Kati Pöllänen

MONITORING OF CRYSTALLIZATION PROCESSES BY USING INFRARED SPECTROSCOPY AND MULTIVARIATE METHODS

Thesis for the degree of Doctor of Science (Technology) to be presented with due permission for public examination and criticism in the Auditorium 1382 at Lappeenranta University of Technology, Lappeenranta, Finland on the 1st of September, 2006, at noon.

Acta Universitatis
Lappeenrantaensis
243

Supervisor Professor Pentti Minkkinen
Department of Chemical Technology
Lappeenranta University of Technology (LUT), Finland

Reviewers Professor John F. MacGregor,
Dofasco Professor of Process Automation and IT
Dept.of Chemical Engineering McMaster University Hamilton, Ontario, Canada

Ph.D. Veli-Matti Taavitsainen,
Institute of Technology
EVTEK University of applied sciences, Vantaa, Finland

Opponent Prof. Torbjörn Lundstedt
Department of Medicinal Chemistry
Uppsala biomedicinska centrum (BMC), Uppsala University, Sweden

ISBN 952-214-234-4
ISBN 952-214-235-2 (PDF)
ISSN 1456-4491

Lappeenrannan teknillinen yliopisto
Digipaino 2006

Abstract

Pöllänen Kati

Monitoring of crystallization processes by using infrared spectroscopy and multivariate methods

Lappeenranta 2006

118 p.

Acta Universitatis Lappeenrantaensis 243

Diss. Lappeenranta University of Technology

ISBN 952-214-234-4, ISBN 952-214-235-2 (PDF), ISSN 1456-4491

Batch cooling crystallization is an important purification unit operation in the pharmaceutical industry. The quality and further usability of the crystalline product depend on the size, shape, size distribution and purity such as the polymorphic form. Product properties can be controlled by controlling the cooling and mixing conditions and the way in how the first crystals are introduced into the crystallization system. To succesfully control the process, the phenomena inside the crystallizer leading to certain product properties should be understood. Therefore, crystallization process should be monitored in real time and the properties of crystalline product should be reliably characterized. Aiming to that the U.S. Food and Drug Administration (FDA) (2004) has addressed an initiative of process analytical technology (PAT), which states that new efficient process monitoring and product quality evaluation tools should be developed to maintain or improve the current level of product quality assurance. Modern analyzers in monitoring as well as multivariate methods to treat the data are recommended to be applied in development of PAT tools.

In this thesis, the application of in-situ attenuated total reflection Fourier transform infrared (ATR-FTIR) spectroscopy and multivariate data analysis methods are applied to study different phenomena during the crystallization process. Partial least scuares (PLS) modeling is applied for prediction of the solute concentration and further the driving force, supersaturation, in the crystallization process. The concentration measurement results are correlated to the quality of the obtained product. Multivariate statistical process control (MSPC) tools are used to monitor the on-set, i.e., primary nucleation of the crystallization process, which provides a new tool for alarming approaching nucleation and and exploring the chemical state prior to nucleation. Parallel factor analysis (PARAFAC) is tested for studying batch-to-batch variations in the crystallization processes. By monitoring crystallization processes using spectroscopic techniques and applying different multivariate methods a wide range of different kinds of information on the phenomena inside the crystallizer can be investigated.

In this thesis the methods to characterize the polymorphic composition using diffuse reflectance Fourier transform infra red (DRIFT-IR) spectroscopy together with multivariate data analysis tools such as PLS, principal component analysis (PCA) and soft independent modeling of class analogy (SIMCA) are tested. The DRIFT-IR method is combined with multivariate methods to provide a tool to rapidly estimate the polymorphic composition of the crystalline product.

Keywords: attenuated total reflection, diffuse reflectance, infra red spectroscopy, batch cooling crystallization, sulfathiazole, process analytical technology, multivariate methods, chemometrics, PLS, MSPC, PCA, SIMCA, PARAFAC, OSC

UDC 543.42 : 665.662.5 : 519.237

ACKNOWLEDGEMENTS

This study has been carried out at Lappeenranta University of Technology in the Laboratory of Inorganic and Analytical Chemistry and Laboratory of the Separation Technology.

I thank my supervisor Professor Pentti Minkkinen for his advice. I would like to thank Professor Emer. Lars Nyström for his guidance. Docent D.Sc. (Tech.) Marjatta Louhi-Kultanen deserves warm thanks for providing me tools for understanding crystallization process systems. I am indebted to Docent D.Sc. (Tech.) Satu-Pia Reinikainen for her invaluable help and guidance with multivariate methods.

I express my appreciation to Ph.D. Samantha Platt for proof-reading this thesis and helping me to improve English language.

I am thankful for the reviewers of this thesis Professor John F. MacGregor and Ph.D. Veli-Matti Taavitsainen for their valuable comments and constructive criticism, which significantly helped me to improve the thesis.

Most of this work has been performed within the Medipros project of Finnish Funding Agency for Technology and Innovation. Within the project, I am grateful to Professor Juha Kallas for all guidance, kind help and patience in all occasions. I am deeply indebted to the researchers in Medipros, without them this stony road would have turned into a rocky mountain. I will always owe a debt of graditude to M.Sc. Antti Häkkinen for the huge amount of data he produced for me to explore and model. This dissertation would not exist without that data. I am also thankful that we have always been able to pull together in during the dissertation research. I thank M.Sc. Mikko Huhtanen for our worthy scientific discussions and for always providing the other point of view. In addition, Mikko's extraordinary sense of humor has made many of my days.

I express my warm graditude to the Medipros partners in the University of Helsinki. Especially, I am thankful to Professor Jukka Rantanen for his guidance in spectroscopic methods, his valuable ideas during this work and his optimistic way of thinking when obstacles were encountered. I thank D.Sc. Milja Karjalainen for XRPD measurement results.

I would also like to acknowldedge our industrial partners within the Medipros project for their interest in this topic and valuable comments.

I am thankful to M.Sc. Haiyan Qu for our co-operation in C15 measurements. I thank Ms. Päivi Hovila for performing DRIFT-IR measurements on crystalline samples. Laboratory Manager Markku Maijanen is acknowledged for his help with experimental set-ups.

I thank all co-workers in Department of Chemical Technology for intriguing working atmosphere.

My warmest graditude goes to my friend and colleague M.Sc. Mari Kallioinen, whose support has had an unestimable meaning when undergoing the ups and downs of this project. I have always been able to count on Mari. Our conversations on the topic and off-topic have prevented many disasters in the work and in personal life.

In this context I also want to acknowledge our entertainment committee: Mari, Mikko, M.Sc. Eero Kaipainen and M.Sc. Pekka Olin, for such a fruitful co-operation in the Christmas party productions; those were the most relaxing tasks I've had in LUT.

Finnish Funding Agency for Technology and Innovation (TEKES), Graduate School of Chemical Engineering (GSCE), Finnish Academy, Orion corporation Fermion, Kemira Oyj., Kemira foundation, Gust. Komppa foundation, Danisco foundation and Foundation for Technology Promotion (Tekniikan edistämissäätiö (TES)) are acknowledged for funding and financial support.

I thank my parents Paula and Alpo for providing me with the good tools of life. My brother Antti deserves sincere thanks for not having a technical view of life, which is relaxing change for me, being normally surrounded by engineers.

Last but not least I express my deepest graditude to my family. My husband Riku kicked my butt every time I was going to give up, and he has always encouraged me to follow my intuitions in scientific issues. Our children Iida and Eetu by being just themselves have prevented life from becoming too serious.

Lappeenranta, June 18, 2006

Kati Pöllänen

Table of contents

List of publications

This thesis is based on the following papers, which are referred to in the text by the Roman numerals I-VI

I) Pöllänen, K., Häkkinen, A., Reinikainen S.-P., Louhi-Kultanen, M., Nyström L., 2005, *ATR-FTIR in monitoring of crystallization processes: Comparison of indirect and direct OSC methods*, Chemometrics and Intelligent Laboratory Systems, Vol. **76**, pp. 25-35

II) Pöllänen, K., Häkkinen, A., Huhtanen M., Reinikainen, S.-P., Karjalainen M., Rantanen J., Louhi-Kultanen M., Nyström L., 2005, *DRIFT-IR for quantitative characterization of polymorphic composition of sulfathiazole*, Analytica Chimica Acta, Vol. **544**, Issue 1-2, pp. 108-117

III) Pöllänen, K., Häkkinen, A., Reinikainen S.-P., Louhi-Kultanen M., Nyström L., 2006, *A study on batch cooling crystallization of sulphathiazole: Process monitoring using ATR-FTIR and product characterization by automated image analysis*, Chemical Engineering Research and Design, Vol. **84**(A1), pp. 47-51

IV) Qu H., Pöllänen K., Louhi-Kultanen M., Kilpiö T., Oinas P., Kallas J., 2005, *Batch cooling crystallization study based on in-line measurement of supersaturation and crystal size distribution*, Journal of Crystal Growth, Vol. **275**, pp. e1857-e1862

V) Pöllänen K., Häkkinen A., Reinikainen S.-P., Rantanen J., Minkkinen P., 2006, *Dynamic PCA based approach on MSPC charts in nucleation prediction in batch cooling crystallization processes*, Chemometrics and Intelligent Laboratory Systems, in press

VI) Pöllänen K., Häkkinen A., Reinikainen S.-P., Rantanen J., Karjalainen M., 2005, Louhi-Kultanen M., Nyström L., *IR spectroscopy together with multivariate data analysis as a process analytical tool for in-line monitoring of crystallization process and solid state analysis of crystalline product*, Journal of Pharmaceutical and Biomedical Analysis, Vol. **38/2**, pp. 275-284

Contribution of the author

The author planned the ATR-FTIR calibration measurements and performed sulfathiazole calibration measurements in publications I, III, IV and VI. The author has planned all of the DRIFT-IR measurements and measured most of the crystal samples in publications II and VI. Crystallization experiments of the sulfathiazole was done by M.Sc. Antti Häkkinen and crystallization experiments of the C15 by M.Sc. Haiyan Qu. Some of the DRIFT-IR measurements were performed by Ms. Päivi Hovila. XRPD measurements in II and VI have been performed by D.Sc. Milja Karjalainen.

The author has performed basically all multivariate analyses for IR data treatment in papers I-VI except in paper II part of the OSC filtering procedures were performed together with M.Sc. Mikko Huhtanen.

The author has been in charge of the preparation of Papers I, II, V and VI. In Paper III the parts concerning concentration measurements were under the responsibility of the author. The author participated in the writing of Paper IV.

Symbols and abbreviations

A	total number of principal components extracted from the data matrix in modeling, [-]
A^*	maximum number of eigenvectors that can be included in the model, [-]
Abs	measured absorbance, [-]
$\mathbf{A}_{PAR}$	loading matrix for the first mode of three dimensional data matrix ($n \times A$), [-]
a	number of principal component, [-]
$\mathbf{B}_{PAR}$	loading matrix for the second mode of three dimensional data matrix ($m \times A$), [-]
$\mathbf{b}$	vector of regression coefficients ($1 \times m$), [-]
b_0	regression coefficient of the zero intercept, [-]
$\mathbf{C}_{PAR}$	loading matrix for the third mode of three dimensional data matrix ($k \times A$), [-]
$Contr_T^2$	T^2 variable contribution, [-]
$Contr_Q$	Q variable contribution, [-]
c	concentration, [mol/dm^3] or [g solute/100 g solvent]
c^*	equilibrium concentration, [mol/dm^3] or [g solute/100 g solvent]
Δc	supersaturation, [mol/dm^3] or [g solute/100 g solvent]
$\mathbf{D}_k$	diagonal matrix of the kth row of $\mathbf{C}$ in three dimensional data matrix ($k \times A$)
d	coefficient for the selected confidence limit level, [-]
d_p	depth of penetration of the radiation into the sample, [m]
$\mathrm{d}Q_{contr,lim}$	difference between 95% confidence limits of Q contributions of calibration stage and prediction stage of the model, [-]
$\mathrm{d}T^2_{contr,lim}$	difference between 95% confidence limits of T^2 contributions of calibration stage and prediction stage of the model, [-]
$\mathbf{E}$	residual matrix of $\mathbf{X}$ decomposition ($n \times m$), [-]
$\mathbf{E}_{\mathbf{OSC}}$	data matrix from which the OSC component(s) are extracted ($n \times m$), [-]
e	residual from the model or score value from non-retained eigenvectors for single sample and single variable, [-]
$\mathbf{F}$	residual matrix of $\mathbf{Y}$ decomposition ($n \times q$), [-]
$\mathbf{f}$	residual vector of $\mathbf{y}$ decomposition ($n \times 1$), [-]
F_{crit}	tabulated one-sided value for (A^*-A) and (A^*-A)(n-A-1) degrees of freedom, [-]
$\mathbf{h}_\mathbf{X}$	leverage vector for $\mathbf{X}$ space ($n \times 1$), [-]
$\mathbf{h}_\mathbf{XY}$	leverage value for $\mathbf{XY}$ space ($n \times 1$), [-]
I_0	intensity of incident energy transmitted, [-]
I	intensity of transmitted light, [-]
I_{PAR}	total number of variables in the first mode of three way array, [-]
i	object/sample number, [-]
i_{PAR}	the sample number in the first mode of three way array, [-]
J_{PAR}	total number of samples in the second mode of three way array, [-]
j	variable number, [-]
j_{PAR}	the sample number in the second mode of three way array, [-]
K_{PAR}	total number of samples in the third mode of three way array, [-]
k_{PAR}	the sample number in the third mode of three way array, [-]
l	path length, [m]
m	number of descriptive variables, [-]

n	number of objects or measured samples, [-]
n_1	refractive indice of the ATR-element, [-]
n_2	refractive indice of the sample, [-]
$\mathbf{P}$	loading matrix of $\mathbf{X}$ data ($A \times m$), [-]
$\mathbf{p_a}$	loading vector of $\mathbf{X}$ data for the component a ($1 \times m$), [-]
$\mathbf{p_{OSC}}$	loading vector of OSC component ($1 \times m$), [-]
$\mathbf{q_a}$	loading vector of Y data for the component a ($1 \times q$), [-]
q	number of response variables, [-]
q_a	chemical loading
Q	squared perpendicular distance of a multivariate observation from the projection space, [-]
Q_{cv}^2	cross validated coefficient of determination, [-]
R	ratio of the sample diffuse reflectance spectrum and a non-absorbent reference sample, [-]
R^2	coefficient of determination, [-]
r_i	studentized residual, [-]
$\mathbf{S}$	diagonal matrix with diagonal elements equal to eigenvalues, [-]
S	relative supersaturation , [-]
s	standard deviation, [-]
$\mathbf{s}$	vector of standard deviations ($n \times 1$), [-]
s_{crit}	confidence limit for distance to the model, [-]
$s_{t_a}^2$	estimated variance of the score vector $\mathbf{t}_a$, [-]
$\mathbf{T}$	score matrix from $\mathbf{X}$ decomposition ($n \times A$), [-]
t	time, [s]
$\mathbf{t}_a$	score vector for component a from $\mathbf{X}$ decomposition ($n \times 1$), [-]
$\mathbf{t_{OSC}}$	OSC component score vector ($n \times 1$), [-]
T	temperature, [K] or [°C]
T^2	distance from the multivariate mean to the operating point on the PC plane, [-]
T_0	initial temperature, [K] or [°C]
T_f	final temperature, [K] or [°C]
Tr	transmittance, [%]
$\mathbf{u}_a$	score vector for component a from $\mathbf{Y}$ decomposition, [-]
$\mathbf{w_{OSC}}$	OSC component weight vector ($1 \times m$), [-]
$\mathbf{X}$	matrix of descriptive variables ($n \times m$), [-]
$\mathbf{x}$	vector of descriptive variables of one sample ($n \times 1$), [-]
$\hat{\mathbf{x}}$	modeled variation of variables of one sample ($n \times 1$), [-]
$\bar{x}$	mean of the population variable, [-]
$\mathbf{Y}$	matrix of response variables ($n \times q$), [-]
$\mathbf{y}$	vector of response variables ($n \times 1$), [-]
$\mathbf{y}_i$	measured response value of sample i, [-]
$\hat{y}_i$	predicted response value of sample i, [-]
$\bar{y}$	mean of measured response values, [-]
$\mathbf{w}$	weight vector from $\mathbf{X}$ decomposition, [$1 \times m$]

Greek letters

λ	wavelength of the incident radiation, [nm]
$\Delta\nu$	
ε	molar extinction coefficient
θ	angle off incidence of the reflected radiation, [°]

τ	batch time, [s]

Subscripts

hom	homogenous nucleation
het	heterogenous nucleation
lim	limit
contr	contribution
met	metastable zone width
surf	surface nucleation

Abbreviations

ATR	attenuated total reflection
CSD	crystal size distribution
DRIFT-IR	diffuse reflectance Fourier transform infra red
DSC	differential scanning calorimetry
FDA	U.S. Food and Drug Administration
FTIR	Fourier transform infra red
IRE	internal reflecting element
MSPC	multivariate statistical process control
MSC	multiplicative scatter correction
NIR	near infra red
OSC	orthogonal signal correction
PARAFAC	parallel factor analysis
PAT	process analytical technology
PCA	principal component analysis
PCR	principal component regression
PLS	partial least squares
PRESS	predicted residual error of sum of squares
QSAR	quantitative structure activity analysis
RMSEC	root mean squared error of calibration
RMSEP	root mean squared error of prediction
SIMCA	soft independent modeling of class analogy
SNR	signal to noise ratio
SNV	standard normal variate filtering
SPE_x	squared prediction errors from **X**-phase decomposition
TG	thermal gravimetry
XRD	X-Ray diffraction
XRPD	X-Ray powder diffraction

1 INTRODUCTION

Nowadays, more and more effort is directed to the use of on-line process monitoring and control in pharmaceutical manufacturing processes in order to enhance the understanding of the unit operations and also to ensure high quality product and small batch-to-batch variations (Yu et al., 2004). The U.S. Food and Drug Administration (FDA) has recently addressed an initiative of Process Analytical Technology (PAT), which states that the new efficient tools that can be used during pharmaceutical manufacturing and quality control are to be developed and implemented while maintaining or improving the current level of product quality assurance (FDA, 2004). The tools to be utilized can be divided into four groups:

I) Multivariate data acquisition and analysis tools
II) Modern process analyzers or process analytical chemistry tools
III) Endpoint monitoring and process control tools
IV) Knowledge management tools

The use of some or all of these tools is either applicable to a single unit operation or to an entire manufacturing process and its quality assurance (Yu et al., 2004). The use of PAT should lead to :

I) Better process understanding and fewer process failures
II) Ensurance of the product quality using optimal design, continuous monitoring and feedback control
III) Reduction in the cycle time, which further improves manufacturing efficiency
IV) Identification for the reasons for deviations within the processes
V) Better process knowledge and scientifically based risk assessment

(FDA, 2004; Yu et al, 2004).

The pharmaceutical manufacturing process of solid state drugs includes several unit operations, which are synthesis, crystallization, mixing, granulation, tabletting and coating. Crystallization is an important purification unit operation within the pharmaceutical industry, and the quality of the crystallized product: size, size distribution, shape and purity all influence on the further processability and usability of the product. It is an essential issue that the desired product quality can be obtained. Traditionally, the pharmaceutical crystallization processes have been followed by laboratory testing and analysis to verify the product quality (Yu et al., 2004). In addition, rather large batch-to-batch variations in the crystallization processes have resulted. Understanding and the control of crystallization processes has not been sufficient. Novel monitoring as well as data management tools are of great importance in enhancing the controllability of the crystallization process and resulting crystalline product.

The concepts within the crystallization process where PAT tools could be applied and what could be obtained using these tools are illustrated in Figure 1.

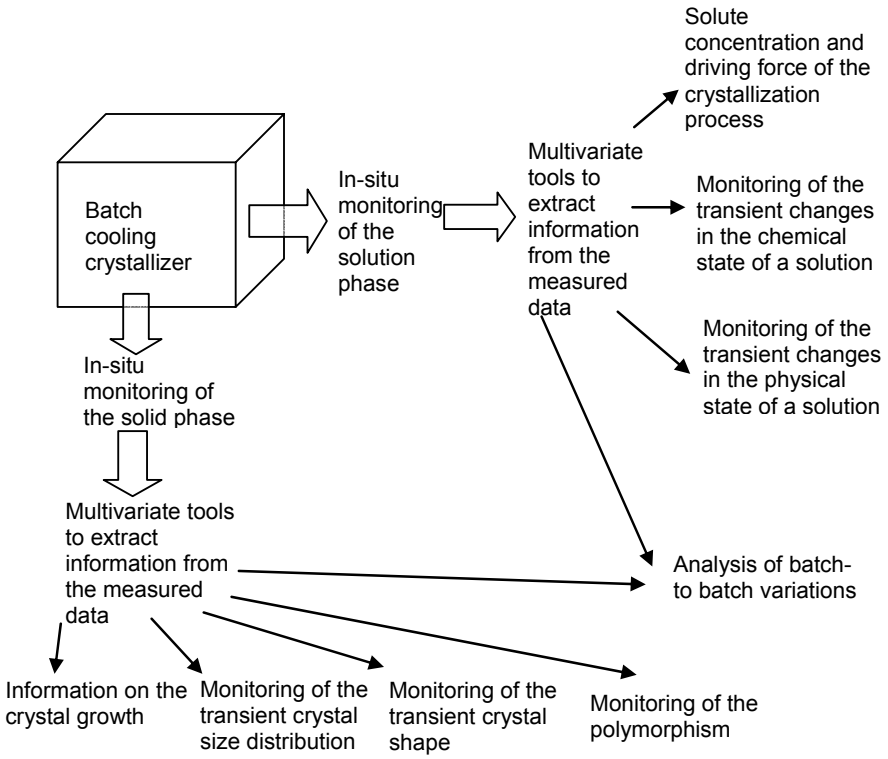

Figure 1 The PAT tools that can be applied in monitoring the crystallization process and the information that can be obtained

Figure 1 shows that a wide range of the different phenomena could be obtained by applying the modern real-time analyzers to collect information from the crystallizer. The obtained information should lead to both a better understanding of the process via the ability of monitoring and visualizing the true transient phenomena inside the crystallizer. Also the ability of monitoring, e.g., the solute concentration and chemical as well as physical changes inside the crystallizer should enhance the control possibilities of the crystallization processes. For example, when the solute concentration can be measured in real time, this leads to the possibility of feedback control of the crystallizer. Also the knowledge of the process can be enhanced via batch-to-batch variation analysis derived from measured data.

PAT tools can and should also be applied to off-line characterization of crystalline samples: Modern image analyzers produce high quality size and shape data, but the full utilization requires proper data treatment methods. The polymorphic composition of the product can be

evaluated using various different analytical techniques. To obtain qualitative and quantitative information from different kind of analytical data multivariate methods can be applied.

The utilized tools to monitor crystallization processes will lead to a deeper understanding of the phenomena in the crystallization process, which further improves the crystallization process design. The utilization of the tools is also economical in the long run, as a better monitored and controlled processes will lead to reduced process failiors and batch-to-batch variations.

2 AIM OF THE STUDY

The aim of the study is to enhance the understanding of batch cooling crystallization processes and improve the product quality by

1) Developing and evaluating methods based on attenuated total reflection Fourier transform infra red (ATR-FTIR) spectroscopy and multivariate analysis techniques for in-situ monitoring of the solution phase in a batch cooling crystallization process

2) Developing methods based on difffuse reflectance Fourier transform infra red (DRIFT-IR) spectroscopy and multivariate analysis technicues for product quality evaluation.

3 OUTLINE

The thesis is divided into theory and experimental parts. In the theory part, the fundamentals of batch cooling crystallization processes are presented concentrating on the issues which are of importance to crystallizaiton process monitoring and crystalline product quality characterization. The basic principles of spectroscopic techniques used in this particular study are introduced, and the focus of a discussion lies on the special issues related to crystallization monitoring using vibrational spectroscopy. The multivariate methods applied within this study are introduced shortly and, again, concentration relies on the use of multivariate methods in spectral data treatment, which is the application in this study.

In the experimental part, the experiments presented in the publications I-VI are shortly explained and the main results obtained and published in I-VI are summarized. In addition to the results presented in publications I-VI the batch-to-batch variation analysis of the batch cooling crystallization process is included in the experimental section as new results.

4 BATCH COOLING CRYSTALLIZATION PROCESS AS A PURIFICATION UNIT OPERATION IN THE PHARMACEUTICAL INDUSTRY

4.1 Quality of the crystalline product

The crystalline product quality is usually defined by the size, crystal size distribution (CSD), habit, polymorphic form and purity of the product (Jones, 2002; Mersmann, 1995) Usually, a certain mean size and narrow size distribution as well as high polymorphic purity is desired. The requirements for the crystal properties are strongly application dependent. However, the mean crystal size, CSD and habit affect the downstream processes of the pharmaceutical crystalline substance such as filtration, drying, milling, blending, granulation and tabletting (Barret et al., 2005). Usually an attempt is made to avoid fines generation, since fines can remarkably retard the filtration process (Barret et al., 2005).

Polymorphism is defined as the ability for a substance to exist as two or more crystalline phases, that have different molecular arrangements (Brittain, 1999). The different polymorphic forms differ from each other by several physical properties (Brittain, 1999):

I) Packing properties: e.g., molar volume and density, hygroscopicity
II) Thermodynamic properties: e.g., thermodynamic activity, solubility
III) Spectroscopic properties: e.g. vibrational transitions
IV) Kinetic properties: e.g., dissolution rate, sates of solid state reactions, stability
V) Surface properties: e.g., surface free energy, interfacial tensions, habit
VI) Mechanical properties, e.g., hardness, tensile strength, compactibility, tabletting, handling, flow and blending

Obviously differences in these properties cause differences in further processability in downstream processes. In addition to this the pharmaceutical performance of the crystalline product, for instance, dissolution, stability and usability in the final dosage form can be different for different polymorphic forms. Usually, the lowest energy crystalline polymorph, *stable* form, is chosen for development (Singhal and Curatolo, 2004). There are some exceptions to that rule: If the most stable form has insufficient solubility to have desired healing effect or drugs for quick relief are developed, the more soluble *metastable* form can be selected for further development. In addition, the ease of manufacturing or economical reasons can be a reason for developing a metastable polymorph (Singhal and Curatolo, 2004).

The product quality in the crystallization process is affected by processes of nucleation, growth, attrition, breakage and agglomeration. The operating conditions, the type of the crystallizer and the material properties of the liquid and solid phases influence in the product quality (Mersmann, 1995). The factors that may have an effect on the polymorphic form of the crystals

include: solvent composition, degree of supersaturation, the temperature range used, additives, seeding, agitation (Brittain, 1999).

4.2 Supersaturation

The fundamendal driving force of the crystallization process is the change in the chemical potential between the prevailing and equilibrium state (Jones, 2002). Chemical potential is not easy to measure, therefore, the concentration of the solute in excess solubility is commonly used to refer driving force of crystallization process

$$\Delta c = c - c^* \tag{1}$$

where Δc is the supersaturation also called as the concentration driving force (Figure 2), c is the concentration present in the crystallizer, c^* is the equilibrium concentration at a certain temperature. In addition, the relative supersaturation in the isothermal system (S) is commonly used:

$$S = \frac{c}{c^*} \tag{2}$$

Typically the concentration is expressed in molar units but also mass concentration or mass ratios can be used (Jones, 2002). The supersaturated state in solution is generated differently depending on the crystallization process used. These different processes include cooling, evaporation, drowning out or a chemical reaction (Mersmann, 1995), of which the cooling crystallization is considered in this study.

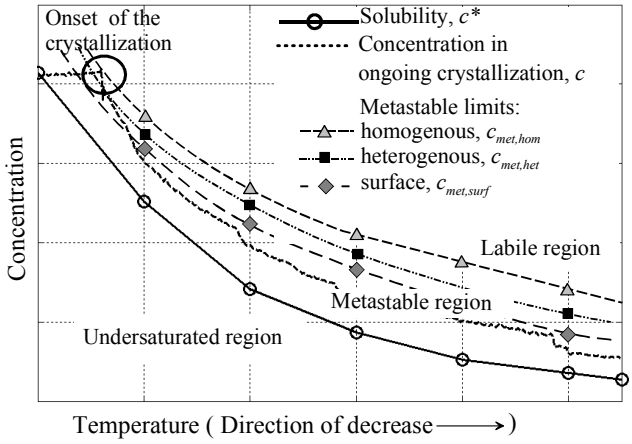

Figure 2 The schematic illustration of a solubility curve, concentration profile during the crystallization and different imaginary metastable regions

In the cooling crystallization systems, the *solubility* of the solute in used solvent increases as the temperature increases. When this solubility concentration is obtained the system is said to be under equilibrium. The *supersaturated* stage is achieved when the system under equilibrium is cooled down. The supersaturated stage can be divided into two groups: the *metastable region*, where the solute molecules tend to transfer onto existing crystals and the *labile region* where in addition to mass transfer onto crystals, new nuclei are formed spontaneously. Basically, the solute to be crystallized is dissolved into the solvent at a certain, elevated temperature. Then the clear solution is cooled down and the supersaturated system is obtained. First crystals are introduced to the system (*onset of crystallization*) by adding seed crystals or exceeding the metastable limit. As the first crystals are introduced, the solute molecules in excess solubility tend to move onto crystals to release the supersaturation and attempt to go for equilibrium. By further cooling, however, the solution is kept supersaturated, which further causes the existing crystals to grow and/or new nuclei to be formed.

The supersaturation level influences on the different mechanisms in the cooling crystallization process: nucleation, growth, agglomeration and aggregation as well as polymorph transitions which are considered in the following four chapters. The supersaturation level can be controlled by controlling the crystallization process conditions, which is dealt with in Chapter 4.3.

4.2.1 Metastable zone width and nucleation processes

The nucleation processes and metastable zone width have clear correlation. In a simplified manner said, exceeding the metastable limit causes the system is under labile region, where new nuclei can be formed (Figure 2). Thermodynamic and kinetic equations for nucleation processes exist in literature (e.g. Mersmann, 1996; Mullin, 2001; Jones, 2001; Myerson, 1993). The concept of metastable zone limit is not well defined neither kinetically nor thermodynamically (Ulrich and Strege, 2002).

Several different kinds of nucleation processes exist. Different nucleation processes dominate at different supersaturation levels. A commonly used classification is to present three different metastable zones (Ulrich and Strege, 2002; Mersmann, 1996), example of which Figure 2 illustrates. The limit of homogenous nucleation ($\Delta c_{\mathrm{met,hom}} = c_{\mathrm{met,hom}} - c^*$) in Figure 2 is the limit where spontaneous nucleation can occur from clear solution without a solid phase present. The limit for heterogeneous nucleation ($\Delta c_{\mathrm{met,het}} = c_{\mathrm{met,het}} - c^*$) is the limit above which the surface nuclei can be formed on possibly existing foreign particles or rough surfaces in the crystallizer. Homogenous and heterogeneous nucleation processes refer to primary nucleation processes and these can take place without any crystalline material present (Ulrich and Strege, 2002; Mersmann, 1996).

The primary nucleation influences on the mean particle size. The higher the supersaturation is at the moment of primary nucleation, the larger is the number of nuclei formed resulting in smaller mean size of the product (Ulrich and Strege, 2002; Mohameed, 2002). The nucleation process affects essentially on the polymorphic form of the product crystals.

Surface nucleation can occur when corresponding metastable zone limit ($\Delta c_{\mathrm{met,surf}} = c_{\mathrm{met,surf}} - c^*$) is exceeded. New nuclei form onto of existing crystals. Attrition and breakage of existing crystals refer to the fourth nucleation type, which unlike the other nucleation types is not concentration dependent. Attrition occurs due to collisions of crystals with each other as well as to crystallizer walls or to impeller (Mersmann, 1996; Ulrich and Strege, 2002). The impeller configuration, impeller speed, crystallizer configuration, as well as properties of the solid and liquid phase influence strongly on the degree breakage of crystals and attrition nucleation. In order to these attrition fragments to grow to nuclei and crystals, the supersaturated stage is required. Attrition nucleation is often the dominating nucleation mode in crystallizations from solution (Virone et al., 2005). Surface nucleation and attrition nucleation are referred to as secondary nucleation processes since those are due to the existing crystals in the system.

CSD and mean crystal size of the product crystals are strongly dependent on the degree of the secondary nucleation processes. Theoretically, the level of surface nucleation during a crystallization process could be avoided by controlling the supersaturation level to the stage where no nucleation processes exist. Consequently, this should lead to narrow CSD. In practice, as the defining the metastable limits unambiguously are not possible, the avoidance of nucleation during ongoing crystallization is difficult to do. The secondary surface nucleation and heterogeneous nucleation can cause also variations in the polymorphic composition due to different solubilities of the polymorphs in different stages of crystallization. The attrition nucleation is not dependent on supersaturation level, but the fragments grow differently at different supersaturation levels: at relatively high supersaturations, larger fragments grow faster than smaller fragments and at low supersaturations, zero growth and fragments can dissolve partially or totally (Virone et al., 2005).

In practice, different nucleation processes, especially, heterogeneous, surface and attrition nucleation can be simultaneously present. Mersmann, 1996, Kim, and Mersmann, 2001 presented equations for theoretically predict metastable zone width. The width of the metastable limit depends on several different factors, e.g., temperature level, cooling rate and mixing conditions (Tähti et al., 1999; Ulrich and Strege, 2002). Increasing the cooling rate broadens the metastable zone. Presence of impurities narrows the metastable zone width. The concentration

limit, where certain nucleation processes begin cannot be unambiguously experimentally determined (Mersmann, 1996).

4.2.2 Nucleation and polymorphs

The nucleation process is the most critical in forming polymorphs (Brittain, 1999). In the nucleation of polymorphs the Ostwald's step rule is assumed to apply. By that rule the polymorph with the highest Gibbs' free energy is the least stable and forms first. Under certain thermodynamical conditions, however, a solution phase mediated fast polymorph conversions from less stable to more stable form can take place which is the principle of the Ostwald's step rule (Brittain, 1999). In that process, the less stable polymorph nuclei first dissolves and after that the more stable polymorph crystallizes out. This rule is not a thermodynamic law, and it is not always obeyed, however. The supersaturation level effect on the nucleating polymorphs can be thought of as follows: the supersaturation level that is present at the primary nucleation can affect the rate of which the Oswald's step rule is taking place.

The nucleation of polymorphs can be also understood when considering, what happens in the process of primary nucleation. As the supersaturation level in a clear solution approaches the homogenous metastable limit, the solute molecules tend to move closer to each other and can form aggregates, which is an attempt to further reduce the Gibbs energy (Brittain, 1999). Different orientation of aggregates likely causes a different polymorphic form to be nucleated. It is assumed that the polymorphic form of nuclei dictates the polymorphic form of the product crystals. Several different aggregates can be present in the solutions simultaneously. It is assumed that the aggregate which has the highest concentration or for which the critical activation energy is the lowest will form the first nucleus leading to the crystallization of this particular polymorph (Brittain, 1999). It is also possible that more than one polymorph may nucleate and, as a consequence, a mixture of polymorphs is obtained as product. The different polymorphs may nucleate during the course of crystallization thorough secondary surface nucleation. The polymorphic form of that is nuclei at this stage can be different than the ones that have been formed in the early stage of crystallization.

Factors, besides supersaturation that may influence on the nucleating polymorphic form and purity of the crystals in addition to supersaturation include, e.g., solvent selection, temperature range of crystallization process and seed crystals. It is not well understood, which of these factors dominate, however.

4.2.3 Crystal growth

The growth of crystals in a supersaturated solution is a very complex process. In general, an increase in supersaturation increases the crystal growth rate, but at the same time the secondary nucleation processes are increased (Ulrich and Strege, 2002; Paul et al., 2005). The balance between the growth and nucleation is a critical issue regarding the product quality (Paul et al., 2005). In order to minimize the width of the CSD, growth should be dominating process over secondary surface nucleation. This can in principle be obtained by maintaining a very low supersaturation level thoroughout the crystallization process. This can, however, lead to the uneconomical operation of the crystallizer. The crystal growth rate is particle size dependent (Mersmann, 1996; Myerson, 1993) and, in practice, the growth rate decreases as the crystal sizes increases. To obtain economical operation of the crystallizer, the optimization of the equal growth in the dynamic transient state process should be considered. Different growth rate can also lead to different crystal shapes (Ulrich and Strege, 2002). In industrial mixed tank crystallizers, crystal breakage due to collisions with each other, with the walls of the crystallizer and to the impeller can be to a great scale. Therefore, in practice the true product outcome in terms of size or habit cannot be evaluated simply by crystal growth rates or directions, or supersaturation level, but it can strongly be altered by mixing conditions.

4.2.4 Crystal agglomeration

Agglomeration is the process where two or more crystals attach to each other by as a result of malgrowth crystals or crystal crystal collisions in supersaturated solutions. Agglomeration is a dominating process for the very small particles in the submicron and micron range and neglible for large particles (Mersmann, 1996). Agglomeration should avoided because it causes reduced affective surface area (Paul et al., 2005).

The agglomeration level depends on the movement of primary particles and liquid as well as the number of collisions in the supersaturated solution. As the very small particles tend to agglomerate, the rate of nucleation should be neglible in order for agglomeration to be prevented, thus the crystallization process should be run within the metastable zone. In a primary nucleation process the number of crystals is related to the supersaturation level, and at low supresaturation level agglomeration is less probable than at high supersaturation levels (Paul et al., 2005).

4.3 Control of cooling crystallization systems

The aim of the control of the batch unit operation is to obtain quality product economically and reliably with negligible batch-to-batch variations. The controllable factors in cooling crystallization process include the crystallization phenomena influencing the operating

conditions: the way of carrying out the on-set of the crystallization, what type of cooling mode and rate is used, what are the mixing conditions, and the solvent selection.

4.3.1 On-set of the crystallization process

The *on-set* of the crystallization process occurs by introducing the first crystals/nuclei into the supersaturated clear solution. The onset of the crystallization process is important since it defines the number of the primary nuclei in the system and most likely, the polymorphic form of the crystals, since the nuclei define the lattice towards which the growth is oriented (Brittain, 1991). There are two approaches for on-set of crystallization: *unseeded* and *seeded*.

In unseeded crystallizations, primary crystal nuclei form spontaneously from solution when the supersaturation level exceeds the either limit of primary nucleation (Figure 2). Consequently, drastic decrease in concentration level decrease drastically. This can be considered as an uncontrolled way of starting the crystallization process, because the homogenous nucleation is stochastic process (Davey and Garside, 2000; Ulrich and Strege, 2002). However, the number of nuclei formed and polymorphic form that nucleates depend on controllable process parameters: cooling rate, nucleation temperature, and mixing conditions. The key into understanding the primary nucleation of the polymorphs, especially the supersaturation level impacts on it, is to be able to monitor the system and to see different aggregates priori to nucleation.

In a seeded crystallization process, crystallization stage begins by introducing a small amount of seed crystals to the metastable system within the metastable range, which causes the supersaturation release on the surfaces of the added seed crystals and the crystallization process begins. This traditionally refers to a controlled way of crystallization onset (Davey and Garside, 2000). The quality of the seeds: amount, mean size, and CSD of seed crystals affect strongly on the crystalline product obtained from batch (Kubota et al., 2001; Doki et al. 2004). The target of seeding is that only the seed crystals grow causing a narrow CSD for the product, and no secondary nucleation occurring. This would require seeding such that the supersaturation level after seeding would not increase dramatically, but the supersaturation release onto seeds only. Kubota et al. (2001) demonstrated that optimization of the amount of seed crystals can prevent secondary nucleation. In the concept of mean, size of the seed crystals Mersmann (1995) states that to obtain the large surface of the seed crystals, which is favorable for growth small seed particles are advantageous over large coarse particles. Presumable, the narrow CSD of the seed crystals would be favorable in order to those have equal growth rate and result in even sized crystals. Introducing the seed crystals into the surely saturated system prevents dissolving of the seeds or presence of the nuclei before seeding (Mersmann, 1995). Real time monitoring of the

change in the state from undersaturated thorough to saturate into supersaturated and distinguishing the correct seeding moment would increase the knowledge and controllability of seeded crystallization processes.

Polymorphic form of the seed crystals strongly influences the polymorphic form of the resulting product (Kitamura, 2004, Beckmann, 2000). In order to obtain a pure polymorphic form seed crystals should be pure polymorphs. The product presumably exhibits the same polymorph as seeds, but to be sure, the operator must be sure that the seed crystals do not dissolve and that secondary surface nucleation is negligible. To obtain product of the pure polymorph understanding of the polymorphism of the whole system is required and process conditions controlled to avoid formation of undesired polymorphic forms.

4.3.2 Cooling modes

The level of supersaturation in the cooling crystallization processes is usually manipulated by changing the cooling policy in the crystallizer. When the cooling rate is very low, the mass transfer of the solute molecules on the surface of existing crystals is fast enough to release most of the supersaturation and consequently the supersaturation level remains low. When the cooling rate increases, the solubility decrease with decreasing temperature becomes greater than the mass transfer rate onto the crystals, which causes the supersaturation level to increase and remain high. The supersaturation level increases as the system cools down faster than is the mass transfer of the solute molecules from the solution onto the crystals (Mersmann, 1995). The constant cooling rate do not necessarily cause the constant supersaturation level throughout the process, because as the crystallization process proceeds the solid phase increases which causes changes in mass transfer rate.

Different cooling modes in the cooling crystallization processes can be divided into three different main categories: natural cooling, constant cooling and programmed cooling rates. The schematic representation of these is presented in Figure 3.

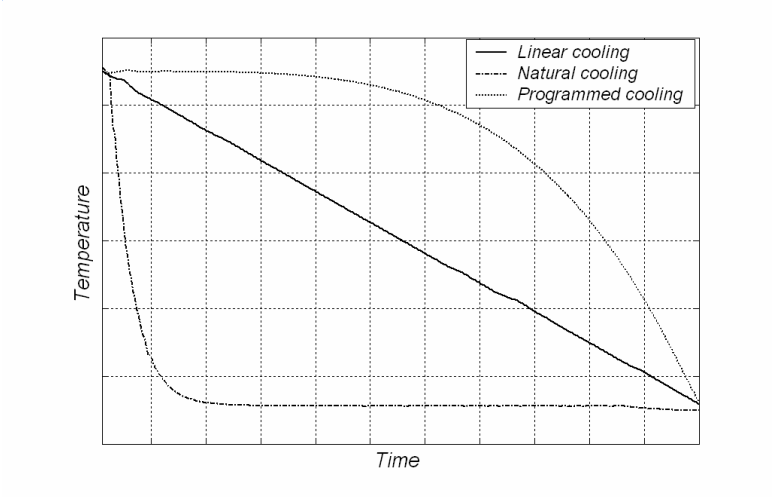

Figure 3 Schematic presentation of different cooling policies usually applied in cooling crystallization processes

In natural cooling, the system is cooled down using a constant temperature coolant. This results in typically very high cooling rates at the beginning of the batch (Figure 3). This further means that it is possible that the system is supersaturated faster than is the desupersaturation due to growth of the crystals. This can cause the metastable limit to be exceeded and a considerable spontaneous nucleation process to take place. This can result in very large number of crystals which are rather small and irregularly shaped.

The constant cooling rate is predetermined constant cooling rate applied throughout the process. If the constant cooling rate is too high, the possibility of exceeding the metastable limit repeatedly throughout the crystallization process and the high level of uncontrolled secondary nucleation can be resulted. Basically, the lower the cooling rate is the more unlikely is to exceed any metastable limit, this principle can, however, lead to uneconomically long batch times. Therefore, the cooling rate should be optimized. The linear cooling rate does not take into account the transient state of the solid phase and thus, the increased mass transfer rate from the liquid phase to solid phase as the process proceeds.

Programmed cooling mode has been developed in order to find the theoretical optimal cooling mode for the transient crystallization system. The objective of such cooling modes is to ensure that the rate of the supersaturation generation is always the same as is the mass transfer onto the growing crystals. This means that when a small amount and/or small crystals exist the cooling rate will be small and as the crystallization process proceeds the solid surface will increase, the cooling will be increased as well. A theoretical equation for ideal cooling in seeded crystallization processes with constant nucleation and growth rates was derived by Mullin and

Nyvlt (1971). This equation was very difficult to apply in practice, since it contains several terms that are time dependent such as the mean seed size at a certain time during the crystallization. Therefore, simplifications of this equation have been proposed (Mullin and Nyvlt, 1971; Jones and Mullin, 1974). The simplified cooling curve in order to maintain constant supersaturation during the whole crystallization process can be presented as follows:

$$\frac{T_0 - T}{T_0 - T_f} = \left(\frac{t}{\tau}\right)^4 \tag{3}$$

where t is time, τ is the total batch time, T_0 is the initial solution temperature, T is the solution temperature at time t and T_f is the final solution temperature. The Equation 3 is for an unseeded crystallization process where the assumption is that the solubility has linear dependence on temperature. This controlled cooling should therefore favor the growth over nucleation and thus lead to a narrower CSD (Choong and Smith, 2004; Costa et al, 2005; Mullin, 2001). However, results indicating that the optimal cooling probably could not be achieved using proposed cooling profile also exist (Kubota et al., 2001; Pöllänen et al., 2005 b, 2006)

Recently, the development of a closed loop control for crystallization systems has become to a hot topic. (Fujiwara et al., 2002; 2005; Grön et al., 2003; Liotta and Sabesan, 2004) The objective of such control procedures is usually to maintain the desired supersaturation level throughout the process by first measuring the supersaturation level and based on that control the cooling of the system. There are two possible approaches to this type of feedback control approaches: measurement based or model based. In the measurement based applications online measurement of the concentration and temperature is used to calculate the control effect on temperature. In the model based control the mathematical model of the crystallization process including heat and mass balances is used as the basis, and the model parameters are updated based on the measurement (Grön et al., 2003).

4.3.3 Mixing conditions

Mixing conditions are essential issues when discussing the control of a batch crystallizer. Sometimes the wrong mixing conditions can overrule the otherwise intelligent control of supersaturation and the on-set of the crystallization process. Changes in the mixing intensity and different impeller types cause different levels of breakage of the crystals, which changes the properties of the solid phase in terms of the number of crystals present and the specific area of the solid phase available to grow (Yu et al., 2004). The increasing mixing rate increases the number of possible collisions in the crystallization, which can increase the possibility of two or more crystals to agglomerate.

Good mixing usually refers to similar mixing conditions throughout the reactor and the enhancement of the mass transfer between the solution and solid phase due to the smaller boundary layer thickness (Virone et al., 2005). The mixing condition effects divide into macroscopic scale effects and microscopic scale effects. The macro mixing means the overall performance of the mixing in the reactor and correspondingly the micro mixing refers to the turbulent mixing on the molecular level (Jones, 2002). Different impeller types give different macro vs. micro mixing distributions and different attrition and breakage levels (Shimizu et al., 1998). Mersmann and Löffelmann (2000), suggest that the final product size can be strongly determined by the attrition for crystals above 100 μm. It can also be further speculated that the largest crystals are likely to suffer most from attrition, because they have the higher collisional probabilities due to their larger area and mass (Matthews and Rawlings, 1998). Davey and Garside (2000) have stated that it is currently recognized that attrition nucleation is the most significant nucleation mechanism in crystallizers for materials with high or moderate solubility.

4.3.4 Solvent selection

The solvent selection is one way to control crystallization process. For example, the solvent can be the dominating factor, e.g., to changes in the polymorphic form. (Bladgen, 2001; Bladgen et al, 1998a, Bladgen et al., 1998b) Some solvents seem to favor the formation of certain polymorphic forms, because these can selectively adsorb to certain faces of some polymorphs and the nucleation or growth of those is inhibited or retarded, which is the advantage for the nucleation or growth of other polymorphs (Brittain, 1999). In addition to this, the solvent can alter the relative solubilities of the polymorphs; therefore, the solvent selection can alter the polymorph nucleation processes (Brittain, 1999). The differences in solution viscosity, density, and diffusivity as well as changes in the solid-liquid interfacial energy cause changes in the growth rate and habit of the crystals (Myerson et al., 1986). In addition to this, the solubility of the solute can be remarkably different in different solvents, which can influence on the variability in nucleation and crystal growth. Solvent is traditionally selected been based on experience, analogy and experimental testing (Kolár et al., 2001). There are also proposals for systematic ways to select solvents. (Kolár et al., 2001; Myerson et al, 1986)

4.3.5 Summary of the crystallization process scheme

The summary of the crystalline product quality measures, different phenomena present in the crystallization process affecting on the product quality and the controllable process parameters are compiled in Figure 4.

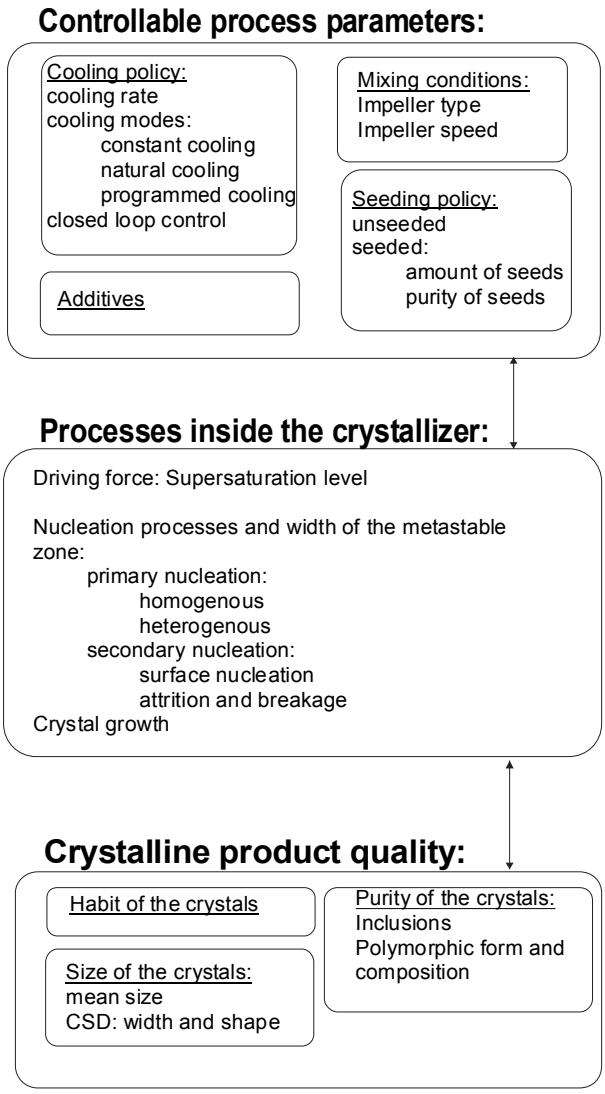

Figure 4 The crystallization process system: Controllable process parameters, processes inside the crystallizer and crystalline product quality measures.

To obtain the desired product quality all the different processes inside the crystallizer should be under control and the correct settings for the process parameters driving the process conditions should be defined. To produce a full monitoring scheme for the crystallization process, the transient phenomena inside the crystallizer should be monitored, i.e., the evolvation of the supersaturation level and the changes in the solid phase should be evaluated real time. The quality of the product should be evaluated fast and reliably. In addition, to enhance the knowledge of the crystallization process in practice, the factors causing batch-to-batch variations should be explored and explained.

5 APPLICATIONS OF MONITORING CRYSTALLIZATION PROCESS AND CRYSTALLINE PRODUCT

The full scale monitoring of the crystallization process and the outcome require different techniques for different purposes, and no single technique to monitor all important phenomena inside the crystallizer exist. In following chapters, the state-of-art of applications used in crystallization process and product monitoring is briefly reviewed. Applications in vibrational spectroscopy together with multivariate methods in the crystallization process and product monitoring is more deeply outlined.

5.1 Process monitoring

5.1.1 Solution phase

Solution phase measurement includes the metastable zone width, on-set of the crystallization and supersaturation measurements. There are several methods tested for on-line measurement of supersaturation. Supersaturation can be determined from the density of a crystal free solution (Gutwald and Mersman, 1990; Qui and Rasmuson, 1994). Density measurements are not accurate in the case of industrial organic solutions where various concentrations of impurities are involved. In addition, the technique requires an external sampling loop, which may cause operating difficulties. Concentration of a crystallizing solution can be determined by measuring the electrical conductivity of the solution (Löffelmann and Mersmann, 2002). This application does not require external sampling, but there exist several limitations in use of conductivity measurement techniques: Organic systems are not usually conductors, undesirable crystallization might occur in the conductivity cell and the results are temperature dependent and sensitive to possible impurities. Calorimetric measurement can be considered as an indirect technique for supersaturation on-line measurements (Févotte and Klein, 1996a, b; Löffelmann and Mersmann, 2002). The heat flow can be associated with a chemical reaction by measuring the temperature change it produces. However, the thermal effects are typically weak and off-line computations are required for reliable results.

Vibrational spectroscopy provides a technique for in-situ crystallization process monitoring using immersion probes. Vibrational spectrum contains lot of information and the phenomena present in crystallization process can be monitored from measured spectral data. Attenuated total reflection Fourier transform infra red spectroscopy (ATR-FTIR) and Raman spectroscopy have proven to be reliable technique for solution phase monitoring (Uusi-Penttilä and Berglund, 1996; Dunuwila and Berglund, 1997; Togkalidou et al., 2001; Lewiner et al., 2001a; Lewiner et al., 2001b; Grön and Roberts, 2001; Fevotte, 2002; Grön et al., 2003; Fujiwara et al., 2002; Doki et al., 2004; Profir et al., 2002; Grön and Roberts, 2004; Feng and Berglund, 2002; Liotta

and Sabesan, 2004; Qu et al., 2004; Pöllänen et al., 2006; Pöllänen et al., 2005a, 2005b). In addition to that applications of in-situ Raman spectroscopy in crystallization monitoring have been proposed (Schwartz and Berglund, 1999; Schwartz and Berglund, 2000; Starbuck et al., 2002; Ono et al., 2004; Hu et al., 2005; Falcon and Berglund, 2004; Tamagawa et al., 2002a; Tamagawa et al., 2002b).

The advantages and disadvantages of vibrational spectroscopy in in-situ monitoring of crystallization process is presented in Table 1.

Table 1 Advantages and disadvantages of the use of vibrational spectroscopy in in-situ monitoring of a crystallization process.

Advantages	**Disadvantages**
No external sampling methods are required	Requires (complex) mathematical methods to be used in order to obtain quantitative information: Requires calibration and extensive calibration measurements
Techniques are widely applicable for various solute solvent systems	Sensitive to mechanical changes, e.g., concussions directed to the probe, changes the performance of the probe
Simultaneous measurement of several species, e.g., impurities	Temperature changes affect on the obtained spectrum
The monitoring of the primary nucleation process can be performed	Non-uniformity in the chemical atmosphere around the probe, e.g., due to imperfect mixing can cause erroneous results
Possibility to build a closed-loop control of the crystallization process	Immersion probe produces an additional flow resistance inside the crystallizer
Forming crystals do not interfere with the measurement of the solution phase (ATR-FTIR)	The strongly aggressive chemical environment can cause the detoration of the probe, e.g., oxidation of the probe
Measurement of transformation of polymorphs during the crystallization process is possible (Raman)	The crystals or air bubbles can attach to the surface of the probe, which gives an erroneous result
Can be used in aqueous media (Raman has very weak water responses and ATR-FTIR can be used due to fixed path-lengths which limits the absorbance to the sample)	Long term drifting causes the need to update calibrations periodically
	Fluoresence problems can ruin the measurement (Raman).

The ATR-FTIR technique has proved to be a suitable technique for monitoring crystallization processes due to its wide applicability to different systems, since most compounds absorb radiation in the mid-IR range. In addition, the measurement takes place at the interface between the ATR element and a sample and therefore the existing crystals are assumed to not disturb the measurements. However, transformation of the spectral data to concentration information is a critical issue in order to obtain reliable results.

Traditionally solute concentration predictions from spectral data measured from crystallization process have been done by correlating the height or area of the specific band to the solute concentration by regular linear regression (Uusi-Penttilä and Berglund, 1996, Dunuwila and Berglund, 1997, Fevotte, 2002, Lewiner et al., 2001a, 2001b, 2002; Ono et al., 2004; Hu et al., 2005). Multivariate calibration for solute concentration prediction from batch cooling crystallization processes using either principal component regression (PCR) or partial least squares (PLS) methods have been proposed by, e.g., Togkalidou et al., 2001; Profir et al., 2002; Liotta and Sabesan, 2004; Pöllänen et al., 2005c; Starbuck et al., 2002; Schwartz and Berglund, 1999; Schwartz and Berglund, 2002; Tamagawa et al., 2002a; Tamagawa et al., 2002b.

ATR-FTIR spectroscopy has been applied in monitoring of the purity of the racemic compound (Profir et al., 2002), the certain polymorphic form (Doki et al., 2004) and the of the level of the specified impurity simultaneously with the solute concentration measurement (Derdour et al., 2003).

The information on the concentration of crystallizing substance allows the possibility of investigating the effects of the driving force changes on the obtained product morphology and CSD. Consequently, the process conditions can be optimized to meet the desired product properties. The results from the studies in which ATR-FTIR was used for concentration measurement during the crystallization experiments using different constant cooling rates showed that the metastable range increased with the cooling rate and also with the overall concentration level, as was expected theoretically. The overall concentration level increased by seeding as well (Fevotte 2002, Lewiner et al., 2001a, 2001b, 2002). The average size of the product was observed to be largest when seeding was used (Lewiner et al., 2001a) and the number of crystals was found to increase with the cooling rate (Lewiner et al., 2001b). The variations in crystal shape in terms of length-to-width ratio were found to be reduced when seeding was applied (Lewiner et al., 2001b).

An application of the closed loop feedback control based on ATR-FTIR concentration measurements has been considered. (Fujiwara et al., 2002; Grön et al., 2003; Liotta and Sabesan, 2004). The results from crystallizations using closed loop control have shown that by

maintaining the concentration level under the metastable limit, larger and more uniform product crystals can be obtained than when the metastable limit is exceeded (Fujiwara et al., 2002). Controlled crystallization processes with constant supersaturation levels have been compared to uncontrolled crystallizations in previous studies (Grön et al., 2003, Liotta and Sabesan, 2004). Crystallization with feedback control with constant supersaturation was found to provide larger crystal sizes than crystallization with a constant cooling rate (Liotta and Sabesan, 2004).

The nucleation process can potentially be studied using vibrational techniques. However, only a few attempts to in predicting nucleation have been published. Falcon and Berglund, 2004 used PCA analysis to track subtle changes in the Raman spectra which could have been used as a warning for the onset of crystallization. They found that by monitoring the loading values of PCA analysis, the systematic changes in a certain spectral variable indicated the upcoming nucleation. Vibrational spectra contain peaks that are characteristics of some of the intermolecular interactions, e.g., hydrogen bondings, thus the molecular clustering could possibly be monitored.

The data obtained from the ATR-FTIR or Raman measurement together with multivariate methods could be used for evaluation of the batch-to-batch variations of crystallization process. Such studies are seldom presented, however. Kourti et al., (1995), used multiblock and multiway PLS methods to analyze batch-to-batch variations. Falcon and Berglund, 2004 used PCA analysis to study batch-to-batch variations in antisolvent crystallization processes. They propose, that the score plots of the PCA analysis were well suited for detecting differences in the batches.

5.1.2 Solid phase

The solid phase measurements include different nucleation processes, agglomeration, growth, breakage and attrition. Ultrasonic spectroscopy has been applied for the on-line measurement of particle size (Mougin et al., 2003). Laser diffraction techniques have been applied for in-situ monitoring of particle size during a crystallization process (Yamamoto et al., 2002; Ma et al., 2000; Ma et al., 2001; Qu et al., 2004). High-speed on-line digital imaging sensors are becoming more popular in the investigation of particle size and especially particle shape (Barret, 2003; Barret and Glennon, 2002; Calderon De Anda et al., 2005a; 2005b, 2005c). UV measurement has been applied for at-line monitoring of the particle size with mean size of 200 nm (Taavitsainen, 2001). The discussion of these techniques are beyond the scope of this thesis, but the successful simultaneous application of solid phase measurement and solution concentration monitoring would provide complete monitoring scheme of the crystallization process. However, solid phase measurement techniques still have many challenges to overcome

before they can be considered reliable and accurate enough techniques for real-time monitoring of solid phase monitoring in crystallization process (Yamamoto et al., 2002; Ma et al., 2000; Ma et al., 2001; Calderon De Anda, 2005). These challenges are currently gaining a lot of attention and should be considered more deeply in the future.

5.2 Off-line characterization of the product

5.2.1 Size and shape

In addition to the in time analysis of the process, also the end product should be characterized. The characterization of the product includes the size characterization, which traditionally has been done by sieving or laser diffraction based techniques. Several applications of image analysis techniques for characterizing crystalline solids has been recently reported, for example, by Lewiner et al. 2001a, 2001b, 2002; Bernard-Michel et al., 2002; Pons et al., 2002; Ålander et al., 2003; Faria et al., 2003; Hurley et al. 2005; Ferreira et al. 2005; Pöllänen et al., 2006. Image analysis is the only way to define size and shape simultaneously. Manual image analysis is time consuming, and its drawback is that the operator bias can affect considerably on the result. The application of automated image analysis provides an opportunity for relatively rapid analysis and eliminates the operator bias. (Pöllänen et al., 2006).

5.2.2 Polymorphic form of the product

The polymorphic form and purity of the product is an important factor expecially in the production of pharmaceuticals. Traditional methods applied are X-Ray diffraction (XRD), thermal gravimetry (TG), differential scanning calorimetry (DSC) and vibrational spectroscopic techniques, e.g., Raman, near infra red (NIR) and FTIR. From these techniques, XRD alone measure truly the structural information, which is therefore the only method which directly measures the fact whether the samples are structurally different (Brittain, 1991). Other techniques can be considered as an additional method to characterize the composition of the product. Vibrational spectroscopy provides a tool for rapid screening of the polymorphic composition, and also the possibility of quantitatively characterizing the product composition, when multivariate calibration methods are included. In addition NIR and FTIR devices are economical to purchase and to use and available in almost all of the laboratories. The drawbacks of vibrational spectroscopy in solid state characterization include the differences in the spectra caused by the size distribution of the product. The DRIFT-IR, ATR-FTIR, NIR and Raman methods together with PCR and PLS calibration as well as lazy learning algorithms and neural networks have been succesfully used in quantitative characterization of the polymorphic composition of powder samples (Kipouros et al, 2005; Agatonovic-Kustrin et al., 2001; Salari et al, 1998 Helmy et al., 2003; Blanco et al., 2000; Auer et al. 2003).

6 SPECTROSCOPIC METHODS

6.1 Fourier transform infra red (FTIR) spectroscopy

Fourier transform spectrometers are mainly based on the Michelson interferometer (Griffiths, 1986). The broadband source transmits the radiation to the beam splitter, which sends part of the radiation to a fixed mirror and other part to a movable mirror (Figure 5). The movable mirror moves in the direction perpendicular to its plane (Griffiths, 1986). The beams reflect from the mirrors and return to the beam splitter. The path length difference between the fixed mirror and moving mirror causes interference between the returning beams, as these beams split into two the first half to a detector and the another half to a light source.

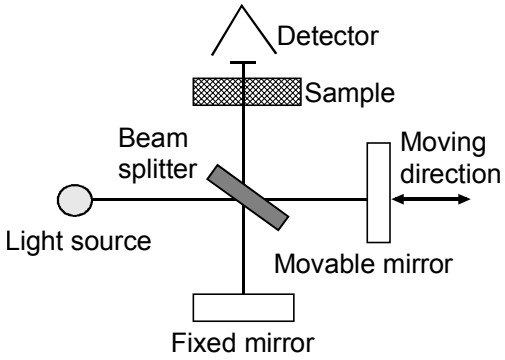

Figure 5 The optical configuration of the Michelson interferometer (modified from the Griffiths, 1986; Workman, 1997)

The intensity of those beams depends on the path length difference, which gives the spectral information in the form of an interferogram (Griffiths, 1986; Workman, 1997). The interferogram is presented in the form of intensity as a function of time or position of the moving mirror. The interferogram is then turned into a spectrum using Fourier transformation. After Fourier transformation, the spectrum intensity is presented as a function of wave number or wavelength (Griffiths, 1986).

The radiation can be transmitted thorough the sample to the detector or it can be reflected from the sample and this reflected light is then transmitted to the detector. The used method depends on the sample/process type, e.g. for a translucent samples transmission methods are suitable whilst for opaque samples reflection methods apply the best. Different accessories are manufactured for both transmission and reflection measurements for different sample types. Two reflection methods applied in experimental part of this thesis are considered in the following two chapters.

IR spectroscopy measures the amplitude of the molecular vibrations, which are caused by the change in the dipole moment of the molecule. Several different types of molecular vibrations exist. The two main categories include stretching and bending vibrations, which form different subgroups. In addition to this, vibrational coupling can exist, which means interactions between vibrations for example those for different bonds attached to the same atom. (Colthup et al., 1975) Vibrational spectra shows not only molecular structure and functional groups but also inter- and intramolecular features such as hydrogen bonding between molecules as these forces influence on molecular bonds (Wülfert et al., 1998)

The physical changes in the sample and its environment, i.e., pressure, temperature and viscosity changes cause also variations to the vibrational spectra. Intermolecular forces, e.g., hydrogen bonding, is affected by the physical conditions such as temperature and pressure and can cause, therefore, changes in the vibrational spectra (Wülfert et al.; 1998). The temperature changes influence in the spectrum especially to the bands from the functional groups that contain –H bonding (Czarnecki et al., 1994; Ozaki et al., 1997; Finch and Lippincott, 1956; Finch and Lippincott, 1957; DeBraekeleer, 1998; Noda et al., 1995; Wülfert et al., 1998). For example, for the stretch mode of hydrogen bonded –OH groups, raising the temperature decreases the average cluster size and relative absorbance of the free groups (Noda et al., 1995; Wülfert et al., 1998). In addition to this, the peak shifts and broadening of the bands occur especially for the functional groups with H-bonding (Wülfert et al., 2000).

There is always noise present in the spectroscopic measurements. In the spectroscopic measurements the level of noise, termed as a signal-to-noise ratio (SNR) is related to the spectrometer and measurement parameters: measurement time, resolution, throughput, and mirror velocity and detector size.

To optimize the quality of the spectrum and the robustness of the measurement, requires a compromise between these parameters, and in practice, the parameters that are usually changed are measurement time, resolution and throughput (Griffiths, 1986).

The SNR value of a spectrum measured with a certain resolution is increased proportional to the square root of measurement time (Griffiths, 1986). Measurement time depends on the number of consecutively measured averaged spectra. With slow scanning interferometers, also the scan speed is variable (Griffiths, 1986). The number of spectra averaged is always the compromise between the robustness of the measurements and the quality of the spectra. If the off-line samples are measured for example in the quality control of bulk powder the number of averaging scans are defined by obtaining reasonably high quality spectrum in a reasonable time, i.e., increasing the number of averaging scans does not significantly improve the SNR. If the

reaction kinetics is measured, the measurement of one spectrum should not take longer than the changes in the process kinetics. If too many averaging scans are included, the acquisition time of a spectrum is too slow to observe instantaneous changes in the system. In addition, resulting averaged spectrum is not representing the one instantaneous moment in a system but the averaged spectrum includes all the changes present in the process during the time the scans of this spectrum are measured. Therefore, in kinetics measurements the robustness of the measurements is a key issue.

Resolution means the capability of a spectrometer to separate bands that are some distance from each other (Workman, 1997). Resolution used in the measurements can be expressed as the increment of wave numbers (cm^{-1}) $\Delta\nu$ on which the resolution is measured, i.e., smaller increments cause better separation of bands close to each other and, thus, the better resolution. An improvement in the resolution increases the measurement time, since retardation is slower with smaller wave number increments. In addition, SNR is halved when the measurement increment is halved. This means that in order to obtain same SNR with higher resolution the number of scans has to be increased, which again increases the measurement time needed. Increasing resolution leads to a better separation of lines (Griffiths, 1986). The rule of thumb says that one should select the lowest possible resolution that is enough to separate the relevant bands. This ensures the most robust measurement possible.

Throughput is defined as the power received at a detector through an optical system (Griffiths, 1986), in terms of effectiveness of an optical system to transmit light relative to amount of light introduced to the system (Workman, 1986).

IR regions include the *n*ear-IR (NIR) 12800-4000 cm^{-1} (700-2500 nm), mid-IR 4000-200 cm^{-1} (2500-$5\cdot10^4$ nm) and very seldom-used far-IR region 200-10 cm^{-1} ($5\cdot10^4$-10^6 nm). In this study, the mid-IR region is used. The advantage and sometimes the disadvantage of the mid-IR region is that almost everything, all of the organics and many inorganic compounds absorb within the mid-IR region. IR spectroscopy is widely usable but the obtained spectrum from mixtures can be rather complex. In addition, the bands obtained can include interactions between different species or the peaks from different constituents can overlap each other. Some strong absorbent can totally cover up the smaller peaks. Therefore, the interpretation of the IR-spectrum or qualitative or quantitative analyses on samples from IR spectrum is not a straightforward task.

The basis of the quantitative analysis in spectroscopic measurements is Beer's law

$$Abs = -\log\left(\frac{I}{I_0}\right) = -\log(Tr) = \varepsilon c l \tag{4}$$

where *Abs* is a measured absorbance I_0 is the intensity of incident energy transmitted, I is the intensity of transmitted light, Tr is the transmittance, ε is the molar absorptivity, c is the concentration and l is the path length. Assumption is that the relationship between absorbance and concentration is linear (Workman, 1997). Ideally, the height of a certain peak in the absorption spectrum (absorption in y-axis) related to the constituent in interest could be linearly related to the concentration of this constituent. However, this is usually complex task when measuring the mixtures of different species due to aspects already discussed in the previous paragraphs: random noise can cause variation to the spectrum, small peak shifts can occur due to different mechanical and physical reasons, changes in the physical environment, e.g., temperature, or pressure changes while measuring. When measuring mixtures of complex species, the interactions between the species cause interaction peaks and the peaks from different species can overlap each other, and therefore, the absorbance of a certain peak is not always due to the constituent of interest.

6.2 Attenuated total reflection (ATR)

ATR accessories are widely used in FTIR spectrophotometers especially in the measurement of liquids and thin films. ATR is also called Internal Reflectance (Workman, 1997). The principle of ATR spectroscopy is illustrated in Figure 6.

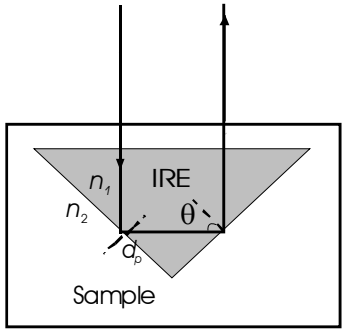

Figure 6 The principle of attenuated total reflection IRE=Internal reflecting element, n_1=refractive indice of the ATR-element, n_2=refractive indice of the sample, dp=depth of the penetration of the beam into the sample, θ= angle of incidence of the reflected and propagating radiation

The radiation from the light source is transmitted thorough an internal reflecting element (IRE) at a certain angle of incidence (θ) that causes the light to undergo internal reflection at the interface of the IRE and sample. Total reflection occurs at the transition from an optically dense IRE element with high refractive index (n_1) and optically less dense sample with much lower refractive index (n_2) than the IRE element. The beam is penetrated into the sample by fixed depth (d_p). The internally reflected radiation is attenuated by the absorption of the sample and the reflected light is a function of this absorption at a certain wavelength. As a result the

reflected light that is transmitted to the detector reveals an absorption spectrum. (Workman, 1997)

The depth of the penetration (d_p) of the radiation into the sample at constant wavelength (λ) and angle can be calculated by

$$d_p = \frac{\lambda}{2\pi\, n_1 \left(\sin^2\theta - \left(n_2/n_1\right)^2\right)^{1/2}} \tag{5}$$

(Harrick, 1979). Intensity in the different wavelengths changes as the radiation penetrates to different depths (Workman, 1997). For typical samples the order of approximate depth of penetration is in one or two micrometers, thus ATR can be considered essentially as surface measurement technique in a macro scale measurements (Workman, 1997). In addition, as the radiation does not penetrate deep into the sample, also the samples containing strong IR absorbents, such as water, can be measured using the ATR technique (Workman, 1997). The sample and the IRE has to be in inherent contact in order for the measurement to be representative. Any contamination in the surface of the element causes the measurement to be taken from this contaminate, not the sample. In addition to this, any scratches or the chemical deterioration on the IRE cause the measurement to be inappropriate and any mechanical distraction can cause misalignment of the element, which further cause absorptivity decrease due to losses in the energy transmitted thorough the IRE (Workman, 1997).

The radiation can undergo one or multiple reflections. The number of reflections that occur depends on the geometry of the IRE (Workman, 1997). The conical element presented in Figure 6, causes the beam to undergo two internal reflections before being passed to a detector. The number of reflections is meaningful, because of the fact that the surface covered by sampling is a product of the number of reflections and the depth of penetration (Workman, 1997). Therefore, two reflections can be considered to cover twice the area as one reflection.

The conical type element shown in Figure 6 is ideal for mounting to a probe and placing in a reactor, and this can therefore be applied for *in-situ* monitoring. There are also several other types of element geometries that can be mounted to probes, for those references the reader is asked to look for instructions from probe manufacturers. Selection of the geometry depends on the application: circle or tunnel cells are widely applied for volatile liquids, constant pressure horizontal ATR-elements are typically used for thin films etc. (e.g. Workman, 1997; Harrick, 1979; Griffiths and de Haseth, 1986). With some IRE geometries, the angle of incidence can be changed, and this property be used to have measurements from different depths of the sample (Workman, 1997). For the element geometry represented in Figure 6 the angle if incidence is fixed to 45°.

There are several different IRE materials available commercially, which have different chemical and mechanical resistancies and different samples/process conditions require different IRE material (Table 1). In addition, these materials have different usable spectral ranges as can be seen in Table 2.

Table 2 Properties of the most common IREs used in immersion probes ([1]Workman, 1997; [2]Axiom Analytical, 2005)

Material	[1]Spectral range (cm^{-1})	[2]Hardness (KNOOP)	[1]Refractive Index at 1000 cm^{-1}	[2]Special considerations
Zinc sulfide (ZnS)	17 000-833	250	2.2	Sensitive to some acids and strong oxidizers
Zinc selenide (ZnSe)	20 000-460	120	2.4	Sensitive to acids and oxidizers
AMTIR (Infrared glass made from germanium, arsenic and selenium)	11 000-625	170	2.5	Sensitive to strong alkalis
Silicon	8 300-660	1500	3.4	Sensitive to HF, HNO_3, NaOH
Germanium (Ge)	5 500-600	780	4.0	Sensitive to hot sulphuric acid

Therefore, the selection of the IRE material is always a compromise between a desired measurement range, the chemical and mechanical conditions where IRE is used and the price of the IRE. The properties presented in Table 1 are of course estimates and source dependent, and cannot be considered as exact values. In addition to the use of these materials alone, there are applications of IRE, where the element itself is produced from one material which is then coated with another, more resistant one, usually with diamond to enhance chemical and mechanical resistance.

6.3 Diffuse reflectance Fourier transform infra red (DRIFT-IR) spectroscopy

In the DRIFT-IR technique the light transmitted is imposed on the surface of the sample. From that surface the light is reflected over a range of angles and part of the radiation is transmitted into the sample or reflected from the sample. Reflected light has two components: specular and diffuse. The diffuse component consists basically of the chemical information. “The diffuse reflected light emergs from random reflections, refractions and scatter inside the sample. Further, more transmission and absorption may occur at the other interfaces” (Workman, 1997). Together these phenomena cause the spectral composition of the reflected radiation transmitted to the detector. As a result reflection spectrum is obtained.

The particle size can affect the reflectance signal. Reduction of particle size by grinding may improve the quality of the spectra. In addition to this, a narrow size distribution and equal shape

is beneficial. Also the properties of the surface affect on the obtained result. If there is strong surface reflections present, the sample might have to be diluted with IR-radiation nonabsorbing media such as KBr (Workman, 1997) to reduce the reflection intensity.

The spectrum obtained from a diffuse reflectance measurement does not have linear relationship between the reflected intensity and composition, i.e., Beer's law does not apply (Workman, 1997). To correct the diffuse reflectance spectrum and to obtain the linearized spectrum the Kubelka-Munk function is used (Workman, 1997; Spragg, 1998; Clark and Roush, 1984)

$$f(R) = \frac{(1-R)^2}{2R} \tag{6}$$

where $f(R)$ is the corrected spectrum R is the ratio of the sample diffuse reflectance spectrum and a reference sample that is non-absorbent, e.g., KBr in mid-IR region (Workman, 1997). Alternatively the spectrum is linearized using $\log(1/R)$, which corresponds to the absorption units (relationship similar to that between the transmission and absorption) (Settle, 1997; Clark and Roush, 1984). However, it is hard to distinguish based on literary which method is preferable and why. Both techniques are reported not to perform true linearization of the data. K-M amplifies strong bands and it is said to be suitable for cases where the absorption level is relatively low and low specular reflectivity is sufficient (Garbassi et al., 1993; Tremblay and Gagné, 2002). Kubelka-Munk has successfully been used in several applications (Tremblay and Gagné, 2002; Greene et al., 2004; Park et al., 2000). Also the R units have been reported to be succesfully used (Kipouros et al., 2005). The unit $\log(1/R)$ has widely been used in several applications (Agatonovic-Kustrin et al., 2001; Blanco et al., 2000; Pöllänen et al. 2005a; Pöllänen et al. 2005c;Andrés and Bona, 2005).

7 MULTIVARIATE TOOLS TO EXTRACT INFORMATION FROM SPECTRAL DATA

The measured data, e.g., spectra from the process or analytical samples form the data matrix of explanatory variables **X** by size is $n \times m$ where n is the number of objects or samples and m is the number of measured variables. The matrix of the response variables is called **Y**, and its size is $n \times q$ where q is the number of response variables. If $q=1$ **Y** is a vector and is then denoted as **y**. **X** data can be explored by itself or correlations between **X** and **Y** can be explored and models for predicting **Y** from **X** can be developed using the multivariate methods.

There are number of different ways to look at spectral data and the application to be used depends strongly on the purpose of the analysis:

- Seeking the similarities and differences between the samples, e.g., use of clustering methods in purity analysis.
- Process faults and drifts measured in time.
- Prediction of a certain property, e.g., the concentration of the constituent in interest.
- Batch-to-batch variation analysis.

To obtain reliable results by multivariate modeling, several steps are required: Detection of the quality of the samples and removing possible outliers, data preprocessing, model building and model validation. In the following chapters, some commonly used multivariate methods for data preprocessing, classifying samples, statistical process control, multivariate calibration and also batch-to-batch variation analysis are introduced. In addition to this, methods for data quality analysis and outlier detection, as well as model validation are briefly discussed.

7.1 Preprocessing methods

7.1.1 Mean centering

Mean centering refers to an operation where the average value for each variable is calculated and these average values are subtracted from each data object. Thus the variance of all the variables have a mean value of zero. Mean centering is done practically in every multivariate analysis.

7.1.2 Scaling

Sometimes the variables exhibit very different numerical ranges. In multivariate methods such as in principal component analysis (PCA) and PLS methods the results are not independent for the units of variables. Therefore, the variables with high numerical values and variance can exhibit a too strong weights in the model whereas the weight of the variables with very small numerical values and variance can be underestimated. Therefore, the variables are often scaled

to unit variance. For each variable, the standard deviation is calculated and the inverse standard deviation for each variable is used to multiply each column, i.e., the variable vector in the **X** matrix. After that treatment each scaled variable has equal unit variance.

Scaling of the spectral data cause the bands from specific constituents and the baseline to be equally weighted, which further results in an unnecessary increase in the noise level. The most meaningful variables, i.e., spectral bands related to modeled phenomena actually should have larger meaning and larger scale variation than do the baseline points. Therefore, scaling all the variables to equal meaning might not be beneficial in this case.

7.1.3 Variable selection

The data used in modeling defines the need of variable selection and the way how it should be performed. All the variables are not automatically equally important in modeling. For instance, the bands in the spectrum that are due to a different phenomena than the modeled ones, can actually be harmful in the modeling since these cannot be related to the modeled phenomenon. In addition, including the baseline does not necessarily give additional value to the modeling procedure. Therefore, it may be beneficial to select certain relevant variables to the model and exclude the variables that do not improve the model (Brereton, 2002). The selection of variables from the spectral data should be based on priori knowledge of the variables. When the spectral data is used in the modeling, the important issue is: what spectral regions reflect the modeled phenomena. However, reducing the number of variables too drastically can cause problems in terms of reduced dimensions in the multivariate matrices (Brereton, 2002).

The in addition the type of the data also the purpose of the modeling define the need and the level of variable selection. If the objective of the model is to obtain the best possible momentary predictions, removing of the variables not related to modeled variable might be beneficial. However, for process monitoring purposes it might be advantageous to include as many of the **X** variables as possible in the modeling. For example, the occurrence of the new phenomena in process can occur in different variables than the usual process variation.

Several methods are proposed to find the optimal number of variables to be included into the model. Different filters to remove irrelevant variables are presented, e.g., in Brereton, 2002 and Höskuldsson, 1996. The variables can be put into order based on their mean, standard deviation, or variance. (Brereton, 2002) The selection of the variables can also be based on an increase in the correlation coefficient (R^2) value or the variable set which gives the best predictions of the calibration set, i.e., the smallest errors in the predictions of the calibration and/or test sets are obtained, which can be written $\min \Sigma \left| y_i - \hat{y}_i \right|$. In addition, e.g., cross validation, squared covariance and H-error criteria can be used in the selection of variables (Höskuldsson, 1996).

7.1.4 Multiplicative scatter correction (MSC)

The MSC method attempts to remove the effects of scattering and most of the random variance. In this method each spectrum is linearized to the "ideal" spectrum of the sample (Workman, 1998). "Ideal" spectrum is calculated as an average spectrum of all samples. Each spectrum is then corrected by first calculating a linear regression against the corresponding points in the ideal spectrum. Then the slope from linear regression is subtracted from the original spectrum, and the result of the subtraction is divided with the intercept from the linear regression (Workman, 1998).

The MSC is only applicable when spectral responses are at least fairly linearly dependent on the concentration of the constituents of interest. However, if there is a large variability in the sample composition and thus in resulting spectra, the mean spectrum does not represent the ideal spectrum. In this case, the MSC filtering can worsen the spectra to be used in further modeling purposes (Workman, 1998).

7.1.5 Standard normal variate (SNV)

The SNV method also tries to remove the different scattering effects. In SNV method each spectrum is normalized by the standard deviation of the responses within the whole spectral range (Workman, 1998). The SNV correction is often followed by detrending, which attempts to remove the baseline shifts and different trends in the baseline. These effects are typical especially for diffuse reflectance spectra (Workman, 1998). A linear least squares regression is used to fit a polynomial to the spectral responses and the curve is subtracted from the spectrum (Workman, 1998).

The method is applicable to the spcetra which responses are linearly dependent on concentration. However, SNV is more widely applicable and more variation in the data can be present than in MSC method, as each spectrum is corrected independently not by using the average of all spectra (Workman, 1998).

7.1.6 Orthogonal signal correction (OSC)

The MSC and SNV methods may remove also information relevant to the predicted variable; hence, those are applied to the **X** matrix. These methods thus ignore the variation within **y**. Wold et al. (1998) introduced the orthogonal signal correction (OSC) filters, which are based on the principle that variation, which is unrelated to the predicted variable is eliminated from the data matrix by ensuring that the removed data is mathematically orthogonal to predicted variable. In addition, the vector to be removed, i.e., OSC component should be of maximal size, which causes that the maximal amount of **y** independent variation is removed with one OSC component. Single component OSC model is derived as follows

$$\mathbf{X} = \mathbf{t}_{\mathbf{OSC}}\mathbf{p}'_{\mathbf{OSC}} + \mathbf{E}_{OSC}\ , \quad \mathbf{t}_{\mathbf{OSC}} = \mathbf{X}\mathbf{w}_{\mathbf{OSC}} \perp \mathbf{y}, \quad \left\|\mathbf{x}_{\mathbf{OSC}}\right\| = 1 \tag{7}$$

where $\mathbf{t}_{\mathbf{OSC}}$ (n×1) , $\mathbf{p}_{\mathbf{OSC}}$ (1×m) and $\mathbf{w}_{\mathbf{OSC}}$ (1×m) are score, weight and loading vectors of the OSC component as are the PLS model components defined later. In the OSC model the score vectors ($\mathbf{t}_{\mathbf{OSC}}$) are orthogonal to **y**. Matrix $\mathbf{E}_{\mathbf{OSC}}$ is one OSC component filtered data matrix. For more OSC components, filtering is applied to $\mathbf{E}_{\mathbf{OSC}}$**.** Eventually $\mathbf{E}_{\mathbf{OSC}}$ is the filtered matrix to be used, e.g., in the PLS calibration.

Several authors have derived OSC methods that differ from each other by the way the orthogonal score vectors are found, and how the orthogonality between score vector and **y** is ensured (Andersson, 1999; Fearn, 2000; Höskuldsson, 2001; Feudale et al., 2002; Trygg and Wold, 2002; Trygg, 2002; Westerhuis, 2001). Different OSC approaches do not necessarily lead to the same result. Developed OSC filters can be divided in two main groups: In indirect approach the score vector orthogonal to **Y** is found by an internal iteration procedure while in direct approach the **Y** orthogonal matrix is directly derived from the data matrices.

The filtering should enhance the modeling performance. In predictive models, this usually means that the model should be simpler and better predictions should be obtained. However, it has been argued that usually equal total number of components both in PLS and OSC-PLS are needed in the optimal model. In addition, it has been presented, that the most optimal form of OSC would not give any better predictions than does the regular PLS when same total number of components is used (e.g., Goicoechea and. Olivieri, 2001). It is also possible to investigate the nature of the **Y** orthogonal components by OSC and possibly evaluate where this variation arises.

7.2 Principal component analysis (PCA)

PCA is a one way to reduce the dimensions of the data. In PCA, the data matrix **X** of size $n \times m$ is decomposed as follows

$$\mathbf{X} = \mathbf{TP}' + \mathbf{E} \tag{8}$$

where **T** (size $n \times A$) is the score matrix, **P** is the loading matrix (size $A \times m$) and **E** is the residual matrix of size $n \times m$. A refers to a number of principal components extracted from the data matrix.

The first principal component is the linear projection of original (preprocessed) variables onto line which maximizes the covariance in the data. The second principal component is similarly derived from the data, which is left after extracting the first component. The principal components are orthogonal to each other thus they are independent. The detailed description of the PCA is presented by, e.g., Wold, (1987), Martens and Naes, (1993) and Jackson, (1991).

The scores (**T**) describe the samples and loadings (**P**) describe the variables in the derived decomposition. Ideally the score and loading matrices contain all linear systematic variation present from the data and the random variation, noise, is left in the residual matrix. The scores are usually plotted against the sample number or one component scores against another component scores (scatter plot) in order to evaluate the quality of the samples within the data. The samples can form clusters in the scatter plots, where the samples that can be considered similar are located in one cluster. The similar plots are derived from the loadings in order to identify the original variables causing the grouping and clustering found by using the scores (Vandeginste et al. 1998; Eriksson et al. 2001).

The number of components selected in a PCA model depends on the purpose, where the model is to be used. The first two or three principal components extracted from the data contain usually most of the systematic variation present in the data. Therefore, the scores and loadings of the first two or three components of the model are often used for interpretation of the main phenomena in the data. For process monitoring purposes, one tries to select the number of components of the PCA model so that the linear systematic variation would be included in the model and the random variation would form the residual matrix. If too few components are extracted, the residual matrix usually contains also systematic variation and correspondingly extracting too many components causes random variation to be included in the last components scores and loadings of the derived model. There are several different methods to select the correct number of principal components in the model. One possible method to select the number of principal components is a cross-validation (CV), which is more closely discussed in Chapter 7.7.

7.3 Multivariate statistical process control (MSPC)

Typically, in process monitoring, the variables should be monitored over time. Traditionally each variable is monitored separately over a time by univariate process monitoring. In many circumstances, however, the number of measured variables is too large to be successfully visualized with univariate methods. In addition, the measured variables are seldom independent from each other (Kourti and MacGregor, 1995). There are typically only a few driving phenomena present in the process, and all the measurements have different ways of representing these phenomena (Kourti and MacGregor, 1995). The spectral data is the special case of this type of multivariate measurement: Single measurement results in hundreds sometimes even thousands of correlated variables. Multivariate methods are capable of treating all of the data simultaneously and look at how all the variables are behaving relative to each other.

MSPC charts are based on the PCA models. Kresta et al. 1991 presented the principle of these techniques. Typically, Hotelling's T^2 statistics calculated from the scores of a PCA model is used:

$$T^2 = \sum_{a=1}^{A} \frac{\mathbf{t}_a^2}{s_{t_a}^2} \tag{9}$$

where $s_{t_a}^2$ is the estimated variance of the score vector $\mathbf{t}_a$. T^2 is a distance from the multivariate mean to the operating point on the PC plane (Wise and Gallagher, 1996), and will detect whether or not the variation in the quality variables in the plane of the first A PCs is greater than can be explained by common cause (Kourti and MacGregor, 1995 and MacGregor and Kourti, 1995).

If a very new type of event occurs during the process, which is not present in the reference data set used for building the PCA model, the new sample will be out of plane (Kourti and MacGregor, 1995). To detect this type of variation the Q statistics can be calculated (Jackson, 1991)

$$Q = (\mathbf{x} - \hat{\mathbf{x}})'(\mathbf{x} - \hat{\mathbf{x}}) \quad \text{where } \hat{\mathbf{x}} = \sum_{a=1}^{A} \mathbf{t_a p_a} \tag{10}$$

Q statistics refers also to the squared prediction errors of the calibration set (SPE_x), because it practically is the squared perpendicular distance of a multivariate observation from the projection space. Q statistics is a measure of the variation in the data not included in the model.

If the new phenomenon is appearing to the system, this should be able to detect using T^2 and Q statistics. It is possible to detect the gradual drifting of the process slowly from a certain level to from T^2 and Q statistics. To obtain a proper process-monitoring scheme, both statistics should be included simultaneously. (Kourti and MacGregor, 1995 and MacGregor and Kourti, 1995)

By T^2 and Q statistics, abnormal samples can be detected, but it is often very important to distinguish what changes in the measured data and more precisely what variables cause the samples to exceed the confidence limits. MacGregor et al. (1994) developed contribution plots for visualization of the variables in the process-monitoring scheme.

The contributions of each process variables to the T^2 statistics of the measured samples are

$$\mathbf{Contr}_{\mathbf{T^2}} = \mathbf{t}\sqrt{\mathbf{S}^{-1}}\mathbf{P}' \tag{11}$$

In which the matrix **S** is a diagonal matrix equal to the eigenvalues of **X** and it normalizes the score values. The contributions of each process variable to the Q statistics of measured samples can be calculated as

$$\mathbf{Contr_Q} = (\hat{\mathbf{X}} - \mathbf{X}). \quad (12)$$

To decide whether the sample is an abnormal sample in either T^2 or Q statistics or which are the acceptable limits for the variation of the variable contributions, and to have, e.g., a limit for the alarm for abnormal sample, the confidence limits can be calculated. There are several different approaches for calculating the confidence limits. It depends on the dataset which confidence limits should be used. Calculation of the confidence limits based on different types of distributions and when certain confidence limits are applicable are presented in the literature, e.g., in Massart et al., (1997).

According to Massart et al., (1997), sample size n is large (n>30), the standard deviation s can be estimated by

$$s = \sqrt{\sum_i (x_i - \bar{x})/(n-1)} \quad (13)$$

where x_i and $\bar{x}$ are descriptor variable i and mean of the descriptor variables, respectively. For large sample sizes, the confidence intervals can be estimated based on normal distribution

$$\mu = \bar{x} \pm d \cdot \frac{s}{\sqrt{n}} \quad (14)$$

where d is the coefficient for the selected confidence limit level, which for standardized normal distribution is, e.g., 1.645 and 1.96 for 90% and 95% confidence limits, respectively.

As the new samples are projected onto the PCA model of the reference samples, the T^2 and Q statistics for the new samples are calculated and those values compared to calculated confidence limits, the sample quality can be evaluated.

Confidence levels for the variable contributions can also be calculated using Eq. 13. When the variable contributions of an abnormal sample are reflected to the confidence limits of the contributions, the variables which cause this particular sample to be abnormal by the T^2 or Q statistics exceed the confidence limits. Thus, the reasons which cause the sample to have been found to be an abnormal one can be detected and a detailed evaluation on the reasons for different faults can be done. This illustration is very practical in spectral data analysis. For example, if there appears a clearly new band in the spectrum during the process, the position of the band will give an indication on the chemical nature of the new phenomenon appearing in the process, and as a consequence the possible contaminant or undesired product can be possibly detected.

The higher the confidence limit the less sensitive is the system alarming the fault or abnormality, but on the other hand, a too sensitive system may lead to unacceptable number of false alarms can be detected (Ramaker et al., 2004). To avoid a false alarm, the correct confidence level should be found and both Q and T^2 statistics applied simultaneously. The number of components selected in the PCA model and the size of the reference set, affect on the performance of the MSPC in terms of false alarms. Ramaker et al. (2004) found that overfitting causes a greater false alarm rate, but underfitting does not have such problems, if the size of the reference data set is sufficient. A too small dataset causes more false alarms than a large dataset as the model stability is lower (Ramaker et al., 2004). The size of a proper dataset is case dependent, however. To find the correct number of components included in the model and to ensure the stability of the model in terms of sufficient amount of data used in the modelig, some validation procedure, e.g., cross validation procedures should be included.

MSPC charts can also be used in outlier detection. The samples exceeding the confidence limits can be considered to represent something different from rest of the samples in the set. The decision whether the sample is an extreme sample representing true variation present and which should actually be included in the model or totally an outlier can be done by looking at the contribution charts. In batch processes, **Y** state is often not available during the batch. In these circumstances, MSPC analysis is only performed in the **X**-space. In this case MSPC does not give an indication about the samples which might have been measured inaccurately, i.e., the problems in the, e.g., calibration model in the **XY**-space cannot be detected. The sensitivity analysis techniques applied for the derived model described in Chapter 7.6 serve this purpose. However, if the Y-space is measured throughout the batch, MSPC charts for Y-space can be derived and used for outlier detection.

7.4 Soft independent modeling of class analogy (SIMCA)

The qualitative analysis of the data involves the purity analysis of the samples. From spectral data it can often be difficult to see whether or not one particular sample is significantly different from an other sample and if this sample is acceptable or not. This is because there is always noise present in the spectra and the differences in the spectra due to, e.g., a small amount of contaminate can be significant. Therefore, statistical methods where the ranges for accepted sample quality are defined and new, unknown samples are reflected can be used for fast quality evaluation of samples.

Soft independent modeling of class analogy (SIMCA) is one of the earliest classification methods proposed in chemometrics by Wold (1976), and this method one of the methods which can be used for qualitative analysis of the samples. The first step of SIMCA analysis is to build

a PCA model and decide on the number of components needed to describe the structure of the data. Each class has a separate principal component model. There are several different statistical criteria which can be used to define the limit for the class. Sharaf et al., (1986) applied the *F*-statistics for calculating the confidence limits for each class from the standard deviations of the samples. Also the T^2 and Q statistics and the confidence limits for these, which were introduced in Chapter 7.3, can be used in defining the limits for the classes.

One of the traditional ways to illustrate the SIMCA model is so called Coomans plot, first introduced by Coomans et al., (1984). In the Coomans plot the confidence limits of two different classes are plotted in the same figure. If the sample lies within the confidence limit of either class this sample belongs to that class. If the specific sample is within the confidence limits of both of the classes it belongs to both classes. If the sample exceeds both of the confidence limits, it is not part of either of modeled class but it is an outlier. The correct number of components selected to the model is a critical issue for obtaining reliable classification of the samples.

7.5 Partial least squares (PLS)

Measured data can be used to predict a quantitative amount of a certain property, e.g., the concentration of the constituent of interest from spectral data. Traditional regression methods should not be applied to correlated variables, which the spectral variables typically represent. In complex chemical systems, the bands in the IR spectrum from different constituents often overlap each other, the absorbencies of specific compounds of interest can be low and, consequently, no single peak can be found to correlate reliably with the concentration. In addition, random variation is always present in the spectra, and for this reason a single peak signal to noise ratio can be extremely low and this can provide inaccurate estimates of the concentration level. Multivariate methods, e.g., PLS regression and PCR calibration models can be applied to overcome these problems. PLS enables the linear modeling of the correlated variables. The collinearity between variables stabilizes the model which can be considered an advantage rather than a problem (Wold et al., 1983). Typically the use of the multivariate methods can lower the noise level (Kourti and MacGregor, 1995).

PLS is a method where the relationship between **Y** and **X** data matrices is derived by a linear multivariate model (Höskuldsson, 1996; Wold et al., 1983). The PLS model applications include a multivariate calibration, process monitoring and optimization and quantitative structure activity analysis (QSAR). (Eriksson et al., 2001) There are two different types of PLS algorithms: PLS1 for the system where there is only one **Y** variable, i.e., the **Y** matrix is actually a vector and PLS2 for the system where there are two or more **Y** variables and these are

modeled simultaneously with one model. Generally, separate PLS1 models for each **Y** variable give better predictions for both calibration and test sets than does one PLS2 model for all **Y** variables unless the **Y** variables are strongly correlated (Höskuldsson, 1996). Algorithms for the two different PLS versions are presented in, e.g., Martens and Naes, (1989).

In PLS modeling the predictor variable matrix **X** and the response variable matrix **Y** are decomposed into the structure and noise parts by latent variables as follows:

$$\begin{aligned} \mathbf{X} &= \mathbf{t}_1\mathbf{p}_1 + \ldots + \mathbf{t}_A\mathbf{p}_A + \mathbf{E}_{A+1} = \mathbf{TP'}+\mathbf{E} \\ \mathbf{Y} &= \mathbf{u}_1\mathbf{q}_1 + \ldots + \mathbf{u}_A\mathbf{q}_A + \mathbf{F}_{A+1} = \mathbf{UQ'}+\mathbf{F} \end{aligned} \qquad (15), (16)$$

where $\mathbf{p}_a$ and $\mathbf{q}_a$ are the loading vectors, $\mathbf{t}_a$ and $\mathbf{u}_a$ are the scores for the component a and matrices **E** and **F** are the residual terms of the **X** and **Y** decompositions respectively. A is the number of PLS components used in the decomposition. The maximal squared covariance of the score vectors for each component is seeked as follows:

$$\max(\mathbf{u't})^2 = \max(\mathbf{q'Y'Xp})^2 \quad \text{for} \quad |\mathbf{p}| = |\mathbf{q}| \qquad (17)$$

There are several different ways to find the PLS solution (Martens and Naes, 1989; Höskuldsson, 1996; Vandeginste et al., 1998).

As an interpration of PLS model scores **t** and **u** describe the samples in **X** space and **Y** space respectively. The similarities and dissilmilarities of the samples within the **X** and **Y** spaces can be explored from the score vectors. Correspondingly the loadings **p** and **q** and weights **w** describe the variables. The loading vectors are found similarly than in PCA model and the weight vector is rotated such that it gives the best possible correlation with Y when the covariance between **t** and **u** is maximized. Basically, it can be seen from the weight vector which variables have the most effect , i.e., have the highest weight in interpreting the **u**.

In practice, a building the calibration model is a multistep procedure which starts with the evaluation of the quality of the samples. Then a proper pretreatment method should be selected and data evaluated after preprocessing. The number of components in the model should be decided based on some validation criterion. The decision of the correct number of components, i.e., the complexity of the model should be carried out by using appropriate validation method. Too few components in the model means that all the variation within **X** explaining **Y** has not been included in the model and predictions of the new samples are not as good as they could be. Too many components in the model mean that irrelevant variation that does not predict **Y** is included in the model, which can result in overfitting problems. Too complex model can be unstable, and calibration samples can fit very well to the model, but if new samples are predicted using the derived model, the predictions of new samples are not accurate.

Calibration samples should be evaluated with sensitivity analysis after building the PLS model to find out possible new outliers in the **XY** space. Some sensitivity analysis methods are described in Chapter 7.6. If outliers are found, they should be left out from the model. A new model should be built from the data where the outlier was removed. Finally, the model should be validated using the proper validation criterion, e.g., using an external test set (Chapter 7.7).

7.6 Outlier detection

Outliers are the samples that exhibit somehow different structure from other samples in the modeled data. PCA and MSPC analyses can be used to evaluate the sample quality in the **X** space (Chapters 7.2 and 7.3). Sensitivity analysis describes how the results of the modeling depend on the data used and those methods can be used for pointing out possible outliers from the calibration set. (Höskuldsson, 1996) It is possible that some samples that are appropriate within the **X** space can be considered as outliers within the **XY** space. There are number of different criteria for sensitivity analysis (Vandeginste et al., 1998). Two sensitivity analysis criteria applied in this thesis, leverages and studentized residuals, are introduced in the following two paragraphs.

Object leverages can be considered as measure of the influence of a point (observation) on the PLS-model. Samples with high leverage values have large influence on the model, and those can be different from other samples to some extent. (Eriksson et al., 2001) There are typically two reasons for high leverage values: either the sample exhibits somehow an extreme sample in the group of samples or it is an erroneous measurement, and thus an outlier. Leverage values can be calculated in **X** space and in **XY** space. The leverage value for object *i* for in the **X** space and in **XY** space denoted as $\mathbf{h_X}$ and $\mathbf{h_{XY}}$ respectively can be calculated for centered data as follows:

$$\mathbf{h}_{\mathbf{X}_i} = \mathbf{t}_i \left(\mathbf{T}'\mathbf{T}\right)^{-1} \mathbf{t}_i' + 1/n \tag{18}$$

$$\mathbf{h}_{\mathbf{XY}_i} = \mathbf{h}_{\mathbf{X}_i} + (\mathbf{y}_i - \hat{\mathbf{y}}_i)^2 / \sum (\mathbf{y}_i - \hat{\mathbf{y}}_i)^2 \tag{19}$$

(Höskuldsson, 1996; Eriksson et al., 2001). The leverage in the **Y** space is thus calculated as the squared error of prediction of the calibration set.

Studentized residuals are defined as follows

$$\mathbf{r}_i = (\mathbf{y}_i - \hat{\mathbf{y}}_i) / \mathbf{s}\sqrt{\mathbf{1} - \mathbf{h}_{\mathbf{X}_i}} \tag{20}$$

The confidence limits for the leverages and studentized residuals can be calculated as presented in Eq. 13. The leverages of **X** space and **XY** space can be studied simultaneously with

studentized residuals. These criteria offer good measures for detecting possible outliers that may affect on the regression parameters of the model. Detailed explanation on the sensitivity analysis criteria is presented in, e.g., Höskuldsson (1996).

7.7 Model validation

Validation procedures include the data quality evaluation, model complexity selection and analyzing the performance of the model in the true purpose where the model is to be used. Model validation procedures are used to define the correct number of components in the model and to test that the model works properly. There are several different statistics available to define the number of components needed in the model and these methods are described, e.g., in Höskuldsson (1996). In this study the cross validation (CV) method is used to define correct number of components used to obtain stabile PCA or PLS model. An validation using external independent test set is considered to test the predictive ability of the model.

CV is an internal validation method, where calibration procedure is repaeated several times by leaving one or more samples out of the calibration, building the calibration without those samples and predicting the excluded samples. In full cross validation this procedure is repeated until all the samples have been left out from the model and predicted once and only once (Martens and Naes, 1993).

The measures defined by cross validation include the cross validated coefficient of determination Q_{cv}^2, which corresponds to the captured variance in cross validation and is defined as

$$Q_{cv}^2 = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y})^2} \qquad (21)$$

where y_i is the response in cross validation, $\hat{y}$ is CV predicted value of y, $\bar{y}$ is the mean value of the ys. The validity of the CV criterion depends on the calibration data. The CV criterion, especially leave-one-out validation, can give over-optimistic estimation on the model validity. In addition, for the autocorrelated data the commonly used random selection of the samples for partial CV is not appropriate.

To obtain the true evaluation of the model performance, it should be tested that all the variation that will be present in future samples is included in the calibration set. Internal validation cannot estimate how good the model is predicting unknown samples. Therefore, as a validation of the true performance of the model root means squared error of prediction (RMSEP) value of an external test set is suggested.

$$PRESS = \sum_i (y_i - \hat{y}_i)^2$$
$$RMSEP = \sqrt{PRESS / n} \qquad (22), (23)$$

PRESS refers to a predicted residual error of sum of squares. In practice test samples are brought into the model as new samples and the responses are predicted using the derived model and, e.g., the RMSEP value is calculated. In practice, it one should evaluate carefully whether the additional component enhances RMSEP value significantly or just nominally. To obtain a reliable predictions of the test set these external samples must not be autocorrelated to the calibration samples in any way. External validation is extremely important especially, when the predictive model is to be used very close to its range and also in the cases where there is the possibility that some extrapolation is needed in predictions of true crystallization measurement points. If extrapolation is needed, it would be beneficial to have an external test set having responses close to true values to be predicted.

7.8 Parallel factor analysis (PARAFAC)

In PARAFAC analysis the data is decomposed into trilinear components based on alternating least squares estimation principle. This decomposition is a generalization of bilinear PCA. In regular PCA the data is divided into scores and loadings, but the decomposition obtained from the PARAFAC consists of one score vector and two loading vectors. However, these vectors are often all called loading vectors since those are treated numerically in a similar manner, i.e., treatment of different modes (directions) in modeling is equal. Another difference between the PARAFAC and PCA is that in the PARAFAC model orthogonality between the components is not required as is the case with PCA model. (Bro, 1998)

The axes of PARAFAC model are unique, which means that no other post processing is needed, because there is no other model with the same fit in the least squares solution. However, this does not necessarily mean that the solution is true: Number of components selected has to be correct, the global minimum should be found, the data must be trilinear, and trilinear model itself should be the correct way of estimating the solution (Bro, 1998). There is not rotational freedom in the PARAFAC model.

The matrix notation of PARAFAC can be written as

$$\mathbf{X}_{k_{PAR}} = \mathbf{A}_{PAR}\mathbf{D}_{k_{PAR}}\mathbf{B}_{PAR}{}^{\mathbf{T}} + \mathbf{E}_{k_{PAR}}, \quad k_{PAR} = 1,..., K_{PAR} \qquad (24)$$

where K_{PAR} is the number of variables in the third mode, k_{PAR} is the frontal slab under consideration (shaded with gray in Figure 7). $\mathbf{A}_{PAR}$ and $\mathbf{B}_{PAR}$ are loading matrices of directions presented in Figure 7 and $\mathbf{D}_k$ is a diagonal matrix of which values are the kth row of $\mathbf{C}_{PAR}$,

which is the loading vector of the third mode (Figure 7). **E** is the matrix of residuals. (Bro, 1998)

In practice, for each k_{PAR} slab (shaded with gray in Figure 7) bilinear model is derived. For different k_{PAR}s, the model parameters $\mathbf{A}_{PAR}$ and $\mathbf{B}_{PAR}$ will be similar but they they will differ thorough the diagonal matrix $\mathbf{D}_{k_{PAR}}$, which is different for different k's. (Bro, 1998) Detailed mathematical background of PARAFAC model decomposition is presented in Bro, 1998. In PARAFAC analysis the data is assumed to exhibit a trilinear structure.

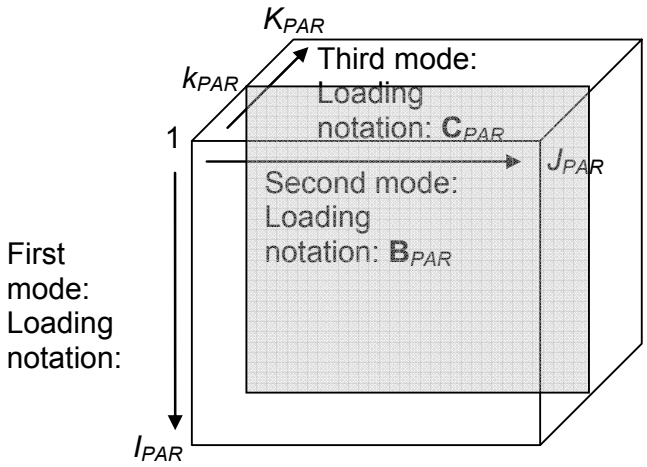

Figure 7 Three way data matrix.

Preprocessing of multi way data is not a straightforward task. Regular preprocessing methods, such as centering and scaling can be applied to multiway data. Centering can be performed against all the modes, but it must be done one mode at a time. However, sometimes centering results in some artifacts in the data, which are not easy to explain with the PARAFAC model. In this case, centering does not simplify the model as would be expected. Scaling has to be done for the whole slabs or submatrices directly; it is not appropriate to unfold the data prior to scaling. It has to be kept in mind, that scaling and centering are interdependent, and the scaling within one mode disturbs the centering across this mode. Therefore, centering should be performed after scaling (Bro, 1998).

To validate the model, different validation criteria can be calculated, and the model validity analyzed thorough these criteria. Validation from a statistical point of view distinguishes how well the properties of the data fit the assumptions of the model. Probably the most important validation procedure in practical applications is to answer the question: How well the model

explains, generalizes or predicts the real problem it is aimed at describing. Different validation methods for N-Way model validation are widely explained in Bro (1998).

The decision of the number of components included in the model can be done using several criteria: rank analysis, split half analysis, residual analysis. In the case of PARAFAC modeling the core consistency diagnostics is a one way to define the number of components needed in the model, and the mathematical background of this method is explained in detail in Bro (1998).

Different constraints can be set up priori to the PARAFAC modeling. Constraints can be helpful, e.g., in avoiding degewneracy and numerical problems or enable the quantitative analysis from qualitative data. (Bro, 1998) Non-negativity constraint is often used within the spectral data as the spectral parameters typically are positive. (Bro, 1998)

The applications of the PARAFAC models include the exploratory analysis, where the interrelationships between the variables and objects are studied. In addition, the PARAFAC models can be used for determining, e.g., a spectra of pure analytes from the spectra measured from mixtures of analytes. (Bro, 1998)

The loadings of the PARAFAC models are used to visualize the features the modeled phenomena. The loadings of certain mode represent the phenomena present in the data towards that mode. In addition, the residuals of the model and contributions of specific variables can be explored to evaluate the batches.

7.8.1 Analysis of batch-to-batch variation

Seeking batch-to-batch variations requires similar data to be gathered from different runs. When this data is gathered over time, and in, e.g., spectral data there are hundreds of variables measured at one time point. The data gathered from one batch exhibits a two dimensional structure. When several batches are evaluated, this causes the data to be a three way array, and special methods are required for analyzing the multiway data. From batch-to-batch variation analysis it should be possible to distinguish the batches which are different from the others and also, perhaps, to evaluate the causes for these differences, e.g., the time where different phenomena in the batch occur and also the variables which cause these variations in a particular batch.

There are several different methods applied to track changes between the batches. These methods differ from each other by the structure of the model. Geladi (2002) and Geladi and Åberg (2001) applied PARAFAC modeling to analyze batches from NIR data and could explain the reaction pattern using this method. In addition, Storgrange et al. 2004 applied PARAFAC to analyze batches from simulated spectral datasets and they could detect the major variation in the

data, evaluate the quality differences between the batches and the addition of the reagent could be pointed out. Other applied methods to study batch-to-batch variation include, e.g., multiway principal component analysis (MPCA) (e.g. Nomikos and MacGregor, 1992, 1994, 1995 a, b; Flores-Cerillo and MacGregor, 2004), structuration des tableauz a truas indices dela statistique (STATIS) (Gourvenec et al., 2005) and Tucker3 (Storgrange, 2004) models. The use of multiblock methods (MBPCA/MBPLS) in batch process analysis has been applied by Kourti et al. (1995, 1996) and Lopes et al. (2002). Boque and Smilde (1999) used PARAFAC based multiway covariates models to analyze batch process. In addition, if the quality measurements are available multiway PLS can also be applied to study batch-to-batch variations (Nomikos and MacGregor, 1995). Multiway PLS is proposed by Wold, 1987 and in this method the multiway data is unfolded to two dimensional structure and regular two dimensional PLS is performed. There exist a N-Way PLS proposed by Bro (1996) which generalizes the two-way PLS algorithm to the data of higher orders using multilinear PARAFAC-type structure. Explanation of these methods is beyond the scope of this thesis, and the reader is advised to look for references mentioned for more information.

Smilde (2001) stated that the practice always has to show, which model works best for the present case. In addition the goal of the empirical process model defines also, which multivariate model performs the best. It has been argued that the PARAFAC in analyzing the batch data is would be unsuitable method. If there are several on-line process measurements included in the data, the data does not have trilinear structure, and this restricts the applicability of PARAFAC method since trilinearity is assumed in PARAFAC. Spectral data is a special case of the process data and it consists of a linear combination of each absorbing species according to Beer-Lambert law, while regular process data (temperature, pressure and flow rate) can often exhibit strongly non-linear structure (Gurden et al., 2002). Therefore, when spectroscopic measurement is used to measure composition with time, and it is the only type of data, PARAFAC could be considered as an suitable method to analyze batches.

8 EXPERIMENTAL

8.1 Materials and methods

The crystallized material was pharmaceutical grade (European Pharmacopoeia/United States Pharmacopoeia) sulfathiazole (Industrias GMB S.A., Castellbisbal, Barcelona, Spain) and the solvent components were deionized water and 1-propanol [purity min. 99.7 % (w/w); Aspokem Oy, Helsinki, Finland] in Papers I-III,V,VI. The another crystallized material in Papers I and IV was organic ingredient C15 and the solvent used was toluene.

8.2 Crystallization experiments

The process conditions applied in sulfathiazole and C15 crystallizations are presented in Table 3. All the process condition combinations in Table 3 for sulfathiazole have not been reported in this study, but the combinations reported are as follows: The batch to batch variations in the crystallizations due to different impeller configurations and impeller speeds was studied using the data gathered from the crystallizations using a 50/50 w-% mixture of water and 1-propanol and a constant cooling rate of 9.2°C/h, which results in number of batches being 27 and this is reported as new results in Chapter 9.2. The effect of solvent composition on the outcome of the crystals in Papers II and VI as well as the on-set of crystallization process in Paper V were studied using different solvents, at different constant cooling rates using a curved blade turbine impeller with an impeller speed of 400 rpm. The different cooling modes and cooling rates were studied using a constant solvent composition of 50/50 w-%, and a curved blade turbine impeller with an impeller speed of 400 rpm in Paper III.

Table 3. Process conditions applied in sulfathiazole crystallizations

	Sulfathiazole (Papers I, III, V, VI)	**C15 (Papers I, IV)**
Reactor:	4 dm^3 jacketed glass reactor	1 dm^3 jacketed glass reactor
Baffles:	4 baffles	4 baffles
Impeller types and speeds:	Pitched blade turbine: 250 rpm, 400 rpm, 550 rpm Anchor impeller: 100 rpm, 175 rpm, 250 rpm Bar turbine: 900 rpm, 1200 rpm, 1500 rpm Curved blade turbine: 250 rpm, 400 rpm, 550 rpm	Pitched blade turbine: 500 rpm
Thermostat:	LAUDA RK 8 KP	LAUDA RK 8 KP
Temperature measurement:	Pt100	Pt100
Solvent:	Water 1-propanol 25/75 w-% water/1-propanol 50/50 w-% water/1-propanol 75/25 w-% water/1-propanol	Toluene
Amount of dissolved solid:	0.9 g stz/100 g water 2.2 g stz/100 g 1-propanol 10.9 g stz/100 g 25/75 w-% water/1-propanol 20.7 g stz/100 g 50/50 w-% water/1-propanol 15.8 g stz/100 g 75/25 w-% water/1-propanol	35 g C15/100 g toluene
Solubility temperature:	80°C	75°C
Initial temperature:	85°C	80°C
Final temperature:	25°C	25°C
Seeding:	Unseeded	Unseeded 1 w-%, 2 w-%, and 3 w-% loads of seed crystals. Seeding was done at 74.5°C
Cooling modes (batch time):	3.9°C/h (14 h), 5.5°C/h (10 h), 9.2°C/h (6 h), 27.5°C/h (2 h) Programmed cooling (6 h) Natural cooling (6 h)	10°C/h (5 h) Programmed cooling (5 h) Step mode 1 (5 h) Step mode 2 (5 h)
Initialization moment of the programmed, natural and step cooling modes:	75°C (1°C above the observed nucleation in the crystallization experiment with 9.2°C/h cooling rate)	75°C (at the solubility temperature)

The solution was kept at 85°C in sulfathiazole crystallization and at 80°C in C15 crystallizations for one hour to make sure that all of the added solute was dissolved before the beginning of the cooling stage. The system was thus undersaturated for the first 5°C in the cooling stage and supersaturated for rest of the process for both crystallization processes. The duration of the crystallization experiments presented in brackets in Table 3 with constant cooling rates used correspond to the cooling in the supersaturated stage, because supersaturation is the prerequisite for the crystallization process to occur.

For programmed and natural cooling, the duration of the experiment presented in the brackets in Table 2 refers for the cooling in the supersaturated stage. The duration of the experiments was fixed similar than that of the linear cooling (one of the linear cooling rates in sulfathiazole case) in order to compare the results with the experiments carried at constant cooling rate. The different cooling modes are illustrated in Figure 3. The programmed cooling was defined using Eq. 3. The experiments with the natural cooling in sulfathiazole crystallizations were performed by rapidly changing the temperature of the cooling medium to a constant temperature of 25°C. The final temperature of the batch, 25°C, was reached approximately in 1.5 hours. The crystals were maintained in the stirred solution at a constant temperature of 25°C for the remaining 4.5 hours.

The stepwise modes in C15 crystallizations resembled somehow natural cooling in the sulfathiazole case, but those were performed by setting the thermostat cooling water temperature to a constant value, and letting the system cool down. Also the set point of the coolant was set to the final temperature of the batch right after the system was supersaturated. Two different stepwise processes were used: In the first case the cooling water temperature was set to 40°C and after the system was cooled to that temperature linear cooling from 40°C to 25°C was performed so that the total batch time was 5 h (Step mode 1). In the second case the temperature of the cooling water was set straight to the final value 25°C (Step mode 2). Crystallizations with linear and controlled modes were performed as unseeded and by using 1 w-%, 2 w-% and 3 w-% load of seed crystals. Step mode 1 was done without seed crystals and using 2 w-% seed crystals and step mode 2 was only performed as an unseeded process. The seed crystals were obtained by sieving, and the fraction under 50μm was used as the seed crystals.

The detailed explanation of the sulfathiazole and C15 experiments is presented in Papers III and IV respectively.

8.3 In-situ ATR-FTIR measurements

ATR-FTIR measurements were taken using an ABB BOMEM MB155S spectrophotometer equipped with a Dipper 210 ATR immersion probe with a conical ZnSe element (in Sulfathiazole measurements Papers I, III, V, VI) or AMTIR element (in C15 measurements I,IV). The probe was placed inside the crystallizer. The probe was manufactured by Axiom Analytical Inc.. The shape of the element causes the light beam to undergo two internal reflections at the interface of the sample and the element before the attenuated reflection proceeds to the detector. Grams32 software was used to collect the spectra. The mid-IR spectrum from 4000 to 750 cm^{-1} was collected. As a compromise between the quality of the spectrum and the robustness of the measurements, the spectral resolution of 16 cm^{-1} was used and each spectrum consisted on average of 20 consecutive scans. Water at 25°C was used as a background spectrum in sulfathiazole measurements (Papers I,III, V, VI) and toluene at 25°C in C15 measurements (Papers I, IV). The measured spectrum was thus actually, the background subtracted spectrum, which is in this study referred to as "raw spectrum" or "measured spectrum".

8.3.1 Calibration measurements

It is important that the calibration routine covers the range of variation where the actual measurements take place, and therefore the design of the calibration measurements is an essential issue. However, in cooling crystallization applications, the concentration measurement range in the calibration measurement is restricted by the solute solubility in the solvent and the calibration measurement with known concentration has to be made in an undersaturated state, to be able to have the exact concentration points for the calibration measurements. True process condition measurements in crystallization processes take place in the supersaturated state, which is unstable and reliable calibration measurements are not possible in that state since any distraction can cause the nucleation to begin and the solution concentration to be erroneous. In addition, changing the temperature and possibly changing the solvent composition cause variation in the absorption spectrum. This variation should also be covered in the calibration measurements.

The calibration measurement plan for the sulfathiazole concentration measurement is presented in Figure 8. Calibration set was measured in the same vessel using similar mixing conditions for the true process conditions. The single sample was measured by weighing the corresponding amount of solute in the solvent, letting that dissolve, and after all the solute had dissolved, the absorption spectrum was collected. In practice, the time needed for dissolving was two hours at maximum. Several separate calibration measurement runs were performed and separate

calibration and test sets were measured. The test set was used to select the proper calibration model. A detailed explanation on calibration measurements is presented in Papers I and III.

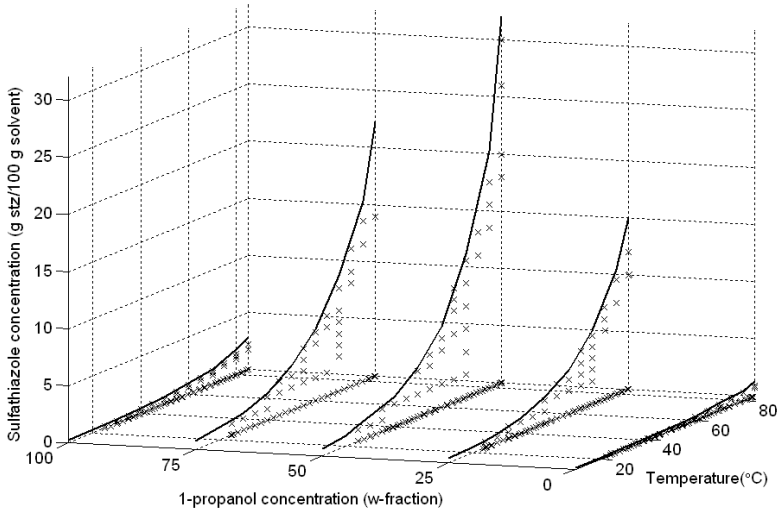

Figure 8 The calibration measurement points (x) and the solubilities of sulfathiazole in different w-fraction mixtures of water and 1-propanol (-).

The calibration measurements were performed from clear solution and in an undersaturated state. As mentioned, the true measurements take place mainly in the system where both liquid and solid phases are present and in supersaturated conditions. Therefore, an additional testing procedure was performed, i.e., where the solubility of the solute in the solvent was measured using ATR-FTIR from the system by introducing more solute to the solvent than would dissolve at any conditions within the process. The results were compared to the gravimetrically measured solubilities, which are described in Paper III. This caused the solid and liquid phases to be present during the whole measurement, which was not the case in calibration measurements. The suspension was kept at a constant temperature until the saturated solution was obtained which took approximately two hours and then the absorption spectrum was measured. This was done for the whole temperature range. Descriptions of the solubility measurement are provided in Papers I, III and IV.

8.3.2 ATR-FTIR measurements from the crystallization process

The absorption spectra were measured throughout the whole crystallization process with a time increment of 1 min. The concentration of dissolved solute in the solvent was predicted using the derived PLS model. In addition, the spectral data gathered from the sulfathiazole crystallizations were used to predict the nucleation moment and analyzing the batch-to-batch

variations in different sulfathiazole crystallizations. The description of the measurement procedure is presented also in Papers III and IV.

8.4 In-situ measurement of the crystal size distribution (IV)

In-situ measurement of the crystal size distribution in C15 crystallizations in Paper IV was done using a MTS PSyA Laser Reflection Analyzer. The system is equipped with an immersion probe placed inside the crystallizer. The measurement procedure in this application is presented in Paper IV.

8.5 Off-line analysis of the polymorphic composition of product crystals

8.5.1 DRIFT-IR measurements

The sulfathiazole samples were measured using a Perkin Elmer IR spectrophotometer accompanied by a diffuse reflectance accessory. The accessorry was conical in shape, and the diameter of the top of the accessory was 9 mm. The measurements were done with spectral resolution of 8 cm^{-1} and collecting 10 averaging scans to one spectrum. The ground KBr powder was used as the background in the measurements. The measured wave number range was from 4000 to 450 cm^{-1}. To avoid measurement errors due to effects from the wide particle size distribution and to make samples more homogenous the samples were ground before measurement and the measurement from one sample was repeated two to five times. A detailed description on the DRIFT-IR measurements are presented in Papers II and VI.

8.5.2 X-Ray powder diffraction (XRPD) measurements

The structure of the obtained sulfathiazole crystals was measured using an X-Ray Powder Diffractometer Bruker axs D8. The samples were ground priori to the measurement to reduce the preferred orientation of the crystals and to make them more homogenous. The XRPD experiments were performed in a symmetrical reflection mode with CuKα radiation (1.54 Å) using Göbel Mirror bent gradient multilayer optics. The scattered intensities were measured using a scintillation counter. The angular range was from 5^o to 40^o with steps of 0.05^o and a measuring time of 1 s/step. A description of the XRPD measurements are presented in Papers II and VI. The XRPD measurements and the estimation of the polymorphic composition from diffractograms were performed in Helsinki University Faculty of Pharmaceutics and the results obtained were the approximate polymorphic composition of the measured samples.

8.6 Off-line analysis of size and shape of the product crystals

The final product size was characterized in C15 crystallization experiments using a Coulter LS 130 laser diffraction analyzer. The description of the measurement procedure is presented in Paper IV.

The automated image analysis PharmaVision 830, Malvern Instruments, Ltd. was used to characterize the size and shape of produced sulfathiazole crystals. The description of automated image analysis technique, and measurement procedure are presented in Paper III. The size and shape of the crystals by crystal length and roundness were described. The crystal length refers to the largest projection of the crystal, whereas the roundness is a measure of the length to width ratio. For a perfect circle, the roundness equals 1, and for a needle shaped crystal, the value of roundness approaches 0. These parameters were readily provided by the PharmaVision-software.

8.7 Data analyses

Several different types of data analyses were applied to the measured spectral data. All the data analyses were done using Matlab ver. 6.5 or 7.0.1. by MathWorks Inc. In two-way analyses Matlab algorithms originally copyright from Eigenvector research were used. Part of the algorithms have been updated to be suitable for newer Matlab versions and also some parameter calculations added by S.-P. Reinikainen. N-Way data analyses were performed using the PLS Toolbox 3.5 by Eigenvector Research Inc..

8.7.1 Analysis of the batch-to-batch variations

The objective of the PARAFAC modeling was to test, whether or not the similarities and dissimilarities beween the batches could be pointed out, and reasons for possible dissimilarities between batches evaluated. In addition, the aim was that the phenomena in the batches in the time scale and spectral variable scale could be distinguished and identified.

PARAFAC analysis was made for the spectral data gathered from sulfathiazole crystallizations to evaluate the batch-to-batch variations in the crystallization experiments. For the validation of the model Core Consistency diagnostics was applied. The model validity was also estimated by examination of how well the model reflected what was already known from the problem.

The obtained spectral data matrices were arranged in the three-way matrix, which represented therefore the structure illustrated in Figure 7. In this study, Mode 1 represented different batches (loading notation $\mathbf{A}_{PAR}$), Mode 2 represented time (temperature) (loading notation $\mathbf{B}_{PAR}$) and Mode 3 measured spectral points, i.e., wave numbers (loading notation $\mathbf{C}_{PAR}$). Also other arrangements for the Mode orders were tested, but the abovementioned arrangement was found to give the most reasonable results.

Raw data was used since centering or baselining did not improve the modeling. Also no restrictions were used since as, e.g. non-negativity for different modes separately was tested.

8.7.2 In-situ monitoring of the onset of the crystallization and forming polymorph

The objective of this analysis was to monitor the crystallization system prior to the onset of the actual crystallization. A dynamically built PCA model together with MSPC charts were applied to the methodology for real time prediction of the on-set of a crystallization process from ATR-FTIR data gathered from sulfathiazole crystallizations. The alarm system for approaching nucleation was proposed. For this purpose, a two-stage procedure was proposed: 1) The PCA model was derived dynamically from the beginning of the process where the system was unsaturated and the MSPC statistics and 95% confidence limits were calculated for that model. 2) The sampled measured under the supersaturated state were predicted using the model derived in stage 1. In addition, MSPC statistics was calculated. As alarm criteria the 95% confidence limits of T^2 and Q statistics were used.

In addition the pseudocolor images of the T^2 and Q contributions of the spectral variables are a function of the temperature (time). They were used to illustrate the changes in the process which can be seen from spectral variation as the nucleation approaches, and to see whether or not this variation can be linked to the polymorphic form of forming crystals. To visualize the small contribution changes from one measurement to another, the difference between the 95% limit of the contributions from the last measured data point and 95% limit of the contributions of the model derived in the calibration stage are calculated and are denoted as dT^2_{lim} and dQ_{lim}. A description of the proposed methodology is presented in Paper V.

8.7.3 The calibration routine for concentration prediction

In the building of the calibration model for solute concentration prediction from 1) in-situ ATR-FTIR measurements and 2) for the quantification of the polymorphic composition of the sulfathiazole bulk material from the off-line DRIFT-IR measurement a multistep calibration routine was applied. Calibration measurements for solute concentration prediction using ATR-FTIR are presented in Chapter 7.3.1 and in Papers I and III. The calibration and test samples for calibration of the off-line powder samples using DRIFT-IR was simply done by measuring the samples obtained from crystallizations with XRPD and DRIFT-IR. The quantification of the polymorphic composition of the sulfathiazole samples from XRPD patterns was done by assuming that measured XRPD pattern is a linear combination of XRPD patterns of the pure polymorph components. The results from XRPD quantification were used as descriptive variables when building the PLS model for the polymorph composition prediction of the powder samples using the DRIFT-IR technique. This procedure is explained in Papers II and VI.

The calibration routine is closely presented in I and the steps and corresponding multivariate methods used are listed below: The data was centered. The MSPC and sensitivity analyses were applied to evaluate the quality of the samples. Improvement in R^2 value and in error of prediction of the calibration and test sets were applied for variable selection. Primary analysis on the spectral knowledge on the importance of the variables was also one main criterion to select the important spectral ranges for multivariate modeling.

OSC filtering methods were applied to preprocess the data. The OSC filtering was selected because its ability of removing Y independent variation from the data. In the studied process conditions Y independent variation can be assumed to exsist, e.g., due to solvent or temperature (solute concentration measurements using ATR-FTIR) or variation related to the particle orientation or size distribution (polymorphic composition measurements using DRIFT-IR). Therefore, the OSC filtering was an obvious option as an pretreatment method. The predictive PLS model was built and the number of components included in the model was selected based on the RMSEP value of an external test set. An additional validation procedure to validate the PLS model for in-situ solute concentration prediction using ATR-FTIR was applied. Solubility was measured using ATR-FTIR and cthe result is compared to gravimetrically measured solubilities in corresponding solute-solvent system. (Chapter 7.3.1 and Papers I, III, IV)

8.7.4 Off-line classification of crystalline samples

In addition to the quantitative characterization of the samples measured from the sulfathiazole crystalline product based on the off-line DRIFT-IR measurements the multivariate qualitative classification methods were applied. The principle method used was PCA, but also the PCA derived methods SIMCA and MSPC statistics were tested for classification of crystalline bulk samples. The description of the use of these methods are applied in Papers II and VI.

9 RESULTS AND DISCUSSION

9.1 Obtained spectral data

9.1.1 ATR-FTIR data

The example spectra from the calibration measurements for solute concentration prediction using ATR-FTIR data are presented in Figure 9.

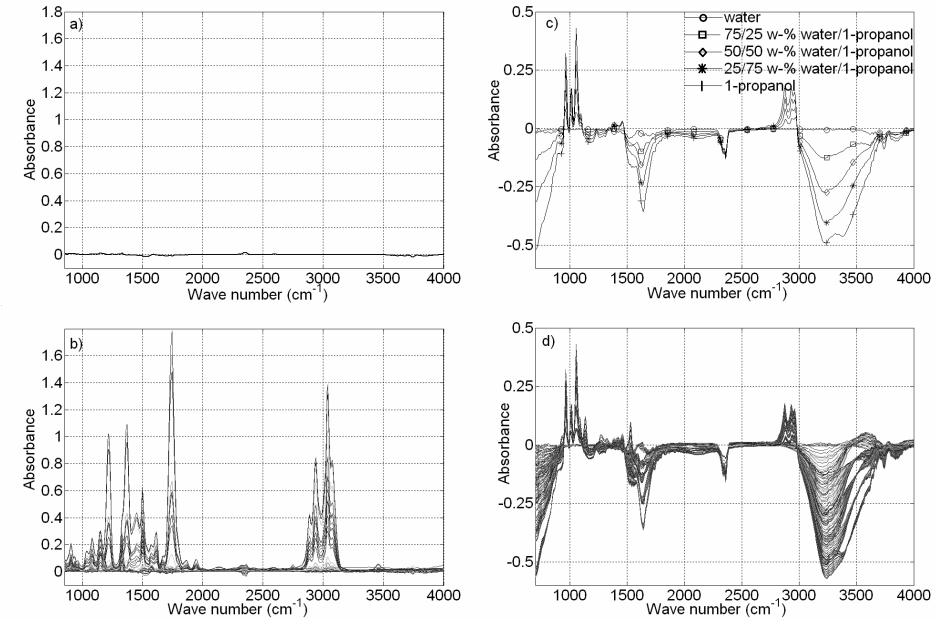

Figure 9 Background subtracted spectra from the solute concentration measurements using ATR-FTIR. Background was water at 25°C. a) and b): C15 measurements.Toluene at 25°C was used as background spectrum and as solvent. a) Toluene spectrum at 25°C b) spectra of different concentrations of C15 dissolved into toluene measured within the temperature range from 25°C to 75°C c) and d): sulfathiazole measurements. Water at 25°C was a background spectrum. c) Spectra of the solvents used d) spectra of different concentrations of sulfathiazole dissolved in different solvents within the temperature range from 25°C to 80°C

Figure 9 clearly shows that for the C15 case the responses are practically due to the dissolved solid, as the pure toluene, which was used as the solvent was used as background. In the sulfathiazole case, however, different solvent compositions cause major effects to the measured spectra when pure water is used as background. When the pure water is used as a background spectrum the spectra of the mixture sample can contain negative values as the absorption from the mixture in certain wave number range is lower than water absorption in that range. The overall absorbance level is ove ten times higher for the C15 system than for the sulfathiazole system. The responses in the spectra due to dissolved sulfathiazole are clearly of a minor scale.

Variation due to sulfathiazole is mainly located within the wave number range from 1000 to 1700 cm^{-1}. This variation can not be seen by the naked eye in Figure 9.

In addition the C15 sysem is binary and the sulfathiazole system is ternary. The C15 crystallization system was the only two-component system C15 and toluene. In the sulfathiazole system, there were two solvent components: water and 1-propanol and the solute sulfathiazole. The C15 system can therefore be considered simpler than the sulfathiazole system. As was discussed in Chapter 6.1 the temperature changes can cause variation in the measured spectra. The solvent spectra for both example cases measured are presented in Figure 10.

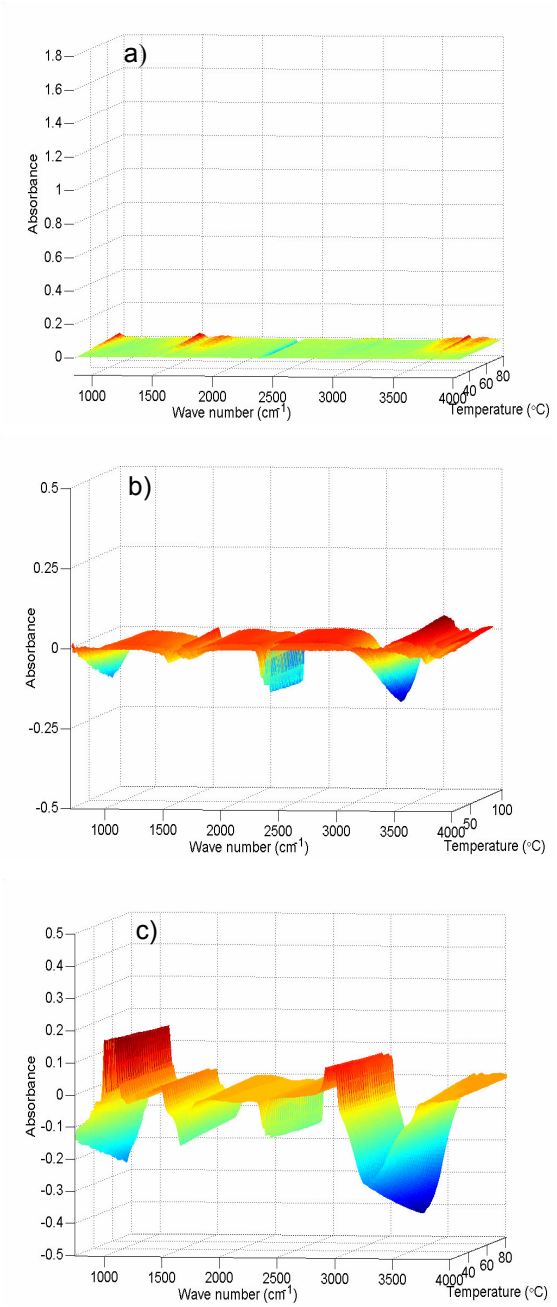

Figure 10 ATR-FTIR spectra of the solvents at different temperatures a) toluene (background: toluene at 25°C), b) water (background: water at 25°C) c) 50/50 w-% mixture water and 1-propanol (background: water at 25°C)

To evaluate the importance of the temperature effect on both cases, the y-axes in Figure 10 are scaled to the maximum variation in the solute or different solvents in the corresponding case

presented in Figure 9. It can be seen that for C15 the temperature effect on the solvent spectrum (absorption variation from -0.03 to 0.03) is minor compared to the responses from the dissolved C15. For the sulfathiazole case the temperature effect is relatively larger than for the C15 case. Especially –OH vibration range from 3200 to 3700 cm^{-1} is temperature sensitive. C15 system has lower temperature effects than does sulfathiazole system.

9.1.2 DRIFT-IR data

An example of DRIFT-IR spectra measured from the powdered sulfathiazole samples is presented in Figure 11.

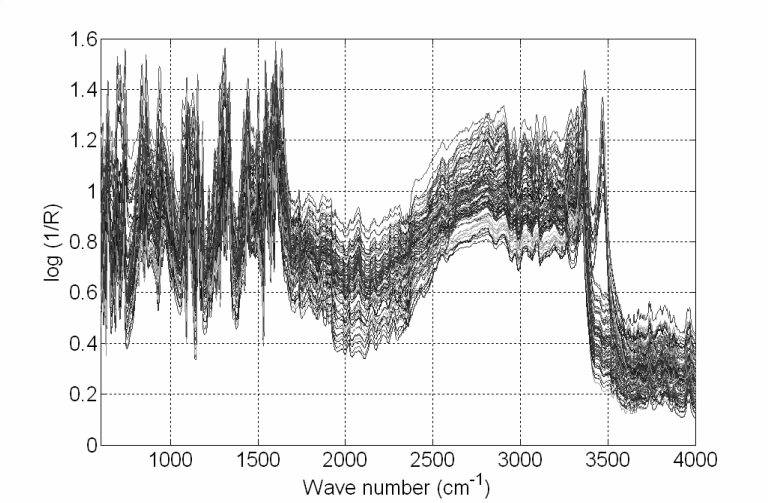

Figure 11 Raw DRIFT-IR spectra measured from powdered sulfathiazole samples. Background spectrum measured from powdered KBr sample.

It can be seen from raw measured spectra in Figure 11 that the responses from the spectral range from 600 to 1500 cm^{-1} are noisy. The variation in the baseline is rather large. There are also clear bands in the range from 3200 to 3600 cm^{-1}.

9.2 PARAFAC modeling

Based on the core consistency diagnostics and the evaluation of how well the model describes what is known from the process, the 2 component PARAFAC model was derived. Core consistency value was 99.96 and the loadings clearly represented true phenomena within the process. Adding a third component caused the core consistency value to crash down (-155,26), which is obviously a sign of the model having too many components. The non-negativity restriction was tested, but that did not improve the model. In fact, non-negativity restriction made interpretation more difficult. This is obviously because in this case, also large number of negative values existed in the measured data as was explained in Chapter 9.1.1.

The PARAFAC model derived for the sulfathiazole crystallizations for a batch time of 6 h with different cooling modes, different solvents and different impeller types illustrates how the different process conditions can be visualized by PARAFAC. The loadings for the different modes are presented in Figure 12. Mode 1 represents different runs, Mode 2 time/temperature during the run and Mode 3 represents the wave numbers.

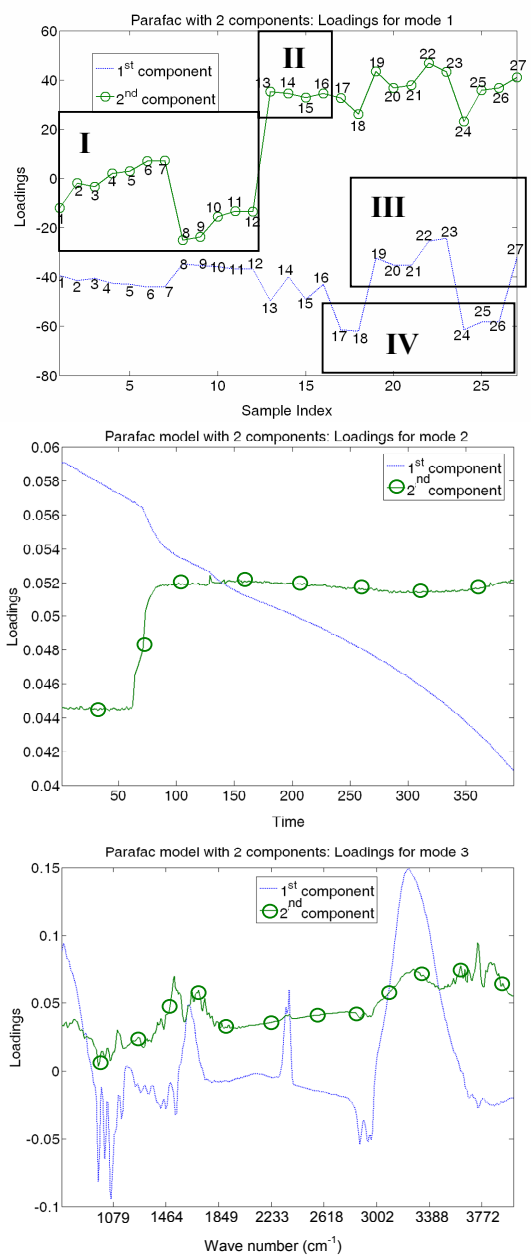

Figure 12 Loadings of the two-component PARAFAC model from sulfathiazole crystallization data. Total batch time 6 h. Different solvent compositions, different cooling modes and different impeller configurations. The squared areas in the up most figure correspond following experiments: I) solvent 50/50 w-% mixture of water and 1-propanol, constant cooling rate, different mixers; II) solvent 50/50 w-% mixture of water and 1-propanol, experiments 13 and 15 with programmed cooling and experiments 14 and 16 with natural cooling, III) solvent 75/25 w-% mixture of water and 1-propanol, IV) solvent 25/75 w-% mixture of water and 1-propanol

Loadings for Mode 1 illustrated in the way as in Figure 12 do not provide a thorough information on the samples. However, the first component seems to separate crystallizations done with different cooling rates. The second component seems to separate the crystallizations done using 50/50 w-% mixture of water and 1-propanol as solvent and constant cooling rate from rest of the crystallizations.

Loadings for Mode 2 represent different aspects of variation that occur as the crystallization process proceeds: The first component reflects the "mean" temperature and may reflect some of the concentration profile. The "mean" here means that both temperature and concentration profiles vary in different runs, and therefore the profile seen here is a combination of those different profiles. However, the number of crystallizations with constant cooling rate is remarkably higher than is the number of crystallizations using programmed or natural cooling, and these can dominate the loadings in Mode 2, as the loadings represent an almost linear structure. There is also a step in Mode 2 loadings for a 2^{nd} component at 70 min which corresponds the the moment of primary nucleation, i.e., the onset of crystallization in every run. As was described earlier, the batches are cooled down from 85°C to the moment of nucleation with a constant cooling rate of 9.2°C/h, this leads to the fact that at 70 min the temperature is thus approximately 74°C, i.e., the temperature where the nucleation is observed. The second component separates clearly the two stages of the crystallization process: low loadings are before nucleation, from 0 to 70 min (from 85°C to 74°C) and high loadings for the part where the crystals grow. The sample separation in the Mode 1 can not be explained by the loading structure in the Mode 2.

Loadings for Mode 3 represent the part of the spectrum that is extracted by that specific component. The first component loadings of the resemble an approximate mirror image of the raw spectral data (Figure 9). The highest loadings in the variable range from 3230 to 3600 cm^{-1} are caused by –OH stretching. This range is sensitive of course to the type of solvent but also to temperature changes. This also explains why the first component separates the samples in Mode 1 based on different solvents, because the most sensitive bands in the first component loadings in Mode 3 are also sensitive to solvent composition changes. Another high loadings in the 1^{st} component are located in the variables from 2230 to 2380 cm^{-1} and where the absorption of CO_2 occurs in the IR spectrum. Variation in this peak exists due to temperature and also other conditions, e.g., different mixing can cause differences in CO_2 absorption. Loadings for the second component for Mode 3 resemble somehow the peaks that are related to the sulfathiazole concentration in the variable range from 1380 to 1700 cm^{-1} but there is also another range in the end of the spectrum variables from 3300 to 3800 cm^{-1} which seem to have an impact on Mode 3 loadings in the second component. The latter range can be related to intramolecular vibrations

(Silverstein et al., 1991), and thus the variation in this range can be due to differently arranged molecules/clusters in the system. This range could thus be sensitive for separating the differences due to different impellers used in crystallizations, but the separation due to different crystallizations could not be seen explicitly from the second component loadings of the Mode 1.

Another way of investigating the PARAFAC model is to visualize one component against another by scatter plot. The 2nd loading against the 1st loading for all three modes in a derived PARAFAC model are illustrated in Figure 13.

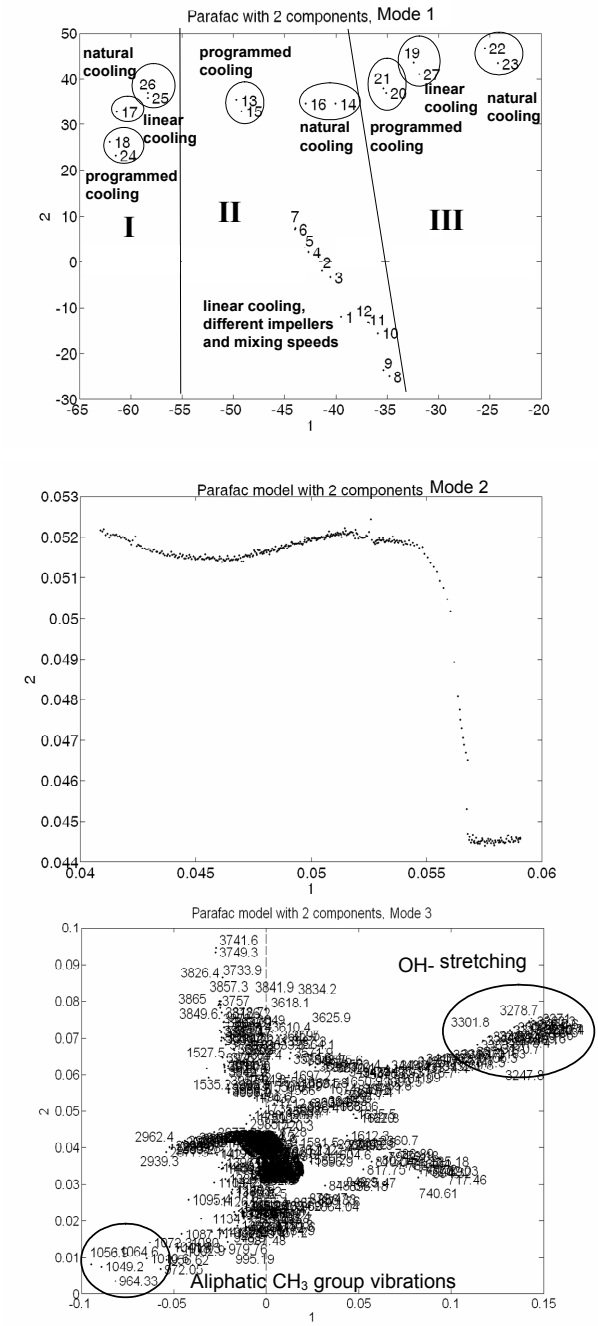

Figure 13 Scatter plots for the modes of a two-component PARAFAC model from sulfathiazole crystallization data. Total batch time of 6 h. Different solvent compositions, different cooling modes and different impeller configurations. The graph on top has been divided into three parts that representing solvent system used in experiments: I) 25/75 w-%, II) 50/50 w-% and III) 75/25 w-% mixture of water and 1-propanol.

The scatter plot for Mode 1 illustrates how the samples are clustered horizontally based on the solvent used (different regions separated with a line in Figure 13 for Mode 1). As the amount of 1-propanol increases the samples are located towards the right in the scatter plot. This result is expected since the 1-propanol concentration has a big impact on spectrum especially in the ranges that represent high loadings in Mode 3 in Figure 12. An further investigation of the scatter plot for Mode 1, it can be seen that inside the specific solvent, the samples cluster based on the cooling profile used, which is also an expected result, since different cooling profiles exhibit different concentration profiles after the nucleation has occurred. In addition, the crystallization with a 50/50 w-% mixture of water and 1-propanol using a linear cooling rate and different mixers form a group of their own.

A scatter plot for the Mode 2 differentiates the situation before and after the nucleation, and the scatter plot for Mode 3 represents the loadings for spectral variables where it seems that the meaning of the –OH stretching group increases towards the right where the samples 75/25 w-% mixture of water and propanol are located. This is a reasonable result since as the 1-propanol concentration decreases the depth of the –OH stretching "absorption dip" increases (Figure 9). This can be seen on the loadings for Mode 3 in Figure 13. Correspondingly, the bands sensitive to 1-propanol concentration in the range from 1000 to 1100 cm^{-1} are located in the very right of the scatter plot which corresponds to increasing 1-propanol concentration in the system. The end of the spectrum from 3700 to 4000 cm^{-1} seems to be sensitive to cooling conditions such that the samples with programmed or natural cooling tend to settle in the upper part of the scatter plot of Mode 1, which corresponds to the high loadings in the range from 3700 to 4000 cm^{-1}. The scatter plot for the Mode 2 simply gives the variation with time of the first component (x-axis direction) and the second component separates the time before nucleation and the time after the nucleation process (y-axis direction).

The model presented above shows how the different process conditions in the batches can be seen in the PARAFAC model derived from the spectral data gathered from the batch processes. Some of the differences in the process conditions, such as differences in the solvent composition cause remarkable changes in the measured spectrum and it is therefore obvious that these effects dominate in the model. It can be argued, therefore, that Figure 13 is not very informative since, for example, the differentiation of different solvent compositions is not clearly very interesting from a practical point of view. Because there was a huge variation present in the experiments, truly interesting small scale variation could not be observed from this set of data.

To truly investigate the subtle variation between batches, the variation in the crystallization experiments should be remarkably smaller than in the experiments presented above. Therefore,

an another model was derived from the experiments done using a fixed solvent 50/50 w-% mixture of water and 1-propanol and using a linear cooling rate but different mixing conditions. Different mixing conditions change the physics in the systems due to the different levels of micro and macro scale mixing. The physical changes can also cause differences in the spectrum as was discussed earlier. Different physical conditions can alter the baseline behaviour and cause differences in position or widths of the bands. To evaluate the possibility of extracting the different mixing conditions from the measured spectra the two component PARAFAC model was derived from the batches where the solvent composition was 50/50 w-% water and 1-propanol and the linear cooling with 9.2°C/h was used, but the impeller types and mixing intensities varied. The results of this model are illustrated as scatter plot in Figure 14.

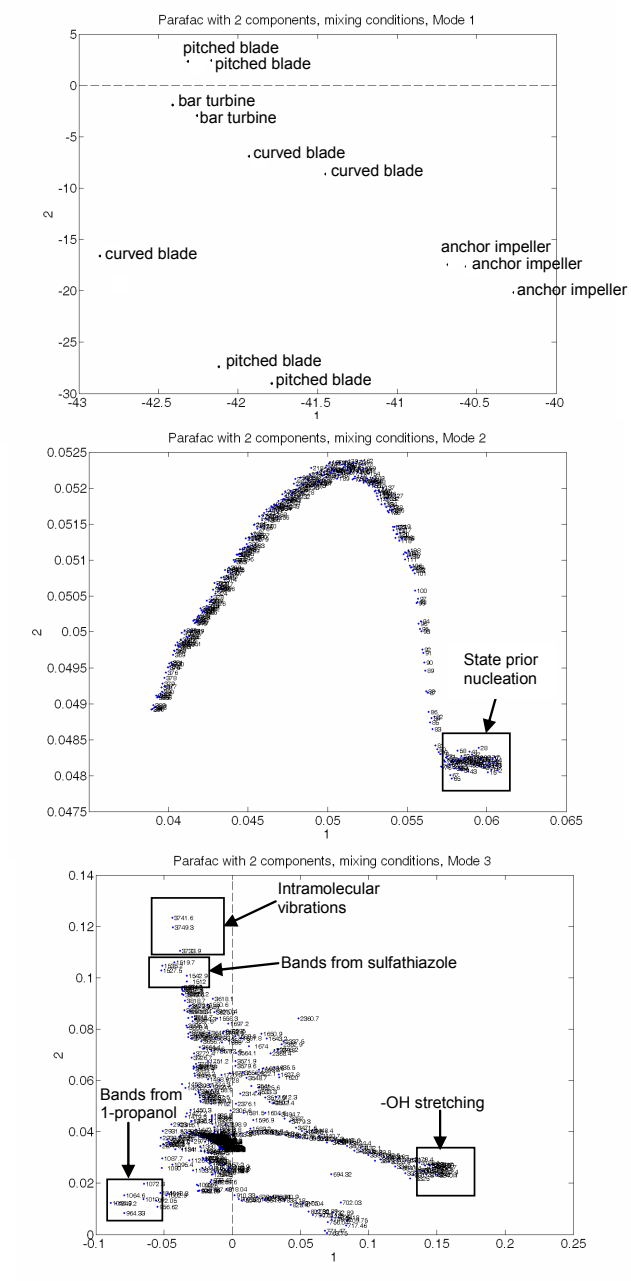

Figure 14 Loadings for two-component PARAFAC model from sulfathiazole crystallization data. Total batch time of 6 h. Solvent composition: 50/50 w-% mixture of water and 1-propanol, constant cooling rate of 9.2°C/h. Different impeller configurations: pitched blade turbine, anchor impeller and bar turbine

Scatter plots of the Figure 14 show that the samples with different mixing conditions cluster into groups in Mode 1. The batches produced using different mixers are mainly clustered together. There are three samples located in the the bottom-left of Figure 14 for Mode 1. These

samples are clearly separated from the other samples. When refletcing on the scatter plots of Mode 1 to the scatter plot of the Mode 3, it can be seen that this horizontal movement is clearly related to the bands of 1-propanol sensitive range and –OH vibrations (Figure 14). This was also seen earlier when the crystallizations using different solvents were used. The vertical position on the other hand is related to the peaks sensitive to sulfathiazole concentration. It is possible, therefore, that the solution composition is slightly different in those three samples in the bottom left of the Mode 1 scatter plot in Figure 14 than in the other samples. Most of the samples, however, show a profile that is linearly decreasing and the samples are clustered based on the impeller used.

The scatter plot for the Mode 3 (Figure 14) shows that the vertical changes in the samples (referring to the 2nd component in the model) are due to the end of the spectrum from 3700 cm^{-1}, in this part there are not clear bands in the spectrum that indicate clearly the concentration species of interest in the system, this change can be caused by the –OH bonds arranging differently in the alcohol molecule or beween alcohol and water molecule since this range refers to alcohol intra- intermolecular bondings. (Silverstein et al., 1991) This change can be due to the fact that different mixers mix the system differently in the micro scale and therefore the solute and solvent molecules can be blended differently. There is also a slight trend in the baseline in the scatter plot in Mode 3.

It can be concluded from PARAFAC modeling results that the runs with different process conditions could be separated and pointed out. Especially when the chemical composition of the mother liquor and, consequently, the amount of dissolved solute was different the batches could be separated and the spectral ranges which were different in these batches could be pointed out. This result is expected, since the spectra measured from the system where different solvent mixtures are used have major differences in several wave number ranges. Therefore, this result could not be considered very interesting. In addition, the differences in cooling mode and impeller type caused the samples cluster into their own groups. The spectral ranges where the changes due to different cooling modes and impellers possibly occur in the measured data could be pointed out.

In the time scale, only the most drastic occurrence, nucleation, could be explicitly separated from the rest of the process. Otherwise, the time scale loadings represented either the “mean” concentration and/or temperature profiles. From the data under investigation and using this PARAFAC method it was not possible to observe the possible differences in the dynamics of different processes, which would have been a very interesting issue to be investigated. The contribution plots for each experiments were compared, but they did not show sufficient

sensitivity to estimate possible differences in the dynamics of the different experiments from this type of data.

It is possible, that from other types of data, e.g., data from crystal phase monitoring or by combining two or more monitoring methods, the differences in the dynamics of the process could be evaluated. In that case, PARAFAC may not be the best suitable technique for this type of analysis as has been argued in the literature and was discussed also in Chapter 7.8. In addition, truly industrially interesting batch-to-batch variation analysis would require large number of batches done with constant process conditions and those to be analyzed using a suitable exploratory analyzing method.

9.3 Prediction of the nucleation and of the forming polymorph

The results of the MSPC analysis of the ATR-FTIR spectra measured from sulfathiazole crystallization processes for monitoring the assumed solute molecule clustering prior to nucleation (described in Chapters 4.2.1 and 4.3.1) are presented in Paper V. The objective of the MSPC charts was to monitor small changes in the solution phase as the nucleation approaches. The PLS modeling derived for solute concentration prediction (Chapter 9.4) is not capable of detecting the approaching nucleation, because the solute concentration should be constant prior the primary nucleation.

9.3.1 Prediction of nucleation

The PCA results presented in Paper V showed that there clearly exists a systematical variation in the measured IR spectrum as the nucleation approaches. The nucleation moment come out clearly from the PCA results made from in-situ measured IR spectra. Thus, the basis for using a PCA derived method for predicting nucleation and creating an alarm system for approaching nucleation exists.

The alarm criterion of the approaching nucleation was set up. It was found that for setting up the alarm criterion only one sample exceeding either the 95% confidence limits of T^2 or/and Q statistics was not sufficient, because the alarm was obtained too early in many tested crystallization experiments and the alarm time was not constant for different tested crystallization processes. Therefore, alarm was set as three subsequent samples 95% exceeded the confidence limit for both T^2 and Q statistics. This criterion performed well leading to the alarm of approaching nucleation 1.5±0.2 ºC before the nucleation was detected.

Possible drawbacks of the proposed alarming methodology can be that the number of samples measured in the undersaturated stage and thus included in the PCA model can be too small for deriving a stable model. In this study, the number of samples measured for the calibration set,

33, and with this amount of samples, the model seemed to perform adequately. However, the number of samples cannot be very much smaller than used 33 and actually greater number of samples measured in the undersaturated stage would be beneficial. This would require more frequent sampling in the process system, which is recommended in future applications.

Besides the drawbacks, the proposed methodology proved to be easily applicable in an automated form. The proposed methodology provides a new way to monitor a crystallization process and enhance the possibilities for crystallization process control. The nucleation prediction scheme could be included to the crystallization process-monitoring scheme. As was mentioned in Chapter 5.1.1, the development of a closed loop control of the crystallization process is an important topic in crystallization process research. To control crystallization processes successfully, information on the nucleation, the most chaotic part of the crystallization process should be obtained in real time. In addition, nucleation monitoring also provides the tool to observe possible faults in nucleation process, e.g., premature nucleation, real time. This enhances the possibility of taking correcting actions in the process control during the batch to prevent bad product quality.

9.3.2 Predicting the polymorphic form of the forming crystals

The contribution charts were used to evaluate the spectral variables and especially the change in the spectral variables as the nucleation approaches. Especially pseudo color image presentation was a useful way of showing the variation within the spectral variables. The method used in Paper V was to visualize the change in the 95% confidence limits with time. By this way the spectral ranges that undergo time dependent changes as the nucleation proceeds could be emphasized. When different polymorphs were resulted, the spectral ranges, that gradually changed prior to nucleation were different.

Observed changes in the spectra can be due to several reasons. The clustering of the solute molecules can change the spectral responses, illustration of such changes was main objective in the study presented in Paper V. However, the temperature changes can cause variation to spectral responses as discussed in Chapter 6.1. In addition, the solvent composition gives its contribution to the spectral responses. The spectrum is, thus, always the combination of the components and also conditions of the system. To minimize the effects from other sources than solute molecules careful variable selection was performed. Several different spectral ranges were tested and the spectral range from 1000 to 1700 cm^{-1} was found to give the most apparent systematic change in the responses as the nucleation proceeded. This spectral range exhibits the responses from sulfathiazole and 1-propanol concentration but have neglible temperature effects.

In this case the different polymorphs were resulted from crystallization from different solvent compositions. As was presented in the Chapter 8.3 water was used as an background in sulfathiazole crystallization experiments and also in the results presented in Paper V. Thus, different solvent compositions can give different contributions to the measured spectra and influence of the solvent composition cannot totally be excluded. If we assume, that the solvent composition gave a response to the spectrum and did not have interations with other components in the system, the effect of solvent could be excluded by subtracting the solvent spectrum from the spectra measured from the process. This assumption is not exactly true since there usually is responses from the interactions present in the spectrum. However, by subtracting the solvent spectrum from the measured spectrum, it can, at least, be estimated whether or not responses from solvent itself result in a major effect on the result. Subtracting the solvent spectrum from the data did not considerably change the result in this case, however. At least, it can be said, that the solvent component solely did not cause the systematic change observed in this dynamic system, but also the other ingredients were involved in these changes.

This technique could be used in the controlling of the polymorph formation, an issue which is currently of great interest within PAT. If the nascent polymorph could be predicted before nucleation, cooling conditions could be changes in order to drive the process in the direction of the desired product structure and thus prevent the formation of the undesired polymorph.

The results presented here can only be considered as preliminary, and testing of the method with different solute-solvent systems is needed to prove the wide applicability for the polymorphic form prediction. A system where different polymorphs were crystallized from one solvent using different cooling conditions would be beneficial. The experiments presented in this study did not result in clearly different polymorphs from one solvent by changing the cooling mode as will be presented in Chapter 9.7.2. In addition, the samples should be taken from the solid phase right after the nucleation process to truly evaluate the structure of the crystals formed in primary nucleation.

ATR-FTIR might not be the best possible technique for this type of analysis, as the measurement area is rather small and restricted to immediate interface of the ATR accessory and the sample. In addition, the phenomenon which is measured, i.e., the differences in the solute molecule clustering might not be the primary information obtained from the IR spectra. In-situ Raman spectroscopy would perhaps be a better suited technique, since the clustering of the molecules can affect to the scattering effects and also the result is obtained from deeper in the sample than what is possible using the ATR-FTIR technique. Raman spectroscopy has already been tested to some extent as discussed in 5.1.1 for monitoring polymorphic systems, but the data was not evaluated using the MSPC technique.

9.4 Calibration modeling routine for predictive models: solute concentration prediction using ATR-FTIR and polymorphic composition prediction from powders using DRIFT-IR

9.4.1 Variable selection

Variable selection improved the calibration model in terms of increased R^2 value and decreased the error of prediction of the calibration and the test sets in C15 concentration prediction model. The wave number range from 1000 to 2000 cm^{-1} gave the best predictions of the calibration and test sets in this case. The RMSEP value for the test set decreased 15% when the selected subset was used instead of total measured spectrum in C15 system. For the model derived for sulfathiazole concentration prediction, there was not clear improvement in R^2, RMSEC, and RMSEP values by variable selection. (Paper I). There can be several possible reasons for this result. The C15 system is simpler than the sulfathiazole system in many ways as was explained in Chapter 9.1.1. Therefore, the obtained spectrum is less complex for the C15 system than for sulfathiazole system. Therefore, the correlation of the spectral variables to the solution concentration is clearer in C15 case than in sulfathiazole case.

Secondly, for the C15-toluene system, the responses appearing from the solute C15 were more intense and precise than for the sulfathiazole-water-1-propanol system, which Figure 9 clearly shows. Figure 9 shows that these responses are sensitive for the C15 concentration, which appears in the wave number range selected to the calibration model from (1000 to 2000 cm^{-1}). When the spectral responses, which correlate to the modeled variable, are intense and precise, these variables can most likely produce a stable and precise model. For the sulfathiazole system, such intense and precise bands indicating the sulfathiazole concentration do not exist and in this case, it was not possible to find a small set of variables, which properly explain the variation in **y**.

The temperature was included in the **X** variables and in addition, for sulfathiazole the solvent composition variable was included. The temperature and solvent composition variables were centered and variance was scaled to the level of spectral responses. However, including the temperature variable in the **X** matrix did not significantly improve the model performance. Actually, OSC filtering seemed to remove at least some of the temperature induced spectral variation (Chapter 9.4.3).

When PLS calibration models were derived to predict the polymorphic composition of powdered sulfathiazole samples in Papers II and VI, the model performance was clearly improved when the proper subset of the variables was selected. Selecting the wave number range from 3500 to 2500 cm^{-1} resulted in 13% smaller RMSEP values than when the model was derived using the whole spectral range. This wave number range selection from 3500 to 3250

cm^{-1} was obvious as there was a visible variation in the spectrum due to different polymorphs (Figure 11).

9.4.2 MSPC charts and sensitivity analysis in data quality evaluation

In this study, MSPC charts were found useful in evaluating the quality of the measurement data e.g. calibration data. MSPC analysis was applied to all of the samples in order to have preliminary evaluation on the sample quality for outlier detection. (Papers I, IV, VI)

The results revealed that abnormal or extreme samples within the model space could be pointed out from T^2 chart and those in the residual space from Q chart. Contribution charts provided information on the variables that cause the particular sample to an outlier. There are several possible reasons for judging the sample as an abnormal one. Figure 15 shows example MSPC charts of the sulfathiazole calibration measurements done with ATR-FTIR. This example presented in Figure 15 is from the calibration set of sulfathiazole concentration measurement using ATR-FTIR. Corresponding MSPC charts for C15 calibration measurements using ATR-FTIR are presented in Paper I and MSPC charts for the polymorphic composition of the powdered sulfathiazole samples using DRIFT-IR is presented in Papers II and VI.

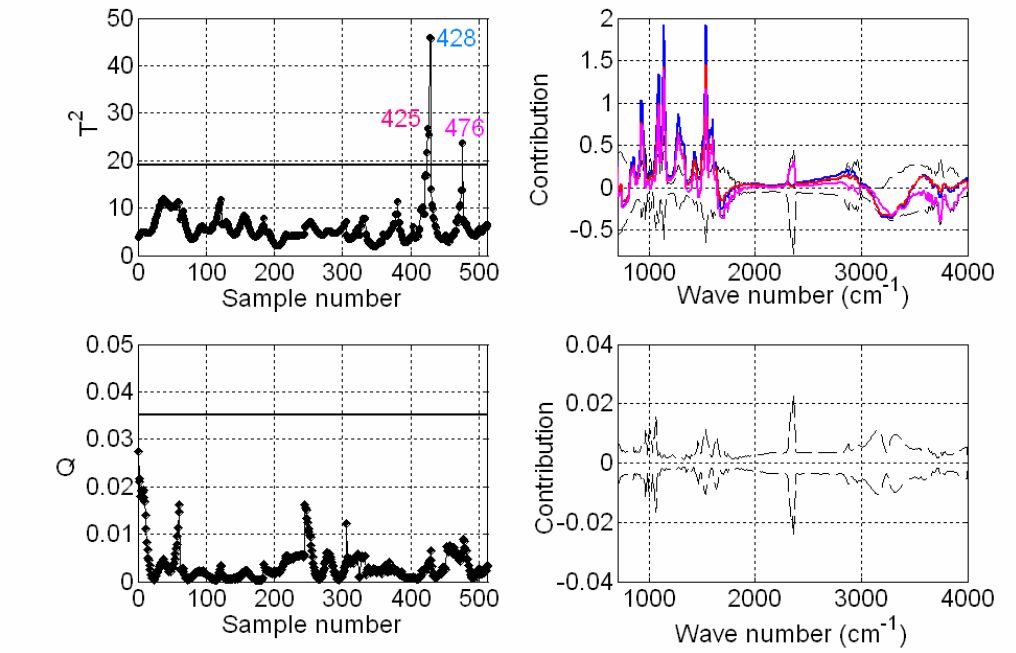

Figure 15 MSPC charts for the ATR-FTIR calibration measurements of sulfathiazole. a) T^2 chart, b) T^2 contributions, c) Q chart, d) Q contributions. The horizontal lines in a) and c) represent 95% confidence limit for the T^2 and Q values and the dashed lines in b) and d) represent the 95% confidence interval.

Baseline problems can cause samples to be abnormal in MSPC charts and these are observed several different ways: The baseline can exhibit otherwise normal baseline at the wrong level. Paper I presents an example of this behavior. Another alternative is that the baseline exhibits a different shape than it should do as presented in Paper I and in a small scale in Figure 15. In these cases, the baseline of an abnormal sample exhibits a linearly increasing trend. In Paper II, the baseline of an abnormal sample resembles a shape of a parabola.

The abnormality can exist by the contamination of the sample, and in this case, additional bands appear in the spectrum. The contribution charts can be applied to detect contamination of the samples. An illustrative example on such detection is in Paper I, where abnormally high CO_2 level together with abnormal baseline caused the sample to be an outlier and it shows in the contribution charts. An exceptionally high noise level can also cause the sample to exhibit a different structure than the other samples. MSPC charts are, therefore, powerful in determining also possible errors within the measurement procedure.

The samples illustrated in Figure 15 exhibit extreme samples and these have high T^2 contribution values in the wave number range where the bands related to sulfathiazole concentration appear. These samples should not be removed, since they are actually essential samples that should be included in the model to ensure that the whole range of variation is included in the calibration. Therefore, previous knowledge on the data, i.e., temperature, dissolved solute concentration, and solvent composition of a particular sample, is essential in deciding whether the sample is an outlier rather than an extreme sample in the data space.

MSPC charts were also found suitable for testing that the new samples represent similar structure to the samples used in the calibration (Papers I, II, VII). It is also important to make sure that the structural part of the data does not significantly change when preprocessing methods are applied. This also could be done by analyzing the filtered data using MSPC to compare the structure of the filtered data to the structure of non-filtered data (Papers I, II, VII).

As was mentioned in Chapter 7.3, only the **X** space can be evaluated using MSPC charts, and therefore, this method cannot point out the samples that have, e.g., in the **y** space as long as the variation in **X** is similar to the rest of the data and is within the range of variation. Studentized residuals together with leverage values were used to evaluate the samples in **XY** space. Figure 16 shows the studentized residuals and leverage values for the same data set as was used in the previous MSPC analysis. The sample numbers in MSPC analysis result in Figure 15 correspond to these presented in Figure 16.

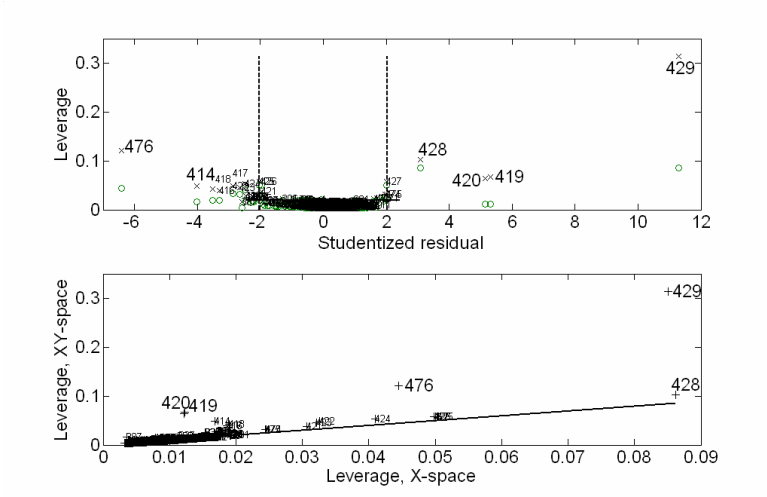

Figure 16 Studentized residuals and leverage values for the PLS model for sulfathiazole concentration prediction from ATR-FTIR data. Vertical lines in the upper image represent 95% confidence limits for studentized residuals. Circles in the upper image represent the leverage values in **X** space and x's represent leverages in the **XY**-space.

The PLS model from which the studentized residuals and leverage values were calculated and illustrated in Figure 16 was derived from the same data used for the MSPC charts in Figure 15. The same samples, which exceeded the 95% confidence limit in the T^2 chart, exist as outliers in the sensitivity analysis charts. The samples 429 and 476 show high leverage values in the **X** space and those increase significantly, when **Y** space is included. There are also some samples, e.g. 420 and 419 exceeding the 95% confidence limit in the Studentized residuals, which did not exceed the 95% confidence limits in T^2. Studentized residual- and leverage analyses point out also the errors in **Y** space, such as weighing errors or other probable inaccuracies in the sample composition. Therefore, both **X** space analysis and **XY** space analysis are required for analyzing the measured data and its usability for modeling purposes.

9.4.3 OSC filtering

The number of the OSC components and number of the PLS components were selected using RMSEP values of the external test set, thus the dimensionality of the model was evaluated based on the true performance of the model. The summary of the results for the different models derived from non-filtered, OSC_{SW} and OSC_{AH} filtered data are presented in Table 5. The detailed description of the results and for solute concentration prediction models using ATR-FTIR are presented in Paper I.

Table 4 The PLS models derived from raw spectral data and OSC_{AH} and OSC_{SW} filtered data

Calibration model (material/ technique)	**R^2Y for one component PLS model**			**Number of components selected into final model (OSC/PLS)**			**R^2Y for the selected model**			**RMSEP for the selected model**		
	Raw data	OSC SW	OSC AH	Raw data	OSC SW	OSC AH	Raw data	OSC SW	OSC AH	Raw data	OSC SW	OSC AH
Stz/ ATR-FTIR	42.0	91.7	59.6	- / 5	2 / 2	3 / 2	99.7	99.9	99.8	0.57	0.55	0.55
C15 / ATR-FTIR	66.4	99.9	94.7	- / 5	2 / 2	3 / 4	99.9	99.9	99.9	0.79	0.42	0.33
SUTHAZ 01/ DRIFT-IR (II)	56.9	76.4	70.4	- / 3	1 / 2	2 / 2	79.8	86.3	84.8	0.06	0.06	0.07
SUTHAZ 02/ DRIFT-IR (II)	46.7	76.7	62.9	- / 3	2 / 2	3 / 2	78.8	99.8	89.7	0.07	0.06	0.08

For the polymorphic composition prediction for powdered material from DRIFT-IR data, two different models were derived to predict SUTHAZ01 and SUTHAZ02 concentrations with separate models, of which the detailed discussion of the results is presented in II. Two separate models gave better predictions of the calibration and test sets overall than did the single model predicting both concentrations. This result is in line with the commonly suggested opinion presented in the literature that usually separate models perform better than one single model predicting several **y**s. Separate models probably performed better than the single model, because the y-variables were not fully correlated and also other than modeled variation was present in the spectra. In addition to that OSC filtering could not have successfully performed with several **y**s.

The number of components used in the final models differs from those presented in Paper I, where number of OSC_{AH} components included in the sulfathiazole model was 6 and in the C15 model 5 while number of PLS components in both models from OSC_{AH} filtered data was 2 in that case. It is true that RMSEP values for those models were slightly better than corresponding ones presented in Table 5, being 0.32 and 0.51 for C15 and sulfathiazole respectively.

Experience showed, however, that simpler models presented in Table 5 performed better in the long run. The sulfathiazole model presented in Table 5 was used in calculation of the results of Papers III and VI. Overall, making model more complex and including more OSC and/or PLS components into the model could slightly improve further RMSEP values, but the improvement is just on the nominal scale and is not significant for the true prediction of the new unknown samples. In addition, a very complex model can suffer from over fitting problems. Therefore, one should select the simplest possible model which gives the predictions for the test set, which are not improved significantly by adding more OSC or PLS components into the model. The

models presented in Table 5 were selected based on this principle and thus do not represent minimum RMSEP values, but the best possible combination.

The effect of filtering can be observed when one component PLS models are derived from non-filtered and filtered data. However, the total number of components when data are filtered is actually two. Therefore, the performances of the models in terms of complexity cannot be evaluated. However, it can be seen whether the filtering removes unnecessary variation. RMSEP value for one PLS component model decreased clearly by OSC filtering. For PLS model with one component $R^2\mathbf{Y}$ value was increased by OSC filtering in the reported cases presented in Table 5. Therefore, OSC filtering clearly removed redundant variation from the data. However, the overall complexity of the model did not decrease by filtering because the included OSC components also increase the complexity of the model. The $R^2\mathbf{Y}$ value of the final models increased clearly by OSC-filtering, but the effect on the RMSEP value was of minor scale. Improvement in the predictions was therefore not significant, and, thus, need for OSC filtering cannot be justified from the prediction point of view. These results are in correspondence with the previous studies presented in literature as mentioned in chapter 7.1.6.

On comparison of the two OSC methods used, the first OSC_{SW} component removes more variation than does the first OSC_{AH} component, which results in higher $R^2\mathbf{Y}$ values after the first PLS component for the model for OSC_{SW} filtered data than those for the model from OSC_{AH} filtered data. To perform appropriate filtering number of OSC_{SW} components needed in filtering was lower than number of OSC_{AH} components needed for filtering. Based on the root mean squared error of calibration (RMSEC) values, which were calculated for the calibration set, the models derived from OSC_{SW} data showed clearly lower values than the models from OSC_{AH} filtered data. The slightly best model in terms of predictive ability was the model derived from OSC_{AH} filtered data for solute concentration prediction models (ATR-FTIR data) and OSC_{SW} for polymorphic composition prediction (DRIFT-IR data). Differences in the RMSEP values for differently treated models were of a nominal scale, however. For DRIFT-IR, the OSC filtering method performed better than did the tested MSC and SNV methods.

When comparing the two solute concentrations prediction models the $R^2\mathbf{Y}$ value increased more after the first component for the C15 model than for sulfathiazole model. In addition, the number of OSC_{AH} components needed in the model was smaller for the C15 model than the sulfathiazole model. RMSEP values overall were lower for the C15 model, than for the sulfathiazole model, even if the concentration range measured was almost the same. These differences can be because the C15 system and obtained spectral data was clearly simpler than for the sulfathiazole system. These differences in the complexities are closely described in Chapter 9.4.1. In sulfathiazole case, the **y** correlates to very small-scale variation within the

spectra, and therefore, it was self-evident that first two components did not explain a lot of variation in **y**. In the case of C15, the variation in the spectra was mostly due to the C15 concentration.

Figure 17 illustrates the effects of OSC_{SW} filtering in the C15 system and in the sulfathiazole system.

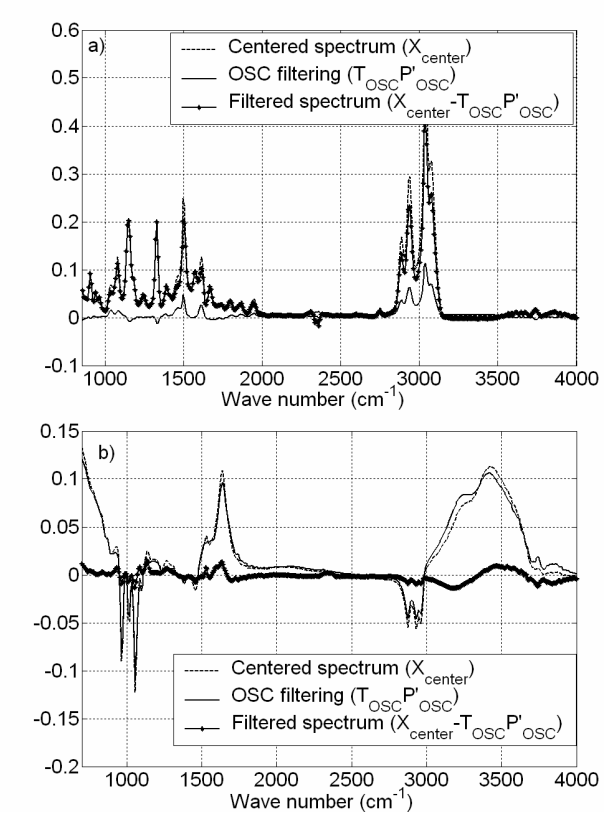

Figure 17 Effect of OSC_{SW} filtering on the measured ATR-FTIR spectrum. a) C15 dissolved in toluene, b) sulfathiazole dissolved in 50/50 w-% mixture of water and 1-propanol. Filtering included 2 OSC_{SW} components in both example cases

Figure 17 shows that for C15 case, only a small amount of data was removed by filtering and the filtered spectrum is quite similar with the centered spectrum. In sulfathiazole case, on the other hand, $\mathbf{T_{OSC}P'_{OSC}}$, i.e., the data removed from centered spectrum resembles very much the centered spectrum. Resulting filtered spectrum in sulfathiazole case has very small-scale variation throughout the measured spectral range. These results are in line with the knowledge on the nature of the data, which was discussed closely in Chapter 9.1.1. Simply, the most variation in C15 spectrum result from the variation in the C15 concentration, i.e., this is correlated to **y**. In addition, the temperature effects on the spectrum in C15 case were of very

minor scale. Consequently, in C15 case there should not be very much **y** orthogonal variation to be removed by OSC filtering. In the sulfathiazole case, the most variation in the spectrum is due to the solvent composition changes and temperature effects, and only a very small variation appears from the sulfathiazole, i.e., modeled **y.** Thus, there should be a major amount of y independent variation present in sulfathiazole case. The variation removed ($\mathbf{T}_{OSC}\mathbf{P}'_{OSC}$) from the centered spectrum in sulfathiazole case shows clearly the spectral ranges where the most variation due to solvent composition and temperature changes were observed to be shown in Chapter 9.1.1.

Based on results in Figure 17, it seems that OSC filtering would be beneficial in the sulfathiazole case, since there is a lot of **y** orthogonal variation present. In the C15 case, however, the amount of **y** orthogonal variation in the spectrum is so small that it is questionable whether the use of OSC filtering is justified or not. The final selection of the modeling procedure: regular PLS or PLS+OSC, should be based on the true performance of the model, i.e., the complexity and true predictive ability.

Filtering removed the baseline effects and filtering emphasized the relevant bands of the modeled phenomenon, especially in the sulfathiazole case where the small bands in the spectral range from 1000 to 1500 cm^{-1} still exist in the filtered data while other intense bands were almost completely removed by filtering. Magnitude of the regression coefficients of the models derived with and without OSC treatments presented in Paper I show which variables are the most important in the model. The most relevant regression coefficients in all different models derived from differently pretreated data are in accordance with the spectroscopic knowledge of the modeled variables.

For selecting the best combination of OSC and PLS components or the issue whether OSC filtering is needed or not, the automated routine to calculate all the possible and sensible model combinations and the criteria for the selection of best performing model is needed. RMSEP value of an external test set was found to be an appropriate way of selecting the complexity of the model.

9.4.4 Model performance validation using solubility measurements

An additional validation with solubility measurement provided a tool to evaluate the flexibility of the solute concentration prediction model to be used for supersaturation predictions from crystallization process. Solubility measurements done using ATR-FTIR were in good agreement with gravimetric solubility measurements and gave an indication. The RMSEP for the solubility measurements between the ATR-FTIR result and gravimetric result was 1.2 and

1.4 for C15 and sulfathiazole cases respectively, when the best performing model based on RMSEP value was used in predictions of new unknown samples.

The solubility concentration agreement between these two methods was poorer for the solubilities at the highest and the lowest ends of temperature scales. In principle, the ATR-FTIR measurement gave higher solubility concentrations at the low end of the temperature range and lower solubilities at the high end of temperature range than did the gravimetric method. The error is a result from both analytical methods used in solubility measurements. (Paper I)

9.5 PLS model performance for solute concentration prediction using ATR-FTIR

The mean RMSEP value for the derived models was 0.33 for C15 model and 0.55 for sulfathiazole model as presented in Table 5.

There are several different aspects to be considered when evaluating the performance of the spectroscopic measurements and quantitative multivariate analysis. In the spectroscopic measurements there are questions of the limit of detection, resolution in the y-direction and short term and even more importantly long term stability of the spectrometer and accessory, which affect the reliability of the measurements and the performance of the derived model. Also mechanical changes in the measurement accessory can alter the performance of the measurement. The process state changes from the conditions where the calibration set have been measured can naturally cause the predictions of unknown samples to be inaccurate, if the model is not derived properly.

The changes in the process conditions can also affect the model usability for the measurements. When the highest cooling rate of 27.5°C/h was used in the case of sulfathiazole, the predicted concentration level seemed exceptionally high. The crystallizing system was probably in the labile region thoroughout the process, thus new nuclei could have been formed constantly. In addition, the concentration level was probably too far away from the conditions where the calibration measurements were done.

For the sulfathiazole crystallizations different impellers were used, and it was found that the model gave unreliable predictions of the samples measured from process when the measurements were conducted under clearly different mixing conditions than the mixing conditions used in the calibration measurements. These problems could have possibly been overcome by applying appropriate calibration transfer methods, which are suggested to be tested in future studies.

The disturbances caused by the solid material or bubbles attaching onto the probe were not observed significantly. However, if the mixing intensity was very low, some bubbles tended to

attach onto the probe and they were not wiped out by the stream caused by mixing, as seemed to be case when higher mixing intensities were used.

The spectra measured from the system, where there was a very small amount of sulfathiazole dissolved were evaluated to estimate the limit of detection. It was found that in the sulfathiazole case reliable detections were obtained above concentrations of 1 g solute/100 g solvent. For the sulfathiazole system, the measurements from the solvents in which the solubility of sulfathiazole was extremely low, especially water and 1-propanol (the solubilities at 85° 1.22 g solute/100 g solvent and 2.86 g solute/100 g solvent respectively), reliable concentration measurements during the crystallization processes could be done. For C15 the measurements below a concentration of 1 g solute/100 g solvent were not attempted.

The short term repeatability could be evaluated from the repeated measurements. Stability of the measurement device could be evaluated from prediction errors in the calibration measurements, more precisely from the errors in a test set, which was around 0.3-0.6 depending on the system in question (Table 5). The concentration range measured was from 1g solute/100 g solvent to 30g solute/ 100g solvent. When the crystallization measurements were repeated, there was always differences in the performed batch present, which did not imply only the performance of the predictive model, but also repeatability of the batches.

The long term repeatability cannot be explicitly defined, but the crystallization experiments of sulfathiazole were performed and concentration levels measured during a 10 month period. During this period 37 experiments were monitored. Appropriate concentration predictions from crystallization process were obtained if the process conditions used were similar as were used in the calibration measurements or the solvents into which the solubility of sulfathiazole was reasonable. Any mechanical disturbance, e.g., change in the probe position caused also possibility for inaccuracies in predictions of the new samples . The throughput from the light source via the ATR-probe to the detector was lowered slightly, approximately 7%, during the ten month measurement period of sulfathiazole crystallizations.

9.6 PLS model performance and other multivariate methods for the characterization of the polymorphic composition using DRIFT-IR

The DRIFT-IR provided a rapid way to analyze the powdered samples. However, sample preparation was needed. In order to minimize the CSD and shape effects, the samples had to be grinded. Non-grinded samples resulted in bad quality spectra, which were not proper for multivariate analysis. On the other hand, samples for XRPD analysis had to be grinded as well. The specific area in the DRIFT-IR measurement is rather low, which reduces representativity of the sampling.

The PLS models for predicting the polymorphic composition of two sulfathiazole polymorphs SUTHAZ01 and SUTHAZ02 reported in Paper II gave satisfactory results with a RMSEP value around 0.06-0.07 (Table 5) depending on which polymorph was modeled and what preprocessing method was used. The concentration range measured was in weight fraction units and varied from 0.1 to 0.8.

The uncertainty level of the quantification results from XRPD measurements were up to ±10%. When XRPD results were used as response variables, the predictive ability suffered from high uncertainty in measured **y** values and not very certain predictions of the unknown samples from derived PLS model could not be expected. Preparation of the physical calibration set is strongly suggested when pure polymorphs are available. The PLS model derived this way without a true physical calibration set can only be considered as a method which confirms the results obtained from the XRPD measurement. Another reason for the inaccuracies in the predictions of the test set can be that the log (1/R) transformation does not represent truly linearized spectrum. This issue as was discussed in Chapter 6.3.

The qualitative analysis methods reported in Papers II and VI summarize that the scatter plots from PCA well separated the samples representing mainly different polymorphs. In addition, PCA derived MSPC charts visualize samples that represented different structure from the other samples. The samples representing undesired polymorphic purity could be pointed out. The contribution plots could serve the purpose of detecting what is the reason for a particular sample being different from the calibration set. In this study, the SIMCA analysis was found as the best way of classifying the samples. It was especially useful to use Cooman's plot to classify samples representing either of the two polymorphs either of which the samples mainly consisted. However, it was also possible to identify the samples that did not belong to either of the groups.

9.7 Supersaturation measurement results and supersaturation effect to product quality

9.7.1 Nucleation moment

For the unseeded crystallization process of sulfathiazole, different cooling conditions affected the nucleation process as the width of the metastable zone increased with increasing cooling rate, which is illustrated for sulfathiazole crystallizations in Figure 18. The width of the metastable zone became wider as the cooling rate increased being at its narrowest at 5°C (80-75 ºC) with a 3.9ºC/h cooling rate and widest at 8.5ºC (80-71.5ºC) with a 27.5ºC/h cooling rate. The concentration level remains higher after nucleation as the cooling rate increases. These results are in correspondence with previously reported results by other research groups as was discussed in Chapter 4.2.1.

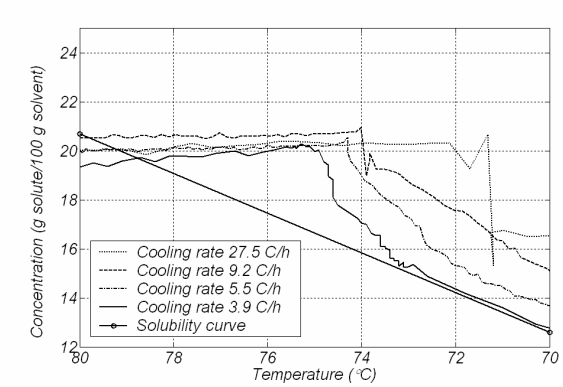

Figure 18 Concentration profiles from sulfathiazole crystallization from 50/50 w-% mixture of water and 1-propanol using different constant cooling rates. Supersaturated state prior to nucleation and the nucleation moment are presented.

Also different cooling modes cause the nucleation process to proceed differently in sulfathiazole crystallizations as illustrated in Figures 19 and 20. Figure 19 shows that the nucleation seems to occur at different temperatures when different cooling modes are applied. The different cooling modes were switched on at the temperature where the nucleation was observed using 9.2°C/h linear cooling rate as was described in Chapter 8.2 and in Paper III.

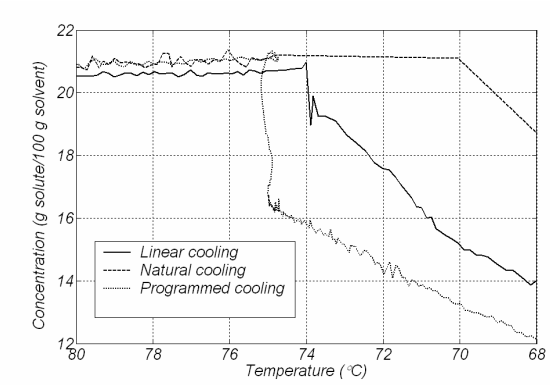

Figure 19 Concentration profiles from sulfathiazole crystallization from 50/50 w-% mixture of water and 1-propanol using different cooling modes. Supersaturated state prior nucleation and nucleation moment.

Thus, the nucleation should have occurred right after switching on the specific cooling mode, which actually is seen in Figure 20. With programmed cooling the temperature decreases very slowly and with natural cooling the temperature decreases rapidly, which causes the differences in the observed nucleation temperature even though the nucleation occurs at same time.

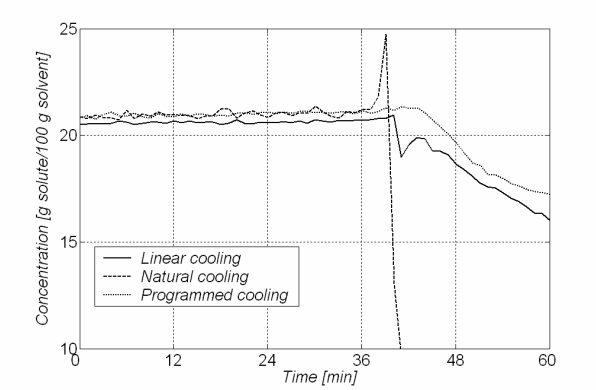

Figure 20 Concentration profiles from sulfathiazole crystallization from 50/50 w-% mixture of water and 1-propanol using different cooling modes. Supersaturated state prior nucleation and nucleation moment are presented.

The concentration profiles after nucleation are different with different cooling modes. As the temperature decreases rapidly with the natural cooling mode, the nucleation process is drastic and thus the concentration decreases fast. With programmed cooling, the temperature decreases very slowly in the beginning, and thus the nucleation moment is less drastic than with natural cooling and the concentration rate decreases slower than with the natural cooling rate. The concentration profile right after the nucleation from the linear cooling mode is in between of those of the other two cooling modes.

Figures 21 and 22 illustrate the nucleation processes with different cooling modes in the C15 case.

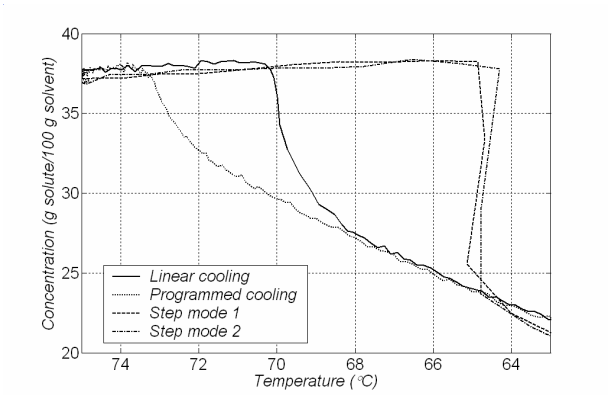

Figure 21 Concentration profiles as a function of temperature from C15 crystallization toluene using different cooling modes. Step modes 1 and 2 refer to the situation where the temperature of the coolant was set to 40°C and 25°C respectively. Supersaturated state prior to nucleation and the nucleation moment

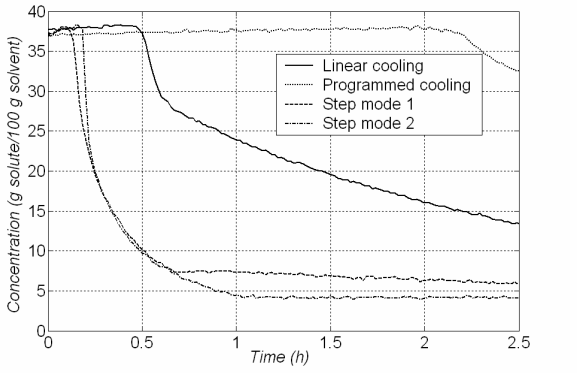

Figure 22 Concentration profiles as a function of time from C15 crystallization toluene using different cooling modes. Step modes 1 and 2 refer to the situation where the temperature of the coolant was set to 40°C and 25°C respectively. Supersaturated state prior to nucleation and the nucleation moment.

The nucleation moment occurs at a different time in the C15 experiments performed with different cooling modes. In this case, the specific cooling mode was switched on right after the supersaturated state was reached. Therefore, the nucleation moment was driven by the instantaneous cooling rate of the specific cooling mode. For example, in stepwise cooling mode, the cooling rate in the beginning of the process was extremely high, and thus the nucleation takes place at lower temperatures. In the programmed cooling on the other hand, the cooling rate was extremely slow in the beginning of the batch and thus the nucleation occurred at the higher temperature than with other cooling modes applied in C15 crystallizations. This result is also in correspondence with a previously presented result, which shows that increase in the cooling rate increases the width of the metastable zone. In time scale, the nucleation process occurred first with the step mode and last with the programmed cooling. Figure 22 illustrates also, that the true crystallization process starts up almost 2 hours before in stepwise process than in programmed cooling experiment.

The seeding caused the nucleation to occur immediately after applying the seed crystals into the system regardless of the cooling mode. Therefore, with seeding, the supersaturation level remained very low also at the nucleation moment and the nucleation occurred exactly at the same moment and at the same temperature in all seeded experiments with different cooling modes.

9.7.2 Cooling mode effect on supersaturation level and product crystals

The cooling mode effect on the supersaturation level and quality of the product crystals in sulfathiazole crystallizations is described in Papers III and VI and the result are briefly summarized in this Chapter. The overall concentration level increased when the cooling rate

was increased. This is an expected and logical result and can be explained by the mass transfer rates of the solute molecules from solution phase on existing crystals (Chapter 4.3.2)

Cooling rate

With the highest cooling rate, 27.5ºC/h, the predicted supersaturation was extremely high. The process conditions differed both chemically (the highest supersaturation) and physically (new crystals forming throughout the process) from the calibration measurements when the highest, 27.5ºC/h, cooling rate in sulfathiazole crystallizations was used. Therefore, the predictions from that crystallization experiment were the most uncertain of all crystallizations. The concentration level was high enough to exceed the metastable limit throughout the crystallization with a 27.5ºC/h cooling rate. This caused the solution to cool down faster than the solute could transfer onto existing crystals, and consequently the spontaneous nucleation took place throughout the process. The level of primary nucleation during the crystallization was most likely small with cooling rates from 9.2 to 3.9ºC/h.

The differences in the concentration levels resulted in changes in the outcome of product crystals. The widest size distribution and the most irregularly shaped crystals were obtained with the highest cooling rate, which implies that the massive nucleation occurred throughout the crystallization process. With slower cooling rates the resulting size distributions of the product crystals became slightly narrower, but overall the size distributions were rather wide. The largest size found in all samples was approximately equal. Therefore, it seems, that attrition nucleation in crystallization of sulfathiazole had a considerable effect. Similar results and actually the concept of maximum obtainable crystal size have been introduced previously in literature and are referred to in Chapter 4.3.3. As the cooling rate was lower than 27.5°C/h, produced crystal shapes changed significantly. The larger crystals were mostly long and thin plates. Small crystals were mostly more rounded than the larger ones. The differences between the shapes of the crystals produced with different the roundness increased slightly with decreasing cooling rate. These results could be explained by the breakage of the large crystals, which was caused by the mixing during crystallization.

Cooling mode

The supersaturation level changed remarkably, when applying different cooling modes. For sulfathiazole crystallizations, the programmed cooling caused a very low supersaturation level at the beginning of the crystallization process, and high supersaturation towards the end of the process. The programmed cooling did not maintain a constant supersaturation level, which has been an objective of that cooling mode. Thus, the simplifying assumptions made to obtain the equation for the programmed cooling profile were not appropriate for the case of sulfathiazole.

Concentration profiles from crystallization with natural cooling show that the concentration level decreased rapidly at the very moment the temperature of the cooling medium was set to a constant value of 25°C. The suspension cooled down rapidly (in 1.5 h) and practically all the crystals were formed in a few minutes. For the rest of the process (for 4.5 h) the system was only mixed in constant temperature and practically no crystal growth existed.

Only small scale differences could be observed in the size distributions between the different samples obtained from the sulfathiazole crystallizations the using different cooling modes. The narrowest size distribution was obtained when the crystallization was carried out using natural cooling and the widest distribution results from the experiments with the linear cooling profile. The crystals obtained from natural cooling mode experiments suffered probably from two processes, which made the size distribution narrower: mixing for 4.5 hours at constant temperature caused significant attrition of the large crystals and simultaneously the Ostwald's ripening caused dissolving of the very small crystals. The size distribution of the crystals obtained by the programmed cooling profile is positioned in the middle of the distributions obtained by the other two cooling modes. As the programmed cooling did not maintain the supersaturation constant throughout the process, the programmed cooling did not result in the narrowest size distribution. There was relatively large amount of small crystals in samples from programmed cooling experiments, which probably nucleated in the end of the batch where the concentration level was rather high. The crystals produced with the programmed cooling profile were clearly more elongated than the crystals obtained by linear or natural cooling. The elongated shape crystals from controlled cooling experiments were probably formed in the final parts of the batch and they did not suffer from attrition The crystals produced by natural cooling had the largest roundness values, which could have been due to the mixing which grinds the crystal edges smoother.

The different cooling modes in the C15 crystallizations (Paper IV) the step modes resulted in very high initial supersaturations, which rapidly decreased after nucleation and reached the equilibrium within the 1 h after the nucleation, while the total batch time was 5 h. Linear cooling resulted in the moderate supersaturation level throughout the process. In the controlled cooling, the nucleation took place after two hours of experiments. There was an increase in the supersaturation level in the very end of the batch with programmed cooling was used. The seeding caused the overall supersaturation level slightly to decrease in all C15 crystallizations compared to the unseeded crystallizations, but the most effective change was the differences in the onset processes of the crystallization experiments. The narrower size distributions of C15 were obtained when seeding was used, as can be expected by the results presented in the literature. With controlled and linear cooling modes, a relatively large amount of fines could

have been produced by the secondary surface nucleation at the end of the batch, as the supersaturation level was found to increase in crystallization with controlled and linear cooling modes towards the end of the batch.

Polymorphic form

Theoretically, the level of supersaturation could influence on the polymorphic outcome of the crystals because the solubilities of the specific polymorphs to a particular solvent are different. In principle, if the cooling takes place at a concentration level that lies between the solubility of two polymorphs, the polymorph with the lower solubility should appear. The results reported in Paper VI show, that the polymorphic composition of the product sulfathiazole crystals seemed not to depend on the cooling rate used although the concentration levels in the crystallization processes were remarkably different. One possible reason for this could be that the relative differences in the solubilities of the different sulfathiazole polymorphs for in the solvents used are rather small. In addition, there can be several other things in the crystallization process, which can drive the polymorphic composition of the product crystals, one being the solvent composition, which seemed to be the dominant one in this study. In addition, according to the Ostwald's law, metastable polymorphs can undergo solution mediated phase changes from a less stable polymorph to a more stable one.

10 CONCLUSIONS AND FUTURE WORK SUGGESTIONS

The aim of the work was to explore different ways to utilize IR spectroscopy together with multivariate data analysis tools in real time monitoring of the solution phase in the batch cooling crystallization process and in off-line characterization of the polymorphic purity from crystalline samples.

The three main approaches in real time monitoring of the solution phase using IR-spectroscopy and different multivariate methods include: 1) Batch-to-batch variation analysis from a cooling crystallization process using PARAFAC. 2) Multivariate monitoring of the on-set of the unseeded crystallization process using MSPC charts. The on-set of the crystallization is important in terms of the number of primary crystals in the system and in the polymorphic form of the product crystals. 3) Calibration model building for prediction of the solute concentration and supersaturation level within the crystallization process. The supersaturation level is the driving force of the crystallization and thus a very important process parameter affecting the product quality.

Polymorphic composition is an essential quality property of the product affecting further the processability and usability of the crystalline product. Off-line characterization of the polymorphic composition of crystalline samples was done using DRIFT-IR technique. PCA, MSPC, and SIMCA methods were tested in qualitative analysis of the crystalline samples. The quantification of the polymorphic composition from crystalline samples using a predictive PLS model was derived.

Analysis of the batch-to-batch variation using a PARAFAC model provided an interesting insight into the phenomena that can be viewed from in-situ measured spectral data. Not only the differences in chemical state but also some changes in the physical conditions inside the crystallizer could be visualized with PARAFAC modeling. To evaluate the causes of the batch-to-batch variation requires expertise knowledge on the spectral analysis; however, interpretation is not a straightforward task. Different types of data should be measured and the data processed using N-Way methods for obtaining a deeper insight of the crystallization process. In order to evaluate the crystallization batches applicability of different N-Way modeling approaches and diagnostics should be tested.

Monitoring the state prior to nucleation and during the nucleation from ATR-FTIR together with MSPC charts provided a new way of observing the changes in the system where the concentration of the solute remains constant, but the molecules are arranging to be nucleated. The approaching nucleation could be predicted and an alarm criterion to be set up, which further enhances the control of the crystallization in the very beginning of the process. On the

other hand, the chemical state prior to nucleation could be evaluated, which opens up the potential possibility to predict a forming polymorph. If this is possible, perhaps the crystallization process could be controlled in such a way that the desired polymorph could always be obtained, which would reduce the number batches of unsatisfactory quality. However, the method requires that the number of samples measured in the beginning of the batch to be large enough to produce a stable model, and to obtain information from the truly transient state, the sampling interval should be therefore as short as possible. This is restricted by the fact that the average of several scans is needed in order to obtain a satisfactory signal to noise ratio of resulting spectra. In the future, more studies on predicting the formation of the polymorphs from various different solute-solvent systems will provide an insight whether this technique is widely applicable in control of polymorphic systems. In future studies, experiments where the crystalline samples are collected and analyzed immediately after the nucleation occurs would provide additional information on the relationship between spectral information obtained prior to nucleation and the polymorphic form of the formed crystals. In addition, the use of Raman spectroscopy for this purpose instead of ATR-FTIR would perhaps provide an even more powerful tool to predict the forming polymorph. The specific area of the measurement with the Raman technique can be significantly larger than with the ATR-FTIR technique, and simultaneous measurement of the solution and solute phases is possible.

Several different research groups have previously proved that the solute concentration measurement using ATR-FTIR and PLS modeling is a versatile technique. In this research, the calibration procedure including the data quality evaluation with MSPC and sensitivity analyses, data pre-processing using OSC filtering and variable selection, modeling, and model validation steps using RMSEP criterion was proposed. This routine resulted in reliable and stable calibration models for real prediction of the unknown samples. OSC filtering can be used for pre-processing of ATR-FTIR, as it seems to remove redundant variation from the measured data. OSC filtering did not significantly enhance the predictive ability of the model, however. The drawback of the calibration routine derived for this particular purpose is that the solute measurement from the crystallization takes place slightly out of the range of the calibration measurements: concentrations slightly higher and the solid phase present. Therefore, an additional validation using solubility measurements from suspension with ATR-FTIR to test the applicability of the model in true process conditions is recommended. The drawback of the method includes that even small changes in the process conditions, e.g., in mixing conditions seem to result in the calibration model not being adequate, and thus the model needs to be updated frequently. Especially, the suitability of the calibration transfer methods in updating the calibration models should be explored. In addition, the sampling issue should be carefully

considered in order to optimize the number of calibration samples needed for the stable calibration model. These improvements would make the calibration procedure more economical. Results showed that by measuring the driving force of the crystallization processes, evaluation of the process condition effects on the obtained product quality was possible. Therefore, a deeper understanding of the factors driving the quality of the resulting product was achieved.

All the applied methods, PCA, MSPC, PLS, PARAFAC were capable of detecting the on-set of the crystallization process. However, only the dynamic PCA derived MSPC diagnostics provided a tool to predict truly the nucleation moment and to evaluate the chemical changes in the solution phase prior nucleation. From the concentration predictions using PLS, observation of the differences in supersaturation profile due to different cooling policies were possible. Thorough the concentration profile the proceeding crystallization process could be evaluated and different batches compared which was not possible to this extent using the other tested multivariate tools. PARAFAC modeling on the other hand provided a method to evaluate the spectral ranges which differentiate the batches where the process conditions, e.g., solvent, cooling mode or impeller type changed. Process dynamics could not be evaluated using PARAFAC model.

ATR-FTIR technique together with multivariate methods proved, thus, to be a versatile technique for monitoring the solution phase. To obtain a full crystallization-monitoring scheme the different in-situ measurement techniques should be compiled. These additional techniques should include the method for monitoring size and shape of the crystals throughout the process using probes developed for that purpose. Additional chemical information from both solution and solid phase could be obtained, e.g., by using Raman spectroscopy. All these additional methods separately and especially when combined together would benefit from the use of multivariate methods to extract the relevant information from the enormous amount of data. In addition, the measurement based closed loop control of a crystallization process, which already has been proposed by few research groups, could well provide an enhancement to the control of crystallization processes. Overall, a combination of the analytical and data analysis techniques proposed here would well improve the control of a crystallization process and increase the understanding of the phenomena present in crystallization, which are the main objectives of PAT.

DRIFT-IR technique together with multivariate methods for characterization of the polymorphic composition of solid samples provides an additional technique to evaluate the purity of the crystalline samples by clustering the samples representing the different composition to different classes. The samples could be classified easily and rapidly based on

their quality. However, the technique cannot be used as the only technique to evaluate the polymorphic composition of the samples. In addition, to obtain truly quantitative information, i.e., the distribution of the polymorphs in a specific sample, the preparation of a calibration set from pure samples is strongly recommended. This does not only hold for the quantification of DRIFT-IR measurements, but also for other techniques used to measure the polymorphic composition of powdered samples.

References

Agatonovic-Kustrin, S., Rades, T., Wu, V., Saville, D., Tucker, I., G., 2001, Determination of polymorphic forms of ranitidine-HCl by DRIFTS and XRPD, J. Pharm. Biomed. Anal., 25, pp. 741-750

Andersson, C., A., 1999, Direct orthogonalization, Chemom. Intell. Lab. Syst. 47 pp. 51-63

Andrés, J., M., Bona, M., T., 2005, Analysis of coal by diffuse reflectance near-infrared spectroscopy, Anal. Chim. Acta, 535, pp. 123-132

Auer, M., E., Griesser, U., J., Sawatzki, J., 2003, Qualitative and quantitative study of polymorphic forms in drug formulations by near infra red FT-Raman spectroscopy, J. Mol. Struct., pp. 661-662

Axiom Analytical, 2005, Technical datasheets, [www.goaxiom.com/datasheets/dipper_ds.html], 15.9.2005

Barret, P., 2003, Selecting in-process particle size analyzers, Chem. Eng. Proc., 99, pp. 26-32

Barret, P., Glennon, P., 2002, Characterizing the metastable zone width and solubility curve using Lasentec FBRM and PVM, Chem. Eng. Res. Des., 80, pp. 799-805

Barret, P., Smith, B., Worlitschek, J., Bracken, V., O'Sullivan, B., O'Grady, D., 2005, Org. Proc. Res. Dev., 9, pp. 348-355

Beckmann, W., 2000, Seeding the desired polymorph: Background, Possibilities, Limitations, and case studies, Org. Proc. Res. Dev., 4, pp. 372-383

Bernard-Michel, B., Pons, M.N. and Vivier, H., 2002, Quantification, by image analysis, of effect of operational conditions on size and shape of precipitated barium sulphate, Chem. Eng. J., 87(2), pp. 135-147.

Bladgen, N., 2001, Crystal engineering of polymorph appearance: the case of sulphathiazole, Powder Tech., 121, pp. 46-52

Bladgen, N., Davey, R., J., Lieberman, H., F., Williams, L., Payne, R., Roberts, R., Rowe, R., Docherty, R., 1998 a), Crystal chemistry and solvent effects in polymorphic systems, Sulfathiazole, J. Chem. Soc., Faraday Trans., 94, pp. 1035-1044

Bladgen, N., Davey, R., J., Rowe, R., Roberts, R., 1998 b), Disappearing polymorphs and the role of reaction by-products: the case of sulfathiazole, Int. J. Pharm., 172, pp. 169-177

Blanco M., Coello J., Iturriaga, H., Maspoch, S., Pérez-Maseda, 2000, Determination of polymorphic purity by near infra red spectrometry, Anal. Chim. Acta, 407, pp. 247-254

Boque, R., Smilde, A., K., 1999, Monitoring and diagnosing batch processes with multiway covariates regression models, AIChe J., 45, pp. 1504-1520

Brittain, H., G., 1999, Polymorphism in pharmaceutical solids, 10th ed., Marcel Dekker Inc., New York

Bro, R., 1996, Multiway calibration, Multilinear PLS, J. Chemom., 10, pp. 47-16

Bro, R., 1998, Multi-Way analysis in food industry, Diss., 311p., Royal Veterinary and Agriculturary University Denmark

Calderon De Anda, J., A., Wang, X., Z., Lai, X., Roberts, K., J., Jennings, K., H., Wilkinson, M., J., Watson, D., Roberts, D., 2005a, Real-time product morphology monitoring in crystallisation using on-line imaging, AIChe J., in press

Calderon De Anda, J., A., Wang, X., Z., Roberts, K., J., 2005b, Multi-scale segmentation image analysis for the in-process monitoring of particle shape with batch crystallizers, Chem. Eng. Sci., 60, pp. 1053-1065

Calderon De Anda, J., Wang, X., Z., Lai, X., Roberts, K., J., 2005c, Classifying organic crystals via in-process image analysis and the use of monitoring charts to follow polymorphic and morphological changes, J. Proc. Contr., 15, pp. 785-797

Choong, K., L., Smith, R., 2004, Optimization of batch cooling crystallization, Chem. Eng. Sci., 59, pp. 313-327

Clark, R. N., Roush, T. L., 1984, Reflectance spectroscopy: Quantitative analysis techniques for remote sensing applications, J. Geophys. Res., Vol. 89, pp. 6329-6340

Colthup, N., B., Daly, L., H., Wiberley, S., E., 1975, Introduction to infrared and raman spectroscopy, Academic Press, New York

Coomans, D., Broeckaert, M., Derde, M., P., Tassin, A., Massart, D., L., Wold, S., 1984, Use of a microcomputer for the definition of multivariate confidence regions in medical diagnosis based on clinical laboratory profiles, Comp. Biomed. Res., 17, pp. 1-14

Costa, C., B., B., da Costa, A., C., Maciel Filho, R., 2005, Mathematical modeling and optimal control strategy development for an adipic acid crystallization process, Chem. Eng. Proc., 44, pp. 737-753

Czarnecki, M., A., Czarnecka, M., Ozaki, M., Iwahashi, M., 1994, Are the integrated absorption coefficients temperature dependent? FT-NIR study of the first overtone of the OH stretching mode of octanoic acid, Spectrochim. Acta, 50A, pp. 1521-1528

Davey, R. and Garside, J., 2000, From molecules to crystallizers, Oxford University Press Inc., New York, USA

DeBraekeleer, K., Cuesta Sánchez, F., Hailey, P., A., Sharp, D., C., A., Pettman, A., J., Massart, D., L., 1998, Influence and correction of temperature perturbations on NIR spectra during the monitoring of a polymorph conversion process prior to self-modelling mixture analysis, 17, pp. 141-152

Derdour, L., Févotte, G., Puel, F., Carvin, P., 2003, Real-time evaluation of the concentration of impurities during organic solution crystallization, Powder Tech., 129, pp. 1-7

Doki, N., Seki, H., Takano, K., Asatani, H., Yokota, M., Kubota, N., 2004, Process control of seeded batch coolint crystallization of the metastable α-form Glycine using an in-situ ATR-FTIR spectrometer and in-situ FBRM particle counter, Cryst. Growth Des., 4, pp. 949-953

Dunuwila, D., D., Berglund, K., A., 1997, ATR FTIR spectroscopy for in sity measurement of supersaturation, J. Cryst. Growth, 179, pp. 185-193

Eriksson, L., Johansson, E., Kettaneh-Wold, N., Wold, S., 2001, Multi- and megavariate data analysis, Principles and applications, Umetrics AB, Umeå

Falcon, J., A., Berglund, K., A., 2004, In situ monitoring of antisolvent addition crystallization with principal component analysis of Raman spectra, Cryst. Growth Des., 4, pp. 457-463

Faria, N., Pons, M.N., Feyo de Azevedo, S., Rocha, F.A. and Vivier, H., 2003, Quantification of the morphology of sucrose crystals by image analysis, Powder Technol., 133(1-3), pp. 54-67.

Fearn, T., 2000, On orthogonal signal correction, Chemom. Intell. Lab. Syst., 50, pp. 47-52

Feng, L., Berglund, K., A., 2002, ATR-FTIR for determining optimal cooling curves for batch crystallization of succinic acid, Cryst. Growth Des., 2, pp. 449-452

Ferreira, A., Faria, N., Rocha, F., Feyo de Azevedo, S., Lopes, A., 2005, Using image analysis to look into the effect of impurity concentration in NaCl crystallization, Chem. Eng. Res. Des., 83(A4), pp. 331-338

Feudale, R., N., Tan, H., Brown, S., D., 2002, Piecewise orthogonal signal correction, Chemom. Intell. Lab. Syst. 63, pp. 129-138

Févotte, G., Klein, J., P., 1996a, A new policy for the estimation of the course of supersaturation in batch crystallization, Can. J. Chem. Eng., 74, pp. 372-384

Févotte, G., Klein, J., P., 1996b, Calorimetric methods for the study of batch crystallization processes: Some key results, Trans. Inst. Chem. Eng., 74, pp. 791-796

Févotte, G., 2002, New perspectives fo the on-line monitoring of pharmaceutical crystallization processes using in situ infrared spectroscopy, Int. J. Pharm., 241, pp. 263-278

Finch, J., N., Lippincott, E., R., 1956, Hydrogen bond systems: temperature dependence of OH frequency shifts and OH band intensities, J. Chem. Phys, 24, pp. 908-909

Finch, J., N., Lippincott, E., R., 1957, Hydrogen bond systems-temperature dependence of OH frequency shifs and OH band intensities, Phys. Chem., 61, pp. 894-902

Flores-Cerillo, J, MacGregor J.F., 2004, Multivariate monitoring of batch processes using batch-to-batch information, AIChe J., 43, pp. 1219-1249

Fujiwara, M., Shan Chow, P., Ma, D., L., Braatz, R., D., 2002, Paracetamol crystallization using laser backscattering and ATR-FTIR spectroscopy: Metastability, agglomeration and control, 2, pp. 363-370

Fujiwara, M., Nagy, Z., K., Chew, J., W., Braatz, R., D., 2005, First principles and direct design approaches for the control of pharmaceutical crystallization, J. Proc. Contr., 15, pp. 493-504

Garbassi, F., Morra, M., Occhiello, E., 1993, Polymer surfaces, from physics to technology, 1st ed., John Wiley & Sons, West Sussex, England

Geladi, P., Forsström, J., 2002, Monitoring of a batch organic synthesis by near-infrared spectroscopy: modeling and interpretation of three-way data, J. Chemom., 16, pp. 329-338

Geladi, P., Åberg, P., 2001, Three-way modelling of a batch organic synthesis process monitored by near infrared spectroscopy, J. Near Infrared Spectr., 9, pp. 1-9

Goicoechea, H., C., Olivieri, A., C., 2001, A comparison of orthogonal signal correction and net analyte preprocessing methods. Theoretical and experimental study, Chemm. and Intell. Lab. Syst., 56, 2, pp. 73-81

Gourvenec, S., Stanimirova, I., Saby, C-A., Airiau, C., Y., Massart, D., L., 2005, Monitoring batch processes with the STATIS approach, J. Chemometrics, 19, pp. 288-300

Greene, E., F., Tauch, S., Webb, E., Amarasiriwardena, D., 2004, Application of diffuse reflectance infrared Fourier transform spectroscopy (DRIFTS) for the identification of potential diagenesis and crystallinity changes in teeth, 76, pp. 141-149

Griffiths, P., R., de Haseth, J., A., 1986, Fourier transform infrared spectrometry, John Wiley & Sons, New York

Grön, H., Roberts, K., J., 2001, Nucleation, growth and pseudo-polymorphic behavior of citric acid ad monitoired in situ by attenuated total reflection Fourier transform infrared spectroscopy, J. Phys. Chem., 105, pp. 10723-10730

Grön, H., Borissova, A., Roberts, K., J., 2003, In-process ATR-FTIR spectroscopy fo closed-loop supersaturation control of a batch crystallizer procucing monosodium glutamate crystals of defined size, Ind. Eng. Chem. Res., 42, pp. 198-206

Grön, H., Roberts, K., J., 2004, An examination of the crystallization of urea from supersaturated aqueous and aqueous-methanol solutions as monitored in-process using ATR FTIR spectroscopy, 4, pp. 929-936

Gutwald, T., Mersmann, A., 1990, Batch cooling crystallization at constant supersaturation: Technique and experimental results, Chem. Eng. Tech., 13, pp. 229-237

Harrick, N., J., 1979, Internal reflection spectroscopy, 1st ed., Harrick Scientific Corporation, New York

Helmy, R., Zhou, G., X., Chen, Y., W., Crocker, L., Wang, T., Wenslow, R., M., Vailaya, A., 2003, Characterization and quantitation of aprepitant drug substance polymorphs by attenuated total reflection Fourier transform infrared spectroscopy, Anal. Chem., 75, pp. 605-611

Hu, Y., Liang, J., K., Myerson, A., S., Taylor, L., S., 2005, Crystallization monitoring by Raman spectroscopy : Simultaneous measurement of desupersaturation profile and polymorphic form in Flufenamic acid systems, Ind. Eng. Chem. Res., 44, pp. 1233-1240

Hurley, L.A., Jones, A.G. and Hammond, R.B., 2004, Molecular packing, morphological modeling, and image analysis of cyanazine crystals precipitated from aqueous ethanol solutions, Cryst. Growth Des., 4(4), pp. 711-715

Höskuldsson, A., 1996, Prediction methods in science and technology, Vol. 1 Basic theory, Colourscan, Warsaw

Höskuldsson, A., 2001 Variable and subset selection in PLS regression, Chemom. Intell. Lab. Syst. 55, pp. 23-38

Jackson, J.E., 1991, A User's Guide to Principal Components, John Wiley & Sons, New York, NY

Jones, A., G., 2002, Crystallization process systems, 1st ed., Butterworth Heinemann, Oxford

Kim, K.-J., Mersmann, A., 2001, Estimation of metastable zone width in different nucleation processes, Chem. Eng. Sci., 56, pp. 2315-2324

Kitamura, M., 2004, Controlling factors and mechanism of polymorphic crystallization, Cryst. Growth Des., 4, pp. 1153-1159

Kipouros, K., Kachimanis, K., Nikolakakis, I., Malamataris, S., 2005, Quantitative analysis of less soluble form IV in commercial carbamatzepine (form III) by diffuse reflectance FTIR spectroscopy (DRIFTS) and lazy learning algorithm, Anal. Chim. Acta, 550, pp. 191-198

Kolár, P., Shen, J.-W., Tsuboi, A., Ishikawa, T., 2001, Solvent selection for pharmaceuticals, Fluid Phase Eq., 4892, pp. 1-12

Kourti, T., MacGregor, J., F., 1995, Process analysis, monitoring and diagnosis, using multivariate projection methods, Chemom. Intell. Lab. Syst., 28, pp. 3-21

Kourti, T., Nomikos, P., MacGregor, J., F., 1995, Analysis, monitoring and fault diagnosis of batch processes using multiblock and multiway PLS, J. Proc. Cont., 5, pp. 277-284

Kourti, T., Lee, J., MacGregor, J., F., 1996, Experiences with industrial applications on projection methods for multivariate statistical process control, Computers in Chemical Engineering, 20, s745

Kresta, J., MacGregor, J., F., Marlin, T., E., 1991, Multivariate statistical monitoring of process operating performance, Can. J. Chem. Eng., 69, pp. 35-47

Kubota, N., Doki, N., Yokota, M., Sato, A., 2001, Seeding policy in batch cooling crystallization, Powder Tech., 121, pp. 31-38

Lewiner, F., Févotte, G., Klein, J., P., Puel, F., 2001a, Improving batch cooling seeded crystallization of an organic weed-killer using on-line ATR FTIR measurement of supersaturation, J. Cryst. Growth, 226, pp. 348-362

Lewiner, F., Klein, J., P., Puel, F., Févotte, G., 2001b, On-line ATR FTIR measurement of supersaturation during solution crystallization processes. Calibration and applications on three solute/solvent systems, Chem. Eng. Sci., 56, pp. 2069-2084

Lewiner, F., Févotte, G., Klein, J., P., Puel, F., 2002, An online strategy to increase the average crystal size during organic batch cooling crystallization, Ind. Eng. Chem. Res., 41, pp. 1321-1328

Liotta, V., Sabesan, V., 2004, Monitoring and feedback control of supersaturation using ATR-FTIR to produce an active pharmaceutical ingredient of a desired crystal size, Org. Proc. Res. Dev., 8, pp. 488-494

Lopes, J., A., Menezes, J., C., Westerhuis, J., A., Smilde, A., K., 2002, Multiblock PLS analysis of an industrial pharmaceutical process, Biotech. Bioeng., 80, pp. 419-427

Löffelman, M., Mersmann, A., 2002, How to measure supersaturation?, Chem. Eng. Sci., 57, pp. 4301-4310

Ma, Z., Merkus, H., de Smet, J., Heffels, C., Scarlett, B., 2000, New developments in particle characterization by laser diffraction : size and shape, Powder Tech., 111, pp. 66-78

Ma, Z., Merkus, H.., Scarlett, B., 2001, Extending laser diffraction for particle shape characterization : technical aspects and application, Powder Tech., 118, pp. 180-187

MacGregor, J., F., Jaeckle, C., Kiparissides, C., Koutoudi M., 1994, AIChe J., 40, pp. 826-838

MacGregor, J., F., Kourti, T., 1995, Statistical process control of multivariate processes, Contr. Eng. Pract., 3, pp. 403-413

Martens, H., Naes, T., 1993, Multivariate calibration, John Wiley & Sons, Chichester

Matthews, H.B. and Rawlings, B., 1998, Batch crystallization of a photochemical: modeling, control and filtration, AIChE J., 44(5), pp. 1119-1127.

Mersmann, A., 1995, Crystallization technology handbook, 10th ed., Marcel Dekker Inc., New York

Mersmann, A., 1996, Supersaturation and nucleation, Trans IChemE, 74, pp 812-820

Mersmann, A. and Löffelmann, M., 2000, Crystallization and precipitation: the optimal supersaturation, Chem. Eng. Technol., 23(1): 11-15.

Mohameed, H., A., 2002, Effect of cooling rate on unseeded batch crystallization, Chem. Eng. Proc., 41, pp. 297-303

Mougin, P., Wilkinson, D., Roberts, K., J., Jack, R., Kippax, P., 2003, Sensitivity of particle sizing by ultrasonic attenuation spectroscopy to material properties, Powder Tech., 359, pp. 243-248

Mujunen, S.-P., 1999, Multivariate monitoring of wastewater treatement processes in pulp and paper industry, Diss., Acta Polytechnica Scandinavica, Chemical technology series, No. 264

Mullin, J., W., Nyvlt, J., 1971, Programmed cooling of batch crystallizers, Chem. Eng. Sci., 26, pp. 369-377

Mullin, J., W., 2001, Crystallization, 4th ed., Butterworth-Heinemann, Oxford, Great Britain

Myerson, A., S., 1993, Handbook of industrial crystallization, Butterworth-Heinemann, Boston, USA

Myerson, A., S., Decker, S., E., Welping, F., 1986, Solvent selection and batch crystallization, Ind. Eng. Chem. Proc. Des. Dev., 25, pp. 925-929

Noda, I., Liu, Y., Ozaki, Y., Czarnecki, M., A., 1995, Two-dimensional Fourier transform near-infrared correlation spectroscopy studies of temperature-dependent spectral variations of oleyl alcohol, J. Phys. Chem., 99, pp. 3068-3073

Nomikos P., MacGregor, J., F., 1994, Monitoring batch processes using multiway principal component analysis, AIChe J., 40, pp. 1361

Nomikos P., MacGregor, J., F., 1995a, Multivariate SPC charts for monitoring batch processes, Technometrics, 37, pp. 41

Nomikos, P., and MacGregor, 1995b, Multi-way partial least squares in monitoring batch processes, Chemom. Intell. Lab. Syst., 30, pp. 97

Ono, T., ter Horst, J., H., Jansens, P., J., 2004, Quantitative measurement of the polymorphic transformation of α-Glutamic acid using in-situ Raman spectroscopy, Cryst. Growth Des., 4, pp. 465-469

Ozaki, Y., Liu, Y., Noda, I., 1997, Two-dimensional infrared and near-infrared correlation spectroscopy: applications to studies of temperature-dependent spectral variations of self associated molecules, Appl. Spectrosc., 51, pp. 526-535

Park, K.-S., Hyeson, L., Jun, C.-H., Park, K., Jung, J.-W., Kim, S.-B., 2000, Rapid determination of FeO content in sinter ores using DRIFT spectra and multivariate calibrations, Chemom. Intell. Lab. Syst., 51, pp. 163-173

Paul, E., L., Tung, H.-H., Midler, M., 2005, Organic crystallization processes, Powder Tech., 150, pp. 133-143

Pons, M.N., Vivier, H., Delcour, V., Authelin, J.-R. and Paillères-Hubert, L., 2002, Morphological analysis of pharmaceutical powders, Powder Technol., 128(2-3), pp. 276-286.

Profir, V., M., Furusjö, E., Danielsson, L.-G., Rasmuson, Å., C., 2002, Study of the crystallization of mandelic acid in water using in situ ATR-IR spectroscopy, Cryst. Growth Des., 2, pp. 273-279

Pöllänen, K., Häkkinen, A., Reinikainen S.-P., Louhi-Kultanen M., Nyström L., 2006, A study on batch cooling crystallization of sulphathiazole: Process monitoring using ATR-FTIR and product characterization by automated image analysis, Chem. Eng. Res. and Des., 84(A1), pp. 47-51

Pöllänen, K., Häkkinen, A., Huhtanen M., Reinikainen, S.-P., Karjalainen M., Rantanen J., Louhi-Kultanen M., Nyström L., 2005a, FTIR for quantitative characterization of polymorphic composition of sulfathiazole, Anal. Chim. Acta, 544, pp. 108-117

Pöllänen K., Häkkinen A., Reinikainen S.-P., Rantanen J., Karjalainen M., Louhi-Kultanen M., Nyström L., 2005b, IR spectroscopy together with multivariate data analysis as a process analytical tool for in-line monitoring of crystallization process and solid state analysis of crystalline product, J. Pharm. Biom. Anal., 38/2, pp. 275-284

Pöllänen, K., Häkkinen, A., Reinikainen S.-P., Louhi-Kultanen, M., Nyström L., , 2005c, ATR-FTIR in monitoring of crystallization processes: Comparison of indirect and direct OSC methods, Chemom. Intel. Lab. Syst., 76, pp. 25-35

Qu H., Pöllänen K., Louhi-Kultanen M., Kilpiö T., Oinas P., Kallas J., 2005, Batch cooling crystallization study based on in-line measurement of supersaturation and crystal size distribution, Journal of Crystal Growth, 275, pp. e1857-e1862

Ramaker, H.-J., van Sprang, E., N., M., Westerhuis, J., A., Smilde, A., K., 2004, The effect of the size of the training set and number of principal components on the false alarma rate in statistical process monitoring, Chemom. Intell. Lab. Syst., 73, pp. 181-187

Salari, A., Young, R., E., 1998, Application of attenuated total reflection FTIR spectroscopy to the analysis of mixtures of pharmaceutical polymorphs, Int. J. Pharm., 163, pp. 157-166

Schwartz, A., M., Berglund, K., A., 1999, The use of Raman spectroscopy for in situ monitoring of lysozyme concentration during crystallizatin in a hanging drop, J. Cryst. Growth, 203, pp. 599-603

Schwartz, A., M., Berglund, K., A., 2000, In situ monitoring and control of lysozyme concentration during crystallization in a hanging drop, J. Cryst. Growth, 210, pp. 753-760

Settle, F., 1997, Handbook of instrumental techniques for analytical chemistry, Prentice Hall, New Jersey

Sharaf, M., A., Illman, D., L., Kowalski, B., R., 1986, Chemometrics, John Wiley & Sons, N. Y.

Shimizu, K., Nomuya, T., Takahashi, K., 1998, Crystal size distribution of aluminum potassium sulfate in a batch crystallizer equipped with different types of impeller, J. Cryst. Growth, 191, pp. 178-184

Silverstein, R., M., Clayton Bassler, G., Morril, T., C.,1991 Spectrometric identification of organic compounds, 5th ed., Wiley, N.Y.,

Singhal, D., Curatolo, W., 2004, Drug polymorphism and dosage form design: a practical perspective, Adv. Drug Del. Rev., 56, pp. 335-347

Spragg, R., 1998, 17. Reflection spectra, Int. J. Vib. Spect., [www.ijvs.com], Vol. 1, Ed. 4, pp. 17-30

Starbuck, C., Spartalis, A., Wai, L., Wang, J., Fernandez, P., Lindemann, C., M., Zhou, G., X., Ge, Z., 2002, Process optimization of a complez pharmaceutical polymorphic system via in situ Raman spectroscopy, 2, pp. 515-522

Stordrange, L., Rajalahti, T., Libnau, F., O., 2004, Multiway methods to explore and model NIR data from a batch process, Chemom. Intell. Lab. Syst., 70, pp. 137-145

Taavitsainen, V.-M., 2001, Strategies for combining soft and hard modelling in some physicochemical problems, Diss., Acta Universitatis Lappeenrantaensis, 245 s.

Tamagawa, R., E., Miranda, E., A., Berglund, K., A., 2002a, Raman spectroscopic monitoring and control of Aprotinin supersaturation in hanging-drop crystallization, Cryst. Growth Des., 2, pp. 263-267

Tamagawa, R., E., Miranda, E., A., Berglung, K., A., 2002b, Simultaneous monitoring of protein and $(NH_4)_2SO_4$ concentrations in aprotinin hanging-drop crystallization using Raman spectroscopy, Cryst. Growth Des., 2, pp. 511-514

Togkalidou, T., Fujiwara, M., Patel, S., Braatz, R., D., 2001, Solute concentration prediction using chemometrics and ATR-FTIR spectroscopy, J. Cryst. Growth, 231, pp. 534-543

Tremblay, L., Gagné, J.-P., 2002, Fast quantification of humic substances and organic matter by direct analysis of sediments using DRIFT spectroscopy, Anal. Chem., 74, pp. 2985-2993

Trygg, J., Wold, S., 2002, Orthogonal projections to latent structures (O-PLS), Journal of Chemometrics 16, 119-128

Trygg, J., 2002, O2-PLS for qualitative and quantitative analysis in multivariate calibration, Journal of Chemometrics, 16, 283-293

Tähti, T., Louhi-Kultanen, M., Palosaari, S., 1999, On-line measurement of crystal size distribution during batch crystallization, Proceedings of the 14th international symposium on industrial crystallization, Cambridge, UK

U.S. Food and Drug Administration (FDA), 2004, Guidance for Industry. PAT — A framework for innovative pharmaceutical manufacturing and quality assurance, Rockville, MD, USA, pp 21.

Ulrich, J. and Strege, C., 2002, Some aspects of the importance of metastable zone width and nucleation in industrial crystallizers, J. Cryst. Growth, 237-239(Pt. 3), pp. 2130-2135.

Uusi-Penttilä, M., Berglund, K., A., 1996, Spectroscopic monitoring of environmentally benign anti-solvent crystallization, J. Cryst. Growth, 166, pp. 967-970

Vandeginste, B., G., M., Massart, D., L., Buydens, L., M., C., De Jong, S., Lewi, P., J., Smeyers-Verbeke, J., 1998, Handbook of Chemometrics and Qualimetrics: Part B, Elsevier, Amsterdam

Virone, C., ter Horst, J., H., Kramer, H..J., M., Jansens, P., J., 2005, Growth rate disperstion of ammonium sulphate attrition fragments, J. Cryst. Growth, 275, pp e1397-e1401

Westerhuis, J., A., 2001, Direct orthogonal signal correction, Chemom. Intell. Lab. Syst., 56, pp. 13-25

Wise, B. M., Gallagher, N., B., 1996, The process chemometrics approach to process monitoring and fault detection, J. Proc. Cont., 6, pp. 329-348

Wold, S., 1976, Pattern recognition by means of disjoint principal components model, Pattern Recogn., 8, pp. 127-139

Wold, S., Martens, H., Wold, H., The multivariate calibration problem in chemistry solved by PLS method, Lect. Notes Math. 973, 1983, pp. 286-293

Wold, S., 1987, Principal Component Analysis, Chemom. Intell. Lab. Syst., 2, pp. 27-52

Wold, S., Geladi, P., Esbensen, K., Öhman, J., 1987, Multi-way principal components-and PLS-analysis, J. Chemom., 1, pp. 41-56

Wold, S., Antti, H., Lindgren, F., Öhman, J., 1998, Orthogonal signal correction of near-infrared spectra, Chemom. Intell. Lab. Syst. 44, pp. 175-185

Workman, J., Springsteen, A., W., 1997, Applied spectroscopy, A compact reference for practitioners, 1st ed., Academic Press Limited, London

Wülfert, F., Kok, W., T., Smilde, A., K., 1998, Influence of temperature on vibrational spectra and consequences for the predictive ability of multivariate models, Anal. Chem., 70, pp. 1761-1767

Wülfert, F., Kok, W., T., de Noord, O., E., Smilde, A., K., 2000, Linear techniques to correct for temperature-inducec spectral variation in multivariate calibration, Chemom. Intell. Lab. Syst., 2000,pp. 189-200

Yamamoto, H., Matsuyama, T., Wada, M., 2002, Shape distinction of particle materials by laser diffraction pattern analysis, Powder Tech., 122, pp. 205-211

Yu, L., X., Lionberger, R., A., Raw, A., S., D'Costa, R., Wu, H., Hussain, A., S., 2004, Applications of process analytical technology to crystallization processes, Adv. Drug Del. Rev., 56, pp. 349-369

Ålander, E.M., Uusi-Penttilä, M.S. and Rasmuson, Å.C., 2003, Characterization of paracetamol agglomerates by image analysis and strength measurement, Powder Technol., 130(1-3), pp. 298-306

ACTA UNIVERSITATIS LAPPEENRANTAENSIS

200. PÖYHÖNEN, AINO. Modeling and measuring organizational renewal capability. 2004. U.s. Diss.

201. RATAMÄKI, KATJA. Product platform development from the product lines´ perspective: case of switching platform. 2004. 218 s. Diss.

202. VIRTANEN, PERTTU. Database rights in safe European home: the path to more rigorous protection of information. 2005. 425 s. Diss.

203. Säädöksiä, systematiikkaa vai ihmisoikeuksia? Oikeustieteen päivät 19. – 21.8.2003. Toim. Marjut Heikkilä. 2004. 350 s.

204. PANTSAR, HENRIKKI. Models for diode laser transformation hardening of steels. 2005. 134 s., liitt. Diss.

205. LOHJALA, JUHA. Haja-asutusalueiden sähkönjakelujärjestelmien kehittäminen – erityisesti 1000 V jakelujännitteen käyttömahdollisuudet. 2005. 201 s., liitt. Diss.

206. TARKIAINEN, ANTTI. Power quality improving with virtual flux-based voltage source line converter. 2005. U.s. Diss.

207. HEIKKINEN, KARI. Conceptualization of user-centric personalization management. 2005. 136 s. Diss.

208. PARVIAINEN, ASKO. Design of axial-flux permanent-magnet low-speed machines and performance comparison between radial-flux and axial-flux machines. 2005. 153 s. Diss.

209. FORSMAN, HELENA. Business development efforts and performance improvements in SMEs. Case study of business development projects implemented in SMEs. 2005. 209 s. Diss.

210. KOSONEN, LEENA. Vaarinpidosta virtuaaliaikaan. Sata vuotta suomalaista tilintarkastusta. 2005. 275 s. Diss.

211. 3rd Workshop on Applications of Wireless Communications. 2005. 62 s.

212. BERGMAN, JUKKA-PEKKA. Supporting knowledge creation and sharing in the early phases of the strategic innovation process. 2005. 180 s. Diss.

213. LAAKSONEN, PETTERI. Managing strategic change: new business models applying wireless technology as a source of competitive edge. 2005. 142 s. Diss.

214. OVASKA, PÄIVI. Studies on coordination of systems development process. 2005. U.s. Diss.

215. YANG, GUANGYU. Control and simulation of batch crystallization. 2005. U.s. Diss.

216. MUSTONEN–OLLILA, ERJA. Information system process innovation adoption, adaptation, learning, and unlearning: a longitudinal case study. 2005. U.s. Diss.

217. SAINIO, LIISA-MAIJA. The effects of potentially disruptive technology on business model – a case study of new technologies in ICT industry. 2005. 205 s. Diss.

218. SAINIO, TUOMO. Ion-exchange resins as stationary phase in reactive chromatography. 2005. 175 s. Diss.

219. CONN, STEFFEN. Software tools to support managers: development and application to new product development and strategic alliance management. 2005. 168 s. Diss.

220. TYNJÄLÄ, TERO. Theoretical and numerical study of thermomagnetic convection in magnetic fluids. 2005. U.s. Diss.

221. JANTUNEN, ARI. Dynamic capabilities and firm performance. 2005. 210 s. Diss.

222. KOLA-NYSTRÖM, SARI M. In search of corporate renewal: how to benefit from corporate venturing? 2005. 190 s. Diss.

223. SARÉN, HANNU. Analysis of the voltage source inverter with small DC-link capacitor. 2005. 143 s. Diss.

224. HUUHILO, TIINA. Fouling, prevention of fouling, and cleaning in filtration. 2005. U.s. Diss.

225. VILJAINEN, SATU. Regulation design in the electricity distribution sector – theory and practice. 2005. 132 s. Diss.

226. AVRAMENKO, YURY. Case-based design method for chemical product and process development. 2005. U.s. Diss.

227. JÄRVINEN, KIMMO. Development of filter media treatments for liquid filtration. 2005. U.s. Diss.

228. HURMELINNA-LAUKKANEN, PIA. Dynamics of appropriability – finding a balance between efficiency and strength in the appropriability regime. 2005. U.s. Diss.

229. LAARI, ARTO. Gas-liquid mass transfer in bubbly flow: Estimation of mass transfer, bubble size and reactor performance in various applications. 2005. U.s. Diss.

230. BORDBAR, MOHAMMAD HADI. Theoretical analysis and simulations of vertically vibrated granular materials. 2005. U.s. Diss.

231. LUUKKA, PASI. Similarity measure based classification. 2005. 129 s. Diss.

232. JUUTILAINEN, ANNELI. Pienen matkailuyrityksen yrittäjän taival. Oppiminen yrittäjyysprosessissa. 2005. 191 s. Diss.

233. BJÖRK, TIMO. Ductility and ultimate strength of cold-formed rectangular hollow section joints at subzero temperatures. 2005. 163 s. Diss.

234. BELYAEV, SERGEY. Knowledge discovery for product design. 2005. U.s. Diss.

235. LEINONEN, KARI. Fabrication and characterization of silicon position sensitive particle detectors. 2006. U.s. Diss.

236. DUFVA, KARI. Development of finite elements for large deformation analysis of multibody systems. 2006. U.s. Diss.

237. RITVANEN, JOUNI. Experimental insights into deformation dynamics and intermittency in rapid granular shear flows. 2006. U.s. Diss.

238. KERKKÄNEN, KIMMO. Dynamic analysis of belt-drives using the absolute nodal coordinate formulation. 2006. 121 s. Diss.

239. ELFVENGREN, KALLE. Group support system for managing the front end of innovation: case applications in business-to-business enterprises. 2006. 196 s. Diss.

240. IKONEN, LEENA. Distance transforms on gray-level surfaces. 2006. 132 s. Diss.

241. TENHUNEN, JARKKO. Johdon laskentatoimi kärkiyritysverkostoissa. Soveltamismahdollisuudet ja yritysten tarpeet. 2006. 270 s. Diss.

242. KEMPPINEN, JUKKA. Digitaaliongelma. Kirjoitus oikeudesta ja ympäristöstä. 2006. 492 s.