KINETIC MODELING OF DISSOLUTION AND CRYSTALLIZATION OF BATCH REACTIONS WITH IN SITU SPECTROSCOPIC MEASUREMENTS

by

Chun H. Hsieh

July, 2012

Director of Thesis: Paul J. Gemperline

Major Department: Department of Chemistry

The use of Process Analytical Technology (PAT) recommended by the Food and Drug Administration (FDA) has significantly increased during the past years in the design, control and monitoring of pharmaceutical or chemical manufacturing processes. Nowadays, PAT is also commonly used in Good Manufacturing Practices (GMPs). Some PAT techniques employ on-line fiber-optic sensors to acquire non-destructive measurements of physical properties. In this thesis we demonstrated that these methods can also be used to obtain kinetic information of dissolved and solid fractions of molecular substances in slurries in real time.

The main objective of this project was to develop a comprehensive model for DuPont's sulfonylurea coupling reaction for monitoring purposes (e.g. detect process upsets, detect endpoints, and forecast changes). The comprehensive model will allow us to estimate the kinetics of the reaction, the kinetics of dissolution, and the kinetics of crystallization.

Before such a complex model can be developed, it was necessary to conduct experiments in a simpler system (e.g. salicylic acid in water-ethanol mixtures). These experiments were designed so the process of dissolution and the process of crystallization could be observed

independently of one another and independently of chemical reactions. Consequently, we were able to elucidate an appropriate model for each process in our small-scale semi-batch reactor.

For the simpler system, this study used attenuated total reflectance ultra-violet visible (ATR UV-vis) spectroscopy for kinetic modeling of the dissolution and crystallization of salicylic acid in ethanol-water. The dissolution model, which relied on a power-law equation, was obtained by adding aliquots of an ethanol-water mixture into a salicylic acid slurry. Near-infrared (NIR) diffuse reflectance spectroscopy was used to detect and quantify the solid fraction present in a slurry. Using a Partial Least-Squares (PLS) calibration, we were able to verify and validate the kinetic model. A temperature probe was also used to monitor heat changes involved in these experiments.

KINETIC MODELING OF DISSOLUTION AND CRYSTALLIZATION OF BATCH REACTIONS WITH IN SITU SPECTROSCOPIC MEASUREMENTS

A Thesis

Presented To The Faculty of the Department of Chemistry

East Carolina University

In Partial Fulfillment of the Requirements for the Degree

Master of Science in Chemistry

By

Chun H. Hsieh

July, 2012

KINETIC MODELING OF DISSOLUTION AND CRYSTALLIZATION OF BATCH REACTIONS WITH IN SITU SPECTROSCOPIC MEASUREMENTS

By

Chun H. Hsieh

APPROVED BY:

DIRECTOR OF THESIS _____

Paul J. Gemperline, PhD

COMMITTEE MEMBER _____

Yu Yang, PhD

COMMITTEE MEMBER _____

Yumin Li, PhD

COMMITTEE MEMBER _____

Mary-Ellen McNally, PhD

CHAIR OF THE DEPARTMENT OF CHEMISTRY

_____

Rickey P. Hicks, PhD

DEAN OF GRADUATE SCHOOL

_____

Paul J. Gemperline, PhD

ACKNOWLEDGEMENTS

I would like to express my great appreciation to my mentor Dr. Paul J. Gemperline who gave me the opportunity of doing my master's degree in his research group, and for his invaluable encouragement and guidance. Without his enthusiasm, dedication, and patience, I could not have completed this research project and my master thesis. I also want to extend special thanks to my thesis committee members Dr. Yu Yang, Dr. Yumin Li, and Dr. Mary-Ellen McNally for their valuable time, assistance, and advice. I want to give my sincere gratitude to former members of our research group, particularly, Dr. Julien Billeter and Mrs. Tess B. Kirchner, for their technical assistance and support. For the development and growth of this project, I would like to thank National Science Foundation (NSF) under grant number CHE-0750287 for Grant Opportunities for Academic Liaison with Industry (GOALI) and E.I. DuPont de Nemours and Co., Inc., Crop Protection Products and Engineering Technologies for their generous funding of this research. I would like to thank all the GOALI principal investigators Dr. Ron Hoffman (DuPont), Dr. Liguo Song (UTK), Dr. Frank Chamber (OSU), and Dr. Kelsey Cook (NSF), for their time and invaluable advice. My thanks are extended to Dr. Rich Davis, David Bailey, Steve Platz, Kevin Day, and Dr. David Cho, for their assistance when I was doing my internship at Stine-Haskell Research Center. Last but not least, my greatest appreciation is reserved for my parents Mr. Andy Hsieh and Mrs. Chin Mei Liao, my sister Annie Hsieh, my aunts Mrs. Lichen Lo and Mrs. Ting Hsieh. Their infinite love, support, positive attitude and encouragement have helped me to be a successful person in life.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

Organization of This Thesis

This thesis is organized as follows. Chapter 1 presents a general introduction to some of the concepts employed in this thesis research. Chapter 2 provides more detailed, theoretical explanation of different principles employed in this thesis. Chapters 3 and 4 are manuscripts to be submitted for publication in peer-reviewed journals. Chapter 5 provides concluding remarks that tie together the work described in chapters 3 and 4.

1.1 Chemometrics

Chemometrics is the application of mathematical, statistical, graphical and symbolic methods to (1) maximize the chemical information that can be extracted from data, (2) obtain information about chemical systems, and (3) design or determine optimal experimental procedures. Chemometrics techniques are useful at any point in a chemical analysis, from the very beginning of an experiment, until the data is discarded. There are several advantages to using chemometrics techniques, and some of these advantages include: (1) the ability to obtain relevant and useful information from less resolved data, (2) fast speed in obtaining real-time information from measured spectra, and (3) the ability to provide clear, unambiguous information and discrimination power when applied to multivariate data sets. The correlative relationship within a data set can be determined by using chemometrics techniques, and these relationships can be used to predict both physical and chemical properties of processes and, therefore, robust models can be developed [1].

There are two types of modeling techniques used in chemometrics; they are soft-modeling and hard-modeling. Soft-modeling is a general technique imposing minimum amount

of constraints regarding peak shape, location, or identity to determine the number of components in a time evolving system such as chromatography where there are overlapped peak as well as the spectral and concentration profiles of each compound [2, 3]. Soft-modeling approaches usually assume the pure component spectra and profiles are both non-negative and unimodal. Soft-modeling approaches attempt the description of an evolving system without the need of an a priori model postulation, physical or chemical. Some typical applications for soft-modeling technique include: (1) characterization of batch reactions [4], (2) end point detection [5], (3) on-line reaction monitoring [6], (4) quantification of trace analytes [7], and (5) hyperspectral fluorescence imaging [8, 9]. However, some disadvantages when using soft-modeling techniques are rotational ambiguities and intensity ambiguities; meaning multiple models fit the data equally well [10]. Soft-modeling separates and resolves complex data sets into simpler components matrices, however, in hard modeling techniques first principles based models are fitted to complex data sets by estimating model parameters [10].

Kinetic model-modeling techniques consist in modeling some part of the data using a chemical or physical first principles model and determining some model on kinetic parameters by optimization [11]. Hard-modeling techniques are used in this study because they are the only approaches that allow estimation and determination of reaction mechanisms, model parameters, concentration profiles, solid profiles, and interpolation and extrapolation outside the conditions used to determine the model and its parameters [10]. In addition, hard-modeling techniques can be used to determine the pure component spectra of all absorbing species in a reaction. Hard-modeling of chemical systems involves fitting a model directly to measured multivariate spectroscopic data. The results obtained are generally reliable when using hard modeling methods if the appropriate reaction model has been postulated [12, 13] and the correct initial

conditions are known. The objective for hard-modeling is to fit a model by reducing sum of squares in the residuals matrix [10].

In order to understand a chemical process involving slurries, it is necessary to determine the reaction mechanisms, which include chemical equilibria, kinetics of dissolution, and the kinetics of crystallization within the process [14]. Furthermore, it is particularly crucial to know or be familiar with kinetics and equilibrium constants of a process in order to predict its chemical behavior under different circumstances [15]. In the past years, reaction kinetics has been studied by taking experimental samples from the reaction mixture at different reaction times and analyzing them externally by offline methods like gas or liquid chromatography [14]. The measured data are then fitted to the kinetic model by non-linear regression [14]. Near infrared (NIR) [16], infrared (IR) [17], Raman [18] and Untraviolet/Visible (UV/VIS) [19, 20] spectroscopy are all now being used to study reaction kinetics by fitting a kinetic model directly to multivariate spectral data.

Multivariate Curve Resolution (MCR) or the Self-Modeling Curve Resolution (SMCR), is one of the commonly used chemometric soft-modeling techniques. SMCR was developed to elucidate the composition of a multivariate set of data taken on mixture, with or without incomplete prior knowledge of the components of the system. The main purpose of SMCR is to reduce the instrumental response for a reaction mixture into the pure contributions of each component involved in the system studied, in another words, the method resolves mixture data matrix into concentration profiles and pure component spectra [21-23]. The pure component spectra and pure concentration profiles usually are non-negative, and the concentration profiles usually are unimodal [6]. SMCR is a calibration free technique that was further developed by Gemperline by integrating either soft constraints or hard constraints. Hard constraints are

constraints that are enforced strictly, whereas soft constraints allow small deviations from constrained values [4]. Gemperline developed soft constraints using penalty functions in order to develop models that allow consideration of non-ideal behavior due to noise, non-ideal chemical response and non-ideal spectroscopic response when using SMCR methods.

SMCR is an example of iterative methods [24-27]. Iterative methods are often used due to their versatility for model manipulation. Iterative methods do not require any assumptions regarding the model due to their ability to handle different kinds of data structures and chemical problems. Additionally, they are also able to incorporate external data into the process. An example for non-iterative method is the orthogonal projection approach (OPA) [28].

A fundamental assumption needed to be made for OPA is that the unadulterated spectra in the data that were recorded during the process are more disparate than the corresponding mixture spectra [29]. OPA defines the number of components existing in the mixture and their corresponding spectra in the first step [29]. After this, the multivariate curve resolution-alternating least-squares (MCR-ALS) technique is applied to resolve the mixture data matrix, meaning that the spectra recorded as a function of time during the process is disintegrated into the concentration profile matrix and the pure component spectral matrix [30-32]. The MCR-ALS technique is commonly used to study chemical reactions. Some of the MCR-ALS applications include: (1) esterification reactions [33], (2) polymerization reactions [34, 35], (3) protein folding [32], and (4) alcohol fermentation [36].

1.2 Process Analytical Technologies

Process Analytical Technology (PAT) refers to analytical methods dedicated to the design, control and monitoring of industrial and pharmaceutical manufacturing processes [37]. In recent years, optimization of chemical processes from a laboratory scale to a pilot scale has

become necessary for optimization of yield, and minimization of reaction time, use of reagents, and quantity of waste produced. Effective optimization techniques are prominent in today's competitive world. PAT typically may employ use of a variety of on-line fiber-optic sensors. These inexpensive sensors can be used to acquire a large amount of non-destructive measurements that can be in turn related to physical and kinetic properties in real time. These measurements are analyzed by a set of multivariate statistical methods collectively known as Chemometric methods.

Under the initiative of the Food and Drug Administration (FDA), PAT has recently gained wider acceptance in the pharmaceutical and chemical industries [38]. PAT contributes to a better understanding of processes and can increase the selectivity and yield of products [39]. Quality control, e.g. Good Manufacturing Practices (GMP), also benefits from the higher rate of sampling of PAT over traditional offline analysis.

In this thesis, a batch reactor coupled with Near-Infrared (NIR) diffuse reflectance spectroscopy and Attenuated Total Reflectance Ultraviolet visible (ATR UV-vis) spectroscopy and kinetic fitting was used as an efficient method for process analysis. Batch reactors provide a means of precise control of reactor conditions and they provide an effective way for conducting experiments to obtain a maximum yield in a minimum time or minimal cost as well as to determine the specific final conditions in terms of quality and quantity [40].

1.3 Purpose of Research

The primary goal of this research project was to use chemometric hard and soft modeling techniques to determine the concentration profiles as a function of time for batch slurries. This approach is unique because kinetic modeling of batch slurries has never been done before. The model postulated should include both crystallization and dissolution. If successful, this project

will allow use of models for monitoring for quality control, which is a part of PAT. Making optical measurements in slurries is difficult, because reflectance measurements include both effects of light-scattering and light-absorption. Therefore, specific data handling techniques needed to be used to separate these effects. Another goal of this project was to be able to use these predictions in real time to estimate the correct amount of reagent to be added during the progress of the reaction and to know the product yields before the process is complete.

This was a collaborative project with East Carolina University, University of Tennessee, Oklahoma State University, and E.I. du Pont de Nemours and Company. Both NIR and UV-Vis measurements and kinetic modeling were done at East Carolina University. The measurements were from batch reactions of (1) salicylic acid in a water-ethanol mixture and (2) slurry-based sulfonylurea coupling reaction to form the product Metsulfuron Methyl (herbicide) manufactured by DuPont. Self-modeling curve resolution (SMCR) was used to characterize the batch information and form composition profiles of the reactants and products. Kinetic models based on information at the beginning of the reaction were used to make predictions at the end of the reaction.

1.3.1 Project 1: Modeling of Dissolution and Crystallization of Salicylic Acid

The first part of this research project was to develop a low-theory kinetic model for the dissolution and crystallization of salicylic acid in a solvent mixture. The solvent mixture was composed of 48% water and 52% ethanol. This kinetic model might help to monitor in real time chemical processes involving slurries [41]. For this study, Attenuated Total Reflectance Ultra-Violet visible (ATR UV-vis) spectroscopy and Near-Infrared (NIR) diffuse reflectance spectroscopy was used.

1.3.2 Project 2: Modeling of Sulfonylurea Coupling Reaction

The second part of this research project was to develop a combined kinetic model including reaction, dissolution and crystallization for the slurry-based sulfonylurea coupling reaction that leads to a herbicide manufactured by DuPont. The mathematical form of the kinetic model obtained in Project 1 for crystallization and dissolution of salicylic acid was used as a starting point in the second project. Ultimately, this combined kinetic model should help DuPont to monitor online the sulfonylurea coupling reaction and to perform endpoint detection, including determination of the appropriate stoichiometric amount of a reagent to be pumped in, as well as fault detection. This study uses NIR diffuse reflectance spectroscopy and Focused Beam Reflectance Measurement (FBRM). Additionally, slurry samples were taken from the reaction mixture and were analyzed by High Performance Liquid Chromatography (HPLC) to validate results obtained from NIR spectroscopy.

2. THEORY

2.1 Batch Slurry Processes

A slurry is a suspension formed when a quantity of powder is mixed into a liquid in which the solid is only slightly soluble or insoluble. Slurries contain a large amount of solid and are viscous and denser than the liquid from which they are formed initially. Some typical examples of slurries are mixtures of cement and water to form concrete, mixtures of wood pulp and water used to make paper, or mixtures of animal waste and water used as fertilizer, etc. Chemical, pharmaceutical, petrochemical, and polymer industries often use slurries, particularly in the chemical manufacturing industry. Spectroscopic techniques are often employed to monitoring chemical reactions in real time; however, making optical measurements in slurries can be very difficult due to light-scattering effects.

Spectroscopy is one of the instrumental tools that is frequently used for the identification and the quantification of chemical compounds of interest. Furthermore, spectroscopic techniques can be used to study the kinetics of chemical reactions and determination of kinetic parameters, such as rate constants. In addition to kinetic parameters, spectroscopy is capable of determining parameters that are related to the measured signal (e.g. absorption spectra). Absorption spectra are particularly valuable parameters because they provide structural information on the reactive species that are suitable for data oriented chemical process development [11]. Nowadays, spectroscopic techniques have become the method of choice for process monitoring. Spectroscopic data is often represented by a spectrum, a plot of the response of interest as a function of wavelength or frequency.

Fiber optics sensors coupled spectroscopic methods provide great advantages in process analysis. Optical fiber sensors are lightweight, small in size, and consist of fine strands of plastic, glass or other transparent fibers that are capable of transmitting radiation [42]. Typical diameters of an optical fiber can range from 0.05 um to 600 um. The fiber is composed of three layers of materials. The core or inner layer conducts the light; the cladding or middle layer is a transparent layer with a refractive index slightly lower than the refractive index of the internal layer. Both of these layers are covered with a protective polymer coating on the outside to provide strength and protection for the fiber. The cladding causes an abrupt refractive index change at the core-cladding interface [43]. One of the advantages for using optical fiber sensors coupled with spectroscopic method is that the sampling errors can be reduced significantly by inserting a fiber optic probe into a reactor. It is then unnecessary to stop and interrupt the system by collecting a sample and transporting it to an off-line analyzer [44]. However, some disadvantages include: (1) low light throughput and light changes that can hinder the measurement process, (2) high sensitivity to temperature changes, (3) sensitivity to perturbation such as change in bending of the optical fibers, and lastly, (4) some fibers can be very fragile and easily broken. Fiber optic sensors are typically made up of single or multiple fibers. Multiple fibers are commonly used for reflectance measurement due to the greater light throughput and better capacity and light collection abilities. Single fiber systems are usually used for transmission measurements [45].

2.2 ATR UV-Vis Spectroscopy

As recommended by the U.S. Food and Drug Administration (FDA), in order to understand and improve the process control and avoid failed batches, real-time concentration information (in liquid or solid phase) can be obtained using PAT [46]. One of the commonly used PAT techniques is Attenuated Total Reflectance (ATR). ATR is a spectroscopic technique

that can be used to measure in-situ the spectroscopic signal of chemical solutions, even in presence of solid particles. This is one of the advantages of ATR over transmission spectroscopy. In-situ spectroscopic techniques avoid the drawbacks encountered when taking samples out of a system for offline analysis, such as changes of volume, of temperature, of pressure or perturbation due to the quenching method.

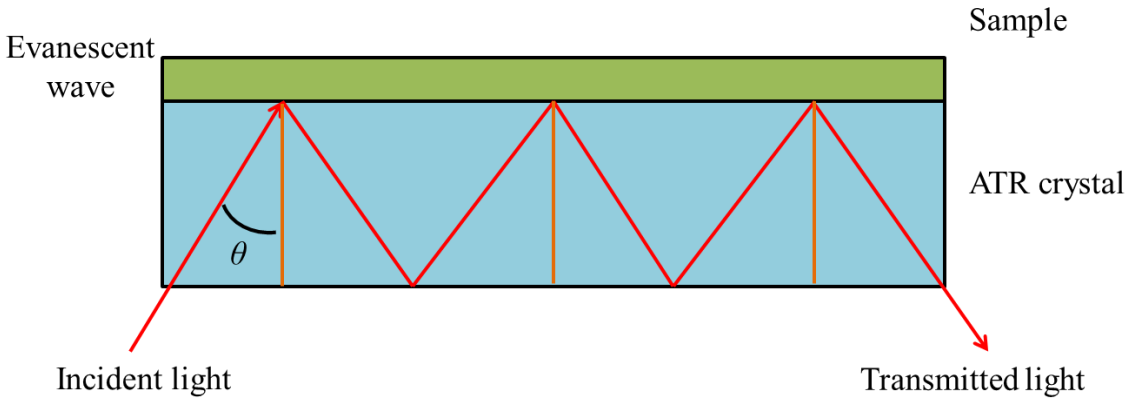

Figure 2.1: Schematic of a multiple reflection ATR system

The principle of attenuated total reflection is based on the behavior of an incident radiation beam that strikes an interface between a transparent material with high refractive index, i.e. the crystal of the ATR probe, and a transparent medium with a lower refractive index, i.e. the sample [47]. Snell's law defines a critical angle at a given wavelength which separates total internal reflection and external reflection. When the angle of incidence is wider than the critical angle, total internal reflection occurs and, at each interface, an evanescent wave penetrates into the sample where absorption can occur. This evanescent wave decays exponentially with the distance from the surface of the crystal. After several total reflections, the radiation leaves the crystal (see Figure 2.1). The interaction between the evanescent wave and the species in solution leads to a linear relationship between absorbance and concentration.

There are a number of crystal materials currently available in the market for UV/Vis ATR spectroscopy. It is very crucial to select an appropriate crystal material for ATR spectroscopy. Diamond crystals, being the most durable and robust, are by far the best ATR crystal material; however, the cost is very high and was not available for this research setting. In this thesis, a three-bounce sapphire ATR crystal was used because it is low-cost, easy to use, and ideal for use in an industrial process environment. A sapphire ATR probe is capable of operating at temperature up to 300 °C and at pressures up to 1200 psi. The three-bounce design of the ATR probe has a total path length of approximately 1.5 microns [48]. Sapphire allows response down to 200 nm in samples with an index of refraction up to 1.5. In our work, at 307 nm the effective path length was approximately 108 nm calculated by Equation 2.1 [49], where λ is the wavelength, d_e is the effective penetration, ϑ is the angle of incidence beam, n_1 is the refractive index of ATR sapphire crystal, n_2 is the refractive index of solvent mixture (48% water, 52% ethanol) [50] at 25 °C.

$$d_p = \frac{\lambda}{2\pi\sqrt{(n_1)^2 \sin^2\theta - (n_2)^2}} \tag{2.1}$$

2.3 NIR Diffuse Reflectance Spectroscopy

NIR spectroscopy is a commonly used spectroscopic technique for quantitative analysis [51]. Some of the typical applications for near infrared spectroscopy include pharmaceutical analysis [52], medical diagnostics (e.g. monitoring of blood sugar concentration) [53], and food and agrochemical quality control [54]. In diffuse reflectance (see Figure 2.2), a beam of radiation strikes the surface of a sample (e.g. dried powder sample or slurry); the radiation is then reflected back in all directions. The intensity of the reflected radiation depends on the size, the number of particles and on the nature of the scattering compound. The NIR diffuse reflectance spectra are

also affected by the absorption of the scattering compounds, and the absorption and light scattering properties of the material interact in a nonlinear way.

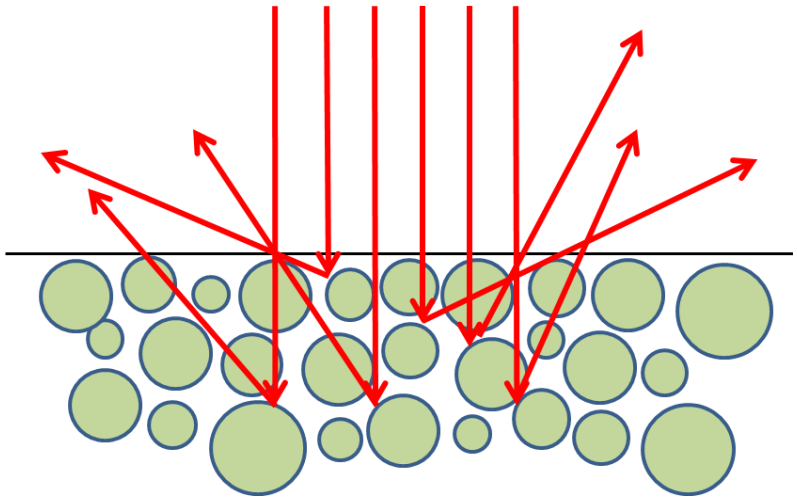

Figure 2.2: Schematic of diffuse reflection

So far, no method for the separation of light scattering and light absorption signals from a measurement at a single point in space and time has been demonstrated in the literature. However, several methods have been proposed to attempt approximate the light absorption signal from reflectance measurements, i.e. kubelka-munk theory [55], multiplicative scatter correction [56], enhanced multiplicative scatter correction [55] and variants of these latter two methods.

The NIR region of electromagnetic spectrum used to be considered useless because of the large amount of overlapping overtone and combination bands in this region [57]. However, the development of cutting-edge chemometrics techniques has proved that the NIR region contains an abundance of information, particularly in the area of kinetic modeling and process control. The introduction of optical light fibers has contributed to an enormous expansion of conventional NIR spectroscopy in terms of the analysis of toxic or hazardous materials in process control

applications [58]. NIR spectroscopy now offers numerous advantages in comparison to other spectroscopic techniques. In addition, NIR spectroscopy offers the possibility of non-destructive, multi-component analysis and identification of solid materials almost independently of their morphologies [58].

NIR spectroscopy measures the intensity of the absorption of near-infrared light at different wavelengths, typically range from 700 nm up to 2500 nm [59]. This region has sufficiently high energy to promote and excite overtones and combinations of molecular vibrations to higher energy levels. Overtones typically occur when a vibrational mode changes from the ground state to the higher level of excited state ($\Delta J > 1$). The overtones of stretching vibrations are usually stronger than bending or rocking modes. The overtone bands observed in the NIR region mainly involve hydrogen atoms bonded to other atoms such as carbon, nitrogen, and oxygen. There are two explanations for this phenomenon. First of all, the vibrations of atoms of similar size are more harmonic than ones of atoms of varying size. Different size atoms result in anharmonic vibrations and have stronger overtones [60]. Second of all, the fundamental stretching vibrations of these C-H, O-H, and N-H bonds occur at the shorter-wavelength, higher frequency end of the MIR spectrum due to their lower reduced mass. Also, NIR bands are broader than the corresponding Mid-Infrared (MIR) bands, although hydrogen-bonding effects do cause broadening in some overtones [60].

The NIR region contains groups of combination bands. Combination bands are formed due to transitions in two vibrational modes as a consequence of absorption of one quantum of radiation. These bands involve hydrogenic vibrations. These bands appear at frequencies that represent the sum of the two vibrations. The stronger the combination bands, the better coupling between the two vibrations. According to Kaye, good coupling is most likely to occur when the

frequencies of the two vibrations are very similar or when they are bound together by a double bond or ring. The NIR region is sensitive to hydrogen bonding and the intensity of OH groups is, enhanced significantly in the NIR region relative to NIR overtones. Overtone bands receive all their intensity from anharmonicity corrections as mentioned previously. The intensity of OH stretching absorbance increases in NIR by an order of magnitude upon hydrogen bonding. Hydrogen bonded hydroxyl groups have a much larger change in dipole moment upon hydrogen bonding.

There are several advantages of using NIR spectroscopy over mid IR spectroscopy. The overtones and combination bands have weak absorption coefficients and the molar absorptivities are low with a detection limit on the order of 0.1%. Due to the fact that NIR absorptions are generally 10 to 1000 times weaker in intensity than the fundamental mid IR and UV-Vis analysis, minimal sample preparation is required. NIR samples can be directly measured whereas mid IR and UV-Vis may require sample dilutions and the presence of solvents. Also, short optical pathlengths or dispersion in non-absorbing matrices is not required for NIR analysis.

2.4 Mathematical Methods and Preprocessing Techniques

All pre-processing techniques have the common goal of reducing the un-modeled variability in the data in order to enhance the feature sought in the spectra. In this thesis, Multiplicative Scatter Correction (MSC) was used because it is one of the most widely used pre-processing techniques for NIR. MSC was used to remove the light scattering effects in the NIR diffuse reflectance spectra. MSC correction is achieved by regressing a measured spectrum against a reference spectrum and then correcting the measured spectrum using the slope and intercept of this linear fit [61]. This pre-processing technique has proven to be effective in minimizing baseline offsets and multiplicative effect [62]. In addition, the Morris-Marten

baseline correction [63] technique was also used. It uses an offset, linear, and quadratic model adjustment. Baseline correction is typically performed through subtraction of a linear or polynomial fit of baseline from the original spectrum to remove tilted baseline variation.

2.4.1 Partial Least Square Regression

In the PLS method, the regressions are computed with least-square algorithms. The goal of the PLS is to create a linear relationship between two matrices, the spectral data **X** and the reference values **Y**. This method of modeling uses both **X** and **Y** in order to determine the variables in **X** matrix that will best describes the **Y** matrix. The general idea of PLS is to try to find the multidimensional direction in the **X** matrix that explains the maximum multidimensional variance direction in the **Y** matrix [64]. This can be explained by the representation of the spectra in the space of wavelengths. All the coding was done using the in-house Matlab PLS-R toolbox written by Lorder et.al [65].

2.4.2 Self-Modeling Curve Resolution (SMCR)

Self-modeling curve resolution (SMCR) is an iterative method that resolves mixture data into concentration profiles and pure component spectra of the constituents without a priori knowledge about the chemical system. The pure component spectra and pure concentration profiles are usually non-negative, and the concentration profiles are usually unimodal [6]. SMCR is a calibration free technique that was further developed by Gemperline by integrating either soft constraints or hard constraints. Hard constraints are constraints that are enforced strictly, whereas soft constraints allow small deviations from constrained values [4]. Gemperline developed soft constraints in order to develop models that take into consideration non-ideal

behavior due to noise, non-ideal chemical response and non-ideal spectroscopic response to minimize the lack of fit commonly encountered when using SMCR methods.

2.4.3 Alternating Least Squares (ALS)

Alternating Least Squares (ALS) is a model free (soft-modeling) iterative technique that can be used to obtain estimates of $\mathbf{C}$ and $\mathbf{A}$. It is based on a refinement by successive linear regressions [4]. Starting with an initial guess for $\mathbf{C}$, ALS calculates $\mathbf{A}$ by linear regression and then uses this $\mathbf{A}$ to recalculate $\mathbf{C}$. These iterative steps are typically repeated several hundreds of times. Between each linear regression, soft constraints are applied, e.g. non-negativity, unimodality and closure. In the following, a simplified description of a typical ALS algorithm is given.

1. Find an initial guess for $\mathbf{C}$
2. Calculate $\mathbf{A}$ based on $\mathbf{C}$ and $\mathbf{Y}$, i.e. $\mathbf{A} = \mathbf{C}^{+}\mathbf{Y}$
3. Apply soft constraints on $\mathbf{A}$
4. Calculate $\mathbf{C}$ based on $\mathbf{A}$ and $\mathbf{Y}$, i.e. $\mathbf{C} = \mathbf{Y}\mathbf{A}^{+}$
5. Apply soft constraints on $\mathbf{C}$
6. Go back to step 2

2.4.4 Principal Component Analysis

Principal component analysis (PCA) is a method for the visualization of complex data by dimension reduction [66]. The most prominent PCA application is the reduction of the number of variables and the representation of a multivariate data table in a low dimensional space. Consequently, the new variables are linear combinations of the original ones and can be interpreted like spectra, and noise is separated from the information of interest [66, 67].

2.4.5 Normalization

Normalization is a row operation that is applied to the data from each sample and comprises methods to make the data from all samples directly comparable with each other [30]. Normalization is the process of isolating statistical error in repeated measured data.

2.5 Crystallization

Crystallization is the process of formation of solid crystals precipitating from a solution. Crystallization is also a chemical solid-liquid separation technique, in which mass transfer of a solute from the liquid solution to a pure solid crystalline phase occurs [68]. The crystallization process consists of two major events; nucleation and crystal growth. Both are described in more detail in below.

2.5.1 Nucleation

The nucleation rate is defined as the birth of new crystals per unit time. The nucleation rate can be further sub-divided into primary nucleation rate and secondary nucleation rate. The main difference between primary nucleation (Equation 2.2) and secondary nucleation (Equation 2.3) is that for primary nucleation, the initial formation of nuclei occurs when there are no other seed crystals present inside a crystallizer [69]. As for secondary nucleation, the formation of nuclei occurs when there are seed crystals already present in a crystallizer. Secondary nucleation is attributable to fluid shear. It is primary due to collisions between existing crystals or with a solid surface of the crystallizer or propeller [68, 70]. In Equations 2.2 and 2.3, J_{NI} and J_{NII} are the primary nucleation and secondary nucleation rates, A_I, A_{II}, B_I, *and* B_{II}, are semi-empirical parameters, N^h is the stirrer speed [68].

$$J_{NI} = A_I \exp\left[\frac{-B_I}{\ln(S)^2}\right] \tag{2.2}$$

$$J_{NII} = A_{II} \Delta c^b N^h \tag{2.3}$$

$$S = c / c_{sat} \qquad \Delta c = c - c_{sat} \tag{2.4}$$

Equation 2.4 describes the degree of supersaturation where S represents the supersaturation ratio, c represents concentration of the analyte in solution and c_{sat} represents the solubility of the analyte.

2.5.2 Crystal Growth

Crystal growth is usually described by a thin film model with a mass transfer process across the diffusion layer surrounding the crystal in competition with solute molecule deposition onto the surface of crystal [70]. In Equation 2.5, r_c is the crystallization rate of the analyte. η_r is the effective size factor introduced by Garside [68], k_c is the kinetic rate constant, MW_s is the molecular weight of analyte, d_s is the density of the analyte, c_{sat} is the solubility of the analyte, c is the analyte concentration, and g is the exponent. Φ_s and Φ_v are the surface and volumetric shape factors.

$$r_c = \frac{\Phi_s MW_s k_c}{3 d_s \Phi_v} \eta_r (c - c_{sat})^g \tag{2.5}$$

The rate of crystal growth can be defined as the function of supersaturation, which means the higher the supersaturation level, the faster the crystal growth rate.

2.6 Dissolution

The process of dissolution occurs when a solute is placed in contact with a solvent and dissolves to form a solution [69]. In Equation 2.6, r_d is the dissolution rate of the analyte, k_d is the kinetic rate constant, MW_s is the molecular weight of analyte, d_s is the density of the analyte, c_{sat} is the solubility of the analyte, c is the analyte concentration

$$r_d = \frac{2MW_s k_d}{d_s}(c_{sat} - c) \tag{2.6}$$

2.7 High Performance Liquid Chromatography (HPLC)

HPLC is the most widely used chromatographic separation technique for the separation, identification and determination of chemical components in complex mixtures [71]. Figure 2.3 shows a generic schematic of a HPLC system. A typical HPLC system includes a pump, injector, column, detector and data system. The heart of the system is the column where separation occurs. Separation by HPLC is based on the difference in migration r

ates of the components through a column (stationary phase) by solvent (mobile phase). HPLC can be further subdivided into two different categories: Normal Phase Chromatography and Reverse Phase Chromatography

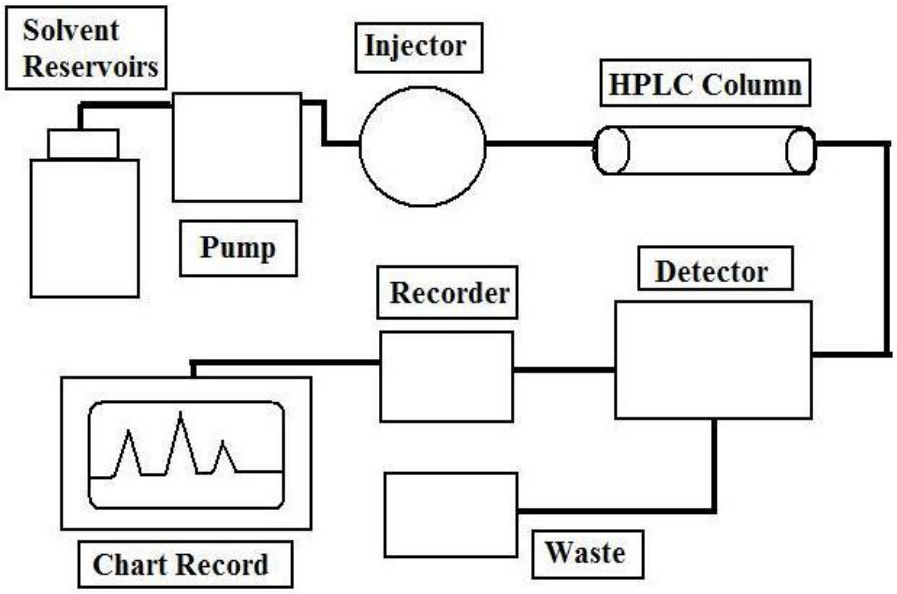

Figure 6.3: Schematic of a typical HPLC system

Normal Phase Chromatography separates molecules on the basis of differences in the strength of their interaction with a polar stationary phase. It uses a polar stationary phase and a nonpolar organic solvent, such as n-hexane, chloroform, and methylene chloride as the mobile phase. Reversed Phase Chromatography uses a nonpolar stationary phase and a polar mobile phase, such as methanol, acetonitrile, water, or mixtures of these solvents. The most common bonded phases are n-octyldecyl (C_{18}) and n-decyl (C_8) chains, and phenyl groups. Reverse-phase chromatography is the most common form of liquid chromatography, primarily due to the wide range of analytes that can dissolve in the mobile phase. Reversed Phase Chromatography was used in this research.

Chromatographic efficiency and flow-rate of the mobile phase are inversely proportional to the silica particle size; backpressure generated by small particles can be a problem. HPLC was used for off-line analysis to validate results obtained from NIR spectroscopy.

2.8 Focused Beam Reflectance Measurement (FBRM)

Focused Beam Reflectance Measurement (FBRM) is a technique that is commonly used for on-line or in-situ measurements to measure particles' chord length and particle distribution in concentrated suspensions and emulsions without sample extraction and sample preparation [72]. In this thesis, the FBRM was used to detect and confirm that supersaturation and un-seeded nucleation and crystallization occurred in the sulfonylurea coupling reaction studies. The FBRM measurement has been used in a wide variety of applications, such as the determination of aggregation of titanium oxide in rheology studies [73], the particle sizing in estuaries and coastal waters [74] and the monitoring of particle population of emulsions [75, 76]. Several research groups have used the method to monitor crystal size distribution on-line [77].

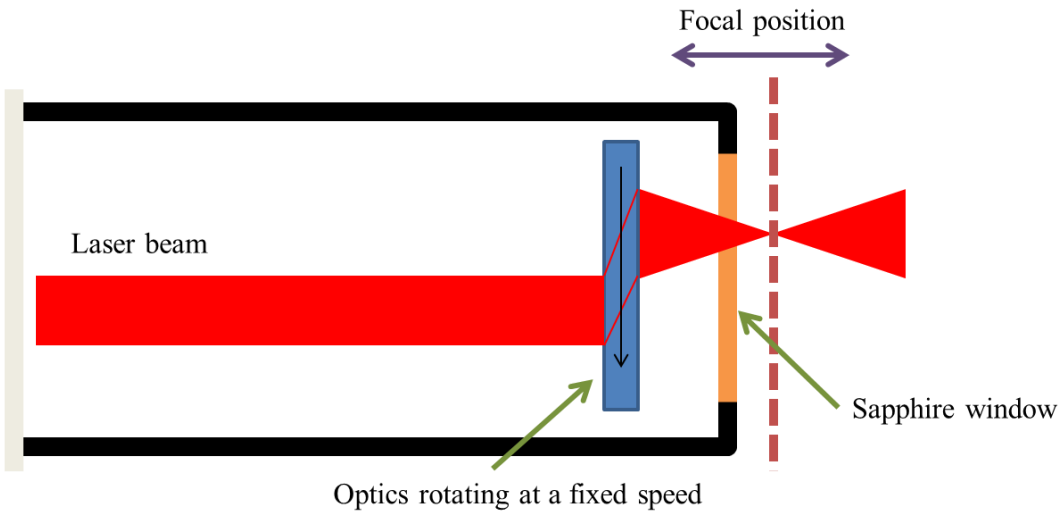

Figure 2.4: Schematic of the FBRM Probe

The FBRM is implemented using a cylindrical probe which can be easily located within the suspension to be characterized. An infrared laser beam rotates at high velocity, propagating into the suspension through a sapphire window on the probe tip (see Figure 2.4). The position of

the focal point is adjustable, depending on the particle properties [78]. When the beam hits a particle, it is reflected and back propagated through the probe window. This optical signal is then processed by the device electronics and the corresponding chord length is calculated as the product of the measured crossing time and the beam velocity. The duration of the reflected light pulse is directly proportional to the width of the particle intersected [74].

2.9 Kinetic Hard-Modeling

Kinetic modeling requires one to postulate a kinetic model and to subsequently fit measured data in the least square sense by optimizing non-linear model parameters [39]. In this work, univariate and multivariate data were used. Kinetic hard modeling consists in modeling some part of the data using a chemical or physical first principles model and determining some kinetic parameters by optimization. In bilinear spectroscopy and according to Beer's law, either concentration profiles (**C**, i.e. time profiles) or pure component spectra (**A**, i.e. wavelength profiles) can be modeled.

2.9.1 Beer-Lambert's Law

According to Beer's law, shown in Equation 2.7, absorbance data collected in matrix **Y** (*nt* x *nw*) is the sum of the contributions from all *nc* absorbing species, *nt* and *nw* being the number of times and wavelengths, respectively, where rows in **Y** represent digitized spectra measured as a function of time.

$$\mathbf{Y} = \mathbf{C}\,\mathbf{A} + \mathbf{R} \tag{2.7}$$

Matrices $\mathbf{C}$ (*nt* x *nc*) and $\mathbf{A}$ (*nc* x *nw*) contain the concentration profiles and pure component spectra (molar absorptivities), respectively. Since measurements are subject to measurement errors, a matrix of residuals $\mathbf{R}$ (*nt* x *nw*) has to be included in Equation 2.7 to take this into account.

2.9.2 Kinetic model and numerical Integration

Kinetic modeling starts by postulating a kinetic model $\wp$, which is based on elementary reactions, i.e. when partial orders equal the stoichiometric coefficients. The kinetic model is translated into a system of Ordinary Differential Equations (ODE), which are numerically integrated and result in the *nc* concentration profiles in $\mathbf{C}$.

$$\dot{\mathbf{C}} = \wp(\mathbf{p}, \mathbf{v}, \dot{\mathbf{v}}, \mathbf{r}, \mathbf{r_d}, \mathbf{r_c}) \tag{2.8}$$

Model $\wp$ is a system of ODEs depending on a vector of model parameters $\mathbf{p}$, volume $\mathbf{v}$, flow rate $\dot{\mathbf{v}}$, and rates $\mathbf{r}$, $\mathbf{r_d}$ and $\mathbf{r_c}$. Model $\wp$ is numerically integrated using initial conditions $\mathbf{c_0}$. The rate of reaction $\mathbf{r}$ collects all rates r_i of each i^{th} reaction step, which are based on their corresponding rate law, as shown in Equation 2.9, with k_i the reaction rate constant and $\boldsymbol{\nu_{i,:}}$ the partial orders (stoichiometric coefficients of reactants) of the i^{th} step. All rate constants k_i are collected in a vector $\mathbf{k}$.

$$r_i = k_i \prod_{j=1}^{nc} c_j^{\nu_{i,j}} \tag{2.9}$$

The rate of dissolution $\mathbf{r_d}$ collects all rates $r_{d,i}$ of each i^{th} dissolution step, which are based on a power law equation, as shown in Equation 2.10, with $k_{d,i}$ the dissolution rate constant,

$c_{sat,i}$ the saturation limit and n_i the power coefficient of the i^{th} step. All dissolution rate constants $k_{d,i}$ and power coefficients n_i are collected in vectors $\mathbf{k_d}$ and **n**, respectively.

$$r_{d,i} = k_{d,i}\,(c_{sat,i} - c_i)^{n_i} \tag{2.10}$$

The rate of crystallization $\mathbf{r_c}$ collects all rates $r_{c,i}$ of each i^{th} crystallization step, which are also based on a power law equation, as shown in Equation 2.11, with $k_{c,i}$ the crystallization rate constant and m_i the power coefficient of the i^{th} step. All crystallization rate constants $k_{c,i}$ and power coefficients m_i are collected in vectors $\mathbf{k_c}$ and **m**, respectively.

$$r_{c,i} = k_{c,i}\,(c_i - c_{sat,i})^{m_i} \tag{2.11}$$

Finally, the vector of model parameters **p** contains vectors $\mathbf{k}$, $\mathbf{k_d}$, $\mathbf{k_c}$, **n** and **m**.

2.9.3 Elimination of the Linear Parameters

At each iteration of the non-linear optimization steps, pure component spectra **A** for ATR UV/Vis measurements are linearly estimated using $\mathbf{C}$ and $\mathbf{Y}$ (Absorbance), as shown in Equation 2.12.

$$\mathbf{A} = \mathbf{C^{+}Y} \tag{2.12}$$

$\mathbf{C^{+}}$ is called the left pseudo-inverse and is computed as $\mathbf{C^{+}} = (\mathbf{C^{T}C})^{-1}\mathbf{C^{T}}$.

2.9.4 Non-linear Optimization

Standard Newton-Gauss-Levenberg/Marquardt (NGL/M), one of the most commonly used non-linear gradient optimizers, was used to optimize the model parameters. NGL/M relies on the calculation of the derivative of the residuals **R** (Equation 2.13). The derivatives of optimizable parameters required in the NGL/M method, were estimated by the technique of finite difference approximation because analytical derivatives were not known. The method of finite difference approximation was used with a relative shift value 1×10^{-6}. For the convergence criterion, we used a relative change of 1×10^{-4} in the sum of square residual from one iteration step to the next.

Boundary conditions on the adjustable parameters were implemented into the NGL/M function. If the values of parameters after the computed the shift exceed a maximum or a minimum value, then the parameters were adjusted to the upper or lower limits, respectively. The convergence of gradient methods is superior to other methods as long as the initial guesses for the parameters are close to their optimal values [79]. If divergence occurs and the direction of the shift vector is too far away from the ideal direction, Marquardt parameter *mp* is being introduced and increased (multiplication by 5 per iteration) until the *ssq* begins to converge (see Figure 2.5). Once the *ssq* converges the magnitude of the Marquardt parameter is reduced (division by $\sqrt{5}$ per iteration) and ultimately set to zero when the desirable criterion is reached [79].

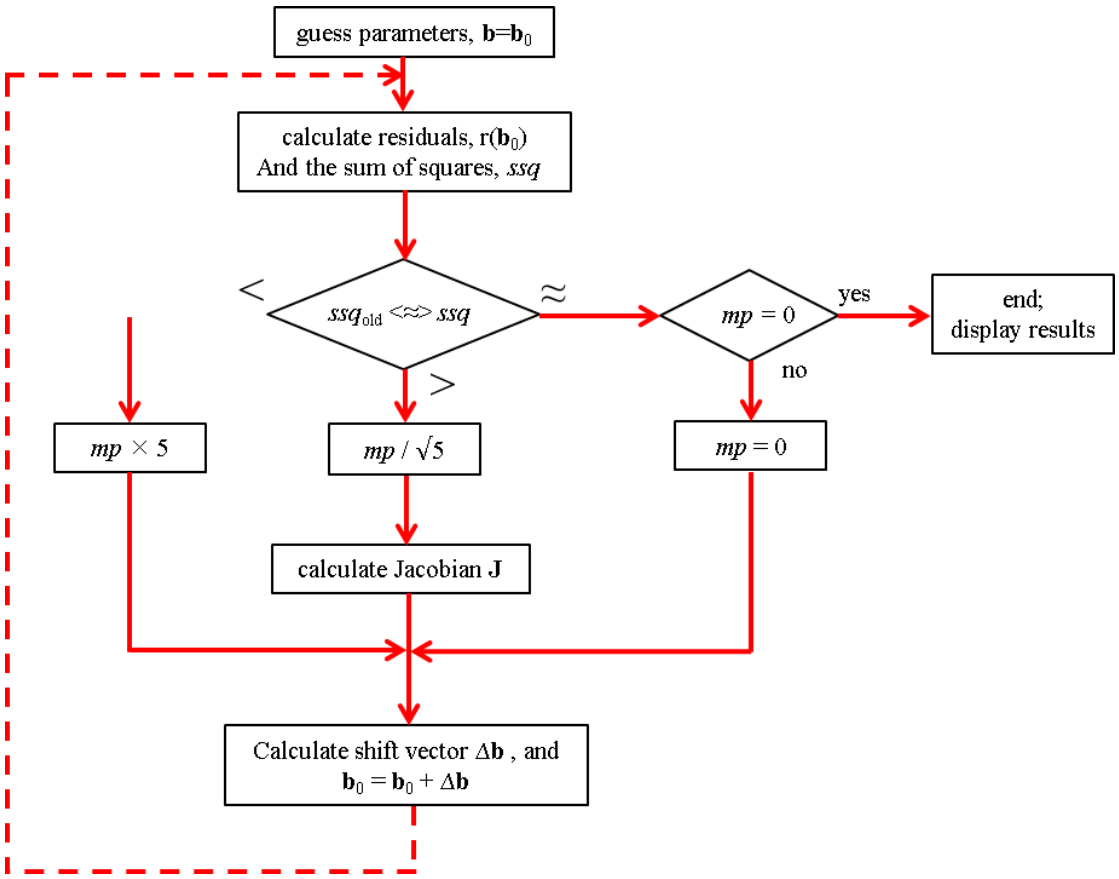

Figure 2.5: Flow diagram of the Newton-Gauss-Levenberg/Marquardt (NGL/M) method

According to Beer's law (Eq. 2.7), residuals **R** are calculated as:

$$\mathbf{R} = \mathbf{Y} - \mathbf{C}\,\mathbf{A} \tag{2.13}$$

The sum of squares, *ssq*, is the sum of all squared residuals **R**.

$$ssq = \sum_{i=1}^{nt} \sum_{j=1}^{nw} \mathbf{R}_{i,j}^2 \tag{2.14}$$

The least-squares optimization problem is then defined as $\min_{\mathbf{p}} ssq$.

3. KINETIC MODELLING OF DISSOLUTION AND CRYSTALLIZATION OF BATCH SLURRY WITH IN SITU SPECTROSCOPIC MEASUREMENT AND CHEMOMETRIC TECHNIQUES

3.1 Abstract

Slurries are often used in chemical and pharmaceutical manufacturing processes. The use of Process Analytical Technologies (PATs) recommended by the Food and Drug Administration (FDA) has significantly increased during the past years in the design, control and monitoring of pharmaceutical or chemical manufacturing processes. Nowadays PAT is also commonly used in Good Manufacturing Practices (GMPs). In this paper, we describe a new method to utilize online fiber-optic sensors to obtain nondestructive measurements of physical properties and kinetic data of the dissolved and solid fractions of molecular substances in slurries in real time. Diffuse reflectance near infrared (NIR) and attenuated total reflectance UV/Visible (ATR UV-Vis) spectra were measured simultaneously as a function of time from a small-scale semi-batch reactor. We then fit a simplified, comprehensive kinetic model to the measured spectroscopic data to determine the kinetics of crystallization and the kinetics of dissolution for online monitoring and quality control purposes (e.g. for detecting process upset, endpoint detection, and forecasting changes). The model comprises a system of ordinary differential equations. Nonlinear parameter optimization is used to estimate the model parameters for this system of ordinary differential equations. The parameters estimated in the model include dissolution and crystal growth rate constants, the dissolution rate exponent, as well as the degree of supersaturation which provides the driving force for crystallization.

KEY WORDS: Kinetic modeling; Spectroscopy; Slurries; Process analytical technology; Nonlinear fitting

3.2 Introduction

Slurries are often used in chemical and pharmaceutical manufacturing processes. Under the initiative of the Food and Drug Administration (FDA), Process Analytical Technologies (PAT) has recently gained increased utilization in the pharmaceutical and chemical industries [38]. PAT contributes to a better understanding of processes and can improve the purity and yield of products [39]. Quality control, e.g. Good Manufacturing Practices (GMP), also benefits from the higher rate of sampling of PAT over traditional offline analysis. In recent years, scale-up of chemical processes from laboratory scale to a pilot scale has become necessary for the optimization of yield, minimization of reaction time and use of reagents, and decreasing quantity of waste produced. Effective optimization techniques are prominent in today's competitive world of scientific and technological advancements. PAT typically requires use of a variety of online measurements such as fiber-optic sensors. These inexpensive sensors can be used to obtain a large amount of nondestructive measurements that can be useful for estimating physical and kinetic properties in real time. These measurements are analyzed by a set of multivariate statistical methods collectively known as Chemometric methods.

Chemometrics is the application of mathematical, statistical, graphical or symbolic methods to (1) maximize the chemical information which can be extracted from data, (2) obtain information about chemical systems, and (3) design or determine optimal experimental procedures. Chemometrics techniques are useful at any point in a chemical analysis, from the very beginning of an experiment until the data is discarded. There are several advantages for using chemometrics techniques, and some of these advantages include: (1) the ability to obtain

relevant and useful information from less resolved data, (2) fast speed in obtaining real-time information from measured spectra, and (3) the ability to provide clear, unambiguous information and discrimination power when applied to multivariate data sets. Correlation within a data set can be treated by using chemometrics techniques and these relationships can be used to predict both physical and chemical properties of processes and, therefore, robust models can be developed [1].

There are two types of modeling techniques used in chemometrics; they are soft-modeling techniques and hard-modeling techniques. Soft-modeling techniques are commonly applied to time-evolving systems such as chromatographic processes or batch chemical reactions. Soft-modeling techniques are used to estimate spectra and concentration profiles of each constituent in these evolving systems [2, 3] without the need to postulate a physical or chemical model. Soft-modeling techniques impose a minimum amount of constraints regarding peak shape, location, and identity to determine the pure component spectra, the number of components and the concentration of profiles in the evolving system. Soft-modeling techniques make minimal assumptions about the pure component spectra and concentration profiles which are usually expressed in the form of constraints. Some commonly used constraints assume the pure component spectra and profiles are both non-negative and unimodal. The multivariate signal is part of the data set which describes the properties of interest [1].

Some typical applications for soft-modeling techniques include: (1) characterization of batch reactions [4], (2) end point detection [5], (3) online reaction monitoring [6], (4) quantification of trace solutes [7], and (5) resolution of hyperspectral fluorescence imaging [8, 9]. However, some disadvantages when using soft-modeling techniques are rotational ambiguities and intensity ambiguities, meaning multiple models fit the data equally well [10]. Soft-modeling

techniques separate and resolve complex data sets into simpler component matrices, however, in hard modeling techniques first principles based models are fitted to complex data sets by estimating model parameters [10].

Kinetic hard-modeling techniques consist in modeling some part of the data using a chemical or physical first principles model and determining some kinetic parameters by optimization [11]. Hard-modeling techniques are used in this study because they are the only approach that allows estimation and determination of reaction mechanisms, model parameters, and concentration profiles in the liquid phase as well as undissolved solutes in the solid phase. Kinetic models can be used for interpolation and extrapolation to conditions different from those used to determine the model and its parameters [10]. In addition, hard-modeling techniques can used to indirectly determine concentration profiles of nonabsorbing species in a reaction. Hard-modeling of chemical systems involves fitting a model directly to measured multivariate spectroscopic data. The results obtained are generally reliable when using hard modeling methods if the appropriate reaction model has been postulated [12, 13]. The objective for hard-modeling is to fit a model by reducing the sum of squares or the residuals matrix [10].

In order to understand chemical processes involving slurries, kinetic models can be used to determine reaction mechanisms, which include chemical equilibria, kinetics of dissolution, and the kinetics of crystallization within the process [14]. Furthermore, it is particularly important to understand the kinetics and equilibrium constants of a process in order to predict a chemical behavior under different conditions [15]. In the past years, reaction kinetics have been studied by taking aliquots from a reaction mixture at different times and analyzing them externally by using techniques such as gas or liquid chromatography [14]. The measured data are then fitted to a kinetic model by non-linear regression [14]. Near infrared (NIR) [16], Infrared

(IR) [17], Raman [18] and UV/Visible (UV-Vis) [19, 20] spectroscopy has been used to study reaction kinetics by fitting a kinetic model directly to multivariate spectral data measured as a function of time during the course of a reaction.

In this paper, a batch reactor coupled with Near-Infrared (NIR) diffuse reflectance spectroscopy and Attenuated Total Reflectance Ultraviolet visible (ATR UV-Vis) spectroscopy was used to produce data for kinetic fitting as an efficient method for process analysis. Batch reactors provide a means of precise control of reactor conditions and they provide an effective way for conducting experiments to obtain a maximum yield in a minimum time or minimal cost as well as to determine the specific final conditions in terms of quality and quantity [40]. The primary goal is to use hard-modeling techniques to determine the concentration profiles as a function of time for batch slurries. The model postulated should include both crystallization and dissolution processes. Making optical measurements in slurries is challenging because reflectance measurements include both effects of light-scattering and light-absorption. Specific data handling techniques were designed to treat these effects.

3.3 Experimental

A series of four batch experiments were conducted using a 50 mL reactor at East Carolina University in Greenville, NC. In-situ NIR and ATR UV/Visible data were collected for all batches. The same chemical conditions, experimental protocol and process conditions were used for all batches. The process studied was the crystallization and dissolution of salicylic acid ($C_7H_6O_3$, molecular weight 138.12, Fisher Scientific ACS certified CAS 69-72-7) in a solvent mixture (52% w/w ethanol (CH_3CH_2OH, molecular weight 46.07, EMD 200 proof CAS 64-17-5), 48% w/w de-ionized water).

3.3.1 Reactor and Instrument Setup

The small-scale batch experiments reported in this paper were performed in a custom jacketed 50 mL reaction vessel made in-house at East Carolina University. As shown in Figure 3.1, the reactor and lid were specifically designed to accept a NIR probe, an ATR probe, a temperature probe, an electrical heater and two PTFE feed lines. The reactor jacket was thermostated using a MGW Lauda RMS6 heater/chiller and the PTFE feed lines were connected to a syringe pump. The reactor temperature, jacket temperature, heater power, and reagent additions from the pumps were monitored and controlled by WinISO software from H.E.L. Inc. (Lawrenceville, NJ). Stirring was controlled by an IKA Labortechnik RCT Basic magnetic stirrer and a PTFE coated stir bar placed in the reaction vessel.

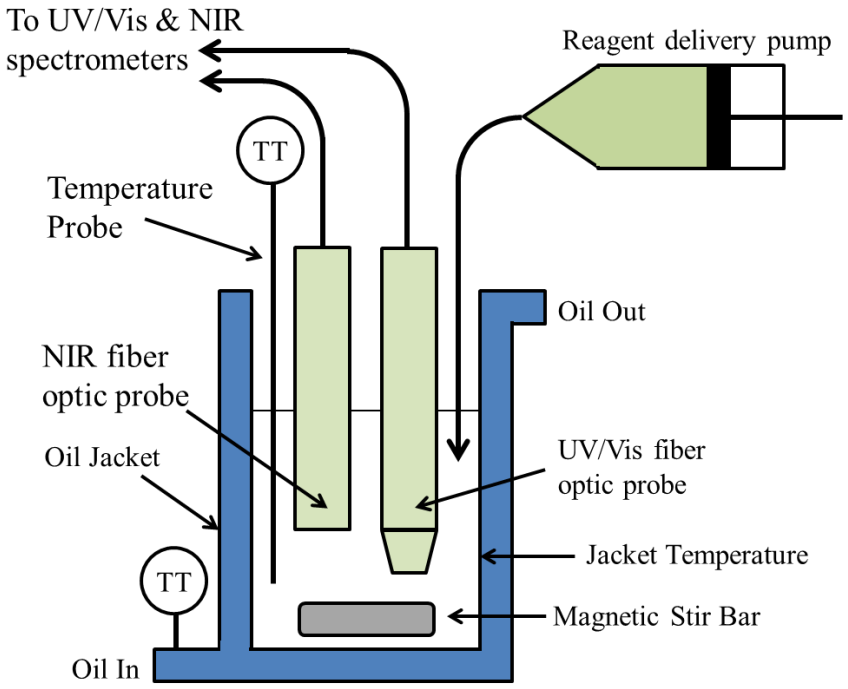

Figure 3.7: Schematic representation of the in-house miniature semi-batch reactor.

Before a batch experiment was initiated, the jacket temperature was set to a precisely controlled constant value of 25 °C. The batch temperature was monitored using a calibrated stainless steel temperature probe in the reactor.

3.3.2 Spectroscopic Instrument and Data Acquisition

NIR and UV-vis spectra were collected in the batch reactors using similar acquisition parameters. NIR measurements were made using a FOSS NIRSystems (Silver Spring, MD) model 6500 scanning spectrophotometer with a resolution of 4 nm in the 1100-2500 nm regions, and a diffuse reflectance fiber optic probe (FOSS, 1.5 m, 1/2" OD x 12", sapphire lense). NIR spectra were measured at a rate of 1 per second. A total of 10 were averaged at every 30 second interval using Vision software designed by FOSS NIRSystems. A ceramic disk was used to measure a reference spectrum before each run.

UV-vis measurements were made using a tec5USA (Plainview, NY) multi-channel spectrophotometer with a resolution of 3 nm in the 190-1100 nm regions, and an ATR fiber optic probe (Hellma 661.821 model, 205 mm, 125 mm OD, sapphire crystal). UV-vis spectra were measured at a rate of 1 per second. A total of 10 were averaged at every 30 second intervals using MultiSpec Pro Process software designed by tec5USA. For all batches, dark current spectra were acquired before the start of each experiment, and ethanol-water mixture was used as the reference spectrum.

3.3.3 Experimental Design

For each batch experiment, the reactor was initially charged with 20 mL of a solvent mixture (48% water, 52% ethanol) and maintained at 25°C. Four precisely measured aliquots of solid of salicylic acid were then introduced into the reactor at ten minute intervals. A slurry of

solid SA in equilibrium with a saturated solution was achieved after the fourth addition. At this stage, an independent measurement was made to estimate the volume of the heterogeneous mixture using a graduate cylinder. A pycnometer was used to measure the density. The volume of the heterogeneous mixture was 22.7 mL and the density was 0.90 g mL^{-1}.

The reactor temperature was then increased to 35°C in order to dissolve all the remaining solid material. Subsequently, the temperature was decreased to 25°C to produce a supersaturated solution at a metastable state. Finally, the solution was seeded with 2.50 g seed crystals (i.e. salicylic acid). An automated syringe pump was then used to deliver fresh solvent in the second dissolution phase of the experiment. The dissolution study was conducted by introducing six aliquots of 4 mL of fresh solvent into the reactor at twenty or forty minute intervals. Collection of NIR and ATR UV/Visible spectra were continued every 30 seconds during the dissolution phase of the experiment.

3.3.4 Kinetic Models for Slurries

Slurries used in reaction mixtures have competing processes of dissolution and precipitation. To model these processes, population balance equations (PBEs), which we refer as high-theory models, are often used to study these kinds of mechanisms within a system. In our application of kinetic modeling, we assumed that our batch processes were well mixed, volume changes were addictive, and the length, surface and volumetric shape factors for the crystals remained approximately constant throughout the duration of the experiment. We also assumed there was no agglomeration of crystal particles or breakage of large crystals due to shear forces from mixing.

Crystal growth is usually described by a thin film model with a mass transfer process across the diffusion layer surrounding the crystal in competition with solute molecule deposition

onto the surface of the crystal. Equations 3.1 and 3.2 describes what we refer to as high-theory models derived from population balance equations (PBEs) that are often used to describe crystallization and dissolution processes.

$$r_c = \frac{\Phi_s MW_s k_c}{3 d_s \Phi_v} \eta_r (c - c_{sat})^g \tag{3.1}$$

$$r_d = \frac{2 MW_s k_d}{d_s} (c_{sat} - c) \tag{3.2}$$

In Equations 3.1 and 3.2, r_c and r_d are the crystallization and dissolution rates of the solute, respectively. η_r is the effective size factor introduced by Garside [68], k_c and k_d are the kinetic rate constants, MW_s is the molecular weight of solute, d_s is the density of the solute, c_{sat} is the solubility of the solute, c is the solute concentration in bulk solution, and g is the rate order. Φ_s and Φ_v are the surface and volumetric shape factors, respectively.

$$r_c = k_c (c - c_{sat})^g \tag{3.3}$$

$$r_d = k_d m (c_{sat} - c)^n \tag{3.4}$$

$$r_{SA,in} = \frac{f_{1,t}}{V_t} C_{SA,in} \tag{3.5}$$

The above high-theory models can be simplified to provide low-theory models under the assumptions made in the previous paragraphs. In this work, the above simplified models were used to describe crystal growth and dissolution, Equation 3.3 describes the crystal growth process where k_c represents the accumulation of all of the parameters and constants in Equation 3.1. Equation 3.4 describes the dissolution process, where m is the remaining undissolved salicylic acid solid. Equation 3.5 describes the seeding step, where $r_{SA,in}$ is the seeding rate per unit time (mol/L), $f_{1,t}$ is the solid volumetric flow-in rate (L per unit time), V_t is the total volume of the slurry (L), $C_{SA,in}$ is the salicylic acid concentration of the solid in (mol/L).

3.3.5 Beer-Lambert's Law

According to Beer's law, shown in Equation 3.6, absorbance data collected in matrix **Y** (*nt* x *nw*) is the sum of the contributions from all *nc* absorbing species, *nt* and *nw* being the number of times and wavelengths, respectively.

$$\mathbf{Y} = \mathbf{C}\,\mathbf{A} + \mathbf{R} \tag{3.6}$$

Matrices **C** (*nt* x *nc*) and **A** (*nc* x *nw*) contained the concentration profiles and pure component spectra (molar absorptivities), respectively. The columns in matric **C** which represents the concentration profiles are produced by numerical integration of the differential equations specified in the first-principles kinetic model. Nonlinear optimization of the model parameters is used to adjust concentration profiles that match the experimental measured data. Since measurements are never flawless, a matrix of residuals **R** (*nt* x *nw*) was included in Equation 3.6 to take into account instrumental noise and model lack of fit.

3.3.6 Surface Enhancement Effect

Under supersaturated conditions, we observed that the measured ATR UV/Visible absorbance signal was higher than expected by our model and calibration curves. We believe this was due to weak adsorption phenomena at the immediate surface of the ATR crystal (discussed in more detail later). Our model was adjusted to incorporate a correction factor *cf* to model this phenomena, using the correction factor *cf* as an optimizable variable.

$$\mathbf{s} = \mathbf{c} - c_{sat} \tag{3.7}$$

$$\mathbf{s} = 0 \text{ when } c_{sat} > \mathbf{c} \tag{3.8}$$

$$\mathbf{c}^{*} = \mathbf{c} + (cf \cdot \mathbf{s}) \tag{3.9}$$

Vector **c** represents concentration of the solute, salicylic acid, in solution, was adjusted according to Equation 3.9 and the vector **s** represents the degree of supersatration as a function of time estimated by the model which is given in Equation 3.8. Scalar *cf* is an adjustable parameter, which is used to produce the resulting compensated concentration profile.

3.3.7 Kinetic Model and Numerical Integration

Kinetic modeling starts by postulating a kinetic model **H** (Equations 3.3-3.5) which are summarized in Equation 3.14. Model **H** is based on elementary dynamic processes, in this case crystallization and dissolution. The kinetic model is translated into a system of Ordinary

Differential Equations (ODE), Equations 3.10-3.12, which are numerically integrated and result in the *nc* concentration profiles in **C**.

$$\frac{dC_{SA}}{dt} = -r_c + r_d - (\frac{f_{2,t}}{V_t} \cdot c) \tag{3.10}$$

$$\frac{dM_{SA}}{dt} = Ms_{SA} \cdot V_t \cdot (r_c - r_d + r_{SA,in}) \tag{3.11}$$

$$\frac{dV_{liq}}{dt} = f_{2,t} - \frac{Ms_{SA} \cdot V_t}{D_{SA}} (r_c - r_d) \tag{3.12}$$

$$\mathbf{b} = [k_d, k_c, n, cf] \tag{3.13}$$

$$\frac{dC_{SA}}{dt} = \mathbf{H(b, v, \frac{dV_{liq}}{dt}, r_d, r_c)} \tag{3.14}$$

Model **H** is a system of ODEs depending on a four-dimensional vector of model parameters **b**, Equation 3.13, volume **v** as a function of time, flow rate $\frac{\mathrm{dV_{liq}}}{\mathrm{dt}}$, and dissolution and crystallization rates $\boldsymbol{r_d}$ and $\boldsymbol{r_c}$. Model **H** is numerically integrated using initial conditions $\mathbf{c_0}$.

The rate of dissolution $\boldsymbol{r_d}$ collects all rates $r_{d,i}$ of each i^{th} dissolution step, which are based on a power law equation, as shown in Equation 3.15.

$$r_{d,i} = k_{d,i} m \, (c_{sat,i} - c_i)^{n_i} \tag{3.15}$$

In Equation 3.15, $k_{d,i}$ is the dissolution rate constant, $c_{sat,i}$ is the saturation limit and n_i the power coefficient of the i^{th} step. All dissolution rate constants $k_{d,i}$ and power coefficients n_i are collected in vectors $\boldsymbol{k_d}$ and $\boldsymbol{n}$, respectively.

The rate of crystallization $\mathbf{r_c}$ collects all rates $r_{c,i}$ of each i^{th} crystallization step, which are also based on a power law equation, as shown in Equation 3.16.

$$r_{c,i} = k_{c,i}\,(c_i - c_{sat,i})^{g_i} \tag{3.16}$$

In Equation 3.16, $k_{c,i}$ is the crystallization rate constant and g_i is the power coefficient of the i^{th} step. All crystallization rate constants $k_{c,i}$ and power coefficients g_i are collected in vectors $\boldsymbol{k_c}$ and g, respectively. Finally, the vector of model parameters **b** contains vectors $\boldsymbol{k_d}, \boldsymbol{k_c}, \boldsymbol{n}$ *a*nd scalar *cf* .

3.3.8 Elimination of the Linear Parameters

At each iteration of the non-linear optimization steps, pure component spectra $\mathbf{A_1}$ for ATR UV/Vis measurements are linearly estimated in the 260-370 nm region using $\mathbf{C}$ and $\mathbf{Y_1}$ (Absorbance), as shown in Equations 3.6 and 3.17. With this concept, we can also separate the linear parameter in the NIR reflectance measurement at 1100 nm. At 1100 nm, absorption by the components in the slurry is negligible thereby providing a measure of light-scattering from particles in the slurry as a function of time, $\mathbf{y_2}$, free from the interference of any absorption signals. In Equation 3.18, **m** represents the model estimated solid fraction profiles as a function of time, and a_2 is the NIR reflectance coefficient.

$$\mathbf{A_1} = \mathbf{C^{+}Y_1} \tag{3.17}$$

$$a_2 = \mathbf{m^{+}y_2} \tag{3.18}$$

$\mathbf{C^{+}}$ and $\mathbf{m^{+}}$ are called the left pseudo-inverse and are computed as $\mathbf{C^{+}} = (\mathbf{C^{T}C})^{-1}\mathbf{C^{T}}$ and $\mathbf{m^{+}} = (\mathbf{m^{T}m})^{-1}\mathbf{m^{T}}$.

3.3.9 Non-Linear Optimization

Standard Newton-Gauss-Levenberg/Marquardt (NGL/M), one of the most commonly used non-linear gradient optimizers, was used to optimize the model parameters. NGL/M relies on the calculation of the derivative of the residuals **R**. The derivatives of optimizable parameters required in the NGL/M method were estimated by the technique of finite difference approximation because analytical derivatives were not all know. The method of finite different approximation was used with a relative shift value 1 x 10^{-6}. For the convergence criterion, we used a relative change of 1 x 10^{-4} in the sum of square residual from 1 iteration step to the next.

Boundary conditions on the adjustable parameters were implemented into the NGL/M function. If the values of parameters after the computed the shift exceed a maximum or a minimum value, then the parameters were adjusted to the upper or lower limits, respectively. The convergence of gradient methods is superior to other methods as long as the initial guesses for the parameters are close to their optimal values [79]. According to Beer's law Equation 3.6, residuals for the liquid phase $\mathbf{R_1}$ and the solid phase $\mathbf{r_2}$ can be calculated as:

$$\mathbf{R_1} = \mathbf{Y_1} - \mathbf{C\,A_1} \tag{3.19}$$

$$\mathbf{r_2} = \mathbf{y_2} - \mathbf{m}\, a_2 \tag{3.20}$$

The sum of squares, *ssq*, is the sum of all squared residuals $\mathbf{R_1}$ and $\mathbf{r_2}$.

$$ssq = \sum_{i=1}^{nt} \sum_{j=1}^{nw} (\mathbf{R_1}_{i,j}^{\,2} + \mathbf{r_2}_{i,j}^{\,2}) \tag{3.21}$$

The least-squares optimization problem is then defined as $\min_{\mathbf{p}} ssq$.

3.4 Results and Discussions

Figure 3.2 shows the entire record of an experimental run, including ATR UV/Visible absorbance at 307 nm (A), NIR reflectance at 1100 nm (B) and temperature (C). Figure 3.2-A shows the UV/Visble absorbance during one experiment at 307 nm. During the time interval from 15 min to 167 min, 4 aliquots of solid salicylic acid are added, and the dissolution gives a stepwise increase in the absorbance profile. These precisely measured amounts were used to generate UV-Vis calibration data. It is well known that changes in temperature produce small changes in the refractive index of the liquid solvent which could affect the reliability of the ATR measurement. In order to evaluate the significance of this effect, at 78 min, the temperature of the solution was increased from 25 °C to 35 °C which produced a corresponding 0.5% decrease in the UV/Visible absorption at 307 nm. This change in absorbance as a function of the temperature was considered negligible for the purpose of our model fitting work.

At 115 min, the temperature was reduced from 35 °C back to 25 °C, and the absorbance returned to the original level. At this point in the experiment, only homogenous solution was present in the reactor. At 167 min, the last addition of salicylic acid produced a saturated solution.

Only about 30% of the added material dissolved and a slurry was formed. At 204 min, the temperature of the batch mixture was increased to 35 °C, and the remaining solid salicylic acid dissolved producing a homogenous solution. At 252 min, the temperature of the batch mixture was cooled to 25 °C producing a supersaturated solution. A corresponding significant increase in the absorption measured by the ATR crystal was observed. This change was much larger (about 3.1 %) compared to the change observed at 78 min. Experimental evidence suggested the increase was most likely due to weak adsorption phenomena in the evanescence field near the ATR crystal surface. Nucleation and crystallization on the surface of ATR crystal can be ruled out by the following logic: Prior to 25 °C the probe was submerged in a solution held at 35 °C for thirty minutes. The solution was then cooled to 25 °C quickly. The residual heat in the probe ensured that it stayed warmer than the surrounding solution, thus preventing formation of nuclei or crystallization of salicylic acid on the surface of ATR crystal.

At 282 min, 2.50 g of seed crystals were introduced at which point the supersaturated solution quickly crystalized and produced a saturated solution in equilibrium with the solid. At 302 min, 4 mL of fresh mixture solvent was added, and a dissolution profile was observed. Another 4 mL of fresh mixture solvent was added at 342 min, 382 min, 422 min, 462 min and 502 min. After the 5^{th} solvent addition at 462 min, an under-saturated solution was produced. At 502 min, only dilution effects were observed and no dissolution.

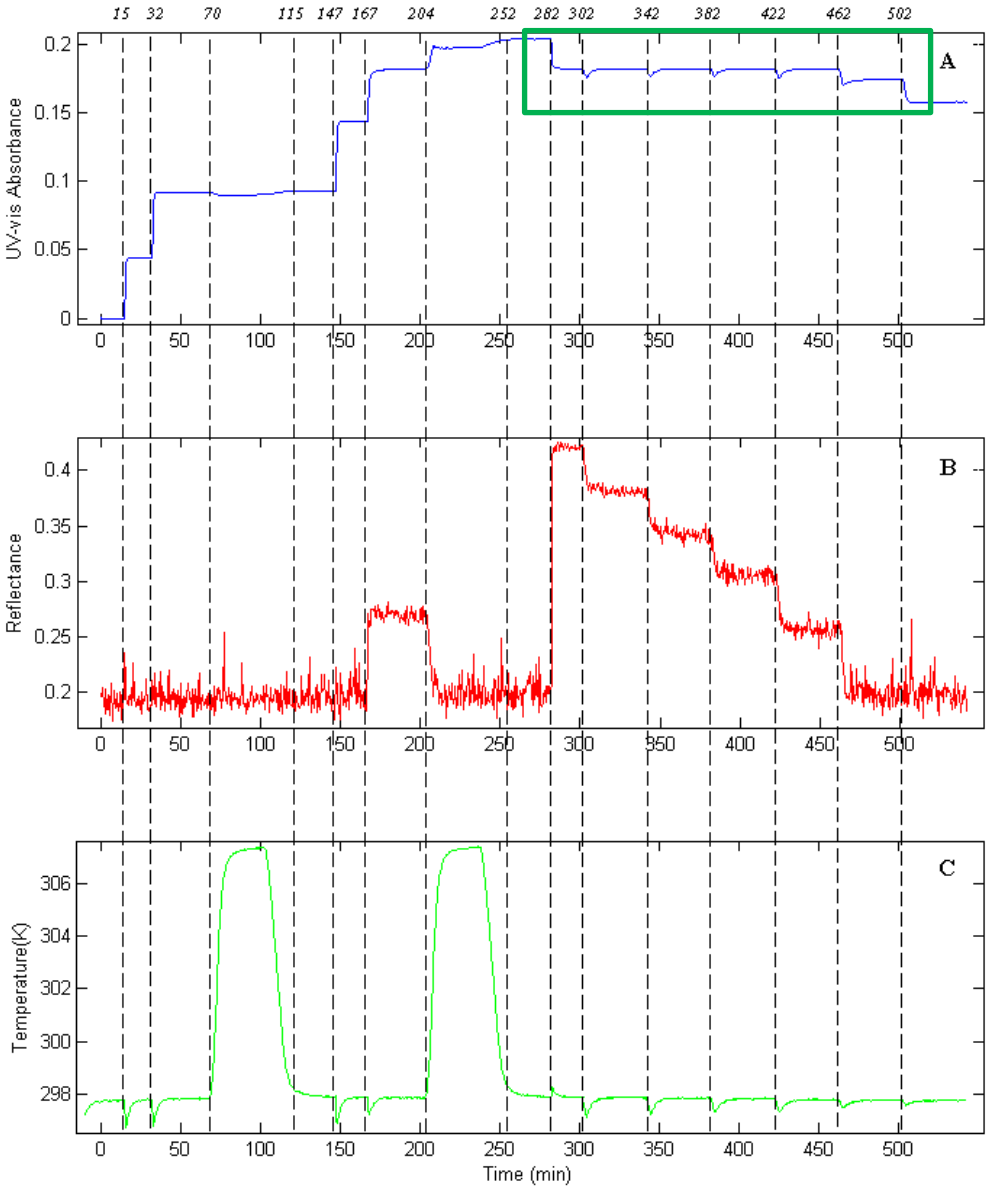

Figure 3.8: A comprehensive record of an experiment run, including time resolved ATR UV/Visible absorbance at 307 nm (A), NIR reflectance with preprocessing treatments at 1100 nm (B) and temperature (C) data.

In Figure 3.2-B, the NIR reflectance at 1100 nm indicates the presence of solid material at 167 min to 204 min, and at 282 min to 462 min. At 282 min when slurry was produced there was a considerable increased in the NIR reflectance signal can observe.

In Figure 3.2-C corresponding changes in the temperature can be observed. When solid materials at 25 °C were introduced a temporarily decrease in temperature was observed as energy was absorbed by the dissolution process. In Figure 3.3, the upper panel shows an expanded view of the region from Figure 3.2-A from 262 min to 512 min. The blue-dotted line shows the ATR UV/Visible measurement at 307 nm, and the red continuous line shows the model-estimated concentration of dissolved salicylic acid obtained by fitting ATR UV/Visible and NIR reflectance measurements. The bottom panel shows the model-estimated amount of solid salicylic acid remaining in the slurry as a function of time.

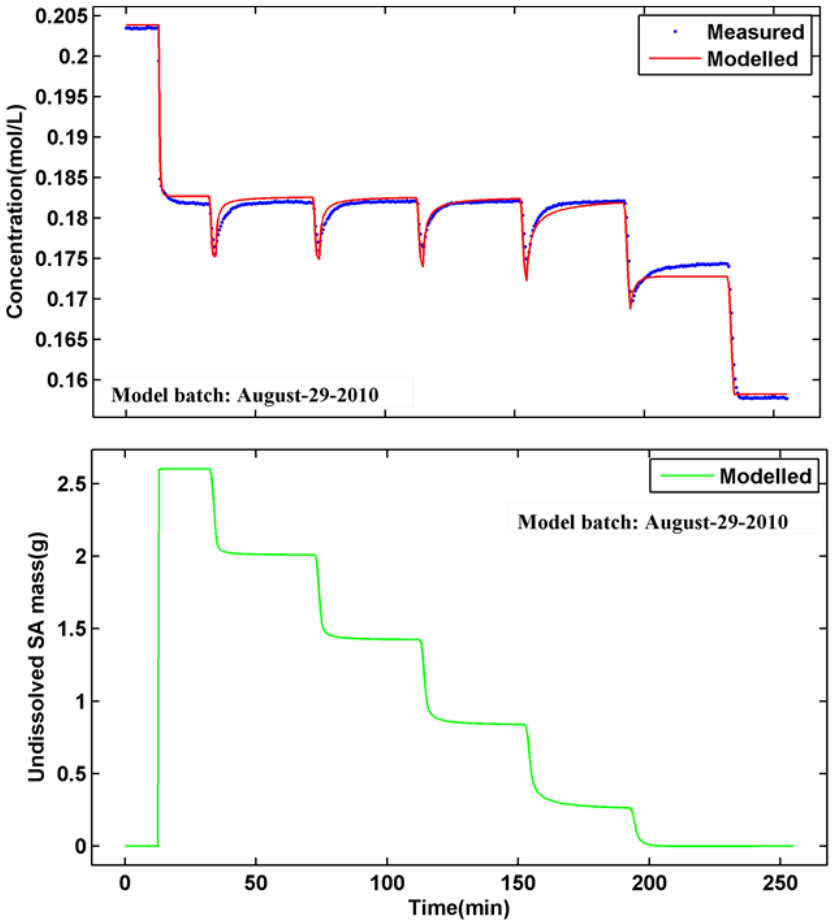

Figure 3.9: Upper panel shows an expanded view of the region from Figure 3.2-A from 262 min. to 512 min, and the bottom panel shows the model estimated solid profile, the remaining salicylic acid solid as a function of time (elapsed time, in min).

In order to evaluate the reproducibility of our experimental protocol and the reproducibility of the model fitting process, three additional replicate batches were performed but at a shorter time frame. This ensures the model was robust with respect to changes in the experimental time frame. Three replicates were run at half of the time scale of the original batch (i.e. fresh solvent was added at twenty minute intervals rather than forty minute intervals).

In Figure 3.4, only the ATR UV/Visible profiles (liquid phase) are shown for brevity's sake. The near-infrared measurements and temperature data are not shown but are comparable to these shown in Figure 3.2-A, 3.2-B and 3.2-C. Figure 3.4 shows good agreement between the modeled dissolution curves and the experimental dissolution curves; however, slight variations might be caused by small departures from the assumptions listed below.

In these experiments, we assumed the slurries were well-stirred. We know that this does not strictly hold because there are small dead zones in the reactor behind the ATR UV/Visible probe and the NIR reflectance probe. We also assumed the slurry is sufficiently well-mixed such that all particles of SA were suspended relatively uniformly throughout the reactor vessel, and that the stirring rate was relatively constant. Small changes in the stirring rate from batch to batch will show up as differences in the estimated rate constants for dissolution and crystallization, respectively. Last but not least, we also assumed the particle size distribution of the dissolving solid material remains approximately the same throughout the experiment. For the purpose of this empirical model fitting process, a reasonably satisfactory fit was obtained.

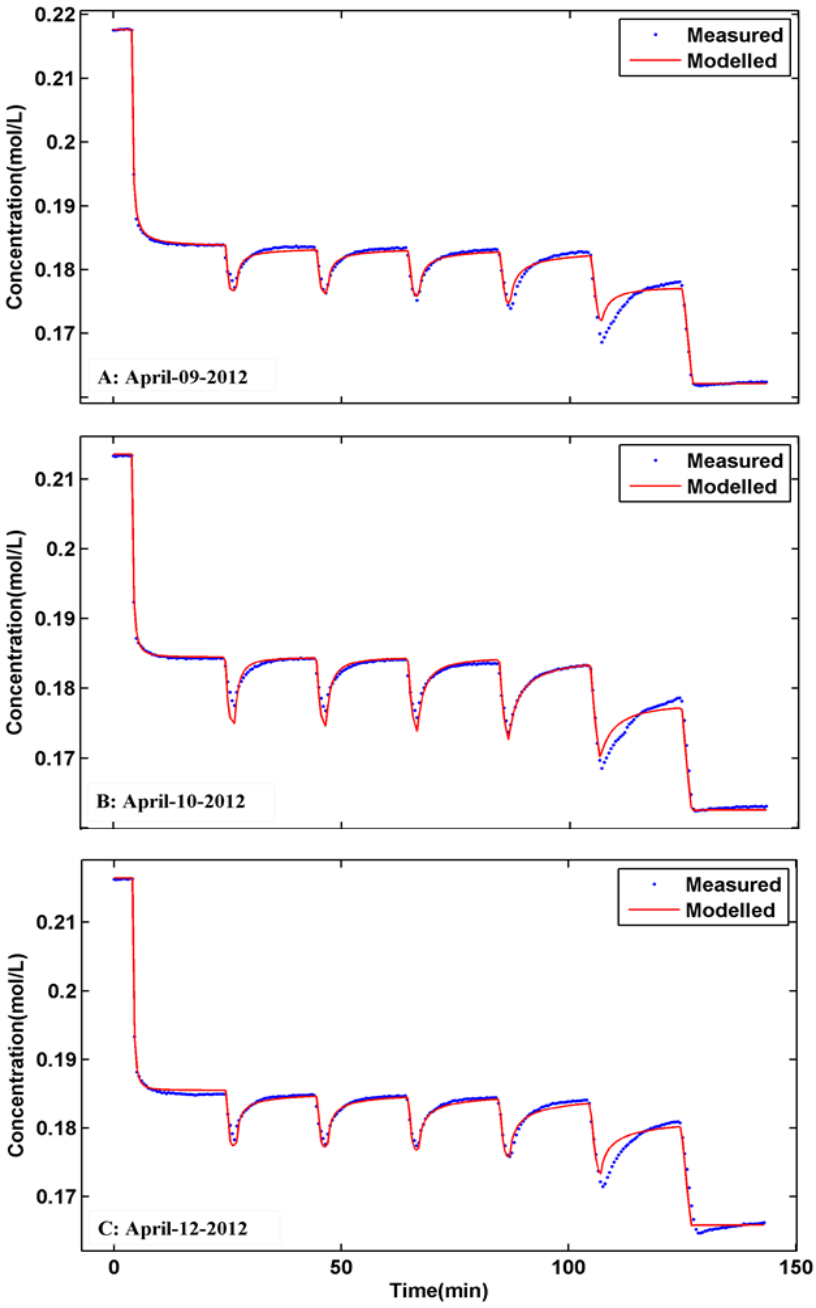

Figure 3.10: Time resolved measured (dots) of three selected replicate batches at a set point of 25 C°, and fitted (line), ATR UV/Visible absorbance data at 307 nm.

Figure 3.5 shows the indirect model estimated profile of undissolved salicylic acid of three selected replicate batches run obtained by fitting ATR UV/Visible pure component spectral

data (see Equation 3.17-3.18). NIR reflectance data was not used to produce these results. Indirect estimation of the solid fraction is possible because the integrated dynamic models enforce strict mass balance based on the precisely specified initial conditions of each batch experiment. In this respect the procedure is somewhat like a dynamic spectrophotometric titration.

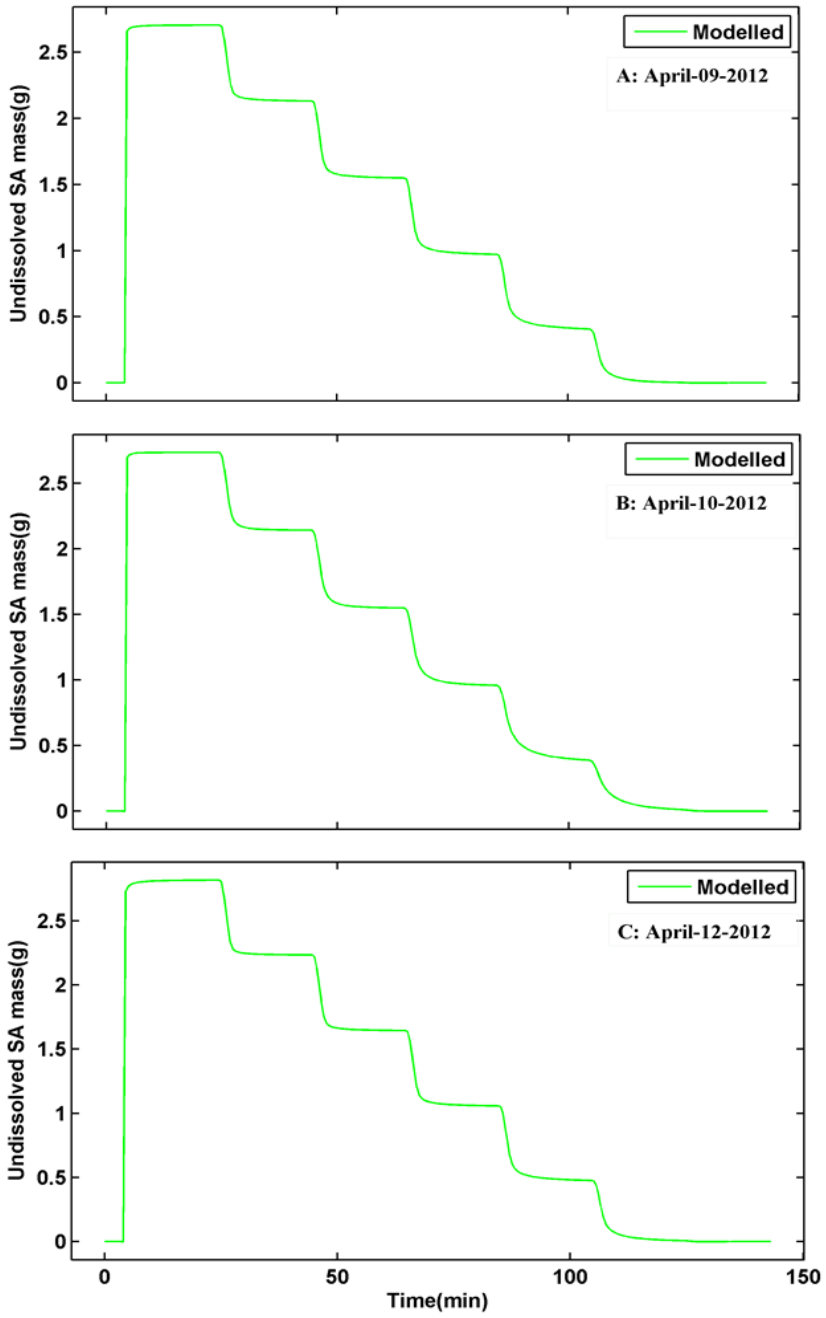

Figure 3.11: Model estimated profiles for the un-dissolved salicylic acid solid of three selected replicate batches run by fitting ATR UV/Visible pure component spectra (Equations 3.18-3.19).

Figure 3.6 shows the comparison of the model-estimated solid fraction profiles with Partial Least Squares (PLS) results for the first three dissolution steps. PLS is used here for validation purposes.

A separate set of experiments were conducted to precisely determine the solubility of salicylic acid in our solvent mixture (48% water, 52% ethanol) at 25 °C. A total of eight calibration mixtures were prepared from a saturated solution of salicylic acid with precisely measured amounts of undissolved salicylic acid. These mixtures were well-stirred and the NIR reflectance spectra were measured from 1100 nm to 1700 nm. PLS calibration was performed using three latent variables and MATLAB code written in-house using the algorithm of Lorder et. al [65]. It should be emphasize that reflectance was used in the PLS calibration rather than conventional log(1/R), as this gave the best linearity with the fewest number of factors.

Dissolution steps 4 and 5 are not included in Figure 3.6, because the amount of solid present in the reactor is too low to give reliable NIR reflectance signal. It is well known that NIR reflectance probes are not suitable for measuring systems with low light scattering properties.

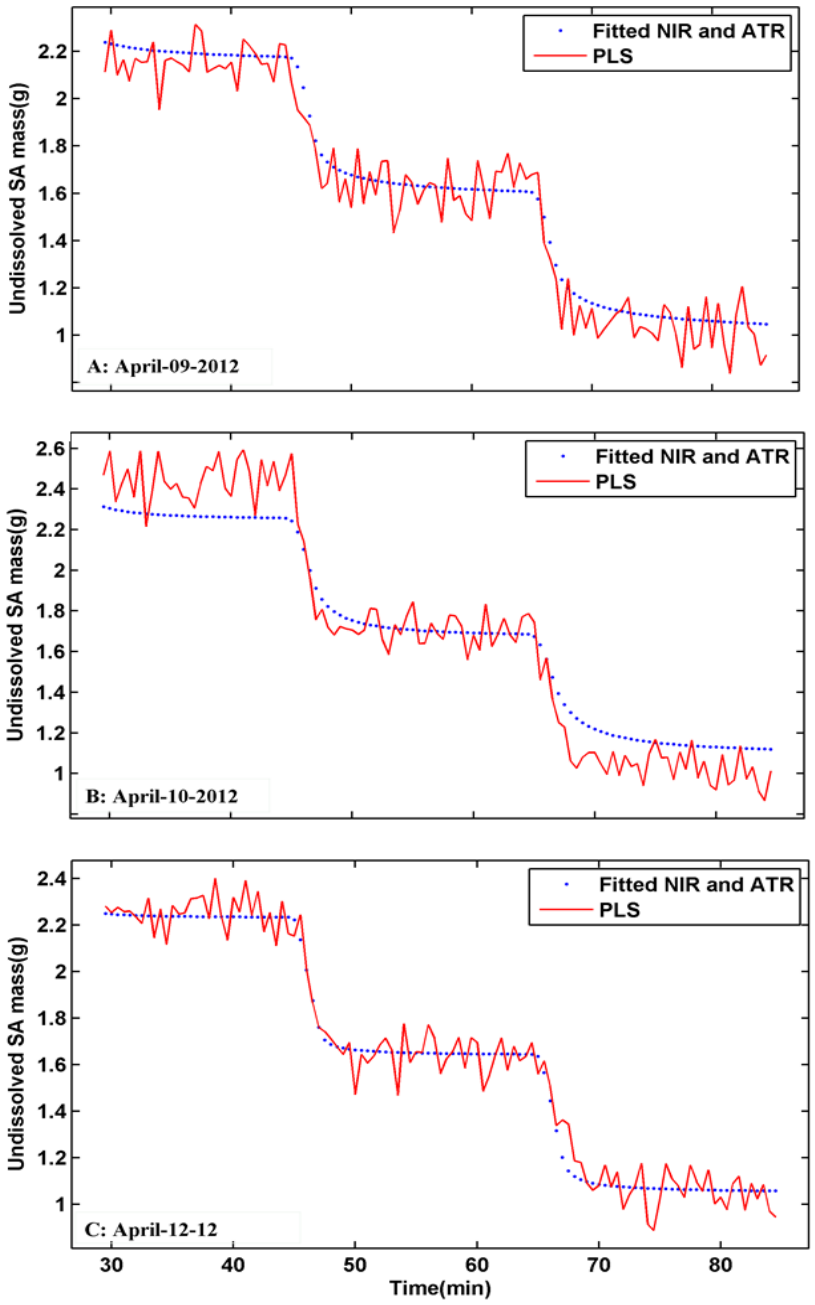

Figure 3.12: Comparison of the model estimated solid fraction profile (dots) as a function of time for three selected replicate batches run with PLS (line). Both ATR UV/Visible absorbance data at 307 nm and NIR reflectance data at 1100 nm is fitted into kinetic models.

Table 3.2: Comparison of Kinetic Model Fitting Results for UV/Visible Only, and UV/Visible fused with NIR Reflectance Data.

	Batch A		**Batch B**		**Batch C**		**Average**	**STD**	**RSD**
	Input	*Optimized*	*Input*	*Optimized*	*Input*	*Optimized*			
UV/Visible spectra only									
k_d ($L^{n-1}/(mol^{n-1}min)$)	79.41	67.90	59.43	35.38	109.9	55.838	53.04	16.44	31.00
k_c ($L^{n-1}/(mol^{n-1}min)$)	35.86	30.04	27.52	27.21	30.27	29.374	28.88	1.481	5.127
n	2.228	2.505	2.426	2.280	2.315	2.453	2.412	0.1181	4.897
g	1.500	-	1.500	-	1.500	-			
$c_{sat\ (mol/L)}$	1.004	-	1.004	-	1.004	-			
UV/Visible & NIR spectra									
k_d ($L^{n-1}/(mol^{n-1}min)$)	52.10	59.43	52.10	52.30	109.9	52.10	54.61	4.174	7.644
k_c ($L^{n-1}/(mol^{n-1}min)$)	28.78	27.52	28.78	32.92	30.27	28.78	29.74	2.821	9.485
n	2.576	2.426	2.576	2.789	2.315	2.576	2.597	0.1822	7.017
g	1.500	-	1.500	-	1.500	-			
$c_{sat\ (mol/L)}$	1.004	-	1.004	-	1.004	-			

Table 3.1 shows the comparison of the kinetic model fitting results. The upper half of the table shows the results obtained by fitting UV/Visible measurements, and the bottom half of the table shows the results obtained by fitted both UV/Visible and NIR reflectance measurements. The averaged fitting results from upper and lower portion of the table are quite similar to one to another. The purpose of the model fitting process is not to determine fundamental constants in properties of dissolution and crystallization rates. Instead, the purpose is to develop an empirical

model that gives a statistically sufficient phenomenological description of the batch process for monitoring and control.

3.5 Conclusion

In this paper, we have demonstrated an advanced measurement and modeling strategy to monitor the dissolution and crystallization of salicylic acid in a mixed solvent. The solid phase was measured using in situ diffuse reflectance NIR spectroscopy, while the liquid phase was measured using in situ ATR UV/Visible spectroscopy, respectively. In order to determine the kinetic parameters involved during the crystal growth and dissolution processes, a low-theory, simplified comprehensive global model describing these processes as a function of time was solved using MATLAB. The model was coupled with a MATLAB nonlinear optimization subroutine. The suggested kinetic laws were sufficient in describing the set of four crystallization and dissolution experiments performed during the study. The parameters estimated in the model included dissolution and crystallization rate constants, dissolution rate exponent, as well as the degree of supersaturation which provides the driving force for crystallization. In addition, the methods required to carry out the fitting of a kinetic model to measured multivariate spectroscopic data of slurries (i.e. NIR and UV-vis spectra) have been demonstrated, from the postulation of the model and the derivation of the differential equations, through the numerical integration of the model to yield the concentration profiles and finally the calculation of the pure component spectra and fitting the models' rate constants to measured data.

3.6 Acknowledgement

This research was supported by the National Science Foundation (NSF) under Grant Number CHE-0750287, Grant Opportunities for Academic Liaison with Industry (GOALI) and

the E.I. DuPont de Nemours and Co., Inc., Crop Protection Products and Engineering Technologies. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

3.7 Notations

b	vector of parameters
c_{sat}	solute solubility (mol/L)
c	solute concentration (mol/L)
$C_{SA,in}$	salicylic acid concentration (mol/L)
d_s	solid density (g/L)
$f_{1,t}$	solid flow-in profile (L/ut)
$f_{2,t}$	liquid flow-in profile (L/ut)
g	crystal growth exponent
k_c	crystallization rate constant $L^{n-1}/(mol^{n-1}min)$
k_d	dissolution rate constant $L^{n-1}/(mol^{n-1}min)$
MW_s	molecular weight of solute (kg mol^{-1})
n	dissolution exponent
r_c	growth rate (mol/L/ut)
r_d	dissolution rate (mol/L/ut)
r_{SA}	seeding rate (mol/L/ut)
S	supersaturation
t	time (s)

V_t total volume (L)

η_r effectiveness factor

Φ_s, Φ_v surface and volumetric shape factors.

4. KINETIC HARD-MODELLING OF BATCH SLURRY REACTIONS

4.1 Abstract

The use of slurry mixtures in chemical and pharmaceutical manufacturing processes is common due to the advantages that this method provides. Kinetic modeling of large scale batch reactions has been shown to be an excellent tool in on-line batch reaction monitoring [41, 80-83]. To date, little work has been done in modeling slurry reactions. Presented here is a kinetic modeling method for a commercially relevant slurry reaction that produces a solid suspension of the product. The model was built using near infrared (NIR) diffuse reflectance measurements taken by a submergible fiber-optic probe inside of a jacketed reactor vessel. Measurements were confirmed by off-line high performance liquid chromatography (HPLC) analysis with samples taken from a recirculation loop, and using focused beam reflectance measurement (FBRM) measurements to observe the product precipitation event. The system chosen was the reaction of a heterocycle triazine suspended in xylene with an isocyanate to form a metsulfuron methyl product precipitant.

KEY WORDS: Kinetic modeling; Spectroscopy; Slurries; Metsulfuron methyl; Nonlinear fitting

4.2 Introduction

The manufacturing of chemical and pharmaceutical products has greatly benefitted from the use of slurry mixtures in the production process. There are many advantages to using slurries; for instance, the production of a product that is simply suspended in a liquid solvent is much easier to separate than a liquid product mixed with a liquid solvent. Simply drying the product

instead of proceeding through a series of separation steps saves the manufacturer time and money. Also, depending on the situation, the amount of solvent used to suspend the materials can be minimized and more easily and recycled for use in subsequent batches, which saves the manufacturer on cost of materials. It is for reasons such as these that the use of slurries in industrial reactions is common. However, batch reaction monitoring becomes more difficult when the materials being measured are in more than one phase.

Modeling of batch reactions in industry is typically done using a Partial Least Squares (PLS) method. However, these methods often require extensive and costly calibration data sets and may not be very robust to small modifications and product formulation. Kinetic modeling, on the other hand can be a more robust system that can compensate for changes in the experimental protocol as long as the appropriate mathematics have been added to the model set-up. For instance, if changes in reactor temperature are frequent, as long as the appropriate calibration has been added to the kinetic model parameters, then the modeled data should reflect the temperature changes, even in real time. If the mathematics in the kinetic model accurately represents the mechanisms of the reaction itself, then the model should be able to approximate a set of spectra that accurately reflects the experimentally measured data. A tool such as this is beneficial in industry for its ability to detect process upsets errors in real time. For example, if the measurements from a production batch do not fall within an acceptable range predicted by the model, then the error can be noticed before the batch reaches the product testing stages. An error that is spotted early could even potentially be rectified. If there were a mistake in the initial charge of starting materials, the batch mixture could be adjusted by the addition of reactants that would rectify the error and prevent the waste of time and materials.

Presented here is a functional kinetic model for a commercially relevant slurry system. The model characterizes the dissolution of a heterocycle triazine slurry that is suspended in xylene, its reaction with an isocyanate solution that is pumped in, and the subsequent precipitation of the metsulfuron methyl product.

4.3 Theory

Kinetic models are made up of a series of rate law equations that are used to build ordinary differential equations (ODEs) necessary to construct concentration profiles for all species present in the reaction system. Initially, certain values, such as the reaction rate constants, are unknown due to the fact that they vary by system and initial conditions. In order to obtain these values, a nonlinear optimization of the parameters must be performed to estimate them. Once they have been calculated, the concentration profiles should reflect the actual changes in concentration for each species present in the reaction and can be used to calculate the model fit. A tutorial for the kinetic model fitting process is described by Puxty [79], and will only be summarized here.

4.3.1 Building Concentration Profiles

The construction of concentration profiles is dependent on the use of a series of ODEs that establish the changes in concentration for each component in the system as a function of time. In order to create the ODEs, however the reaction mechanisms of the system must be known. These mechanisms are necessary to form the rate laws for each reaction and the ODEs are calculated based on these. A concentration profile that accurately portrays the changes that the system undergoes is necessary to establish a good model fit, so these rate laws and ODEs must describe the data as accurately as possible.

Construction of the ODEs can be quite simple, even if the reaction is somewhat complex. For instance, consider a two-step reaction mechanism:

$$A + B \xrightarrow{k_1} C \tag{4.1}$$

$$2\mathrm{C} \xrightarrow{k_2} \mathrm{D} \tag{4.2}$$

The rate laws for these reactions can easily be established.

$$r_1 = k_1 [A][B] \tag{4.3}$$

$$r_2 = k_2 [\mathrm{C}]^2 \tag{4.4}$$

The rates of change for each component of the reaction, A, B, C, and D, are established by these rate laws and are combined in the ODEs.

$$\frac{d[A]}{dt} = \frac{d[B]}{dt} = -r_1 \tag{4.5}$$

$$\frac{d[\mathrm{C}]}{dt} = r_1 - r_2 \tag{4.6}$$

$$\frac{d[\mathrm{D}]}{dt} = r_2 \tag{4.7}$$

In order to calculate the concentration at any given moment, these differentials must be integrated with respect to time. Unfortunately, for such a large and complex data set, this is impossible to do explicitly. For this reason, a numerical integration method must be relied upon to calculate the concentration profiles. In this work, the fourth order Runge-Kutta method of numerical integration is used [84].

4.3.2 Calculating the Residuals

In order to establish how well the modeled concentration profiles fit the experimental data, an estimation of the model estimated spectra must be made for comparison. This estimated set of spectra is calculated using a restatement of the Beer-Lambert law:

$$\mathbf{Y}_{\mathrm{calc}} = \mathbf{CA} \tag{4.8}$$

Where $\mathbf{Y}_{\mathrm{calc}}$ is the estimated set of spectra, **C** is a matrix containing the concentration profiles for each component of the reaction, and **A** is a matrix containing the molar absorptivity for each component. Because the matrix **A** is also unknown, it must be calculated using the calculated concentration profile and the measured experimental spectra.

$$\mathbf{Y} = \mathbf{CA} \Rightarrow \mathbf{A} = \mathbf{C}^{+}\mathbf{Y} \tag{4.9}$$

If **A** is replaced within the equation, it can be shown that this value is not necessary to compute the estimated spectra.

$$\mathbf{Y}_{\text{calc}} = \mathbf{C}\mathbf{C}^{+}\mathbf{Y} \tag{4.10}$$

A matrix of residuals **R** is computed by taking the difference of **Y** and $\mathbf{Y}_{\text{calc}}$.

$$\mathbf{R} = \mathbf{Y} - \mathbf{Y}_{\text{calc}} \tag{4.11}$$

In order to establish how well the modeled spectra fit the experimental spectra, the sum of squares of the residuals is calculated.

$$ssq = \sum r_{ij}^{\ 2} \tag{4.12}$$

In the unrealistic case that the two were a perfect match, the value of *ssq* would be equal to the measurement error in the experiment.

4.3.3 Nonlinear Optimization of Parameters

As previously mentioned, it is very unlikely that certain values, such as the reaction rate constants, are known when computing the concentration profiles. Fortunately, there are routines that can be used to optimize these values and give a close approximation to their true values. The

routine that is used here is the Newton-Gauss-Levenberg-Marquardt (NGLM) algorithm [85, 86]. This routine requires initial guesses for each of the unknown experimental parameters. Using these initial guesses, asset of modeled spectra is computed and the value for *ssq* is calculated. The algorithm then goes through a series of iterations wherein a shift for each parameter is calculated and a new set of model spectra is calculated. This is done repeatedly until the value of *ssq* converges.

4.4 Experimental

A series of three batch reactions were conducted at Stine-Haskell Research Center in Delaware, DE. In-situ NIR reflectance measurements were collected for all batches. The same chemical conditions, experimental protocol and process conditions were used for all batches. In this paper, the slurry-based synthesis of Metsulfuron Methyl (see Figure 4.1), a herbicide manufactured by E.I. DuPont was studied. The studies included the chemical coupling reaction and the dissolution of metsulfuron methyl in xylene. The sulfonylurea coupling reaction is a challenging system for reaction monitoring because one of the reactants and the product is only partially soluble, and being present in both solid and liquid phase of the slurry.

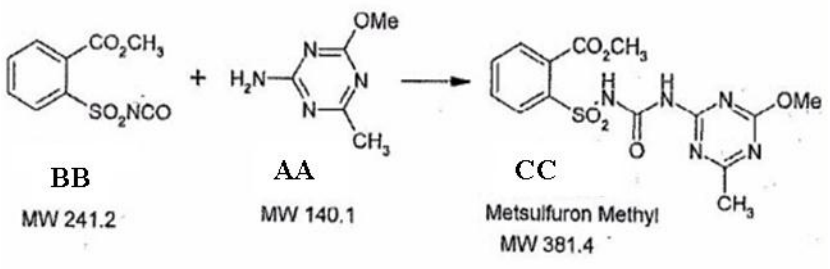

Figure 4.1: Sulfonylurea coupling reaction

4.4.1 **Reactor and Apparatus Setup**

The batch reactions reported in this paper were performed in a custom jacketed 1000 mL reaction vessel made at Stine-Haskell Research Center. As shown in Figure 4.2, the reactor and lid were specifically designed to accept a NIR reflectance probe, an overhead stirrer, a thermocouple, a FBRM probe (not shown), and an additional opening for the recirculation tube. The reactor jacket was thermostated using a heater/chiller oil bath. A calibrated balance and a peristaltic pump were used for CMBSI reagent delivery.

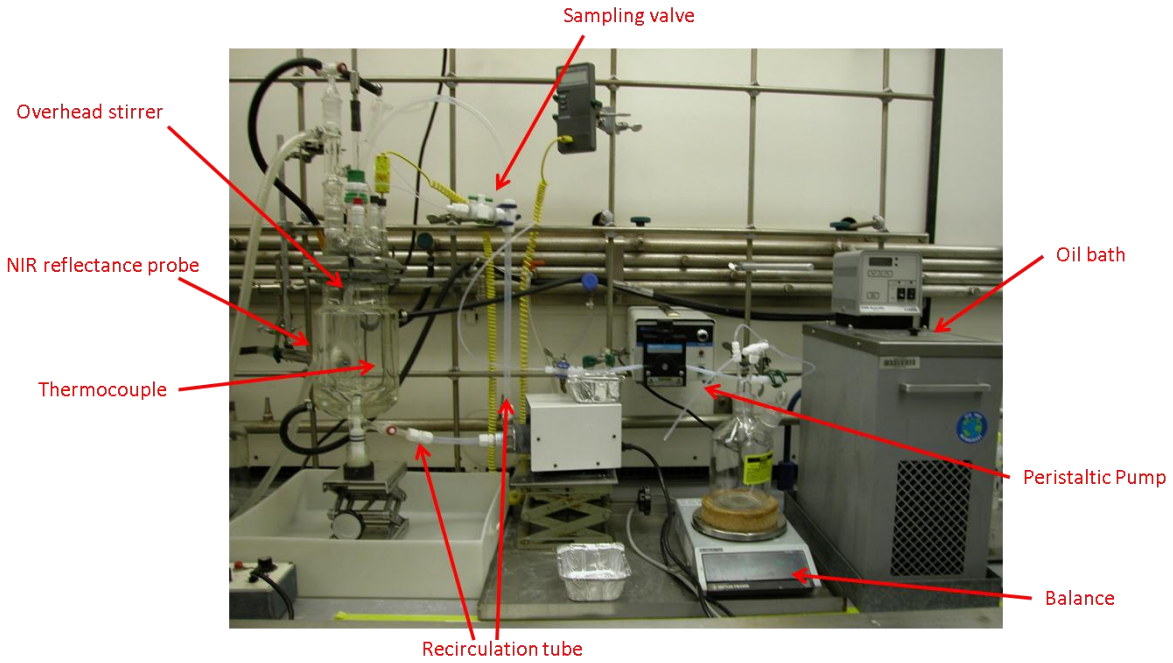

Figure 4.2: A picture of the lab-scale jacketed reactor and apparatus setup.

Before a batch reaction was initiated, the jacket temperature was set to a precisely controlled constant value of 85°C. The batch temperature was monitored using a calibrated thermocouple in the reactor.

4.4.2 Spectroscopic Instrument and Data Acquisition

NIR measurements were collected using a FOSS NIRSystems (Silver Spring, MD) model XDS Process Nema 4X monochromator in the 850-2200 nm regions, and a diffuse reflectance fiber optic probe (FOSS, 1.5 m, ½" OD x 12", sapphire lense). NIR spectra were measured at a rate of 1 per second. A total of 10 were averaged at every 30 second interval using Vision software designed by FOSS NIRSystem. For all batches, an internal reference spectrum was measured before each run.

4.4.3 Experimental Design

The sulfonylurea coupling reaction was carried out in a 1000 mL jacketed glass reactor. The reactants were 1,3,5-Triazin-2-amine, 4-methoxy-6-methyl- (AA, $C_5H_8N_4O$, molecular weight 140.14, CAS 1668-54-8) and benzoic acid, 4-isocyanato-2-[[[[(4-methoxy-6-methyl-1,3,5-triazin-2-yl)amino]carbonyl]amino]sulfonyl]-, methyl ester (BB, $C_{15}H_{14}N_6O_7S$ in C_8H_{10}, molecular weight 241.2, CAS 129346-33-4) and the product was benzoic acid, 2-[[[[(4-methoxy-6-methyl-1,3,5-triazin-2-yl)amino]carbonyl]amino]sulfonyl]-, methyl ester (CC, molecular weight 381.4, CAS 74223-64-6). Both AA and CC have a very limited solubility in benzene, dimethyl- (xylene, C_8H_{10}, molecular weight 106.16, CAS 1330-20-7). As BB is sensitive to moisture and water, and also harmful, it was quenched with isopropyl alcohol to form a derivative form of BB, benzoic acid, 2-[[[(1-methylethoxy)carbonyl]amino]sulfonyl]-, methyl ester, (DD, $C_{12}H_{15}NO_6S$, CAS 140617-87-4) when it was needed for calibration purpose.

The reactor was initially charged with 750 mL xylene solvent and 103 g of AA at 85°C. BB was dissolved in xylene and four additions of BB (115 g of CMBSI per addition) were delivered via peristaltic pump. Each BB addition corresponded to a quarter of the reaction, i.e. 25% of reaction (addition 1). After an equilibration time of 20 min, triplicate slurry samples were

taken from the reaction mixture and analyzed by HPLC to validate results obtained by NIR spectroscopy.

Figure 4.3 shows the HPLC elution order for AA, CC and DD at 230 nm using a gradient method (water and acetonitrile) [87].

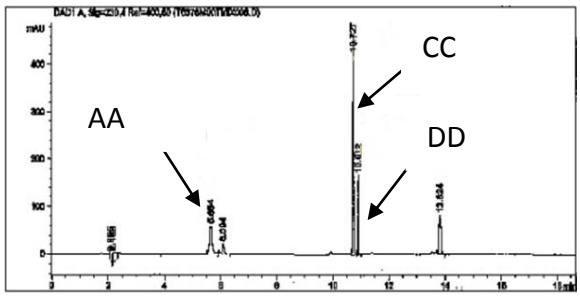

Figure 4.3: HPLC chromatogram showing the elution order of AA, CC and DD at 230 nm using a gradient method (acetonitrile and pH 3.0 water adjusted with phosphoric acid H_3PO_4). Column: Agilent Zorbax Eclipse C-18 (25 cm x 4.6 mm, 5um), temperature: 40 °C, injection volume: 10 uL, flow rate: 1.5 mL min^{-1}

Several different sampling techniques were tested, to ensure sufficient precision and reproducibility. In order to determine the minimum sample size that did not disturb the course of the reaction but was representative of the batch, two sampling techniques were compared. Sampling method (1), which only used a 20 µL slurry sample, was compared with sampling method (2), which used a 3.5 mL slurry sample. The reproducibility of triplicate aliquots was approximately the same for both sampling methods in terms of standard deviation, thus sampling method (1) was used in subsequent work.

4.4.4 Partial Least Squares Calibration

Before beginning the kinetic modeling of the system, it was needed to be confirmed that data would be consistent from batch to batch. Taking measurements of a slurry mixture can yield inconsistent results due to the inconsistencies in the sampling. Each measurement will have different levels of reflectance signal attributed to light absorption and light scattering. So, a partial least squares (PLS) calibration was performed to ensure that one batch could be used as a predictor for another. This process ensures that kinetic modeling is a viable option for modeling the system. A description of this process can be found in the literature, so only a brief summary of the process will be provided here [88-91].

To begin, a PLS regression was performed on individual data sets. Each of these was constructed by adding a known amount of reactant to a known amount of suspended triazine sample. Samples were taken from the reactor vessel at specified points in time and the concentrations at those intervals were calculated using HPLC data from each sample. The goal of this regression was to relate the measured spectra to the measured concentration information. The measured concentration data was used to construct weighing factors for the PLS regression, creating a concentration profile for the spectroscopic data. After this, a PLS calibration was performed. A PLS calibration involves the use of one data set as the predictor (the training set) for another (the test set). In this case, the weights applied to the first data set were applied to the test set, and the new concentration profile was then compared to the measured concentration data. The results from one such analysis are displayed in Figure 4.4. As can be seen in the plot, the agreement between the predicted concentration profile and the actual concentrations agree, demonstrating that the measurements are consistent enough that kinetic modeling is a viable approach for characterizing the spectroscopic data.

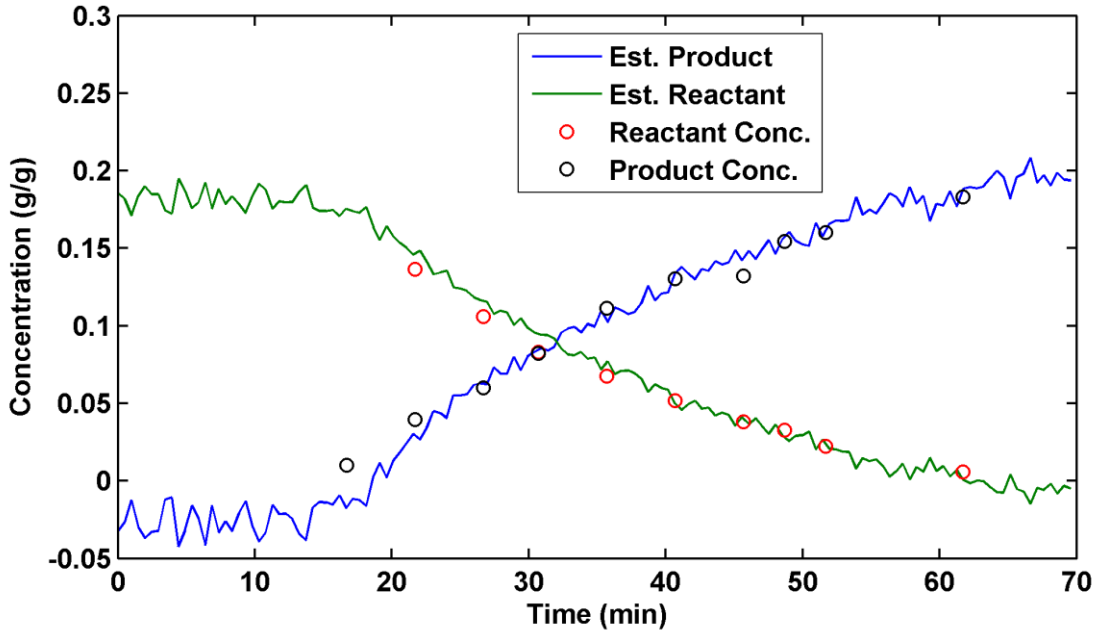

Figure 4.4: PLS calibration. Lines indicate the estimated concentrations of the reactant and product based on the calibration. Circles indicate the measured concentration determined by HPLC.

4.4.5 **Kinetic Models for Slurries and Numerical Integration**

In order to fit a kinetic model, it was necessary to construct a system of Ordinary Differential Equations (ODEs) that represent a given reaction mechanism. In this paper, the system of ODEs used is shown in Equations 4.17-4.21. Equations 4.13-4.14 describe the reactions of AA with BB producing the product CC in a batch reactor. Additional terms were required for reagent flow-in conditions used in a fed batch reactor.

$$A_s \xrightarrow{k_1} A \tag{4.13}$$

$$A + B \xrightarrow{k_2} P \tag{4.14}$$

Equations 4.13 represents the dissolution process where the starting reagent AA, a solid material went from solid form A_s into solution form A after it was added into xylene, k_1 is the dissolution rate constant. Equation 4.14 represents the sulfonylurea coupling reaction, where the dissolved form A reacts with reagent BB B to form product P, k_2 is the coupling reaction rate constant.

$$r_1 = k_1(C_A - C_{Asat})^d \tag{4.15}$$

$$r_2 = k_2 C_A C_B \tag{4.16}$$

From Equations 4.13-4.14, we constructed two rate laws (Equations 4.15-4.16) where r_1 and r_2 represent the rate of dissolution and the rate of the coupling reaction, respectively, k_1 and k_2 are the corresponding rate constants, C_A is the dissolved AA concentration in bulk solution, C_{Asat} is the solubility of AA, C_B is the BB concentration, and d is dissolution the rate order. The kinetic model was then translated into a system of Ordinary Differential Equations (ODEs), Equations 4.17-4.21, which were numerically integrated producing the resulting in the *nc* concentration profiles in **C**.

$$\frac{dV}{dt} = F \tag{4.17}$$

$$\frac{dC_{As}}{dt} = -r_1 - \frac{C_A}{V}F \tag{4.18}$$

$$\frac{dC_A}{dt} = r_1 - r_2 - \frac{C_A}{V}F \tag{4.19}$$

$$\frac{dC_B}{dt} = -r_2 + \frac{C_{Bin} - C_B}{V}F \tag{4.20}$$

$$\frac{dC_P}{dt} = r_2 - \frac{C_P}{V}F \tag{4.21}$$

In Equations 4.17-4.21, F is the flow-in rate (L/min) of BB solution, V is the total volume of the slurry (L), C_{As} is the undissolved AA concentration (mol/L), C_B is the BB concentration inside the reactor (mol/L), C_{Bin} is the BB concentration in the reagent reservoir (mol/L), and C_P is the product CC concentration (mol/L).

4.5 Results and Discussions

At the beginning of the reaction, the reactor was charged with 750 mL xylene and 103 g of solid triazine heterocycle (AA), which did not completely dissolve. During the reaction, A4098 solid materials were suspended in xylenes and was heated at 85 °C and agitated while a concentrated isocyanate solution (BB) was added via peristaltic pump to the reactor. At the end of an experiment, the mass of product CC in the slurry was about 25-30% solid by weight (g/g) and the slurry had the consistency of a thick milk-shake.

Figure 4.5 shows the entire record of an experimental run. Figure 4.5-A shows the NIR reflectance data (log 1/R) for the sulfonylurea coupling reaction as a function of wavelength from 1100 nm to 2200 nm region. Figure 4.5-C shows the reflectance data treated with basic baseline correction only, Figure 4.5-D shows the reflectance data treated with baseline correction and Extended Multiplicative Scatter Correction (EMSC) (discussed in more detail later).

The disappearance with time of the N-H overtone band at 2010 nm is shown in Figure 4.6 and shaded bars show the times during which the peristaltic pump was turned on to deliver the BB reagent solution. The NIR reflectance signal at 2010 nm increased when the first addition of BB was introduced (until an elapsed time of 20 min). We suppose that reactant AA, which was initially solid, dissolved and was subsequently consumed by BB. At the same time the product, CC, formed a supersaturated solution. After time 20 min, the concentration of product CC exceeded the metastable supersaturated limit and spontaneously crystallized as shown by the steep drop or relaxation of the signal. Further relaxations, however were less pronounced than the first one, and were observed for the next three additions of BB. A separate experiment was conducted to include an FBRM probe inside the reactor. FBRM measurement confirmed that the supersaturation and un-seeded nucleation and crystallization events occurred in the reactor during the course of an experiment.

In order to evaluate the reproducibility of our experimental protocol and the reproducibility of the model fitting processes, two additional replicate batches were performed but at a shorter time frame. This ensures the model was robust with respect to changes in the experimental time frame.

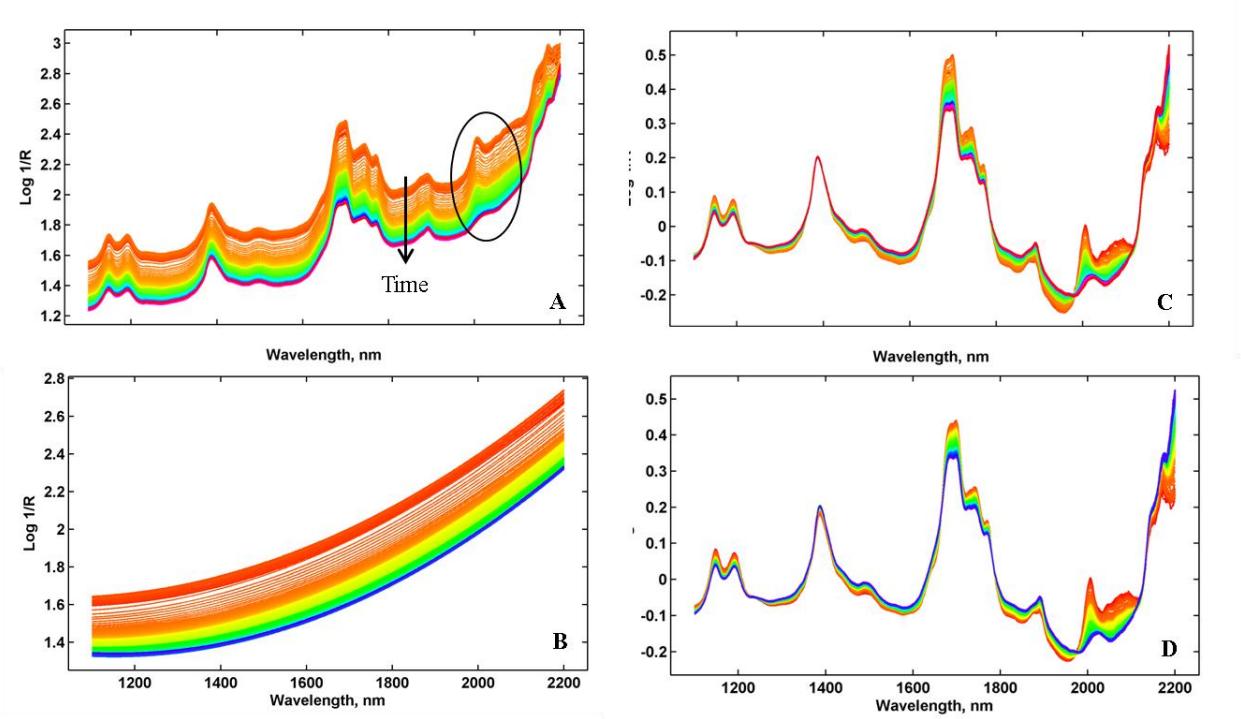

Figure 4.5: NIR reflectance data (log 1/R) as a function of wavelength from 1100 to 2200 nm region. Panel (A) shows the unaltered reflectance data, panel (B) shows the baseline offset, panel (C) shows the reflectance data applied with baseline correction, panel (D) shows the reflectance data applied with baseline correction and Extended Multiplicative Scatter Correction (EMSC).

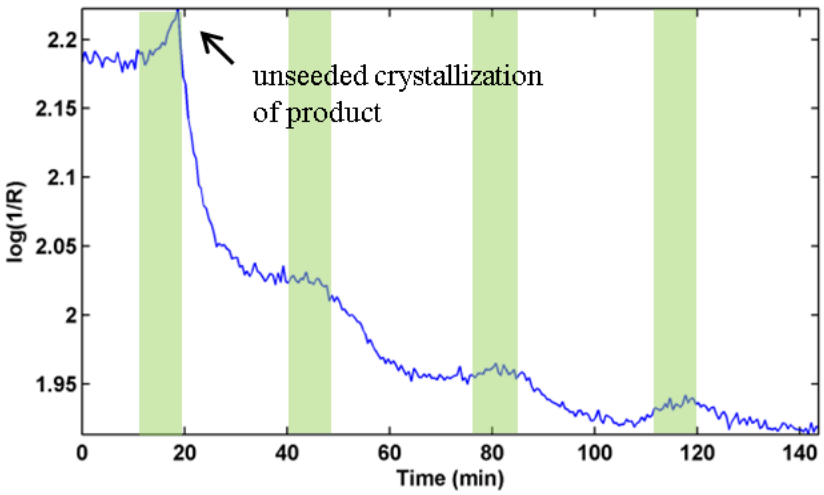

Figure 4.6: Time resolved NIR reflectance data at 2010 nm for the sulfonylurea coupling reaction involving four additions of BB.

Preprocessing treatments were applied to the NIR reflectance spectra using the Extended Multiplicative Scatter Correction (EMSC) technique [56] before fitting a kinetic model. Raw NIR spectra (log(1/R), Figure 4.5-A) were first treated with EMSC baseline correction (Figure 4.5-B), which models baseline offsets in near-infrared spectra as a combination of an offset, a function linear with wavelength and a function quadratic with wavelength. Figure 4.5-C was obtained by subtracting the baselines in Figure 4.5-B from the log(1/R) NIR reflectance spectra in Figure 4.5-A. The final step in the EMSC technique is normalization of each spectrum to the average spectrum in the data as shown in Figure 4.5-D. As xylene is the dominant signal in these spectra, the EMSC technique effectively adjusts the NIR spectra to a constant effective path length for xylene absorption.

Figure 4.7 shows the preprocessed near-infrared spectra used in the kinetic fitting process, wavelength region 1800 nm to 2100 nm was used. The left panel shows the measured near infrared spectra as a function of wavelength, and the right panel shows the model estimated spectra as a function of wavelength. Good agreement can be seen to exist between the measured and modeled spectra.

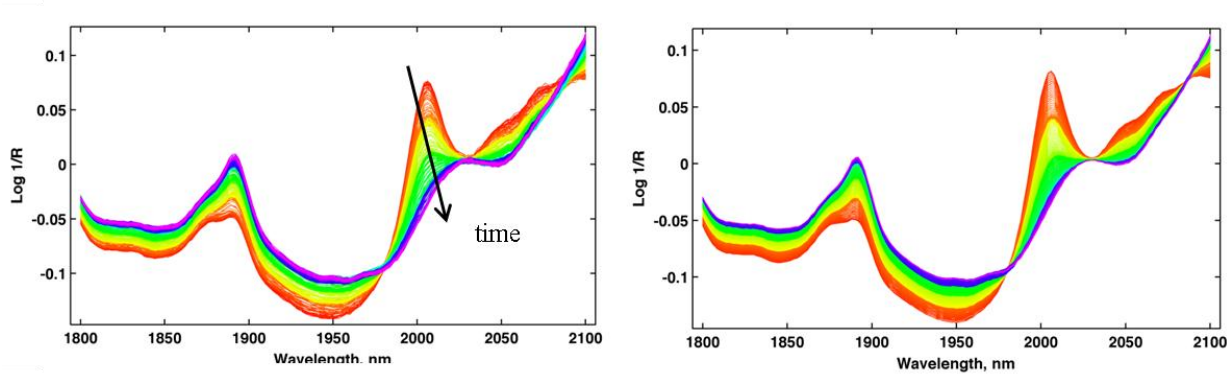

Figure 4.7: Comparison of the measured spectral data (left) with model estimated spectral data (right) as a function of wavelength to show the quality of fitting.

In these experiments, we assumed the slurries were well-mixed such that all particles of AA and CC were suspended relatively uniformly throughout the reactor vessel, inside the recirculation loop, and that the stirring rate was relatively constant. We also assumed the particle size distribution of the solid materials remains approximately the same throughout the experiment.

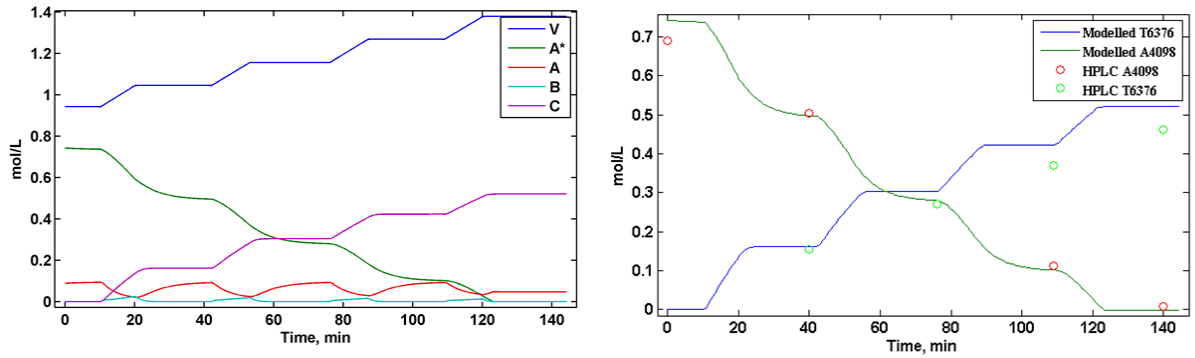

Figure 4.8: Comparison of the model estimated concentration profiles as a function of time.

In Figure 4.8, the left plot shows the model estimated concentration profiles. The blue continuous line shows the total volume (V), green, red continuous lines shows the solid (A*) and liquid (A) form of AA, respectively. The turquoise continuous line shows the flow-in profile for the BB reagent (B), the purple continuous line shows the CC concentration profile (C). When BB reagent (B) was first introduced, it gets consumed by AA in liquid phase (A) and form product CC (C) immediately. The right plot shows the comparison of model estimated concentration profiles with HPLC data as a function of time for validation purposes. The green curve shows the model estimated AA concentration profile, the blue curve shows the model estimated CC concentration profile, and the red and green circles show the AA concentration and the CC concentration profiles measured by off-line HPLC, respectively.

Good agreement was observed between the off-line HPLC results (red circles) and the model estimated concentration of starting material (green curve) as a function of time; however, significant discrepancies were observed for the product. The source of error was most likely from the HPLC validation work, because we know the stoichiometry of the reaction and the initial concentrations of the reagents precisely. When we were running the protocol, in certain batches we were running low of reagent BB, and the BB reagent reservoir on the balance was tilted slightly. Therefore, the amount of mass delivered in the last addition may not be very accurate, and thus we used the estimated amount for the last addition. We know the calibrated pump was pumping at a constant speed and how much reagent pump delivered in the first three additions. Consequently, we took the average of first three additions and assumed that the pump continued to deliver the same amount for the last addition because it never actually ran dry.

We have additional data from May 2010 that has not been analyzed yet. We have PLS work that indicates some of the HPLC validation data from May 2010 is useful (see Figure 4.4) and as a follow up, kinetic model needs to be estimated for the May batch data in the hopes we can gain and have useful validation data that could lead to a journal article publication.

Table 4.1: Kinetic model fitting results fitted with NIR reflectance data.

Parameters	**Batch A**		**Batch B**		**Batch C**		**Average**	**STD**	**RSD**
	Input	*Optimized*	*Input*	*Optimized*	*Input*	*Optimized*			
K_1	0.2000	0.5802	0.2000	0.1604	0.2000	0.2341	0.3249	0.2241	68.99
K_2	6.000	6.1261	6.000	26.55	6.000	64.94	32.54	29.86	91.78
A_{sat}	0.08800	0.06380	0.08800	0.09490	0.08800	0.07020	0.07630	0.01642	21.52
SSQ		0.1278		0.3310		0.3011	0.2533	0.1097	43.31

Table 4.1 shows the comparison of the kinetic model fitting results, obtained by fitted NIR reflectance measurements. For the purpose of this empirical model fitting process, reasonably satisfactory fitting results were obtained. The purpose of the model fitting process is

not to determine fundamental constants in properties of dissolution or chemical reaction rates. Instead, the purpose is to develop an empirical model that gives a statistically sufficient phenomenological description of the batch process for monitoring and control.

4.6 Conclusion

In summary, we have developed a novel kinetic modeling strategy to monitor commercially relevant slurry reactions (i.e. dissolution and coupling reaction). NIR diffuse reflectance spectral data were measured using a fiber-optic probe inside of a jacketed reactor vessel. FBRM measurements confirmed the observation of the product precipitation event in the NIR diffuse reflectance spectra. In order to determine the kinetic parameters involved during the dissolution and coupling reaction processes, a low-theory, simplified comprehensive global model describing these processes as a function of time was solved using MATLAB. The model was coupled with a MATLAB nonlinear optimization subroutine. The parameters estimated in the model included dissolution and coupling reaction rate constants, and solubility of A4098 in xylene at 85 °C. At the present time, the HPLC validation does not support in submission of publication of a journal article. Future work includes analyze batch data from May 2010 that has not been analyzed yet, and expanding the existing kinetic model by incorporating nucleation and crystallization processes. This will help improve the quality of fitting results.

4.7 Acknowledgement

All authors are grateful to the financial support from the National Science Foundation (NSF) under Grant Number CHE-0750287, Grant Opportunities for Academic Liaison with Industry (GOALI) and the E.I. DuPont de Nemours and Co., Inc., Crop Protection Products and Engineering Technologies. Chun Hsieh thanks Dr. Rich Davis, Dr. David Cho, David Bailey,

Steve Platz, and Kevin Day and for their assistance for the experimental set up and technical supports at Stine-Haskell Research Center.

4.8 Notations

C_A dissolved AA concentration (mol/L)

C_{As} undissolved AA concentration (mol/L)

C_{Asat} solubility of AA (mol/L)

C_B BB concentration in bulk solution (mol/L)

C_{Bin} BB concentration in reagent reservoir (mol/L)

C_P CC concentration in bulk solution (mol/L)

d dissolution rate order

F flow-in profile (L/min)

k_1 dissolution rate constant $L^{n-1}/(mol^{n-1}min)$

k_2 coupling reaction rate constant $L^{n-1}/(mol^{n-1}min)$

r_1 dissolution rate (mol/L/ut)

r_2 coupling reaction rate (mol/L/ut)

V total volume (L)

5. CONCLUSION

In this thesis, we have demonstrated an advanced measurement and modeling strategy to monitor (1) the dissolution and crystallization of salicylic acid in a mixed solvent, and (2) the reaction and dissolution sulfonylurea coupling reaction in xylene. ATR UV-Vis and NIR diffuse reflectance spectroscopy can be used as tools for the monitoring of the liquid and solid phases. The solid phase was measured using in situ diffuse reflectance NIR spectroscopy, while the liquid phase was measured using in situ ATR UV/Visible spectroscopy, respectively.

In order to determine the kinetic parameters involved during the chemical reaction, crystal growth and dissolution processes, a low-theory, simplified comprehensive global model describing these processes as a function of time was solved using MATLAB. The model was coupled with a MATLAB nonlinear optimization subroutine. The suggested kinetic laws were sufficient in describing the set of seven crystallization and dissolution experiments performed during the study (four from salicylic acid experiments, three from sulfonylurea experiments). The parameters estimated in the model included dissolution and crystallization rate constants, dissolution rate exponent, as well as the degree of supersaturation which provides the driving force for crystallization.

We have evaluated the reproducibility of our experimental protocol and the reproducibility of the modeling fitting process by shortening the experimental time by half. This ensures the model was robust with respect to changes in the experimental time frame. In addition, the methods required to carry out the fitting of a kinetic model to measured multivariate spectroscopic data of slurries (i.e. NIR and UV-vis spectra) have been demonstrated, from the

postulation of the model and the derivation of the differential equations, through the numerical integration of the model to yield the concentration profiles and finally the calculation of the pure component spectra and fitting the models' rate constants to measured data.

References

1. Brown, S.D., et al., *Chemometrics.* Analytical Chemistry, 1996. **68**(12): p. R21-R61.
2. Delaney, M., *THE INTERPRETATION OF ANALYTICAL CHEMICAL-DATA BY THE USE OF CLUSTER-ANALYSIS - MASSART,DL, KAUFMAN,L.* Analytical Chemistry, 1984. **56**(4): p. A613-A614.
3. Hamilton J.C., G.P.J., *Mixture Analysis Using Factor Analysis. II: Self-Modeling Curve Resolution.* Journal of Chemometrics, 1990. **4**: p. 1-13.
4. Gemperline, P.J. and E. Cash, *Advantages of soft versus hard constraints in self-modeling curve resolution problems. Alternating least squares with penalty functions.* Analytical Chemistry, 2003. **75**(16): p. 4236-4243.
5. De Braekeleer, K., et al., *On-line application of the orthogonal projection approach (OPA) and the soft independent modelling of class analogy approach (SIMCA) for the detection of the end point of a polymorph conversion reaction by near infrared spectroscopy (NIR).* Chemometrics and Intelligent Laboratory Systems, 1999. **46**(2): p. 103-116.
6. Ma, B., et al., *Characterizing batch reactions with in situ spectroscopic measurements, calorimetry and dynamic modeling.* Journal of Chemometrics, 2003. **17**(8-9): p. 470-479.
7. Tauler, R., S. Lacorte, and D. Barcelo, *Application of multivariate self-modeling curve resolution to the quantitation of trace levels of organophosphorus pesticides in natural waters from interlaboratory studies.* Journal of Chromatography A, 1996. **730**(1-2): p. 177-183.

8. Cutler, P.J., et al., *Methods for Kinetic Modeling of Temporally Resolved Hyperspectral Confocal Fluorescence Images.* Applied Spectroscopy, 2009. **63**(2): p. 153-163.

9. Cutler, P.J., D.M. Haaland, and P.J. Gemperline, *Systematic Method for the Kinetic Modeling of Temporally Resolved Hyperspectral Microscope Images of Fluorescently Labeled Cells.* Applied Spectroscopy, 2009. **63**(3): p. 261-270.

10. Garrido, M., F.X. Rius, and M.S. Larrechi, *Multivariate curve resolution-alternating least squares (MCR-ALS) applied to spectroscopic data from monitoring chemical reactions processes.* Analytical and Bioanalytical Chemistry, 2008. **390**(8): p. 2059-2066.

11. Billeter, J., *Chemometric Methods for Prediction of Uncertainties and Spectral Validation of Rank Deficient Mechanisms in Kinetic Hard-Modelling of Spectroscopic Data*, 2009, Swiss Federal Institute of Technology Zurich: Switzerland.

12. de Juan, A., et al., *Combining hard- and soft-modelling to solve kinetic problems.* Chemometrics and Intelligent Laboratory Systems, 2000. **54**(2): p. 123-141.

13. De Braekeleer, K., et al., *Determination of the end point of a chemical synthesis process using on-line measured mid-infrared spectra.* Applied Spectroscopy, 2000. **54**(4): p. 601-607.

14. Furusjo, E., O. Svensson, and L.G. Danielsson, *Estimation of kinetic parameters from non-isothermal batch experiments monitored by in situ vibrational spectroscopy.* Chemometrics and Intelligent Laboratory Systems, 2003. **66**(1): p. 1-14.

15. Norman, S. and M. Maeder, *Model-based analysis for kinetic and equilibrium investigations.* Critical Reviews in Analytical Chemistry, 2006. **36**(3-4): p. 199-209.

16. Molloy, K.J., M. Maeder, and M.M. Schumacher, *Hard modelling of spectroscopic measurements. Applications in non-ideal industrial reaction systems.* Chemometrics and Intelligent Laboratory Systems, 1999. **46**(2): p. 221-230.

17. Bayada, A., et al., *ATR-IR spectroscopy for the investigation of solution reaction kinetics: Hydrolysis of trimethyl phosphate.* Applied Spectroscopy, 1995. **49**(12): p. 1789-1792.

18. Dyson, R.M., et al., *Modern tools for reaction monitoring: hard and soft modelling of 'non-ideal', on-line acquired spectra.* Journal of Chemometrics, 2000. **14**(5-6): p. 737-750.

19. Bijlsma, S. and A.K. Smilde, *Estimating reaction rate constants from a two-step reaction: a comparison between two-way and three-way methods.* Journal of Chemometrics, 2000. **14**(5-6): p. 541-560.

20. Bijlsma, S., et al., *Estimating reaction rate constants: comparison between traditional curve fitting and curve resolution.* Analytica Chimica Acta, 2000. **419**(2): p. 197-207.

21. Windig, W. and D.A. Stephenson, *SELF-MODELING MIXTURE ANALYSIS OF 2ND-DERIVATIVE NEAR-INFRARED SPECTRAL DATA USING THE SIMPLISMA APPROACH.* Analytical Chemistry, 1992. **64**(22): p. 2735-2742.

22. Windig, W., et al., *SELF-MODELING MIXTURE ANALYSIS OF CATEGORIZED PYROLYSIS MASS-SPECTRAL DATA WITH THE SIMPLISMA APPROACH.* Chemometrics and Intelligent Laboratory Systems, 1992. **14**(1-3): p. 195-207.

23. Windig, W., *SELF-MODELING MIXTURE ANALYSIS OF SPECTRAL DATA WITH CONTINUOUS CONCENTRATION PROFILES.* Chemometrics and Intelligent Laboratory Systems, 1992. **16**(1): p. 1-16.

24. Tauler, R., B. Kowalski, and S. Fleming, *MULTIVARIATE CURVE RESOLUTION APPLIED TO SPECTRAL DATA FROM MULTIPLE RUNS OF AN INDUSTRIAL-PROCESS.* Analytical Chemistry, 1993. **65**(15): p. 2040-2047.

25. Tauler, R. and D. Barcelo, *MULTIVARIATE CURVE RESOLUTION APPLIED TO LIQUID-CHROMATOGRAPHY DIODE-ARRAY DETECTION.* Trac-Trends in Analytical Chemistry, 1993. **12**(8): p. 319-327.

26. Tauler, R., *Multivariate curve resolution applied to second order data.* Chemometrics and Intelligent Laboratory Systems, 1995. **30**(1): p. 133-146.

27. Diazcruz, J.M., et al., *APPLICATION OF MULTIVARIATE CURVE RESOLUTION TO VOLTAMMETRIC DATA .1. STUDY OF ZN(II) COMPLEXATION WITH SOME POLYELECTROLYTES.* Journal of Electroanalytical Chemistry, 1995. **393**(1-2): p. 7-16.

28. Sanchez, F.C., et al., *Orthogonal projection approach applied to peak purity assessment.* Analytical Chemistry, 1996. **68**(1): p. 79-85.

29. Gourvenec, S. and D.L. Massart, *Orthogonal projection approach (OPA) and related methods in process monitoring.* Analytical and Bioanalytical Chemistry, 2004. **380**(3): p. 373-375.

30. de Juan, A. and R. Tauler, *Chemometrics applied to unravel multicomponent processes and mixtures - Revisiting latest trends in multivariate resolution.* Analytica Chimica Acta, 2003. **500**(1-2): p. 195-210.

31. Kessler, W. and R.W. Kessler, *Multivariate curve resolution: a method of evaluating the kinetics of biotechnological reactions.* Analytical and Bioanalytical Chemistry, 2006. **384**(5): p. 1087-1095.

32. Navea, S., A. de Juan, and R. Tauler, *Modeling temperature-dependent protein structural transitions by combined near-IR and mid-IR spectroscopies and multivariate curve resolution.* Analytical Chemistry, 2003. **75**(20): p. 5592-5601.

33. Blanco, M., et al., *Application of multivariate curve resolution to chemical process control of an esterification reaction monitored by near-infrared spectroscopy.* Applied Spectroscopy, 2006. **60**(6): p. 641-647.

34. Ozaki, Y., et al., *Self-modeling curve resolution analysis of on-line vibrational spectra of polymerisation and transesterification.* Macromolecular Symposia, 2002. **184**: p. 229-247.

35. Sasic, S., et al., *Comparison of various chemometric evaluation approaches for on-line FT-NIR transmission and FT-MIR/ATR spectroscopic data of methyl methacrylate solution polymerization.* Analytica Chimica Acta, 2002. **452**(2): p. 265-276.

36. Blanco, M., A.C. Peinado, and J. Mas, *Elucidating the composition profiles of alcoholic fermentations by use of ALS methodology.* Analytica Chimica Acta, 2005. **544**(1-2): p. 199-205.

37. Abebe, S.B., et al., *The information content in NIR spectral data for slurries of organic crystals.* Powder Technology, 2008. **179**(3): p. 176-183.

38. Fevotte, G., *New perspectives for the on-line monitoring of pharmaceutical crystallization processes using in situ infrared spectroscopy.* International Journal of Pharmaceutics, 2002. **241**(2): p. 263-278.

39. Maeder, M. and A.D. Zuberbuhler, *NONLINEAR LEAST-SQUARES FITTING OF MULTIVARIATE ABSORPTION DATA.* Analytical Chemistry, 1990. **62**(20): p. 2220-2224.

40. Aziz, N. and I.M. Mujtaba, *Optimal operation policies in batch reactors.* Chemical Engineering Journal, 2002. **85**(2-3): p. 313-325.

41. Puxty, G., et al., *Modeling of batch reactions with in situ spectroscopic measurements and calorimetry.* Journal of Chemometrics, 2005. **19**(5-7): p. 329-340.

42. Lee, B., *Review of the present status of optical fiber sensors.* Optical Fiber Technology, 2003. **9**(2): p. 57-79.

43. Seippel, R.G., *Fiber Optics*. 1984: Reston Publishing Co. Inc. .

44. Peterson, J.I. and G.G. Vurek, *FIBER-OPTIC SENSORS FOR BIOMEDICAL APPLICATIONS.* Science, 1984. **224**(4645): p. 123-127.

45. Utzinger, U. and R.R. Richards-Kortum, *Fiber optic probes for biomedical optical spectroscopy.* Journal of Biomedical Optics, 2003. **8**(1): p. 121-147.

46. Anderson, J.E., et al., *Determination of the onset of crystallization of N-1-2-(thiazolyl)sulfanilamide (sulfathiazole) by UV-Vis and calorimetry using an automated reaction platform; subsequent characterization of polymorphic forms using dispersive Raman spectroscopy.* Spectrochimica Acta Part a-Molecular and Biomolecular Spectroscopy, 2001. **57**(9): p. 1793-1808.

47. Billot, P., M. Couty, and P. Hosek, *Application of ATR-UV Spectroscopy for Monitoring the Crystallisation of UV Absorbing and Nonabsorbing Molecules.* Organic Process Research & Development. **14**(3): p. 511-523.

48. Billot, P., M. Couty, and P. Hosek, *Application of ATR-UV Spectroscopy for Monitoring the Crystallisation of UV Absorbing and Nonabsorbing Molecules.* Organic Process Research & Development, 2010. **14**(3): p. 511-523.

49. *ATR Theory and Applications (Application Note - 0402),* PIKE TECHNOLOGIES Spectroscopic Creativity.

50. Rioboo, R.J.J., et al., *Concentration and temperature dependence of the refractive index of ethanol-water mixtures: Influence of intermolecular interactions.* European Physical Journal E, 2009. **30**(1): p. 19-26.

51. Roggo, Y., et al., *A review of near infrared spectroscopy and chemometrics in pharmaceutical technologies.* Journal of Pharmaceutical and Biomedical Analysis, 2007. **44**(3): p. 683-700.

52. Villringer, A. and B. Chance, *Non-invasive optical spectroscopy and imaging of human brain function.* Trends in Neurosciences, 1997. **20**(10): p. 435-442.

53. Moros, J., et al., *Univariate near infrared methods for determination of pesticides in agrochemicals.* Analytica Chimica Acta, 2006. **579**(1): p. 17-24.

54. Vargas, W.E. and G.A. Niklasson, *Applicability conditions of the Kubelka-Munk theory.* Applied Optics, 1997. **36**(22): p. 5580-5586.

55. Geladi, P., D. Macdougall, and H. Martens, *LINEARIZATION AND SCATTER-CORRECTION FOR NEAR-INFRARED REFLECTANCE SPECTRA OF MEAT.* Applied Spectroscopy, 1985. **39**(3): p. 491-500.

56. Martens, H., J.P. Nielsen, and S.B. Engelsen, *Light scattering and light absorbance separated by extended multiplicative signal correction. Application to near-infrared transmission analysis of powder mixtures.* Analytical Chemistry, 2003. **75**(3): p. 394-404.

57. Burns, D.A.M., M., *Handbook of Near-Infrared Analysis.* 1992, New York: Marcel Dekker, Inc.

58. Hildrum, K.I.I., T.; Naes, T.; Tandberg A., *Near-Infrared Spectroscopy. Bridging the Gap between Data Analysis and NIR Applications*. 1992, New York.

59. Ehly, M., et al., *Scale-up of batch kinetic models.* Analytica Chimica Acta, 2007. **595**(1-2): p. 80-88.

60. Weyer, L.G.L., S. C., *Handbook of Vibrational Spectroscopy*. 2002, New York: John Wiley & Sons, Inc.

61. Sjostrom, M., et al., *A MULTIVARIATE CALIBRATION-PROBLEM IN ANALYTICAL-CHEMISTRY SOLVED BY PARTIAL LEAST-SQUARES MODELS IN LATENT-VARIABLES.* Analytica Chimica Acta, 1983. **150**(1): p. 61-70.

62. Rinnan, A., F. van den Berg, and S.B. Engelsen, *Review of the most common pre-processing techniques for near-infrared spectra.* Trac-Trends in Analytical Chemistry, 2009. **28**(10): p. 1201-1222.

63. E.B. Martin, A.J.M., C. Kiparissides, *Manufacturing performance enhancement through multivariate statistical process control.* Monitoring and control of industrial system, 1999. **23**: p. 35-44.

64. Shao, X.G., et al., *Multivariate calibration methods in near infrared spectroscopic analysis.* Analytical Methods, 2010. **2**(11): p. 1662-1666.

65. Lorber A., W.L.E., Kowalski, B.R., *A THEORETICAL FOUNDATION FOR THE PLS ALGORITHM.* Journal of Chemometrics, 1987. **1**: p. 19-31.

66. Blanco, M., et al., *Strategies for constructing the calibration set in the determination of active principles in pharmaceuticals by near infrared diffuse reflectance spectrometry.* Analyst, 1997. **122**(8): p. 761-765.

67. Berntsson, O., et al., *Quantitative determination of content in binary powder mixtures using diffuse reflectance near infrared spectrometry and multivariate analysis.* Analytica Chimica Acta, 2000. **419**(1): p. 45-54.

68. Garside, J., *INDUSTRIAL CRYSTALLIZATION FROM SOLUTION.* Chemical Engineering Science, 1985. **40**(1): p. 3-26.

69. Tavare, N.S., *CHARACTERIZATION OF CRYSTALLIZATION KINETICS FROM BATCH EXPERIMENTS.* Separation and Purification Methods, 1993. **22**(2): p. 93-210.

70. Li, R.F., X.Z. Wang, and S.B. Abebe, *Monitoring Batch Cooling Crystallization Using NIR: Development of Calibration Models Using Genetic Algorithm and PLS.* Particle & Particle Systems Characterization, 2008. **25**(4): p. 314-327.

71. Skoog D., H.F., Crouch S., *Principles of Instrumental Analysis.* Thomson Higher Education, 2007.

72. Heath, A.R., et al., *Estimating average particle size by focused beam reflectance measurement (FBRM).* Particle & Particle Systems Characterization, 2002. **19**(2): p. 84-95.

73. Richmond, W.R., R.L. Jones, and P.D. Fawell, *The relationship between particle aggregation and rheology in mixed silica-titania suspensions.* Chemical Engineering Journal, 1998. **71**(1): p. 67-75.

74. Law, D.J., A.J. Bale, and S.E. Jones, *Adaptation of focused beam reflectance measurement to in-situ particle sizing in estuaries and coastal waters.* Marine Geology, 1997. **140**(1-2): p. 47-59.

75. Sparks, R.G. and C.L. Dobbs, *THE USE OF LASER BACKSCATTER INSTRUMENTATION FOR THE ONLINE MEASUREMENT OF THE PARTICLE-SIZE*

DISTRIBUTION OF EMULSIONS. Particle & Particle Systems Characterization, 1993. **10**(5): p. 279-289.

76. Sparks, R.G. and C.L. Dobbs, *THE USE OF LASER BACKSCATTER INSTRUMENTATION FOR THE ONLINE MEASUREMENT OF THE PARTICLE-SIZE DISTRIBUTION OF EMULSIONS (VOL 10, PG 279, 1993).* Particle & Particle Systems Characterization, 1994. **11**(5): p. 404-404.

77. Monnier, O., et al., *Model identification of batch cooling crystallizations through calorimetry and image analysis.* Chemical Engineering Science, 1997. **52**(7): p. 1125-1139.

78. Monnier, O., et al., *Particle size determination by laser reflection: Methodology and problems.* Particle & Particle Systems Characterization, 1996. **13**(1): p. 10-17.

79. Puxty, G., M. Maeder, and K. Hungerbuhler, *Tutorial on the fitting of kinetics models to multivariate spectroscopic measurements with non-linear least-squares regression.* Chemometrics and Intelligent Laboratory Systems, 2006. **81**(2): p. 149-164.

80. Puxty, G., et al., *Multivariate kinetic hard-modelling of spectroscopic data: A comparison of the esterification of butanol by acetic anhydride on different scales and with different instruments.* Chemical Engineering Science, 2008. **63**(19): p. 4800-4809.

81. Cornel, J. and M. Mazzotti, *Estimating Crystal Growth Rates Using in situ ATR-FTIR and Raman Spectroscopy in a Calibration-Free Manner.* Industrial & Engineering Chemistry Research, 2009. **48**(23): p. 10740-10745.

82. Caillet, A., N. Sheibat-Othman, and G. Fevotte, *Crystallization of monohydrate citric acid. 2. Modeling through population balance equations.* Crystal Growth & Design, 2007. **7**(10): p. 2088-2095.

83. Westerhuis, J.A., S.P. Gurden, and A.K. Smilde, *Spectroscopic monitoring of batch reactions for on-line fault detection and diagnosis.* Analytical Chemistry, 2000. **72**(21): p. 5322-5330.

84. William H. Press, S.A.T., William T. Vetterling, Brian P. Flannery, *Numerical Recipes in C: The Art of Scientific Computing*. 1992: Cambridge University Press. 1035.

85. Levenberg, K., *a method for the solution of certain problems in least squares.* Quart. Applied Math, 1944. **2**: p. 164-168.

86. Marquardt, D.W., *AN ALGORITHM FOR LEAST-SQUARES ESTIMATION OF NONLINEAR PARAMETERS.* Journal of the Society for Industrial and Applied Mathematics, 1963. **11**(2): p. 431-441.

87. DuPont HPLC methods Y6266.200.01.BE, T.E., T6376.220.05.ES, 2000, 1999, 2004.

88. Ni, W.D., S.D. Brown, and R.L. Man, *Stacked partial least squares regression analysis for spectral calibration and prediction.* Journal of Chemometrics, 2009. **23**(9-10): p. 505-517.

89. Gemperline, P., *Practical Guide to Chemometrics*. 2006: CRC/Taylor & Francis.

90. Geladi, P. and B.R. Kowalski, *PARTIAL LEAST-SQUARES REGRESSION - A TUTORIAL.* Analytica Chimica Acta, 1986. **185**: p. 1-17.

91. Xu, Q.S., Y.Z. Liang, and H.L. Shen, *Generalized PLS regression.* Journal of Chemometrics, 2001. **15**(3): p. 135-148.