

**STUDIES ON CRYSTALLIZATION PROCESS; MONITORING AND
CONTROL**

Thesis Submitted by

APARNA A M (212CH1072)

In partial fulfillment for the award of the degree of

MASTER OF TECHNOLOGY

IN

CHEMICAL ENGINEERING

Under the guidance of

Dr. MADHUSREE KUNDU



DEPARTMENT OF CHEMICAL ENGINEERING

NATIONAL INSTITUTE OF TECHNOLOGY

ROURKELA-769008

May, 2014



**DEPARTMENT OF CHEMICAL ENGINEERING
NATIONAL INSTITUTE OF TECHNOLOGY
ROURKELA-769008, INDIA**

CERTIFICATE

This is to certify that the thesis entitled “Studies on Crystallization Process; Monitoring and Control” submitted by Ms. Aparna A M to National Institute of Technology, Rourkela, India for the award of degree of Master of Technology in engineering, is a bonafide record of investigation carried out by her in Department of Chemical Engineering, under the guidance of Dr. M. Kundu. The thesis is up to the standard of fulfillment of M. Tech degree as prescribed by regulation of this institute.

Date: 23/05/2014

**Dr. Madhusree Kundu
National Institute of Technology
Rourkela, Orissa-769008**

ACKNOWLEDGEMENT

I wish to express my heartily gratitude to my guide **Dr. M. Kundu** for her invaluable guidance with constant flow of existing new ideas and contagious enthusiasm through my work. My association with my guide will remain a memorable part of my life. I wish to thank her for the readiness to discuss problems of any kind at any time. She is always a source of inspiration to me. I would like to express my indebtedness to madam for the wonderful ideas and an amazing path she paved for my work.

I want to acknowledge the support and encouragement given by Mr. Seshu Kumar Damarla throughout the period of my work.

I express my sincere thanks to Dr. R. K. Singh, Head of Department of Chemical Engineering for providing me with all the necessary facilities to complete this research work. I am also grateful Dr. B. Munshi and all other faculty members of department for their help and support.

I am obliged to all my friends for their support and encouragements. Finally, the thesis would not have been completed without the support of the most important people in my life – my family. I sincerely wish to thank my parents, for their unconditional love and constant encouragement.

Date- 23/05/2014

**APARNA A M
212CH1072**

CONTENTS

	Page No.
Certificate	
Acknowledgement	
Abstract	i
List of Figures	ii
List of Tables	iii
	iv
Chapter 1	
INTRODUCTION	
1.1. Introduction	1
1.2. Objective	3
1.3. Organization of the thesis	3
References	4
Chapter 2	
PRINCIPLES OF CRYSTALLIZATION	
2.1. Introduction	5
2.1.1. Supersaturation and Solubility	5
2.1.2. Metastable zone	6
2.1.3. Crystallization kinetics	7
2.1.3.1. Nucleation	7
2.1.3.2. Crystal growth	8
2.1.3.3. Seeding	9
2.1.3.4. Control of Crystallizers	10

2.1.3.4.1. The Direct Design Approach	10
2.1.3.4.2 The Model Based Control Approach	11
Reference	12
Chapter 3	
BATCH CRYSTALLIZER	
3.1. Introduction	13
3.2. Mathematical modeling of batch cooling crystallizer	14
3.2.1. Moment transformation	16
3.2.2. Energy balance	17
3.2.2.1. Natural cooling	17
3.2.2.2. Linear cooling	18
3.2.2.3. Controlled cooling	18
3.3. Dynamic simulation of batch crystallization process	19
3.4. Crystal size distribution	22
3.5. Conclusion	23
References	24
Chapter 4	
CONTINUOUS CRYSTALLIZER	
4.1. Introduction	25
4.2. Mathematical modeling of non-isothermal continuous cooling crystallizer	26
4.3. Dynamic simulation of continuous crystallization process	29
4.4. Closed loop responses of continuous crystallization process	32
4.5. Crystal size distribution	36
4.6. Conclusion	36

References	37
Chapter 5	
MULTIVARIATE PROCESS MONITORING AND CONTROL	
5.1. Introduction	38
5.1.1. Data based models	39
5.2. Theoretical postulates	40
5.2.1. Principal component analysis (PCA)	40
5.2.2. Similarity factors	42
5.2.2.1. PCA Similarity	43
5.2.2.1. Distance similarity	43
5.2.2.3. Combination of Similarity factors	44
5.2.2.4. Dissimilarity index	44
5.3. Clustering time series data	45
5.3.1. Modified K mean clustering algorithm	46
5.3.1.1. Using similarity factors	47
5.3.1.2. Using dissimilarity index	48
5.3.2. Cluster Performance Evaluation	48
5.4. Pattern matching	48
5.4.1. Moving window based pattern matching	50
5.5. Monitoring of continuous crystallizer	51
5.6. Results & Discussions	52

5.6.1. Clustering of time series data	52
5.6.2. Moving window based pattern matching	54
5.7. Conclusion	55
References	56
Chapter 6	
CONCLUSIONS AND RECOMMENDATION FOR FUTURE WORK	
6.1. Conclusion	57
6.2. Recommendation for Future Work	59

ABSTRACT

Crystallization is an oldest unit operations used by industries for the separation as well as purification of a solid product. It is a popular operation in majority of industries, including pharmaceutical, food, microelectronics and bulk and fine chemicals. The widespread use of crystallizers is attributed to numerous advantages provided by the process like formation of pure solids from impure solution, less energy required for separation, availability of products in a suitable form for packaging, transportation and storage. The mathematical modelling, control and monitoring of crystallization process is a significant research area both from academic as well as industrial point of view.

A temperature trajectory that improves the crystal size distribution in a batch crystallizer is proposed and compared with some of the basic cooling modes like natural, linear and controlled cooling. The properties of the crystalline product and dynamic behavior of the crystallizer, obtained by numerical experimentation are presented and analyzed here.

The Mathematical modeling and simulation of a non-isothermal continuous cooling crystallizer followed by a control strategy to improve the crystal size distribution has been proposed. The model developed here is used to monitor the process using various multivariate statistical process control techniques. Monitoring and control of chemical processes are highly challenging due to their non-linear, multivariate and highly correlated nature. Present dissertation furnishes the successful implementation of various chemometric techniques. Two methods were chosen for process monitoring and control: Clustering of time series data and moving window based pattern matching. PCA similarities and dissimilarity index were chosen as the index of clustering as well as pattern matching. These methods have been successfully implemented in continuous crystallizer for fault detection and to differentiate among various operating conditions. Both the methods produced promising results.

Keywords: Crystallization, control, MSPC, PCA, Monitoring, fault detection.

LIST OF TABLES

Table No.	Table Caption	Page No
Table 3.1	The Physical Parameters and Initial Conditions of the Seeded Batch Crystallizer	19
Table 3.2	Effect of different cooling policies on product CSD	23
Table 4.1	The Physical Parameters and Initial Conditions of the Seeded continuous Crystallizer	29
Table 4.2	Crystal Size Distribution	36
Table 5.1	Operating Conditions	51
Table 5.2	Clustering Performance based on Combined similarity factor $0.67 S_{PCA}^{\lambda} + 0.33 S_{dist}$	52
Table 5.3	Clustering Performance based on Combined similarity factor $0.67 S_{PCA} + 0.33 S_{dist}$	52
Table 5.4	Dissimilarity Index based clustering performance	53
Table 5.5	Clustering performance using different methodologies	53
Table 5.6	Pattern matching using combined similarity factor	54
Table 5.7	Pattern matching using combined similarity factor	54
Table 5.8	Pattern matching using dissimilarity index	54

LIST OF FIGURES

Figure No.	Figure Caption	Page No.
Figure 2.1	The Phase Diagram	7
Figure 2.2	Mechanisms of Crystal Nucleation	8
Figure 2.3	Seeding Policy	9
Figure 3.1	The Direct Design Approach	10
Figure 3.2	Comparison of operating policy for batch cooling crystallization.	20
Figure 3.3	Effect of cooling curves on transient supersaturation	20
Figure 4.1	Effect of cooling curves on seed crystal length.	21
Figure 4.2	Zero order moment of crystallization	30
Figure 4.3	First order moment of crystallization	30
Figure 4.4	Second order moment of crystallization	30
Figure 4.5:	Third order moment of crystallization	31
Figure 4.6:	Fourth order moment of crystallization	31
Figure 4.7	Fifth order moment of crystallization	31
Figure 4.8	Concentration profile	32
Figure 4.9	Crystallizer Temperature Profile	32
Figure 4.10	Jacket Temperature profile	32
Figure 4.11	Zero order moment of crystallization (closed loop).	33
Figure 4.12	First order moment of crystallization (closed loop).	33
Figure 4.13	Second order moment of crystallization (closed loop).	34

Figure 4.14	Third order moment of crystallization (closed loop).	34
Figure 4.15	Fourth order moment of crystallization (closed loop).	34
Figure 4.15	Fifth order moment of crystallization (closed loop).	35
Figure 4.16	Concentration profile (closed loop).	35
Figure 4.17	Temperature profile (set point tracking)	35

NOMENCLATURE

A	Total heat transfer surface area (cm^2)
b	Nucleation rate exponent
B	Nucleation rate
C	Solution concentration (g/l)
C_{mz}	Metastable concentration
C_p	Heat capacity of the solution ($\text{cal/g}^{\circ}\text{C}$)
C_{pj}	Heat capacity of cooling water ($\text{cal/g}^{\circ}\text{C}$)
C_s	Saturation concentration (g solute g solvent ⁻¹)
E	Activation energy
g	Growth rate exponent
G	Growth rate (cm s^{-1})
ΔH	Heat of crystallization (cal/g)
k_b	Birth rate coefficient ($\text{s}^{-1} \text{cm}^{-3}$)
p	Population density of crystals
F_j	Cooling water flow rate
k_g	Growth rate coefficient
k_v	Volumetric shape factor
L	Characteristic crystal length (μm)
M	Mass of solvent in the crystallizer (kg)

T	Reactor temperature ($^{\circ}\text{C}$)
T_j	Cooling jacket temperature ($^{\circ}\text{C}$)
$T_{j\text{sp}}$	Set point of the jacket temperature ($^{\circ}\text{C}$)
UA	Overall heat transfer coefficient ($\text{cal}/\text{min}^{\circ}\text{C}$)
U_0A_0	Overall heat transfer coefficient ($\text{cal}/\text{min}^{\circ}\text{C}$)
V_j	Jacket volume (cm^3)
V	Crystallizer volume (cm^3)

GREEK SYMBOLS

ρ	Density of crystals (g/cm^3)
ρ_c	Density of cooling water (g/cm^3)
μ_0	Zeroth moment of the CSD
μ_1	First moment of the CSD
μ_2	Second moment of the CSD
μ_3	Third moment of the CSD
τ	Time

Chapter 1

INTRODUCTION

Chapter 1

INTRODUCTION

Crystallization is an oldest unit operations used by industries for the separation as well as purification of solid products. It is a popular operation in majority of industries, including pharmaceutical, food, microelectronics and bulk and fine chemicals. The production process of the majority of all solid products includes at least one crystallization step during the synthesis or purification of intermediates or the final product.

Since crystallization is most of the time the first step when a solid product is separated from the liquid solution it is a vital process to monitor the solid properties, like crystal size distribution (CSD), shape, purity etc. Since these physical characteristics has a strong influence on the end product characteristics as well as on the effectiveness of the downstream processes (such as drying, filtration etc.), the proper design and control of crystallization processes can have a noteworthy effect on the overall efficiency of the solid production process and the value of the final product.

During the past two decades significant effort has been made to develop a better understanding of the crystallization mechanisms like nucleation, growth etc., as well as on the modeling and control of crystallization systems. The advancement in process analytical technology and computing power has made this task much easier. The key developments have occurred in novel crystallization concepts, modeling, monitoring and control. A major development in the modeling area is the use of multi-dimensional population balance models which gives the morphological model of crystallization processes, as well as in the better understanding of crystallization in impure media. Developments in the monitoring of crystallization processes are highly attributed to the advances in process analytical technologies and advanced chemometrics and statistical process control concepts. The developments in the modeling and monitoring of crystallizers have also led to significant progress in the field of advanced control concepts.

Monitoring using various process analytical techniques not only utilize real-time measurement tools but also incorporates chemical, physical, microbiological, and statistical

analysis like chemometrics. Chemometrics is usually applied for the interpretation of multivariate data provided by various in situ process monitoring devices. Chemometrics can be used to construct calibration models for sensors, to derive mathematical relationships between the anticipated product properties, like crystal size distribution, and various other subsidizing variables, such as the solute concentration. These mathematical relationships are constructed using multivariate data analysis methods such as Principal Component Analysis (PCA), Partial Least Squares (PLS), Principal Component regression (PCR), Multiple Linear Regression (MLR) and variations of these approaches

Various statistical process monitoring (SPM) techniques based on time series measurements have been developed and proposed for the monitoring of steady-state processes. They are used to spot unwanted departures from steady-state conditions. Thus, these approaches are well suited for monitoring of continuous crystallization systems. The purpose of statistical monitoring tools is to distinguish between random events and systematic changes of the process. Some of the frequently used statistical monitoring methods are the control charts e.g. Shewhart, CUSUM, EWMA. But, these methods are not suited for highly correlated data and in crystallization, the process variables are highly correlated and it may cause false alarms. This can be addressed by the use of control charts which take into account the autocorrelation and modify the control limits. A further approach is to develop a time series model that models the autocorrelation and then use the residuals for monitoring. But, univariate control charts are not efficient due to the interactions between variables. This is due to the fact that in univariate analysis the affiliation to the response variable are visualized one at a time and thus does not disclose the multivariate configurations between the variables, which is vital for industrial processes both for interpretation and prediction.

Multivariate charts like Hotelling's T^2 , MEWMA, MCUSUM, take this relationship into account for monitoring the mean vector or variance-covariance matrix. Univariate control charts are more popular than Multivariate charts. Apart from the difficulty involved in the computation, of multivariate charts there are some additional difficulties which diminish their appeal: Unlike the univariate case, the scale of the values presented on the multivariate chart is not correlated to that of the monitored variables. If an out of control signal is indicated by a multivariate chart, it

becomes difficult to determine which variable responsible for the signal. More complicated methods are needed to decide the origin of the signals.

Therefore, Multivariate Statistical Process Control (MSPC) techniques, which consider all the variables of interest simultaneously and can, extract information on the behavior of each variable or characteristic relative to the others is highly essential. MSPC research is having high value in theoretical as well as industrial application and is certainly a benefit to process monitoring, fault detection and process identification and control.

1.1. OBJECTIVE

In the context of aforesaid discussion, it is worthy to mention that control of crystallization process and application of multivariate statistical process monitoring and control deserves extensive research before its widespread use and commercialization. For very complex and non-linear processes, data based process monitoring, identification & control seems to be unequivocally superior to physical model based approaches. Expert systems or knowledge based system can be developed by integration of process knowledge derived from the process data base. In view of this, the objectives of the present dissertation are as follows:

- Development of a temperature trajectory for control of batch crystallizer CSD.
- Modeling and dynamic simulation of continuous cooling crystallization process.
- Design of optimal control for CSD of continuous cooling crystallizer.
- Monitoring of continuous crystallizer using MSPC techniques.

1.2. ORGANIZATION OF THE THESIS

Chapter 1: This chapter of the thesis presents a brief introduction on monitoring and control of crystallization process. The objective of this thesis with the chapter layout is also included in this chapter.

Chapter 2: Theory of crystallization process along its kinetics and various control approaches are presented in this chapter.

Chapter 3: In this chapter a mathematical model of batch crystallizer was resolved and simulated for three different cooling methods: natural, linear and controlled cooling.

Chapter 4: Modeling, open loop simulations along with the closed loop simulations are the traits of the present chapter.

Chapter 5: Theoretical postulates of the various MPC techniques and the application of these techniques on continuous crystallizer are presented in this chapter.

Chapter 6: On an ending note this chapter concludes the thesis with future recommendation.

REFERENCES

1. Apte C, "Data mining an industrial research perspective." Institute of Electrical and Electronic Engineers Trans. Comput. Sci. Eng., 5, 6-9, 1997.
2. Douglas J M, "Conceptual design of chemical processes", McGraw Hill, New York, 1988.
3. Downs J J, Vogel E F, "A plant-wide industrial process control problem." Comput. Chem. Engr., 17: 245-255, 1995.
4. Luyben W L, Tyreus B D, "Plant-wide process control", McGraw Hill, New York, 1999.
5. Stephanopoulos G, "Chemical process control", Prentice Hall, Englewood Cliffs, 1984.
6. Vollmer U, "Control of crystallization process based on population balance", Phd thesis, Forschungsberichte aus dem Max-Plank Institute, 1-123, 2005.

Chapter 2

PRINCIPLES OF CRYSTALLIZATION

Chapter 2

PRINCIPLES OF CRYSTALLIZATION

2.1 Introduction

Crystallization is a process of formation of solid crystal from a uniform solution. It is used in pharmaceutical as well as bulk chemical industries. Its widespread use is due to the fact that in a single operation enables both separation as well as purification, resulting in a crystalline solid which is sequestered with high level of purity and low capital and operating cost. Crystallization process provides the chance to govern the size and amount of crystals formed. In the case of pharmaceutical industries, the size and structure of the crystals formed are significant. Product quality and efficiency are often dependent on satisfactory control of particle size and structure. The size of crystal produced in a crystallizer can have an unfavorable effect on downstream operations.

Therefore, it is desirable to propose a crystallization process that prevents extreme fines generation and thus can reduce downstream operation problems. The degree of supersaturation can be calculated using the population balance together with the understanding of the nucleation and growth kinetics. The population balance method has a vital part in crystallization and other particulate processes.

2.1.1 Supersaturation and solubility

Solubility determination constitutes the first step in study of crystallization process. The knowledge of solubility gives the rate of supersaturation which is the driving force for crystallization. Solubility measurement is performed by heating a suspension and discerning the temperature corresponding to complete dissolution. Cooling of a saturated solution results in a solution that is supersaturated.

Supersaturation can be defined as the difference in the chemical potential of solute concentration in a supersaturated to that of a saturated solution. Various methods like cooling, antisolvent addition, chemical reactions, solvent evaporation etc can be used supersaturation generation. The most commonly employed method for supersaturation generation is cooling. this method can be applied to materials whose solubility is a function of temperature. the final

product properties and economic traits forms the basis for the selection of a particular method for supersaturation generation.

$$\Delta c = c - c_s \quad (2.1)$$

The solution is said to be supersaturated when Δc is greater than zero.

2.1.2 Metastable zone

The initial stage in the design of crystallization process is to determine the width of metastable zone. Metastable zone is a region of supersaturated solutions where even in the absence of stability no solid phase separates out. The width of the metastable zone depends on several features like coolant flow rate rate, solute concentration, impeller speed, presence of impurities, etc. It can be demarcated as the difference in saturation temperature to the temperature at which first crystals are formed. This temperature difference is named as the maximum undercooling, ΔT_{\max} . The solubility and metastable limit of a system cooled from under saturation to supersaturation is illustrated in Figure 2.1

According to [Ostwald](#) (1897), the phase diagram for crystallization can be divided into three distinct regions namely under saturated, Metastable supersaturated and unstable supersaturated zone.

1. Under saturated zone – In this zone the crystals present in the solution get dissolved making crystallization process infeasible.

2. Metastable supersaturated zone – This zone facilitates the growth of crystals as there are no nucleation of the supersaturated solution.

3. Unstable supersaturated zone – In this zone spontaneous nucleation take place.

In cooling crystallizations solute concentration profile is maintained well within the metastable zone and close to the solubility curve in order to promote crystal growth and to avoid secondary nucleation. In industrial crystallizers the level of supersaturation is retained at about half the metastable width.

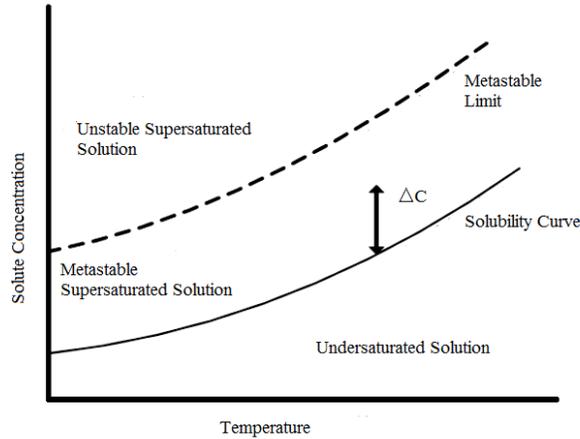


Figure 2.1: The Phase Diagram

]

2.1.3 Crystallization kinetics

Dynamic evolution of crystal population to the state variables of the continuous phase is related by crystallization kinetics. Crystallization kinetics can be explained by crystal growth and nucleation. Nucleation and growth takes place only if supersaturation is positive, indicating a thermodynamically favorable condition for transition of the solute into the solid phase.

2.1.3.1 Nucleation

Nucleation is the process of formation of a new crystalline phase. Nucleation can be broadly categorized as primary nucleation and secondary nucleation. Primary nucleation takes place in the absence of crystalline surfaces. Primary nucleation can further classified as homogeneous and heterogeneous nucleation. Homogeneous nucleation takes place spontaneously in a pure solution through the formation of clusters of solute molecules. On the other hand, heterogeneous nucleation is induced by the presence of impurities. Primary nucleation when the level of supersaturation is high typically when the metastable limit is exceeded. A common practice in industrial crystallization is to dodge primary nucleation since it is often unfavorable to the product quality. This can be done by maintaining the supersaturation within the metastable zone.

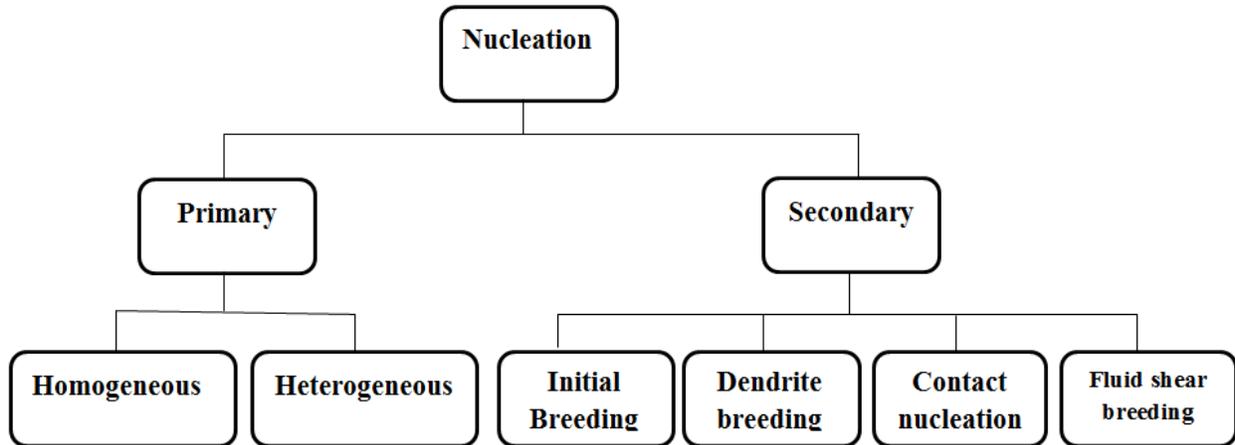


Figure 2.2: Mechanisms of Crystal Nucleation

Secondary nucleation takes place at low levels of supersaturation. Figure 2.2 shows a variety of secondary nucleation mechanisms. Initial breeding happens when crystalline dust fragments adhered to the surface of dry seed crystals are washed off upon introduction into a supersaturated solution. Breakage of dendrites from the parent crystals leads to dendrite breeding, whereas fluid shear breeding transpires due to the breakage of a growing crystal face under hydrodynamic shear forces. Whereas, contact nucleation is a result of mechanical forces, namely vigorous crystal-crystal, crystal-impeller, and crystal-hardware collisions.

Among the various secondary nucleation mechanisms, contact nucleation generally dominates in industrial crystallization. Attrition is a two-step process consisting of the generation of attrition fragments and their subsequent growth. They are created mainly due to crystal-crystal and crystal-hardware collisions. The formation of secondary nuclei is dependent on the properties of the crystalline material and the hydrodynamic conditions inside the crystallizer. Mechanical stress can result in distortion of the crystal lattice of the attrition fragments. The lattice strain intensely affects the existence and the outgrowth of these fragments.

2.1.3.2 Crystal growth

Crystal growth begins once stable nuclei have been shaped in a supersaturated solution. These stable nuclei begin to grow into crystals of observable size. The surface of crystals increases as the solute molecules are assimilated into the accessible and energetically favorable kinks or steps of the crystal lattice.

Crystal growth is primarily a two-step process. In the first step the solute molecules diffuse towards the crystal surface and integration of these growth units into a crystal lattice takes place followed by mass transfer of the solute molecules from bulk through a stagnant layer around the crystal surface. In this stage surface diffusion, spatial orientation, and the subsequent surface integration reaction takes place. The individual steps of the crystal growth are often described in terms of power-law functions of supersaturation.

In this study, the crystal growth rate $\left(= \frac{dL}{dt}\right)$, is represented by the empirical expression

$$G = \alpha L^b \exp\left(-\frac{E}{RT}\right) \Delta c^g \quad (2.3)$$

2.1.4 Seeding

Seeding is a technique that is used in a crystallization process to control the size and amount of crystals produced. Addition of seed crystals occurs prior to crystallization. It provides surface for nucleation and crystal growth. Hence, can control the inception of crystallization. The key step in seeding process is to ensure that the seeds are not added too early. This information is very significant in context of production routine. This highlights why the information of the solubility curve and metastable width are important.

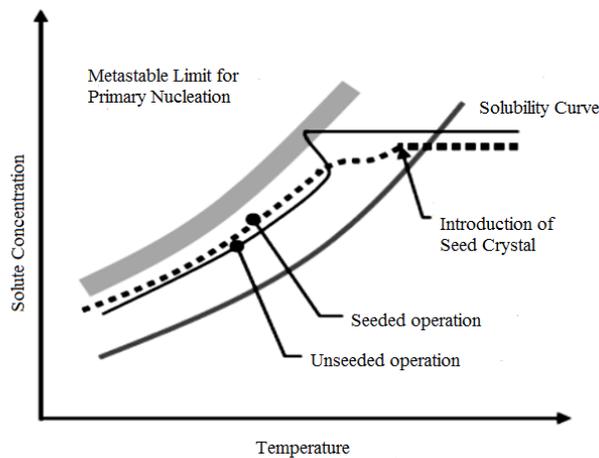


Figure 2.3: Seeding Policy

Some of the key variables to be considered in seeding include seed loading, size of the seeds and when and how they are added. Naturally, smaller seeds provide higher surface area per

unit mass and hence more opening for growth-controlled process. Smaller crystals provide more area for growth as compared to larger ones and thus the process becomes secondary nucleation controlled. Typically, higher amount and smaller size of seeds makes the crystallization process easier to control with minimum process disturbances. Seeds are added in dry or in wet form. When seeds are added as slurry, it becomes easier for the seeds to disperse to their prime crystal size. Whereas adding dry seeds results solvent entrapment and agglomeration.

2.2 Control of Crystallizers

With the advancement in process analytical technology and computing power various interesting control approaches have been proposed to meet the increasing demand of high quality product. These control approaches can be broadly categorized as ‘model based controlled approach’ and ‘direct design approach’.

2.2.1 The Direct Design Approach

The root of the direct design approach for controlling a crystallizer is the knowledge that the desired area of operation of crystallizers are within the metastable zone (Fig.2.4), which is constrained by the solubility curve and the metastable curve.

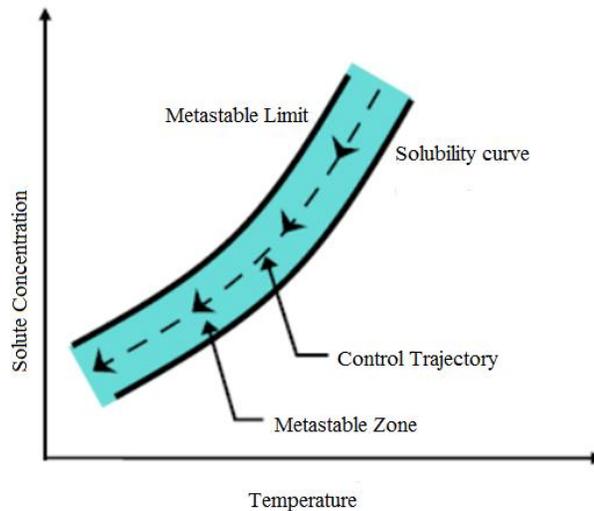


Figure 2.4: The Direct Design Approach

Here, the set point trajectory (solute concentration) is intended to lie inside the metastable zone and a feedback control system is provided to track the solute concentration. The set point is fixed based on the fact that crystals grow at a faster rate near the metastable curve, and the level

of nucleation is low near the solubility curve. This method stands with the advantage that accurate kinetics of growth and nucleation need not be derived. Implementation of this method, call for knowledge of the metastable zone and solute concentration. The popularity of this method is stirred by recent advances in the in situ sensors, which permits automatic evaluation of the metastable zone.

2.2.2 The Model Based Control Approach

Basis of this method is a first-principles model constructed from material and energy balances. Probability balance equation (PBE) gives the solid balance and defines the effect of crystallization kinetics on crystal size distribution. The distributed nature of PBE makes the model based control difficult to handle. This problem is solved by approximating PBE by method of moments. And these crystal moments describe various physical properties of crystal growth like number, length, area and volume. But knowledge of CSD is not given by PBE. The primary objective of model based control approach is production of crystals having large mean crystal size and a narrow crystal size distribution.

First-principles models provide insights into crystallization processes by enabling systematic simulation and optimization of various kinetic phenomena. The drawback of this method is that it requires accurate knowledge of crystallization kinetics. The arrival of in-situ sensors and the use of multivariate statistics have largely improved reliable data collection and analysis for identification of the kinetic model. However, modeling of more complex kinetic phenomena, e.g. agglomeration, dendritic growth, polymorphic transitions, etc., yet remains to be a challenge.

Moreover, kinetic parameters are typically determined at laboratory-scale. Achieving the same supersaturation and localized mixing in an industrial-scale crystallizer is almost impractical due to different heat transfer and mixing characteristics. To add to this the kinetic parameters are largely sensitive to feed impurities. The shortfalls of the identified kinetic models can severely degrade the effectiveness of a model-based control approach at industrial level. In this thesis model based control approach has been used for the batch as well as continuous crystallizer.

REFERENCES

1. Alatalo H, "Supersaturation -controlled crystallization", PhD thesis, Lappeenranta University of Technology, Finland, 17-27, 2010.
2. Doki N, Kubota N, Sato A, "Effect of Cooling Mode on Product Crystal Size In seeded Batch Crystallization of Potassium Alum", Chemical Engineering Journal ,81 ,313-316, 2001.
3. Griffiths H, "Mechanical crystallization". Journal of the Society of Chemical Industry, 44, 7–18, 1925.
4. Mullin J W, "Crystallization", Butterworth's and Heinemann London, 3rd Edition, 7, 354,1993.
5. Mesbah A , "Optimal Operation of Industrial Batch Crystallizers -A Nonlinear Model-based Control Approach", Journal of the Chinese Institute of Chemical Engineers , 26, 147-156, 2010.
6. Wohlk W, Hofmann G, "Types of Crystallizers", International Chemical Engineering, 27, 197-201, 1987.

Chapter 3

BATCH CRYSTALLIZER

Chapter 3

BATCH CRYSTALLIZER

3.1 Introduction

Batch crystallization from solution is used in pharmaceutical and agrochemical industries for solid–liquid separation as well as purification. They are used for the production of both low-volume and high-value chemicals. A batch crystallizer produces crystals of large mean crystal size and narrow crystal size distribution which has direct effect on the final product qualities. Therefore, to increase the batch crystallization performance and product quality it is very important to control the CSD. Supersaturation trajectory which is a function of temperature has direct effect on CSD. Thus the crystallizer temperature serves as the manipulated variable for cooling crystallizer. But the temperature of the crystallizer cannot be manipulated. Only the inlet temperature of the cooling medium can be influenced. So in most of the cases slurry temperature is set as the manipulated variable which controls supersaturation.

The need to produce products with specified crystal size distribution calls for an optimized crystallization process. Some of the objectives of an optimized batch crystallizer are, maximize the mean crystal size, maximize the terminal seed crystal size, minimize the coefficient of variation of crystals and to reduce seed crystal mass to the nucleated crystal mass ratio etc. (Debasis et al., 2003) . Griffiths (1925) suggested maintaining the cooling profile of a batch crystallizer well within the metastable zone results in better CSD. However, no computable procedures for optimal operation of a batch crystallizer was proposed. Mullin and Nyvlt (1971) ,developed a cooling method and it was considered to be the optimal method for batch crystallization in the programmed cooling mode. Further, a more systematic method that mathematically calculates the programmed cooling profiles that takes nucleation also into account was suggested by Jones and Mullin (1974).According to the studies conducted by Jagadesh et al.(1996) using a laboratory crystallizer, a unimodal distribution of the crystals can be produced with virtually no nucleation under natural cooling mode if enough seeds are added. Almost same results were reported by Doki et al. (2001) for a large scale crystallizer. The cooling mode suggested here is very easy to operate. But it is not considered to be optimal for attaining large / desired product size.

A single optimal control policy was generated by [Chang and Epstein \(1982\)](#). [Zhang and Rohani \(2003\)](#) proposed an optimal policy by combining the objectives of achieving CSD with a large mean size and small CV into one objective by a weighted sum approach. Similarly, single-objective optimization approach using a penalty for time of operation was used to solve the problem of achieving the largest possible final crystal size in the shortest possible batch time ([Choong and Smith, 2004](#)). The effect of different cooling policies on the supersaturation level in batch crystallizers has been widely investigated. The effect of cooling profile on supersaturation level for a batch crystallizer that produces ammonium sulfate was investigated by [Hojjati and Rohani \(2005\)](#). Four cooling strategies including natural, linear, controlled, and impulse change in natural cooling policy were investigated and the results showed that at low seed loading, control of cooling policy is must to ensure narrow crystal size distribution with large mean size.

In this work a mathematical model of batch cooling crystallizer was resolved for the three different cooling approaches, natural, linear and controlled cooling using terminologies for nucleation and crystal growth rates and solubility determined for potassium sulphate solutions. These theoretical considerations were used to predict the temperature, supersaturation and the product distribution. The population balance equation together with the mass and energy balances is converted into the set of moment equations for the first four moments having physical meaning. The properties of the crystalline product and dynamic behavior of the crystallizer obtained by numerical experimentation carried out by using a computer program in MATLAB environment are presented and analyzed.

3.2 MATHEMATICAL MODELING OF BATCH COOLING CRYSTALLIZER

Crystallization process is complex process. To predict the performance of a crystallizer proper understanding of the phenomenon like nucleation, growth, attrition, breakage and dissolution are required. A mathematical outline for modeling a batch crystallization process comprises of the population balance equation (PBE) which describes a conservation equation for number of crystals in a population.

In order to simplify the mathematical model of the process following assumptions were made: the system is well mixed in both phases, no volume changes, crystals or solution are

neither removed nor added, and all same sized crystals grow at same rate. With these assumptions population balance for a batch crystallizer with one specific growth dimension is given as:

$$\frac{\partial p}{\partial t} + \frac{\partial(pG)}{\partial L} = B_n = b_n \varphi_n(L); \quad L_n(0) > 0 \quad (3.1)$$

$$\frac{\partial p}{\partial t} + \frac{\partial(pG)}{\partial L} = 0 \quad ; \quad L_n(0) = 0 \quad (3.2)$$

Were p is the population density; B and G are the birth and growth rate respectively. When the growth rate is independent of crystal size ($G \neq G(L)$) Eq. (3.2) can be written as:

$$\frac{\partial p}{\partial t} + G \left(\frac{\partial p}{\partial L} \right) = 0 \quad (3.3)$$

This equation was derived by [Branson and Dunning \(1949\)](#)

The population density p in equation is demarcated as “unit mass of solvent” basis as we have taken that the volume changes are considered insignificant and an anhydrous salt is considered. The solution of eq. (3.1) comprises of 2 parts, the population balance of the seed crystals as well as generated nuclei. In order to distinguish between the two, former is denoted as S crystals and the latter as N crystals, and their corresponding number densities are continuous functions of size and time.

Mass balance for the batch cooling crystallizer is given as:

$$-\frac{dc}{dt} = \frac{3W_{s0}}{L_{s0}^3} L_s^2(t) G_s(L, t) + \alpha \rho_c \left[3 \int_0^\infty p_n(L, t) G_n(L, t) L_n^2(t) dL + J_n(t) L_{n0}^3 \right] \quad (3.4)$$

$$J_n = k_n \Delta c^n \quad (3.5)$$

Equation (3.4) defines the rate of change of solute concentration in a supersaturated solution due to the growth of the S -crystals, the generation and subsequent growth of the N -crystals. J_n is the number rate, k_n nucleation rate constant and Δc supersaturation.

The saturation rate, is written as:

$$R_s = \frac{dc_s}{dt} = \frac{dT}{dt} \cdot \frac{dc_s}{dT} = C \frac{dc_s}{dt} \quad (3.6)$$

Adding $\frac{dc_s}{dt}$ to both sides of the equation (3.4) gives the supersaturation equation:

$$-\frac{d\Delta c}{dt} = C(T, t) \frac{dc_s}{dT}(t) + \frac{3W_{s0}}{L_{s0}^3} L_s^2(t) G_s(L, t) + \alpha \rho_c [3 \int_0^\infty p_n(L, t) G_n(L, t) L_n^2(t) dL + J_n(t) L_{n0}^3] \quad (3.7)$$

3.2.1 Moment Transformation

The method of moments is employed as a model reduction technique to obtain the analytical solution of PBE. This method allows us to condense the PBE to a set of nonlinear ordinary differential equation. These ordinary differential equations enable us to efficiently solve the crystallization model to pre specified accuracy. The moment model is particularly suitable for control applications. The moment transformation describes the j^{th} moment with regard to population density function:

$$\mu_j = \int_0^\infty p L^j dL \quad j = 0, 1, 2, 3 \dots \quad (3.8)$$

In addition, the four moments of crystal size distribution have physical significance:

μ_0 = total crystal number per unit volume

μ_1 = total crystal length per unit volume

μ_2 = total crystal surface area per unit volume

μ_3 = total crystal volume per unit volume.

$$\frac{d\mu_{n0}}{dt} = J_n = k_n \Delta c^n \quad (3.9)$$

$$\frac{d\mu_{n1}}{dt} = \mu_{n0} \varphi_n \phi + \frac{d\mu_{n0}}{dt} L_{n0} \quad (3.10)$$

$$\frac{d\mu_{n2}}{dt} = 2\mu_{n1} \varphi_n \phi + \frac{d\mu_{n0}}{dt} L_{n0}^2 \quad (3.11)$$

$$\frac{d\mu_{n3}}{dt} = 3\mu_{n2} \varphi_n \phi + \frac{d\mu_{n0}}{dt} L_{n0}^3 \quad (3.12)$$

The supersaturation balance Eq. 3.7 may be written in terms of Eq. 3.8 where $j = 3$, to give:

$$-\frac{d\Delta c}{dt} = C(\theta, t) \frac{dc_s}{dT}(t) + \frac{3W_{s0}}{L_{s0}^3} L_s^2(t) G_s(L_s, t) + \alpha \rho_c \frac{d\mu_{n3}}{dt} \quad (3.13)$$

These moment equations is used in the solving the operating policies considered.

The crystal growth rate, $G(= \frac{dL}{dt})$, may be represented by the empirical expression:

$$G = \alpha L^b \exp\left(-\frac{E}{RT}\right) \Delta c^g \quad (3.14)$$

or can be written as

$$G = \varnothing(L) \phi(c, \theta) \quad (3.15)$$

$$\varnothing(L) = L^b \quad (3.16)$$

$$\phi(c, \theta) = \alpha \exp\left(-\frac{E}{RT}\right) \Delta c^g \quad (3.17)$$

3.2.2 Energy Balance

For the programmed batch cooling crystallizer the energy balance is given by:

$$\frac{dT}{dt} = C(T, t) \quad (3.18)$$

here the temperature decreases with time and the heating periods are neglected.

Here three basic cooling modes have been considered

- Natural cooling
- Linear cooling
- Controlled cooling

3.2.2.1 Natural Cooling

In natural cooling, the batch is cooled from a prominent temperature to room temperature by natural convection (Newton's law of cooling) and crystallization is uncontrolled. Here there is a rapid drop in the solution temperature early in the batch and the solution supersaturates at a faster rate than it is de supersaturated by growth of crystals. Thus supersaturation swiftly crosses the metastable curve and excessive nucleation takes place. Finally these nuclei grow and as a consequence an irregular product distribution is obtained.

The cooling rate for natural cooling $C(T, t)$ is described by Newton's law (i.e., the cooling rate is proportional to the excess temperature):

$$\frac{dT}{dt} = -K_1\{(T(t) - T_w)\} \quad (3.19)$$

where T_w , Is the coolant temperature. K_1 , is associated to the initial and final boundary conditions on temperature, T_0 and T_f , and the batch time, τ , by the equation :

$$K_1 = \frac{1}{\tau} \ln \left\{ \frac{T_0 - T_w}{T_f - T_w} \right\} \quad (3.20)$$

3.2.2.2 Linear Cooling

In Linear cooling the temperature tracks a constant cooling rate. The cooling rate is given by:

$$\frac{dT}{dt} = -K_2 \quad (3.21)$$

K_2 , is associated to the initial and final boundary conditions on temperature, T_0 and T_f , and the total batch time, τ , by the equation :

$$K_2 = \frac{T_0 - T_f}{\tau} \quad (3.22)$$

3.2.2.3 Controlled Cooling

Modern crystallizers are often operated with some degree of control over cooling as uncontrolled cooling usually degrades the quality of crystals. In order to control the crystal size distribution of the product a proper control of supersaturation throughout the operation is must. Such a control is possible by maintaining the solute concentration well within the metastable zone. This results in a better product crystal size as nucleation becomes negligible (3.5) and growth essentially occurs on the seeded crystals. In order to obtain large crystals the temperature should be reduced gradually in the early stages and more rapidly towards the end of the batch. This mode of operation is termed controlled cooling.

Here the controlled cooling is illustrated by the case of constant nucleation rate. If the rate of nucleation is assumed to be reliant only on the supersaturation

$$R_n(t) = k_n \Delta c^n(t) \quad (3.23)$$

then for controlled cooling :

$$\Delta c(t) = \Delta c_0 = \text{constant} \quad (3.24)$$

$$\frac{d\Delta c}{dt} = 0 \quad (3.25)$$

The equation (3.30) can be rearranged to obtain cooling rate

$$\frac{dT}{dt} = - \left[\frac{d\Delta c}{dt} + \frac{3W_{s0}}{L_{s0}^3} L_s^{2+b} \phi + \alpha \rho_c \frac{d\mu_{n3}}{dt} \right] / \left(\frac{dc_s}{dT} \right) \quad (3.26)$$

3.3 DYNAMIC SIMULATION OF BATCH CRYSTALLIZATION PROCESS

The cooling curves, transient supersaturation, nucleation rate and product CSD were computed from the equations given above. The seeded batch crystallization of potassium sulfate is considered. The following physical properties and initial conditions were used to obtain the dynamic simulation results of batch cooling crystallizer. (Table 3.1)

Table 3.1: The physical parameters and initial conditions of batch crystallizer

Parameters	Value	Parameters	Value
K_n	6.67×10^{10}	n	7.63
α	100	g	2
$-E/R$	2.4×10^3	L_{n0}	6×10^{-5} m
L_{s0}	550×10^{-6} m	$\Delta C(0)$	7.32×10^{-3} kg/kg
W_{s0}	2.36×10^{-3} kg/kg	T_0	333 K
T_f	296 K	T_w	288 K
τ	10800 sec	ρ_c	2.3×10^{-3} kg/m ³
s_0	0.088	s_1	0.00156

Dynamic simulation is obtained by using ODE 45 solver in MATLAB software. Dynamic responses of state variables are shown in figure 3.1.

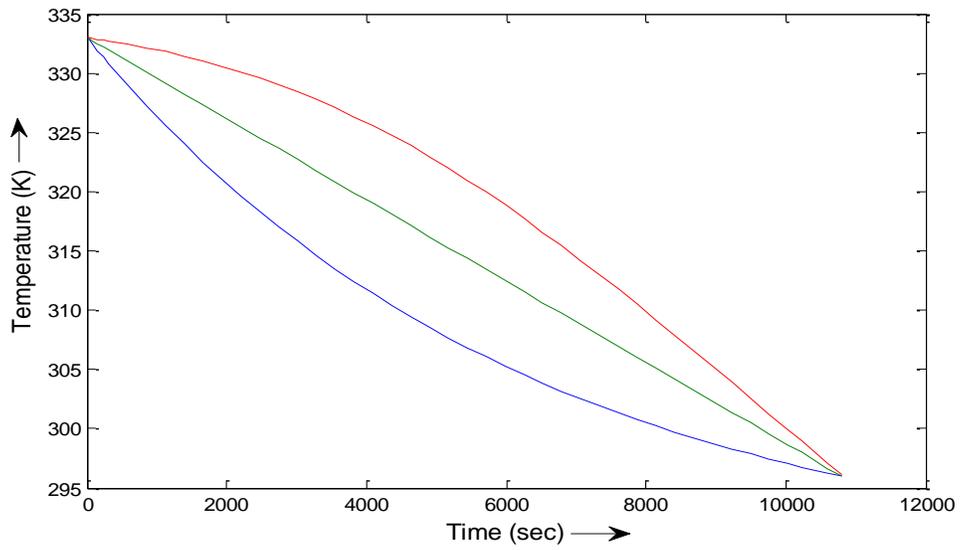


Figure 3.1 Comparison of operating policy for batch cooling crystallization. (— Natural, — Linear, — Controlled)

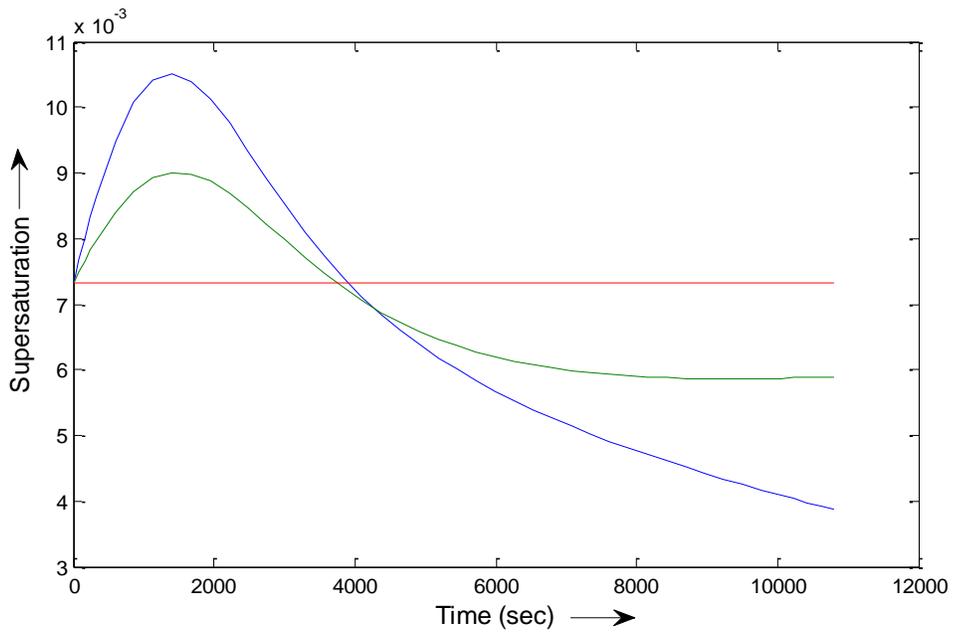


Figure.3.2 Effect of cooling curves on transient supersaturation. (— Natural, — Linear, — Controlled)

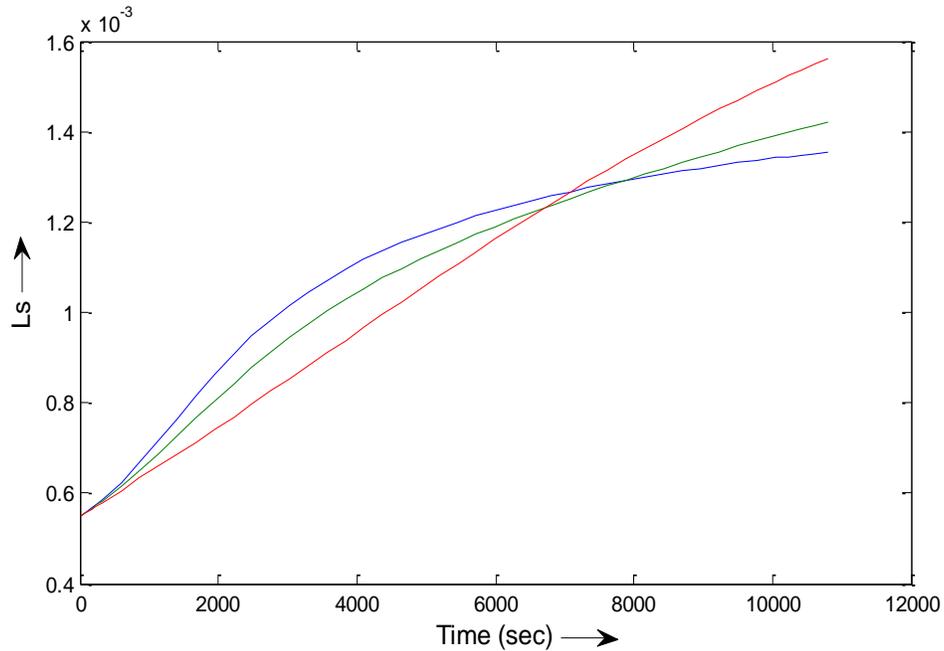


Figure.3.3: Effect of cooling curves on seed crystal length. (— Natural, —Linear, —Controlled)

It is quite clear from the figure 3.1 that the temperature profile for natural cooling is concave in nature. i.e. the slope of the temperature profile is maximum at the commencement of the batch. So in order to attain natural cooling profile there should be a very swift drop in the crystallizer temperature early in the batch. In case of controlled cooling the temperature profile is convex in nature, i.e. temperature drop is maximum at the end of the batch, and for linear cooling the temperature drop remains constant throughout.

The degrees of supersaturation corresponding to various cooling profiles studied are shown in figure 3.2. Initial supersaturation peaks are obtained in case of natural as well as linear cooling. The effect is more prominent in natural cooling. We can see that the terminal supersaturation and rate of nucleation are in the reverse order. That is the final supersaturation for natural cooling is lesser than that of linear, which is lower than that for controlled cooling. Thus we can say that natural cooling profile corresponds to lowest residual supersaturation and thus produces more amount of solute is been formed. Thus, natural cooling does not give the optimal crystal size distribution.

Constant nucleation rate cooling shows an improvement in the terminal size of the seed crystals compared with both linear and natural cooling (Fig.3.3). This is due to the fact that, natural cooling results in the formation of huge amount of nuclei as the supersaturation crosses the metastable curve and these nuclei grow and compete for solute with the seeded crystals. Whereas, in the case of controlled cooling supersaturation is maintained at a level within the metastable limit throughout the batch, so nucleation would be negligible and an better product crystal size is be obtained.

3.4 Crystal Size Distribution

The primary objective of crystallization process in industries is to produce solid products with desired purity and crystal size distribution. Crystal size distribution is a very important factor that determines the efficiency of downstream operations as well as the effectiveness of the product. In the case of batch cooling crystallization quality of the product is emphasized from the crystal size distribution obtained at the end of a batch. The final CSD is dependent on supersaturation profile which is a function of temperature. Hence a proper control strategy that gives the desired supersaturation trajectory is needed. As the performance of model base control approach is highly prone to uncertainties, a model that can effectively explain the process and accurately estimate the model parameters is of utmost importance. The CSD of batch crystallizer is accounted by considering primary nucleation, secondary nucleation, crystal growth, breakage and aggregation. Due to the unavailability of an accurate theoretical model to express the crystallization kinetics, nucleation and growth rates are expressed as empirical power law expressions.

In this study following properties of final CSD is investigated: number average size, final mean size of crystals and coefficient of variation.

$$\text{Number Average Size} = \frac{\mu_1}{\mu_0} \quad (3.27)$$

$$\text{Final mean size} = \frac{\mu_3}{\mu_2} \quad (3.28)$$

A crystal size distribution with a large mean size and minimum coefficient of variation (CV) is highly desirable for the end product.

Table 3.2: Effect of cooling policies on product CSD

Operating Policy	Terminal Size of 'S' Crystals (μm)	Number Average Size (μm)	Final Mean Size (μm)	Coefficient of Variation (%)
Natural Cooling	1355	380	1156	50
Linear Cooling	1421	421	1270	48
Controlled Cooling	1561	491	1500	44

The effect of various cooling policies on the product CSD's is shown in Table 3.2 where the potential improvement of controlled cooling profile over the other policies is highly visible. Constant nucleation rate cooling results in the increase in final size of the seed crystals, the mean crystal size and final mean size. A further improvement in the terminal size of seed crystal is possible if we eliminate nucleation completely.

3.5. CONCLUSION

The CSD from a batch cooling crystallization process was prophesied by the numerically solving the mathematical model using the crystallization kinetics like nucleation and crystal growth. The results clearly portray the prospective benefits of controlled cooling for the betterment of the product crystal size over that gained by natural and linear cooling.

Controlled cooling significantly minimizes the extent of nucleation and improves the size of the terminal crystals, compared to natural cooling. The ultimate outcome is significant improvement in the number average size, final mean size and coefficient of variation of the whole crystal distribution as well as the terminal size of the seed crystals.

REFERENCES

1. Ajinkya M.B., Ray W.H., "On the Optimal Operation of Crystallization Processes", *Chemical Engineering Communications*, 1, 181-186, 1974.
2. Chang, Epstein C.T., "Identification of Batch Crystallization control Strategies Using Characteristic Curves", *A.I.Ch.E Symposium Series*, 78, 68-75, 1982.
3. Choong K.L., Smith R., "Optimization of Batch Cooling Crystallization", *Chemical Engineering Science*, 59, 313–327, 2004.
4. Doki N., Kubota N., Sato A, "Effect of Cooling Mode on Product Crystal Size In seeded Batch Crystallization of Potassium Alum", *Chemical Engineering Journal* ,81 ,313-316, 2001.
5. Jones A.G., "Optimal Operation of a Batch Cooling Crystallizer", *Chem. Eng. Sci.*, 29, 1075-1086, 1974.
6. Jones A.G., Mullin J.W., "Programmed Cooling Crystallization of Potassium Sulfate Solutions", *Chem. Eng. Sci.*, 29, 105-117, 1974.
7. Jagadesh D., Kubota N., Yokota M., "Large and Mono-size Product Crystals from Natural Cooling Mode Batch crystallization", *Journal of Chemical Engineering of Japan*, 29, 865-873, 1996.
8. Griffiths, H., "Mechanical crystallization". *Journal of the Society of Chemical Industry*, 44, 7–18, 1925.
9. Hojjati H., Rohani S., "Cooling and Seeding Effect on Supersaturation and Final Crystal Size Distribution (CSD) of Ammonium Sulphate in a Batch Crystallizer", *Chemical Engineering and Processing*, 44, 949-957, 2005.
10. Mullin J.W., Nyvlt J., "Programmed Cooling of Batch Crystallizers", *Chem. Eng. Sci.*, 26, 369, 1971.
11. Sarkar D., Rohani S., Jutan A. , "Multi Objective Optimization of Seeded Batch Crystallization Processes", *Chemical Engineering Science*, 61, 5282-5295 , 2006.
12. Zhang G., Rohani, S., "On-line Optimal Control of a Seeded Batch Crystallizer", *Chemical Engineering Science*, 58, 1887–1896, 2003.
13. Zoltan K., Chewb J., Fujiwara M., "Comparative performance of concentration and temperature controlled batch crystallizations", *Journal of Process Control*, 18, 399–407, 2008.

Chapter 4

CONTINUOUS CRYSTALLIZER

Chapter 4

CONTINUOUS CRYSTALLIZER

4.1 INTRODUCTION

Continuous crystallizers are used in industry for the manufacture of chemical products as well as for recovery of valuable products. They are suited for the production of large tonnage bulk commodity crystalline products as they can be operated at desired operating condition. The production rate of continuous crystallizers vary from 1ton/day to 50 ton/day.

In continuous crystallizers, full use of existing plant capacity is possible as no down-time is needed for filling and emptying of crystallizers as is required for batch operation. The often quoted advantages for continuous crystallizers compared to batch crystallizers are better optimal operating conditions, reduced size of crystallizers and tankage. And some of the drawbacks are continuous crystallizers take time to attain steady state, instability of operation and accumulation problems. Continuous crystallizers are intended to create and maintain constant growing conditions for the crystal product. Due to the large dimension of the equipment it is difficult to attain uniform distribution of the temperature, solute concentration and crystal suspension. Therefore, a proper control strategy is must.

Control of crystallizers has been the topic of many researches since early 1960s. In the previous studies control schemes were mainly concentrated on the control of crystal size distribution (CSD). A feedback control scheme that can overcome instabilities observed in the CSD of continuous crystallizers was proposed by [Randolph et al. \(1977\)](#). Many works emphasizing on the choice of a proper measured and a manipulated control variable for control of crystal size distribution are reported in literature ([Rohani 1986](#), [Randolph et al. 1987](#)). But in majority of these studies complexities in controlling the process have not been considered. In crystallization process apart from CSD there is other variables like crystal purity, crystal shape and productivity that are to be considered.

Studies were conducted by [Sherwin \(1967\)](#) and [Randolph \(1977\)](#) on the stability of well mixed crystallizers. [Vollmera and Raisch \(2002\)](#) based on stability analysis of simple model crystallizer transfer function designed stabilizing feedback H_{∞} controller. [Yin et al. \(2003\)](#)

investigated the stability and dynamics of linearized model of a continuous crystallizer were crystal growth rate was size-dependent. Analyzes of controllability and observability of MSMMPR crystallizer as well as the coupling between the inputs and the outputs by simulation using the linearized moment equation model were reported by [Moldoványi et al.\(2005\)](#)The dynamic behavior of isothermal CMSMPR together with their bifurcation behavior, period doubling cascades, quasi periodic oscillations etc. were considered by [Lakatos et al.\(1995\)](#).

[Ramanathan \(1988\)](#) investigated the stability and dynamic behavior of isothermal CMSMPR in Laplace domain. The characteristics of these models, considering the zeroes and poles location and their effects on system stability and frequency response were discussed in detail. [Shirvani et al.\(1995\)](#)developed a simplified model for complex transfer function models in frequency response domain. These frequency response data of a process system provides excellent recognizable information about structure of a lumped process system as well as a distributed process system such as the crystallizer systems which can be detected clearly in a heuristic manner. [Alberto et al \(2012\)](#) designed a feedback control algorithm for non-isothermal seeded continuous crystallizer. The control objectives was steadying of the third moment of the crystal size distribution and the temperature of the crystallizer considering impeller speed and the jacket flow rate as the manipulated variables. [Heidari et al.\(2012\)](#) contributed a linearized dynamic model for calculating Cumulative Mass Fraction (CMF) of Potassium Nitrate-Water Seeded Continuous Mixed Suspension Mixed Product Removal (CMSMPR) crystallizer by approximated simplified model in frequency domain.

This chapter covers the mathematical modeling and simulation of a non-isothermal continuous cooling crystallizer. Analyzes of the operating conditions were done based on the combined effects of secondary nucleation and cooling temperature. Finally, a closed loop simulation was developed using traditional PID controller. The model developed here was used to generate the database for performing various multivariate statistical process control techniques

4.2. MATHEMATICAL MODELING OF NON ISOTHERMAL CONTINUOUS COOLING CRYSTALLIZER

Presence of two nonlinear discrete kinetic steps: nucleation and growth makes the continuous crystallization a highly complex and nonlinear process. In addition, there are interactions between kinetic, fluid dynamics, and crystal size distribution. A mathematical

background for modeling a continuous crystallization process consists of the population balance equation (PBE) which describes a conservation equation for the amount of crystals in a population.

In order to streamline the mathematical model of the process following assumptions were made: crystallizer is operated in a continuous mode, perfectly mixed, constant volume and infinitesimal new generated particles size. With these assumptions population balance for a continuous crystallizer is given as:

$$\frac{\partial n}{\partial t} + \frac{\partial(nG)}{\partial L} = - \sum_k \frac{(Q_k p_k)}{V} + B\delta(L - L_0) \quad (4.1)$$

Where n is the population density, L is characteristic crystal length, G is the crystal growth rate, B is the nucleation rate, and $\delta(L - L_0)$ is the Dirac delta function acting at $L = L_0$. The initial and boundary conditions are given as:

$$\text{BD: } n(0, t) = 0 \quad n(\infty, t) = 0 \quad t \geq 0 \quad (4.2)$$

$$\text{IC: } n(L, 0) = f(L) \quad 0 \leq L \leq \infty$$

Mass balance for the continuous cooling crystallizer is given as:

$$\frac{dM_s}{dt} = \frac{d(V_L C)}{dt} = \frac{d(\epsilon V C)}{dt} - \sum_k \epsilon_k q_k C_k - R' \quad (4.3)$$

The mass balance gives the flow of solute in and out of the crystallizer and the mass transfer to the crystal phase through nucleation and growth. Here M_s is the mass of solute in the solvent, V is the total suspension volume, V_L is the suspension solids-free volume, C is the solute concentration, q is the volumetric flow rate, R' is the global mass transfer of solute and ϵ is the void fraction.

The energy balance gives the differences in enthalpy of the flow streams, the heat of crystallization and heat removed by the coolant. The following equation denotes the energy balance:

$$\rho V C_p \frac{dT}{dt} = \sum_k q_k H_k - \Delta H_c R' - H_{\text{ext}} \quad (4.4)$$

Where C_p stands for specific heat capacity, ρ is the slurry density, T temperature of the crystallizer, is H_k the enthalpy of the k^{th} stream, ΔH_c is the heat of crystallization and H_{ext} is the net heat removed.

In order to, develop complete model of the crystallization process empirical equations for growth and nucleation are required. Both of them are function of supersaturation and agitation rate as reported by [Salcedo-Estrada\(2000\)](#).

$$G = k_g \Delta c^g N_r^h \quad (4.5)$$

$$B = k_b \Delta c^g M_T^0 N_r^p \quad (4.6)$$

The driving force for both phenomena and is called the relative super-saturation Δc given by the equation (3.7):

$$\Delta c = \frac{(c - c_{\text{sat}})}{c_{\text{sat}}} \quad (4.7)$$

The crystallizer mentioned here is developed for operating in the metastable zone limit. Hence, knowledge of the limits of metastable zone is very important. The polynomial equation for the metastable zone limit (c_{mz}) for the ammonium sulfate-water system as function of crystallizer temperature is reported by Lugo-Martinez:

$$c_{\text{mz}} = 1.317 * 10^{-7} T^3 + 2.6136 * 10^{-6} T^2 + 0.70758 \quad (4.8)$$

The lower limit, saturation concentration (c_s) was estimated using the polynomial equation given by Perry et al. Both of these equations are effective in the range from 20 to 50°C.

$$c_s = 4 * 10^{-5} T^2 + 2 * 10^{-4} T + 0.73 \quad (4.9)$$

To avoid high computational time required in solving of the PBE, these equations are converted into a set of ode by the method of moments. In this context, the method of moments states the i^{th} moment in terms of the population density function as:

$$\mu_j = \int_0^{\infty} p L^j dL \quad j = 0,1,2,3 \dots \quad (4.10)$$

The moment transformation resulted in the following set of differential equations:

$$\frac{d\mu_0}{dt} = \frac{\mu_0^e}{\tau} - \frac{\mu_0}{\tau} + k_b \Delta c^b M_T^0 N_r^p \quad (4.11)$$

$$\frac{d\mu_j}{dt} = \frac{\mu_j^e}{\tau} - \frac{\mu_j}{\tau} + k_b \Delta c^g N_r^h j \mu_{j-1}, \quad j = 1,2 \dots 5 \quad (4.12)$$

$$\frac{dC}{dt} = \frac{(\varepsilon C)_e - C + C\mu_3^e - 3\tau k_g \Delta c^g N_r^h \mu_2 (\rho_c - C)}{\tau(1 - \mu_3)} \quad (4.13)$$

$$\rho V C_p \frac{dT}{dt} = \rho q C_p (T_{in} - T) - \Delta H_c 3\rho_c k_v V k_g \Delta c^g N_r^h \mu_2 - UA(T - T_j) \quad (4.14)$$

$$\rho_j V_j C_{pj} \frac{dT_j}{dt} = \rho_j q_j C_{pj} (T_{jin} - T_j) + UA(T - T_j) - U_0 A_0 (T_j - T_\infty) \quad (4.15)$$

These equations were solved using the SIMULINK tool box of MATLAB to get the open loop simulation of the continuous cooling crystallizer.

4.3. DYNAMIC SIMULATION OF CONTINUOUS CRYSTALLIZATION PROCESS

The dynamics of continuous crystallizer was computed from the above equations. The seeded batch crystallization of ammonium sulfate is considered. The following physical parameters and initial conditions were used, Table 3.1.

Table 4.1: The physical parameters and initial conditions of the seeded continuous crystallizer

PARAMETERS	VALUE	UNIT	PARAMETERS	VALUE	UNIT
ρ_c	1.769	g/cm ³	C(0)	0.772	g/l
ρ	1.244	g/cm ³	τ	2.6	min
C_p	2.85	cal/g ⁰ C	ΔH_c	-12.0354	cal/g
C_{pw}	1	cal/g ⁰ C	UA	1000	cal/min ⁰ C
V	3000	cm ³	$U_0 A_0$	100	cal/min ⁰ C
V_c	820	cm ³	T_{jin}	13	⁰ C
k_g	9.09×10^{-4}	(cm/min)(min/rev) ^h	T_{in}	30	⁰ C
k_b	180	(cm ³ /g) ⁰ (min/rev) ^h /min cm ³	μ_0	0.9772	
g	1.5		$\mu_1(0)$	1.044×10^{-2}	
h	1.337		$\mu_2(0)$	1.8213×10^{-4}	
b	.562		$\mu_3(0)$	2.3407×10^{-7}	
o	.001		$\mu_4(0)$	3.0159×10^{-8}	
p	.050		$\mu_5(0)$	3.8891×10^{-10}	

The following graphs show the open loop simulation of the state variables of non-isothermal continuous cooling crystallizer.

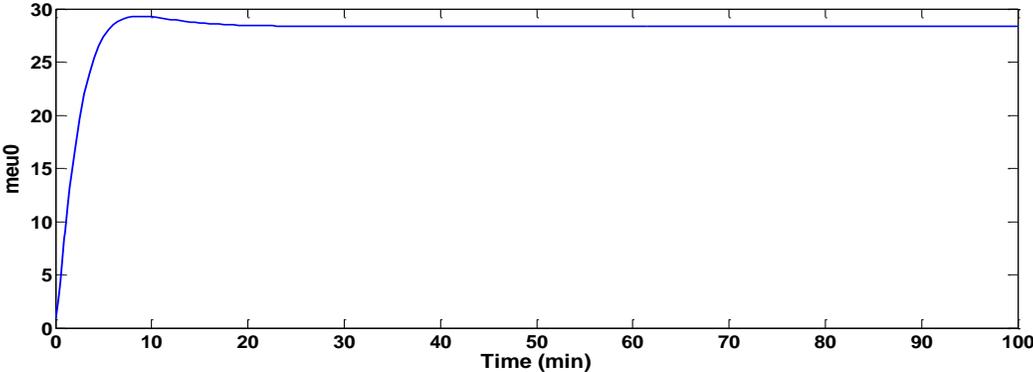


Figure 4.1: Zero order moment of crystallization

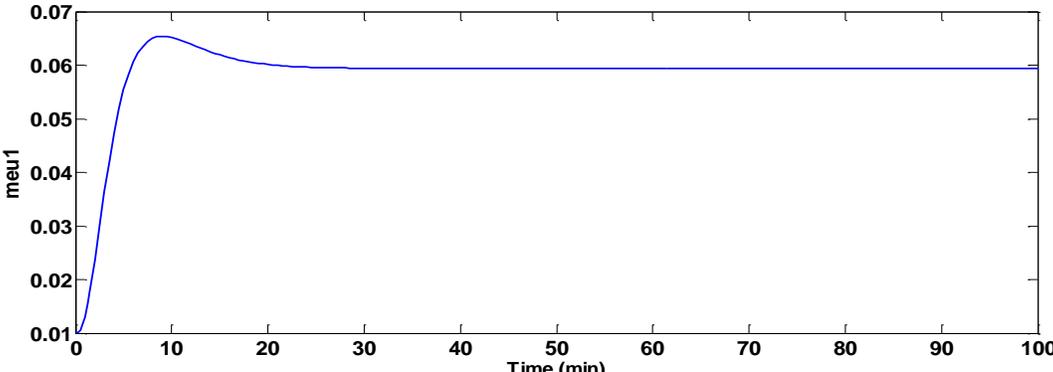


Figure 4.2: First order moment of crystallization

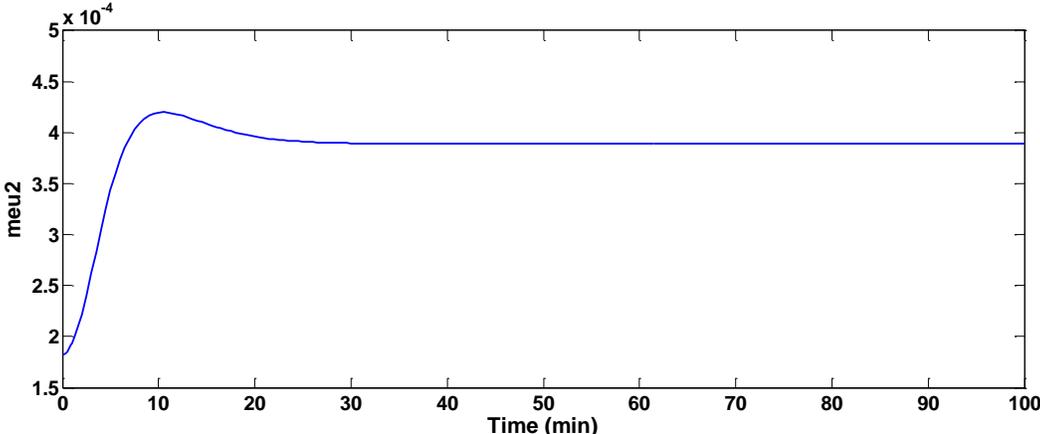


Figure 4.3: Second order moment of crystallization

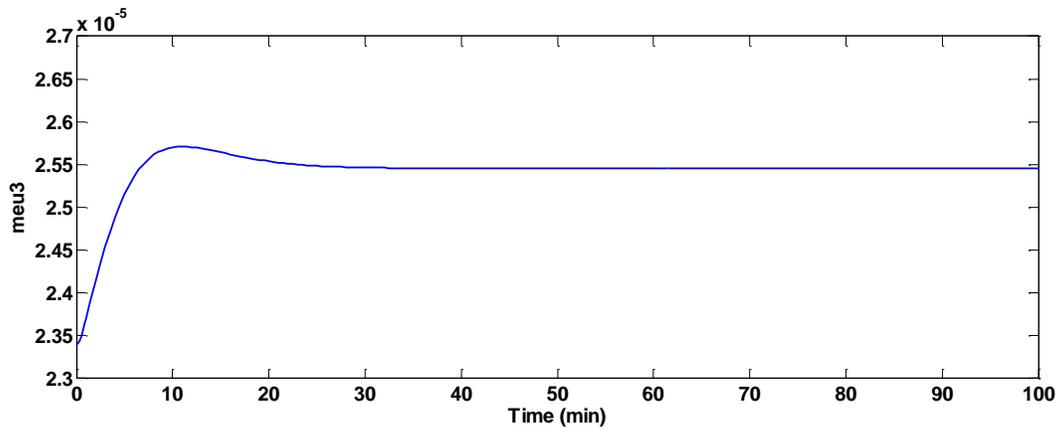


Figure4.4: Third order moment of crystallization

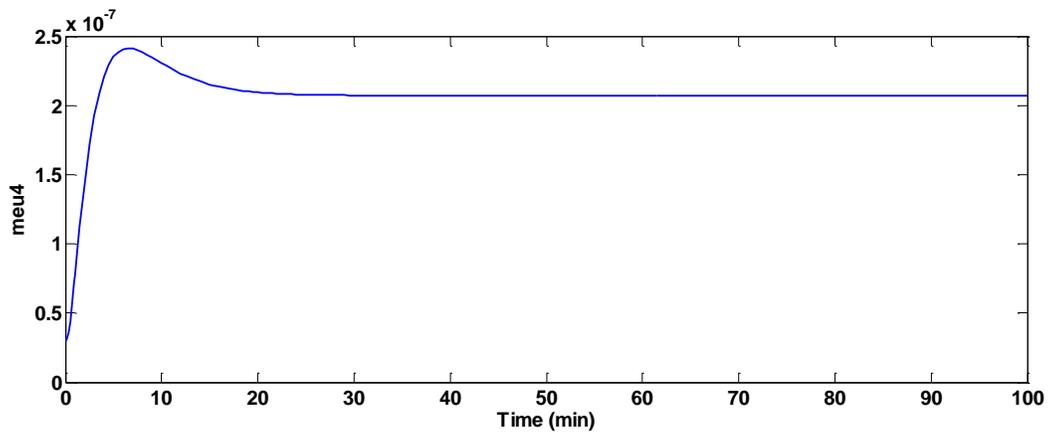


Figure4.5: Fourth order moment of crystallization

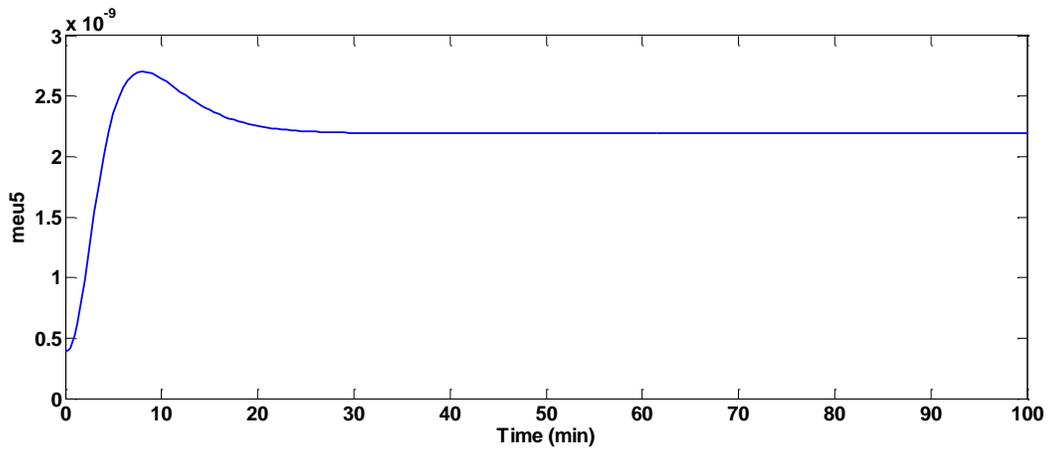


Figure4.6: Fifth order moment of crystallization

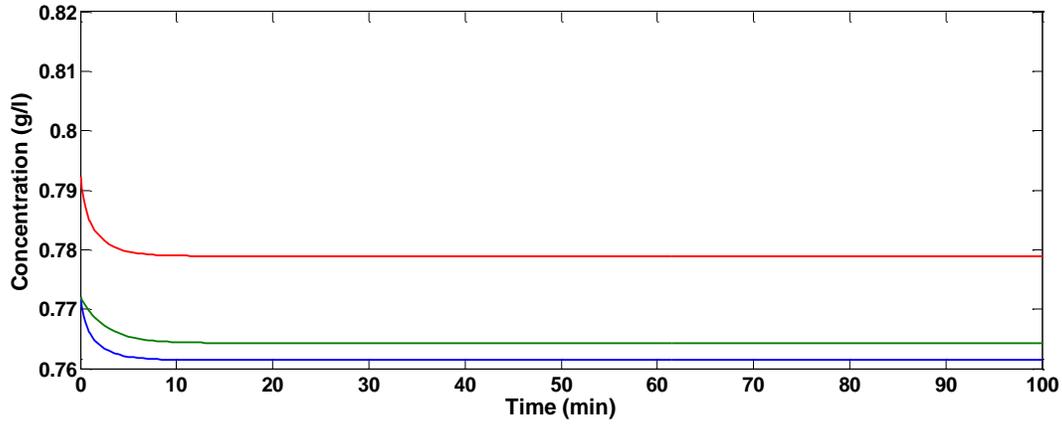


Figure4.7: Concentration profile. ($-C_s$ Saturation Limit, $-C_{mz}$ Metastable Limit)

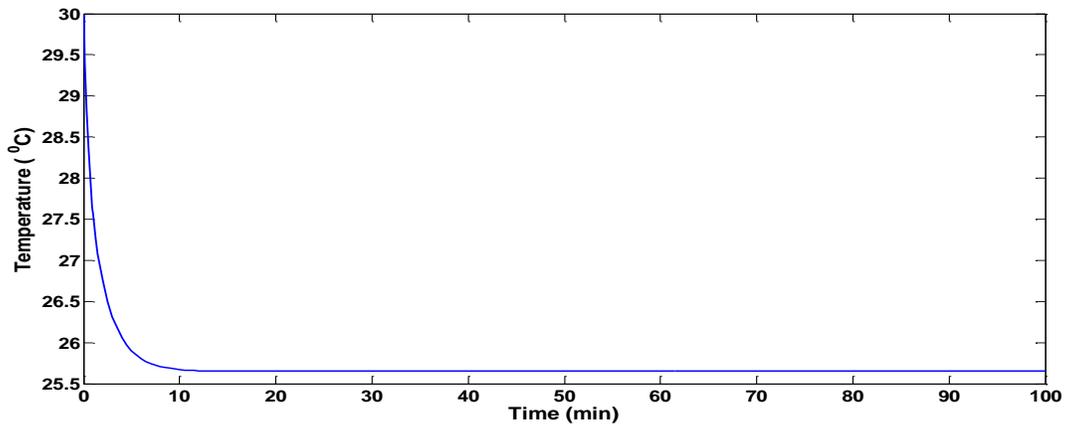


Figure4.8: Crystallizer Temperature Profile

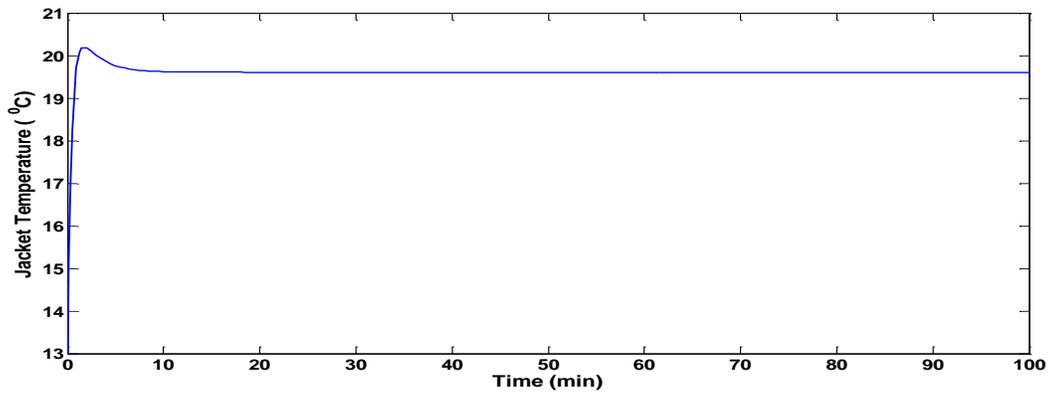


Figure4.9: Jacket Temperature profile

4.4. CLOSED LOOP RESPONSES OF CONTINUOUS CRYSTALLIZATION PROCESS

Crystallization is an important operation in industry owing to the extensive range of supplies that are produced in crystal form. Its widespread use is due to several advantages that

the process provides: a crystal formed from an impure solution is essentially pure, little energy is required for separation, and it provides the product in a suitable condition for packaging and storage. These profits are possible only when a good control strategy is tailored and all control intentions are accomplished. At industrial level, the objective of crystallizers is to obtain crystals with definite characteristics of crystal size and purity.

In this thesis, an attempt has been made to develop a good control strategy that improves the crystal size distribution of a non-isothermal continuous crystallizer that produces ammonium sulphate. Here, a traditional PID controller to control the temperature of the crystallizer has been proposed. Inlet temperature of the coolant is taken as the manipulated variable to control the crystallizer temperature. The parameters values for the PID controllers were obtained minimizing the integral square error (ISE). The PID parameters obtained for the temperature control are $K_c = 78.956$, $\tau_I = 0.7014$, $\tau_D = 0.3174$ and $\tau_F = 12.44$. The following graphs shows the closed loop responses obtained.

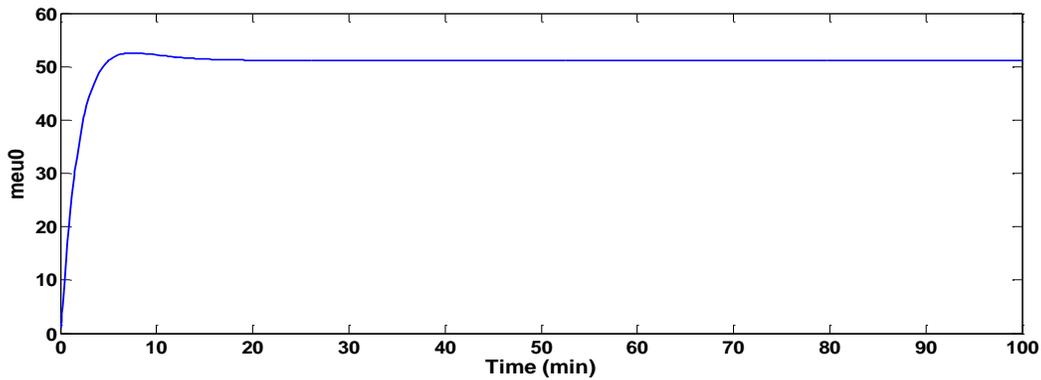


Figure 4.10: Zero order moment of crystallization

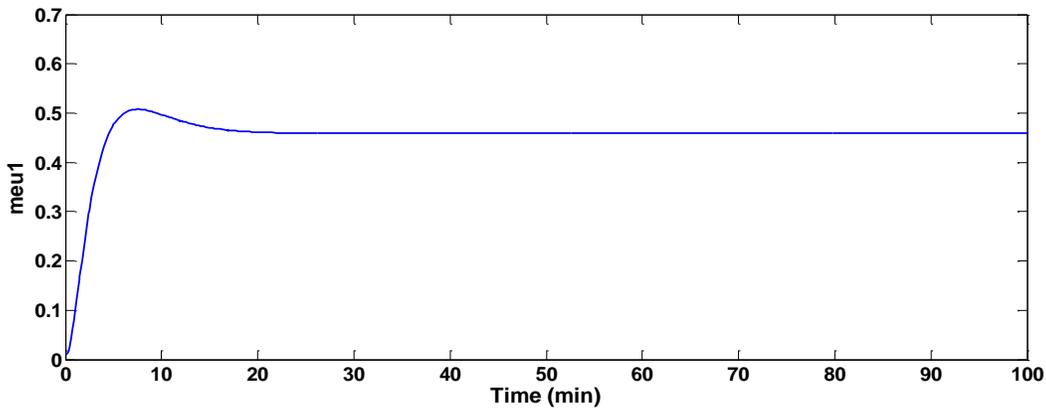


Figure 4.11: First order moment of crystallization

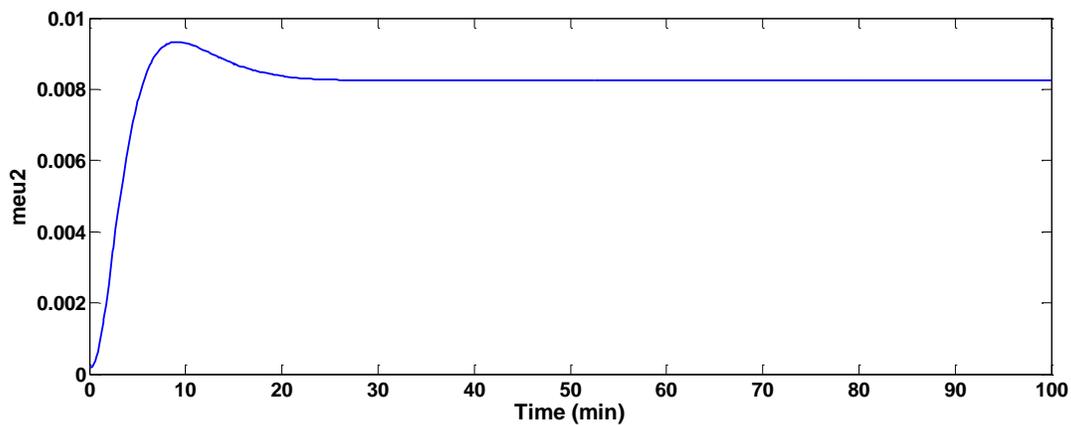


Figure 4.12: Second order moment of crystallization

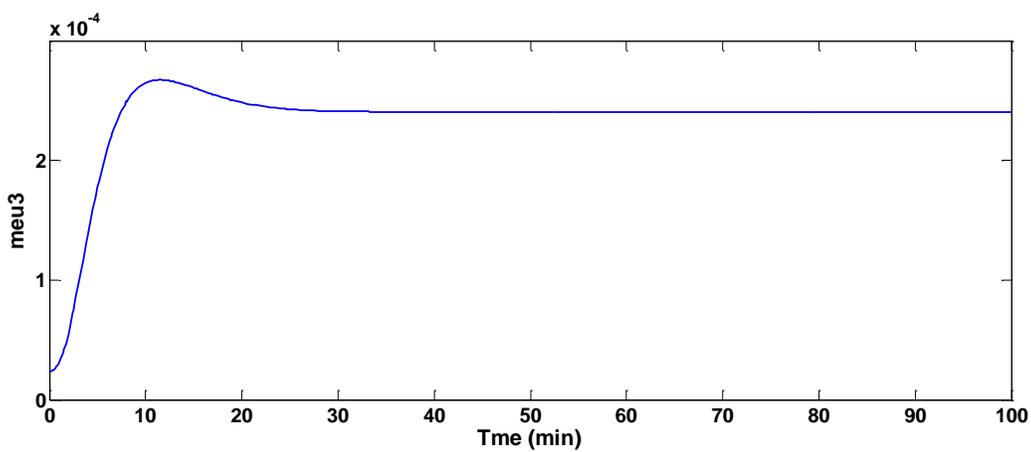


Figure 4.13: Third order moment of crystallization

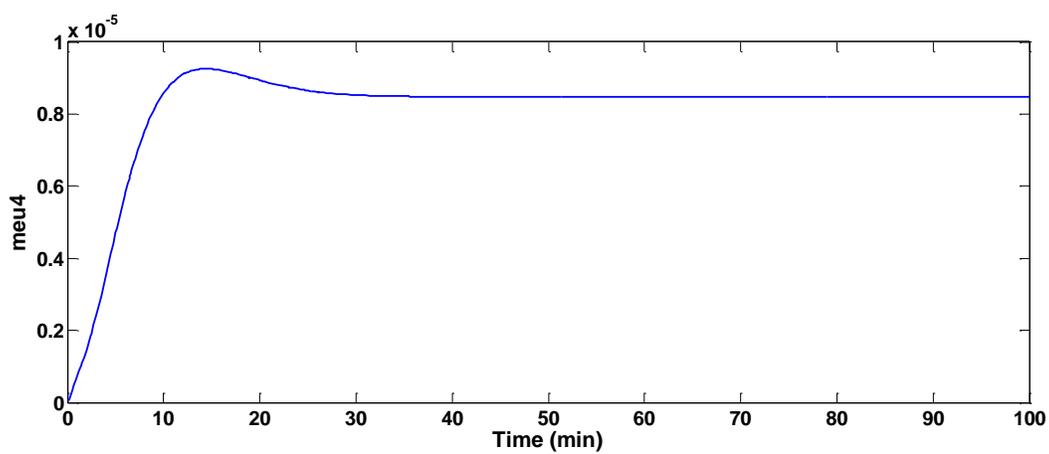


Figure 4.14: Fourth order moment of crystallization

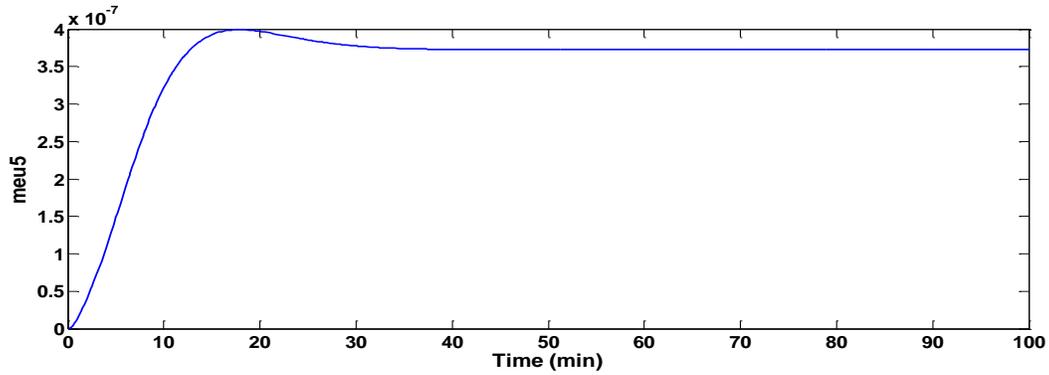


Figure 4.15: Fifth order moment of crystallization

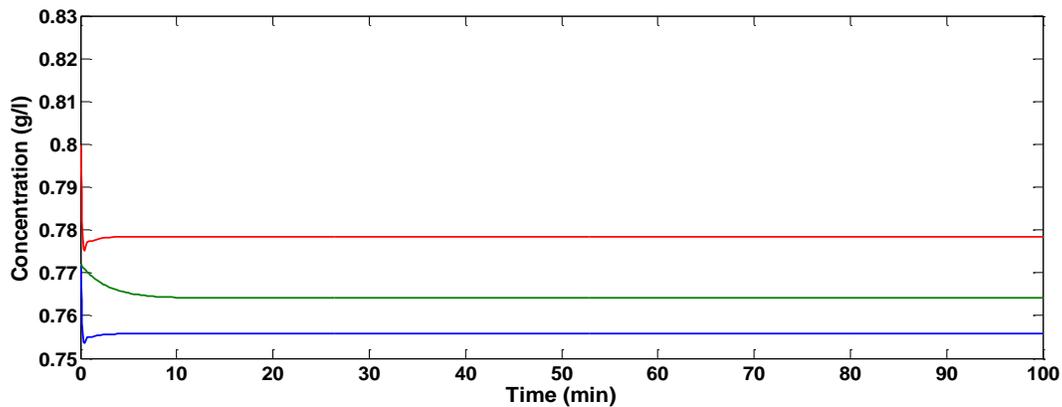


Figure 4.16: Concentration profile (— C_s Saturation Limit — C_{mz} Metastable Limit)

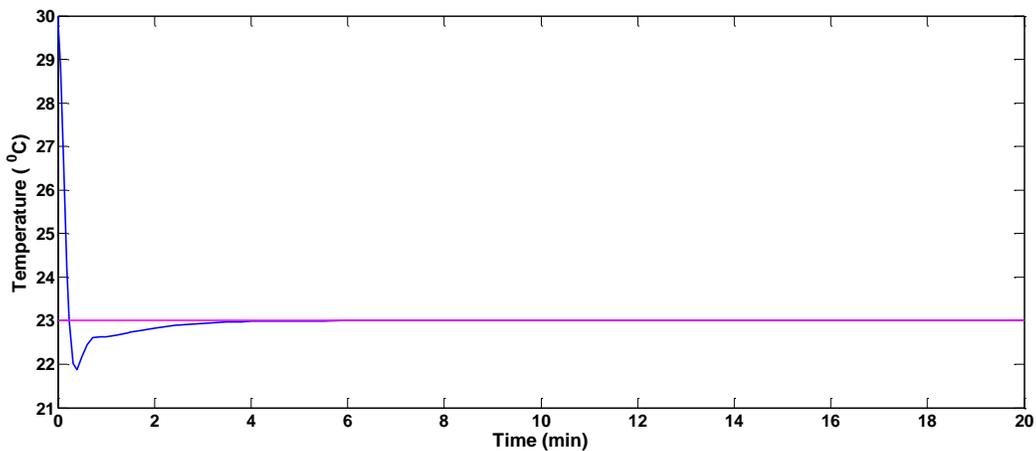


Figure 4.17: Temperature profile (—set point tracking)

From the figure 4.17 we can see that the PID controller enables perfect set point tracking and From Figure 4.17 and 4.16 we can see that open as well as in closed loop

crystallizer concentration is maintained well within the metastable zone limit. But in case of open loop simulation the crystallizer concentration is very close to saturation concentration limit and by temperature control we are able to maintain the concentration within acceptable limits.

4.5. CRYSTAL SIZE DISTRIBUTION

The mean crystal size and coefficient of variation can be calculated using the following relationships of the moments in order to obtain the complete Crystal size distribution:

$$L_{43} = \frac{\mu_4}{\mu_3} \quad (4.14)$$

$$C V = \sqrt{\frac{\mu_5}{\mu_3} - \left(\frac{\mu_4}{\mu_3}\right)^2} \times \frac{\mu_3}{\mu_4} \quad (4.15)$$

The crystal size distribution was calculated using the equation 4.14 and 4.15 for both the open loop and closed loop and are tabulated. The results show significant improvement in the CSD for closed loop.

Table 4.2: Crystal Size Distribution

OPERATING POLICY	MEAN CRYSTAL SIZE(μm)	CV (%)
Open loop	82.67	57
Controlled	354.2	46

4.6. CONCLUSION

The mathematical modeling and simulation of continuous crystallizer was performed to obtain the dynamics. Then a PID controller to control the crystallizer temperature which in turn leads to the improvement in crystal size distribution has been proposed. The overall effect is that the crystal concentration is maintained well within the metastable zone and there is significant improvement in the crystal mean size and coefficient of variation of the whole distribution.

REFERENCES

1. Alberto P, Raul O, "Nonlinear MIMO Control of a Continuous Cooling Crystallizer". *Modelling and Simulation in Engineering*, 11, 1-12,2012.
2. Heidari, Amir, Shirvani, Mansour, "Frequency Domain Model Simplification of Cumulative Mass Fraction in CMSMPR Crystallizer", *Iran. J. Chem. Chem. Eng.*,31, 1-12, 2012.
3. Lakatos B G, Blickle T, "Nonlinear Dynamics of Isothermal CMSMPR Crystallizer: A Simulation Study", *Comput. Chem. Eng.*, 19,501-504, 1995.
4. Moldovanyi N, Lakatos B G, Szeifert F, "Model Predictive Control of MSMPR Crystallizers", *J. Cryst. Growth*, 275, 1347-1349 (2005).
5. Motz S, Mitrovi A, Gilles E D, Vollmer U, "Modeling, Simulation and Stabilizing H^∞ Control of an Oscillating Continuous Crystallizer with Fines Dissolution", *Chem. Eng. Sci.*, 58, 3473-3490 ,2003.
6. Ramanathan S, "Control of Quasi Rational Distributed Systems with Examples on Control of Cumulative Mass Fraction of a Particle Size Distribution" Ph.D. Thesis, Michigan University,131-153,1988.
7. Randolph A D, Beckman J R , Krajevich Z I , *Crystal Size Distribution Dynamics in a Classified Crystallizer Part I. Experimental and Theoretical Study of Cycling in a Potassium Chloride Crystallizer*, *AIChE J.*, 23, 500-505 ,1977.
8. Rohani S, Haeri H C," Modeling and control of a continuous crystallization process Part 1. Linear and non-linear modeling", *Wood Computers and Chemical Engineering*,23, 263-277,1999.
9. Rohani S, Haeri H C," Modeling and control of a continuous crystallization process Part 2. Model predictive control ", *Wood Computers and Chemical Engineering* ,23, 279-286,1999.
10. Shirvani M, Inagaki M, Shimizu T, *Simplification Study on Dynamic Models of Distributed Parameter Systems*, *AIChE J.*, 41, 2652-2658, 1995.
11. Shirvani M, Doustary M A, Shahbaz M, Eksiri Z, "Heuristic Process Model Simplification in Frequency Response Domain", *Int. J. Eng. Transaction B*, 17, 18- 19, 2004.

Chapter 5

MULTIVARIATE STATISTICAL PROCESS MONITORING AND CONTROL

Chapter 5

MULTIVARIATE STATISTICAL PROCESS MONITORING AND CONTROL

5.1. INTRODUCTION

Multivariate Statistical Process Monitoring (MSPM) - also known as Multivariate Statistical Process Control or MSPC - is a valuable tool for warranting reliable product quality and safety in the process industry. With the advent of mature data capture technology, multisensory array, advancement in data collection, compression and storage, data driven process monitoring including product quality monitoring, fault detection and diagnosis are gaining due attention and wide spread acceptance. However, many industries are still lagging behind in fully using their potential to make noteworthy improvements in their production.

The MSPC make use of several multivariate models to instantaneously capture the data from process variables. It can vary from two variables to thousands of variables. This method provides major profits to the process engineers such as:

- Early fault detection
- Improved process understanding
- Process optimization
- Online prediction of quality

Various MSPC approaches, enables monitoring of data at the final product quality stage, as well as at different periods of the process, to recognize the fundamental systematic changes in the process variables. The measured variables in a process are often correlated to a certain extent, for example, by measuring several temperatures in a crystallizer events in a process can be pictured in a reduced subspace and this can give a direct chemical/physical interpretation. To keep such a process under control, traditional univariate control charts are not efficient due to the covariance between variables. This is because in univariate analysis the relationship between the response variable are pictured one at a time and thus does not disclose the multivariate patterns between these variables, which is vital for industrial processes both for interpretation and prediction.

5.1.1. DATA BASED MODELS

Data based modeling is one of the very recently added techniques in process identification, monitoring and control. Here the model parameters are determined by experimental results /wet labs, hence these models are called data based models or experimental models. Unlike the white box models derived from first principles, the black box/data based models or empirical models do not describe the systematic phenomena of the process; they are based on input-output data and only describing the overall behavior of the process. The data based models are especially appropriate for problems that are data rich and information poor. Sufficient numbers of quality data points are essential to propose a good model. Quality data is defined by noise free data; free of outliers and is ensured by data pre conditioning.

The phases in the Data based modeling are:

- System analysis
- Data collection
- Data conditioning
- Key variable analysis
- Model structure design
- Model identification
- Model evaluation

In this era of data explosion, rational as well as potential conclusions can be drawn from the data, and in this regard, multivariate statistics is of profound importance. Several Data driven multivariate statistical techniques such as Principal Component Analysis (PCA), Canonical Correlation Analysis (CCA), Partial Least Squares (PLS), Principal component regression (PCR) and Canonical Variate and Subspace State Space modeling have been proposed for these purposes.

Data based models are divided into two major categories namely:

- Unsupervised models: These are the models which try to extract the different features present in the data without any prior knowledge of the arrays present in the data. Examples are Principal Component Analysis (PCA), Hierarchical Clustering Techniques (Dendrograms), non-hierarchical Clustering Techniques (K-means).

- Supervised models: These are the models which try to learn the features of the data under the guidance of a supervisor who trains these models with inputs along with their corresponding outputs. Examples include Artificial Neural Networks (ANN), Partial Least Squares (PLS) and Auto Regression Models etc.

Efficient data mining, hence, efficient data based modeling enables to exploit the huge database available; in newer dimensions and perspective; embraced with never expected possibilities.

5.2. THEORETICAL POSTULATES

5.2.1. Principal Component Analysis (PCA)

Principal component analysis is a statistical procedure in which a set of measurements which are possibly the correlated variables are converted into a set of values of linearly uncorrelated variables using orthogonal transformation. The set of linearly uncorrelated variables produces are called principal components and first PC gives the direction of largest variation in a data set. PCA technique can be applied on an auto-scaled data matrix to determine the eigenvectors, related eigenvalues and scores or the transformed data along the principal components. The drawbacks of the PCA method are that the new latent variables produced have no physical meaning and this can result in loss of information.

Generally, PCA is a mathematical method used to find correlations and explain variance in a data set. The aim is to map a raw data vector A onto vectors B , where, the vector x represent a linear combination of a set of m orthonormal vectors u_i

$$x = \sum_{i=1}^m z_i u_i \tag{5.1}$$

where the coefficients z_i is given by $z_i = u_i^T x$. This corresponds to the rotation of the coordinate system from the original x to a new set of coordinates given by z . In order to reduce the

dimensions of the data set, only a subset ($k < m$) of the basic feature vectors are conserved. Rest of the coefficients are substituted by constants and each vector x is then given as

$$x' = \sum_{i=1}^m z_i u_i + \sum_{i=1}^d b_i u_i \quad (5.2)$$

Here, u_i are the principal components which are the eigenvectors of the covariance matrix of the data set. The coefficients b_i and the principal components are so chosen that they provide best approximation of the original vector on an average. However, the reduction of dimensionality from m to k can lead to approximation error. By selecting the vectors that correspond to the largest Eigen values of the covariance matrix we can reduce the sum of squares of the errors over the whole data. PCA transformation reduces dimensionality of the original data set (typically 2-3) and the measurements can be plotted in the same coordinate system. These plots show the relation between different observations or experiments. Grouping of data points in those bi-plots suggest some common properties and those can be used for classification. Considering the following matrix:

$$X = \begin{bmatrix} x_{11} & \dots & x_{1m} \\ \vdots & \dots & \vdots \\ x_{n1} & \dots & x_{nm} \end{bmatrix} \quad (5.3)$$

where, each row in X represents one measurement and the number of columns m gives the length of the measurement sequence. After that the covariance matrix $C = \text{cov}(X)$ and its eigenvalues λ was calculated. Its eigenvectors u_i form an orthonormal basis

$U = [u_1, u_2, u_3, \dots, u_m]$; that is $U^T U = 1$. The original data set can be denoted in the new basis using the relation: $Z = U^T X$. After this renovation, a new data matrix of reduced dimension can be erected with the help of eigenvalues of the matrix C . This is done by choosing the highest values λ since they correspond to the principal components with maximum significance. The no. of PCs to be contained within should be high enough to guarantee good separation between the groups. Principal components with low values of λ are ignored. Let the first k PCs as new features be selected disregarding the remaining $(m \times k)$ principal components. In this way, a new data matrix D of dimension $n \times k$ was obtained.

$$D = \begin{bmatrix} z_{11} & \dots & z_{1k} \\ \vdots & \dots & \vdots \\ z_{n1} & \dots & z_{nk} \end{bmatrix} \quad (5.4)$$

The PCA score data sets are grouped into number of classes following the rule of nearest neighborhood clustering algorithm. In the present work PCA based similarity was used for process monitoring purposes.

5.2.2. Similarity Factors:

5.2.2.1 PCA similarity:

Principal component analysis is a multivariate statistical tool that measures the variations in data, and renovates the original set of correlated variables into a new set of uncorrelated variables. The new set of variables thus produced is a linear combination of the original set of variables and they are called as principal components which signify the most vital directions of variations in a dataset.

[Krzanowski \(1979\)](#) developed a method of using similarity factor in the place of Euclidean distance to measure dissimilarity between two datasets. In this method the similarity between two data sets containing n variables are measured using PCA similarity factor, S_{PCA} . The PCA model for each dataset is assumed to contain k principal components and the number of principal component are so chosen that they define at least ninety five percentage variance in each dataset. These principal components are the eigenvectors of the covariance matrix of the multivariate dataset. The dissimilarity between the two data sets is computed by comparing their PCs and is given by a single factor, S_{PCA} .

Two data sets X_1 and X_2 having n variables each are considered. It is assumed that the PCA contain k principle components each. Then the corresponding $(n \times k)$ subspaces are denoted by L and M respectively. The S_{PCA} that compares these datasets is given as:

$$S_{PCA} = \frac{\text{trace}(L^T M M^T L)}{k} \quad (5.5)$$

Geometrically, S_{PCA} is the sum of the squares of the cosines of the angles between each principal component L and M . Thus,

$$S_{PCA} = \frac{1}{k} \sum_{i=1}^k \sum_{j=1}^k \cos^2 \theta_{ij} \quad (5.6)$$

θ_{ij} is the angle between the i^{th} PC of X_1 and the j^{th} PC of X_2 .

But, here all the PCs are weighed equally and thus might not be efficient in capturing the degree of dissimilarity between the datasets when only one/two PCs describe the variance. Thus, it calls for a modified PCA similarity factor; S_{PCA}^λ that weights each PC by its explained variance. The S_{PCA}^λ is defined as:

$$S_{PCA}^\lambda = \frac{\sum_{i=1}^k \sum_{j=1}^k (\lambda_i^{(1)} \lambda_j^{(2)}) \cos^2 \theta_{ij}}{\sum_{i=1}^k \lambda_i^{(1)} \lambda_j^{(2)}} \quad (5.7)$$

where $\lambda_i^{(1)}$ and $\lambda_j^{(2)}$ are the i^{th} eigenvalues of the first and second datasets respectively.

5.2.2.1. Distance similarity:

The distance similarity compares 2 datasets that having same kind of spatial alignment but are positioned at distance from each other. This similarity factor is used when two datasets are having similar PCs but the values of the process variables are dissimilar due to different operating conditions. We can define the distance similarity factor, S_{dist} as:

$$S_{\text{dist}} \triangleq 2 \times \frac{1}{\sqrt{2\pi}} \int_{\phi}^{\infty} e^{-\frac{z^2}{2}} dz = 2 \times \left[1 - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\phi} e^{-\frac{z^2}{2}} dz \right] \quad (5.8)$$

where, ϕ is the Mahalanobis distance from the midpoint of the first dataset (X_1) to the center the second dataset (X_2) and is defined as:

$$\phi = \sqrt{X'_1 - X'_2 \sum_1^{*-1} (X'_1 - X'_2)^T} \quad (5.9)$$

X'_1 and X'_2 are sample mean vectors, \sum_1 is the covariance matrix of dataset X_1 , and \sum_1^{*-1} is the pseudo-inverse of X_1 calculated using SVD (singular value decomposition). Here X_1 is taken as the reference dataset. In equation 5.8, one sided Gaussian distribution is used as $\phi \geq 0$

.The error function in equation 5.8 is calculated using standard tables or softwares. Integration normalizes S_{dist} between 0 and 1. As the relative values of S_{dist} is used for clustering as well as pattern matching, any mapping from ϕ to S_{dist} can be used.

5.2.2.3. Combination of similarity factors

A combined similarity factor which is a combination of S_{PCA}^λ and S_{dist} can be used as a similarity measure. The key issue in this case is how these similarities should be united to yield a single parameter of similarity. Here the similarity factors are jointed into a sole similarity SF by means of weighted average of the two similarities and are given as:

$$\text{SF} \triangleq \alpha_1 S_{\text{PCA}}^\lambda + \alpha_2 S_{\text{dist}} \quad (\alpha_1 + \alpha_2 = 1) \quad (5.10)$$

The value of weighting factors can be chosen based on the prominence of each similarity factor for a particular application. Here, we selected the values α_1 of and α_2 as 0.67 and 0.33 respectively. It is very important to note that there is no relation between the number of readings in each dataset and similarity factors and can be evaluated with comparatively less computational exertion.

5.2.2.4. Dissimilarity Index

In addition to the above mentioned similarity measures for clustering and pattern matching another factor called the dissimilarity index is also used in this work. [Kano et al.](#) proposed a dissimilarity index to match datasets for process monitoring. Here eigenvalue or the eigenvector decomposition of the covariance matrix of the combined data set is used to calculate the dissimilarity index. The covariance matrix of the combined data set X, is defined as:

$$X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \quad (5.11)$$

The datasets X_1 and X_2 are then projected on the eigenvector matrix of the covariance matrix of X and scaled by the corresponding eigenvalues to give the transformed data sets X_1' and X_2' respectively. Then, the covariance matrices of the transformed data sets are calculated. Finally, the eigenvalues and eigenvectors of the covariance matrices are calculated. The covariance matrices of both transformed data sets have similar eigenvectors, and their eigenvalues are associated as:

$$\lambda_1^{1'} = 1 - \lambda_2^{2'} \quad (5.12)$$

where $\lambda_1^{1'}$ and $\lambda_2^{2'}$ are the eigenvalues of the covariance matrices of the modified data. If the data sets are similar to each other their eigenvalues will be near to 0.5, but if the data sets are dissimilar, then the largest and the smallest eigenvalues will be near to 0 and 1 respectively. Finally, a dissimilarity index D, defined to calculate the dissimilarity between two datasets.

$$D = \frac{4}{P} \sum_{j=1}^P (\lambda_j - 0.5)^2 \quad (5.13)$$

The dissimilarity index D varies from zero to one. If two data sets are similar in any form D will be close to zero. However, D will be close to one when data sets are quite different from each other.

In view of this, two MSPC methods clustering time series data and moving window based pattern recognition technique was adapted for process monitoring. Continuous crystallization processes were taken for the implementation of the proposed monitoring techniques. Instead of using first hand plant data, the crystallization processes was modeled using first principles and processes were perturbed at industrially applicable operating conditions including faulty conditions to create vector time series datasets.

5.3. CLUSTERING TIME SERIES DATA

Cluster analysis is a process of finding groups among the data. The advent of various data recording systems has enabled the collection of large amount of valuable data about the normal as well as abnormal plant behavior. Classification of these datasets into different groups of operating condition can serve to be highly beneficial for decision making in fault identification and analyses. Clustering is a process of finding groups among datasets that are similar to each other. These clusters are further evaluated to gain process knowledge from the similarity among the datasets. The knowledge gained from the clustering can be very important for purposes like process enhancement, fault detection, where each new dataset is grouped as either a known condition or a new condition. In this thesis clustering of time series data based on similarity and dissimilarity index have been studied.

Though clustering is a very important research area for process monitoring and control, only very few researchers have reported the clustering of dynamic and high dimensional multivariate time series data based on similarity measures. [Johnston](#) and [Kramer](#) used a

probabilistic approach together with expectation maximization algorithm for the clustering of time series data. But this approach is not suited for batch processes as in such cases process dynamics have the difference between different operating conditions. [Huang et al.](#) performed clustering by dividing huge clusters of multivariate data sets into smaller ones using PCA models. [Wang](#) and [McGreavy](#) successfully classified different operating conditions of a simulated fluid catalytic cracking unit using the Auto class algorithm. The drawback of this approach is that it calls for each dataset to possess same number of readings; or else different datasets may contain different no. of features. [Wang](#) and [McGreavy's](#) used another method where Euclidean distance between datasets was used to measure similarity.

[Wang](#) and [Li](#) described conceptual clustering for crafting state space based monitoring of a process. [Kavitha](#) and [Punithavalli \(2010\)](#) claimed that all traditional clustering or unsupervised algorithms are incongruous for real time data. [Xiaozhe Wang et al. \(1998\)](#) proposed clustering methodologies based on extraction of structure based statistical characters from a multivariate time series data. [Sudjianto et al.](#) and [Trouve et al. \(1996 & 2000\)](#) used PCA to extract similarity from the large datasets. Such feature extracting serves as dimensionality reduction of the datasets. [Singhal and Seborg \(2002\)](#) used PCA and Mahalanobis distance similarity measures for location of similar operating condition in large multivariate database. They proposed modified k-means cluster algorithm using similarity factors. In many instances, it was proved that use of PCA for grouping of objects that possess similar characteristics in multivariate time series data would be beneficial.

Intention of this work is to locate the abnormality of the process by distinguishing between the states of operating conditions involved in large database. Clustering of the different operating conditions including the faulty operating conditions of a simulated continuous cooling crystallizer is presented here.

5.3.1. MODIFIED K MEAN CLUSTERING ALGORITHM

5.3.1.1. Using Similarity Factors

Clustering technique is simpler of grouping structures as no priori assumptions are made. Classification of the data are made base on similarities or distances (dissimilarities). The number

of clusters K can be prespecified or can be determined iteratively by the clustering procedure. In general, the K -means clustering proceeds in three steps, which are as follows:

1. Data sets are clustered in K clusters.
2. A dataset is assigned to a cluster having nearest centroid. Then these centroids are recalculated for cluster on reception of a new item and for the cluster which loses the same item
3. The step 2 is repeated until no more reassignment of datasets takes place or until stable cluster are generated.

The K -means clustering ensures a faster computation as it does not require the distance matrix as required in hierarchical clustering. The time series data belonging to various operating conditions were discriminated and classified using modified K -means algorithm.

- The K mean clustering algorithm is given as:

We have Q datasets, $\{X_1, X_2, \dots, X_q, \dots, X_Q\}$ to be clustered into K clusters.

Step1: The j^{th} dataset in the i^{th} cluster is denoted by $X_j^{(i)}$. The aggregate dataset $\mathcal{X}_i (i=1,2,\dots,K)$ for each of the K clusters is computed as:

$$\mathcal{X}_i = \left[\left(X_1^{(i)} \right)^T \dots \left(X_j^{(i)} \right)^T \dots \left(X_{Q_i}^{(i)} \right)^T \right]^T \quad (5.14)$$

where, Q_i is the number of datasets \mathcal{X}_i and $\sum_{i=1}^K Q_i = Q$

Step 2: The dissimilarity value between the datasets $X_q (q=1,2,\dots,Q)$ and each of the K aggregate datasets $\mathcal{X}_i (i=1, 2,\dots,K)$ are calculated as:

$$d_{iq} = 1 - SF_{i,q} \quad (5.15)$$

where, $SF_{i,q}$ is the similarity factor between the q^{th} dataset and the i^{th} cluster (equations 5.10).

The aggregate dataset \mathcal{X}_i (equation 5.14) is taken as the reference. The dataset X_q is allocated to the cluster to which it is least dissimilar, that is to the cluster having smallest value of d_{iq} .

Step 3: Repeated the aforesaid steps for all Q datasets.

5.3.1.1. Using Dissimilarity Index

In order to cluster the time series data based on the dissimilarity index replace equation 5.15 of the algorithm by 5.13

5.3.2. CLUSTER PERFORMANCE EVALUATION

To gauge the performance of varied clustering methodologies two parameters specifically cluster purity, p , and clustering efficiency, η has been outlined. If datasets comprises of N_{op} operating conditions and N_{DBj} datasets of operating condition j within the database and the datasets have been divided into K clusters, then a cluster purity p , is outlined to validate the purity of every cluster in terms of what number of datasets of a selected operating condition are present in that cluster. The cluster purity for the i^{th} cluster is defined as:

$$p_i \triangleq \frac{(\max_j N_{i,j})}{N_{P_i}} \times 100\% \quad (5.16)$$

Here, $N_{i,j}$ is the number of datasets of operating condition j within the i^{th} cluster, and N_{P_i} is the number of datasets in the i^{th} cluster. The dominant operating condition of cluster is decided by viewing the number of times a particular operating condition is repeated in a cluster.

The second parameter is the cluster efficiency, η , which measures the extent of distribution of datasets pertaining to particular operating condition into different clusters. If the cluster efficiency is high then all data sets pertaining to a specific operating condition will be classified into one cluster. Thus, this measure avoids large K values during clustering. The clustering efficiency for the j^{th} operating condition is defined as:

$$\eta_j \triangleq \frac{(\max_i N_{i,j})}{N_{DBj}} \times 100\% \quad (5.17)$$

The p and η provide good balance between cluster purity and the concentration of operating conditions into separate clusters.

5.4. PATTERN MATCHING

With the advent of modern data recording systems it is possible collect and store huge amount of data in industries. Industrial plants collect huge number of process variables, production, and product quality and maintenance information on continuous basis. Thus a huge amount of data is available for analyzing the process and earlier existences of abnormal conditions. A historical record contains much valuable information pertaining to the process, but

extraction of useful data from them is a difficult task. It would be highly beneficial if these relevant datasets could be extracted from a huge historical database. Even though the terms like data processing and data discovery have gained much attention in variety of activities, conferences, and commercial products where analysis of large databases is required, comparatively few engineering applications are reported.

Pattern matching is a vital answer to the present downside. It is a method of locating historical record that is similar to a definite set of data. Though manual hunts of historical database is possible, it is very tedious and time intense. So, it is highly desired to formulate a machine driven pattern matching. This method can be used to locate abnormal plant conditions if a similar type of abnormal condition has arisen in the past. Thus this method can ease two important purposes like recognizing the source of the occurrence of abnormality and to develop an efficient solution that can inhibit or minimize future abnormal operation and their impact.

Matching techniques can be broadly classified into two: unsupervised and supervised techniques. Supervised methods necessitate the knowledge about training data for every type of pattern that is being taken. Many of the classical pattern matching methods are termed as supervised techniques. Some of the well-known supervised pattern recognition techniques are statistical approaches, neural networks and rule-based systems. This thesis is strictly restricted to unsupervised learning techniques, where previous occurrences of the present pattern will serve as training data. Some of the existing unsupervised patterns matching techniques are based on neural network methods like adaptive resonance theory (ART), self-organizing maps etc.

Pattern matching techniques has been successfully applied in areas of machine learning, image processing, and speech and character identification by [Fukunaga \(1990\)](#), [Bishop \(1995\)](#), [Shurmann\(1996\)](#). Process monitoring and data investigation based on pattern matching techniques from a chemical engineering viewpoint are also found in literature ([Stephanopoulos and Han, 1994](#), [Kramer and Mah, 1994](#),[Davis et al., 1999](#)).Pattern matching of simulated response data and experimental data. i.e. a model of the process is used to simulate faulty situations which are not found in the current database. [Basseville](#) and [Nikiforov \(1993\)](#) have reported various methods for sensing changes in the dataset using multivariable statistics and time-series models. But, efficiency of these methods reduces as the dimensionality of the historical datasets decreases.

A number of methods for univariate pattern recognition have been reported. These methods are based on the evaluation of Euclidean distances of modified data. For multivariate applications pattern classification or fault diagnosis are based on PCA. Thus, principle component analysis can also be used for fault detection and analysis.. Raich and Cuinar (1995) used the regular principal component metrics and the PCA similarity factor for differentiating between different forms of faults. Johannesmeyer et al. (2002) developed a pattern matching method that uses PCA similarity and limit violation similarity factor. Kano et al. (2001) proposed dissimilarity factors based on eigenvalue decompositions of the datasets for monitoring. Ashish Singhal et al. (2002) used Q and T^2 statistics based on PCA similarity for pattern matching with a case study of batch fermenter. He also proposed two different similarity measures and a moving-window method to depict the degree of dissimilarity between the current period of concern and windows of historical data of continuous CSTR. Damarla and Kundu (2011) proposed PCA and distance based combined similarity together with the moving window approach to discriminate between the normal operating conditions as well as fault detection of bioreactor and drum boiler process. Deng and Tian (2013) used statistics kernel PCA similarity factor for fault pattern recognition of nonlinear process. Li and Wen (2014) used pattern matching method for detecting faults in air handling unit system.

An unsupervised multivariate statistical technique based on a combined similarity factor and a dissimilarity index along with a moving-window based approach to differentiate between the normal operating conditions and faulty condition of a continuous crystallizer is proposed in this work.

5.4.1. MOVING WINDOW BASED PATTERN MATCHING

The basic aim of the pattern matching method is to identify periods of operation from a historical database whose operational conditions are identical to the target operating condition. In case of automated pattern matching methodology only minimal information from the user is required. i.e. the user only has to supply the relevant process variables and their corresponding measurements for a period of current examination data. This data is called as the snapshot data and serves as a stencil for searching the historical record.

In moving window based pattern matching, the significant historical data are subdivided into data windows having size as that of snapshot dataset. These historical datasets are structured

by placing data windows side by side along the time axis resulting in datasets which are having same length and are non-overlapping. The snapshot or template data moves in sample wise manner through the historical template and the similarity between them is explained by distance and PCA based combined similarity factor and dissimilarity index. The historical data windows with the highest similarity factors are taken and analyzed by the process Engineer.

5.5. MONITORING OF CONTINUOUS CRYSTALLIZER

A case study on simulated continuous cooling crystallizer was performed to evaluate the proposed clustering and pattern matching techniques. Dynamic simulation of continuous cooling crystallizer with cooling jacket dynamics was performed. Operating conditions that contain faults of variable degree and disturbances was simulated for the crystallizer. Three process variables were recorded. The details of the dynamic simulation and control study of continuous crystallizer are given in chapter 4.

The four operating conditions (Table 5.1) were generated including the faulty operating conditions (opt. condi. 2) as well. Simulation was performed for a period of 100 min. with a sampling time of 0.5 sec. for each variable. Thus 18 different datasets with 201 data points each for each of the 3 measured variables was produced.

In moving window based pattern matching proposed here, the historical data are divided into four data windows having same size as the snapshot dataset. These historical datasets are then structured by placing windows side by side. Four snapshot data sets have been considered here. The snapshot data are then moved through the historical template and the similarity between them is evaluated by distance and PCA based combined similarity factor and dissimilarity index. The historical data windows with the largest values of similarity factors are collected.

Table 5.1: Operating Conditions

OPERATING CONDITION	NO OF DATASETS
Step change in T_{jin}	3
Ramp input to inlet concentration (Fault)	6
Step input to inlet concentration	6
Step input to T_{in}	3

5.6. RESULTS & DISCUSSIONS

5.6.1. Clustering of time series data

The K-means clustering procedure was performed using two similarity factor combinations: $SF = 0.67 S_{PCA}^{\lambda} + 0.33 S_{dist}$ and $SF = 0.67 S_{PCA} + 0.33 S_{dist}$ and a dissimilarity index. The performance of these clustering methodologies are evaluated using the parameters cluster purity, p , and clustering efficiency, η . The combined similarity factor $SF = 0.67 S_{PCA}^{\lambda} + 0.33 S_{dist}$ has clustered four operating conditions into five clusters in which the second operating condition i.e. the fault has been clustered into two clusters all the other operating conditions are properly clustered resulting in 100% cluster purity, p , and 91.66% clustering efficiency, η . (Table 5.2)

Table 5.1: Clustering performance based on combined similarity factor based

Cluster no.	N_p	p(%)	Dominant Op. cond.	Operating condition:			
				1	2	3	4
1	3	100	1	3	0	0	0
2	4	100	2	0	4	0	0
3	2	100	2	0	2	0	0
4	6	100	3	0	0	6	0
5	3	100	4	0	0	0	3
Avg	NA	p=100	NA	$\eta=100$	$\eta=66.66$	$\eta=100$	$\eta=100$

$$SF = 0.67 S_{PCA}^{\lambda} + 0.33 S_{dist}$$

Table 5.2: Clustering performance based on combined similarity factor based

Cluster no.	N_p	p(%)	Dominant Op. cond.	Operating condition:			
				1	2	3	4
1	3	100	1	3	0	0	0
2	3	100	2	0	3	0	0
3	1	100	2	0	1	0	0
4	2	100	2	0	2	0	0
5	9	66.66	3	0	0	6	3
Avg	NA	p=93.32	NA	$\eta=100$	$\eta=50$	$\eta=100$	$\eta=100$

$$SF = 0.67 S_{PCA} + 0.33 S_{dist}$$

The combined similarity factor $SF = 0.67S_{PCA} + 0.33S_{dist}$ also clustered four operating conditions into five clusters where the second operating condition was clustered into three clusters and the 3rd and 4th operating conditions being clustered as one. This method showed 93.32% cluster purity, p , and 87.5% clustering efficiency, η (Table 5.3). The dissimilarity index gave better results than the above mentioned combined similarity factors. The dissimilarity index is capable of detecting faulty operating condition as cluster 2 and datasets pertaining to various other operating conditions were also identified correctly. Thus, resulting in 100% cluster purity p , and 100% clustering efficiency η (Table 5.4).

Table 5.3: Dissimilarity Index based clustering performance

Cluster no.	N_p	$p(\%)$	Dominant Op. cond.	Operating condition:			
				1	2	3	4
1	3	100	1	3	0	0	0
2	6	100	2	0	6	0	0
3	6	100	3	0	0	6	0
4	3	100	4	0	0	0	3
Avg	NA	$p=100$	NA	$\eta=100$	$\eta=100$	$\eta=100$	$\eta=100$

A summary of clustering performance obtained using different methodologies are presented in Table 5.5. The results shows that dissimilarity index provides the best clustering performance followed by the combined similarity factor $SF = 0.67 S_{PCA}^\lambda + 0.33S_{dist}$.

Table 5.5: Clustering performance using different methodologies.

Clustering method	$p_{avg.}(\%)$	$\eta_{avg.}(\%)$
Dissimilarity Index	100	100
$0.67 S_{PCA}^\lambda + 0.33 S_{dist}$	100	91.66
$0.67 S_{PCA} + 0.33 S_{dist}$	93.32	87.5
S_{PCA}^λ	88	100
S_{PCA}	94.44	83.33
S_{dist}	NA	NA

5.6.2. Moving window based pattern matching

The proposed moving window based pattern matching technique perfectly identified all the snapshot data including normal and faulty data which find their similarity with the same process data present in the historical database. The technique was tested in sample wise as well as dataset wise manner. The similarity factors obtained in dataset wise pattern matching using combined similarity factors $0.67 S_{PCA}^{\lambda} + 0.33 S_{dist}$, $0.67 S_{PCA} + 0.33 S_{dist}$ and dissimilarity index are shown in Table 5.5, 5.6 and 5.7 respectively.

Table 5.6: Pattern matching using combined similarity factor

	Opt. 1	Opt. 2	Opt. 3	Opt. 4
Opt. 1	1	0.3950	0.9441	0.7084
Opt. 2	0.3950	1	0.5745	0.6871
Opt. 3	0.9441	0.5745	1	0.9952
Opt. 4	0.7084	0.6871	0.9952	1

$$0.67 S_{PCA}^{\lambda} + 0.33 S_{dist}$$

Table 5.7: Pattern matching using combined similarity factor

	Opt. 1	Opt. 2	Opt. 3	Opt. 4
Opt. 1	1	0.6707	0.7845	0.7805
Opt. 2	0.6707	1	0.6697	0.6799
Opt. 3	0.7845	0.6697	1	0.9970
Opt. 4	0.7804	0.6799	0.9970	1

$$0.67 S_{PCA} + 0.33 S_{dist}$$

Table 5.8: Pattern matching using dissimilarity index

	Opt. 1	Opt. 2	Opt. 3	Opt. 4
Opt. 1	3.466×10^{-26}	0.6039	0.4372	0.474
Opt. 2	0.6039	3.1061×10^{-30}	0.1818	0.3230
Opt. 3	0.4372	0.1818	3.8621×10^{-31}	0.0974
Opt. 4	0.4746	0.3230	0.0974	1.4626×10^{-30}

In the pattern matching based on similarity datasets sets which are similar have similarity factor as 1. whereas in pattern matching based on dissimilarity index if two data sets are similar to each other D value will be close to zero. Thus we can see from the above tables that all the snapshot data sets have been correctly identified.

5.7. CONCLUSION

With the proposed clustering technique, the time series data corresponding to various operating conditions including abnormal ones of the continuous cooling crystallization processes were clustered/classified efficiently using a similarity factor and dissimilarity index based modified K-means clustering algorithm. A Sample wise moving window based pattern matching technique developed for the process monitoring of the crystallizer was also implemented successfully. Efficient modeling and simulation of the crystallization process taken up was a key factor behind the generation of databases required for the successful implementation of the proposed monitoring techniques. The machine learning algorithms developed here with their encouraging performances deserves immense significance in the current perspective of process monitoring and control.

REFERENCES

1. Basseville M., "On-Board Component Fault Detection and Isolation Using the Statistical Local Approach", *Automatica*, 34, 1391-1415, 1998,
2. Bishop, C. M., "Neural Networks for Pattern Recognition", Clarendon Press: Oxford, U.K., 1995.
3. Damarla. S. K and Kundu .M, "Monitoring of Bioreactor using Statistical Techniques", *Res.J.Chem.Sci.*, 3, 114-117, 2011.
4. Damarla.S .K and Kundu .M, "Monitoring of Drum-boiler Process Using Statistical Techniques", *International Journal of Chemical Engineering and Applications*, 2, 172-180, 2011.
5. Davis, J. F., Piovoso M. J., Kosanovich K. A., Bakshi B. R., "Process Data Analysis and Interpretation". In *Advances in Chemical Engineering*, 25, 1-199, 1999.
Deng X, Tian X, "Nonlinear process fault pattern recognition using statistics kernel PCA similarity factor", *Neurocomputing*, 121, 298–308, 2013.
6. Fukunaga, K., "Introduction to Statistical Pattern Recognition", Academic Press: New York, 1990.
7. Huang Y, McAvoy TJ, Gertler J., " Fault isolation in nonlinear systems with structured partial principal component analysis and clustering analysis", *Can. J. Chem. Engr.*, 78, 569–577, 2000.
8. Johannesmeyer M. C., Singhal A, "Pattern Matching in Historical Data", *AIChE Journal*, 48, 2022-2038, 2002.
9. Johnston LPM, Kramer MA., "Estimating state probability functions from noisy and corrupted data", *AIChE J.*, 44, 591–602, 1998.
10. Kano, M., Hasebe, S., Hashimoto, I., Ohno, H., "A New Multivariate Statistical Process Monitoring Method Using Principal Component Analysis"., *Comput. Chem. Engr.*, 25, 1103-1113, 2001.
11. Kavitha V. and Punithavalli M., " Clustering time series data stream-A literature survey", *IJCSIS*, 8, 289-294, 2010.
12. Li S, Wen J, "Application of pattern matching method for detecting faults in air handling unit system", *Automation in Construction*, 43, 49–58, 2014.

13. Lin YT, Chen Y K, Kung SY., " A principal component clustering approach to object-oriented motion segmentation and estimation", J. VLSI Signal Proc. Syst. for Signal, Image and Video Tech.,17,163–187,1997.
14. Owsley L, Atlas L, Bernard G. "Automatic clustering of vector time-series for manufacturing machine monitoring", In Proceeding IEEE International Conference on Acoustics, Speech, and Signal Processing, Munich, Germany, 3393–3396,1997
15. Raich. A, Cuinar, "A Statistical Process Monitoring and Disturbance Diagnosis in Multivariate Continuous Processes", AIChE J., 42, 995-1009,1996.
16. Rosen C, "Supervisory control of wastewater treatment plants by combining principal component analysis and fuzzy c-means clustering", Water Sci. and Tech.
17. Singhal. A, Seborg, D. E. "Matching Patterns from Historical Data Using PCA and Distance Similarity Factors", Proceedings of the 2001 American Control Conference; IEEE: Piscataway, NJ, 1759-1764, 2001.
18. Singhal A, Seborg D. E., "Pattern Matching in Historical Data Using PCA", IEEE Control Systems Magazine, 2, 1759-1764, 2002.
19. Srinivasan R, Wang C, "Dynamic principal component analysis based methodology for clustering process states in agile chemical plants", Ind. Eng. Chem. Res. ,43, 2123–2139,2004.
20. Wang X Z, Mc Greavy C." Automatic classification for mining process operational data", Ind. Eng. Chem. Res, 37, 2215–2222, 1998.
21. Wang X Z, Li RF, "Combining conceptual clustering and principal component analysis for state space based process monitoring", Ind. Eng. Chem. Res., 38, 4345–4358, 1999.

Chapter 6

CONCLUSIONS AND RECOMMENDATIONS FOR FUTURE WORK

Chapter 6

CONCLUSIONS AND RECOMMENDATIONS FOR FUTURE WORK

6.1. CONCLUSION

Industrial crystallizers are usually operated in two modes continuous and batch wise. Extensive research has been carried in this field. Continuous crystallizers are operated at the steady state. The steady state CSD gives determines the product quality. In batch crystallization, the quality of the product quality is determined by CSD obtained at the end of the batch, which influenced by the operation of the process during the entire batch. In the batch crystallization, the cooling rate is manipulated to obtain the desired CSD. Thus the problem of trajectory planning comes in.

Multivariate Quality Control and univariate SPC models known at present are not entirely accurate. They do give a good estimation of the process or system, but they make so many assumptions about the data. Data sets are mostly auto correlated. Therefore, Multivariate Statistical Process Control (MSPC) techniques, which consider all the variables of interest simultaneously and can, extract information on the behavior of each variable or characteristic relative to the others. MSPC research is having high value in theoretical as well as practical application and is certainly beneficial to process monitoring, abnormal condition detection and process identification & control. From this perspective, the current project was taken up.

Present work could successfully develop a proper control strategy for both batch and continuous crystallizer as well as successfully implemented various MSPC techniques in process identification, monitoring and control of crystallization process. The deliverables of the present dissertation are summarized as follows:

- The crystal size distribution of a batch crystallizer was projected by the numerically solving of a mathematical model using crystallization kinetics like nucleation and crystal growth. The results clearly show the prospective benefit of controlled cooling (constant nucleation rate) for improving the crystal size of the product compared to that obtained by natural and linear cooling.

- The mathematical modeling and simulation of continuous crystallizer was performed to obtain the dynamics. Then a PID controller to control the crystallizer temperature which in turn leads to the improvement in crystal size distribution has been proposed. The overall effect is that the crystal concentration is maintained well within the metastable zone and there is significant improvement in the crystal mean size and coefficient of variation of the whole distribution.
- Implementation of clustering time series data and moving window based pattern matching for detection of faulty conditions as well as differentiating among various normal operating conditions of continuous crystallizer.

6.2. RECOMMENDATION FOR FUTURE WORK

- MPC control of batch and continuous crystallizer.
- Identification of batch and continuous crystallization process dynamics in latent subspace using partial least squares (PLS)
- Development of PLS controllers for crystallization process
- Integration of statistical process monitoring and control for crystallization process.