/SUM_Teta;
end

%-------------- Combinatorial part of Gamma   --------------
L_Gamma_C(1:NC) = 0;
for i=1:NC
    L_Gamma_C(i) = log(Phi(i)/x(i)) + 5*q(i)*log(Teta(i)/Phi(i)) + L(i) - Phi(i)/x(i)*SUM_L;
end
%----------------------------------------------------------
TETA(1:K) = 0;
SUM_TETA = 0;
X(1:K) = 0;
SUM_X(1:K) = 0;
SUM_X_mixture = 0;
SAI(1:K,1:K) = 0;
SUM_SAI(1:K) = 0;
SUM_SAI_SAI(1:K) = 0;
for i=1:K
    for j=1:NC
        SUM_X(i) = SUM_X(i) + v(i,j)*x(j);
        SUM_X_mixture = SUM_X_mixture + v(i,j)*x(j);
    end
end
for i=1:K
    X(i) = SUM_X(i)/SUM_X_mixture;
    SUM_TETA = SUM_TETA + Q(i)*X(i);
end
for i=1:K
    TETA(i) = Q(i)*X(i)/SUM_TETA;
end
for i=1:K
    for j=1:K
        SAI(i,j) = exp(-a(i,j)/T);
    end
end
for i=1:K  
for j=1:K  
    SUM_SAI(i) = SUM_SAI(i) + TETA(j)*SAI(j,i);  
end  
end  

for i=1:K  
for j=1:K  
    SUM_SAI_SAI(i) = SUM_SAI_SAI(i) + TETA(j)*SAI(i,j)/SUM_SAI(j);  
end  
end  

L_Gamma_R_K(1:K) = 0;  
for i=1:K  
    L_Gamma_R_K(i) = Q(i)*(1 - log(SUM_SAI(i)) - SUM_SAI_SAI(i));  
end  

%--------------------  For calculation of Ln_Gamma in component i for  
%--------------------  subgroup j  
TETA_Component(1:K,1:NC) = 0;  
SUM_TETA_Component(1:NC) = 0;  
X_Component(1:K,1:NC) = 0;  
SUM_X_mixture_Component(1:NC) = 0;  
SUM_SAI_Component(1:K,1:NC) = 0;  
SUM_SAI_SAI_Component(1:K,1:NC) = 0;  

for i=1:NC  
    for j=1:K  
        SUM_X_mixture_Component(i) = SUM_X_mixture_Component(i) + v(j,i);  
    end  
end  

for j=1:K  
    for i=1:NC  
        X_Component(j,i) = v(j,i)/SUM_X_mixture_Component(i);  
    end  
end  

for i=1:NC  
    for j=1:K  
        SUM_TETA_Component(i) = SUM_TETA_Component(i) + Q(j)*X_Component(j,i);  
    end  
end  

for j=1:K  
    for i=1:NC  
        TETA_Component(j,i) = Q(j)*X_Component(j,i)/SUM_TETA_Component(i);  
    end  
end  

for i=1:NC  
    for j=1:K  
        for z=1:K  
            SUM_SAI_Component(j,i) = SUM_SAI_Component(j,i) + TETA_Component(z,i)*SAI(z,j);  
        end  
    end  
end  

for i=1:NC  
    for j=1:K  
        for z=1:K  
            SUM_SAI_SAI_Component(j,i) = SUM_SAI_SAI_Component(j,i) + TETA_Component(z,i)*SAI(j,z)/SUM_SAI_Component(z,i);  
        end  
    end  
end  

L_Gamma_R_K_Component(1:K,1:NC) = 0;  
for i=1:NC  
    for j=1:K  

L_Gamma_R_K_Component(j,i) = Q(j)*(1 - log(SUM_SAI_Component(j,i)) - SUM_SAI_SAI_Component(j,i));
end
end

L_Gamma_R(1:NC) = 0;
for i=1:NC
    for j=1:K
        L_Gamma_R(i) = L_Gamma_R(i) + v(j,i)*(L_Gamma_R_K(j) - L_Gamma_R_K_Component(j,i));
    end
end

L_Gamma = exp(L_Gamma_R + L_Gamma_C);

A.3. Bubble point calculation for the NRTL-SAC model

function out = BUBBLE_NRTLSAC(P_tot,x,r,Antoine)
    N = numel(x);
    A = Antoine(1:N,1);
    B = Antoine(1:N,2);
    C = Antoine(1:N,3);
    G = gamma(x,r);
    for i=1:N
        T_sat(i) = B(i)/(A(i) - log10(P_tot)) - C(i);
    end
    T_sol = 0;
    for i=1:N
        T_sol = T_sol + x(i)*T_sat(i);
    end
    T0 = T_sol;
    options=optimset('Display','iter');
    T = fsolve(@(T)fun(T,Antoine,x,G,P_tot),T0,options);
    for i=1:N
        P_sat(i) = 10^(A(i) - B(i)/(T + C(i)));
    end
    for i=1:N
        y(i) = x(i)*G(i)*P_sat(i)/P_tot;
    end
    out = [T,y];
end

A.4. Dew point calculation

function out = DEW(P_tot,y,r,Antoine)
    N = numel(y);
    A = Antoine(1:N,1);
    B = Antoine(1:N,2);
    C = Antoine(1:N,3);
    for i=1:N
        T_sat(i) = B(i)/(A(i) - log10(P_tot)) - C(i);
    end
    T_sol = 0;
    for i=1:N
        T_sol = T_sol + y(i)*T_sat(i);
    end
    T0 = T_sol;
    options=optimset('Display','iter');
    T = fsolve(@(T)fun(T,Antoine,x,G,P_tot),T0,options);
    for i=1:N
        P_sat(i) = 10^(A(i) - B(i)/(T + C(i)));
    end
    for i=1:N
        y(i) = x(i)*G(i)*P_sat(i)/P_tot;
    end
    out = [T,y];
end
P_sat(i) = 10^{(A(i) - B(i)/(T_sol + C(i)))};
end

P_sat_new = 0;
for i=1:N
    P_sat_new = P_sat_new + P_tot*(y(i)*P_sat(1)/(P_sat(i)));
end

T = B(1)/(A(1) - log10(P_sat_new(1))) - C(1);
for i=1:N
    P_sat(i) = 10^{(A(i) - B(i)/(T + C(i)))};
end

for i=1:N
    x(i) = y(i)*P_tot/P_sat(i);
end
G = gamma(x,r);
P_sat_new = 0;
for i=1:N
    P_sat_new = P_sat_new + P_tot*(y(i)*P_sat(1)/(G(i)*P_sat(i)));
end

T_old = B(1)/(A(1) - log10(P_sat_new)) - C(1);
error_1 = 1;
while (error_1>0.00000001)
    gamma_old = gamma(x,r);
    error_2 = 1;
    for i=1:N
        P_sat(i) = 10^{(A(i) - B(i)/(T_old + C(i)))};
    end
    while (error_2>0.000000001)
        for i=1:N
            x(i) = y(i)*P_tot/(gamma_old(i)*P_sat(i));
        end
        %--- x's are normalized
        S = sum(x);
        for i=1:N
            x(i) = x(i)/S;
        end
        gamma_new = gamma(x,r);
        error_2 = sum(abs(gamma_new - gamma_old));
        gamma_old = gamma_new;
    end
    G = gamma(x,r);
P_sat_new = 0;
    for i=1:N
        P_sat_new = P_sat_new + P_tot*(y(i)*P_sat(1)/(G(i)*P_sat(i)));
    end
    T_new = B(1)/(A(1) - log10(P_sat_new)) - C(1);
    error_1 = abs(T_new - T_old);
    T_old = T_new;
end
T = T_new;
out = [T,x];
A.5. Bubble point calculation for the UNIFAC model

```matlab
function out = Bubble_UNIFAC(P_tot,x,Antoine,a_parameter,R,Q,v,K)

N = numel(x);
A = Antoine(1:N,1);
B = Antoine(1:N,2);
C = Antoine(1:N,3);

for i=1:N
    T_sat(i) = B(i)/(A(i) - log10(P_tot)) - C(i);
end

T_sol_old = 0;
for i=1:N
    T_sol_old = T_sol_old + x(i)*T_sat(i);
end

error = 1;
j = 1;
while (error > 0.00000001)
    G = UNIFAC_new(T_sol_old,x,a_parameter,R,Q,v,K);
    for i=1:N
        P_sat(i) = 10^(A(i) - B(i)/(T_sol_old + C(i)));
    end
    for i=1:N
        y(i) = x(i)*G(i)*P_sat(i)/P_tot;
    end
    Sum = 0;
    for i=1:N
        Sum = Sum + x(i)*G(i)*P_sat(i)/P_sat(j);
    end
    P_sat(j) = P_tot/Sum;
    T_sol_new = B(j)/(A(j) - log10(P_sat(j))) - C(j);
    error = abs(T_sol_new - T_sol_old);
    T_sol_old = T_sol_new;
end
out = [T_sol_old,y];
```

B. Matlab codes for solvent screening

B.1. Optimization for single solvent

```matlab
clc
clear all

load('solvent_segment_number');
load('solid');
P_tot = 1.013;
count = 1;
for j=2:9
    for i=1:62
        T_sat(i) = r_solvent(i,7)/(r_solvent(i,6) - log10(P_tot)) - r_solvent(i,8);
        T_max(i) = 40 + 273.15;
        T_min(i) = 20 + 273.15;
        if T_max(i) >= solute(j,1) - 5 || T_min(i) >= solute(j,1) - 5 || T_max(i) > T_sat(i) - 10
            Mass_per_solvent_1(j,i) = 0;
            Mass_per_solvent_2(j,i) = 0;
        else
            Mass_per_solvent_1(j,i) = Sol(T_max(i),0,[solute(j,4:7)' r_solvent(i,1:4)'
                solute(j,3),solute(j,1),[solute(j,8) r_solvent(i,5)]'),...
                solute(j,3),solute(j,1),[solute(j,8) r_solvent(i,5)]);
        end
    end
end
```
\[
\text{Mass per solvent}_2(j,i) = (\text{Sol}(T_{\text{max}}(i),0,\text{solute}(j,4:7)'r_{\text{solvent}}(i,1:4)',\text{solute}(j,3),\text{solute}(j,1),\text{solute}(j,8) r_{\text{solvent}}(i,5))) - ...
\]
\[
\text{Sol}(T_{\text{min}}(i),0,\text{solute}(j,4:7)'r_{\text{solvent}}(i,1:4)',\text{solute}(j,3),\text{solute}(j,1),\text{solute}(j,8) r_{\text{solvent}}(i,5)))
\]
\[
S_1(j,i) = \text{Sol}(T_{\text{max}}(i),0,\text{solute}(j,4:7)'r_{\text{solvent}}(i,1:4)',\text{solute}(j,3),\text{solute}(j,1),\text{solute}(j,8) r_{\text{solvent}}(i,5))
\]
\[
S_2(j,i) = \text{Sol}(T_{\text{min}}(i),0,\text{solute}(j,4:7)'r_{\text{solvent}}(i,1:4)',\text{solute}(j,3),\text{solute}(j,1),\text{solute}(j,8) r_{\text{solvent}}(i,5));
\]
\[
\text{if} \ (\text{Mass per solvent}_1(j,i)<0)
\]
\[
\text{Mass per solvent}_1(j,i) = 0;
\]
\[
\text{elseif} \ (\text{Mass per solvent}_2(j,i)<0)
\]
\[
\text{Mass per solvent}_2(j,i) = 0;
\]
\[
\text{end}
\]
\[
[C_1, I_1] = \max(\text{Mass per solvent}_1(j,:));
\]
\[
[C_2, I_2] = \max(\text{Mass per solvent}_2(j,:));
\]
\[
\text{Max}_1(j,:) = [C_1, I_1];
\]
\[
\text{Max}_2(j,:) = [C_2, I_2];
\]
\[
count = \text{count} + 1
\]
\[
\text{end}
\]
\[
count = 1;
\]
\[
i = 1;
\]
\[
\text{for} \ T = 273+15:273+50
\]
\[
\text{SS}_1(i) = \text{Sol}(T,0,\text{solute}(3,4:7)'r_{\text{solvent}}(51,1:4)',\text{solute}(3,3),\text{solute}(3,1),\text{solute}(3,8) r_{\text{solvent}}(51,5));
\]
\[
\text{SS}_2(i) = \text{Sol}(T,0,\text{solute}(3,4:7)'r_{\text{solvent}}(59,1:4)',\text{solute}(3,3),\text{solute}(3,1),\text{solute}(3,8) r_{\text{solvent}}(59,5));
\]
\[
i = i + 1;
\]
\[
\text{end}
\]
\[
T = 15:50;
\]
\[
\text{plot}(T,\text{SS}_1);
\]
\[
\text{hold on}
\]
\[
\text{plot}(T,\text{SS}_2,'--');
\]
\[
\]
B.2. Optimization for binary solvent combination

clc
clear all

load('solvent_segment_number');
load('solid');
P_tot = 1.013;
starting_solvent = 1;
ending_solvent = 62;
solute_ID = 9;
tic;
for i = starting_solvent:ending_solvent-1
  for j = i + 1:ending_solvent
    x0 = [273+40;1;273+20;1];
    lb = [273.15+30;0;273.15+15;0];
    ub = [273.15+min(solute(solute_ID,1)-273.15-5,65);1000;273.15+min(solute(solute_ID,1)-273.15-5,20);1000];
    A = [-1,0,1,0]; b = 0;
    \[
    \text{x,fval} = \text{fmincon}(@x(-1).*\text{binary}(x,\text{solute}_ID,1),x0,A,b,[]),lb,ub,\text{@}(x)\text{binary}_\text{const}(x,P_t,\text{solute}(i,1:4)'r_{\text{solvent}}(i,1:4)',\text{solute}(i,6:8) r_{\text{solvent}}(j,6:8)));
    \]
    Initial_Temperature(i,j) = x(1);
    Final_Temperature(i,j) = x(3);
    Initial_Solvent_Ratio(i,j) = x(2);
    Final_Solvent_Ratio(i,j) = x(4);
    Yield_per_Solvent(i,j) = fval;
  end
end
toc;
CPU_time = toc/3600;
function F = binary(y,solute_ID,solvent_1_ID,solvent_2_ID)
load('solvent_segment_number');
load('solid');
F = (Sol(y(1),y(2),[solute(solute_ID,4:7)' r_solvent(solvent_1_ID,1:4)'
 r_solvent(solvent_2_ID,1:4)'],solute(solute_ID,3),solute(solute_ID,1),[solute(solute_ID,8)
 r_solvent(solvent_1_ID,5) r_solvent(solvent_2_ID,5)]))...
 - Sol(y(3),y(4),[solute(solute_ID,4:7)' r_solvent(solvent_1_ID,1:4)'
 r_solvent(solvent_2_ID,1:4)'],solute(solute_ID,3),solute(solute_ID,1),[solute(solute_ID,8)
 r_solvent(solvent_1_ID,5) r_solvent(solvent_2_ID,5)]));
end

B.2.1. Constraints for binary solvent combination

function [c,ceq] = binary_const(x,P_tot,r,Antoine)
c(1) = x(1) - 10 - BUBBLE_NRTLSAC(P_tot,[x(2)/(1+x(2));1/(1+x(2))],r,Antoine);
c(2) = x(3) - 10 - BUBBLE_NRTLSAC(P_tot,[x(4)/(1+x(4));1/(1+x(4))],r,Antoine);
ceq = 0;
end

B.3. Optimization code for ternary solvent combination

clc
clear all
load('solvent_segment_number_1');
load('solid');
P_tot = 1.013;
% solvent_1_ID = 5;
% solvent_2_ID = 6;
starting_solvent = 1;
ending_solvent = 30;
tic;
% count_first_solvent = 1;
% count_second_solvent = 1;
z = 1;
for solute_ID = 9:9
    for i = starting_solvent:ending_solvent - 2
        for j = i + 1:ending_solvent
            for k = j + 1:ending_solvent
                x0 = [273+40;1;1;273+20;1;1];
                lb = [273.15+30;0;0;273.15+15;0;0];
                ub = [273.15+min(solute(solute_ID,1)-273.15-
                    5,65);500;500;273.15+min(solute(solute_ID,1)-273.15-5,20);500;500];
                A = [-1,0,0,1,0,0]; b= 0;
                [x,fval] = fmincon(@(x)(-
                    1).*ternary(x,solute_ID,i,j,k),x0,A,b,[],[],lb,ub,@(x)ternary_const(x,P_tot,[r_solvent_1(i,1:4)',
                    r_solvent_1(j,1:4)', r_solvent_1(k,1:4)'],[solute(solute_ID,4:7)' r_solvent_1(i,1:4)'
                    r_solvent_1(j,1:4)'],solute(solute_ID,3),solute(solute_ID,1),[solute(solute_ID,8)
                    r_solvent_1(i,5) r_solvent_1(j,5) r_solvent_1(k,5)]));
                Initial = Sol(x(1),[x(2),x(3)],[solute(solute_ID,4:7)' r_solvent_1(i,1:4)'
                    r_solvent_1(j,1:4)']);
                Final = Sol(x(4),[x(5),x(6)],[solute(solute_ID,4:7)' r_solvent_1(i,1:4)'
                    r_solvent_1(j,1:4)']);
                Initial_Temperature(i,j,k,z) = x(1);
Final_Temperature(i,j,k,z) = x(4);
Initial_Solvent_Ratio_1(i,j,k,z) = x(2);
Initial_Solvent_Ratio_2(i,j,k,z) = x(3);
Final_Solvent_Ratio_1(i,j,k,z) = x(5);
Final_Solvent_Ratio_2(i,j,k,z) = x(6);
Yield_per_solvent_1(i,j,k,z) = -fval;
end

% count_first_solvent = count_first_solvent + 1;
end
toc;
CPU_time(z) = toc/3600;
z = z + 1;
tic;
end

for solute_ID = 1:1:
[c,i] = max(Yield_per_solvent_1);
[t,j] = max(c);
y,k] = max(t);
Index(solute_ID,:) = [ i(:,j(:,k(:,solute_ID)),solute_ID),k(:,solute_ID),solute_ID];
max_solubility(solute_ID) =
Yield_per_solvent_1(Index(solute_ID,1),Index(solute_ID,2),Index(solute_ID,3),solute_ID);
Initial_Temp(solute_ID) =
Initial_Temperature(Index(solute_ID,1),Index(solute_ID,2),Index(solute_ID,3),solute_ID);
Final_Temp(solute_ID) =
Final_Temperature(Index(solute_ID,1),Index(solute_ID,2),Index(solute_ID,3),solute_ID);
Initial_Sol_Ratio(solute_ID,:) =
[Initial_Solvent_Ratio_1(Index(solute_ID,1),Index(solute_ID,2),Index(solute_ID,3),solute_ID),
Initial_Solvent_Ratio_2(Index(solute_ID,1),Index(solute_ID,2),Index(solute_ID,3),solute_ID)];
Initial_Volume_Ratio(solute_ID,:) =
[Initial_Solvent_Ratio_1(Index(solute_ID,1),Index(solute_ID,2),Index(solute_ID,3),solute_ID)*r_solvent_1(Index(solute_ID,1),5)/r_solvent_1(Index(solute_ID,2),5)*r_solvent_1(Index(solute_ID,3),5)*r_solvent_1(Index(solute_ID,3),10)/r_solvent_1(Index(solute_ID,1),10)
Initial_Solvent_Ratio_2(Index(solute_ID,1),Index(solute_ID,2),Index(solute_ID,3),solute_ID)*r_solvent_2(Index(solute_ID,1),5)/r_solvent_1(Index(solute_ID,2),5)*r_solvent_1(Index(solute_ID,3),5)*r_solvent_1(Index(solute_ID,3),10)/r_solvent_1(Index(solute_ID,1),10)]
Initial_Volume_Ratio(solute_ID,:) =
[Initial_Solvent_Ratio_1(Index(solute_ID,1),Index(solute_ID,2),Index(solute_ID,3),solute_ID)*r_solvent_1(Index(solute_ID,1),5)/r_solvent_1(Index(solute_ID,2),5)*r_solvent_1(Index(solute_ID,3),5)*r_solvent_1(Index(solute_ID,3),10)/r_solvent_1(Index(solute_ID,1),10)
Initial_Solvent_Ratio_2(Index(solute_ID,1),Index(solute_ID,2),Index(solute_ID,3),solute_ID)*r_solvent_2(Index(solute_ID,1),5)/r_solvent_1(Index(solute_ID,2),5)*r_solvent_1(Index(solute_ID,3),5)*r_solvent_1(Index(solute_ID,3),10)/r_solvent_1(Index(solute_ID,1),10)]
for i=1:11
V_initial(i,:) = [Initial_Volume_Ratio(i,1,1)*Initial_Volume_Ratio(i,1,2)*1000./(1 + Initial_Volume_Ratio(i,1,2)*Initial_Volume_Ratio(i,1,1) + Initial_Volume_Ratio(i,1,2));
Initial_Volume_Ratio(i,1,2)*1000./(1 + Initial_Volume_Ratio(i,1,2)*Initial_Volume_Ratio(i,1,1) + Initial_Volume_Ratio(i,1,2));
V_initial(i,:) = [Initial_Volume_Ratio(i,1,1)*Initial_Volume_Ratio(i,1,2)*1000./(1 + Initial_Volume_Ratio(i,1,2)*Initial_Volume_Ratio(i,1,1) + Initial_Volume_Ratio(i,1,2));
Initial_Volume_Ratio(i,1,2)*1000./(1 + Initial_Volume_Ratio(i,1,2)*Initial_Volume_Ratio(i,1,1) + Initial_Volume_Ratio(i,1,2));
Initial_Volume_Ratio(i,1,2)*1000./(1 + Initial_Volume_Ratio(i,1,2)*Initial_Volume_Ratio(i,1,1) + Initial_Volume_Ratio(i,1,2));
V_final(i,:) = [Initial_Volume_Ratio(i,1,1)*Initial_Volume_Ratio(i,1,2)*1000./(1 + Initial_Volume_Ratio(i,1,2)*Initial_Volume_Ratio(i,1,1) + Initial_Volume_Ratio(i,1,2));
Initial_Volume_Ratio(i,1,2)*1000./(1 + Initial_Volume_Ratio(i,1,2)*Initial_Volume_Ratio(i,1,1) + Initial_Volume_Ratio(i,1,2));
Initial_Volume_Ratio(i,1,2)*1000./(1 + Initial_Volume_Ratio(i,1,2)*Initial_Volume_Ratio(i,1,1) + Initial_Volume_Ratio(i,1,2));
end

function F = ternary(y,solute_ID,solvent_1_ID,solvent_2_ID,solvent_3_ID)
load('solvent_segment_number');
load('solvent_segment_number_1');
load('solid');
F = Sol(y(1), [y(2), y(3)], [solute(solute_ID, 4:7)' r_solvent_1(solvent_1_ID, 1:4)'
  r_solvent_1(solvent_3_ID, 1:4)'], solute(solute_ID, 3), solute(solute_ID, 1), [solute(solute_ID, 8)
  r_solvent_1(solvent_1_ID, 5) r_solvent_1(solvent_2_ID, 5) r_solvent_1(solvent_3_ID, 5)])
  - Sol(y(4), [y(5), y(6)], [solute(solute_ID, 4:7)' r_solvent_1(solvent_1_ID, 1:4)'
  r_solvent_1(solvent_3_ID, 1:4)'], solute(solute_ID, 3), solute(solute_ID, 1), [solute(solute_ID, 8)
  r_solvent_1(solvent_1_ID, 5) r_solvent_1(solvent_2_ID, 5) r_solvent_1(solvent_3_ID, 5)));
end
function F = ternary_1(y, solute_ID, solvent_1_ID, solvent_2_ID, solvent_3_ID)
load('solvent_segment_number');
load('solid');
F = (Sol(y(1), [y(2), y(3)], [solute(solute_ID, 4:7)' r_solvent_1(solvent_1_ID, 1:4)'
  r_solvent_1(solvent_3_ID, 1:4)'], solute(solute_ID, 3), solute(solute_ID, 1), [solute(solute_ID, 8)
  r_solvent_1(solvent_1_ID, 5) r_solvent_1(solvent_2_ID, 5) r_solvent_1(solvent_3_ID, 5)])
  - Sol(y(4), [y(5), y(6)], [solute(solute_ID, 4:7)' r_solvent_1(solvent_1_ID, 1:4)'
  r_solvent_1(solvent_3_ID, 1:4)'], solute(solute_ID, 3), solute(solute_ID, 1), [solute(solute_ID, 8)
  r_solvent_1(solvent_1_ID, 5) r_solvent_1(solvent_2_ID, 5) r_solvent_1(solvent_3_ID, 5)));
end

B.3.1. Constraints for ternary solvent combination

function [c, ceq] = ternary_const(x, P_tot, r, Antoine)
c(1) = x(1) - 10 -
  BUBBLE_NRTLAC(P_tot, [x(2)*x(3)/(1+x(2)*x(3)+x(3)), x(3)/(1+x(2)*x(3)+x(3)), 1/(1+x(2)*x(3)+x(3))],
  r, Antoine);
c(2) = x(4) - 10 -
  BUBBLE_NRTLAC(P_tot, [x(5)*x(6)/(1+x(5)*x(6)+x(6)), x(6)/(1+x(5)*x(6)+x(6)), 1/(1+x(5)*x(6)+x(6))],
  r, Antoine);
  ceq = [];
end

C. Matlab codes for MOMC of polymorphic transformation

C.1. Batch crystallization simulation for polymorphic system

function ANS = batch_crystal(T_F1, T_F2, S, U, U1, a, m, L0, LN, Ostwald_Factor, seed_constant_stable, seed_constant_unstable, seed_STD_stable, seed_STD_unstable, seed_ave_size_stable, seed_ave_size_unstable, rho_stable, rho_unstable, rho_solvent, kv_stable, kv_unstable, C_0, T_0, batch_time, KB1, KB2, KBB1, KBB2, KB_1, KD, ID, KD_1, KGI, KG2, KG_1, G_1)
  %------------------------------------------------------------------
  %  ANS(t,1) = zeroth moment metastable
  %  ANS(t,2) = first moment metastable
  %  ANS(t,3) = second moment metastable
  %  ANS(t,4) = third moment metastable
  %  ANS(t,5) = zeroth moment stable
  %  ANS(t,6) = first moment stable
  %  ANS(t,7) = second moment stable
  %  ANS(t,8) = third moment stable
  %  ANS(t,9) = concentration

164
% ANS(t,10) = fourth moment metastable
% ANS(t,11) = fourth moment stable
% ANS(t,m+12 - 2*m+11) = number of particles in each interval for metastable
% ANS(t,m+12 - m+11) = number of particles in each interval for metastable

% a = length(Temperature);
% m = 500;
% L0 = 0;     %--- [micrometer]
% LN = 150*10^-6;   %--- [micrometer]
% ostwald_factor = 0;

% -- seeding parameters

% lamba = [8.227*10^7 , 3.877*10^8 , 3.731*10^8 ; 2.483*10^10 , 1.63*10^10 , 1.501*10^10];
% sigma = [8.608 , 12.127 , 12.115 ; 27.289 , 27.989 , 28.131]*10^-6;
% miu = [214.977 , 127.269 , 127.427 ; 155.069 , 155.017 , 155.007]*10^-6;
CSD_Alpha(1,:) = seed_csd(seed_constant_unstable,seed_STD_unstable,seed_ave_size_unstable,m,L0,LN)';
CSD_Beta(1,:) = seed_csd(seed_constant_stable,seed_STD_stable,seed_ave_size_stable,m,L0,LN)';

%-----------------------
% rho_crystal = 1540;
% alpha_v = .031;
% rho_crystal_1 = 1540;
% alpha_v_1 = .48;
% rho_solvent = 990;
phi = (LN-L0)/m;
f0 = CSD_Beta(1,:);
f0_1 = CSD_Alpha(1,:);
L=(L0+phi/2):phi:(LN-phi/2);

%----------------------- For the stable polymorph

%----- Initial conditions
miu_0 = trapz(L,f0);
miu_1 = trapz(L,f0.*L);
miu_2 = trapz(L,f0.*L.^2);
miu_3 = trapz(L,f0.*L.^3);
miu_4 = trapz(L,f0.*L.^4);
miu_5 = trapz(L,f0.*L.^5);

%----------------------- For the metastable polymorph

%----- Initial conditions
miu_0_1 = trapz(L,f0_1);
miu_1_1 = trapz(L,f0_1.*L);
miu_2_1 = trapz(L,f0_1.*L.^2);
miu_3_1 = trapz(L,f0_1.*L.^3);
miu_4_1 = trapz(L,f0_1.*L.^4);
miu_5_1 = trapz(L,f0_1.*L.^5);

Mass_of_Alpha_seed = rho_unstable/rho_solvent*kv_unstable*1000*miu_3_1;
Mass_of_Beta_seed = rho_stable/rho_solvent*kv_stable*1000*miu_3;

%----------------------- For the first polymorph

N0 = [ miu_0_1 miu_1_1 miu_2_1 miu_3_1 miu_0 miu_1 miu_2 miu_3 C_0 miu_4_1 miu_4 0 0 0 0 0 0 0 0 0 0 miu_5_1 miu_5 0 0 ];

% ANS(11) = fourth moment metastable
% ANS(11) = fourth moment stable
% ANS(m+12 - m+11) = number of particles in each interval for metastable
% ANS(m+12 - 2*m+11) = number of particles in each interval for metastable

% t_final = batch_time*3600;
delta_t = t_final/a;

result(1,1:15) = [N0(1:11),N0(22),N0(23),0,0];
MIU_OF_PARTICLES(1,:) = N0;

CSD_stable(1:m,1) = f0';
CSD_metastable(1:m,1) = f0_1';

z = 2;
for i=1:a
    if (i<=S)
        T(i) = T_0 - (T_0 - T_F1)*(delta_t*i/(delta_t*S))^U;
    else
        T(i) = T_F1 - (T_F1 - T_F2)*(delta_t*(i-S)/(t_final - delta_t*S))^U_1;
    end
end
T = [T_0 T];
Supersaturation_stable(1) = result(1,9)/solubility(T(1));
Supersaturation_metastable(1) = result(1,9)/solubility_1(T(1));
Number_ave_size_stable(1) = result(1,6)/result(1,5)*10^6;
Number_ave_size_metastable(1) = result(1,2)/result(1,1)*10^6;
Volume_ave_size_stable(1) = result(1,11)/result(1,8)*10^6;
Volume_ave_size_metastable(1) = result(1,10)/result(1,4)*10^6;
Mass_stable_form_produced(1) = Mass_of_Beta_seed;
Mass_unstable_form_produced(1) = Mass_of_Alpha_seed;
tic;
for i=2:a+1
    kb1 = KB1;
    kb2 = KB2;
    kbb1 = KBB1;
    kbb2 = KBB2;
    kb_1 = KB_1;
    kbb_1 = KBB_1;
    kd = KD;
    id = ID;
    kd_1 = KD_1;
    kg1 = KG1;
    kg2 = KG2;
    kg_1 = KG_1;
    g_1 = G_1;
    tspan = 0:delta_t/3:delta_t;
    if (result(end,9) >= (1 + Ostwald_Factor)*solubility_1(T(i)))
        kd_1 = 0;
        kd = 0;
        kb1 = 0;
        kb2 = 0;
        kbb1 = 0;
        kbb2 = 0;
        kg1 = 0;
    elseif (result(end,9) > solubility_1(T(i)) && result(end,9) < (1 + Ostwald_Factor)*solubility_1(T(i)))
        kd_1 = 0;
        kd = 0;
    elseif (result(end,9) > solubility(T(i)) && result(end,9) <= solubility_1(T(i)))
        if (result(end,1)<=0 || result(end,2)<=0 || result(end,3)<=0 || result(end,4)<=0 ||
            result(end,10)<=0)
            N0(1) = 0;
            N0(2) = 0;
            N0(3) = 0;
            N0(4) = 0;
            N0(10) = 0;
            N(22) = 0;
            N(12) = 0;
            N(13) = 0;
            N(14) = 0;
            N(15) = 0;
            N(16) = 0;
            N(24) = 0;
            kd_1 = 0;
            kbb_1 = 0;
            kg_1 = 0;
            kd = 0;
            Number_ave_size_metastable(z) = 0;
        end
    end
Volume_ave_size_metastable(z) = 0;
else
  kb_1 = 0;
  kg_1 = 0;
  kd = 0;
end
else
  kb1 = 0;
  kb2 = 0;
  kbb1 = 0;
  kbb2 = 0;
  kg1 = 0;
  kg2 = 0;
  kb_1 = 0;
  kg_1 = 0;
end

options=optimset('Display','iter');
t N = ode45(@(t,N) batch_sim1(t,N,T(i),rho_stable,kv_stable,rho_unstable,kv_unstable,rho_solvent,kb1,kb2,kbb1,kbb2,kb_1,kbb_1,kd,kd_1,kg1,kg2,kg_1,g_1),tspan,N0,options);
for q=1:numel(N)
  if N(q)<0
    N(q) = 0;
  end
end

N = real(N(end,:));
N0 = N;
MIU_OF_PARTICLES(z,:) = N;

result(z,1) = N(1) + N(12);
result(z,2) = N(2) + N(13);
result(z,3) = N(3) + N(14);
result(z,4) = N(4) + N(15);
result(z,5) = N(5) + N(17);
result(z,6) = N(6) + N(18);
result(z,7) = N(7) + N(19);
result(z,8) = N(8) + N(20);
result(z,9) = N(9);
result(z,10) = N(10) + N(16);
result(z,11) = N(11) + N(21);
result(z,12) = N(22) + N(24);
result(z,13) = N(23) + N(25);

B_stable(z) = Birth(kb1,kb2,T(i),result(end,9),kbb1,kbb2,result(end,4));
B_metastable(z) = Birth_1(kb_1,T(i),result(end,9),kbb_1);
Dissolution_stable(z) = Dissolution(kd,T(i),result(end,9));
Dissolution_metastable(z) = Dissolution_1(kd_1,T(i),result(end,9));
Growth_stable(z) = growth(kg1,kg2,T(i),result(end,9));
Growth_metastable(z) = growth_1(kg_1,g_1,T(i),result(end,9));
G_stable(z) = Growth_stable(end) - Dissolution_stable(end);
G_metastable(z) = Growth_metastable(end) - Dissolution_metastable(end);

Supersaturation_stable(z) = result(end,9)/solubility(T(i));
Supersaturation_metastable(z) = result(end,9)/solubility_1(T(i));
Number_ave_size_stable(z) = result(end,6)/result(end,5)*10^6;
Number_ave_size_metastable(z) = result(end,2)/result(end,1)*10^6;
Volume_ave_size_stable(z) = result(end,8)/result(end,7)*10^6;
Volume_ave_size_metastable(z) = result(end,4)/result(end,3)*10^6;
result(z,14) = B_metastable(z)/G_metastable(z);
result(z,15) = B_stable(z)/G_stable(z);

Number_stable = f0.*phi;
Number_metastable = f0_1.*phi;
NUM_STABLE(z,1:m) = Number_stable;
NUM_UNSTABLE(z,1:m) = Number_metastable;
C.2. Equations governing the polymorphic crystallization process

```matlab
function dN = batch_sim1(t,N,T,rho_crystal,alpha_v,rho_crystal_1,alpha_v_1,rho_solvent,kb1,kb2,kbb1,kbb2,kb_1,kbb_1,kd,kd_1,kg1,kg2,kg_1,g_1)

dN = zeros(25,1);

dN(1) = 0;  %---- N(m+1) = zeroth moment for Alpha phase, metastable

dN(2) = -1*Dissolution_1(kd_1,T,N(9))*N(1) + 1*growth_1(kg_1,g_1,T,N(9))*N(1);  %---- First moment for Alpha phase, metastable

dN(3) = -2*Dissolution_1(kd_1,T,N(9))*N(2) + 2*growth_1(kg_1,g_1,T,N(9))*N(2);  %---- Second moment for Alpha phase, metastable

dN(4) = -3*Dissolution_1(kd_1,T,N(9))*N(3) + 3*growth_1(kg_1,g_1,T,N(9))*N(3);  %---- Third moment for Alpha phase, metastable

end
```

```matlab
Mass_balance = C_0 + Mass_of_Alpha_seed + Mass_of_Beta_seed - (Mass_unstable_form_produced(end) + Mass_stable_form_produced(end) + result(end,9));

cputime_MOMC = toc;

Mass_stable_form_produced(z) = rho_stable/rho_solvent*kv_stable*1000*result(end,8);
Mass_unstable_form_produced(z) = rho_unstable/rho_solvent*kv_unstable*1000*result(end,4);
```

```matlab
CSD_stable(1:m,z) = ff;
CSD_metastable(1:m,z) = ff_1;
```

```matlab
f0 = CSD_stable(1:m,z);
f0_1 = CSD_metastable(1:m,z);
```

```matlab
Mass_stable_form_produced(z) = rho_stable/rho_solvent*kv_stable*1000*result(end,8);
Mass_unstable_form_produced(z) = rho_unstable/rho_solvent*kv_unstable*1000*result(end,4);
```
dN(6) = 1\times \text{growth}(kg_1,kg_2,T,N(9)) \times N(5) - 1\times \text{Dissolution}(kd,T,N(9)) \times N(5); \quad \text{#### First moment for Beta phase, stable}

dN(7) = 2\times \text{growth}(kg_1,kg_2,T,N(9)) \times N(6) - 2\times \text{Dissolution}(kd,T,N(9)) \times N(6); \quad \text{#### Second moment for Beta phase, stable}

dN(8) = 3\times \text{growth}(kg_1,kg_2,T,N(9)) \times N(7) - 3\times \text{Dissolution}(kd,T,N(9)) \times N(7); \quad \text{#### Third moment for Beta phase, stable}

dN(11) = 4\times \text{growth}(kg_1,kg_2,T,N(9)) \times N(8) - 4\times \text{Dissolution}(kd,T,N(9)) \times N(8); \quad \text{#### Fourth moment for Beta phase, stable}

dN(23) = 5\times \text{growth}(kg_1,kg_2,T,N(9)) \times N(11) - 5\times \text{Dissolution}(kd,T,N(9)) \times N(11); \quad \text{#### Fifth moment for Beta phase, stable}

dN(9) = \frac{-3 \times 1000}{\rho_{solvent} \times \rho_{crystal_1} \times \alpha_v_1} \times \text{growth}_1(kg_1,g_1,T,N(9)) \times (N(3)+N(14)) + \\
\frac{3 \times 1000}{\rho_{solvent} \times \rho_{crystal_1} \times \alpha_v_1} \times \text{Dissolution}_1(kd_1,T,N(9)) \times (N(3)+N(14)) - \\
\frac{3 \times 1000}{\rho_{solvent} \times \rho_{crystal} \times \alpha_v} \times \text{growth}(kg_1,kg_2,T,N(9)) \times (N(7)+N(19)) + \\
\frac{3 \times 1000}{\rho_{solvent} \times \rho_{crystal} \times \alpha_v} \times \text{Dissolution}(kd,T,N(9)) \times (N(7)+N(19)); \quad \text{#### N(m+5) = Concentration}

C.3. CSD generation for each polymorph

function df = CSD(t,f,G,D,B,m,phi)

%% the first size is L0 and the last is Ln.

%% if the number of class sizes are m, then each class size (phi) would
%% be (Ln-L0)/m

df = zeros(m,1);

df(1) = -(G)/(2*phi)*f(1) + D/(2*phi)*f(2) - D/(2*phi)*f(1) + B;

for i=2:m-1

df(i) = -(G)/(2*phi)*f(i) - (D)/(2*phi)*f(i) + G/(2*phi)*f(i-1) + D/(2*phi)*f(i+1);
end

df(m) = -(D)/(2*phi)*f(m) - G/(2*phi)*f(m) + G/(2*phi)*f(m-1);

C.4. Seed size distribution of both polymorphs
function CSD = seed_csd(lamda, sigma, miu, m, L0, LN)

phi = (LN-L0)/m;

z = 1;
for L=(L0+phi/2):phi:(LN-phi/2)
    CSD(z) = lamda/((2*pi)^.5*sigma)*exp(-(L - miu)^2/(2*sigma^2));
    z = z + 1;
end

C.4. Birth, growth, and dissolution of the stable polymorph

function B = Birth(kb1, kb2, T, C, kbb1, kbb2, moment_2_met)
C_star = solubility(T);
B = kb1.*(abs(C./C_star)).^(7/3).*exp(-kbb1./(log(C./C_star)).^2) + kb2.*moment_2_met.*exp(-kbb2./(log(C./C_star)));

function G = growth(kg1, kg2, T, C)
C_star = solubility(T);
G = kg1.*abs(C./C_star-1).^(5/6).*exp(-kg2./(abs(C./C_star-1)));

function D = Dissolution(kd, T, C)
C_star = solubility(T);
D = kd.*(abs(C./C_star-1));

C.5. Birth, growth, and dissolution of the metastable polymorph

function B = Birth_1(kb_1, T, C, kbb_1)
C_star_1 = solubility_1(T);
B = kb_1.*(abs(C./C_star_1)).^(7/3).*exp(-kbb_1./(log(C./C_star_1)).^2);

function G = growth_1(kg_1, g_1, T, C)
C_star_1 = solubility_1(T);
G = kg_1.*abs(C./C_star_1-1).^(5/6).*exp(-g_1./(abs(C./C_star_1-1)));

function D = Dissolution_1(kd_1, T, C)
C_star_1 = solubility_1(T);
D = kd_1.*(abs(C./C_star_1-1));

C.6. Solubility of both polymorphs

function C_star = solubility(T)
C_star = 7.644*10^-3.*T.^2 - 1.165*10^-1.*T + 6.222;

function C_star = solubility_1(T)
C_star = 8.437*10^-3.*T.^2 + 3.032*10^-2.*T + 4.564;

C.7. Optimal process calculation

%--------- Kinetic parameters for stable polymorph
KB1 = 5.4*10^4;
KB2 = 6*10^4;
KBB1 = 15;
KBB2 = 0.001;
KD = 0;
ID = 0;
KG1 = 6.5*10^-8;
KG2 = 0.16;

%-------- Kinetic parameters for metastable polymorph
KB_1 = 8*10^-5;
KBB_1 = 0.1;
KG_1 = 2.5*10^-7;
G_1 = 0.09;
KD_1 = 2*10^-8;

%-------- Seeding parameters for stable polymorph
lambda_1 = 1*10^10;
sigma_1 = 2*10^-6;
m_1 = 40*10^-6;
k_v_stable = 0.031;
rho_stable = 1540;

%-------- Seeding parameters for metastable polymorph
lambda_2 = 5*10^10;
sigma_2 = 2*10^-6;
m_2 = 40*10^-6;
k_v_metastable = .48;
rho_metastable = 1540;

run_time = cputime;
%-------- Initial condition and segment numbers in length and time domain
L0 = 0;
LN = 100*10^-6;
m = 300;
a = 40;
Ostwald_Factor = 1;
rho_solvent = 996;
T_0 = 45;
t_batch = 4; %--- hours
initial_concentration = 20; %--- gr/kg
delta_t = t_batch/a;
time = 0:delta_t:t_batch;
phi = (LN - L0)/m;
L = L0 + phi/2:phi:LN - phi/2;
L = L.*10^6;
lb = [20 , 20, 0, 0, 0];
ub = [T_0, 20, a, 5, 5];
for i=1:a-1
    for j=1:a
        if j == i
            A(i,j) = -1;
        elseif j == i+1
            A(i,j) = 1;
        else
            A(i,j) = 0;
        end
    end
end
for i=a:2*a-2
    for j=1:a
        if j == i - a + 1
            A(i,j) = -1;
        elseif j == i+1-a + 1
            A(i,j) = 1;
        else
            A(i,j) = 0;
        end
    end
end
b = [R_max*ones(a-1,1);-R_min*ones(a-1,1)];
b = R_max*ones(a-1,1);
x0 = [T_F1, T_F2, S, U, U_1];
\[ x, fval = \text{fmincon}(@f(x) - 1).^\text{objective_1}(x(1), x(2), x(3), x(4), x(5), a, m, L0, LN, \text{Ostwald\_Factor}, \text{lamda}_1, \text{lamda}_2, \text{sigma}_1, \text{sigma}_2, \text{miu}_1, \text{miu}_2, \text{rho\_stable}, \text{rho\_metastable}, \text{rho\_solvent}, \ldots) \]

\[ \text{k\_v\_stable}, \text{k\_v\_metastable}, \text{initial\_concentration}, T_0, t\_batch, K\text{B}1, K\text{B}2, K\text{B}B1, K\text{B}B2, K\text{B\_1}, K\text{B\_2}, K\text{D}, K\text{D\_1}, K\text{G}_1, K\text{G}_2, K\text{G\_1}, G_1, x0, [], [], [], [], [lb, ub]); \]

\[ \text{AA} = \text{batch\_crystal}(x(1), x(2), x(3), x(4), x(5), a, m, L0, LN, \text{Ostwald\_Factor}, \text{lamda}_1, \text{lamda}_2, \text{sigma}_1, \text{sigma}_2, \text{miu}_1, \text{miu}_2, \text{rho\_stable}, \text{rho\_metastable}, \text{rho\_solvent}, \ldots) \]

\[ \text{Conc}(1,:) = \text{AA}(:,9); \]

\[ \text{Mass\_met}(1,:) = \text{AA}(:,\text{end}); \]

\[ \text{Mass\_st}(1,:) = \text{AA}(:,\text{end} - 1); \]

\[ \phi = (LN - L0)/m; \]

\[ L = L0 + \phi/2:phi:LN - \phi/2; \]

\[ \text{for} \ i = 1:a \]

\[ \text{if} \ (i<=x(3)) \]

\[ T(i) = T_0 - (T_0 - x(1))*(\delta_t*i/(\delta_t*x(3)))^x(4); \]

\[ \text{else} \]

\[ T(i) = x(1) - (x(1) - x(2))*(\delta_t*(i-x(3))/(t\_batch-\delta_t*x(3)))^x(5); \]

\[ \text{end} \]

\[ \text{end} \]

\[ T = \text{real}([T_0, T]); \]

\[ \text{time} = 0:\delta_t:t\_batch; \]

\[ \phi = (LN - L0)/m; \]

\[ L = L0 + \phi/2:phi:LN - \phi/2; \]

\[ L = L.*10^6; \]

\[ \text{figure}(1) \]

\[ \text{plot}(\text{time}, \text{solubility}(T), '-o', \text{time}, \text{solubility\_1}(T), '-\--', \text{time}, \text{Conc}); \]

\[ \text{title('concentration');} \]

\[ \text{figure}(2) \]

\[ \text{plot}(\text{time}, \text{Mass\_st}, \text{time}, \text{Mass\_met}, '-\--'); \]

\[ \text{title('Mass');} \]

\[ \text{figure}(3) \]

\[ \text{plot}(\text{time}, \text{AA}(1:end,32)./\text{AA}(1:end,31).*10^6); \]

\[ \text{figure}(4) \]

\[ \text{plot}(L, \text{AA}(1,\text{end}-m-1:end-2)); \]

\[ \text{toc}; \]

\[ \text{run\_time} = \text{cputime} - \text{run\_time}; \]

\[ \text{run\_time} = \text{run\_time}/60; \]
Resume

Surname: Sheikholeslamzadeh
Name: Ehsan

Objective Summary

- Chemical and environmental process design, analysis, and optimization
- Chemical waste management and recovery process design
- Quality assurance and process reliability

Academic Background

- University of Western Ontario, London, Ontario, Canada (2009-2013)
  Ph.D. of Chemical Engineering, Reaction and Process Systems Engineering
- Sharif University of Technology (SUT), Tehran, Iran (2006-2008)
  M.Sc. of Chemical Engineering, Transport Phenomena and Separation Processes
- Shiraz University, Shiraz, Iran (2002-2006)
  B.Sc. of Chemical Engineering, Petrochemical Industries

Interests

- Simulation, modeling, and optimization of chemical and physical processes using powerful programs (Aspen, Hysys, Pro II, Matlab, LabVIEW, Comsol Multiphysics)
- Conceptual and detailed design of various chemical processes and equipment such as pumps, compressors, separators, reactors, etc
- Process control and automation in dynamic and steady-state operation and its application in start-up, shut-down, and process safety
- Laboratory procedure and operation using variety of equipment (XRPD, DSC, TGA, Raman Spectroscopy, FTIR (solid-state and solution), FBRM (size distribution of particles), Malvern Mastersizer, Titration, pH metering)
- Thermodynamics evaluation and calculation of complex systems (VLE, LLE, SLE, and VLLE) and their application in various chemical process systems
- Pharmaceutical process design and control, solvent screening, solubility prediction, characterization, and production processes
- Modeling and design of experiments (DOE) and analysis using ANOVA procedure for chemical and environmental processes
Software and Other Skills

- Excellent knowledge of simulation and HAZOP study software (Aspen Dynamics, Aspen Zyqad, Aspen Hysys, Aspen Properties, Aspen B-JAC, Pro II)
- Excellent knowledge of software for preparation of documents and reports (MS-Word, MS-Excel, MS-PowerPoint, MS-Visio, MS-Access, Autocad P&ID and Plant3D, Minitab)
- Excellent knowledge of engineering and drawing software (Autocad P&ID and Plant 3D, Matlab, Matlab GUI, C#, LabVIEW)
- Familiar with characterization equipment, Lasentec FBRM, Raman spectroscopy, ATR-FTIR (solution and solid-state characterization), XRPD, DSC, TGA
- Preparing, management, and organizing documents and reports for laboratory and industrial scale chemical process systems
- Experienced in project scheduling and supervising a project team
- Well organized in planning and performing assigned job
- Excellent in presenting the project and communicating with people in related professions
- WHIMS, General Safety & Hazardous Waste Management, and Employee Health and Safety certificates from Western University

Working Experience

- Instructor, Lecturing the LabVIEW software to the last-year undergraduate student for the process dynamics and control course and implementing the software for air temperature control system. The University of Western Ontario, London, Canada, Since 2011 until now.
- Four years of teaching assistantship of Bachelor’s courses in the University of Western Ontario, London, Canada (Since Sep. 2009) such as chemical engineering thermodynamics, chemical reactor design, process dynamics and control, chemical process simulation, and Oil production, refining, and processing.
- Two years of teaching assistantship of Bachelor’s courses in the Sharif University of Technology, Tehran, Iran.
**Journal Publications and Conference Presentations**


**Other Activities**

- Technical program chair of the particle technology research centre (PTRC) and 61th Canadian Chemical Engineering Conference member, London, Ontario, Canada, Oct. 2011.
- Member of Canadian Society for Chemical Engineers (CSChe).
- Member of CGS (chemical graduate society) of graduate students of the University of Western Ontario

**Honors**

- Recipient of Western Graduate Research Scholarship and NSERC (Natural Sciences and Engineering Research of Canada) grant, since Sep. 2009.
- Rank 3rd in the national entrance exam of M.Sc. program, Iran, 2006.
- Rank 2nd in the Olympiad of Chemical Engineers in Tehran, Iran, 2006.