Process Analytical Technology (PAT) Control Tools for High Cell Density Cultivation of Glycoengineered *Pichia pastoris* for Human Interferon α2b Production

A Thesis

Submitted for the Partial Fulfilment of the

Requirements for the Degree of

Doctor of Philosophy

By

Allampalli Satya Sai Pavan

Department of Biosciences and Bioengineering

Indian Institute of Technology Guwahati

Guwahati 781 039, Assam, India

February 2024

Dedicated to my wife

and parents

Indian Institute of Technology Guwahati
Guwahati 781 039, Assam, India

DECLARATION

I, hereby declare that the research findings in this thesis titled "Process Analytical Technology (PAT) Control Tools for High Cell Density Cultivation of Glycoengineered *Pichia pastoris* for Human Interferon α2b Production" is the result of research work carried out by me under the supervision of Professor Senthilkumar Sivaprakasam, Department of Biosciences and Bioenginering, Indian Institute of Technology Guwahati, for the award of the degree of Doctor of Philosophy. This work has not been submitted elsewhere for any degree or membership of any institute or university to the best of my knowledge and belief. Also, due acknowledgements have been made, wherever the research findings of other researchers have been cited in this thesis.

Date: 23/02/2024

Place: Guwahati

Allampalli Satya Sai Pavan

Roll No. 176106014

Indian Institute of Technology Guwahati
Guwahati 781 039, Assam, India

CERTIFICATE

It is certified that the work described in this thesis entitled "Process Analytical Technology (PAT) Control Tools for High Cell Density Cultivation of Glycoengineered *Pichia pastoris* for Human Interferon α2b Production" by Mr. Allampalli Satya Sai Pavan for the award of degree of Doctor of Philosophy is an authentic record of the results obtained from the research work carried out under my supervision in the Department of Biosciences and Bioengineering, Indian Institute of Technology Guwahati, India. This work has not been submitted elsewhere for the award of any other degree.

Date: 23/02/2024

Prof. Senthilkumar Sivaprakasam

(Thesis Supervisor)

Acknowledgements

In embarking upon the journey of doctoral research, I have been both humbled and enriched by the myriad individuals whose unwavering support and contributions have played an instrumental role in shaping this dissertation.

As I navigate the intricate landscape of academia, I express my sincere gratitude to my supervisor Prof. Senthilkumar Sivaprakasam for providing the opportunity to pursue research and a platform to learn many exciting things. Throughout this intellectual expedition, he has been offering sage guidance, insightful feedback and encouraging me in difficult times. I am thankful to him for providing all the necessary facilities to pursue my research work.

I extend my gratitude to my doctoral committee members, Professor K. Pakshirajan, Professor G. Pugazhenthi and Dr. S. Narayanasamy for their encouragement since state-of-art-seminar, constructive suggestions and discussions that helped me to execute the work successfully.

My sincere gratitude goes to my senior colleague Dr. Naresh Mohan who has generously helped me in familiarising with the PAT topic. Discussions with him has enriched my bio-thermodynamics knowledge and my interest has further piqued in the area of PAT.

I would like to thank my senior colleague Dr. Srikanth Katla who was instrumental in providing the necessary skillset to execute tasks related to *Pichia pastoris*.

I thank my lab colleagues Dr. Ganesh Nehru, Dr. Subbi Rami Reddy, Dr. Rengesh, Dr. Kiran Kumar, Dr. Nivedhitha Swaminathan, Ms. Sandhya Sekhar, Ms. Ankita Saikia, Mr. Sivakumar, Ms. Shikha, Mr. Sanjeev Kumar, Mr. Manoj Kumar for their camaraderie, support, insights, all joyous and memorable tea breaks/outings that we spend together.

I extend my sincere gratitude to my junior colleague Ms. Sandhya Sekhar, whose steadfast presence and solace during difficult times helped me in overcoming obstacles.

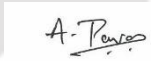
I thank my friends, Dr. Rakesh, Hanumanth, Suraj, Dr. Bapi, Dr. Manoj, Kamal, Deepa, Anjali, Mayur, Arman, Alok, Pratik, Barlina, Muktasree, Dr. Ajith, Dr. Bharat, Dr. Mallikarjun, Mukesh, Late Kranthi, Dr. Viswanth, Naveen Krishna, Deepak Bhargav, Deepak, Prem Kumar, Vamsi Krishna, Sandeep, Srinivas Kumar, Dr. Raghu Vamsi, Raja Sekhar for the direct and indirect support and the collective spirit of friendship has made the journey not only bearable but memorable.

I also thank Department of Science and Technology (DST), Government of India for providing grant (CRG/2019/002882) to carry out the research work. I want to thank the Ministry of Education (MoE), Government of India for providing the fellowship during my Ph.D. tenure.

I am thankful to the all support staff of BSBE department, IIT Guwahati for their help. My sincere thanks to Mr. Nurul Islam, for granting access to various facilites round the clock which helped me to accomplish the work.

My special thanks to Ambica Ma'am for the support, concern and delicious food that has been provided on various occasions.

Most importantly, whatever I have accomplished today is due to the unconditional love and encouragement of my parents, brother throughout my Ph.D. endeavour. As a new entrant into my life, my beloved wife Navya has become an indispensable pillar of encouragement and resilience. I would like to thank my aunt and uncle for their timely support during the fag end of my Ph.D. journey. Finally, I am indebted to the almighty who stood with me all the time.



Allampalli Satya Sai Pavan

Roll No. 176106014

Table of Contents

List of Figures

List of Tables

Nomenclature

Variables

C : Capacitance (pF/cm)

$C_{p,r}$: Specific heat of reaction broth (kJ/kg °C)

F_0 : Initial feed flow rate (mL/h)

F : Feed flow rate (mL/h)

k_c : Controller gain

K_P : Process gain

$\dot{m}_g$: Mass flow rate of gas (kg/s)

$\dot{m}_w$: Mass of water (kg)

q_C : Compensation heat rate (W/L)

q_J : Heat flow from reaction broth to jacket (W/L)

q_E : Environmental heat loss (W/L)

q_{ST} : Agitation heat rate (W/L)

q_S : Specific methanol consumption rate (g / g h)

q_A : Aeration heat loss (W/L)

q_B : Biological heat rate (W/L)

q_P : Specific productivity (mg/g h)

S_0 : Initial substrate concentration (g/L)

S_{in} : Inlet substrate concentration fed into the fermenter (g/L)

S : Substrate concentration in the fermenter (g/L)

T_r : Temperature of reaction broth (°C)

$T_{j,in}$: Jacket inlet temperature (°C)

$T_{j,out}$: Jacket outlet temperature (°C)

V_R : Volume of reaction broth (L)

V_0 : Volume of reaction broth at the end of batch phase (L)

X : Biomass concentration (g/L)

X_0 : Biomass concentration at the end of batch phase (g/L)

$Y_{Q/X}$: Biomass heat yield coefficient (kJ / g of biomass produced)

$Y_{X/gly}$: Biomass yield per glycerol consumed (g. biomass / g. glycerol)

$Y_{X/met}$: Biomass yield per methanol consumed (g. biomass / g. methanol)

$Y_{huIFN\alpha 2B/met}$: Product yield per methanol consumed (mg. huIFN α2b / g. methanol)

$y_{CO_2,in}$: Mole fraction of CO_2 in the air inlet stream

$y_{CO_2,out}$: Mole fraction of CO_2 in the air outlet stream

$y_{N_2,in}$: Mole fraction of N_2 in the air inlet stream

$y_{N_2,out}$: Mole fraction of N_2 in the air outlet stream

Greek symbols

ε : Error

τ : Torque (N. m)

τ_i : Integral time constant (s)

τ_d : Derivative time constant (s)

μ : Specific growth rate (h^{-1})

μ_{est} : Specific growth rate estimated online (h^{-1})

μ_{SP} : Specific growth rate at the setpoint (h^{-1})

μ_{met} : Specific growth rate in methanol induction phase (h^{-1})

μ_{ip} : Specific growth rate in induction phase (h^{-1})

Acronyms

ANN : Artificial Neural Networks

CER : Carbon-dioxide Evolution Rate

CGMP : Current Good Manufacturing Practices

CHO : Chinese Hamster Ovary

CQA : Critical Quality Attributes

CPP : Critical Process Parameter

DAQ : Data Acquisition

DS : Dielectric Spectroscopy

DO : Dissolved Oxygen

FB : Feedback

FDA : Food Drug Administration

FF : Feed Forward

FNN : Fuzzy Neural Network

GS : Gain Scheduling

GP : Gaussian Process

HA : Hyaluronic Acid

huIFN α2b : Human Interferon α2b

ISE : Integral Square Error

MA : Moving Average

MAE : Mean Absolute Error

MANN : Multiphase Artificial Neural Networks

MBC : Model Based Control

MPC : Model Predictive Control

MRAC : Model Reference Adaptive Control

MSE : Mean Square Error

NI : National Instruments

PAT : Process Analytical Technology

PI : Proportional Integral

PID : Proportional Integral Derivative

PTM : *Pichia* Trace Metals

QbD : Quality by Design

RBF : Radial Basis Function

RMSE : Root Mean Square Error

RNN : Recurrent Neural Networks

SCADA : Supervisory Control and Data Acquisition

SG : Savitzky Golay

SGR : Specific Growth Rate

OUR : Oxygen Uptake Rate

YPD : Yeast Extract Peptone Dextrose

YPG : Yeast Extract Peptone Glucose

Abstract

Interferons are a group of multifunctional secreted proteins categorized as cytokines involved in intracellular signalling. Human interferon α2b (huIFN α2b) is a type I interferon and one of 13 variants of interferon α that have been reported. huIFN α2b triggers many biological activities such as antiviral, antiproliferative, and immunomodulatory functions. As a result, it is used in the treatment of hepatitis B and C, hairy cell leukaemia, melanoma, and AIDS-related Kaposi's sarcoma. In the context of recombinant protein production, yeast expression systems, particularly *Pichia pastoris*, offer advantages over bacterial hosts. They facilitate the synthesis of glycosylated recombinant proteins through appropriate post-translational modifications and entail lower operational costs than mammalian cell lines. However, challenges arise in managing higher expression, leading to the accumulation of improperly folded proteins in the Endoplasmic Reticulum (ER). Inappropriately/unfolded protein fractions in culture supernatants require extensive purification steps and are strongly discouraged from the Quality by Design (QbD) perspective. In accordance with the Process Analytical Technology (PAT) initiative, the critical process parameters (CPPs) and critical quality attributes (CQAs) of the process need to be identified and monitored in real-time to achieve improved product quality. Therefore, optimization of the protein titer from *P. pastoris* could be achieved by developing relationships between the various CPPs/CQAs and product titers. Cultivation of *P. pastoris* by manipulating the specific growth rate (μ), a CPP, has a stronger influence at the metabolic level. Therefore, controlling μ in real-time leads to enhanced product output. However, deconvolution of the soft sensor/online sensor input into CPPs is at the nascent stage for most therapeutic protein production. Therefore, this thesis focuses on the appropriate utilization of soft sensors and the development of control strategies for real-time monitoring and control of CPPs governing huIFN α2b production.

Chapter 1 of this thesis deals with the introduction and state-of-the-art literature on real-time monitoring and control of μ.

Chapter 2 of this thesis deals with the appropriate selection of metabolic heat rate based soft sensors for the real-time estimation of μ. Accurate and reliable estimation of μ in real-time is pivotal for reliable monitoring of bioprocesses and the subsequent implementation of advanced control strategies. Gibb's free energy dissipation is imminent for any biological system, and the

metabolic heat rate measurements (calorimetry) formed the basis for estimating μ. However, the rationale behind selecting a suitable estimator model from a calorimetric perspective remains unexplored. The current chapter addresses the notion behind the selection of an appropriate estimator for μ, and the assessment of the estimator models was illustrated using different types of energy metabolism, that is, high exothermic and low exothermic processes. The results indicate that the μ values from the instantaneous heat rate estimator deviated significantly (10-fold higher) from the offline values for highly exothermic processes. Notably, the cumulative heat-based estimator accurately estimated μ values for both types of energy metabolism with performance metrics of < 0.005 h^{-1}.

Chapter 3 focused on the robust control of μ in glycoengineered *P. pastoris* for the production of huIFN α2b. Reliable real-time μ values from the cumulative heat-based soft sensor were considered as input values for the control strategies employed in this chapter. The feedforward strategy employing a predetermined exponential feeding of methanol during the induction phase was applied at defined setpoint values μ_{SP}. A standard PID controller with predetermined gain values regulated the methanol feeding in accordance with the deviation from the μ_{SP} value. An adaptive PID (gain scheduling) significantly minimized the deviation of μ from its μ_{SP} value, reduced the amplitude of oscillation, and achieved long-term controller stability. Robust control of methanol feeding by adaptive PID resulted in a 1.5- and 2.2-fold increase in the productivity of huIFN α2b compared to standard PID and feedforward controls, respectively. Moreover, adaptive PID control facilitated narrow-range control of μ for longer durations (> 20 h) with a low average tracking error (< 6 %), enumerating its scope of application in therapeutic protein production.

Chapter 4 focuses on the effect of methanol concentration on huIFN α2b production. Similar to any other recombinant protein produced in the methylotrophic yeast *P. pastoris,* huIFN α2b production in *P. pastoris* depends on methanol utilization during the induction phase. The present chapter investigated the impact of varying residual methanol concentrations on huIFN α2b productivity and assessed control strategies for the tight control of residual methanol concentration in fed-batch experiments carried out in a fermentation calorimeter. Real-time monitoring of *P. pastoris* metabolism was facilitated by metabolic heat rate measurements and an exhaust gas analyser. Screening of optimal methanol concentrations (1.5, 3, 5, and 7 g/L) revealed the highest huIFN α2b productivity at 3 g/L (0.21 mg/g h). The methanol sensor provided input data for three different control strategies (on/off, proportional integral (PI), and

model-based adaptive PI) to maintain the optimal methanol levels (3 g/L). Model-based adaptive control resulted in the highest huIFN α2b yield, productivity (244.34 mg/L, 0.31 mg/g h), surpassing on/off (188.7 mg/L, 0.16 mg/g h) and PI (202.3 mg/L, 0.28 mg/g h) control strategies. The model-based adaptive PI control framework developed in this chapter can be employed for heterologous protein production in *P. pastoris*.

Chapter 5 deals with a combinatorial approach of controlling μ through optimal feeding of methanol during the induction phase Since μ and residual methanol concentration influenced huIFN α2b productivity. Model predictive control (MPC) has emerged as a robust approach to realise enhanced control over process parameters in bioprocesses. The effectiveness of the MPC hinges on the availability of a robust and workable process model. Although mechanistic models are preferred, practical constraints may limit their feasibility, and researchers have adopted data-driven approaches. In this work, we demonstrate the applicability of two data-driven models, namely, Artificial Neural Networks (ANN) and the Gaussian Process (GP), in MPC applications in fed-batch cultivation of glycoengineered *P. pastoris* for the production of huIFN α2b. Experimental verification was carried out by fed-batch cultivation in a fermentation calorimeter. Real-time monitoring of *P. pastoris* metabolism was facilitated by metabolic heat rate, capacitance, and exhaust gas analyser measurements. GP-based MPC demonstrated better control, efficient utilization of substrates, and 1.1-fold enhanced huIFN α2b productivity compared to ANN-based MPC. Moreover, optimal feeding adapted by the GP-based MPC resulted in a 14% decrease in methanol utilization compared with the ANN-based MPC.

In summary, the approach presented in this thesis follows the measuring, modelling, monitoring, and control (M3C) strategy. The integration of the M3C strategy in bioprocess development will impart quality to the process and significantly improve product quantity and quality.

Chapter 1 Introduction and Review of Literature

1.1 Human Interferon α2b

In the chronicles of scientific history, 1957 marked a pivotal juncture with the collaborative efforts of Alick Isaacs, a distinguished British bacteriologist, and Jean Lindemann, an accomplished Swiss microbiologist. Together, they unveiled a momentous revelation, identifying a soluble factor emanating from chick cells afflicted by the influenza virus. This discerned factor demonstrated a remarkable capacity to impede the inherent vulnerability of neighbouring cells within the culture medium to both homologous and heterologous viral assaults. Given its inhibitory prowess, the nomenclature "interferon" was judiciously bestowed upon this newfound entity, signifying its profound interference with viral propagation dynamics [1]. Interferons (IFNs) are cytokines elicited in response to diverse synthetic and biological triggers. Constituting integral constituents of the foremost defensive echelons, they emanate from a multigene family endowed with the capacity to instigate robust antiviral, antiproliferative, and immunomodulatory responses [2]. IFNs are categorized into three classes, Type I, II, and III IFNs, based on the type of receptor through which they signal. Human interferon α2b (huIFN α2b) is a Type I IFN that encodes a protein polypeptide of 165 amino acids [3]. Characterized by four cysteine residues forming two disulfide bridges [4], huIFN α2b possesses a molecular architecture that underlies its functional diversity. This cytokine can elicit a spectrum of responses encompassing antiviral, antibacterial, antiproliferative, and immunomodulatory effects [5]. Consequently, huIFN α2b, administered independently or in conjunction with other interferons, is extensively employed in prophylaxis against type B and C hepatitis [6,7]. Furthermore, its therapeutic utility extends to the management of diverse malignancies, including melanoma [8,9], AIDS-related Kaposi sarcoma [10], chronic myeloid lymphoma [11], and angioblastoma [12]. The multifaceted applications of huIFN α2b underscore its significance as a versatile therapeutic agent in addressing a range of pathological conditions.

1.2 Microbial production of huIFN α2b

The intrinsic scarcity and high cost associated with natural IFN production prompted a paradigm shift with the advent of recombinant DNA technologies in the 1980s. Recombinant protein production, facilitated by prokaryotic, eukaryotic, and mammalian expression systems, emerged as a compelling alternative to the conventional extraction of IFNs from natural sources [13].

In pursuit of scalable production, recombinant huIFN α2b has been investigated across various expression systems, including bacteria, yeasts, insect cell lines, mammalian cell lines, and

plants. Among these, the Gram-negative bacterium *Escherichia coli* has attained pre-eminence due to its rapid growth, facile cultivation, and the availability of genetic tools [14]. This favoured status led to the development of the first FDA-approved recombinant huIFN α2b drug, Intron A, utilizing the *E. coli* expression system. However, despite its preference, the *E. coli* expression system is not without shortcomings. Codon bias, the formation of inclusion bodies, and the absence of post-translational modifications are notable drawbacks associated with this host. While mammalian cells offer the advantage of producing glycoproteins with a glycan profile closely resembling that of humans, several drawbacks impede the widespread adoption of this expression platform. The inherent heterogeneity in glycans obtained, coupled with batch-to-batch variation, low yield, challenges in scale-up, and associated production complexities, pose significant impediments. Additionally, the high production costs associated with mammalian cell expression further underscore the limitations of this platform [15].

In efforts to surmount the limitations associated with other expression platforms, yeasts have emerged as a particularly favoured and attractive host for the production of various recombinant proteins and biopharmaceuticals within the spectrum of eukaryotic microorganisms. Yeast-based expression systems confer several advantages that distinguish them as a preferred choice. Notably, yeasts exhibit superior fermentation characteristics, including the capacity for high cell density cultivation, lower nutritional requirements, and the ability to secrete produced recombinant proteins extracellularly [16]. These attributes, coupled with high protein yields, contribute to the economic viability of the production process, rendering yeast-based systems an optimal solution for the scalable and cost-effective generation of recombinant proteins.

Pichia pastoris, a methylotrophic yeast, stands out as one of the prominent expression systems for recombinant protein production [17]. *P. pastoris* demonstrates the capacity to oxidize methanol through a methanol utilization pathway for both energy generation and biomass formation. The presence of robust and tightly regulated inducible AOX1 (alcohol oxidase) promoter is well suited for controlled expression of foreign genes [18,19]. Moreover, recent strides in glycoengineering represent a significant advancement, offering a transformative tool for enhancing protein properties by selectively introducing or eliminating glycans. This innovative approach has empowered *P. pastoris* to generate complex humanized recombinant proteins [20].

Owing to the characteristics above and advantages, our research group has successfully expressed huIFN α2b in *P. pastoris* [21].

1.3 Process Analytical Technology

Process Analytical Technology (PAT), an initiative spearheaded by the United States Food and Drug Administration (FDA), serves as a regulatory framework for overseeing operations in chemical, biochemical, pharmaceutical, and biopharmaceutical industries. PAT has been defined as a "system for designing, analysing, and controlling manufacturing through timely measurements (i.e., during processing) of critical quality and performance attributes of raw and in-process materials and processes, with the goal of ensuring final product quality" [22]. One of the principal objectives of PAT implementation is to design a process that can handle variability to deliver consistent product quality [23].

Product quality is greatly influenced by the Critical Quality Attribute (CQA), which is defined as "a physical, chemical, biological or microbiological property or characteristics that should be within an appropriate limit, range or distribution to ensure the desired product quality" [22]. From the PAT implementation perspective, it operates on measuring, modelling, monitoring, and control methodology (Fig. 1.1).

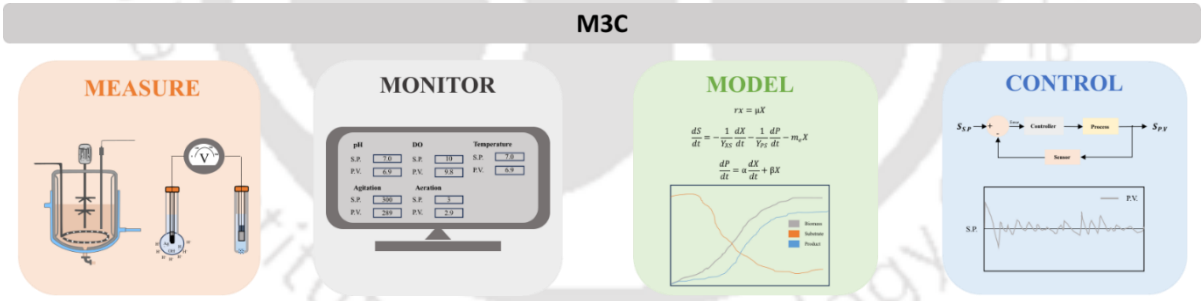


Figure 1.1: Elements of M3C technology to implement PAT framework

It is pivotal to identify the CQAs and their intricate relationship with the Critical Process Parameters (CPPs). The next step involves selecting a suitable process analyser to monitor CPPs and CQAs. Based on the method of analysing the sample, PAT process analysers are segregated as at-line (sample removed, isolated from, and analysed close to process stream), online (sample removed for analysis from the process stream and returned to process stream again), in-line (sample not removed but analysed in place), and off-line (sample removed and

analysed away from the process stream). The application of these PAT tools provides us insights into the physical, chemical, and biological attributes of the process [24].

Nevertheless, the application of PAT tools and the establishment of reliable online sensors present an ongoing challenge, particularly in the context of biotherapeutic production [25]. Addressing this challenge involves the development and deployment of cutting-edge technologies capable of providing instantaneous and accurate data on CPPs and CQAs. These PAT tools and sensors must be robust, adaptable, and capable of operating seamlessly within the dynamic environment of the microbial cultivation process [26].

Soft sensors represent a sophisticated approach to process monitoring, harnessing mathematical algorithms to analyse measurement device signals and generate valuable information about the current state of a dynamic process [27,28]. The increasing computational capacity of computers and advancements in signal processing have not only made soft sensors attractive but also positioned them as a powerful tool for monitoring and controlling various industrial processes [29]. For soft sensors to be truly valuable in bioprocess monitoring, their capacity to measure critical parameters beyond the scope of existing methods is paramount [30]. The observation that numerous industrial bioprocesses operate under sub-optimal conditions underscores the importance of focusing soft sensor monitoring efforts on generating information conducive to control towards optimal process conditions [31]. This strategic emphasis aims to leverage soft sensor outputs within control loops, where CPPs estimated by the soft sensors are actively managed to remain at optimal levels throughout the entire operational span of the bioprocess.

1.4 Significance of Specific Growth Rate (SGR) control

Today, the majority of the industrial-scale fermentations for the production of a wide array of products operate via fed-batch fermentation, owing to numerous advantages ranging from attaining high cell density and higher product yields to avoiding overflow metabolism. In the context of fed-batch fermentation control, feeding substrates and additional nutrients that influence the specific growth rates are carried out via prior expertise and optimal open-loop control [32]. SGR stands as a crucial variable in the cultivation process designated as a CPP, reflecting the physiological state of the cell culture. Its significance also extends to the biosynthesis of the intended product, with the SGR often profoundly influencing the final product quality [33–35]. In certain instances, the maximum specific product formation rates manifest at SGR that is lower than the maximal SGR (μ_{max}). Conversely, in other cases, the

specific product formation rate rises steadily with the growth rate. Notably, for Crabtree-positive microorganisms, exceeding critical specific growth rates triggers a metabolic shift to diauxic growth because of surpassing respiratory capacity. Consequently, the metabolism shifts results in the production of overflow metabolites, subsequently impeding growth, diminishing productivity, and negatively impacting product quality [33].

However, SGR control is not straightforward as there is no direct information in real-time about SGR due to the absence of online sensors. To surmount this challenge, simple controllers like open-loop control are widely used in the industry (Fig. 1.1). The inherent problem associated with simple control, like open-loop control and standard PID control, may not exhibit high efficiency in the control of SGR due to non-linearity [35]. To achieve batch-to-batch reproducibility and transferability, dynamic feedback control methods become crucial. Even though PID controllers have been established as workhorses for the control of various process parameters, the effectiveness of PID controllers with fixed tuning parameters is severely impeded due to significant variations in the process dynamics. Therefore, several approaches have been proposed to tune PID control parameters in microbial cultivations. These methods include gain-scheduling, first-principle models, and fuzzy systems. Other advanced control systems, such as adaptive controllers and model predictive controllers, have been implemented to control various CPPs that vary dynamically during cultivation [35–37]. A brief discussion about multiple control strategies and soft sensors implemented to control SGR is summarised in the preceding sections.

1.5 Control strategies for SGR control

1.5.1 Open loop control

Open loop control methodology is the most prevalent approach in industrial-scale fed-batch fermentation processes. The predetermined substrate feeding is executed according to the initial process conditions and specific operating parameters, primarily governed by batch kinetic principles. For the initial growth stage of a process, predetermined exponential feed profiles are frequently discussed, where an exponential profile can be calculated based on the initial conditions and strain-specific parameters such as the maximum SGR. Several studies reported the implementation of conventional open-loop control for various microbial systems [38–43]. The predetermined feeding of the limiting substrate is carried as represented in Eq. (1.1) and (1.2)

$$F(t) = F_0 e^{\mu_{sp} t} \tag{1.1}$$

$$F_0 = \frac{X_0 V_0 \mu_{sp}}{Y_{X/S}.\ S_0} \tag{1.2}$$

Where $F(t)$ is the feed rate (mL/h) at the time t, F_0 (mL/h) is the initial feed flow rate, X_0 (g/L) is the biomass concentration at the start of the fed-batch phase, $Y_{X/S}$ (g biomass/g substrate) is the batch phase biomass yield on the substrate, V_0 (L) is the culture volume at the start of the fed-batch phase, and S_0 (g/L) is the inlet concentration of the limiting substrate.

To guarantee the quality of recombinant therapeutic proteins produced in mammalian cell cultures, batch-to-batch reproducibility must be maintained. A simple open-loop strategy was devised to facilitate glutamine feeding for CHO cells, and the process deviations were minimized by operating the SGR well below the maximum SGR [42]. Conventional feeding strategies don't account for the loss of biomass through foaming; this is a significant problem, especially in the production of biosurfactants. To address this issue, a control law was formulated to maintain the SGR at the desired setpoint for the production of an antibiotic surfactant using *Bacillus subtilis*. The control law accounted for the loss of biomass, and the feeding rate equation was formulated as represented in Eq. (1.3).

$$V \frac{Y_{X/S} S_{in}}{\mu V} \frac{dF_{in}}{dt} = \mu V \frac{Y_{X/S} S_{in}}{\mu V} F_{in} - X_{out} F_{in} \tag{1.3}$$

Where, F_{in} represents the feed rate (mL/h), X_{out} represents biomass loss due to foam formation (g/L). Even in the case of an open loop strategy, the new feeding control law could maintain SGR at 0.05 $h^{-1,}$ leading to enhanced biosurfactant production [44]. Two evolutionary approaches were developed to optimize the fed-batch fermentation of *S. cerevisiae*. The optimization variants included maximizing the biomass concentration and controlling SGR. An optimal feed profile was attained with the aid of genetic algorithms, where the feed rate is designed by applying constraints [45].

The primary advantages of utilizing this method for SGR control on a large scale are its straightforwardness in application and the absence of the necessity for information regarding measured variables. However, open-loop control has a significant limitation as in scenarios where defined feeding strategies don’t consider any variation in feed concentration, open-loop control fails to adapt to disturbances.

1.5.2 Closed loop control with PI/PID controllers

The inherent disadvantages associated with open-loop control systems are circumvented with the introduction of closed loop control systems in the form of PI and PID controllers. The feedback mechanism enables the system to adjust the deviations from the setpoint by setting a regulatory action (Fig. 1.2). Indeed, the PID controller has become a cornerstone in both academic research and industrial applications as the control law integrates both feedforward and feedback terms. The PID controller's popularity stems from several advantages, including its simplicity in implementation and ease of tuning [36,46].

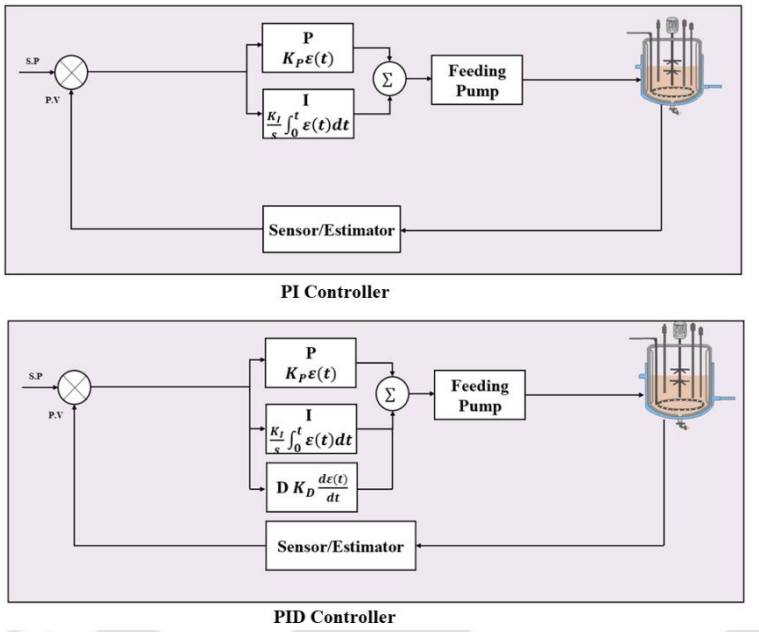


Figure 1.2: Block diagram of PI and PID controller manipulating substrate feed rate

In feedback control, real-time measurement data is employed directly to rectify deviations of the process from its intended trajectory. Within this context, there exists a distinction between control schemes implemented for the control of routinely controlled variables such as temperature and pH. However, in the context of controlling physiological variables like SGR, more intricate regulatory mechanisms are required to manage the parameters at the desired values. Several researchers comprehended the complexity of SGR control and designed the cultivation processes that relied on achieving optimal or at least quasi-optimal SGR profiles, which lead to enhanced productivity [47–50]. A generic model control with a feedforward – feedback part has been developed to deal with the dynamics of fed-batch in recombinant protein production. The approach employed in this work was distinct from typical PI

controllers. In this method, rather than adjusting the action variable directly, the focus was on the rate of change of the control variable, which is dependent on two proportional components. One is proportional to the deviation of the control variable from its setpoint, and another is proportional to the time integral of this deviation, as represented in Eq. (1.4).

$$\frac{d\mu_{set}}{dt} = k_1(\mu_{set} - \mu) + k_2 \int_0^t (\mu_{set} - \mu)\, dt \tag{1.4}$$

Where, k_1 and k_2 are adjustable tuning parameters.

The final control law was based on the Eq. (1.5)

$$F = \left(\frac{\mu}{Y_{X/S}} X + \frac{k_1(\mu_{set} - \mu) + k_2 \int_0^t (\mu_{set} - \mu)\, dt}{Y_{X/S}\sigma_{max} \frac{(K_S - S^2/K_I)}{(K_S + S + S^2/K_I)^2}} \right) \frac{V}{S_F - S} \tag{1.5}$$

Where, σ_{max} represents maximum substrate uptake rate, K_S represents Monod saturation constant, K_I represents inhibition constant. The control mechanism adapted to the deviations in the setpoint by adjusting the feeding rate appropriately and thereby minimizing the fluctuations in desired SGR profiles for the production of recombinant proteins [51]. The non-linearity exhibited during the regulation of SGR was successfully dealt with by the application of a non-linear PI control algorithm was proposed [52]. The control system was based on a minimalist model approach, necessitating solely the measurement of biomass and volume, in addition to certain limitations on the reaction rate. The controller is designed as a partial state feedback mechanism with an adaptable gain. It utilizes a PI algorithm based on the notion of invariant control. The nonlinear integral action was devised to maintain the invariance of a target manifold as specified by the reference model dynamics. Subsequently, a proportional output error feedback component is integrated into the control law to expedite the convergence process. At the same time, PI/PID controllers are robust in terms of controlling physical process parameters and variables where linearity is involved. However, the control action is impeded by the introduction of non-linear dynamics, alterations in operating conditions, and disturbances. To overcome these limitations, more advanced control strategies such as adaptive control, fuzzy control, model predictive control, and data-driven control could be employed for improved performance and robust control.

1.5.3 Fuzzy control

Fuzzy logic serves as the basis of fuzzy control, designed to address system uncertainties without the necessity of complex models. As fed-batch fermentations exhibit highly non-linear behaviour a rigorous approach is necessary to capture the system dynamics for the control of state variables. Nevertheless, fuzzy logic employs a fundamentally different approach compared to other model-based techniques, as it doesn't require any prior information regarding the process dynamics. Based on the current state of the process and the user's experience with the process, fuzzy logic rules are formulated to achieve the desired control objectives [47]. The transformation of quantitative data into qualitative characteristics is the foundation of fuzzy control. The process of establishing the input-output relationship within a system using fuzzy logic involves a set of if-then rules and an interface mechanism (Fig. 1.3).

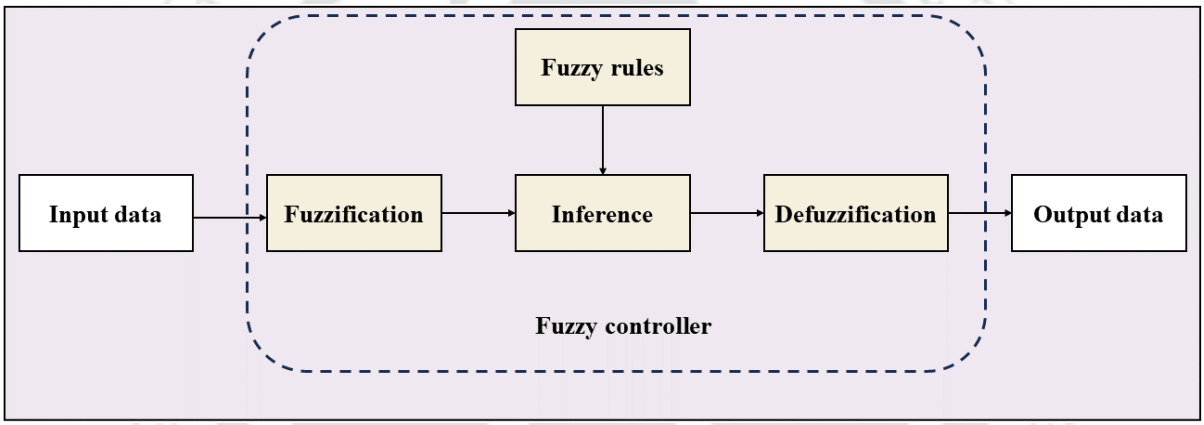


Figure 1.3: Working principle of fuzzy logic controller

Initially, numerical input data is transformed into 'membership functions' by fuzzification, which evaluates the degree to which these numerical values align with a fuzzy set (typically represented by a value between 0 and 1) [53]. Fuzzification entails defining a range of potential values for each input variable, thereby establishing a scale of comparison. Each variable is characterized using a fuzzy set, assigning it a degree of membership based on its conformity to this set. Fuzzy control relies on a series of fuzzy rules that articulate the system's conditions. These rules are derived from the operator's experience with the process and are structured as conditional statements using terms like 'if' and 'then.' For instance, a rule might state: "If the substrate concentration (S) is deemed 'High,' then the substrate feed rate (F) should be 'low'. In this scenario, the controller evaluates the substrate concentration as high and, as a consequence, adjusts the substrate feed rate to a lower level. Since it operates based on linguistic rules rather than rigid mathematical systems, it becomes more adaptable to various

processes or scales. Due to the greater flexibility and ease of operation fuzzy controllers have been successfully employed in bioprocess applications: for example, to the control the fermenter temperature, [54] to control the substrate concentration [55,56]. However, to control the challenging bioprocess state variable, i.e., the SGR, only a few reports of fuzzy logic being used to solve the control problem are available. The first application of fuzzy logic for specific growth rate control was demonstrated on *Saccharomyces cerevisiae* strain [57]. Fuzzy rules were based on values of dissolved oxygen, respiratory quotient (RQ) values. The optimal specific growth rate was maintained by regulating the substrate feed rate, as mentioned in Eq. (1.6) and (1.7), respectively.

$$F_{in}^{*} = (1 + \alpha)F_{in}\ (RQ < 1) \tag{1.6}$$

$$F_{in}^{*} = (1 - \alpha)F_{in}\ (RQ > 1) \tag{1.7}$$

Where, F_{in}^{*} is practical feed rate, F_{in} is the ideal feed rate, and α is the positive fuzzy factor.

A feedforward-feedback control strategy with a five-layer fuzzy neural network (FNN) was implemented for the control of fed-batch fermentation of recombinant *E. coli* JM 103 harbouring plasmid pUR 2921. Variations in SGR and pH were used as inputs for FNN to calculate the glucose feeding rate (output). A four-fold increase in the relative activity of β-galactosidase was obtained by employing two FNNs [58]. An optimized feeding rate of carbon substrate was achieved by implementing an FNN controller on yeast strain for fed-batch fermentation. The five-layer FNN controller included cell concentration, glucose concentration, and increment in cell concentration as the input data and optimized glucose feed rate as the output data. It is well established that values estimated from soft sensors are highly susceptible to deviations from the original value due to process-related disturbances, and to counteract such deviations, a robust control mechanism is necessary to achieve tight control over SGR. An adaptive fuzzy logic-based control algorithm coupled to the PI controller was implemented to control μ on recombinant *E. coli* BL21 strain harbouring pBR322 plasmid [59]. Interestingly, the fuzzy logic was associated with the PI controller, and the input variables were oxygen uptake rate (OUR) and weight of the culture broth (w), the output variables being PI controller parameters, controller gain K_C and integration time constant T_i. The fuzzy rules were created based on the heuristic knowledge as follows

IF OUR/w is poor, THEN T_i is big

IF OUR/w is excellent, THEN T_i is small

IF OUR/w is poor, THEN K_C is small

IF OUR/w is excellent, THEN K_C is big

The developed controller demonstrated better stability over a wide setpoint range than the gain scheduling adaptive controller. Thus, fuzzy logic, in combination with any controller or observer, proves to be a promising control strategy for tight control over various bioprocess state variables. Nevertheless, there are significant limitations linked with fuzzy logic systems. These limitations encompass challenges such as imprecise parameter estimation due to improper fine-tuning of rules, a lack of adaptability for dynamic process states, alterations in process variables due to sensor failure, and inexperience of process operators. Consequently, it is challenging to have robust control over bioprocess state variables when a fed-fermentation undergoes any afore-mentioned problem.

1.5.4 Adaptive control

The application of a standard PID controller is effective in regulating process parameters like temperature, pH, and dissolved oxygen. However, for highly dynamic variables, the fixed tuning parameters of the PID controller (k_C, τ_I, τ_D) might not inherently provide the most appropriate control action. Effectively controlling a dynamic variable necessitates adjusting the tuning parameters of the PID controller based on the response of the controlled variable. Adaptive control strategies encompass non-linear control algorithms that adjust the controller parameters automatically during the ongoing process (Fig. 1.4). These algorithms incorporate various strategies with a common goal of modifying specific control parameters to address the non-linear dynamics and system uncertainties more effectively. Adaptive control proves to be a valuable alternative in cases where the structure of kinetics and the precise kinetic parameters are inaccurately or imprecisely known.

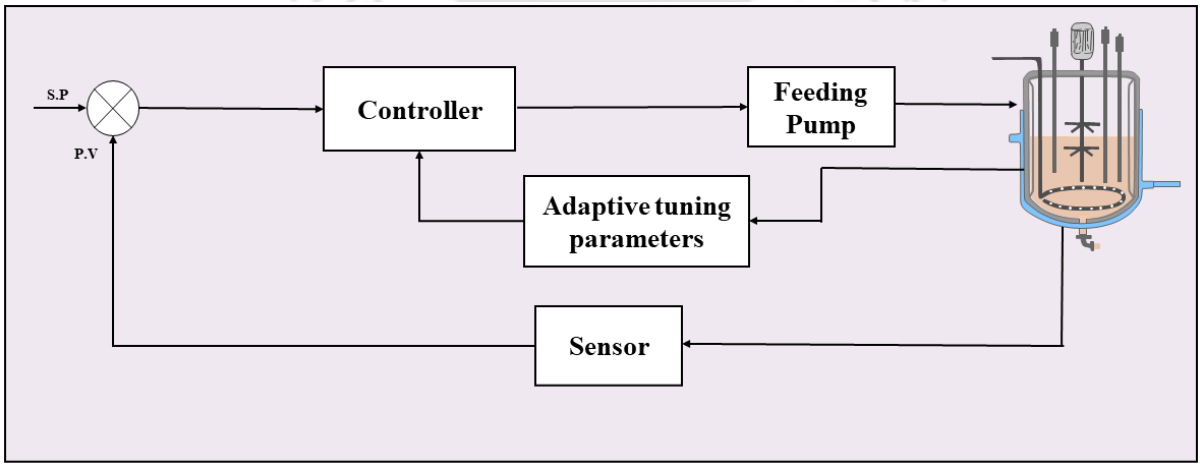


Figure 1.4: Block diagram of an adaptive control

Based on the method used for parameter estimation and adaptation, the controller types are segregated. As fed-batch fermentations are associated with non-linear dynamics and system uncertainties, the adaptive nature of various controller types makes them intriguing for the fed-batch fermentation process. Gain scheduling represents one approach within adaptive control. In this method, the controller tuning parameters aren't fixed but instead adapted based on prior knowledge of the system, allowing it to accommodate and address changing dynamics within the system. Based on data analysis from previous batches and examination of the controller's response over time, the pre-programmed tweaking can be executed. Gain scheduling was used to manage the feed rate for a model baker's yeast system. When the tuning was off, it also led to oscillatory behaviour at specific times [60]. An enhancement to the traditional gain scheduling method involves the online adaptation of controller parameters utilizing available measurement data. Adaptation allows the controller to adjust to non-predictable system dynamics dynamically. In specific scenarios, the available measurement data is initially processed to predict a particular system parameter or state of interest. Subsequently, this predicted parameter or state is integrated into an adaptation algorithm, further refining the controller's parameters and improving its responsiveness to the system's changing dynamics [59,61]. In the aforementioned adaptive control methods, the control law is iteratively adjusted either via a predetermined gain or through updates to the process model, where changes in the model directly impact the control law. However, an alternative approach involves defining the optimal control action required to regulate a process variable. The control action is aimed at reducing the discrepancy between the ideal model and the process outputs. Known as model reference adaptive control (MRAC), this technique has been utilized to address the challenge of feed rate control [62]. A stable MRAC was designed to control SGR for *Bordetella pertussis* fed-batch fermentations. A key advantage of this design is that the only essential online measurement necessary is that of dissolved oxygen. The validity of the design was examined by performing the fed-batch cultivation of *B. pertussis.* The controller effectively maintained the SGR at the desired set point, consistently managing the increased demands for substrates and oxygen throughout an extended fed-batch cultivation duration [63].

A comparative study focussed on designing two distinct adaptive control algorithms for regulating SGR in fed-batch biotechnological processes using *E. coli* for recombinant protein production. The results indicated that both controllers, i.e., a gain scheduling-based adaptive PI controller and a model-free adaptive controller utilizing artificial neural networks, showed comparable control performance. They were particularly effective when employing the

substrate limitation approach and manipulating substrate feeding rates. The study involved testing the controller performance under realistic ranges of feedback signal sampling intervals and measurement noise intensities. Considering the efforts engaged in controller design and tuning, especially in the development of adaptation/learning algorithms, the model-free adaptive control algorithm emerged as more suitable for industrial applications, especially in scenarios with limited knowledge of the process and its mathematical model. The investigation revealed that the model-free adaptive controller displayed better control quality, particularly under low specific growth rate conditions observed during the phase of recombinant protein production. Throughout the simulation runs, the average tracking error did not exceed 0.01 h^{-1} and the temporary overshoots caused by maximal disturbances stayed within the range of 0.025 - 0.11 h^{-1} [64]. Adaptive control is particularly well-suited for dynamic systems characterized by significant disturbances in control parameters. Unlike other control methods that reject disturbance variables, adaptive control inherently adjusts to accommodate disruptions in the control parameters [65]. Consequently, this approach is highly relevant to fermentation systems with unpredictable system dynamics and erratic disturbances, as it possesses the inherent capability to adapt to these unforeseen changes and maintain effective control over the system.

1.5.5 Model Predictive Control (MPC)

Due to its capacity to manage intricate multivariate systems, model predictive control is a widely utilised control technique [66]. MPC involves assessing the variance between predicted and reference values of the controlled variable (SGR) to ascertain the appropriate course of control action. A robust predictive process model must be in place for simulating the fermentation process, foreseeing future timeframes, and predicting current outputs and forthcoming system states. MPC relies on these predictions, considering an optimization process across the entire duration of fermentation (Fig. 1.5).

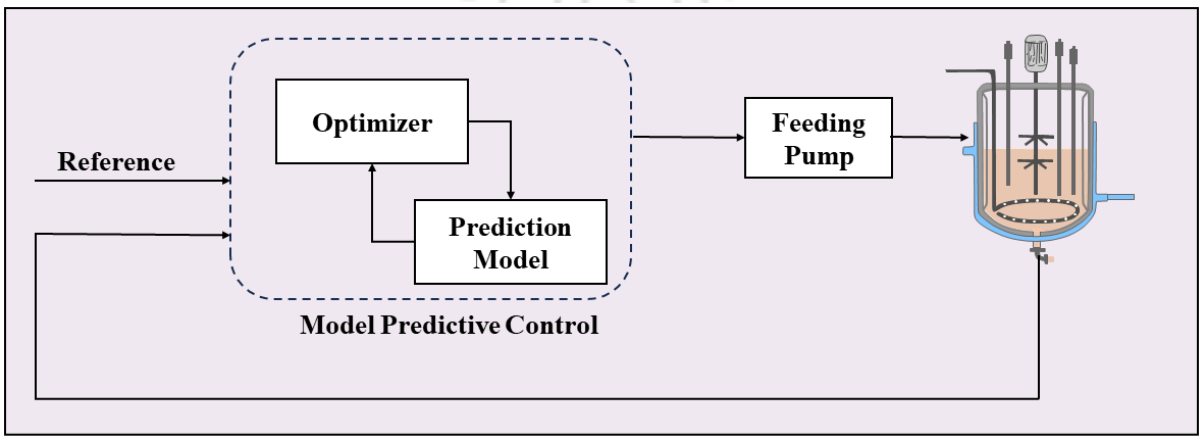


Figure 1.5: Block diagram of a model predictive control manipulating substrate feed rate

The optimization of a predefined cost function determines the most suitable control action at the current time, often in the form of corrections to the feed rate. The method enables the system to dynamically adjust the control actions in response to discrepancies between predicted and desired outcomes, offering a proactive and precise means of control in fermentation.

In utilizing MPC to regulate feed rates, the selection of an appropriate optimization function holds pivotal importance. Two relevant approaches involve tracking a reference trajectory for either biomass concentration or substrate concentration throughout the cultivation process. These reference trajectories can be derived from various sources, including previously recorded data mathematical models. Employing standard sensors such as pH and dissolved oxygen (DO), along with at-line measurements of biomass and glucose, the substrate feed rate was adeptly manipulated through a facile integration of MPC within an industrial bioreactor system. Through this approach, they have successfully demonstrated the MPC's capability to adhere to a predetermined biomass growth profile [67]. MPC system was developed to attain consistent batch-to-batch reproducibility in an animal cell culture, specifically targeting recombinant therapeutic protein production from CHO cells. The primary control objective was to regulate the SGR optimally by managing the oxygen mass consumed by cells through adjustments in the glutamine feed rate. A favourable assessment of the controller's performance was reported, particularly highlighted by the high batch-to-batch reproducibility achieved in cultures operating under this control system. The outcome underscores the effectiveness of the MPC in ensuring consistent and reliable outcomes across multiple batches in the production process [68]. A comparative analysis of two distinct model-based control strategies for concurrently managing two individual substrate uptake rates via two substrate feeds within an *E. coli* fed-batch process was conducted. The controllers evaluated were an elemental balance controller (EBC) and MPC based on a mechanistic model. Both controllers exhibited similar performance and demonstrated competence in fulfilling their designated tasks. The authors determined that, for the specific application under investigation, the EBC was favoured due to its straightforward nature. However, they acknowledged that the potential of the MPC lies in its predictive capabilities and adaptability to diverse objective functions. Consequently, the MPC was deemed a more suitable choice when optimizing for product-related objectives [69]. In a *Penicillium chrysogenum* fed-batch process, a nonlinear Model Predictive Controller (MPC) was implemented and contrasted with a PI(D) controller and an open-loop feedback control scheme referred to as model-based control (MBC). The control actions were guided by state

estimation and predictions derived from a kinetic model, adjusted to suit the specific control objectives. State estimations were facilitated through modifications, including the simplification of the hyphal compartmentalization and the incorporation of measurable OUR (oxygen uptake rate) and CER (carbon dioxide evolution rate). Experimental validation revealed unstable behaviour in the PI(D) controller due to nonlinear process dynamics.

In contrast, the MPC showcased efficient avoidance of by-product formation, leading to enhanced substrate utilization and an overall productivity gain of 14% compared to both the PI(D) and MBC approaches [70]. MPC is a robust closed-loop control approach specifically designed for controlling non-linear processes. It optimizes control actions over the entire process duration, not merely at the current time instant, and incorporates disturbance modelling as an integral part of the optimization problem. Although MPC is extensively used in various industries, transitioning to the bioprocess industry demands a notable leap. To facilitate its adaptation in the bioprocess domain, substantial effort is required to develop dependable and accurate process models. Moreover, the effectiveness of MPC is contingent upon the accuracy of the process model and its capability to manage unforeseen disturbances. Nonetheless, certain drawbacks of this method exist. One primary limitation is its reliance on robust process models, which may not always be readily available or completely accurate. Additionally, MPC tends to be computationally intensive, requiring significant computational resources, which can pose challenges in real-time implementation.

1.5.6 Artificial Neural Networks (ANN) based control

Neural networks can deal with non-linear systems and also adaptively learn about process dynamics; hence, they are extensively used to solve various engineering problems [71]. A complex non-linear system can be described using an Artificial Neural Network (ANN), a data-driven method, without the use of explicit model equations, and its applications are extensively increasing in bioprocess monitoring and control (Fig. 1.6). Different variants of neural networks were implemented in bioprocess monitoring and control purposes. An online state estimator was developed through a Radial Basis Function (RBF) neural network and these estimates were used for the closed loop control which resulted in minimum tracking error compared to open loop strategy [72]. Multiphase Artificial Neural Networks (MANN) has been designed to predict the biomass concentration in various phases of fungal cultivation [73]. Data availability is the major limitation associated with data-driven modelling approach.

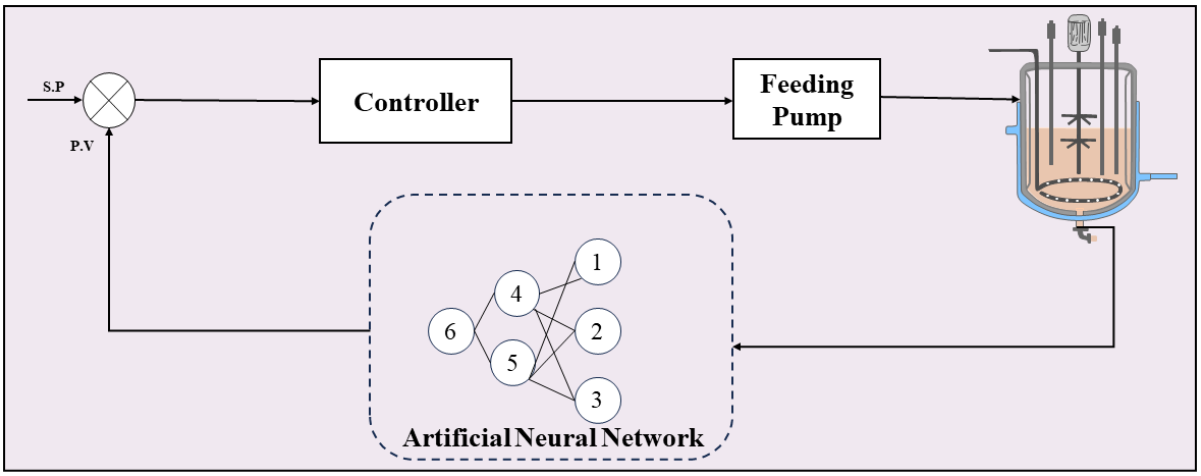


Figure 1.6: Block diagram for artificial neural network-based feedback control

Substantial amount of process data is necessary to train the networks and establish a relationship between the trained data and SGR. Biomass concentration, which is an offline measurement and to generate a huge amount of data for this variable some interpolation techniques were employed. Owing to this reason, there is limited literature for SGR control using data-driven approach. A constant, optimal SGR value was maintained by regulated feeding of methanol in *P. pastoris* for the production of alpha 1- antitrypsin. A Multi-Layer Perception 3 (MLP3) neural network was developed to prevent the time loss and controller divergence. Based on the weight matrices the controller mitigated the error rate which resulted in enhanced production of alpha 1- antitrypsin [74]. A Recurrent Neural Network (RNN) with a topology of 3:4:1 was used to estimate biomass concentration and SGR in real-time for the production of HBsAg in *P. pastoris*. DO concentration, CER and methanol concentration served as inputs for RNN. Constant SGR was maintained by combining the RNN with a PID controller and methanol feeding was regulated accordingly (Table 1.1) which resulted in higher productivity compared to conventional open loop control [75]. ANNs and other data-driven approaches have been used for control applications, since it has been demonstrated that they can accurately predict the behaviour of fermentation systems based on the recorded data. Particularly, in situations when the first principle-based models and empirical models couldn't be used to represent the complex, nonlinear biological systems, data-driven based control renders advantages in terms of dealing with nonlinearity, adapting to the continuously evolving systems. However, the inability to read the resulting network in order to comprehend linkages between variables makes this method less advantageous than other control strategies. As a result, little process knowledge is gained. Additionally, it should be noted that the network is not scalable because it was only trained for one scale and operation of the process, and it must

be retrained for other scales and operations. Nonetheless, the limitations of data-driven approaches in deciphering the resultant network structure to comprehend variable linkages pose a drawback. Consequently, limited process knowledge is acquired, hindering a deeper understanding of the relationships between variables. Moreover, it is important to highlight the scalability of data-driven approach is a concern as they are primarily trained for a specific process scale and operational configuration. Retraining is necessary when transitioning to different scales or operations, making it less adaptable across diverse operational contexts [76,77].

Table 1.1: Overview of various control strategies for SGR control

Controller	Control Law	Objective	Reference
Open loop	$F(t) = F_0 e^{\mu_{sp} t}$	Maximize productivity	[40,78–85]
	$V\frac{Y_{X/S}S_{in}}{\mu V}\frac{dF_{in}}{dt} = \mu V\frac{Y_{X/S}S_{in}}{\mu V}F_{in} - X_{out}F_{in}$	Account for biomass loss due to foaming	[44]
Feedback with PI&PID	$F(t) = F_0 e^{(\mu_{SP} - \mu_{PV})t}$	To avoid overflow metabolism and maximize productivity	[70,75,84, 86,87]
Adaptive Control	$F_{in} = \frac{(\lambda_1 + \lambda_2 X)(E_{set} - E) + \theta_P X}{Y_{S,fer,E}S_{in} - E}$	To avoid overflow metabolism and maintain SGR at desired setpoint	[85]
	$F = \lambda X \nu$	The feedback gain is continuously adapted to maintain SGR	[88]

Fuzzy Logic	$F_{in}^{*} = (1 + \alpha)F_{in}\ (RQ < 1)$ $F_{in}^{*} = (1 - \alpha)F_{in}\ (RQ > 1)$	To maintain SGR closer to the maximum SGR	[57]
Fuzzy Logic with Neural Networks	Linguistic rules	To control SGR and glucose concentration	[58]
Fuzzy Logic + PI controller	Linguistic rules	Maximize productivity with a stable SGR	[89]
Probing Control	$F_k = \frac{K}{h}((C_{sp} - C_k) + \frac{h}{T_i}\sum_{1 \le j \le k}(C_{sp} - C_j))$	Ensure continuous exponential feeding of substrate with DO as input signal for SGR control	[90]
ANN based control	$F(t) = F_0 e^{\mu t}$	Predict the changes in SGR through neural network and regulate the feed to maintain optimal SGR	[74]

RNN based control	$F(t) = K_p\varepsilon(t) + K_I\int_0^t \varepsilon(t)dt + K_D d\varepsilon(t)/dt$ $\varepsilon(t) = \mu_{set} - \mu_{observed,ANN}$	Tight control of SGR with real-time prediction from RNN based observer	[75]
MPC	$J_k = \sum_{i=1}^{P} \frac{\lvert y_{k+1,j} - y_{sp,k+1,j}\rvert}{y_{k+1,j}} w_j + \sum_{i=1}^{M} \frac{\lvert u_{k+1} - u_k\rvert}{\Delta u}$	To track the predefined SGR	[70]
	$J_k = \sum_{i=1}^{P}\left(q_{Glu} - q_{Glu,sp}\right)^2 + \sum_{i=1}^{P}\left(q_{Lac} - q_{Lac,sp}\right)^2 + \sum_{i=1}^{M}\lVert\gamma\Delta F^2\rVert$	To control the substrate uptake rate which has direct influence on SGR	[69]

1.6 Soft sensors for real-time estimation of SGR

The ability to adjust control actions based on the deviations from the setpoint becomes feasibly only when measurements/estimates of the controlled variable are readily available. Information obtained from offline and at-line measurements is often delayed for process control due to labour-intensive manual steps and is suspectable to various sources of error [91]. Soft sensors are amalgamation of online measurements from process analysers and mathematical models employed to provide the estimates of difficult-to-measure variables. Soft sensors can be categorized into knowledge-driven and data-driven approaches. Knowledge-driven strategies are developed from fundamental principles, detailing the relationships between process variables and the quality attributes. However, the accuracy of these models relies heavily on the available process knowledge (e.g., first-principles understanding). Soft sensors are amalgamation of online measurements from process analysers and mathematical models employed to provide the estimates of difficult-to-measure variables. Soft sensors can be categorized into knowledge-driven and data-driven approaches. Knowledge-driven strategies are developed from fundamental principles, detailing the relationships between process variables and the quality attributes. However, the accuracy of these models relies heavily on the available process knowledge (e.g., first-principles understanding) [92]. In contrast, data-driven soft sensors use multivariate data analysis tools, such as principal component analysis or partial least squares, to derive models based on the available data [93]. In the following sections soft sensors utilized for real-time estimation of SGR from various process analysers are discussed.

1.6.1 Soft sensor based on off-gas analysis for SGR estimation and control

Accurate estimates of SGR might be impeded when there is no reliable sensor or model to determine biomass concentration online. Additionally, when the biomass concentrations are low the accuracy of SGR is hampered with oscillations due to instability in SGR estimator values. An alternative approach to estimate SGR is to use other directly measurable stoichiometric variables such as the substrate uptake rate, oxygen uptake rate, carbon dioxide evolution/production rate and base consumption rate [33]. Oxygen uptake rate (OUR) and carbon dioxide evolution rate (CER) measurements provide an accurate estimate of the cellular demands, as cells replicate more strongly, their OUR and CER increase. Based on this fact a simple closed-loop control for SGR was developed with CER as input signal. SGR estimator was modelled based on a linear relationship between CER and SGR as represented in Eq. (1.8)

$$CER(t) = \alpha\mu(t)X(t) \tag{1.8}$$

A predefined SGR profile corresponds to a unique OUR/CER profile. Thus OUR/CER can be used as controlled variable for indirect control of SGR. Utilizing this approach glutamine feed rate for CHO cell cultivation was manipulated according to a modified P control algorithm as represented in Eq. (1.9)

$$F_{Gln} = (1 + K_P E)F_{set} \tag{1.9}$$

Where, F_{set} is the reference feed rate, i.e., the substrate feed rate for the undisturbed process, E is the error term. Compared to the basic P controller, which solely relies on the error to determine the control action, an adaptation to the process dynamics were considered by incorporating the desired substrate feed rate. A more comprehensive approach for accounting the fluctuations and modifications within the system allowed for a refined control where the viability of CHO cells never dropped below 93% when SGR was maintained at 0.02 h^{-1} [78]. With a minimalistic mathematical modelling and a no prior knowledge of the microbial system an accurate SGR control system was proposed [94]. The substrate feeding rate was manipulated to control OUR during the cultivation process in such a way that OUR-based ratio R (Eq. (1.10)) was stabilized at the desired SGR setpoint, then the SGR will asymptotically approach the setpoint.

$$R = \frac{dOUR}{dt}\frac{1}{OUR} \tag{1.10}$$

For the control of R, PI controller was implemented and the controller gain was adapted to the time varying dynamics using gain scheduling approach with the feeding rate as scheduling variable. The dependence of different control algorithms on the accuracy of biomass growth models for the online estimation and control of SGR in fed-batch fermentations can sometimes introduce challenges due to discrepancies between the plant and process models. To overcome this issue, soft sensors that do not rely on a specific biomass model can be employed. A soft sensor was developed for the real-time estimation of SGR. This soft sensor utilized online measurement data of OUR dynamics. By considering the dynamics of oxygen consumption and a tuning parameter ((β)/(α)) specific to the microbial strain, the study achieved the real-time estimation of SGR without the necessity of a biomass growth model [95]. The results from the specific growth rate (SGR) control study reported in reference [63] have demonstrated that a conventional control system utilizing Proportional-Integral (PI) controllers can effectively achieve acceptable control quality. However, applying gain-scheduling algorithms, as discussed in the adaptive control section, and using the CER signal as a scheduling variable

offers a more flexible approach to adapt controller parameters to the time-varying dynamics of the controlled process, resulting in improved control quality. It's worth noting that a well-established method involves implementing a Luedeking-Piret type correlation between CER and biomass growth rate, which has been shown to provide more accurate results.

Another option for real-time SGR estimation involves using OUR data (Table 1.2). Nevertheless, when additional oxygen is introduced into the aeration air to maintain critical dissolved oxygen levels at higher cell densities, the accuracy of OUR data can be compromised due to variations in off-gas composition, pressure, and gas flow rate. As a result, it is advisable to employ CER data in SGR estimation relationships, particularly when controlling high-density cultivation processes.

Table 1.2: Summary of soft sensor-based SGR estimation and control

Process Analyser	Input signal for estimator model	Estimator Model	Controller	References
Exhaust Gas Analyser	OUR	$\mu_{est} = \frac{r_X Y_{O_2/X}}{X Y_{O_2/X}}$	-	[96]
	CER	$\mu_{est} = \frac{r_X Y_{CO_2/X}}{X Y_{CO_2/X}}$	-	[96]
	OUR	$\mu_{est} = \frac{dOUR}{dt}\frac{1}{OUR}$	PI	[94]
	CER	$\mu_{est} = \frac{CER\,(t)}{\int_0^t CER(t)dt}$	-	[47]
Dielectric Spectroscopy	Capacitance	$\mu_{est}(t) = \frac{ln\,(C_{X,t}/C_{X,t-1})}{t_t - t_{t-1}}$	PID	[97–99]
	Cell Number	$N(t) = N_0 e^{\mu t}$	-	[100]

NIR Spectroscopy	Biomass concentration	$\mu_{est} = \frac{r_{X,NIR}}{X_{NIR}}$	-	[101,102]
Biocalorimeter	Metabolic heat rate	$\frac{dq_P}{dt} = \mu q_P$	PI	[103–106]
	Metabolic heat rate	$\mu_{\text{est}} = \frac{q_{\text{P}}}{X_0 V_0 Y_{\text{Q/X}} + (Q_{\text{t}} - Q_0)}$	PI/PID	[33,107,108]

1.6.2 SGR control based on DO concentration

To attain high volumetric productivity, it is important to have an optimized feed rate. However, determining the optimal feed rate involves finding a balance that maximizes the SGR while avoiding saturation of oxidative capacity or other rate-limiting phases in respiratory metabolism. One significant challenge in regulating the feed rate is managing the oxygen demand required by microorganisms to sustain exponential growth at the desired SGR. At sufficiently high biomass concentrations, this demand can surpass the oxygen transfer capacity of the fermenter system, presenting a limitation to continued growth and productivity. The oxygen demand of the microorganisms for maintaining exponential development at the appropriate SGR could exceed the oxygen transfer capacity of the fermenter system at sufficiently high biomass concentration. Therefore, the maximum volumetric oxygen transfer rate was be used to determine the maximum permitted substrate feed rate. A different approach known as "probing control" bypasses aforementioned challenges by assessing the system's response to variations in the feed rate. In this method, the dissolved oxygen (DO) saturation within the fermentation broth is a valuable response variable due to its direct correlation with overflow metabolism and oxygen consumption. Regular and precise measurements of DO content in the fermentation broth is are frequently obtained through reliable probes. Leveraging these measurements provides a practical and effective means of controlling the feed rate in the fermentation process. A probing control approach was demonstrated on industrially relevant hosts such as *S. cerevisiae*, *P. pastoris* and *E. coli*. Initially the limiting substrate was feed at a predetermined rate in exponential manner. Once the system attained stability the feeding is either stopped or reduced to observe the response of the DO signal. If there is no overfeeding, DO signal rises as the substrate is limiting and once feeding is initiated the DO signal declines rapidly. The size of the peak is function of DO level, the level of perturbation of the feed rate, and the perturbation duration. The detection of under and over-feeding is simple with this method and very sensitive because little accumulation of overflow metabolites/substrate leads to a significantly smaller signal. Typically, a piecewise linear feed rate trend results from this approach. A final biomass concentration of 118 g/L was achieved with *P. pastoris*, 93 g/L for *E. coli* and 85 g/L for *S. cerevisiae* [39]. To address the limitations of a linear feed rate trend an alternative strategy was devised, where a different method of analysis of the response to the external perturbations was used. In this method, calculation of the frequency spectrum of the DO signal was performed and the power spectral density for frequencies close to that of the externally applied disturbances was then employed to derive the control variable. The current

approach enabled the continuous tracking of optimal feed rate, effectively sidestepping the inherent risks involved with probing control. The developed method was implemented on pilot scale fermenter for the production of amylase using a model strain of *Bacillus licheniformis*. The limiting substrate was fed into the fermenter based on the control law with a gain scheduler as presented in Eq. (1.11) and (1.12).

$$F_k = \frac{K}{h}\left(\left(C_{sp} - C_k\right) + \frac{h}{T_i}\sum_{1\le j\le k}\left(C_{sp} - C_j\right)\right) \tag{1.11}$$

$$K = K_0\sqrt{\frac{F_{k-1}}{F_{min}}} \tag{1.12}$$

Where, F_k is the feed rate at current sampling point k, T_i is integral time constant, h is sampling rate, K is the gain constant. Overflow metabolism was successfully curtailed with this methodology and an increment of 24% in biomass concentration was attained as compared to standard controller [90]. The probing control approach offered by the DO signal offers the easiest methodology for the control of SGR as there is no additional requirement of costly online sensors and mathematical models. However, in the case of high cell density cultivation where oxygen limitation is quite common, the DO signal will be subjected to lot of perturbations and may interact with SGR control loop leading to erroneous feeding of the substrate.

1.6.3 Soft sensor based on dielectric spectroscopy for SGR estimation and control

In the context of estimating biomass concentration in real-time, spectroscopy-based monitoring including fluorescence probes and dielectric spectroscopy were widely employed over a class of microbial systems [99,109,110]. However, there were limitations found with fluorescence probes when used in cultures complex media. The changes in composition of the medium influenced the fluorescence, rendering these probes less effective in such circumstances [111]. Dielectric spectroscopy emerged as better alternative due to the fact that it could measure only viable cell concentration. The principle for biomass measurement via dielectric spectroscopy is based on the function of cells as capacitors. When an electric field is introduced to a cell suspension, charge separation or polarization transpires at the cell poles. Polarization happens due to the movement of intracellular ions along the electric field, hindered by the low conductivity of the plasma membrane, which serves as an insulating barrier. Consequently, ions from both the cytoplasm and culture medium migrate towards the electrode with the opposite charge. The measured signal is a function of the volume fraction of the cells. Only

cells with an intact membrane potential are recorded with dielectric spectroscopy. Hence, the method is insensitive to dead cells and only measures viable biomass [98,112,113]. Controlling the SGR in filamentous cultures is particularly challenging due to the relatively low growth rate levels compared to organisms like *E. coli*. The low SGRs create difficulties in both controlling and assessing the control strategy, primarily because of the lower signal-to-noise ratio [96]. Additionally, fluctuations in biomass yields throughout the production process have emerged as a significant concern in SGR control. To mitigate potential error propagation in calculating the SGR, the control strategy prioritizes tracking biomass trends instead of the SGR directly. The cell growth rate was estimated by linearizing the growth rate equation for an exponential phase which is shown in Eq. (1.13)

$$\mu_{est} = \frac{\ln\left(\frac{C_{X,t}V_t}{C_{X,\Delta t}V_{X,\Delta t}}\right)}{\Delta t} \tag{1.13}$$

Where, C_X is the current biomass concentration, V_t is the culture volume at current t. Due to the fed-batch dynamics the changes in culture volume are accounted. Based on estimates derived from the dielectric signal, the controlled variable was incorporated into a suitable control algorithm. The designed control strategy was intended to be adaptable during both the growth and decline phases of the process, enabling automatic adjustments to changing biomass yields. It utilized a feedforward-feedback scheme where the error term was integrated into the exponential segment of the feed rate equation (Table 1.1). This approach significantly improved controller performance by reducing high-frequency oscillations in both the controlled and manipulated variables [97,98]. A decline in the SGR arising from factors such as nutrient depletion, accumulation of toxins, or other physiological and chemical stresses could significantly have a negative impact cellular energy production. SGR variation might affect the quantity and quality of monoclonal antibodies (mAbs). In the context of mAb production, the improper execution of post-translational modifications (PTMs) might become more frequent when SGR is not effectively regulated [114]. Real-time estimation of SGR through online monitoring of biomass growth has proven invaluable and applied to the cultivation of recombinant immunoglobulin G (IgG) on CHO cells, allowed for the avoidance of nutrient starvation. By renewing the medium based on the online estimation of SGR, a consistent value of SGR was maintained throughout the entire cultivation process [114].

1.6.4 Soft sensor based on metabolic heat rate production for SGR estimation and control

Microbial growth is facilitated by the breakdown of high-energy nutrients into lower energy components via catabolism. The energy derived from the catabolism drives the replication of cells through anabolic process and the excess energy from the catabolism process is exported as entropy from the cell membrane. According to second law of thermodynamics, this process is spontaneous and the cellular proliferation is fostered by dissipation of Gibb's energy to sustain the metabolism [115]. The quantification of net metabolic heat dissipation is effectively measured through biocalorimeter [116]. The interplay between Gibb's energy dissipation and the microbial growth has been successfully investigated, elucidating fundamental principles governing metabolic regulation and cellular proliferation [117]. Utilizing the concept of Gibb's energy dissipation, multiple frameworks have been established to analyse various types of microbial metabolism, enabling the real-time monitoring of diverse bioprocess systems. In the context of industrial-scale fermenters, their low surface-to-volume ratio leads to substantial metabolic heat generation, rendering it a crucial online signal for monitoring and controlling various bioprocess systems [118]. Catabolite repression is quite notably significant in industrially important strains like *S. cerevisiae* and consequently, the principal objective in any fed-batch fermentation involving this yeast is to mitigate the excessive production of ethanol. Calorimetric signal was as input parameter to regulate the glucose feeding to avoid catabolite repression. The feedback mechanism to the controller was based on decrement of heat signal (glucose exhaustion) and implementing controlled feeding of glucose resulted in a 10% higher biomass yield as compared to that achieved in batch fermentation [119]. A major hinderance for a lab-scale reaction calorimeter (RC1 Mettler-Toledo) is the inability to measure very low heat production rates (< 10 mW/L). Hence a modified version of the calorimeter was successfully developed and rendered it suitable for biological applications. Using the modified calorimeter (bio-RC1) fed-batch cultures of *Bacillus sphaericus* 1593M was performed to produce parasporal insecticidal crystal proteins. The process heat flow (q_P) was used to calculate the substrate consumption rate and a correlation was developed to regulate the substrate feed rate based on process heat evolved during the fed-batch cultivation [120]. The process heat flow was directly used to estimate SGR in real-time as presented in Eq. (1.14) for the cultivation of a recombinant *E. coli* K12 TB1 strain bearing the plasmid pGLO.

$$\frac{dq_P}{dt} = \mu q_P \tag{1.14}$$

The estimated SGR values were integrated into a feedforward-feedback PI controller for the purpose of regulated feeding. In the initial hour of fed-batch fermentation, only feedforward component was engaged, allowing the system to stabilize. Subsequently, both components of the controller were activated to maintain the SGR at 0.2 h^{-1}. The applied control strategy led to the production of a notably high biomass concentration, reaching 120 g/L [103]. A similar control strategy has been applied to *S. cerevisiae* and the SGR was maintained below the critical value to prevent the formation of ethanol due to Crabtree effect. With this calorimetric based control strategy high biomass concentrations of 110 g/L were obtained consistently in a 30L lab-scale biocalorimeter [104]. SGR estimator used in the previously cited literature works well for the microbial systems when the metabolic heat rate production is on lower side (< 10 W/L). However, when dealing with microbial systems that exhibit higher heat production an alternate SGR estimator becomes essential for successful control applications. A SGR estimator that considers the cumulative heat generated during the fermentation was developed as represented in Eq. (1.15).

$$\mu_{\text{est}} = \frac{q_{\text{P}}}{X_0 V_0 Y_{\text{Q/X}} + (Q_{\text{t}} - Q_0)} \tag{1.15}$$

Where, $Y_{\text{Q/X}}$, biomass heat yield coefficient (kJ/g) and Q_{t}, cumulative heat at that time instant (kJ/L) and Q_0 is cumulative heat at the end of batch phase. The developed estimator was tested on three different Crabtree-negative yeast strains *Candida utilis*, *Kluyveromyces marxianus* and *P. pastoris*. A simple PI feedback strategy, where the SGR estimated as presented in Eq. (1.15) and the error term of the controller were looped into the exponential term of the feed rate equation as shown in Eq. (1.16).

$$F_{PI}(t) = F_0 e^{((\mu_{est} + K_P \varepsilon(t) + K_I \int_0^t \varepsilon(t) dt) t)} \tag{1.16}$$

The implemented PI controller demonstrated higher stability, yielding reduced controller errors across all designated setpoints when compared to feedforward controller. The average tracking error didn't exceed 0.08 h^{-1} indicating the effectiveness of the PI controller [33]. The issue of noisy signals stemming from the SGR estimator, influenced by online estimations based on hardware sensor inputs and mathematical models, was addressed through the application of digital signal filtering methods. Filters such as moving average, and Savitzky-Golay were applied to process the raw signal emanated from the SGR estimator. The refined signal was then integrated into the PI controller for the regulation of glucose feed in fed-batch phase on recombinant *E. coli* HMS 174(DE3) strain producing green fluorescence protein (GFP). The utilization of stable input signal resulted in the satisfactory performance of the PI controller in

fed-batch phase which resulted in 12 g/L of GFP [105]. A novel predictive process control strategy employing calorespirometry, which involves simultaneous measurements of heat and CO_2 evolution rates, was developed to optimize methanol feeding for the production of biopolymer polyhydroxy butyrate on methylotrophic bacterium *Methlyobacterium extorquens*. By applying this innovative methodology, methanol concentration was allowed to maintain below the critical limit and the fed-batch fermentation was carried out at higher growth rates (0.2 h^{-1}) which resulted in production of 6.5 g/L of polyhydroxy butyrate. Moreover, the calorespirometry control strategy also facilitated detection of shifts in microbial metabolism [121]. The heat signal emanating from the biocalorimeter is the resultant of substrate consumption by the cells and SGR is estimated from this signal by applying derivative over heat signal. However, the derivative term is prone to noise and ultimately provides noisy estimates if data is not pre-treated. To avoid data pre-treatment steps and to capture the entire dynamics of the fermentation, a novel control strategy was developed which was based on directly using the heat signal. The study involved modelling time trajectories of the heat signal setpoint corresponding to different SGRs. The controller was programmed to track this dynamic setpoint during aerobic cultivations of *S. cerevisiae*. Reliable SGR control was achieved within the range of 0.075 to 0.2 h^{-1} and the average root mean square errors were observed to be 15±3%, demonstrating the accuracy and reliability of this novel control strategy [106]. Production of high molecular weight hyaluronic acid requires a tight control of SGR in a very narrow range. A PID feedback controller with metabolic heat rate as input signal for the controller was used employed to produce hyaluronic acid in *Streptococcus zooepidemicus*. The developed feedback strategy successfully established a robust control system capable of maintaining SGR in close proximity to the setpoint with minimal tracking error. The utilization of an exponential feed rate, executed at the lower SGR value (0.05 h^{-1}) resulted in enhanced molecular weight of hyaluronic acid at 2.98 MDa. Moreover, calorimetric signal-based SGR control by PID controller mitigated adverse effects resulting from the secretion of other end products, while consistently maintaining regular metabolic activities [107].

1.7 Problem statement

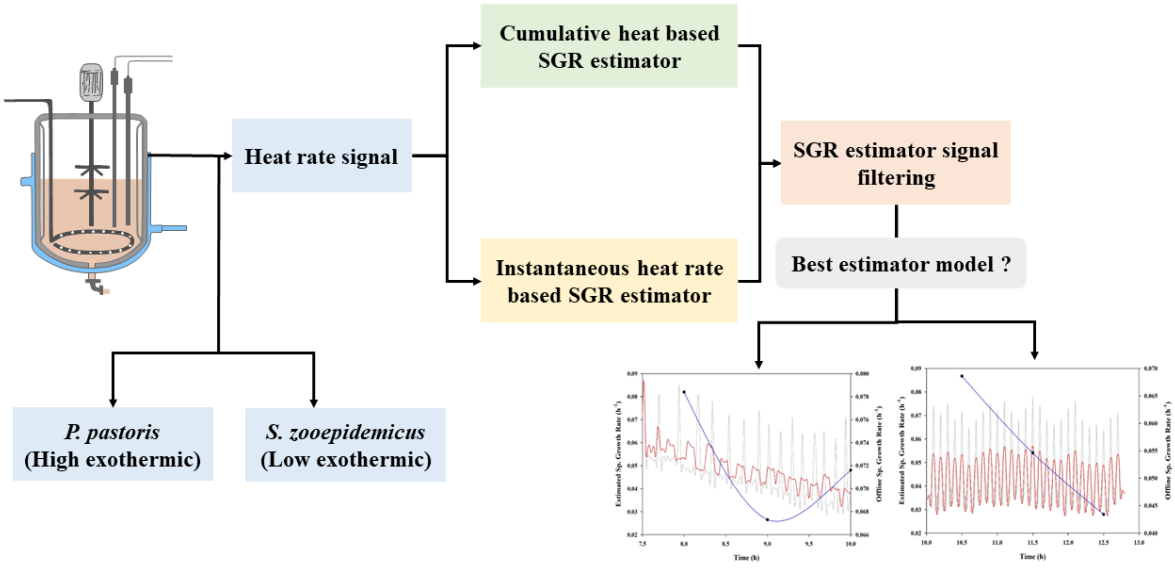
PAT initiative advocates for the identification and real-time monitoring of CPPs and CQAs of the process to achieve improved product quality. In the context of optimizing protein titer from *P. pastoris*, establishing correlations between various CPPs/CQAs and product titers is paramount. One influential CPP in *P. pastoris* cultivation is SGR which has significant impact at the metabolic level and real-time monitoring and control holds promise for augmenting

product yields. However, deconvolution of the soft sensor/online sensor input into CPPs is at the nascent stage for most therapeutic protein production. A recent study successfully utilized a fermentation calorimeter to fingerprint the growth dynamics of *P. pastoris* cultivation producing huIFN α2b and elucidate metabolic changes stemming from various carbon substrates. Consequently, this thesis endeavours to utilize fermentation calorimeter as process analyser and leverage soft sensors effectively for real-time estimation of μ using metabolic heat rate as input signal. However, the reported literature for real-time monitoring of μ with fermentation calorimeter as process analyser didn't focus on notion of selection appropriate soft sensor model. Since a reliable estimator model is of paramount importance to achieve process control related tasks as these models serve as input for the controllers. The initial part of the thesis focuses on selection of appropriate soft sensor for real-time estimation of μ. Reported literature suggests that an appreciable amount of increment in the productivity of recombinant proteins by controlling μ at the optimal values. With the aid of advanced control strategies μ will be controlled in a very narrow range for enhanced huIFN α2b production. Moreover, it is well reported in literature that the residual methanol concentration has a pronounced effect on AOX1 transcription level thereby, influencing the state variables such as specific substrate uptake rate, productivity and μ. In order to have a multi-faceted approach towards CPPs involved in the huIFN α2b a direct control of residual methanol concentration will be carried out with the aid of methanol sensor and different control strategies. Recent literature stated the growing prominence of huIFN α2b as a therapeutic agent. To meet the surge a process strategy needs to be devised where the production is optimized without compromising on the product quality. In order to meet such stringent demand literature suggests that fermentation should be regulated by advanced control strategies like MPC. Therefore, an effective control strategy using MPC will be implemented in this thesis for enhanced production of huIFN α2b.

1.8 Objectives

1. Selection of appropriate metabolic heat rate based specific growth rate estimator for reliable real-time monitoring and control of specific growth rate
2. Deployment of metabolic heat rate based soft sensor for real-time estimation and control of specific growth rate in glycoengineered *Pichia pastoris* for the production of huIFN α2b
3. Design of a model based adaptive PI control strategy for the residual methanol concentration control to enhance the production of huIFN α2b
4. Development of data-driven based model predictive control to design efficient feeding strategies for enhanced productivity

Chapter 2 Metabolic heat based specific growth rate estimators: does the choice of estimation model influence the state of bioprocesses?



2.1 Abstract

Accurate and reliable estimation of specific growth rate (μ) in real-time is pivotal for reliable process monitoring of bioprocess and subsequent implementation of advanced control strategies. Gibb's free energy dissipation is imminent for any biological system, and the metabolic heat rate measurements (calorimetry) formed the basis for estimating μ. However, the rationale behind selecting a suitable μ estimator model based on calorimetric perspective remains unexplored. This present chapter addresses the notion behind the selection of an appropriate estimator for μ and the assessment of the estimator models was illustrated using different types of energy metabolism viz., high exothermic and low exothermic processes. The results indicated that the μ values from the instantaneous heat rate estimator significantly deviated (10-fold higher) from the offline values for highly exothermic process. Notably, the cumulative heat-based estimator accurately estimated μ values on both types of energy metabolism with performance metrics < 0.005 h^{-1}.

2.2 Introduction

The microbial route for synthesizing high-value chemicals and biopharmaceutical products is increasingly gaining widespread traction due to advancements in recombinant DNA technology. Moreover, due to the expiration of existing patents, the surge for alternate drugs is being witnessed, ultimately opening up avenues for a lucrative market of biopharmaceutical products [122]. Fed-batch mode of fermentation is touted as a workhorse in bioprocess industry in order to achieve enhanced yields and productivity [84]. Besides maximizing the productivity, maintaining consistent product quality is of paramount importance especially in the case of biopharmaceutical products. The QbD initiative strongly advocates for the process control of CPPs under the ambit of PAT framework [31]. The importance of controlling CPPs that impact the product's CQAs is unequivocally well reported [25]. However, the process dynamics and the difficulty in measuring the quality-related key process variables make monitoring, optimization and control of fed-batch fermentation an arduous task [25]. The specific growth rate of an organism is often designated as a key bioprocess variable due to its direct influence on the cell physiological state and the rate of product synthesis. Moreover, the specific growth has pronounced effect on the final quality of the product [123,124]. Therefore, an accurate and reliable estimation of μ is the pre-requisite for the tight regulation of μ in desired set point value range. The caveat of the absence of a direct method to estimate μ on real-time encourages the use of indirect methods. Online estimation of μ is inferred through soft sensors which is an amalgamation of a hard sensor and a software implemented mathematical model. Without breaking the sterility barrier the online estimation of μ was progressed by combining in-situ monitoring techniques based on fluorescence spectroscopy [125], optical density probe [29] and dielectric spectroscopy [126] with an appropriate mathematical model. Major hindrance accompanying with in-situ monitoring is perturbations in the stirring rate (to maintain critical dissolved oxygen concentration and break the foam), frequent foam formation and the interference of bubbles in the proximity of sensor measuring tip. Furthermore, the morphological changes to the cell during the course of fermentation often leads to skewed online estimation μ. Other alternative soft sensor approach was based on real-time monitoring of off-gas activity [127]. Nonetheless, the manipulations in aeration and agitation rates to maintain the critical dissolved oxygen (DO) amplifies the noise in DO signal prompting to distorted off-gas measurements. Non-uniform mixing of the culture broth often results in gas hold-up of culture broth and head space of large-scale fermenter thereby, impeding the usage of off-gas based tracking at industrial scale. To address the afore-

mentioned pitfalls a non-invasive sensor capable of tracking the cellular metabolism instantaneously is required. According to the second law of thermodynamics, the microbial growth is spontaneous and to sustain the metabolism the heat is continuously dissipated [115]. The microbial growth is proceeded by the breakdown of higher energy nutrients into lower energy components through the process of catabolism and the energy generated from the catabolism fuels the replication of cells through anabolic reactions and the excess energy is exported from the cell surface as entropy. From thermodynamics perspective, the microbial growth can be viewed as a Gibb's energy transduction process and the excess entropy is exported as chemical entropy (metabolite synthesis) and dissipated as heat energy to the environment [128]. The measurement of net metabolic heat dissipation is facilitated by the means of biocalorimeters [116]. The relationship between Gibbs energy dissipation and growth was successfully explored to uncover plausible underlying principles governing metabolic regulation and growth [129,130]. Based on the Gibb's energy dissipation, several frameworks have been reported different types of microbial metabolism facilitating real-time monitoring of various bioprocess systems [131]. Industrial scale fermenters offer low surface-to-volume ratio, which attributes to significant metabolic heat generation enumerating its reliability as key online signal for monitoring and control applications [132]. Metabolic heat flow rate as an online indicator of metabolic activity was successfully deployed as PAT tool for the control of μ on Crabtree-negative *S. cerevisiae* [133]. Commercial reaction calorimeter RC1 was modified to enhance the sensitivity of the heat signal and calorimetry based control system was successfully deployed on *Bacillus sphaericus* [120]. Frequent problem of overflow metabolism in *E. coli* during fed-batch fermentation was addressed by calorimetric control resulting in high cell density culture [134]. Production of fermentation metabolites was evaded through calorimetric control of *S. cerevisiae* fed-batch fermentation [135]. The potential of biocalorimeter was further explored by controlling μ on three different Crabtree-negative yeast cultivation processes [136]. Moreover, if the industrial scale fermenters are endowed with high precision temperature probes, they could be successfully adapted as fermentation calorimeters at minimal cost. Besides, the fermentation calorimeters could serve as appropriate R&D tool for mitigating scale-up problems related to temperature gradients. Therefore, the feasibility of μ control using calorimetric measurements could be well realised as it is ubiquitous across various microbial systems. Based on published literature, the calorimetric based estimation of μ could be typically segregated into two estimator models (i) cumulative heat based, and (ii) instantaneous heat rate based. Nevertheless, the studies addressing the choice of calorimetric μ

estimator models for reliable process monitoring remains unattempted till date. Measured instantaneous heat rate is universal by-product of ongoing metabolic activities in a cell growth process and hence, it is always associated with fluctuations related to metabolic perturbations. Utilization of raw measured heat rate signal and the influence of smoothing techniques on measured signal (in case of instantaneous heat based μ estimator) are not addressed effectively in published literature. The role of heat yield in cumulative heat based μ estimator for reliable prediction of μ is not elucidated well. Especially, no literature reports are available till date addressing the comparative performance assessment of afore-cited estimators for real-time monitoring applications for e.g., high cell density cultivation, highly and weakly exothermic bioprocess systems.

This present chapter is aimed at investigating the robustness of both instantaneous and cumulative – heat based specific growth estimator models derived from real-time metabolic heat rate measurements obtained from an indigenous fermentation calorimeter. Two different bioprocess systems exhibiting high (aerobic metabolism) and low exothermicity (facultative anaerobe) were deployed in this work to address the versatility of the estimators. A comparative assessment of the estimator models was dealt in detail both from signal smoothing and calorimetric perspective. The findings of this chapter would serve as the guidance for right choice of a heat based μ estimator model which is a pre-requisite for implementing a robust soft sensor-based monitoring and control of fed-batch fermentation.

2.3 Material and Methods

2.3.1 Strains

A glycoengineered *P. pastoris* strain (bearing the plasmid SuperMan5 pep4Δ prb 1Δ procured from Biogrammatics Inc., USA) of Mut^+ phenotype expressing huIFN α2b was used in this study [21]. The glycerol stocks were maintained at -80°C in yeast peptone dextrose (YPD) media containing 20% v/v glycerol. The other strain used was *Streptococcus equi. subsp. zooepidemicus* 3523 (procured from Microbial type cell culture (MTCC), Institute of Microbial Technology, India) is a wild type strain producing hyaluronic acid [137]. It was preserved in Todd-Hewitt (TH) agar slants.

2.3.2 Media

The composition of various media used for the cultivation of glycoengineered *P. pastoris* were as follows:

Starter culture medium (Yeast extract, Peptone, Dextrose; YPD) in g/L: yeast extract 10, peptone 20, dextrose 20.

Pre-culture medium (Yeast extract, Peptone, Glycerol; YPG) in g/L: yeast extract 10, peptone 20, glycerol 20.

Optimized fermentation medium (Basal Salt Medium, BSM) [126]: glycerol 48.84 g/L, K_2SO_4 18.2 g/L, $MgSO_4$ 7.28 g/L, KOH 4.13 g/L, $CaSO_4$. $2H_2O$ 0.93 g/L, $(NH_4)_2SO_4$ 8.42 g/L, 85% H_3PO_4 26.7 mL/L PTM4 salts 4.4 mL/L.

The production media for the cultivation of *S. zooepidemicus*: Yeast extract 10 g/L, Tryptone 5 g/L, K_2HPO_4 7.95 g/L, KH_2PO_4 7.39 g/L, $MgSO_4.7H_2O$ 0.5 g/L.

2.3.3 Inoculum preparation

The glycerol stock of *P. pastoris* strain preserved at -80°C was streaked on to a YPD plate and incubated at 30°C for 24 h. A single colony from the incubated plate was inoculated as starter culture in a test tube containing YPD media. Seed culture of *S. zooepidemicus* was prepared by picking a single mucoid colony and cultured in TH broth supplemented with 5 g/L brain heart infusion and incubated overnight at 37°C under shaking condition (180 RPM).

2.3.4 Fermentation calorimeter design and principle

An indigenously built fermentation calorimeter was employed to measure the enthalpy disseminated from the macro-scale bioprocess (Fig. 2.1). The mode of operation is based on the heat compensation calorimetry, where the compensation heater wattage is regulated proportionally to the metabolic heat generated during the biochemical reaction. The heat compensation fermentation calorimeter is a specialized setup comprising a double jacketed bioreactor equipped with 4 high-sensitive temperature probes (Isothermal Technology Ltd, Merseyside, England) of 1 mK sensitivity. The inner jacket is provisioned for the circulation of silicon oil to maintain the isothermal conditions of the fermentation calorimeter. The outer jacket is sealed with vacuum to prevent the heat loss to the environment. Two high sensitive temperature probes are placed in the reactor, and one each is placed at the jacket/inlet points. Two PID controllers were in action and maintained isothermal conditions in the fermentation calorimeter. In response to the metabolic heat generated due to the cultivation of microorganisms, the first PID controller adjusts the output of compensation heater's wattage and second PID controller regulates the cryostat (Julabo GmbH, Seelbach, Germany) in response to the dynamic variations in jacket inlet temperature ($T_{j,in}$). Unravelling of metabolic heat rate measured by fermentation calorimeter was accomplished by segregation of total heat

rate into non-biological and biological heat rates. Various mechanical operations contribute to the non-biological heat rates and were estimated for each operational parameter as described in table 2.1 [137]. A dynamic energy balance was applied over fermentation calorimeter to calculate the biological heat rate. Eq. (2.1) describes the unsteady state heat rate balance contributed by different operating parameters.

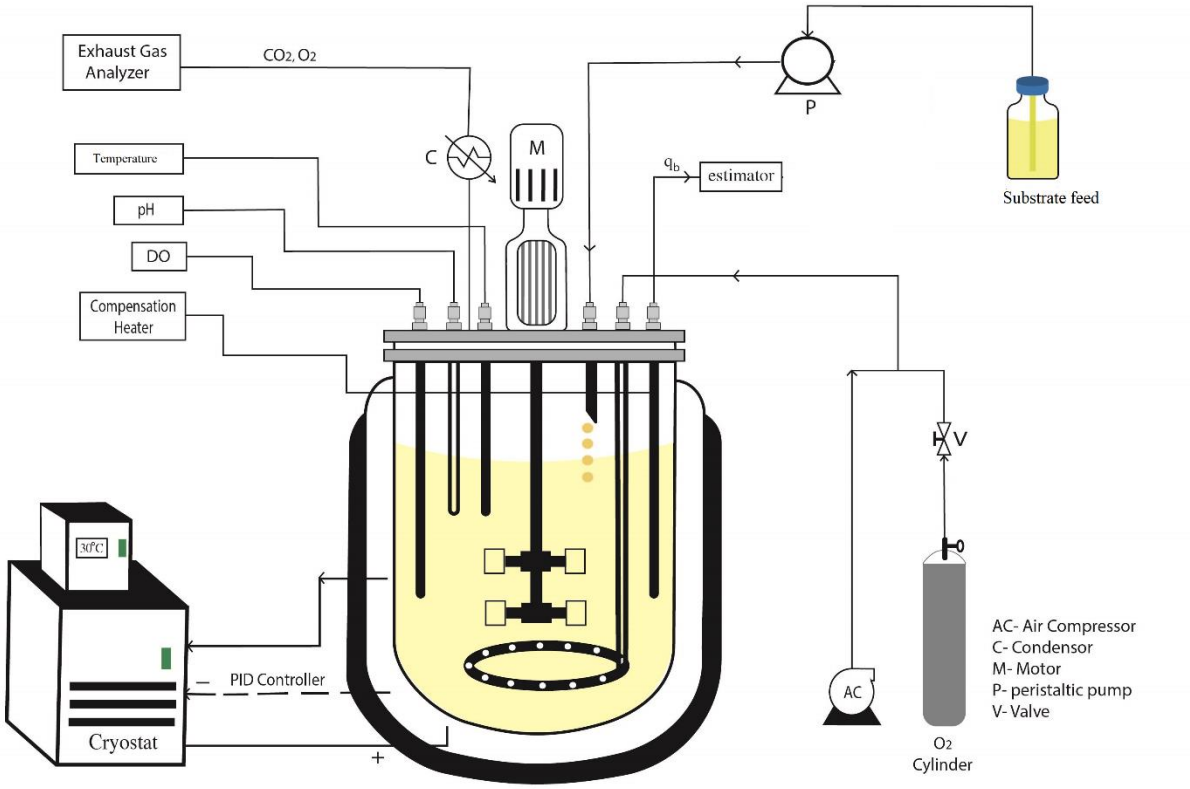


Figure 2.1: Schematic representation of heat compensation calorimeter set-up employed in this study

$$q_{\mathrm{C}} - q_{\mathrm{J}} - q_{\mathrm{E}} + q_{\mathrm{S}} - q_{\mathrm{A}} + q_{\mathrm{B}} = m_{\mathrm{w}} C_{\mathrm{P}} \frac{dT_{\mathrm{r}}}{dt} \qquad (2.1)$$

Where q_{B} (biological heat rate, W/L), q_{C} (Power generated by compensation heater, W/L), T_{r} (reactor temperature), m_{w} (mass of water kg). Heat rate contributions due to non-biological activities were found to be environmental heat loss (q_{E}, W/L), the heat flow from reaction broth to jacket (q_{J}, W/L), agitation heat rate (q_{S}, W/L), and aeration heat loss (q_{A}, W/L) and were lumped as baseline heat rate (Table 2.1). Finally, the biological heat rate is calculated from the output of compensation heater and baseline heat rate (q_{baseline}) as given in Eq. (2.2)

$$q_B = q_{baseline} - q_C \quad (2.2)$$

Table 2.1: Non-biological heat rates contributed due to various process operations in the fermentation calorimeter (Heat rate measurements are reported from [137])

Heat flow terms	Measured/Estimated parameters	Driving force	Governing equation
$T_r, T_{J,in}, T_{j,out}$	K		NA
q_S	RPM	RPM	$q_S = 2 \times 10^{-8} RPM^{-0.312}$
q_A	LPM	LPM	$q_A = -0.878 * LPM - 0.099$
q_E	$U_E A_E$	$(T_r - T_a)$	$q_E = U_E A_E (T_r - T_a)$
q_J	$U_J A_J$	$(T_r - T_{j,out})$	$q_J = U_j A_j (T_r - T_{j,out})$
q_C	$V_C . I_C$	NA	$q_C = V_C . I_C$

The baseline heat which comprises of non-biological heats is represented as per Eq. (2.3)

$$q_{baseline} = q_s - q_E - q_A - q_J \quad (2.3)$$

2.3.5 Determination of carbon dioxide evolution rate (CER)

Off gas activity was tracked by the measurement of gaseous phase concentration of CO_2 evolved during the course of fermentation. The exhaust gas analyser (Ultramat 23, Siemens AG, Berlin, Germany) detects the CO_2 through infrared (IR) absorption. As depicted in Fig. 2.2, the central component comprises Non-Dispersive IR (NDIR) consisting of an IR source, a measuring chamber, a wavelength filter, and a detector. Within the measuring chamber, any CO_2 molecule present will selectively absorb a particular wavelength of light emitted by the IR source. Subsequently, the intensity of the light detected by the IR detector is directly proportional to the quantity of CO_2 molecules within the chamber and the measure intensity is transmitted as an output signal in mole fractions ($y_{CO_2,out}$). CER is calculated from the measured values by using Eq. (2.4)

$$\text{CER}, \text{m. mol}/\text{L. h} = \frac{\dot{m}_g}{V_R}\left[y_{CO_2,out}\left(\frac{y_{inert,in}}{y_{inert,out}}\right) - y_{CO_2,in}\right] \quad (2.4)$$

Where $\dot{m}_g$ the mass flow rate of gas in Kg/s, V_R is the volume of the reaction broth in L, $y_{CO2,in}$ is the mole fraction of CO_2 in air inlet stream, $y_{CO2,out}$ is the mole fraction of CO_2 in air outlet stream, $y_{inert,in}$ is the mole fraction of N_2 in air inlet stream and $y_{inert,out}$ is the mole fraction of N_2 in the air outlet stream.

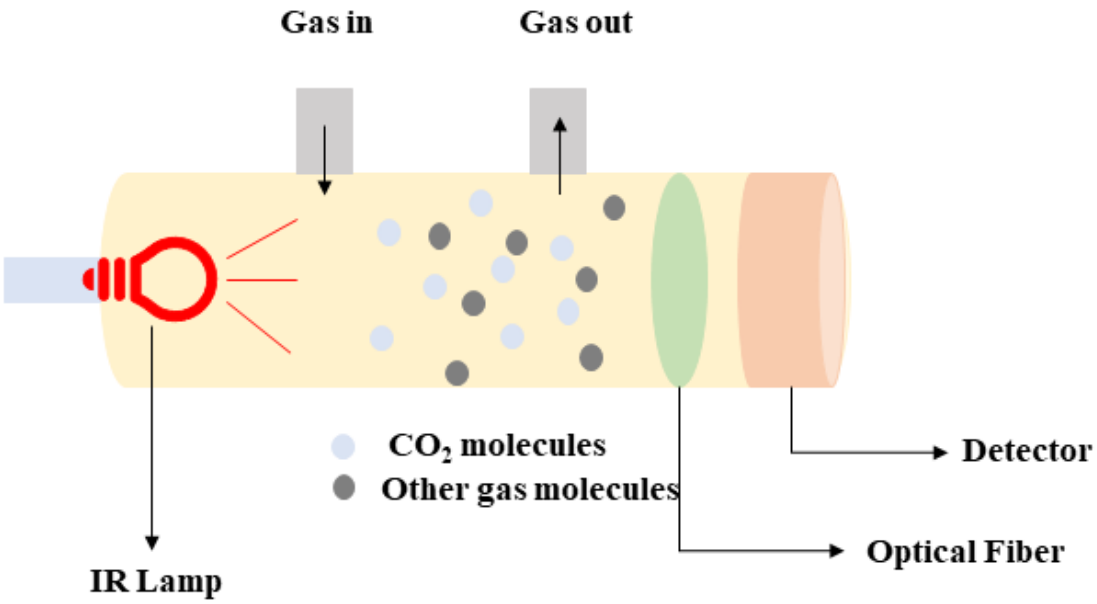


Figure 2.2: Schematic representation of CO_2 analyser working principle (Adapted from [138])

2.3.6 Supervisory Control and Data Acquisition (SCADA) for monitoring and control

Data acquisition hardware (cRIO 9075, National Instruments, Austin, TX, USA) was deployed to acquire all the process variables ($q_C, q_B, q_{baseline}, T_r, T_{j,in}, T_{j,out}, y_{CO_2,in}, y_{CO_2,out}$, CER, pH, and DO) from the fermentation calorimeter (Annexure A1). The data from all sensors were read and processed through LabVIEW 13.0 (National Instruments, Austin, TX, USA), a graphical programming language installed on a standalone PC. Real timesignals for monitoring and control of all process variables and process parameters were logged at every 10 s interval and and real timevisualization of these variables was facilitated by graphical plots.

2.3.7 Development of specific growth rate estimator

2.3.7.1 Instantaneous heat based specific growth rate estimator

The formation of biomass is associated with the coupling mechanism (catabolism and anabolism) and is accompanied by the metabolic heat generation. Therefore, the cellular heat flow is intimately linked to the cell growth process. Based on this notion, the specific growth

rate estimation was carried out by replacing the biomass concentration (X) in the expression for calculating specific growth rate with the biological heat rate/cellular heat flow (q_B) deciphered from the biocalorimeter [134,135,139]. The cellular heat flow measured by the calorimeter is related to specific growth rate based on the following Eq. (2.5)

$$\mu_{est} = \frac{ln\,(q_{b,n} - q_{b,n-1})}{t_n - t_{n-1}} \tag{2.5}$$

Where, $q_{b,n}$ is the biological heat rate at current time point, $q_{b,n-1}$ is the biological heat rate at $n-1$ time point, t_n is current time, t_{n-1} is the time at $n-1$ time point

2.3.7.2 Cumulative heat based specific growth rate estimator

The conversion rates with regard to substrate utilization and biomass production forms the basis for the estimation of yield coefficients related to cell growth. Specific rate for biomass production is calculated from Eq. (2.6).

$$r_X = \mu . X \tag{2.6}$$

Where r_X is biomass production rate and X is the biomass concentration.

The heat yield coefficient due to biomass growth ($Y_{Q/X}$) is fairly consistent for any organism provided stoichiometry of its growth remains constant [140]. $Y_{Q/X}$ is calculated from the slope of the linear plot between cumulative metabolic heat production (Q) and biomass concentration (X). Since, yield coefficient is the ratio of respective specific rates $Y_{Q/X}$ could be related to the specific growth rate (by substituting Eq. (2.6)) as given in Eq. (2.7)

$$Y_{Q/X} = \frac{q_B}{r_X} = \frac{q_B}{\mu\, X} \tag{2.7}$$

The biomass growth profile can be deduced by integrating Eq. (2.6) between the limits as represented here X: $X_0 \rightarrow X_t$ and t: $0 \rightarrow t$,

$$X(t) = X_0 + \frac{Q_t}{Y_{Q/X}} \tag{2.8}$$

Where q_B, biological heat rate (W/L), X_o, biomass concentration at the end of the batch phase (g/L), $Y_{Q/X}$, biomass heat yield coefficient (kJ/g) and Q_t, cumulative heat at that time instant (kJ/L).

Rearranging Eq. (2.6), the specific growth rate can be estimated as a function of metabolic heat rate and cumulative heat, as shown in Eq. (2.9)

$$Y_{Q/X} = \left[\frac{q_B}{\mu\left(X_0 + \frac{Q_t}{Y_{Q/X}}\right)}\right] \quad (2.9)$$

Estimator (Eq. 2.9) could be helpful when the volume changes in a bioreactor are negligible. In the case of fed-batch fermentation, it is essential to account for the variation in the culture volume due to the continuous substrate addition [136]. Hence, Eq. (2.9) requires a modification by incorporating the volume term (V_r).

$$dX.V_r = \frac{q_B}{Y_{Q/X}} dt \quad (2.10)$$

Integrating Eq. (2.10) and rearranging yield a specific growth rate estimator (Eq. (2.11)) for variable volume.

$$\mu_{est} = \frac{q_B}{X_0 V_0 Y_{Q/X} + (Q_t - Q_0)} \quad (2.11)$$

Where V_0 is the volume of reaction broth at the end of batch phase and Q_0 is cumulative heat at the end of batch phase. The modelled μ_{est} was programmed in the LabVIEW utilizing the real-time input of instantaneous (q_B) and cumulative (Q_t) values from SCADA.

2.3.8 *P. pastoris* cultivation (high exothermic process)

Cultivation of glycoengineered *P. pastoris* was carried out in a 5 L lab-scale fermentation calorimeter. In-situ sterilization was performed with pre-filled distilled water at 121°C and 1 bar pressure for 15 minutes. Sterilized BSM medium was transferred to the calorimeter aseptically. A 10% v/v of pre-culture was transferred into the calorimeter aseptically, making the final working volume 2.5 L. Reaction was carried at 30°C, pH was maintained at 5.4 by the addition of ammonia (25% v/v), and orthophosphoric acid (85% v/v) respectively. The dissolved oxygen of the culture broth was maintained above 10% by combining air flow with pure oxygen (total airflow of 1 vvm) and manipulating the agitation rate between 400 to 800 RPM. Antifoam (10% v/v silicone oil) was added manually when required to prevent the formation of excess foaming. The high cell density cultivation of *P. pastoris* was carried out in the following phases: The glycerol growth phase, methanol adaptation phase, and induction phase (production phase). During the growth phase, a significant amount of biomass will be generated with glycerol as a carbon source. The exhaustion of the C-source halts the metabolic activity, as indicated by a drop in the DO signal, heat rate signal, and off-gas activity. Then minimal doses of methanol in pulse were fed as the organism adapted to the new carbon source. A rise in heat rate signal and off-gas activity indicated *P. pastoris* adapted to methanol as a carbon source. The induction/production phase was initiated by continuous feeding of

methanol governed by the substrate balance equation. A predetermined feeding of the substrate in an exponential manner was implemented based on kinetics derived from the previous fed-batch experiments. Samples were collected at regular intervals to estimate biomass, substrate and product concentrations offline.

2.3.9 *S. zooepidemicus* cultivation (low exothermic process)

All the runs were performed in the same 5 L fermentation calorimeter. The production media and glucose were autoclaved separately and transferred aseptically into the sterilized fermenter. Inoculum (10% v/v) was transferred aseptically, making the working volume 2.5 L. The reaction broth was maintained at a temperature of 37°C and a pH of 7.0 using 3 N NaOH, respectively. The DO content of the reactor broth was maintained above 10% by regulating the agitation speed between 400 to 700 RPM. Upon glucose exhaustion (indicated by the drop in the heat rate signal), the fed batch was initiated by continuous glucose feeding. A predetermined feeding rate based on the kinetics derived from the batch experiments was fed exponentially. Samples were collected regularly to determine biomass, product, and substrate concentrations.

2.3.10 Offline Analysis

2.3.10.1 Determination of dry cell weight, huIFN α2b and substrate concentration

Dry cell weight (DCW) measurement was carried out by aliquoting 1 mL of sample in a centrifuge tube and it was centrifuged at 10,000 RPM for 10 min. The pellets were washed twiced with distilled water to remove the salts present and dried overnight at 80°C. Optical density measurements at 600 nm (OD_{600}) using a UV-visible spectrophotometer (GE Healthcare, UK) were carried out to measure the cell growth. Previously developed OD vs DCW correlation (1 OD = 0.21*DCW) was used to convert the absorbance values into DCW (in g/L). Further, the supernatant was collected in a separate vial, stored at -20ºC for C-substrate consumption and huIFN α2b formation analysis. The concentrations of C-substrates (glycerol and methanol) were measured by using HPLC employing reverse phase analytical chromatography (Rezex RHM- monosaccharide H+ Column, Phenomenex Inc, California, USA). Column operating conditions including parameters such as mobile phase, flow rate and temperature were maintained as reported in our previous study [126]. The concentration of huIFN α2b in the collected samples was determined by performing Enzyme linked immunosorbent assay (ELISA) analysis using huIFN α2b ELISA kit (Mabtech AB, Sweden).

2.3.10.2 Determination of dry cell weight, HA and substrate concentration

S. zooepidemicus is known to produce HA a capsular product, the culture samples were treated with an equal volume of 0.1 % (v/v) Sodium Dodecyl Sulphate (SDS) to strip off the HA from the cellular membrane. The cells were pelleted down at 10,000 rpm for 10 min. Then pellets were washed with 0.9% NaCl solution and the cell growth was determined by measuring the absorbance at 600 nm (OD600) using UV-visible spectrophotometer (GE Healthcare, UK). Absorbance was converted into DCW (g/L) using a previously developed OD vs DCW relationship (1 OD = 0.58*DCW). CTAB (Cetyl-trimethyl-ammonium-bromide) turbidity assay was performed to the determine the concentration of HA in the collected samples. HA titers were determined from the previously developed standard plot of pure HA (1 OD = 0.164*HA). Glucose concentration was determined by GOD-POD assay and measuring the absorbance at 546 nm. The absorbance values were converted into glucose concentrations using a developed correlation (1 OD = 0.16*GLU) [141].

2.3.11 Data smoothing of specific growth values from the estimator models

The high frequency of input signal is prone to noise which leads to significant fluctuations in the estimated values. The noise emanating from the metabolic heat rate signal was treated using various data smoothing filters [139]. To evaluate the performance of various filters Moving average (MA), Savitzky Golay (SG), Butterworth and Chebyshev were applied to the previous experimental data. Based on the signal-to-noise ratio (Annexure A2), two filtering methods MA, SG were selected for the experimental work carried out in this study. The data loss was negligible and the loss of shape in the real-time plots wasn't observed in MA, SG due to high SNR compared to other filtering methods.

2.3.12 Feeding strategy

To assess the performance of the specific growth rate estimators, the feeding of carbon (limiting) substrates in both *P. pastoris* and *S. zooepidemicus* cultivation was not coupled with feedback control. Simple, feedforward control using a predetermined feeding rate of the substrate was employed for all experimental runs. The specific growth rate set points (μ_{sp}) of 0.035 h^{-1} and 0.1 h^{-1} for *P. pastoris* and *S. zooepidemicus* were used, respectively. These optimal setpoints were selected from the published literature [141,142]. The predetermined feeding rate was based on Eqn. (2.12)

$$F(t) = F_0 e^{\mu_{sp} t} \tag{2.12}$$

where, F_0 is determined from the previous batch data kinetics for *S. zooepidemicus* cultivation and induction phase kinetics for *P. pastoris* cultivation (Eq. 2.13).

$$F_0 = \frac{X_0 V_0 \mu_{sp}}{Y_{X/S}.\ S_0} \tag{2.13}$$

$F(t)$ is the feed rate (mL/h) at the time t, F_0 (mL/h) is the initial feed flow rate, X_0 (g/L) is the biomass concentration at the start of the fed-batch phase, $Y_{X/S}$ (g biomass/g substrate) is the batch phase biomass yield on the substrate, V_0 (L) is the culture volume at the start of the fed-batch phase, and S_0 (g/L) is the inlet concentration of the carbon substrate. Feeding of the limiting substrate was regulated by a high-precision pump (Watson Marlow 120U/DV, Falmouth, Cornwall, U.K).

2.3.13 Performance indices for specific growth rate estimators

The developed model-based estimators were evaluated through the performance indices root mean square error (RMSE) and mean absolute error (MAE), respectively. RMSE is calculated according to the Eq. (2.14).

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(y_{\text{treated}} - y_{\text{actual}})^2}{N}} \tag{2.14}$$

Where y_{treatted} is the online filtered value of the specific growth rate and y_{actual} is the raw online signal of the specific growth rate, and N is the total number of observations.

MAE is calculated according to the Eq. (2.15)

$$MAE = \frac{\sum_{i=1}^{N} |y_i - x_i|}{N} \tag{2.15}$$

Where y_i are the online filtered values of the specific growth rate, x_i are the raw online signal values of the specific growth rate, and N is the number of observations.

2.4 Results and Discussion

2.4.1 Real-time monitoring of *P. pastoris* and *S. zooepidemicus* growth

Real-time dynamic profiles emanated from online process analysers viz., fermentation calorimeter and CO_2 analyser are depicted for *P. pastoris* and *S. zooepidemicus* cultivation (Fig. 2.3(A), 2.3(B), 2.3(C) and 2.3(D)), respectively. Real-time signals, viz., the heat rate and CER profiles elucidated well the different phases of the growth in *P. pastoris* and *S. zooepidemicus* cultivation process. As biomass production increased, the corresponding metabolic heat rate generated and CER during the cultivation increased during the growth

phase (*P. pastoris* cultivation). The sudden fall in thermogram and respirogram (≈ 22-24 h in Fig. 2.3(A), 2.3(B); ≈ 7-8 h in Fig. 2.3(C), 2.3(D)) was attributed to the exhaustion of the limiting substrates, glycerol (*P. pastoris*) and glucose (*S. zooepidemicus*), respectively. A significant drop-in metabolic heat rate signal was observed during the end of batch phase, nearly 50% decrement of the peak heat rate value (15 W/L to 9 W/L) in case of *P. pastoris* and from 2.5 W/L to 1 W/L for *S. zooepidemicus* owing to exhaustion of their limiting substrates. During induction phase, *P. pastoris* adaptation with the induction substrate (methanol) resulted in steady rise in the metabolic heat rate in proportion with increase in biomass concentration (corroborated well with the offline biomass conc. values). *S. zooepidemicus* fermentation in the fed-batch phase, it could be observed that the heat rate signal was sustained around 1.5 to 2 W/L (Fig. 2.3(C) and 2.3(D)) depicting that cells were metabolically active, and in contrast, the CER signals were dropped from 15 m.mol/L.h to 3 m.mol/L.h (Fig. 2.3(C) and 2.3(D)). The drop in the CER signal could be ascribed to the fact that *S. zooepidemicus* is a facultative anaerobe, and a similar trend was observed [143]. Interestingly, the peaks generated from the heat rate signal indicated that cells were continuously utilizing the substrate for their anabolic activities, i.e., biomass and hyaluronic acid production. The sustained heat signal is due to the biopolymer production (HA) The ability of the fermentation calorimeter to detect the sensitive metabolic heat rate signal, aids in unravelling the metabolism of the organism on real-time which proves the robustness and versatility of the fermentation calorimeter as a monitoring tool for the PAT framework.

2.4.1 Influence of data smoothing on estimator signal

The raw signal data emanating from the sensor are generally associated with noise, leading to inaccuracies in the measured/estimated data. Consequently, this phenomenon impedes the real-time tracking of metabolic activity, and the process could be subjected to severe instabilities by prompting improper control action. In the present study, data smoothing filters were deployed for an appropriate real-time estimation of specific growth rate values. The selection criteria for data smoothing filters were based on factors such as preserving the shape of the underlying signal, high signal-to-noise ratio, the frequency content of the signal, and computation efficiency [144] . Hence, the screening of MA, SG, Butterworth, and Chebyshev filters was performed using the previous experimental data. Based on the aforementioned selection criteria, MA and SG were observed to be suitable and used in this study to smoothen the data (Annexure A1).

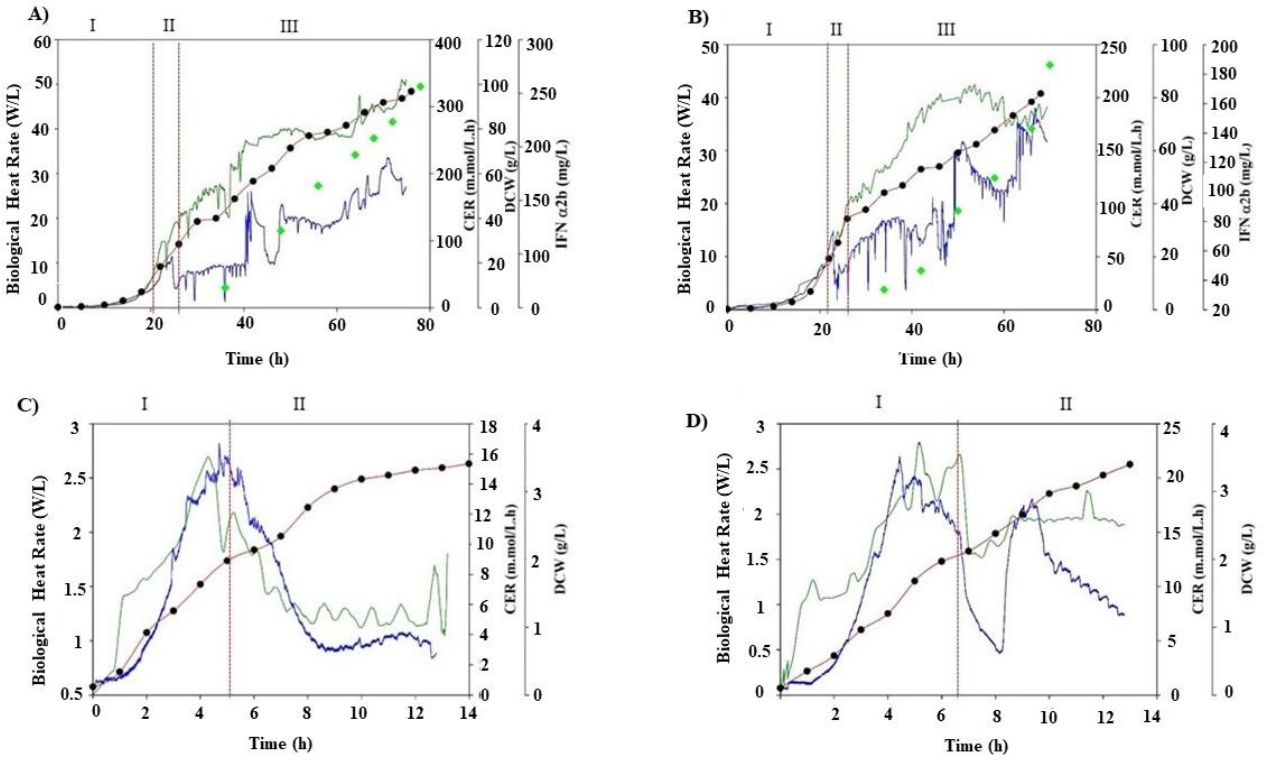


Figure 2.3: Real-time dynamic profiles depicting the heat flow rate from (green continuous), CER (blue continuous), DCW (black circle) and huIFN α2b (green diamonds) for *P. pastoris* cultivation. Phases I, II, III indicate glycerol growth phase, methanol adaptation phase and induction phase respectively. A) Cumulative heat-based specific growth rate estimator run, B) Instantaneous heat rate based specific growth rate estimator run. Real-time dynamic profiles depicting the heat flow rate (green continuous), CER (blue continuous), DCW (black circle) for *S. zooepidemicus* run. Phases I and II represent the batch and fed-batch phases respectively. C) Cumulative heat-based specific growth rate estimator run, D) Instantaneous heat rate based specific growth rate estimator run.

Due to the high frequency of the input signal, a smaller window size (MA 100 data points, SG 51 data points) was selected to preserve the underlying trends displayed by the raw signal. Moreover, such a frequent signal shouldn't be affected by the lag, and for this purpose, a lower order (2) filter pass was implemented in the SG filter. The SG filter of frame length (data points) of 51 and an order of two resulted in effective performance (Fig. 2.3) with a signal-to-noise ratio of 2.2 for *S. zooepidemicus* fermentation. Whereas, the signal-to-noise ratio for MA (100 data points) emerged out to be 1.8. For *P. pastoris* cultivation the application of data smoothing filters on cumulative heat-based estimator resulted in a RMSE of less than 0.004 and MAE of less than 0.0048 h^{-1} respectively. It is evident that the instantaneous heat rate estimator resulted

in large positive and negative values of μ_{est} in *P. pastoris* cultivation (Fig. 2.4(C) and 2.4(D)) which is discussed elaborately in "assessment of specific growth estimators for high exothermic process" (section 2.4.2). Due to the higher SNR once again, the SG filter outperformed the classical MA filter, with lower values of RMSE and MAE being recorded (Table 2.2) for both estimator models. The propensity to attenuate a substantial portion of high frequency signals combined with the noise makes SG filter a suitable choice for data smoothing of high frequency signals. The underlying trend of the data was captured more precisely (Fig. 2.4(B) and 2.4(D)) in the SG filter owing to the fact that SG employs higher-order polynomial regression to smoothen the data. Thus, the main goal of data smoothing was to capture the actual values of the specific growth rates estimated from the production of metabolic heat from growing cultures during the fed-batch phase. Ultimately, the estimated specific growth values based on the heat based (instantaneous and cumulative) μ estimators can be used as appropriate input values for regulating the substrate feed during the fed-batch operation.

2.4.2 Assessment of specific growth rate estimators for *P. pastoris* cultivations (high exothermic process)

Gibb's free energy dissipation from the coupling of anabolism and catabolism dictates the driving force of cell growth among the various classes of microorganisms (aerobic, anaerobic, methanogenesis) [145]. Hence, excess Gibb's energy dissipated as heat energy is intimately linked with the cellular metabolism and real-time measurement of metabolic heat rate (quantitative calorimetry) could fingerprint ongoing physiological activity including cell growth. Various configurations of high sensitive biocalorimeters were successfully employed to measure the metabolic heat rate [146]. The indigenous heat compensation fermentation calorimeter used in this study can detect microbial heat rapidly (10 to 15 s) faster than other heat flux calorimeters [141]. Both cumulative and instantaneous heat rate based μ estimators were tested on *P. pastoris* cultivation, capable of generating the biological heat rate of 35 to 60 W/L during high cell density cultivation in a 5 L lab scale fermentation calorimeter [147]. As *P. pastoris* is an obligate aerobe and hence, the Gibb's free energy ($\Delta_r G_X^0$, driving force for growth) of its growth process almost entirely dissipated as enthalpy change exhibiting a high metabolic heat rate [115], Moreover, *P. pastoris* is a methylotrophic yeast, feeding a limiting substrate (methanol) with high degree of reduction ($\gamma_s = 6$) also contributes to increase in enthalpy change.

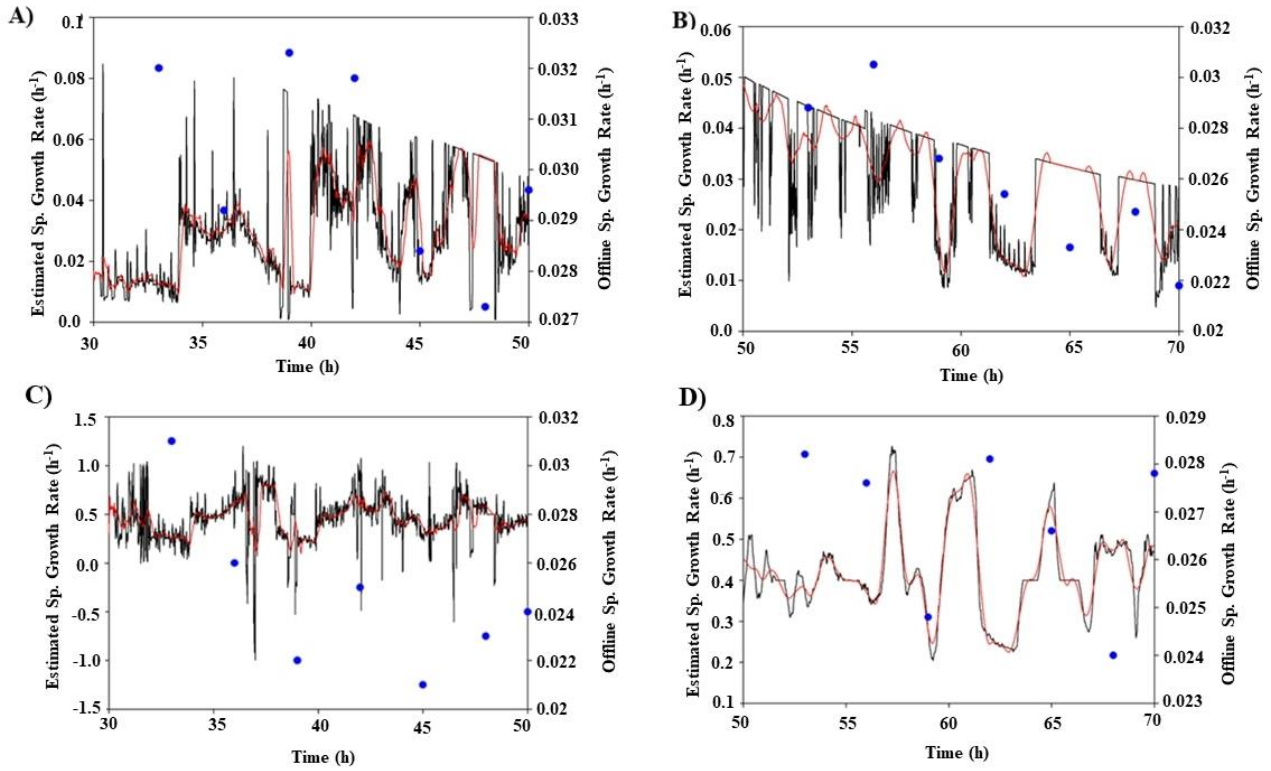


Figure 2.4: Influence of data smoothing on real-time signals of specific growth rate values based on cumulative heat for *P. pastoris* A) Moving average with 100 data points raw signal from the estimator (black continuous), smoothened signal (red continuous), and offline values (blue circles). B) Savitzky Golay filter with 51 data points raw signal from the estimator (black continuous), smoothened signal (red continuous), and offline values (blue circles). Influence of data smoothing on real-time signals of specific growth rate values based on instantaneous heat rate for *P. pastoris* C) Moving average with 100 data points raw signal from the estimator (black continuous), smoothened signal (red continuous), and offline values (blue circles), D) Savitzky Golay filter with 51 data points raw signal from the estimator (black continuous), smoothened signal (red continuous), and offline values (blue circles).

But a compromise has to emerge between $\Delta_r G_X^0$, (i.e., μ) and biomass yield ($Y_{X/S}$) during high cell density cultivation in accordance with the energy transducer model. In conformity with this model, Gibb's energy dissipation is higher at slower growth rates and vice versa (Fig. 2.5) [140]. It is well known that under the tight regulation of the AOX1 promoter in *P. pastoris* with a Mut^+ phenotype is bound to grow at very low growth rates during the induction phase facilitating high cell density cultivation. Hence, the glycoengineered strain used in this study exhibited a maximum specific growth rate of 0.062 h^{-1} in induction phase [126]. Persistent assimilation of methanol by *P. pastoris* requires a perfect balance between substrate uptake

rate and biomass production. However, continuous feeding of methanol is perilous (accumulation of residual methanol) to cellular metabolism; therefore, a tight regulated feeding of methanol is necessary. The induction phase was carried out based on a predetermined feeding starting with 8 mL/h of methanol, and was progressed to 32 to 35 mL/h of methanol in an attempt to cultivate the cells at the desired μ ($\mu_{sp} = 0.035\ h^{-1}$). To ascertain the performance of developed specific growth rate estimators, the models were not integrated with any robust controls (PID or adaptive feedback control) for the regulated feeding of methanol. Initially, the *P. pastoris* cultivation was carried out to evaluate the estimator model based on cumulative heat rate as described by Eq. (2.11) and the estimator predicted μ values in the range of 0.02 to 0.05 h^{-1}, which is corroborated well with the offline μ values ranged from 0.035 to 0.01 h^{-1}(Fig. 2.4(A) and 2.4(B)). The driving force for growth ($\Delta_r G_X^0$) decreases with increase in biomass yield $Y_{X/S}$ value (energy transducer model) and therefore, the associated enthalpy change also decreases. Consequently, slower growth rate is preferred for high cell density cultivation of *P. pastoris* leading to enhanced huIFN α2b yield. Hence, robust control of μ at a low value in a narrow range (0.03 – 0.04 h^{-1}) would be advantageous in huIFN α2b production perspective, but a serious limitation on calorimetric monitoring perspective (because low enthalpy change). In analogous to $Y_{X/S}$, $Y_{Q/X}$ also establishes the basis for the coupling between anabolism and catabolism as per bioenergetics perspective. Moreover, under balanced growth conditions the heat energy released per unit biomass ($Y_{Q/X}$) remains highly consistent for almost all organisms [115]. Hence, the presence of biomass heat yield coefficient ($Y_{Q/X}$) term in cumulative heat-based estimator played a vital role in the reliable estimation of specific growth rate at low range. Despite slower growth rate observed during high cell density cultivation, the measured enthalpy change (heat rate) was significant (highly exothermic) owing to aerobic metabolism of *P. pastoris* utilizing reduced substrate (methanol). In contrast, the application of instantaneous heat rate-based estimator resulted in prediction of μ values ranging from 0.4 to 0.6 h^{-1} exhibiting approximately 10-fold over prediction from the actual offline μ values (from 0.035 to 0.02 h^{-1}). Such a large discrepancy (Fig. 2.4(C) and 2.4(D)) between the estimated and offline μ values could be attributed solely to the estimator model relying on sensitive heat rate signal manifesting the instantaneous metabolic activity as well the process related perturbations. Usually, the high cell density cultivation of *P. pastoris* is accompanied by oxygen limitation and this operational challenge is surmounted by addition of pure oxygen to the air stream to maintain the critical dissolved oxygen level in the culture broth.

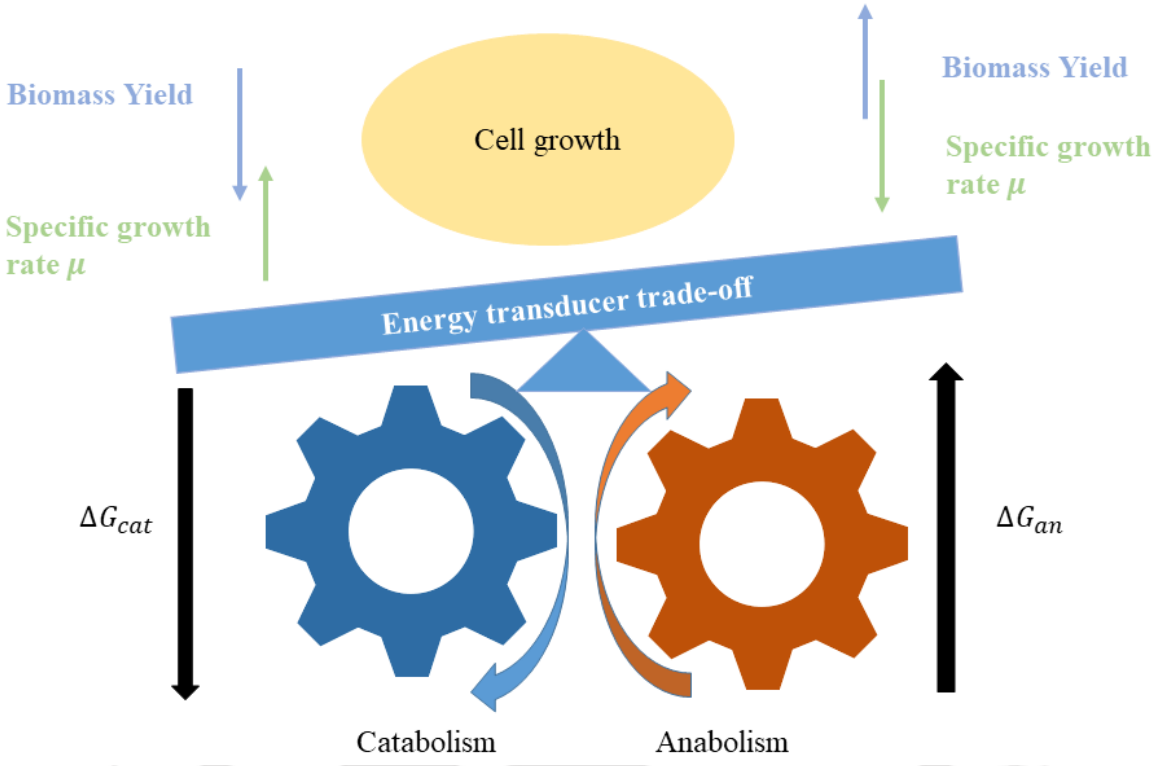


Figure 2.5: Schematic representation of energy transducer trade-off for cell growth

The supply of oxygen during methanol feeding leads to high metabolic heat rate due to higher enthalpy of O_2 utilization (830 kJ/mol of O_2) in addition to contribution by higher of enthalpy of combustion for methanol (727 kJ/C mol) [148]. Hence, the resultant instantaneous heat rate signal exhibited oscillatory behaviour fluctuating between two extremes (high positive and negative heat rate values) owing to momentary exhaustion/addition of methanol/dissolved during the course of high cell density cultivation of *P. pastoris*. Hence, instantaneous estimator yielded poor μ estimates making it highly unsuitable for employing it as soft-sensor for bioprocess monitoring applications. It is evident that application of data smoothing techniques was also not helpful to address the perturbations in the measured signal and extensive filtering of signal would lead to loss of vital process information.

2.4.3 Evaluation of specific growth estimators for *S. zooepidemicus* fermentation (low exothermic)

Considering the sensitivity posed by the heat rate signals, the fermentation calorimeters must be rendered with higher resolution to detect the smaller heat generated by the biological systems [149]. Fermentation calorimeter employed in this present study is encompassed with a high resolution of 6.73 mW/L, which aids in detecting lower metabolic heat. Predetermined glucose feeding was initiated at the end of the batch phase based on the kinetics derived from

the batch phase. Feed rate of 6 mL/h was obtained in order to achieve the desired specific growth rate value ($\mu_{sp} = 0.1\ h^{-1}$). As depicted in the Fig. (2.6), specific growth rate was maintained in the range of 0.05 to 0.12 h^{-1} facilitated by continuous feeding of glucose in the fed-batch phase. Due to the weak metabolic heat generation by *S. zooepidemicus* (2.5 to 3 W/L), the quantitative value of the cumulative heat term in the Eqn. (2.11) has little bearing on the estimator model. It is clearly evident from Fig. (2.6(A) and 2.6(B)) that the predicted specific growth rate values based on the cumulative heat estimator were in the range of offline values (0.04 to 0.09 h^{-1}). The biomass heat yield coefficient ($Y_{Q/X}$) is driven by two other important parameters, i.e., biomass yield on substrate and synthesis of fermentation products in the case of anaerobic systems [150]. Pertaining to the above facts the $Y_{Q/X}$ value is lower in the case of *S. zooepidemicus* fermentation 9.26 ± 0.18 kJ/g of biomass compared to that of 18.23 ± 0.21 kJ/g of biomass produced in *P. pastoris*. Since, *S. zooepidemicus* is a facultative anaerobe, the carbon flux is channelized towards the production of lactic acid (LA) apart from the production of biomass and HA. Low biomass yield and heat of neutralization associated with the production of LA resulted in enthalpy change although the growth process is under facultative anaerobic metabolism. A lower biomass yield generally leads to higher heat dissipation since more energy available in the substrate is converted into thermal energy leading to inefficient growth. The estimator model based on instantaneous heat predicted the specific growth rate values ranging from 0.05 to 0.12 h^{-1}. The offline determined μ values were in the range of 0.11 to 0.065 h^{-1} (Fig. 2.6(C) and 2.6(D)) and the deviation from the estimated values measured in terms of RMSE, MAE was higher compared to that of cumulative heat-based estimator (Table 2.2). The enthalpy change observed in HA fermentation was steady without any significant fluctuations, which could be linked to the weak exothermicity of facultative anaerobic metabolism. In contrast, the respiratory metabolism of *P. pastoris* is very sensitive to dissolved oxygen availability exhibiting frequent metabolic perturbations i.e., highly sensitive fluctuating heat rate signal limiting the performance of instantaneous estimator.

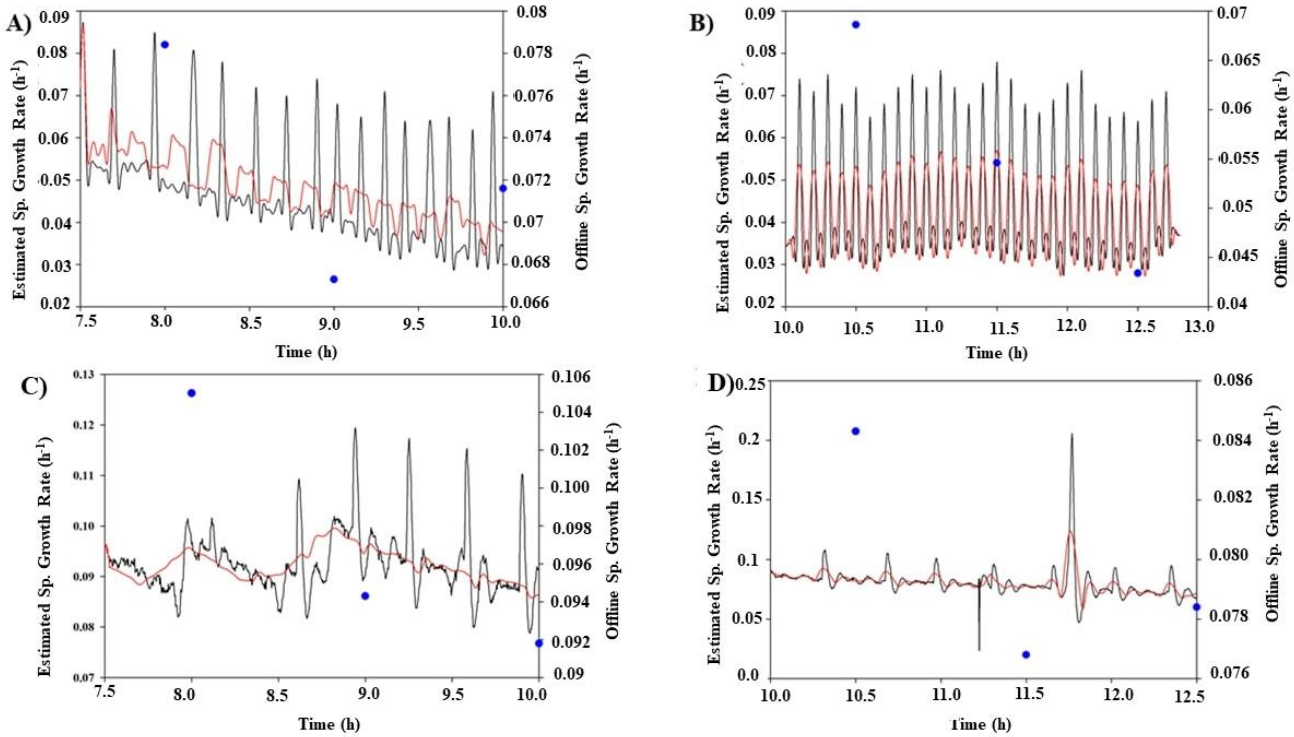


Figure 2.6: Effect of data smoothing on real-time signals of specific growth rate values based on cumulative heat for *S. zooepidemicus* A) Moving average with 100 data points raw signal from the estimator (black continuous), smoothened signal (red continuous), and offline values (blue circles), B) Savitzky Golay filter with 51 data points raw signal from the estimator (black continuous), smoothened signal (red continuous), and offline values (blue circles). Effect of data smoothing on real-time signals of specific growth rate values based on instantaneous heat rate for *S. zooepidemicus* C) Moving average with 100 data points raw signal from the estimator (black continuous), smoothened signal (red continuous), and offline values (blue circles), D) Savitzky Golay filter with 51 data points raw signal from the estimator (black continuous), smoothened signal (red continuous), and offline values (blue circles).

2.4.4 Comparative assessment of estimators

It is evident from Table 2.2, the predictions resulted from the cumulative heat-based estimator were reliable on both high and low exothermic process. Additionally, the estimator could predict reliably the μ values for a longer duration (> 40 h in the case of high exothermic process) (Fig. 2.4) amidst the presence of process related perturbations. This attribute facilitates the robust monitoring and control of μ especially for high cell density cultivations as they proceed with lower μ values. Contrastingly, the instantaneous metabolic heat rate estimator resulted in very poor predictions for high exothermic process.

Table 2.2: Comparative assessment of estimator models on high and low exothermic process

S. No	Estimator Model	Process	RMSE (h^{-1})	MAE (h^{-1})	Average offline μ values (h^{-1})
1	Cumulative heat based	*P. pastoris* (High exothermic)	0.0034	0.0046	0.028±0.005
2	Cumulative heat based	*S. zooepidemicus* (Low exothermic)	0.036	0.045	0.08±0.02
3	Instantaneous metabolic heat rate based	*P. pastoris* (High exothermic)	6.72	7.21	0.025±0.008
4	Instantaneous metabolic heat rate based	*S. zooepidemicus* (Low exothermic)	0.059	0.084	0.074±0.02

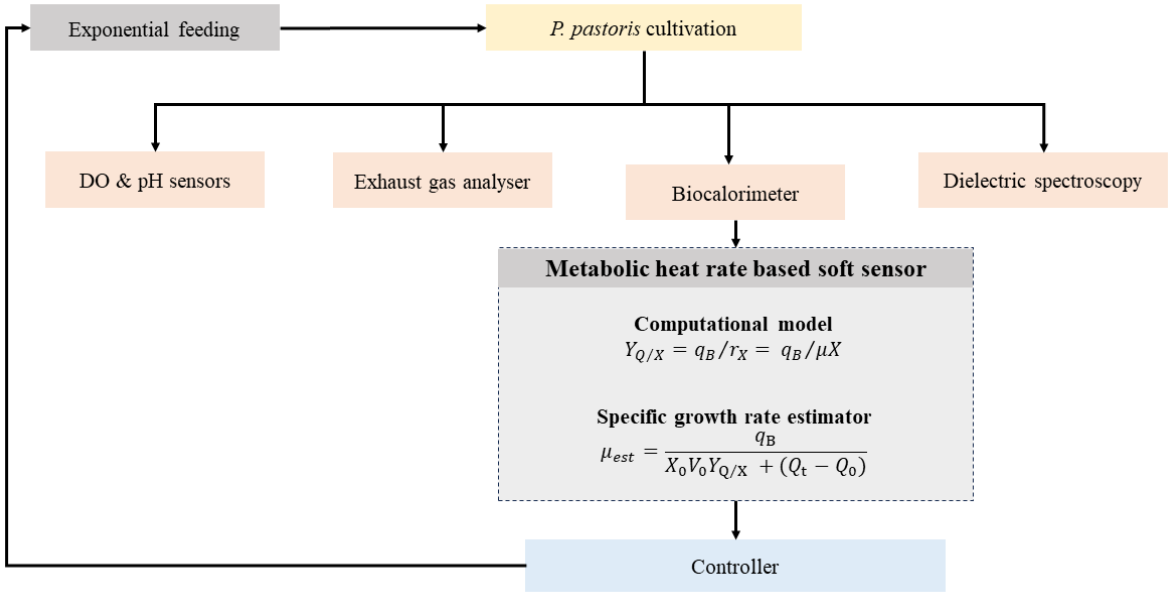
The instantaneous metabolic heat rate is highly sensitive to the metabolic changes combined and process related perturbations which in turn yielded unrealistic μ values and ultimately propelling RMSE and MAE to higher values (6.72 and 7.21 h^{-1}), respectively. Consequently, the application of instantaneous metabolic heat rate based μ estimator is a not suitable as it relies on the momentary metabolic changes leading to rampant oscillations in the estimation of μ values. On the other hand, cumulative heat based μ estimator demonstrated its versatility by estimating reliable μ values on both processes thus proving to be an exemplary μ estimator model for monitoring and control of bioprocess.

2.5 Conclusions

The present investigation explored the potential of bioenergetic models for the real-time prediction of one of the key bioprocess state variables (specific growth rate). This is the first-of-its-kind report to enumerate the right choice of μ estimator model that needs to be implemented for real-time bioprocess monitoring from a calorimetric perspective. The estimator model based on the cumulative heat accurately estimated the μ values for high and low exothermic process. Data smoothing of estimator values did not influence the instantaneous heat rate-based estimator for the high exothermic system, leading to a more than 100-fold increase in performance indices. The versatility and accuracy of cumulative heat-based estimator could be deciphered from the thermodynamic perspective, with the focal point

being biomass heat yield coefficient which accounts for energy dissipation from anabolic payload (i.e., formation of biomass) activities. The research outcome of this chapter is highly crucial for a careful selection of a specific growth rate estimator model depending on the exothermicity and the scale of cultivation, which in turn serves as a reliable input to implement advanced control strategies in fed-batch fermentation and ultimately enhancing the yield, productivity, and quality of the product. As seen earlier in the results and discussion section (section 2.4.5) due to reliable performance of the cumulative heat-based estimator model is selected for real-time estimation of μ in the next chapter and control strategies will be implemented to investigate the influence of μ on huIFN α2b productivity.

Chapter 3 Deployment of metabolic heat rate-based soft sensor for estimation and control of specific growth rate in glycoengineered *Pichia pastoris* for human interferon alpha 2b production



3.1 Abstract

Lack of appropriate process models, reliable online sensors, and process variability in bioprocess systems are poising challenges in real-time monitoring and control of CPPs. This present chapter deals with the utilization of a non-invasive soft sensor emanating from fermentation calorimeter where metabolic heat rate acts as an input signal for online estimation of specific growth rate (μ_{est}) during the induction phase of glycoengineered *P. pastoris* for huIFN α2b production. The feedforward strategy employing a predetermined exponential feeding of methanol during the induction phase was dealt with at defined setpoint values (μ_{SP}). Standard PID controller with predetermined gain values regulated methanol feeding in accordance with the deviation from the μ_{SP} value. An adaptive PID (gain scheduling) significantly minimized the deviation of μ from its μ_{SP} value reduced the amplitude of oscillation, and achieved long-term controller stability. Robust control of methanol feeding by adaptive PID resulted in a 1.5 and 2.2-fold increase in productivity of huIFN α2b compared to standard PID and feedforward controls respectively. Moreover, adaptive PID control facilitated narrow range control of μ for longer durations (> 20 h) with a low average tracking error (< 6 %) enumerating its scope of application in therapeutic protein production in the near future.

3.2 Introduction

Interferon α2b (165 Amino acid residues & 19 kDa Molecular weight) is the most prominent cytokine molecule, interacts with the specific IFN α/β receptors to instigate JAK-STAT mediated signalling pathway to activate immunomodulatory, antiproliferative and antiviral proteins. Commercially known as Intron®A, they are approved by the FDA for the treatment of chronic Hepatitis-C, Hepatitis-B, Hairy cell leukaemia, Malignant melanoma, Follicular lymphoma, Condyloma acuminata, and AIDS-related Kaposi sarcoma. More recently, the demand has surged for IFN α2b due to its usage along with other antiretroviral and antimalarial drugs in treating SARS-CoV-2 (COVID-19) disease [151]. Upon approval by regulatory authorities, biopharmaceutical brands go expired after 20 years from filing, further allowing competitive manufacture of IFN α2b biosimilars by other industrial players. Economically manufacturing generic drugs is vital for their sale which turns on the concerns about process intensification. But also laying emphasis on selecting the most suitable, stable expression hosts is equally important. The microbial route of synthesis of therapeutic proteins is widely acknowledged and the safest route for therapeutic protein production [152].

Pichia pastoris is advantageous in terms of rendering effective post-translational modifications similar to the mammalian host, showing exceptional cultural stability and synthesizing proteins in larger quantities. Also, unlike mammalian cell lines lower nutritional costs and less likely to be contaminated by adventitious agents are beneficial to the biopharma industry for the low-cost biologics production. Most favoured for its role and serving as a major industrial workhorse for biomanufacturing, *P. pastoris* has garnered widespread attention among academic researchers and scientists [153]. Glycosylated IFN α2b increases the half-life of the drug in the bloodstream (18-22 h) and also aids in its removal from the renal system following its function. N-Glycosylation of IFN α2b starts at the endoplasmic reticulum and matures in the Golgi complex (addition of mannose residues), its attachment to the native protein is known to influence the affinity of the protein molecule to its receptor. In our previous study, we achieved cloning IFN α2b gene with N-Glycosylation site in the pPICZα plasmid backbone, followed by homologous recombination into *P. pastoris* genome to exhibit the synthesis of humanized IFN α2b bearing N-Glycan $Man_5GlcNac_2$ attached to 104 Asparagine residue [21]. Complexity in the expression of recombinant proteins is attributed to the combinatorial effect of different factors at biochemical, metabolic, and physiological levels altogether [154]. The process variables viz., pH, temperature, and dissolved oxygen concentration were reported to have minimal impact on productivity and product quality in glycoengineered yeast [155].

Significant impact at the metabolic level can be brought about by altering the methanol feed at different rates [155–157]. Monitoring processes by obtaining information in real-time about estimated process variables, viz., specific growth rate, and productivity, can be helpful to track and course corrections can be taken at failures. PAT lays a common framework to establish a correlation between process parameters and product quality. Insights of metabolic intricacies can be obtained by effectively monitoring the changes in the trend of process parameters [23]. Imparting quality checks at various stages of a process and enabling consistent product quality forms the basis of the QbD approach [23].

Specific growth rate, μ was well known to influence the physiological state of the culture, growth of biomass, rate of product synthesis, and product quality [35]. It is identified to be a CPP and its regulation was reported to enhance product titer and quality. However, acquiring real-time values of μ would be difficult to achieve as no direct sensors are available to measure [84]. The first step in controlling the specific growth rate involves devising a reliable estimator that provides reliable biomass concentration or other factors that can be correlated to μ. Estimating biomass concentrations in real-time is accomplished through a synergistic effect of acquiring data from hard-type sensors (physical sensors) and passing them into a suitable computational model that deciphers the growth rate. Soft sensors utilize this combinatorial approach for real-time estimation of bioprocess state variables. This soft sensor approach toward indirect estimation of biomass concentration is an existing methodology and has been widely adopted from published literature [158].

In-situ monitoring techniques such as dielectric spectroscopy [159], optical density probe [160], and fluorescence spectroscopy [125] were used for the online estimation of biomass concentration and subsequently used for the estimation of the specific growth rate by correlating with a mechanistic model [160]. However, morphological changes in cell populations could lead to distorted correlations [158]. Developing a generalized correlation between the probe response and cell concentration considering process variability comes at the cost of estimation accuracy. To address the shortcomings associated with in-situ monitoring, a soft sensor for estimating specific growth rates from off-gas activity was reported [127,161]. However, changes in aeration and agitation rates to maintain critical dissolved oxygen concentrations (DO) could increase the noise in the measured DO signal, prompting fluctuation in off-gas measurements [33]. Due to these instabilities in online signals, estimators will disseminate erroneous values leading to inadequate control of μ. A non-invasive, instantaneous

estimator is required to surmount this challenge for tight control of specific growth rates in fed-batch cultivations.

The second law of thermodynamics states that microbial growth is a spontaneous reaction, and a fraction of available internal energy in nutrients is dissipated as heat to sustain the metabolism. Quantifying metabolic heat rate (biocalorimeter) provides valuable inputs on biological reactions [115]. In a recent study, the growth dynamics of *P. pastoris* cultivation and change in metabolism based on different carbon substrates were fingerprinted successfully using a fermentation calorimeter [147]. The design of a specific growth rate estimator, μ_{est} using the measured calorimetric signal for monitoring and controlling fed-batch cultivation was already reported [135,162]. Furthermore, the control of μ for Crabtree–negative yeasts was performed using the commercial version of the reaction calorimeter, BioRC1 [136]. However, to date, there is no significant reported study illustrating the potential of calorimetric signals for monitoring and controlling bioprocess, leading to high value-added product synthesis (e.g., biotherapeutics).

The outcome of the previous chapter (section 2.5), provided information about the reliable estimation of the designated CPP i.e., μ for microbial systems that are associated with higher metabolic heat rate generation. With an appropriate process analyser in place for reliable estimation of μ, the controlling of CPP can be applied to enhance productivity and product quality [158,159]. Conventional feedforward controllers are well known to achieve higher cell densities at a predetermined feed rate. However, they lack robust control of μ in the absence of online μ measurements [84]. Additionally implementing a PID control strategy to control the specific growth rate through regulating substrate feed rate has been performed for various microbial systems, *viz*., bacteria [87], and yeast [163]. Although regular PID control offers an advantage in terms of ease of implementation in real-time, maintaining CPPs at the desired μ_{SP} value becomes problematic in the presence of external disturbances, high noise in process signal, and nonlinearity of the process. To resolve this issue, various researchers have attempted different control strategies such as model reference adaptive control [102], model predictive control [164], and Fuzzy control [165]. Nevertheless, implementing these model-based control strategies requires deep knowledge of the process model, or else it will lead to plant model mismatch, and difficulty in process implementation (computation efforts) impedes the extensive usage of model predictive control [84]. An adaptive control strategy (gain scheduling PID) that is relatively easier to implement in real-time will address the problem of

nonlinearity in the process. The gain scheduling-based adaptive control was deployed to control dynamic parameters such as dissolved oxygen [166] and pH [167] were reported.

This present chapter addresses the scope of application of the specific growth rate estimator (μ_{est}) designed based on real-time metabolic heat rate signal for monitoring and controlling huIFN α2b production by *P. pastoris*. Moreover, this chapter also addressed maintaining the specific growth rate in narrow range value (0.03 – 0.04 h^{-1}) by implementing feedforward control, simple PID control, and adaptive PID control using the gain scheduling technique. A comparative assessment of the effect of μ control using three various control strategies influencing the huIFN α2b titer was dealt effectively in the present chapter.

3.3 Material and Methods

3.3.1 Strain

A humanized glycoengineered *P. pastoris* strain (Glycoswitch ® SuperMan5 pep4Δ prb1Δ obtained from Biogrammatics Inc., USA) of Mut^+ phenotype expressing huIFN α2b was used in this work. A previous study described the clone constructs, huIFN α2b productivity, and characterization of the purified product [21]. Glycerol stocks were maintained at -80 ˚C in yeast peptone dextrose (YPD) media containing 20% (v/v) glycerol. The compositions of various media used here are as follows:

- Starter culture medium (Yeast extract Peptone Dextrose, YPD), g/L: yeast extract 10; peptone 20; dextrose 20
- Pre-culture medium (Yeast extract Peptone Glycerol, YPG), g/L: yeast extract 10; peptone 20; glycerol 20.
- Optimized fermentation medium (Basal Salt Medium, BSM) [168]: glycerol 48.84 g/L; K_2SO_4 18.2 g/L; $MgSO_4$ 7.28 g/L; KOH 4.13 g/L; $CaSO_4.2H_2O$ 0.93 g/L; $(NH_4)_2SO_4$ 8.42 g/L; 85% H_3PO_4 26.7 mL/L; PTM4 salts 4.4 mL/L.

3.3.2 Inoculum preparation

Glycerol stock stored at -80˚C was streaked on a YPD plate for the growth of colonies. A single colony from this YPD plate was inoculated as a starter culture into a 5 mL test tube containing YPD media. Pre-culture was prepared by inoculating starter culture into a baffled flask containing 300 mL of YPG media. The pre-culture was grown at 30˚C and 220 RPM for 24 h to a final OD_{600} of 4 absorbance units.

3.3.3 Dielectric Spectroscopy

Dielectric spectroscopy (Impedance spectroscopy/Capacitance probe) measures viable cell concentration (VCC) based on dielectric properties exhibited by the cells suspended in the conductive medium. The tip of the capacitance probe (Aber Instruments, Aberystwyth, United Kingdom) is encircled by two annular rings on a ceramic covering that generates an electric field to the applied frequency. The influence of the electric field on the cells forces ions across the cell membrane to polarize as intracellular/extracellular ions move toward the non-conducting plasma membrane [169]. The charge gradient across the membrane builds up a capacitance and each cell act as a tiny capacitor holding permittivity (ε) or capacitance (pF/cm). This can be distinguished from the dead cells having damaged cell membranes, presumed to be lysed, damaged, or fragmented. The cellular cytoplasm exhibits a notable degree of conductivity owing to the abundance of salts and nucleic acids, while the plasma membrane, comprising a lipid bilayer, demonstrates relatively low or negligible conductivity. Upon the application of an electric field to a suspension of cells, cellular polarization ensues as a consequence of intracellular ion migration along the imposed electric field. However, this movement encounters hindrance from the plasma membrane, functioning as an insulative barrier between intracellular ions and those within the culture medium. Consequently, ions from these respective regions migrate towards their oppositely charged electrodes. The non-conductive nature of the plasma membrane confines ions proximate to its surface, thereby facilitating cells to function as capacitors. Consequently, deceased cells are unable to undergo polarization due to the absence of an intact cell membrane. The presence of an intact cell membrane is thus imperative for achieving charge separation, a process integral to the determination of measured permittivity (or capacitance) values (Fig. 3.1). Consequently, deceased or compromised cells lacking intact plasma membranes, alongside other particulate matter devoid of such membranes, do not contribute to overall permittivity values. The escalation in the volume fraction of cells correlates with an increase in available membranes for polarization, consequently augmenting the suspension's capacitance. Thus, the measured capacitance value serves as an indicator of the total volume fraction of cells within the suspension. Besides biomass, the remaining suspended particles in the cultivation broth do not produce capacitance as they lack a plasma membrane [99,112,169]. In this chapter, the capacitance was recorded at dual frequencies (580 kHz and 15.560 MHz) by Futura software (Aber Instruments, Aberystwyth, United Kingdom), and the difference in value between the frequencies (ΔC) represents the overall capacitance value and related to dry cell weight (Annexure A3).

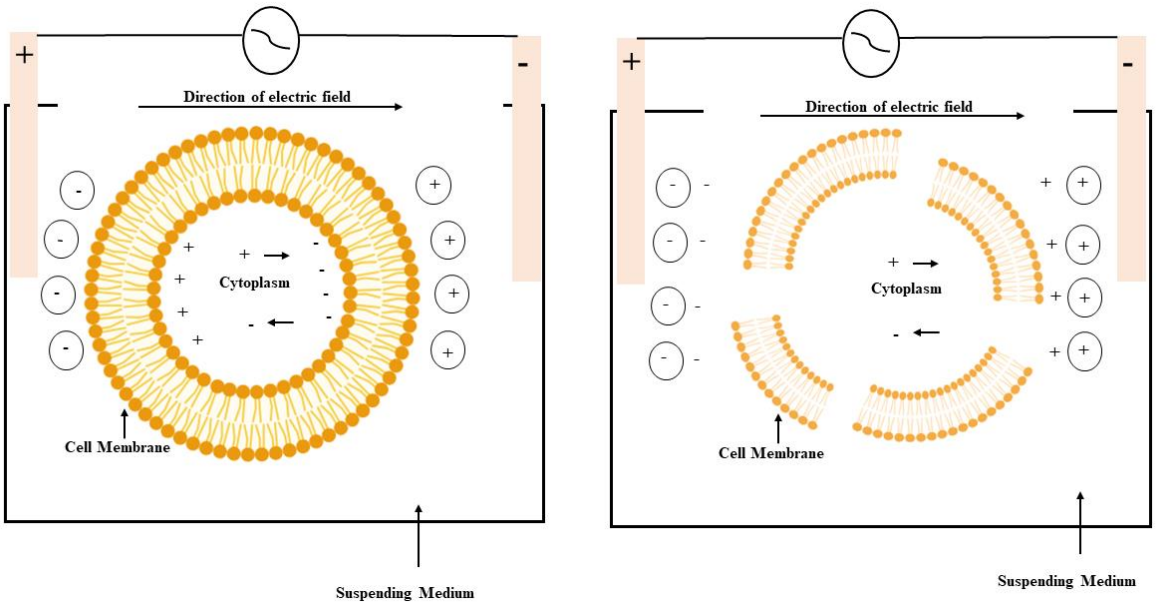


Figure 3.1: Principle of dielectric measurement in a cell suspension medium. Left side represents intact cell membrane and right side represents a lysed cell membrane (Adapted from [169]).

This chapter also explores the application of DS as a supporting PAT process analyzer to elucidate the potential of the metabolic heat rate signal for monitoring *P. pastoris* cultivation.

3.3.4 huIFN α2b production in fermentation calorimeter

huIFN α2b production by glycoengineered *P. pastoris* was carried out at 5L scale. The calorimeter setup working principle described previously (section 2.3.4) was *in-situ* sterilized, prefilled with distilled water at 121°C and 1 bar pressure for 20 min, followed by aseptic transfer of separately autoclaved optimized BSM medium. Reaction temperature was set to 30°C, and the pH of the culture was maintained at 5.4 by adding ammonia (25 % v/v solution) and ortho-phosphoric acid (85 % v/v solution). Dissolved oxygen was maintained above 15 % by combining airflow with pure oxygen (total flow of 1 vvm) and controlling agitation rate between 400 to 800 rpm. A 10% v/v of preculture was transferred into the calorimeter aseptically making the final working volume to 2.5 L. The course of growth of the organism was executed in the following order: Glycerol, Adaptation, and Induction phases, respectively. The first phase utilizing glycerol as chief carbon source generates higher amount of biomass. Depletion of C-source, which brings metabolic activity to a standstill is indicated by a sudden spike in DO signal and complemented with simultaneous drop in other process signals like metabolic heat rate, capacitance, and off-gas activity were noticed. Adaptation phase engaged in adjusting the carbon metabolism towards methanol, was carried out by supplying its minimal

doses to the organism. Subsequent doses gradually activates AOX transcription elements, which can be tracked by the response in capacitance, metabolic heat rate, and off-gas activity. Induction/Production phase involves continuous feeding of methanol to the organism in a mechanism governed by appropriate process models, in our case by exponential feeding. The underlying mechanism of feeding the methanol was accomplished through substrate balance equation and regulation of methanol feeding by implementing various control strategies were explained extensively under 'Feed rate regulation' (Section 3.3.10). Samples were collected at regular time intervals for offline analysis.

3.3.5 Supervisory Control and Data Acquisition (SCADA)

All the process inputs and outputs were integrated by a data acquisition (DAQ) hardware (cRIO 9075, National Instruments, Austin) with specific slots harboring Analog I/O and Digital I/O cards. All the signals were acquired, preprocessed and conditioned by a SCADA built on LabVIEW-based programming codes. Real-time signals, Controlling actuators, Calculated variables (q_B, CER, ΔC, etc.) and were logged at every 10 s interval and can be tracked in real timein graphical plots.

3.3.6 Offline Analysis

Collected samples were stored at 4ºC before the analysis. Cells were separated by centrifuging 1 mL of the sample into pre-weighed 1.5 mL tubes at 10,000 RPM for 10 min at 4ºC. Further, the supernatant was collected in a separate vial and stored at -20ºC for C-substrate, and huIFN α2b analysis. Cells were washed twice with deionized water to remove any salts present and dried at 80ºC for dry cell weight (DCW) determination. Cell growth was measured through optical density measurements at 600 nm (OD_{600}) using a UV-visible spectrophotometer (GE Healthcare, UK). Absorbance can be converted to DCW (in g/L) values by the OD vs DCW correlation (1 OD = 0.21*CDW) developed previously.

Residual glycerol and methanol concentrations in the culture broth were determined using HPLC by reverse phase analytical chromatography (Rezex RHM- monosaccharide H+ Column, Phenomenex Inc, California, USA). Mobile phase containing degassed 5 mM Sulphuric acid was set to flow through the column at 0.6 mL/min in isocratic mode, and column chamber was maintained at 50°C. Determination of huIFN α2b concentration in the collected samples was performed by Enzyme linked immunosorbent assay (ELISA) analysis using huIFN α2b ELISA kit (Mabtech AB, Sweden).

3.3.7 Development of specific growth rate estimator

In the previous chapter (section 2.4.5) we have concluded that cumulative heat-based estimator provides reliable estimations for microbial systems exhibiting high exothermicity. Cumulative heat-based estimator was developed on the basis of yields emerged from the entities' conversion rates with respect to their substrate utilization and biomass formation. Specific rates are calculated from Eq. (3.1).

$$r_X = \mu . X \tag{3.1}$$

Where r_X is biomass production rate and the biological heat yield coefficient ($Y_{Q/X}$), which is obtained from the slope of the linear plot between cumulative metabolic heat production and biomass concentration, is related to the specific growth rate as shown in Eq. (3.2).

$$Y_{Q/X} = {q_B}/{r_X} = {q_B}/{\mu . X} \tag{3.2}$$

The profile of dynamic biomass concentration (X) can be deduced by integrating Eq. (3.2) between the limits $X: X_0 \rightarrow X_t$ and t: $0 \rightarrow t$,

$$X = X_0 + \frac{Q_t}{Y_{Q/X}} \tag{3.3}$$

Where q_B, biological heat rate (W/L), X_o, biomass concentration at the end of the batch phase (g/L), $Y_{Q/X}$, biomass heat yield coefficient (kJ/g) and Q_tCumulative heat at that time instant (kJ/L). Rearranging Eq. (3.3), the specific growth rate can be estimated as a function of metabolic heat rate and cumulative heat, as shown in Eq. (3.4)

$$Y_{Q/X} = \left[\frac{q_B}{\mu \left(X_0 + {Q_t}/{Y_{Q/X}} \right)} \right] \tag{3.4}$$

Estimator (Eq. 3.4) could be helpful when the volume changes in a bioreactor are negligible. In the case of fed-batch fermentation, it is essential to account for the variation in the culture volume due to the continuous substrate addition [170]. Hence, Eq. (3.4) requires a modification by incorporating the volume term (V_r).

$$dx . V_r = \frac{q_B}{Y_{Q/X}} dt \tag{3.5}$$

Integrating the above relation and rearranging will yield a specific growth rate estimator for varying volumes.

$$\mu_{est} = \frac{q_B}{x_0 V_0 Y_{Q/X} + (Q_t - Q_0)} \quad (3.6)$$

Where V_0 is the volume of reaction broth at the end of batch phase. The modeled μ_{est} was programmed in the LabVIEW utilizing the dynamic q_B *and* Q_t values from DAQ, looped into the PID controller to actuate methanol feed rate as control output.

3.3.8 Statistical Analysis

The reliability of model-based μ_{est} values need to evaluated in order to rate the performance of specific growth rate estimator and it was done by comparing the values of μ from offline measurements by a performance index, Root Mean Square Error (RMSE). It can be calculated as shown in Eq. (3.8)

$$\text{RMSE} = \sqrt{\frac{\sum_{i=N}^{N}(y_{predicted} - y_{actual})^2}{N}} \quad (3.8)$$

Where $y_{predicted}$ is online estimated value of the state variable (specific growth rate values from estimator in this case), y_{actual} is the offline value of the state variable (specific growth rate determined offline), and N is the number of observations

Assessment of controller performance was accomplished through calculation of performance index ISE. Eq. (3.9) determines the integral square error of the controller.

$$ISE = \int_0^t e^2 dt \quad (3.9)$$

Where e is the error of the controller, defined as the difference between setpoint and process variable, and t is the operation time of the controller.

3.3.9 Process control strategy

Continuous methanol feeding helps to sustain active functioning of recombinant protein expression under AOX promoter. Control strategies lay its focus on the induction phase μ, which directly relates to the protein production rates. This chapter highlights the effect of different control strategies viz. Feedforward, Feedback (PID control) and Feedback with gain scheduling (Adaptive PID) on μ_{SP} during the induction phase. Growth, C-source utilization, huIFN α2b titer and specific productivities were compared and reported for every control strategy. Feedforward control function as an open-loop system, which dictates exponential increase in feed rate with respect to time, and predicted to represent an increasing trajectory of biomass. Batch phase data aids in computing the initial feed rate, and further programming the precision pumps would enable exponential methanol feeding into the reactor. But predetermined feed unknown of microbial growth may turn out to be lethal for growing cells. Feedback control overcomes this limitation and attaches a sensor input to the control circuit.

Continuous update of μ_{est} will allow the controller to correct its output in terms of feed rate to assuage the process variable within a buffered range. A link between sensor input, controller and actuator established a closed-loop system of functioning and described in detail in 'Feed rate regulation' (section 3.3.10). The final strategy with gain scheduling is a special case that helps to reform the controlled operation by changing the PID parameters dynamically, which in turn updates the feed rate for stable control.

3.3.10 Feed rate regulation

Feed rate, which also is a manipulating variable that affects metabolic rates, viz., specific growth rate, specific IFN α2b productivity, etc. Time-based exponential feeding rate guided by a PID control strikes a balance between starvation and toxicity due to methanol. This work employed three feeding strategies to control the specific growth rate. Feed forward (FF) strategy involved open loop control of specific growth rate by feeding methanol. The feeding rate can be initiated with the usage of the batch-phase constants, *viz*., biomass concentration (X_0), the volume of reaction broth at the end of the batch phase (V_0), initial substrate concentration of the feed (S_0), and biomass yield per substrate consumed ($Y_{X/S}$) as shown in Eq. (3.10)

$$F_O = {X_0 V_0 \mu_{set}}\Big/{Y_{X/S} S_0} \tag{3.10}$$

The time-dependent feeding regulated by the feedforward equation is given in Eq. (3.11)

$$F_{FF}(t) = F_0 e^{\mu t} \tag{3.11}$$

Where F_{FF} represents feeding using feedforward strategy, t is the duration of feeding. Fermentation calorimeters possess the advantage of deciphering heat rate signal into real timeμ_{est}, which enables us to convert it into a closed-loop control system for the desired μ setpoints. The fidelity of the μ_{est} to capture the dynamics of the process enables reliable control over the process parameter. The exponential term gets updated based on the process dynamics, and the applied PID parameters would transform into the control action as a feed rate, as shown in Eq. (3.12). In this chapter, an exponential feeding strategy combining feedforward and feedback terms (Eq. (3.12)) was employed to have tighter control of μ over its μ_{SP} value.

$$F_{FB} = F_0 e^{(PID\ (O/P)).t} \tag{3.12}$$

The error of the controller was computed through Eq. (3.13)

$$\varepsilon(t) = \mu_{SP}(t) - \mu_{est}(t) \tag{3.13}$$

The output of the controller ($PID\ (O/P)$ was based on Eq. (3.14)

$$PID\ (O/P) = k_p((\varepsilon(t) + \frac{1}{\tau_i}.\int_0^t \varepsilon(t).dt + \tau_D.\frac{d}{dt}.\varepsilon(t)) \tag{3.14}$$

Where k_p, τ_i and τ_D are proportional gain, integral time (s), and derivative time (s) respectively. During the FB operation, the tuning parameters were maintained for a fixed period and regularly varied, and the scheme for changing tuning parameters is detailed in the 'Results and discussion' (section 3.4). The adaptive PID control (FB+GS) experiment updated the control law through the gain scheduling technique. Tuning parameters for the gain scheduling algorithm were obtained from fixed gains PID-based experimental run. The gains were updated according to the change in error value. High precision pump (Watson Marlow 120U, Falmouth, Cornwall, U.K) feeding methanol responds to the controller output actuated by the PID loop. A schematic of the entire process is depicted in Fig. 3.2.

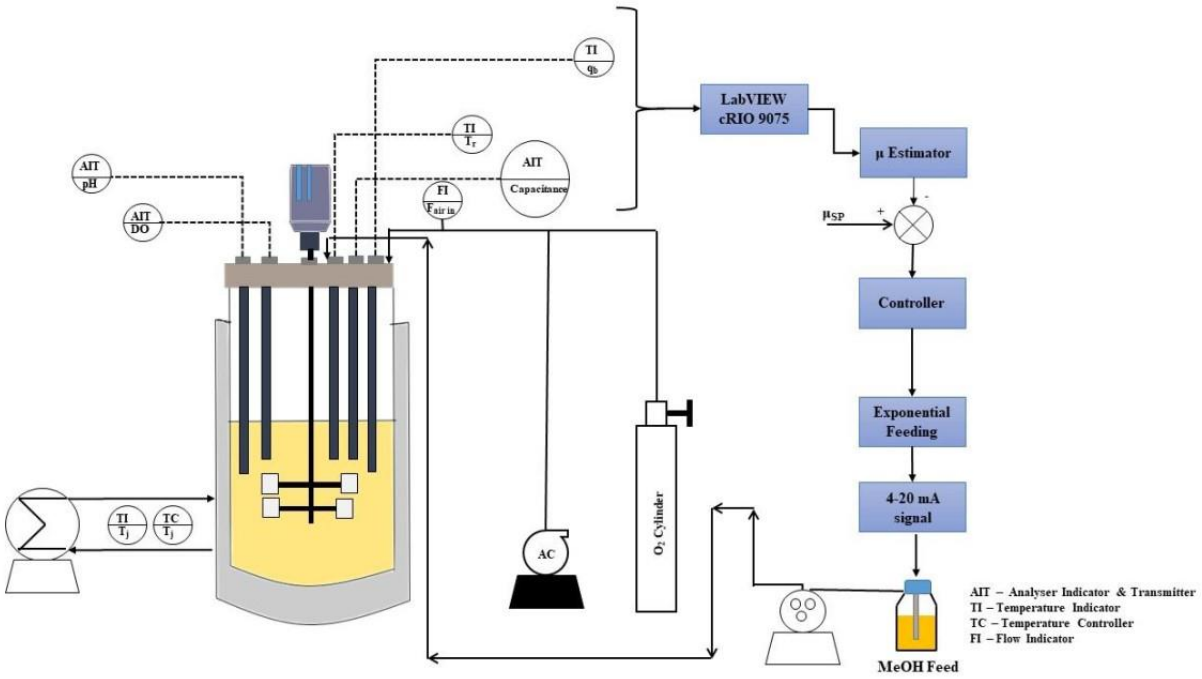


Figure 3.2: Schematic representation of fermentation calorimeter setup for real-time monitroing and control of specific growth rate of glycoengineered *P. pastoris* cultivation for the production of huIFN α2b.

3.3.11 N – glycan analysis

The glycosylated huIFN α2b underwent N-glycan profiling through MALDI-TOF MS analysis [21]. The procedure involved several steps: first, the protein sample underwent reduction, carboxymethylation, and tryptic digestion. Subsequently, PNGase F digestion was utilized to release N-glycans, followed by permethylation of these released N-glycans. The permethylated N-glycans were then separated using a reverse-phase C18 cartridge. Different elution fractions (15%, 35%, 50%, and 75% acetonitrile) were collected and mixed with 15 mg/mL of 2,5-dihydroxybenzoic acid (DHB) matrix in 70:30 methanol/water solution. These mixtures were then spotted onto a target plate and dried under vacuum. MALDI-MS (MALDI TOF/TOF 5800, AB SCIEX) was employed to capture the molecular ion signals from the permethylated glycans corresponding to $[M+Na]^+$ adducts. The predominant peak obtained from the MALDI-MS analysis underwent MS/MS analysis to acquire crucial structural information, such as branching patterns, antennae structures, and linkage positions. The MS/MS spectra of the fragment ions were analysed using the GlycoWorkbench suite, facilitating the annotation of glycan fragment spectra.

3.4 Results and Discussion

3.4.1 Comparative assessment of process analysers and μ_{est} deduced from heat-rate signal

Combined deployment of process analysers such as the Fermentation calorimeter, Capacitance probe, and CO_2 analyser represented the increase in biomass, CER, and other metabolic activities of *P. pastoris* growth. Calorimetry coupled with the respirometric signals interprets the correlation between heat fluxes of the bioprocesses associated with the change in metabolic flux. In *P. pastoris* cultivation, investigating calorimetric profiles with respirometry provides a glimpse of energy metabolism. The synchronization between the afore mentioned process analyser signals and its consistent increase could be attributed to the volumetric increase in the biomass during the high cell density cultivation of *P. pastoris*. The drop in process analyser signals at the end of the glycerol phase and slower adaptation of methanol utilization followed by the induction phase could be noticed in the thermogram and respirogram (Fig. 3.3) at different stages of the cultivation. The μ_{est} depicts the real-time estimation of μ during the huIFN α2b production phase, i.e., the induction phase. For FF and FB strategies, 3 different setpoints i.e. 0.04, 0.035 & 0.03 h^{-1} were carried out sequentially. From the outcomes of previous control strategies implemented, PID with gain scheduling operation was carried out with only 0.035 h^{-1}. The estimated offline specific growth rate (μ_{off}) values could be found as

high as 0.05 h^{-1} and to the lowest 0.025 h^{-1} in all the fed-batch operations. The heat rate-based estimator developed using Eq. (3.6) was reliable in estimating specific growth rates. Root Mean Square Error (RMSE) computed between real-time and offline μ values determine the quality of the μ_{est} signal. Lowermost RMSE indexes are more favoured for setpoint-process variable combinations to achieve robust control operation and the average RMSE was found to be 0.0051 h^{-1}. The utilized methanol undergoes oxidation to formaldehyde in the peroxisome, further leading to stronger oxidation to synthesize energy intermediates for cellular functioning by evolving CO_2 gas. Higher CO_2 synthesis during the induction phase hints at the cell's uptake of methanol majorly diverted towards energy metabolism and, to a lesser extent, towards growth resulting in lower μ values. Calorimetry coupled with respirometric signals interprets the correlation between heat fluxes of the bioprocesses associated with the change in metabolic flux. In *P. pastoris* cultivation, investigating calorimetric profile with respirometry provides a glimpse of energy metabolism. Higher cumulative CO_2 production in induction phases is a good indicator of the ongoing AOX expression activity, which could be elucidated in real-time by estimating dynamic $Y_{CO_2/S}$ values. Continuous feeding of methanol in the FB control run achieved a high huIFN α2b titer of 309.4 mg/L with concomitant highest CER (241 m.mol/L.h) got recorded. CER profile also helps to identify the tracking of methanol feeding such that it does not reach toxic levels. In Fig. 3.4 B, it is essential to note that a perilous level of methanol was accumulated from 65th h. Simultaneously, the CER profile excessively increased from 187 to 241 m.mol/L.h in the same duration, and this increase was observed due to extensive oxidation of methanol to CO_2. This observation will be helpful in the future in tracking the aberrations leading to the destabilization of the process. However, subsequent FB + GS control checked residual methanol concentration under control until its completion. PID controller action and methanol feed rates, manipulated the change in process inputs that were found reflected in modelled μ values. Compared offline values (μ_{off}) at various PID tunings for respective μ_{SP} are presented in Table 3.1. As Gibb's free energy dissipation is imminent to any living system, the change in enthalpies for various metabolism could be successfully deciphered by fermentation calorimeter, and that resulted in better tracking of μ_{est} for all set points. The RMSE value achieved for the metabolic heat rate-based estimator (0.0051 h^{-1}) is significantly low compared to specific growth estimators (RMSE range: 0.2 to 0.4 x 10^{-3} h^{-1} & 0.063 to 0.01 h^{-1}) previously reported in the literature [158,171].

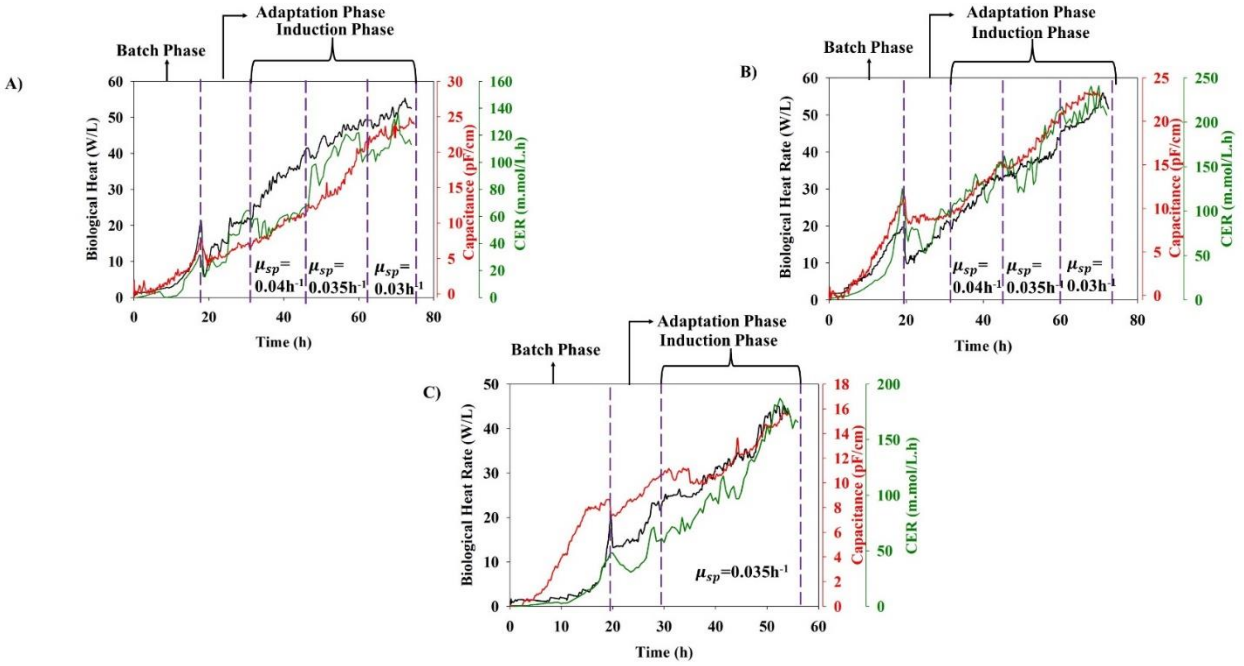


Figure 3.3: Dynamic plots representing metabolic heat rate (Black continuous), capacitance (red continuous) and CER (green continuous) segegrated into glycerol batch phase, adaptation phase and induction phase for 3 different control stratgies. A) Feedforward control (phase I $\mu_{SP} = 0.04\ \mathrm{h}^{-1}$, phase II $\mu_{SP} = 0.035\ \mathrm{h}^{-1}$, phase III $\mu_{SP} = 0.03\ \mathrm{h}^{-1}$), B) Feedback using regular PID control (phase I $\mu_{SP} = 0.04\ \mathrm{h}^{-1}$, phase II $\mu_{SP} = 0.035\ \mathrm{h}^{-1}$, phase III $\mu_{SP} = 0.03\ \mathrm{h}^{-1}$), C) Feedback using adaptive PID control with gain scheduling strategy (phase I $\mu_{SP} = 0.035\ \mathrm{h}^{-1}$).

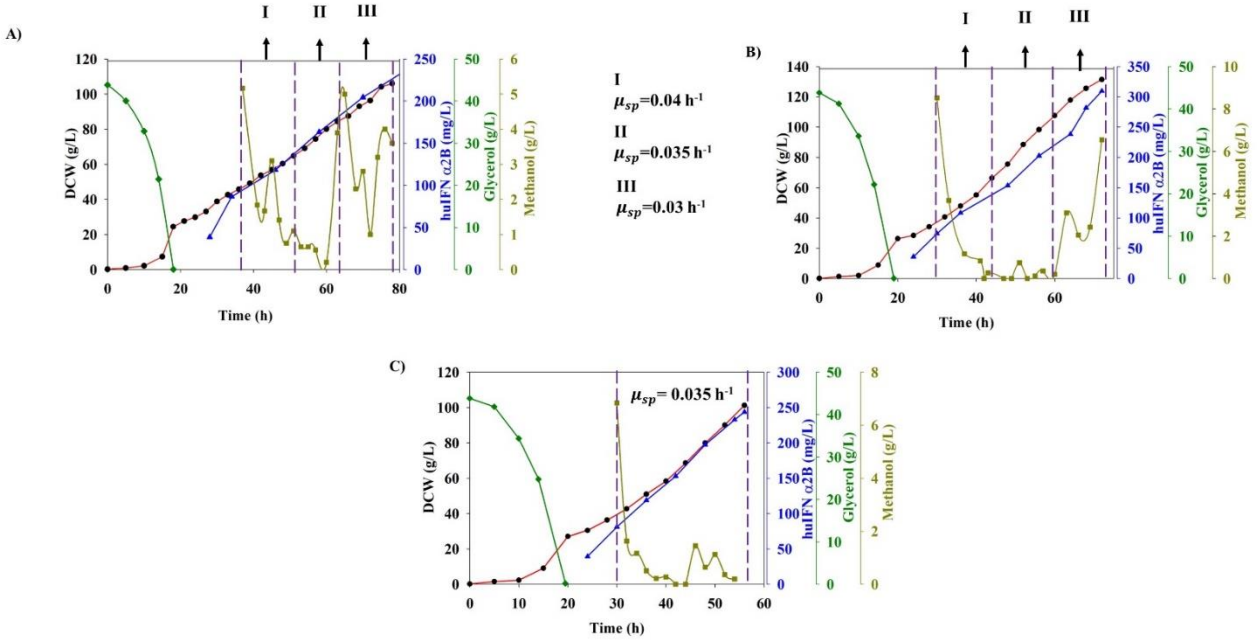


Figure 3.4: Time courses representing cell concentration (Red Circles), consumption of glycerol (Green diamonds), consumption of methanol (Light Yellow squares), and production of huIFN α2b (Blue triangles). A) Feedforward control (phase 1 $\mu_{SP} = 0.04\ \mathrm{h}^{-1}$, phase 2 $\mu_{SP} = 0.035\ \mathrm{h}^{-1}$, phase 3 $\mu_{SP} = 0.03\ \mathrm{h}^{-1}$) , B) Feedback using regular PID control (phase 1 $\mu_{SP} = 0.04\ \mathrm{h}^{-1}$, phase 2 $\mu_{SP} = 0.035\ \mathrm{h}^{-1}$, phase 3 $\mu_{SP} = 0.03\ \mathrm{h}^{-1}$), C) Feedback using adaptive PID control with gain scheduling strategy (phase 1 $\mu_{SP} = 0.035\ \mathrm{h}^{-1}$).

Table 3.1: Effect of tuning parameters on controlled variable in feedback PID control

Controller gain, k_C	Integral time, τ_I (s)	Derivative time, τ_D (s)	$\mu_{sp}(h^{-1})$	$\mu_{Off}(h^{-1})$
3.5	60	0.05		0.0237
2	75	0.05	0.03	0.0256
1.5	90	0.1		0.0286
4.5	45	0.05		0.0386
3	60	0.1	0.035	0.0361
2	90	0.1		0.0356
6	30	0.05		0.0427
4.5	60	0.5	0.04	0.0352
2.5	120	0.5		0.0385

3.4.2 Implications of different control strategies on huIFN α2b production

The batch growth phase was continued until the glycerol was exhausted, which was manifested by a sudden rise in the DO signal. Glycerol exhaustion occurs roughly between 18 to 20 h after inoculation and yields cell density in the range of 22 to 25 g/L. The batch growth phase limited by glycerol supports primarily biomass synthesis, which is evident from the 75% of the carbon flux distribution to facilitate achieving higher biomass yield ($Y_{X/S} = 0.756 \frac{g.biomass}{g.glycerol}$). Thus, substantial biomass was generated to sustain for extended periods in the methanol induction phases subsequently. Initially, few shorter pulses of methanol (5 g) were fed into the reactor, allowed to metabolize and slowly activate the AOX transcription elements. The transition phase is characterized by a negative growth rate at the initial stages and gradually proceeds to a positive slower growth rate in repetitive pulsing. During the induction phases of the FF and FB strategy, methanol feeding was done at three different set point values (μ_{SP}) of the order 0.04, 0.035, and 0.03 h^{-1} respectively, and the feeding was continued for 72 h. Specific productivity (q_P), biomass and huIFN α2b yields exhibited the highest values at the optimal μ_{SP} = 0.035 h^{-1} (Table 3.2). The gain scheduling approach was carried out at μ_{SP} = 0.035 h^{-1}

owing to optimal huIFN α2b productivity concluded from our previous operations. In this approach, μ_{SP} setpoints are dynamically changed responding to the previous trend of the process variables. The effect of PID parameters on feed rate over a time period was recorded for a fixed μ_{SP} and this relationship was helpful for gain scheduling approach.

Table 3.2: Kinetic parameters Biomass concentration, huIFN α2b concentrations, Yield coefficients ($Y_{X/gly}, Y_{X/met}$ *and* $Y_{huIFN\alpha2b/met}$) and specific productivity, q_P was computed for different feeding strategies (FF, FB and FB+GS)

Run	μ_{sp} (h^{-1})	Biomass conc. (g/L)	Final biomass conc. (g/L)	huIFN α2b (mg/L)	$Y_{X/gly}$ (g/g)	$Y_{X/met}$ (g/g)	$Y_{huIFN\alpha2b/met}$ (mg/g)	q_P (mg/g.h)
FF		33.12[a]			0.77[e]			
	0.04	14.07[b]				0.11[f]	0.52[g]	0.18±0.02[h]
	0.035	19.41[b]				0.13[f]	0.61[g]	0.216±0.04[h]
	0.03	12.54[b]	96.92	218.6		0.105[f]	0.58[g]	0.18±0.04[h]
		17.4[c]						
FB		32.2[a]			0.75[e]			
	0.04	35.11[b]				0.12[f]	0.55[g]	0.24±0.04[h]
	0.035	41.3[b]				0.15[f]	0.63[g]	0.35±0.03[h]
	0.03	23.09[b]	131.7	309.4		0.11[f]	0.57[g]	0.28±0.04[h]
FB+GS		32.47[a]			0.75[e]			
	0.035	68.53[b]	101.2	243.65[d]		0.18[f]	0.66[g]	0.56±0.08[h]

a. Final dry cell weight achieved during batch growth phase with glycerol as sole carbon source

b. Final dry cell weight achieved during induction phase at end of control run for the given setpoint

c. Final dry cell weight at end of uncontrolled feeding during induction phase

d. huIFN α2b titer obtained at the end of 56th hour.

e. Yield coefficient values computed at the end of batch phase

f. Yield coefficient values computed at the end of control run for given setpoint

g. Yield coefficient values computed at the end of control run for given setpoint

h. Productivities calculated at end of control run for given setpoint

FF control strategy was accomplished through a predetermined feeding of methanol. The initial feeding rate (F_o) computed using Eq. (3.10) and were observed to be 5.5 mL/h, 5 mL/h, and 5 mL/h for FF, FB, and& FB+GS approaches, respectively. F_o was determined from the batch-phase data as shown in Eq. (3.10), and the FF strategy was continued independently with respect to time, neglecting the dynamics of μ_{met} values. Open loop control systems are least concerned about the perturbations and deviations occurring in the process, and control over desired process variables is difficult to achieve. Incorporating the FF strategy resulted in organisms growing at lower specific growth rates than the desired setpoint as the process progressed (Fig. 3.5). Predetermined feeding was not favourable due to the accumulation of methanol at various stages of huIFN α2b production (Fig. 3.4A). This led to pseudo-continuous feeding of methanol to avoid the toxic levels. Continuous drop in μ_{est} can be noticed (Fig. 3.5), which could be attributed to excess feeding of methanol. Preliminary investigation on the pulsed feeding strategy at different dosage rates suggested that the glycoengineered *P. pastoris* strain is highly unlikely to influence the functional AOX machinery. A similar trend of continuous drop in μ_{off} values were observed (Fig. 3.6) due to intermittent feeding of methanol due to the presence of excess methanol which resulted in nonlimiting conditions. Even though this particular strain can accumulate methanol up to 14 g/L without its cellular mechanism getting affected. However, the accumulation of methanol was consistent in the FF strategy and this resulted in lower huIFN α2b productivity as AOX1 machinery gets 3 to 5 times slower than its regular output in the presence of excess methanol [172]. In the case of the FB control strategy, the feeding rate was determined more specifically based on μ_{est} as shown in Eq. (3.12). Though the FB strategy regulates the feeding independently, a constraint was laid that the feeding does not exceed 50 mL/h to avoid methanol toxicity (> 1.4% v/v). The optimal μ_{SP} was found to be 0.035 h^{-1} resulting in a significant increment of huIFN α2b productivity, which was repetitively observed in both FF and FB strategies. In all the FB-based runs, exponential feeding rates accomplished the overall metabolic requirements of cells by concurrent substrate utilization (residual methanol concentration maintained closer to zero) for various activities more than 14 hours, as shown in Fig. 3.4. The trajectory of the CER signal corroborated well with heat rate and capacitance. Still, the cumulative CO_2 formation in the induction phase was proportionately three times higher than the glycerol batch phase. Higher volumetric CO_2 generation occurred during the induction phase due to the highest degree of reduced carbon source, i.e., methanol (Degree of reduction, γ_S = 8) employed. Based on the experimental data,

simple adaptive control significantly enhanced the q_P value, which is 1.5-fold higher than the conventional FB approach with PID control and 2.2-fold higher than FF control.

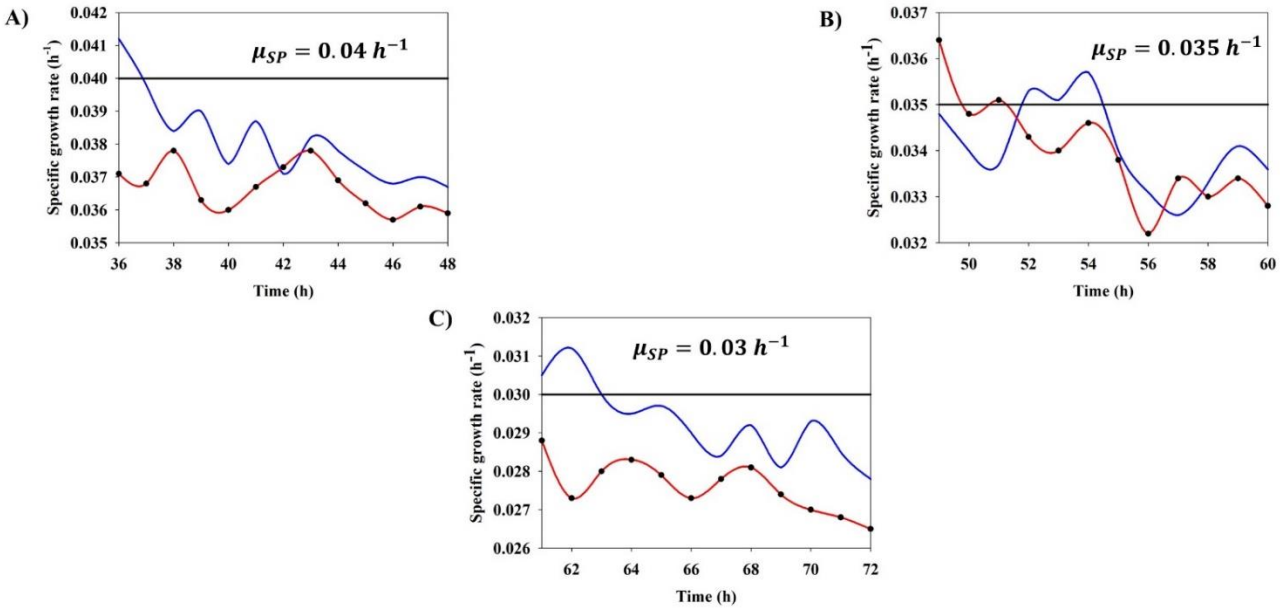


Figure 3.5: Feedforward controller performance based on comparative evaluation of μ_{est} (Blue continuous) μ_{off} (Red continous and Black filled circles) and μ_{SP} (Black continuous) at various setpoints A) $\mu_{SP} = 0.04\ \mathrm{h}^{-1}$, B) $\mu_{SP} = 0.035\ \mathrm{h}^{-1}$, C) $\mu_{SP} = 0.03\ \mathrm{h}^{-1}$

3.4.3 Specific growth rate control of *P. pastoris*

AOX transcription is relatively more potent than the constitutive expression system and most preferred by the industry for the large-scale production of heterologous proteins. The promoter being sensitive to the inducer, control of expression can be handled at ease by the regulated feed of methanol. Unlike previous approaches, the feedback control is challenging to deploy in the induction phase due to lower values recorded, and a significant fraction of methanol was oxidized to CO_2 to drive energy for the expression system-related activities. Therefore, the μ values span between 0.01 to 0.05 h^{-1} during the induction phase, and controlling at a narrower range is vital to sustain the derepressed function of the AOX expression system. Application of soft sensors and elucidation of μ by capacitance probe was already reported by our research group [159], and feedback control of μ_{SP} resulted in a proportional increase of huIFN α2b titer.

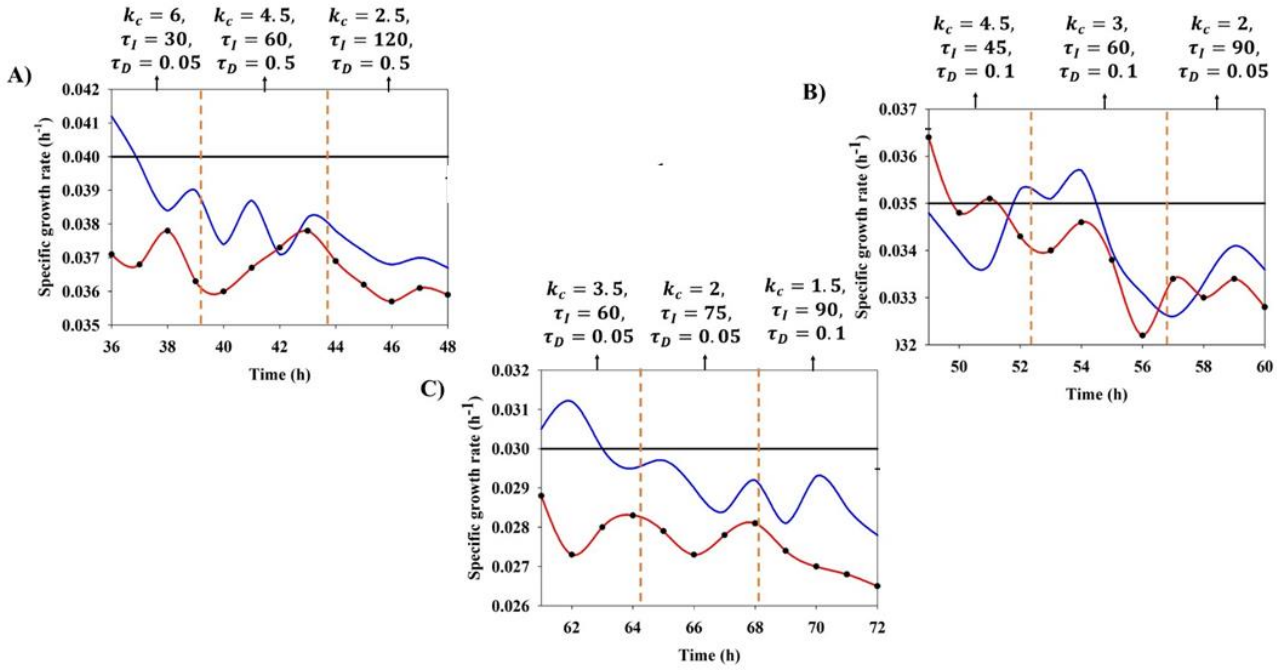


Figure 3.6: Feedback strategy based PID controller performance based on comparative evaluation of μ_{est}(Blue continuous) μ_{off} (Red continous and Black filled circles) and μ_{SP} (Black continuous) at various setpoints A) $\mu_{SP} = 0.04\ h^{-1}$, B) $\mu_{SP} = 0.035\ h^{-1}$, C) $\mu_{SP} = 0.03\ h^{-1}$

Especially, fed-batch operation at μ_{SP} = 0.04 h^{-1} yielded 1483 mg/L of huIFN α2b. This broader investigation led to estimating final huIFN α2b titers at the end of high cell density cultivation at different μ_{SP} characterizes the inherent relationship between μ and protein titers. However, a detailed investigation at the optimal setpoint say μ_{SP} = 0.04 h^{-1} would further open up the avenue for tighter control and regulation of the process. Calorimetry coupled with respirometry has the merits of investigating finer metabolic perturbations in addition to measuring online μ values in a high cell density cultivation. A closer look on μ_{SP} at 0.04 h^{-1} at its feeding profile, and regulation of the residual methanol concentration during the induction phase is critical for the optimal feeding of methanol. Control of methanol-based substrate flow into the reactor poses an additional risk of inflammable nature at higher production scales. Therefore, catering to the operational instabilities, potential toxicity, and narrowest control setpoints in this approach classifies the present investigation to be more challenging than the previous approaches based on glucose feeding. Robust control of specific growth rates was achieved

using the FB strategy through PID control. The control at every setpoint was carried out for 12 hours with a set of PID tuning parameters. Methanol was fed with a smaller controller gain and gradually increased every 3 hours (Table 3.1). Adjusting the tuning parameters improved the controller's performance, as the process variable annealed towards the setpoint (Fig. 3.6). The changes adopted in the PID control elements were determined to reflect in the manipulation of feeding profiles (Fig. 3.8). Gradual increase in integral time (> 50s) was best suited for the controller in reducing the offset and resetting the process variable deployed to the system. This observation can be confirmed by the lower average tracking errors at different setpoints, as shown in Table 3.1. Due to a concomitant increase in methanol feeding, the specific growth rates did not drop below the setpoint values compared to the FF strategy. Accumulation of methanol was lower due to the reliable feedback action of the controller. However, continuous feeding of methanol (limiting substrate) has proven to influence the AOX expression system in regulating growth and other metabolic activities, which has resulted in higher huIFN α2b productivity in contrast with the FF strategy.

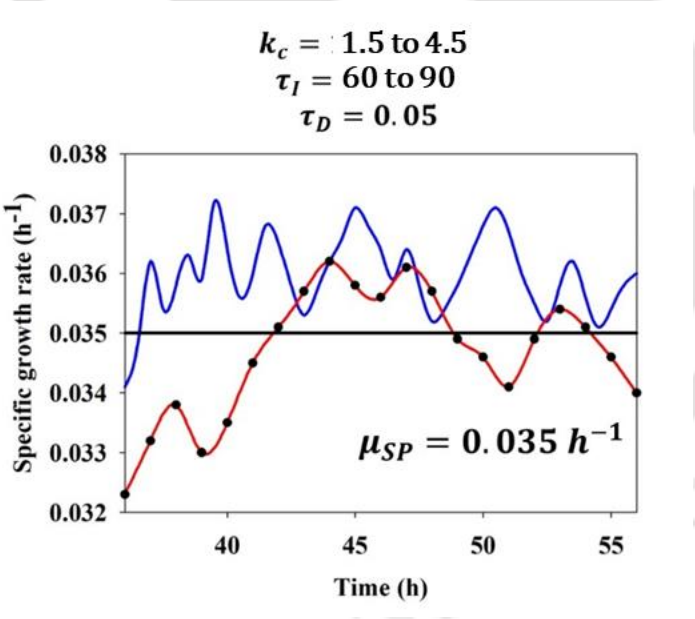


Figure 3.7: Adaptive PID controller performance based on comparative evaluation of μ_{est}(Blue continuous) μ_{off} (Red continous and Black filled circles) and μ_{SP} (Black continuous) at setpoint $\mu_{SP} = 0.035\ h^{-1}$

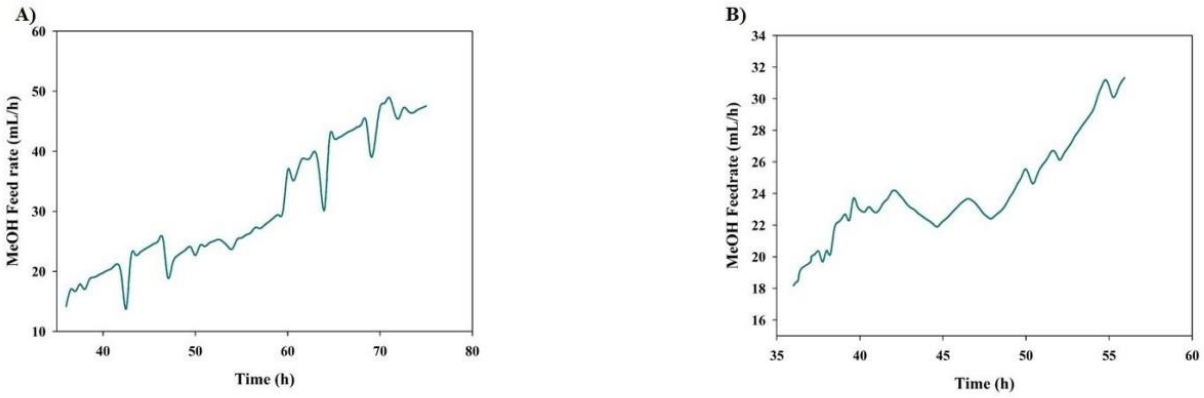


Figure 3.8: Influence of feeding rate (Blue continuous) due to changes in controller error A) Feedback using standard PID control, B) Feedback using adaptive PID control with gain scheduling strategy

The later stage of FB control was problematic due to the accumulation of methanol, and controller action needs reconciliation of feed which provoked the feeding pump to suddenly reduce the feed at two instances between 65^{th} to 72^{nd} hour as depicted in Fig. 3.8A. The gain scheduler managed to incorporate appropriate tuning parameters obtained through FB experimental run and actuated the feeding pump minimize the offset. During the control runs using standard PID and adaptive PID, rapid methanol consumption demonstrates that dynamic mass balance was established between methanol availability (feed stream) and subsequent consumption by *P. pastoris.* This action of tight substrate limitation and ongoing product formation ensured a balance between residual methanol concentration and huIFN α2b titers. Moreover, N-glycan analysis revealed that the most prominent peak corresponded to $Man_5GlcNAc_2$ at 1575.5 m/z (Fig. 3.9) with more than 85% homogeneity being achieved with tight control of growth rate. Whereas, the reduction in glycan homogeneity could be observed (Fig. 3.9 A) for the samples that were collected from the simple feedforward run. This once again underscores the importance of controlling CPPs to have a significant improvement in product quality. In addition, methanol feeding in the adaptive control strategy is less than the other two control strategies, which is a promising observation from the perspective of safety and economy. Hence, a tightly regulated feed rate limits the synthesis of by-products as overflow metabolism is avoided and reinstates housekeeping activities more effectively.

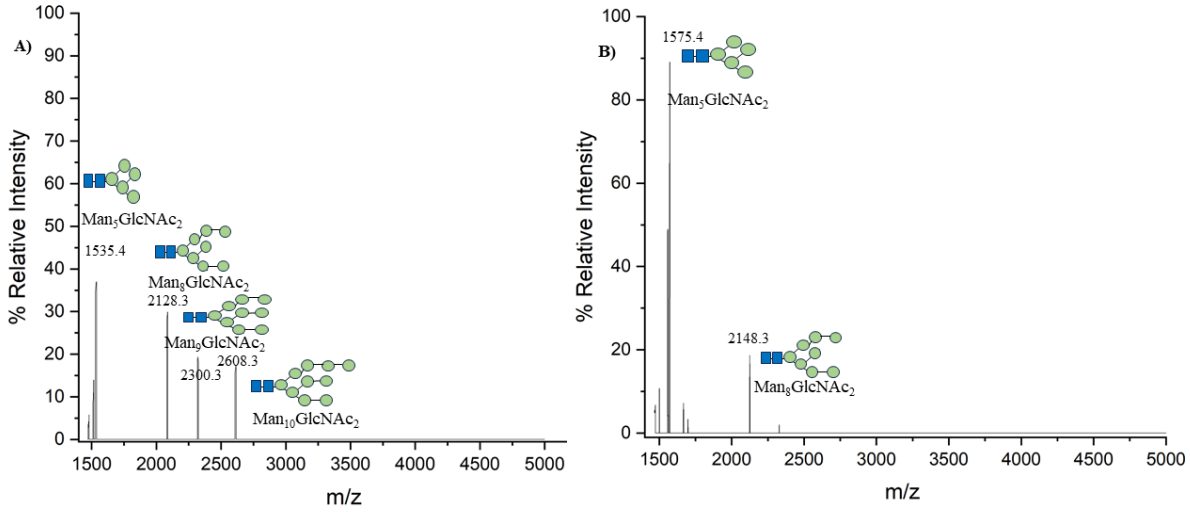
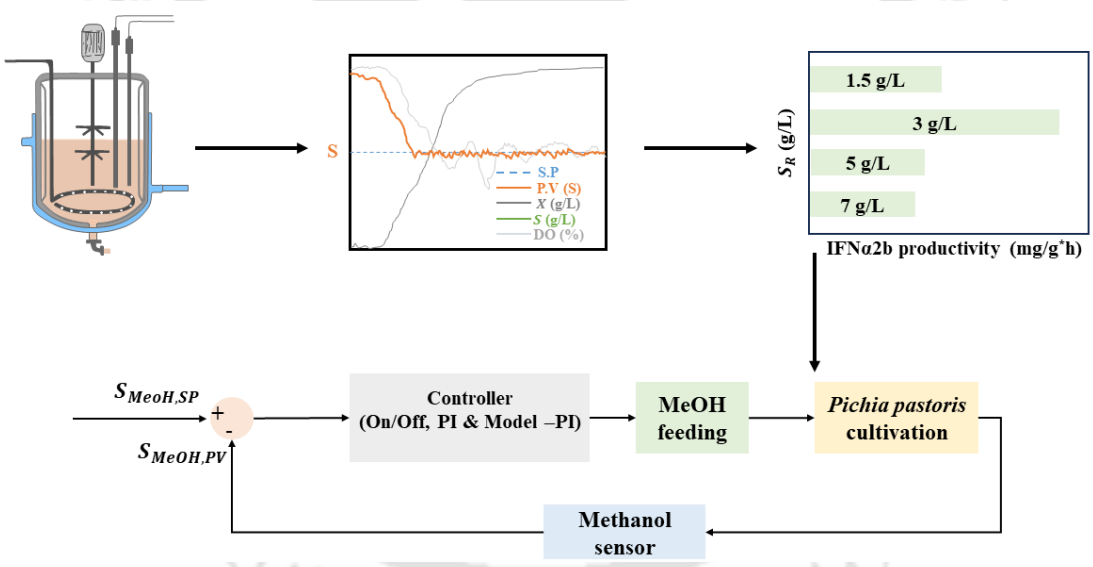


Figure 3.9: N – glycan analysis of purified huIFN α2b samples A) Sample from fed-batch fermentation where specific growth rate is not controlled. B) Sample from adaptive PID control run.

3.5 Conclusions

This chapter encompasses a detailed report on deciphering heat rate signals to monitor and stringent control of specific growth rates of *P. pastoris* in a narrow range (0.03 to 0.04 h^{-1}) for the production of huIFN α2b. This presented approach was the first to illustrate adaptive PID control with optimal tuning parameters, effectively controlling the μ at 0.035 $h^{-1}$with minor errors. Stringent control ensured that the methanol was efficiently uptaken by cells and channelized for growth and protein production, resulting in specific huIFN α2b productivity of 0.56 ± 0.08 mg/g.h. Process control approaches grants a deeper understanding of the nonlinear nature of cellular metabolism, and advances in control program may further influence the precision in quantifying process variables. Also, this present chapter proves that *P. pastoris* as a cell factory acts as a chassis for studying the multiple variables; their control and biological significance can be investigated further. Tight control of μ resulted in carbon limiting conditons as described in results and discussions (section 3.4). However, it is interesting to see the influence of optimal residual methanol concentration and direct control of the same on this glycoengineered *P. pastoris* strain and the impact of residual methanol concentration on various state variables such as μ, huIFN α2b productivity, and substrate uptake rate.

Chapter 4 Enhanced production of human interferon α2b in glycoengineered *Pichia pastoris* by robust control of methanol feeding and implications of various control strategies



4.1 Abstract

Like any other recombinant protein produced in methylotrophic yeast *P. pastoris*, huIFN α2b production in *P. pastoris* depends on methanol utilization during the induction phase. This chapter investigates the impact of varying residual methanol concentrations on huIFN α2b productivity and assesses control strategies towards the tight control of residual methanol concentration in fed-batch experiments carried out in the fermentation calorimeter. Real-time monitoring of *P. pastoris* metabolism was facilitated by metabolic heat rate measurements and exhaust gas analyser. Optimal methanol concentration screening performed at different concentrations viz., 1.5, 3, 5, and 7 g/L, revealed the highest huIFN α2b productivity at 3 g/L (0.21 mg/g h). Methanol sensor provided input data for three different control strategies (on/off, proportional integral (PI), and model-based adaptive PI) to maintain optimal methanol levels (3 g/L). Model-based adaptive control resulted in the highest huIFN α2b yield, productivity (244.34 mg/L, 0.31 mg/g h), surpassing on/off (188.7 mg/L, 0.16 mg/g h) and PI (202.3 mg/L, 0.28 mg/g h) control strategies. The model-based adaptive PI control framework developed in this chapter could be employed for other heterologous protein production in *P. pastoris*

4.2 Introduction

With the advent of microbial techniques, there is a diverse array of expression systems accessible for the production of recombinant proteins. The selection of appropriate microbial chassis is heavily influenced by the characteristics of the target protein, including its structure and biological activity [173]. Over the past thirty years, *P. pastoris* expression system has proven highly effective in producing a wide range of recombinant proteins, therapeutic molecules and high-value added chemicals. *P. pastoris* is a workhorse for microbial synthesis of various products as it is bestowed with several key characteristics such as ease of genetic manipulation, highly efficient DNA transformation, cloning through functional complementation, the ability to yield substantial amounts of intracellular and extracellular proteins and the capability to perform post-translational modifications (glycosylation, disulfide bond formation, and proteolytic processing). Attributing to the myriad of advantages *P. pastoris* is considered as the best choice for heterologous protein expression [174].

Notably, the demand for IFN α2b has recently surged due to its utilization alongside other antiretroviral and antimalarial drugs in the treatment of SARS-CoV-2 (COVID-19) disease, highlighting its significance in combating the pandemic [175]. A previous literature has reported successful cloning of the IFN α2b gene with an N-glycosylation site (glycosylation extends the drug's presence in the bloodstream by increasing its half-life) into the pPICZα plasmid backbone. The construct was then homologously recombined into the *P. pastoris* genome, resulting in the production of humanized IFN α2b carrying the N-glycan $Man_5GlcNac_2$ attached to the 104-asparagine residue [168].

Efficient expression in *P. pastoris* is envisaged due to the presence of a strong inducible alcohol oxidase 1 (AOX1) promoter. It is interesting to note that AOX1 promoter has demonstrated successful application in heterologous protein expression with over 600 recombinant proteins being produced, reaching yields in the range of milligrams to grams per liter of culture [19]. AOX1 promoter exhibits robust induction in response to methanol while being effectively repressed by several carbon sources, such as glycerol and glucose. During the induction phase, both an abundance and a deficiency of methanol can negatively impact the efficiency of transcription and consequently the rate of protein expression. Excess methanol can supress AOX1 promoter and disrupt the metabolism due to formaldehyde buildup, impairing cellular function. Conversely, an insufficient supply of methanol can also lead to a reduction in cellular metabolic activity. It is widely acknowledged that under methanol induction, glycolysis route and TCA cycle are weakened or blocked. The sole energy source for cell/protein synthesis is

from formaldehyde dissimilation pathway [176]. Therefore, controlling the methanol feeding rate is a crucial aspect of the recombinant expression of proteins governed by the AOX1 promoter. A plethora of methanol feeding strategies has been reported for numerous recombinant proteins with appreciable yields in *P. pastoris* [177]. The feeding strategies are broadly segregated into co-feeding and -stat strategies. Co-feeding strategies involve supplementing the methanol feed with other substrates such as glycerol, sorbitol, mannitol [178,179]. Alternatively, *-stat* strategies focus on regulating the methanol feed to maintain a predetermined value for a specific cultivation parameter (such as specific growth rate, dissolved oxygen level, or methanol concentration). In chapter 3, one such strategy (μ – stat) was demonstrated with the aid of gain scheduling PID control. Results of previous chapter (section 3.4.2 and 3.4.3) indicated a significant influence of controlling μ on residual methanol concentration and substrate uptake rate of *P. pastoris*. As mentioned above, the residual methanol concentration strongly influences AOX1 transcription and other metabolic activities related to recombinant protein production. In this chapter, we intend to assess the effect of directly controlling the residual methanol concentration during the fed-batch cultivation of glycoengineered *P. pastoris*.

In this strategy, the residual methanol concentration is meticulously controlled at a predefined setpoint value during the methanol induction phase. This technique is particularly effective for process and optimizing the production of desired products [180,181]. Substrate feed rate regulation facilitated by the offline method relies on unstructured models and lacks the adaptability to handle the process dynamics involved in the fed-batch fermentation. In contrast, the online strategy employs a methanol sensor, which provides real-time information about the residual methanol concentration of the culture medium. Strict regulations for the production and marketing of biopharmaceutical products, aligned with good manufacturing practice (GMP) guidelines, are enforced by the EU and USFDA. Given the intricate nature of biotechnological microorganism cultivation processes and their sensitivity to optimal parameter control, these directives stipulate the utilization of process monitoring and control techniques [182]. Application of different process analysers such as fermentation calorimeter, exhaust gas analyser along with the methanol sensor offer a comprehensive understanding of the process dynamics. Fermentation calorimeter is a non-invasive tool capable of monitoring and measuring the heat rate signals emanated during the fermentation. The heat energy dissipation during microbial growth is the outcome of energy being either released/absorbed due to catabolic and anabolic activities. Thus, heat rate measurements are viable indicators of

ongoing metabolism and has been successfully employed as a process analyser to gain real-time insights over different microbial systems [117]. The degree of oxidation (methanol oxidation to CO_2) i.e., methanol utilization by *P. pastoris* can be tracked in real-time employing exhaust gas analysers. Overall, the application of auxiliary process analysers not only aid in gaining insights about the process but also serve as validation tools for the deployed process control strategies. Control algorithms for controlling the residual methanol concentration in the culture medium have been successfully employed for other recombinant proteins [183,184]. There are reports of using simple on/off control for controlling the methanol concentration [177]. However, the influence of residual methanol concentration on huIFN α2b productivity and *P. pastoris* growth rate and various control strategies' implications remain unexplored till date.

In this chapter, we establish a cultivation protocol based on methanol-*stat* for *P. pastoris* cultivation to produce huIFN α2b. Methanol-*stat* was accomplished by employing various control strategies, including a model-based adaptive control strategy where the tuning parameters are updated automatically based on the process dynamics. Deploying suitable tuning parameters resulted in successful control of residual methanol concentration, thereby enhancing the huIFN α2b productivity.

4.3 Material and Methods

4.3.1 Strain and Media

A glycoengineered *P. pastoris* with a Mut^+ phenotype, harbouring the plasmid SuperMan5 pep4Δ prb 1Δ procured from Biogrammatics Inc., USA, was used for expressing huIFN α2b [21]. Glycerol stocks were maintained at -80°C in yeast extract, peptone, and dextrose (YPD) media containing 20% v/v glycerol.

The composition of various media used for the cultivation of glycoengineered *P. pastoris* was as follows:

Starter culture medium (Yeast extract, Peptone, Dextrose; YPD) in g/L: yeast extract 10, peptone 20, dextrose 20.

Pre-culture medium (Yeast extract, Peptone, Glycerol; YPG) in g/L: yeast extract 10, peptone 20, glycerol 20.

Optimized fermentation medium (Basal Salt Medium, BSM) [168]: glycerol 48.84 g/L, K_2SO_4 18.2 g/L, $MgSO_4$ 7.28 g/L, KOH 4.13 g/L, $CaSO_4$. $2H_2O$ 0.93 g/L, $(NH_4)_2SO_4$ 8.42 g/L, 85% H_3PO_4 26.7 mL/L PTM4 salts 4.4 mL/L.

4.3.2 Inoculum preparation

Glycerol stocks of the strain stored at -80°C was used to streak a YPD plate to initiate colony growth. From the YPD plate, a single colony was selected and inoculated into a 5 mL test tube containing YPD media to establish a starter culture. The starter culture was then used to inoculate a baffled flask containing 300 mL of YPG media, creating a pre-culture. The pre-culture was cultivated at 30°C and 220 RPM for a duration of 24 hours, during which time it reached a final optical density at 600 nm (OD_{600}) of 4 absorbance units.

4.3.3 Supervisory Control and Data Acquisition (SCADA) for monitoring and control

A data acquisition (DAQ) hardware, cRIO 9075 (National Instruments, Austin, Texas, USA), was employed to integrate all process inputs and outputs seamlessly. The DAQ featured dedicated slots housing analog I/Os and digital I/Os. A SCADA system developed using LabVIEW (LabVIEW 13.0, National Instruments, Austin, Texas, USA) acquired all the signals. The data acquired from real timesignals was subjected to data pre-treatment by LabVIEW programming. The raw data was smoothened through the implementation of moving simple average filter with a window size of 100 data points to minimize the noise from the input signals [137,139]. Additionally, for all the acquired analog input/output signals (temparature, pH, DO, off-gas analyser, metabolic heat rate), scaling factors were selectively applied where deemed necessary to align the acquired data with calibrated values [137,185]. The pre-treated real timesignals acquired were visualized in the form of graphical plots and used in the control strategies.

4.3.4 Real-time monitoring of methanol concentration

Methanol concentration was continuously monitored in real-time through an online sensor (Raven Biotech, Vancouver, BC, Canada) immersed within the culture broth. The detection mechanism encompasses a semiconductor chip integrated with a sintered tin oxide film, designed to discern the existence of combustible vapor through a discernible elevation in electrical conductivity upon adsorption of reducing gas onto the sintered surface. This particular detector configuration finds extensive utility in gas quantification applications, owing to its inherent stability and reliability. Within the context of biological cultures, methanol, when dissolved within the culture medium, undergoes diffusion across a specialized

probe membrane (Fig. 4.1). Subsequently, a stream of meticulously filtered, desiccated air intercepts the diffused methanol, conveying it towards the detector assembly. The resultant output from the sensor module corresponds directly to the concentration of methanol present within the culture medium, thereby offering valuable insights into the metabolic activities or environmental conditions impacting the biological system under scrutiny. The output yielded by the sensor is intricately linked to the concentration of methanol present in the culture medium. A calibration curve between the detector voltage and methanol concentration ($1\ V = 0.643 * S_M$) was developed (Annexure A5) to infer the sensor data and data was continuously acquired through SCADA.

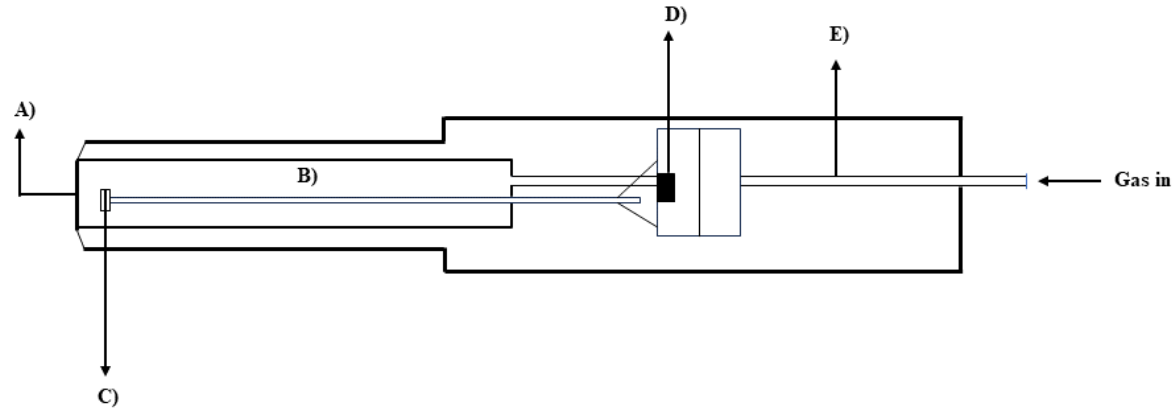


Figure 4.1: Schematic representation of methanol sensor working principle A) Membrane for sample entry and exit, B) Diffusion chamber, C) Semi-permeable membrane, D) Electronic detector, E) Carrier gas chamber. (Adapted from [163]).

4.3.5 huIFN α2b production

Cultivation of glycoengineered *P. pastoris* for huIFN α2b production was carried out in a 5 L fermentation calorimeter (Fig 4.2). In-situ sterilization of the fermentation calorimeter was accomplished using pre-filled distilled water at 121°C and 1 bar pressure for 15 minutes. The sterilized BSM was aseptically transferred into the calorimeter. Subsequently, a 10% v/v pre-culture inoculum was transferred aseptically into the calorimeter, resulting in a final working volume of 2.5 L. The high cell density cultivation was carried out at a reaction temperature of 30°C, and the pH was maintained at 5.4 by carefully adding 25% v/v ammonia and 85% v/v orthophosphoric acid as needed. Dissolved oxygen (DO) was maintained above 10% by supplying the air at 1 vvm and regulating the agitation rate between 400 to 800 RPM. Manual feeding of 10% v/v by supplying antifoam solution (silicon oil) was done in order to supress

the excessive foaming. The high cell density cultivation of *P. pastoris* underwent three distinct phases: the glycerol growth phase, the methanol adaptation phase, and the production phase. A substantial amount of biomass was generated during the growth phase, with glycerol being the primary carbon source. The depletion of the carbon source (glycerol) was evident through the real-time signals from process analysers. A drop in the heat rate signal, off-gas activity, and the rise of DO concentration indicate the complete utilization of glycerol. During the adaptation phase, the organism was exposed to small doses of methanol, and this gradual dosing enables the organism for effective adaptation into the new carbon source (glycerol to methanol). With subsequent doses, the AOX transcription elements were gradually activated [179]. The organism's response to this adaptation process could be observed through changes in metabolic heat rate and off-gas activity. The induction/production phase involved a continuous supply of methanol

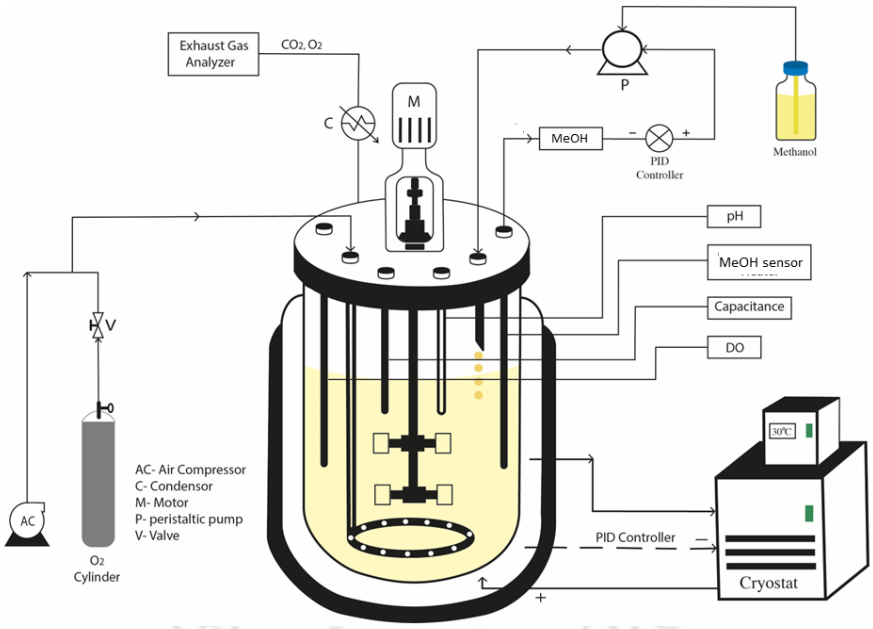


Figure 4.2: Schematic representation of *P. pastoris* cultivation adapted with methanol sensor for the production of huIFN α2b in fermentation calorimeter.

to the organism governed by a substrate balance equation. The regulation of methanol feeding was achieved by employing various strategies, which are elaborately discussed in the 'Methanol feed rate regulation' section. Throughout the cultivation process, samples were collected at regular time intervals. These collected samples served as the basis for offline analysis, allowing for the estimation and monitoring of important parameters related to the process, including biomass growth, substrate consumption, and product formation.

4.3.6 Offline analysis

Samples collected at regular intervals were stored at 4°C for further analysis. Cells from the samples were separated by centrifuging at 10,000 rpm for 10 minutes at 4°C. The resulting supernatant was collected and stored at -20°C for subsequent carbon substrate and huIFN α2b analysis. The cells were then washed twice with deionized water to eliminate any salt residues and subsequently dried at 80°C to determine dry cell weight (DCW). Cell growth was quantified through optical density measurements at 600 nm (OD_{600}), performed using a UV-visible spectrophotometer (GE Healthcare, UK). Absorbance readings were converted into DCW values (g/L) based on the following correlation developed ($1\ OD\ =\ 0.21 * DCW$). The residual glycerol and methanol concentrations present in the culture broth were analysed using High-Performance Liquid Chromatography (HPLC) through reverse-phase analytical chromatography using a Rezex RHM-monosaccharide H+ Column (Phenomenex Inc., California, USA). The mobile phase, which consisted of degassed 5 mM sulfuric acid, was set to flow through the column at a rate of 0.6 mL/min in isocratic mode. The column chamber was maintained at a temperature of 50 °C. The concentration of huIFN α2b was determined by Enzyme-Linked Immunosorbent Assay (ELISA) analysis employing a huIFN α2b ELISA kit (Mabtech AB, Sweden).

4.3.7 Process control strategy

Continuous methanol feeding is instrumental in maintaining the robust functionality of recombinant protein expression under the AOX promoter. Control strategies predominantly focus on regulating the induction phase's specific growth rate or methanol concentration, which correlates directly with protein production rates. Three different control strategies were deployed in order to assess the impact on huIFN α2b productivity. The first control strategy is based on/off control, where the controller compares the setpoint value with the measured process variable, and the error term dictates the controller's status. The PI algorithm regulates the methanol feed rate in the second control strategy, detailed in the 'Methanol feed rate regulation' section. Third control strategy regulates the methanol feed based on the automatic adaptation of controller tuning parameters. The controller tuning parameters were modelled from the substrate balance equation and the mathematical treatment of the same is described in ‘Methanol feed rate regulation’ (section 4.3.8)

4.3.8 Methanol feed rate regulation

The methanol sensor continuously monitors residual methanol concentration in the culture medium. The sensor data serves as an input to the methanol balance equation (Eq. 4.1) to

calculate desired methanol feed rate. The substrate mass balance equation for the fed-batch cultivation is described as per Eq. 4.1

$$FS_O - r_s V = V\frac{dS}{dt} + S\frac{dV}{dt} \tag{4.1}$$

Where F is substrate feed rate in mL/h, $S_O$initial substrate concentration in g/L, r_s substrate consumption rate in g/L.h, V is working volume of the bioreactor in L, S substrate concentration present inside the bioreactor in g/L.

The change in methanol concentration is represented in the Eq. 4.2

$$\frac{dS}{dt} = -r_s + \frac{F(S_O - S)}{V} \tag{4.2}$$

Eq. 4.2 can be represented in terms of deviation variables indicated by $'$ as shown in Eq. 4.3

$$\frac{dS'}{dt} = -r_s + \frac{F'(S_O - S')}{V'} \tag{4.3}$$

S' (g/L) represents the substrate concentration in terms of deviation variable, F' (mL/h) represents the feed rate in terms of deviation variable and V' (L) represents the working volume of bioreactor in terms of deviation variable.

When a sufficiently short time interval is utilized, the fluctuations in methanol concentration and methanol consumption rate can be regarded as negligible. To compute the derivative of methanol concentration, the smoothing spline technique is used [186], and methanol feeding rate is calculated as per Eq. (4.4)

$$F_{t+\Delta t} = F_t - \frac{V}{(S_O - S)}\frac{dS}{dt} \tag{4.4}$$

Where, F_t is feed rate of the substrate at time instant $'t'$ in mL/h and $F_{t+\Delta t}$ is feed rate of the substrate at next time instant $'t + \Delta t'$ in mL/h

In order to maintain the methanol concentration at the desired setpoint, the methanol feed rate needs to be regulated using feedback PI control and the feed rate equation is represented in Eq. (4.5) [183].

$$F_{t+\Delta t} = F_t - \frac{V}{(S_O - S)}\frac{dS}{dt} + K_c\left\{(\varepsilon_t - \varepsilon_{t-1}) + \frac{\Delta t}{\tau_I}\varepsilon_t\right\} \tag{4.5}$$

Where, K_c represents controller gain (dimensionless), ε_t is error at time instant $'t'$ in g/L, ε_{t-1} is error at time instant $'t - 1'$, τ_I is integral time constant in minutes.

The basis for methanol feed rate regulation using model based adaptive PI control is obtained through a substrate balance equation as given in Eq. 4.6

$$\frac{dS}{dt} = \frac{F}{V}(S_f - S) - q_s X \tag{4.6}$$

Where, S is the substrate concentration present inside the bioreactor in g/L, S_f is the substrate feed concentration in g/L, q_s is the specific substrate consumption rate in g/g.h, X is the biomass concentration in g/L.

Fed-batch fermentation is said to be run at quasi-steady state and therefore the substrate balance equation becomes as shown in Eq. (4.7)

$$\frac{dS}{dt} = 0 \tag{4.7}$$

Linearizing the differential equation (Eq. 4.6) at the operating point $'S'$ gives Eq. (4.8)

$$\frac{d\Delta S}{dt} = -\frac{F}{V}(\Delta S) + \frac{(S_f - S)(\Delta F)}{V} \tag{4.8}$$

Rearranging the linearised equation and applying the Laplace transform yields and can be represented in the form of 1st order system as represented in Eq. (4.9)

$$V(s) = \frac{\Delta S(s)}{\Delta F(s)} = \frac{K_P}{\tau s + 1} \tag{4.9}$$

$$\text{Where, } K_P = \frac{S_f - S}{F}, \tau = \frac{V}{F} \tag{4.10}$$

Where, K_P is defined as process gain and τ as time constant

The feed rate F is derived from the substrate mass balance for fed-batch fermentation and is described as shown in Eq. (4.10)

$$F = \frac{X_0 V_0 \mu}{S_{in} Y_{X/S}} \tag{4.10}$$

Where, X_0 is biomass concentration at the end of batch phase in g/L, V_0 is working volume in L, μ is specific growth rate in h^{-1}, S_{in} is the initial substrate concentration of feed entering the reactor in g/L, $Y_{X/S}$ is biomass yield coefficient in g/g.

Therefore, the tuning parameters K_P and τ change automatically according to the variations in the process behaviour. However, the process gain present in Eq. (4.10) doesn't take into

account of controller gain directly. To circumvent this shortcoming, an internal model control based tuning parameters were employed for the process gain as shown in Eq. 4.11 [59].

$$K_c = \frac{\tau}{0.5K_P} \text{ and } \tau_i = \tau \tag{4.11}$$

The model-based adaptive control thereby uses the adapted tuning parameters and regulate the feeding precisely with the help of a high precision pump (Watson Marlow 120U, Falmouth, Cornwall, U.K.) to maintain the residual methanol concentration at the desired setpoint.

4.3.9 Statistical analysis

The controller performance was assessed using a performance index; integral square error (ISE). Eq. (4.12) determines the stability of the controller by integrating the square of error over time. The control systems with minimal ISE tend to minimise the large errors.

$$ISE = \int_0^t e^2\, dt \tag{4.12}$$

Where, e is the error i.e., difference between the setpoint and process variable. All the offline samples were analysed in duplicates and their average values are reported.

4.4 Results and Discussion

4.4.1 Real-time monitoring of high cell density cultivation of glycoengineered *P. pastoris*

Deployment of real-time process analysers, including the fermentation calorimeter, CO_2 gas analyser, and methanol sensor, illustrated the augmentation in biomass production and CER and methanol utilization in *P. pastoris* growth. The coherent progression observed across heat rate and CO_2 signals, characterized by consistent elevation, can be attributed to the corresponding volumetric rise in biomass concentration during the high cell density cultivation of *P. pastoris*. Distinct patterns in the process analyser signals became evident in the thermogram and the respirogram (Fig. 4.3) as the cultivation progressed through various stages. Specifically, a decline in the real-time signals was notable at the conclusion of the glycerol batch phase (22 to 24 h). Subsequently, a gradual adaptation to methanol utilization was characterised by a slower increase in these signals, followed by an adequate depiction of induction phase dynamics. The S_m depicts the real-time concentration of methanol measured by the methanol sensor since the start of the methanol adaptation phase. During screening of optimal residual methanol concentration experiment the concentration was varied in the range viz., 1.5, 3, 5 and 7 g/L. Optimal concentration was found to be 3 g/L and all the three different

process control strategies were run at optimal concentration of 3 g/L. The measured S_m values could be found as high as 8 g/L and to the lowest, 1.5 g/L.

Since the dissipation of Gibb's free energy is inherent to all living systems, the alteration in enthalpies associated with different metabolic processes can be effectively unveiled through the use of a fermentation calorimeter. The customised fermentation calorimeter employed here exhibits a sensitivity of 6.73 mW/L [137,142]. The microbial heat rate generated during the cultivation of *P. pastoris* for the production of huIFN α2b ranged from 38 to 50 W/L, which is in accordance with earlier reported literature [179]. Combining the calorimeter signal with the CO_2 signal allows for interpreting the interrelation between the heat fluxes in bioprocesses and corresponding changes in the metabolic flux. Within the context of *P. pastoris* cultivation, monitoring the calorimetric profile alongside the respirometry provides insights into energy metabolism dynamics. Methanol consumed during the induction phase undergoes oxidation within the peroxisome, producing formaldehyde, which is then further oxidized to generate energy intermediates essential for cellular functions and this oxidation gives rise to the evolution of CO_2 gas. The induction phase is characterized by higher CER values (80 to 170 m.mol/L.h) than the glycerol batch phase, where the maximum CER generated ranges from 50 to 80 m.mol/L.h [142]. An increase in the cumulative CER production in the induction phase indicates the ongoing AOX expression activity. The induction phase, characterized by higher CER production, indicates that a significant portion of the assimilated methanol is directed primarily toward energy metabolism and less toward biomass production than the glycerol batch phase. However, an excessive or sudden increase in CER production is an indication of extensive oxidation of CO_2, which leads to the accumulation of formaldehyde in the culture broth leading to cytotoxic conditions. Application of process analysers along with offline measurements facilitated to calculate the thermochemical energetic yields. The heat yield coefficient due to biomass generation ($Y_{Q/X}$) and the heat yield coefficient due to methanol utilization ($Y_{Q/S}$) were obtained from the slopes of the plots between the cumulative heat (Q) and biomass generated, methanol consumed. The values of ($Y_{Q/X}$) under balanced growth condition remains highly consistent and subjective to the type of the undergoing metabolism. The induction phase $Y_{Q/X}$ values for all experimental runs ranged from 18.56 to 19.72 kJ/g which is in accordance with the reported literature [179,187].

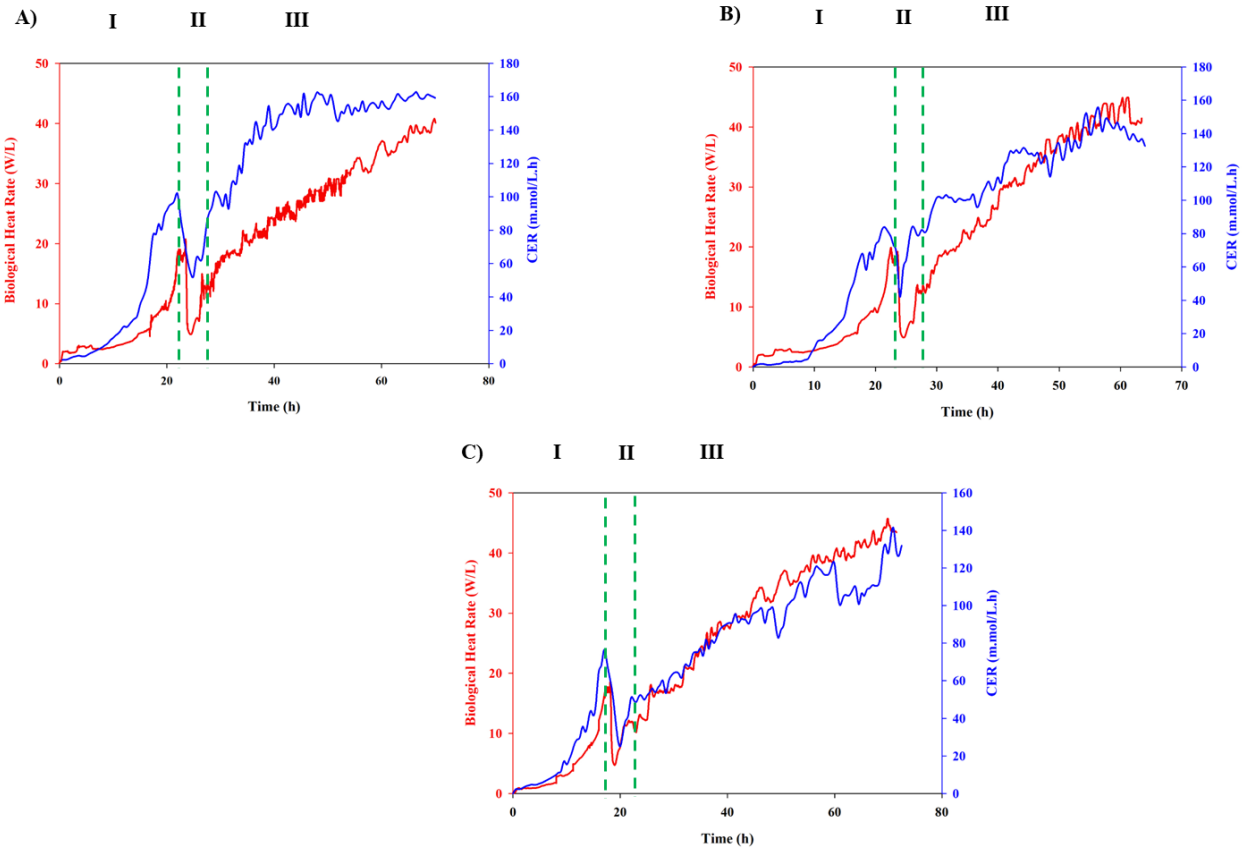


Figure 4.3: Real-time dynamic profiles of biological heat rate (Red continuous) and CER (Blue continuous) segregated into I) Glycerol batch phase, II) Adaptation phase, III) Induction phase for three different control strategies A) On/off control, B) PI control, C) Model-based adaptive PI control

Similarly, $Y_{Q/S}$ values were in the range of 662 to 702 kJ/C.mol which corroborated well the combustion enthalpy of methanol (727 kJ/C.mol). Thus, the real-time measurements from the process analysers were successful in capturing the interplay between the CER production, methanol uptake, and energy metabolism which, further enabled to regulate the methanol feeding systematically for enhanced production of huIFN α2b.

4.4.2 Screening of optimal residual methanol concentration for enhanced productivity of huIFN α2b

Given the heightened sensitivity of *P. pastoris* to methanol concentration, the production of recombinant proteins regulated by the AOX1 promoter necessitates implementing various fed-batch strategies to manage the levels of methanol concentration. Previous research has established that heuristic methanol control strategies result in substantial fluctuations of methanol concentration around the designated setpoint. These fluctuations exert a considerable

impact on the levels of recombinant protein expression and, consequently, on the specific production rate [188]. To assess the influence of methanol concentration on both biomass growth and huIFN α2b production, as well as their specific rates, methanol non-limited fed-batch (MNLFB) fermentation at methanol concentrations of 1.5, 3, 5, and 7 g/L was carried out by varying methanol concentration alone keeping other relevant factors at a constant value adapting one factor at a time (OFAT) approach. The methanol concentration at all setpoints was maintained by a simple ON/OFF control. A summary of specific rates and specific productivities is presented in Table. 4.1. A drop in μ_{off} was observed (Table 4.1) when the methanol concentration in the culture was increased step-wise from 1.5 to 7 g/L. Furthermore, a similar pattern for the methanol uptake rate q_s was obtained. Precisely, while q_s was about 0.056 g/g.h when the residual methanol concentration was 3 g/L, and a reduction to 0.032 g/g.h was observed when the residual methanol concentration was maintained at 7 g/L (Table 4.1). These lower values of both μ_{off} and q_s could potentially be attributed to the secretion of recombinant proteins seems to exhibit a metabolic burden on *P. pastoris*, leading to adverse effects on cell growth and substrate uptake capacity [189]. The biomass growth and huIFN α2b productivity significantly reduced (Table 4.1) for methanol concentration greater than 3 g/L. At higher concentrations of methanol, the accumulation of formaldehyde and hydrogen peroxide inside the cells is a common phenomenon observed in Mut^+ strains of *P. pastoris*. To minimise the toxic effect, the cell might undergo changes in physiological conditions to adapt to the higher concentration, thereby compensating for the recombinant protein production [190]. The highest huIFN α2b production and specific productivity were obtained for a residual methanol concentration of 3 g/L. Lower methanol concentrations (< 3 g/L) yielded lower product accumulation, and at higher methanol concentrations (> 5 g/L) resulted in growth inhibition. In a quest to enhance huIFN α2b titer, different control strategies were employed to maintain the optimal methanol concentration (3 g/L) during the cultivation of glycoengineered *P. pastoris.*

Table 4.1: Screening of optimal methanol concentration for the production huIFN α2b

Methanol concentration (g/L)	Biomass concentration [a] (g/L)	huIFN α2b (mg/L)	Mean q_P (mg/g.h)	Mean q_s (g/g.h)	Mean μ (h^{-1})
1.5	16.4	37.45	0.12±0.02	0.044±0.02	0.027
3	23.2	68.66	0.21±0.01	0.056±0.04	0.032
5	12.6	26.8	0.09±0.02	0.038±0.02	0.024
7	8.4	15.45	0.06±0.01	0.032±0.03	0.018

[a] Each setpoint was run for 10 hours

4.4.3 Influence of various control strategies on huIFN α2b production

Monitoring and regulation of residual methanol concentration within the cultivation medium are of paramount importance because of their dual impact on both the growth and expression of heterologous proteins. To achieve this, it is imperative to implement a suitable methanol sensor and algorithm for controlling the methanol feed rate. In this work, we implemented three different control strategies. The first control strategy was a simple on/off control, where the methanol feed was regulated based on the error value computed as the difference between the setpoint and process value measured by the methanol sensor. A deviation of ±0.75 g/L methanol concentration from the predetermined setpoint of 3 g/L was consistently observed throughout the regulated feeding period. The results indicated that, at the start of huIFN α2b production, the cells exhibited maximum methanol consumption($q_s = 0.07\ g/g.h$). and as the induction phase proceeded, there was a decline in the methanol consumption rate. Similarly, huIFN α2b productivity (q_P) was maximal at the start of regulated methanol feeding, and a drop in q_P was observed at the end of the induction phase. It is clearly evident that simple on/off control couldn't cope with the complex non-linear dynamics exhibited by *P. pastoris* even when there was a provision for direct measurement of the residual methanol concentration. It is worth noting that a significant drop in the growth rate from ($\mu_{ip} = 0.042\ h^{-1}$) to ($\mu_{ip} = 0.028\ h^{-1}$)was observed at the end of the induction phase, and a similar trend was reported for other recombinant proteins produced under Mut^+ phenotype strains [183,190]. The second control strategy was based on a simple model coupled with a PI control algorithm. The integration of the model and PI algorithm enabled the usage of tuning parameters (K_c, τ_I) facilitating a more stable control of methanol concentration. When juxtaposing the trends of q_s, q_P and μ_{ip}between the simple on/off control and the current

model-based PI control algorithm, it becomes evident that the process behaviour is akin to both strategies. However, from a quantitative point of view, a 0.75-fold increment in q_s, q_P was registered for model-based methanol feeding (Table 4.2). An average deviation of ±0.45 g/L of methanol concentration from the predetermined setpoint of 3 g/L was observed in this case. The improvement in the control behaviour can be attributed to controller tuning parameters (K_c, τ_I) and their influence on the process behaviour is entailed in the next section (section 4.4.4). Consequently, a higher huIFN α2b concentration of 202.4 mg/L was achieved under this control strategy. In the final control strategy, a model-based adaptive control algorithm regulated the methanol feed, resulting in an average deviation of ±0.20 g/L (Fig. 4.4) of methanol concentration from the predetermined setpoint of 3 g/L. The adaptation of controller tuning parameters (K_c, τ_I) to the changes in the process dynamics yielded a 1.1 and 1.9-fold increment in huIFN α2b productivity compared to model-based PI control and simple on/off control. The final huIFN α2b titer achieved with this control strategy was 244.34 mg/L. In all the control strategies dealt in this chapter, q_s, q_P and μ_{ip} values declined at the end of induction phase and similar trends were observed for other recombinant proteins produced using this methylotrophic yeast [163,177,183,186].

Table 4.2: Kinetic parameters for various control strategies implemented for the production of huIFN α2b

Control strategy	**Methanol conc. Setpoint (g/L)**	**Biomass concentration (g/L)**	**huIFN α2b (mg/L)**	**Mean q_P (mg/g.h)**	**Mean q_s (g/g.h)**	**Mean μ (h^{-1})**
On/Off	3	68.3	188.7	0.16±0.01	0.041±0.02	0.027
Model-based PI	3	75.4	202.3	0.28±0.02	0.051±0.01	0.031
Model-based adaptive PI	3	84.5	244.34	0.31±0.01	0.063±0.02	0.036

However, the decline in specific rates at the end of induction phase were relatively lower (0.8-fold) for model-based adaptive control in comparison with other two control strategies.

Previous research has demonstrated that the induction of heterologous gene expression triggers a rise in intracellular product levels, along with an elevation in binding protein levels and a concurrent reduction in μ_{ip} values. The cellular stress response is likely to play a substantial role in the context of heterologous protein production [191]. This is particularly relevant to complex proteins such as huIFN α2b, which possesses two disulfide bonds and N-glycosylation site.

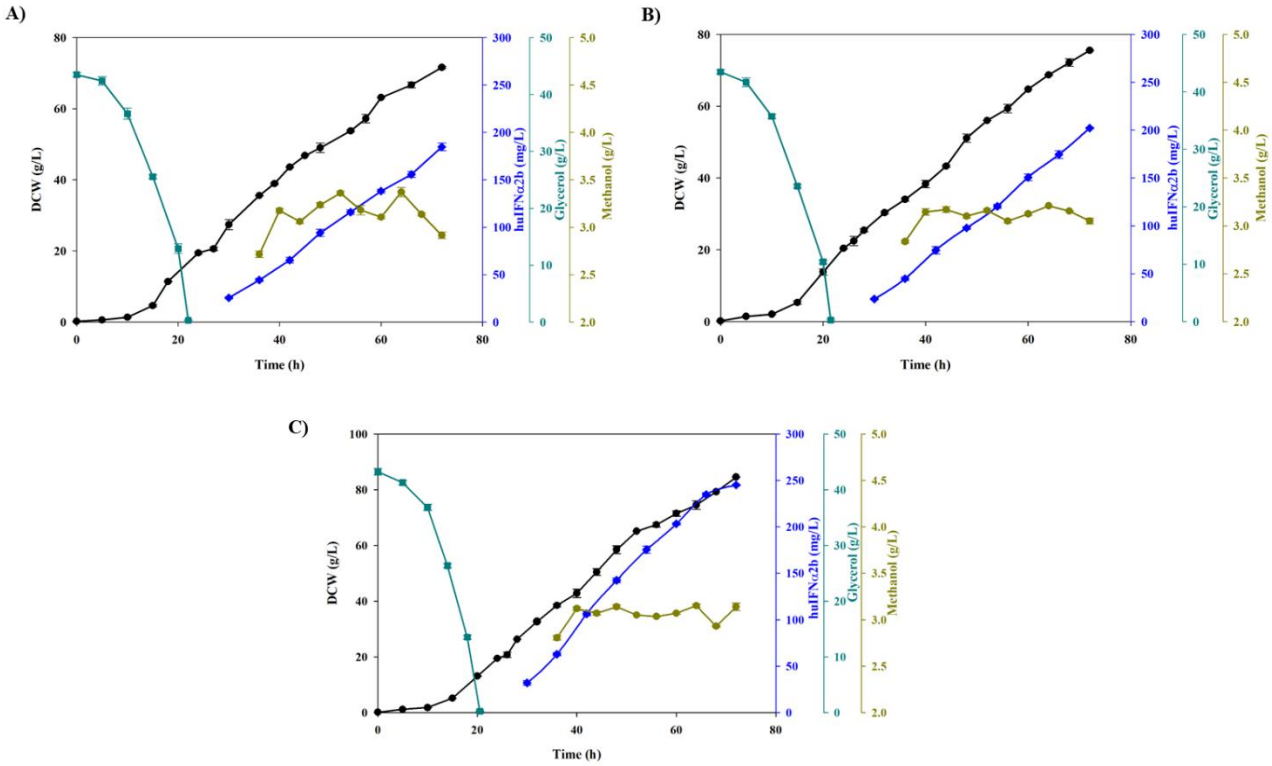


Figure 4.4: Dynamic profile representing the biomass concentration in terms of DCW (Black circles), glycerol concentration (Green squares), production of huIFN α2b (Blue diamonds), methanol concentration (Light yellow circles) for three different control strategies A) On/off control, B) PI control, C) Model-based adaptive PI control

4.4.4 Effect of tuning parameters over control characteristics and process

Simple on/off control is free from tuning parameters, and operating with such a control system is quite perilous when there is no information about process kinetics (i.e., safe residual methanol concentration). An accumulation of methanol in the culture medium is bound to happen, resulting in lower productivity as AOX1 machinery gets 3-5 times slower [176]. To counteract the accumulation of methanol and facilitate continuous feeding of methanol, a

model-based PI control with proper tuning parameters was used in this chapter. The controller gain (K_c) defines the extent to which the control action (feed rate) responds to changes in the process variable (increase or decrease). Larger values of the PI controller gain yield heightened sensitivity, while smaller values lead to decreased sensitivity. Thus, the selection of the optimal gain values is crucial for robust control applications. Open loop tests assist in identifying a range of values conducive to stable control action.

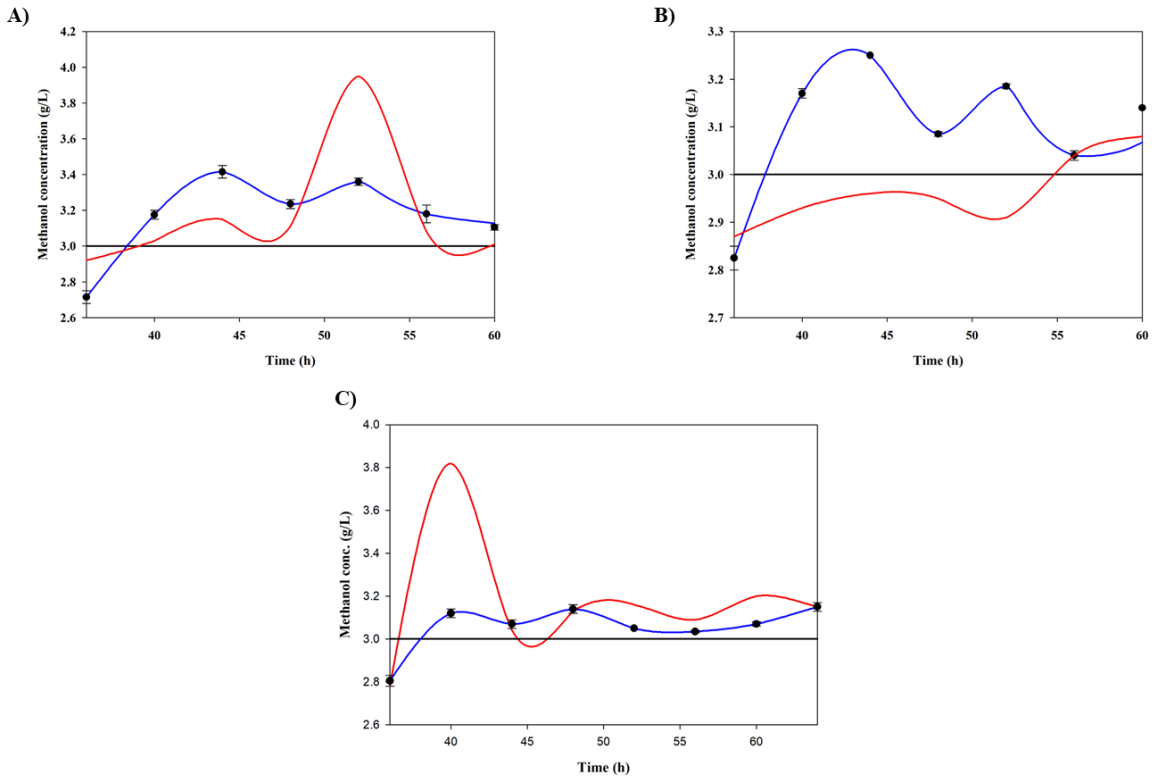


Figure 4.5: Controller performance based on online methanol concentration from sensor (Red continuous), offline methanol concentration (Blue continuous with black circles) and setpoint is represented by black continuous for three different control strategies A) On/off control, B) PI control, C) Model-based adaptive PI control.

In practical terms, longer response times tend to mitigate oscillations in the system's responses. The goal of achieving a stable output, reflected through the feed rate, can be accomplished by implementing optimal PI tuning parameters. Optimal tuning parameters were identified based on Zeigler-Nichols tuning settings and previously published literature [177,186]. The combinations of the tuning parameters implemented ($K_c = 3.5, 2.8, 1.6$ & $\tau_I = 0.02, 0.05, 0.1$). A single set of tuning parameter value was implemented for over 12 hours and based on the tracking error % which is a performance index for rating the controller

performance [19]. It is defined as the difference between setpoint and online value divided by setpoint and other performance index ISE values, the controller gain values were slowly reduced and the increase in integral time showed a positive effect on the process of reducing offset. Optimal tuning parameters yielded a better tracking error of 12.8% and lower ISE value (42.55) as compared to that of simple on/off control (Table 4.3). However, there were instances of methanol accumulation (around 52^{nd} and 64^{th} h) and controller action needed a re-adjustment of methanol feeding. In response to this, the feeding pump was prompted to abruptly decrease the feed rate at certain instances as depicted in Fig. 4.6. Model-based adaptation of tuning parameters ensured that no significant perturbations in the feeding profile (Fig. 4.6). The tuning parameters were changed dynamically with the variation in residual methanol concentration. As a result, the feed rate exhibited lesser oscillatory behaviour in comparison to PI control. Additionally, the updated tuning parameters also aided in maintaining the residual methanol concentration closer to the setpoint approximately for 18 to 22 h (Fig. 4.5). Consequently, a lower ISE and better tracking error was observed for model-based adaptive control strategy (Table 4.4). Hence, a tightly regulated methanol feeding reinstated the housekeeping activities more effectively which resulted in enhanced production of huIFN α2b.

Table 4.3: Assessment of controller performance

Control strategy	Tracking error %	ISE
On/Off	21.3 %	67.8
Model-based PI	12.8 %	42.5
Model-based adaptive	6.2 %	24.34

Table 4.4: Comparison of different controller performances on various *P. pastoris* strains

Control Strategy	*P. pastoris* strain type	Product	Tracking error %[b,c]	Reference
PID	Mut^+ strain	Growth factor	12.43	[24]
PI	Mut^S strain	Rhizopus oryzae lipase	19.34	[14]
PID	Mut^+ and Mut^S strains	ScFvA33	11.84	[27]
Model-based adaptive	SuperMan5	huIFN α2b	6.2	This work

b: Average tracking error was computed using the relationship $\frac{(S_{MeOH,sp}-S_{MeOH,mes})\times 100}{S_{MeOH,sp}}$

c: Tracking error computed from the literature

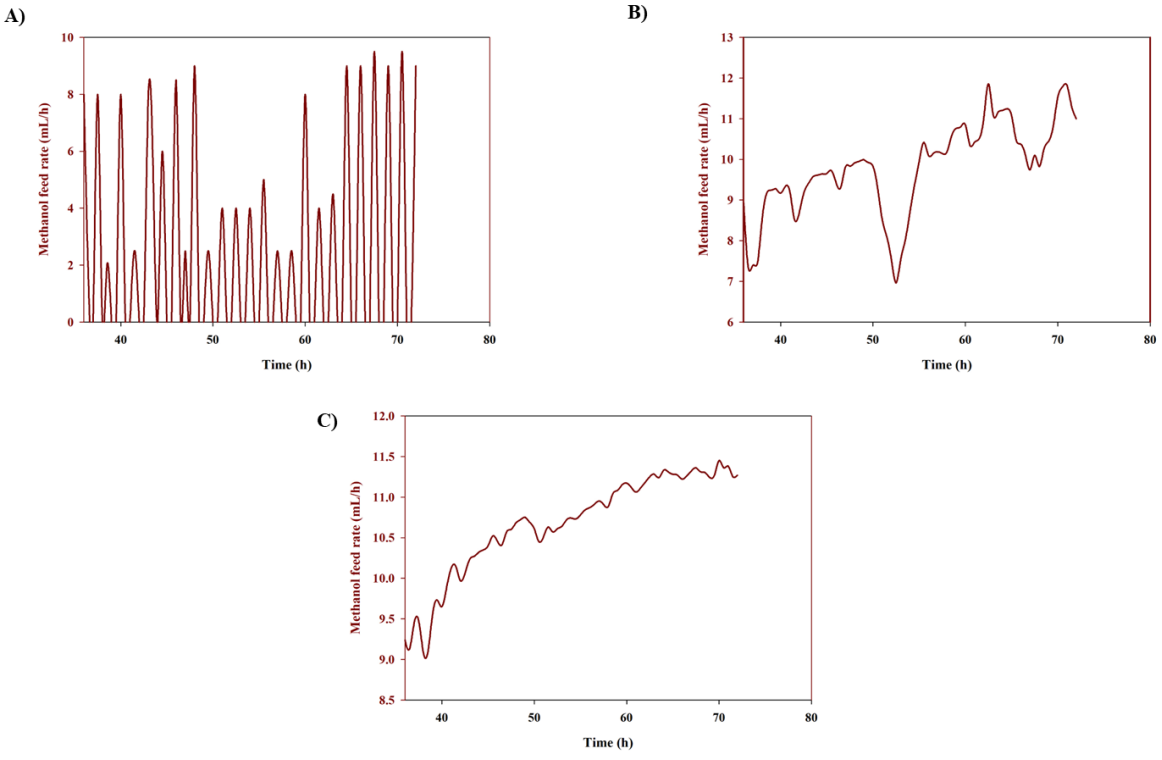
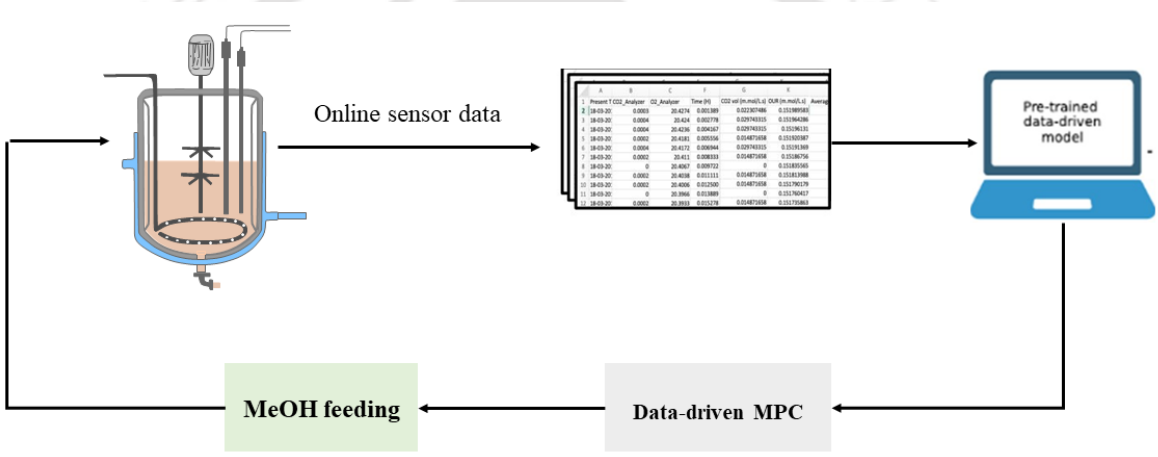


Figure 4.6: Impact of controller on methanol feeding rate A) On/off control, B) PI control C) Model-based adaptive PI control.

4.5 Conclusions

This present chapter involves a comprehensive investigation about screening optimal methanol concentration for enhanced production of huIFN α2b. The residual methanol concentration was controlled stringently at the desired setpoint with the aid of a model-based adaptive control. The approach dealt in this chapter was first to demonstrate the model-based adaptive control of substrate concentration for methanol limited fed-batch fermentation of *P. pastoris*. Tight control of methanol resulted in enhanced productivity compared to simple on/off control. Process control methodologies provide an in-depth insight into the intricate non-linear characteristics of cellular metabolism. Progress in control algorithms has the potential to enhance the accuracy of measuring process parameters, thereby contributing to refined quantification of these variables. Moreover, combinatorial approach of real-time monitoring and control accelerates the evaluation of process performance, ultimately contributing to the enhancement of product quality, adherence to regulatory standards, and commercialization of the product. Also, the model-based adaptive control framework presented in this chapter can be adopted for successful enhancement of recombinant proteins expressed in *P. pastoris*. From chapter three and four we have observed that along with the residual methanol concentration, μ had greatly influenced the state variables huIFN α2b productivity and substrate uptake rate of *P. pastoris*. A combinatorial approach of controlling both variables (μ and residual methanol concentration) by employing advanced control strategies to regulate optimal methanol feeding could result in multiple benefits as demonstrated in Chapter 5.

Chapter 5 Integrating data-driven strategies into model predictive control for enhanced production of human interferon α2b production



5.1 Abstract

The biopharmaceutical industry is witnessing exponential growth and continuously seeking for a scalable optimisation and control strategy to meet the process constraints and objectives. Model predictive control (MPC) has emerged as a robust approach to realise enhanced control over process parameters in bioprocesses. The effectiveness of MPC hinges on the availability of a robust and workable process model. Although mechanistic models are preferred, practical constraints may limit their feasibility and researchers have resorted to data-driven approaches. In this work, we demonstrate the applicability of two data-driven models namely Artificial Neural Networks (ANN) and Gaussian Process (GP) in MPC applications in fed-batch cultivation of glycoengineered *P. pastoris* for the production of huIFN α2b. The experimental verification was carried out by performing fed-batch cultivation in a fermentation calorimeter. Real-time monitoring of *P. pastoris* metabolism was facilitated by metabolic heat rate, capacitance, and exhaust gas analyser measurements. GP-based MPC demonstrated better control, efficient utilization of substrate and 1.1-fold enhanced huIFN α2b productivity compared to ANN-based MPC. Moreover, optimal feeding adapted by GP-based MPC resulted in a 14% decrease in methanol utilization compared with ANN-based MPC.

5.2 Introduction

With the advancements emerging in gene manipulation techniques various microbial systems are emerging as efficient chassis for the production of multitude of biopharmaceutical products through microbial fermentation route. Initiatives like process analytical technology (PAT) and quality by design (QbD) by US FDA and European Medicines Agencies (EMA) not only ensure safer production but also promote to enhance process knowledge by imparting real-time process monitoring and control [46]. Biopharmaceutical industries generally operate in fed-batch mode as it is bestowed with plethora of advantages. However, in industrial practice, the control objectives of fermentation process/bioreactors are limited to simple regulation loops, including pH, temperature and dissolved oxygen. Indeed, there is unequivocal evidence that the implementation of advanced control strategies in managing the biochemical state variables holds the potential to substantially enhance process performance [33]. The formidable challenge in regulating biochemical state variables stems from the inadequate availability of suitable process analysers and robust process models. The intricacies of biological systems, characterized by dynamic and nonlinear behaviour, pose hurdles in developing accurate models [37].

Numerous research endeavours have actively pursued the development of automatic regulation of biochemical state variables in fed-batch fermentation. Presently, two fundamental approaches for biological process control are prevalent: open-loop and closed-loop control. In the open-loop control, the feeding paradigm adheres to a predefined feeding rate which is independent of the output parameters from the process [32,84]. Even when deviations from the targeted state variable trajectory are identified, no corrective control actions, such as adjustments to the feeding profile, are implemented. Conversely, closed-loop control systems integrate real-time feedback, enabling dynamic adjustments to the feeding rate based on observed deviations from the desired process outcomes. The implementation of closed-loop control in bioprocess is facilitated by widely applied techniques including proportional integral derivative (PID) control, fuzzy control, artificial neural networks (ANN) based control and statistical process control. Currently, in fed-batch process the PID controller stands out as the most common closed loop control method and often implemented through indirect feedback control schemes. These schemes link the nutrient feeding rate with the measurements of process parameters viz., pH, dissolved oxygen and operate as pH-stat, DO-stat fed-batch strategies [32]. PID control demonstrates favourable results with enhanced process performance as compared to open-loop control. Despite the advantages posed by PID

controller, the usage of standalone PID controller to regulate biochemical state variables is limited due to the non-availability of sensors to accurately measure the parameters such as biomass or substrate concentration. Control strategies based on empirical models viz., ANN, fuzzy logic and statistical process control demonstrate favourable results if the system parameters do not deviate significantly from the kinetic parameter values that were used during the model creation. The principal limitation inherent in these methodologies resides in the imperative to construct a model bespoke to each novel process, mandating the acquisition of substantial quantities of statistical data [84]. As seen in previous chapters 3 and 4 different process analysers were employed for *P. pastoris* cultivation, the data obtained from these analysers could be used for the model development to predict the biochemical state variables. Moreover, we intend to use a combinatorial approach of controlling μ and optimally regulating methanol feed and determine the effect of such process strategy on huIFN α2b productivity. In order to accomplish a multi objective based control strategy it is imperative to utilize advanced control strategies.

Model Predictive Control (MPC) stands as a widely accepted and popular control technique in industrial applications [66,193]. It is an advanced model-based process control strategy characterized by the solution of an optimization problem, either online or offline, within a set of predefined constraints. The core capability of MPC lies in its ability to predict, utilizing an assumed dynamic process model, over a finite horizon. This predictive capability allows MPC to proactively optimize control actions based on future predictions, enabling the system to anticipate and respond to changes in the process dynamics. The versatility of MPC has contributed to its broad acceptance, making it a valuable tool for enhancing the efficiency and performance of various industrial processes [66]. The reference trajectory is obtained from an existing process or from a mathematical model. Classical mechanistic models or data-driven models are employed for developing the prediction model. Applications of MPC to bioprocess for various microorganism has been reported in literature [68,164,194]. Most of the experimental investigations were based on simulation studies and only a smaller number of studies were reported towards real-time implementation [195].

In this present chapter, MPC was developed with the aid of data-driven modelling. Two data-driven models (Artificial Neural Networks and Gaussian Process) were employed for the prediction of difficult to measure biochemical state variable (i.e., specific growth rate μ). Here, we demonstrate the effectiveness of the developed MPC to control fed-batch cultivation of glycoengineered *P. pastoris* for the production of huIFN α2b which is bestowed with multitude

of biological properties as discussed in section 3.2 and 4.2. This molecule's multifaceted therapeutic applications and its potential to contribute significantly to public health demands for high yield with consistent product quality. Therefore, the protein needs to be produced in larger quantities with enhanced productivity and the production method can benefit from the optimization implemented by the developed MPC.

5.3 Material and Methods

5.3.1 Strain and Media

A glycoengineered strain of *P. pastoris* with a Mut^+ phenotype, containing the plasmid SuperMan5 pep4Δ prb 1Δ procured from Biogrammatics Inc., USA, was employed for the expression of huIFN α2b [21]. Glycerol stocks were maintained at -80°C in yeast extract, peptone, and dextrose (YPD) media containing 20% v/v glycerol.

The composition of various media used for the cultivation of glycoengineered *P. pastoris* was as follows:

Starter culture medium (Yeast extract, Peptone, Dextrose; YPD) in g/L: yeast extract 10, peptone 20, dextrose 20.

Pre-culture medium (Yeast extract, Peptone, Glycerol; YPG) in g/L: yeast extract 10, peptone 20, glycerol 20.

Optimized fermentation medium (Basal Salt Medium, BSM) [168]: glycerol 48.84 g/L, K_2SO_4 18.2 g/L, $MgSO_4$ 7.28 g/L, KOH 4.13 g/L, $CaSO_4$. $2H_2O$ 0.93 g/L, $(NH_4)_2SO_4$ 8.42 g/L, 85% H_3PO_4 26.7 mL/L PTM4 salts 4.4 mL/L.

5.3.2 Inoculum preparation

Glycerol stocks of the strain, stored at -80°C, were utilized to streak a Yeast Extract Peptone Dextrose (YPD) agar plate, initiating the growth of colonies. From the YPD plate, a single colony was carefully selected and introduced into a 5 mL test tube containing YPD media, establishing a starter culture. This starter culture was subsequently employed to inoculate a baffled flask containing 300 mL of Yeast Extract Peptone Glycerol (YPG) media, thereby creating a pre-culture.

The pre-culture underwent cultivation at 30°C with agitation at 220 RPM for a period of 24 h. Over this duration, the pre-culture achieved a final optical density at 600 nm (OD_{600}) of 4 absorbance units, indicative of the growth and biomass accumulation during the cultivation process.

5.3.3 Data acquisition for monitoring and control

For the seamless integration of all process inputs and outputs, a cRIO 9075 data acquisition (DAQ) hardware, sourced from National Instruments in Austin, Texas, USA, was employed. DAQ hardware featured dedicated slots accommodating both analog and digital input/output modules. A Supervisory Control and Data Acquisition (SCADA) system, developed using LabVIEW 13.0 from National Instruments, was utilized to acquire and process all signals.

The real-time signals obtained were subjected to meticulous data pre-treatment through LabVIEW programming. To enhance signal quality, a moving simple average filter with a window size of 100 data points was implemented on the raw data. This filtering mechanism aimed to reduce noise originating from input signals, ensuring a smoother and more accurate representation of the acquired data [137]. Furthermore, for all acquired analog input and output signals, including temperature, pH, dissolved oxygen (DO), off-gas analyser data, and metabolic heat rate, scaling factors were selectively applied. This step was taken to align the acquired data with previously calibrated values, enhancing the accuracy and reliability of the measurements. The pre-treated real-time signals were then visualized through graphical plots, facilitating a comprehensive understanding of the data.

5.3.4 Gaussian Process (GP)

GP is a sophisticated statistical method used for modelling and predicting data points. It belongs to the family of Bayesian non-parametric regression techniques, which are particularly useful when dealing with limited or noisy data and when uncertainty in predictions is a concern. At its core, GP operates on the principles of Gaussian processes, which are a collection of random variables, any finite number of which have a joint Gaussian distribution. In essence, a GP defines a distribution over functions, where each function is characterized by its mean and covariance structure [196]. In GP, the training data is used to infer these parameters of the Gaussian process, allowing for the construction of a probabilistic model that captures the underlying relationship between the input and output variables. This model not only provides point estimates but also quantifies the uncertainty associated with each prediction, making it particularly valuable in decision-making processes.

One of the key advantages of GP is its flexibility in modelling complex, non-linear relationships without making explicit assumptions about the functional form of the data [196]. Additionally, GP naturally handles noisy observations and can seamlessly incorporate prior knowledge or constraints into the modelling process through the choice of appropriate

covariance functions. In this chapter along with ANN described in section 1.5.6, GP was used to predict μ.

5.3.5 Data curation for ANN and GP model training

Machine learning algorithms require enormous data to accurately predict the parameters that characterize the process (in this case, μ), accounting for the key changes during product formation phase. In the present chapter to capture the nonlinear dynamics two algorithms i.e., ANN and GP were used for the estimation of state variable. A two layered feedforward NN with an activation function of *tanh* was developed. GP is a robust regression technique for modelling nonlinear relationships inherent in input and output variables. A GP is distinctly characterized by a mean function and a covariance function, commonly known referred to as a kernel. The mean function encapsulates the average value of the function, while the covariance function delineates the interdependence among function values at distinct input points. The GP model harnesses the Radial Basis Function (RBF) kernel, a choice that significantly influences the shape of the GP [197]. Five fed-batch fermentation experiments with each experiment comprising of 28,000 data points were selected for the purpose of data-driven model development. The dataset contained online data of DO concentration, metabolic heat rate, capacitance and CER. The incorporation of historical data plays a pivotal role in assessing the model's ability to uncover analogous patterns in forthcoming runs. In the dataset, each column, symbolizing an attribute, underwent a scaling process, adjusting to a zero mean and unit variance. This normalization procedure effectively mitigates issues arising from disparate scales across different attributes in the dataset. By standardizing the data in this manner, the model is better equipped to discern meaningful patterns, fostering improved generalization and predictive capabilities across varied attributes. Before training the networks, acquired data from all experiments were clubbed and the final dataset was split randomly using the “randperm” function (available in MATLAB) into training, validation, and testing datasets in the ratio of 75%, 15%, and 15%, respectively (Annexure, A6). The root mean squared error (RMSE) of prediction was used as a performance index to decide the best network for implementation in the MPC.

5.3.6 Control methodology

The objective of this work is to maximize the huIFN α2b productivity from the high cell density cultivation of glycoengineered *P. pastoris*. By maintaining the organism at optimal growth rate and ensuring the majority of the carbon substrate is channelized towards biomass and product

formation enhance productivity is achieved. To achieve the stated objective, MPC was formulated using a cost function as stated in Eq. (5.1)

$$\phi = \sum_{i=1}^{N_P} (\hat{\mu} - \mu_{sp})^2 + \lambda \sum_{i=1}^{N_C} (F_{in} - F_{ref})^2 \tag{5.1}$$

Where, $\hat{\mu}$ is the data-driven based observed specific growth rate (h^{-1}), μ_{sp} is the setpoint of the specific growth rate (h^{-1}), F_{ref} is the pre-calculated reference feeding profile (mL/h), F_{in} is the control input (mL/h), λ is the control penalty gain N_P and N_C are prediction and control horizons.

The reference feeding trajectory is pre-calculated analytically based on the substrate balance equation and it is mathematically represented as shown in Eq. (5.2)

$$F_{ref} = \frac{X_0 V_0 \mu_{sp}}{Y_{X/S}(S_{in} - S)} e^{\mu_{sp} t} \tag{5.2}$$

Where X_0 represents initial biomass concentration (g/L), V_0 represents volume of the culture in the fermenter (L), μ_{sp} is the setpoint of specific growth rate (h^{-1}), $Y_{X/S}$ is the biomass yield per substrate consumed, S_{in} is the concentration of the substrate fed into the fermenter (g/L), S is the concentration of substrate present in the fermenter.

A prediction horizon of 1 h and control horizon of 0.5 h was chosen to minimize the quadratic cost function based on tracking error over a finite moving horizon of length N_P and the trajectory tracking (tracking a pre-defined setpoint) is achieved. If the process variables are beyond the specification, the process will deviate from its norm, and thereby hampering the productivity. Moreover, it doesn't make sense to have the process variables of interest involved in this work to have negative values. Consequently, MPC was subjected to constraints on prediction and control horizons as shown in Eq. (5.3) & (5.4) respectively.

$$0 \leq F_{in,k} \; k = 1, \ldots, N_c \tag{5.3}$$

$$0 \leq \mu_k \; k = 1, \ldots, N_P \tag{5.4}$$

MPC follows receding horizon principle, where the optimization problem subjected to the constraints is solved online at instant k. The overall workflow of the data-driven MPC is depicted in Fig. 5.1.

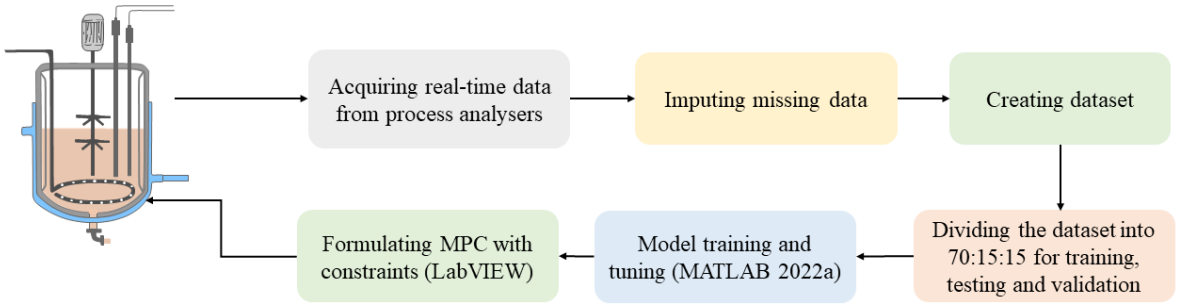


Figure 5.1: Schematic representation of data-driven MPC workflow

5.3.7 huIFN α2b production

Sterilization of fermentation calorimeter was performed in-situ using prefilled distilled water at 121°C and 1 bar pressure for 15 minutes. Following sterilization, BSM was aseptically transferred into the calorimeter. Subsequently, a 10% v/v pre-culture inoculum was aseptically introduced into the calorimeter, resulting in a final working volume of 2.5 L. The high cell density cultivation was performed at 30 °C, with pH maintained at 5.4 through the controlled addition of 25% v/v ammonia and 85% v/v orthophosphoric acid. Dissolved oxygen (DO) levels were sustained above 10% by supplying air at 1 vvm, coupled with agitation rate regulation between 400 to 800 rpm. To mitigate excessive foaming, manual feeding of 10% v/v antifoam solution (silicon oil) was added when necessary. The high cell density cultivation of *P. pastoris* is characterized with three distinct phases: the glycerol growth phase, methanol adaptation phase and the recombinant protein production phase. The glycerol growth phase resulted in significant biomass generation, utilizing glycerol as primary carbon source. Real-time signals from process analysers indicated the depletion of glycerol, manifested through a decline in heat rate signal, off-gas activity, and increase in DO concentration. During the methanol adaptation phase, gradual methanol dosing facilitated effective adaptation to the new carbon source. Activation of AOX transcription elements indicative of the adaptation phase, was reflected in changes in real-time signals of metabolic heat rate and off-gas activity. Production phase involved continuous supply of methanol, regulated by employed process control strategies as detailed in the 'control methodology section'. Throughout the cultivation, regular sample collection facilitated offline analysis, enabling the estimation and monitoring of key process parameters such as biomass growth, substrate consumption, and product formation.

5.3.8 Offline analysis

Samples collected at regular intervals were preserved at 4 °C for offline analysis. Cells were subjected to centrifugation at 10,000 RPM for 10 minutes at 4 °C for extracting cellular components. The resulting supernatant was carefully collected and stored at – 20 °C for subsequent analysis of carbon substrate and huIFN α2b. The separated cells were subjected to a double wash with deionized water to eliminate any residual salts and subsequently dried at 80 °C to determine the dry cell weight (DCW). Cell growth was quantified through optical density measurements at 600 nm (OD_{600}) using a UV visible spectrophotometer (GE Healthcare, UK). Obtained absorbance readings were converted into DCW values (g/L) based on the following correlation developed ($1\ OD\ =\ 0.21 * DCW$).

5.3.9 Performance index

The prediction performance of the data-driven models was assessed through the root mean square error (RMSE) metric which is defined as shown in Eq. (5.5)

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(y_{\text{predicted}} - y_{\text{actual}})^2}{N}} \quad (5.5)$$

Where $y_{\text{predicted}}$ is the model predicted value, y_{actual} is the experimental value and N is the number of data points.

5.4 Results and Discussion

5.4.1 Real-time monitoring of high cell density cultivation of glycoengineered *P. pastoris*

The deployment of real-time process analysers such as fermentation calorimeter, off gas analyser and dielectric spectroscopy exacerbated the monitoring and control action by accurately depicting the dynamics of high cell density cultivation of *P. pastoris.* The consistent elevation in metabolic heat rate, CER and capacitance signal corresponded to a concurrent increase in biomass concentration during high cell density cultivation. Clear demarcation of real-time signals was observed (Fig. 5.2) as the cultivation progressed through different stages. Particularly, a decline in the real-time signals was noticeable at the end of glycerol batch phase (20 to 23 hours). Subsequent to this, a gradual adaptation to methanol utilization was evidenced by a slower increase in these signals, followed by depiction of dynamics during the induction phase.

As Gibb's free energy dissipation is inherent to all living systems, the enthalpic changes associated with diverse metabolic processes can be effectively elucidated through the application of fermentation calorimeter. The customized fermentation calorimeter utilized in this work demonstrates a sensitivity of 6.73 mW/L. The maximum microbial heat generated during the cultivation of glycoengineered *P. pastoris* for the production of recombinant huIFN α2b ranged from 40 to 52 W/L, aligning with previously reported literature [179]. The integration of calorimeter signal with CO_2 signal facilitates the interpretation of the interrelation between heat fluxes in bioprocesses and corresponding alterations in the metabolic flux. In the context of *P. pastoris* cultivation, methanol consumption during the induction phase undergoes oxidation within the peroxisome, yielding formaldehyde, which is subsequently further oxidized to generate essential energy intermediates for cellular functions. This oxidation results in the evolution of CO_2 gas and as a result the induction phase is characterized by higher CER values (80-170 m.mol/L.h) compared to the glycerol phase where the maximum CER generated ranges from 50-80 m.mol/L.h. An increase in cumulative CER production during the induction phase signifies ongoing activity in the expression of AOX.

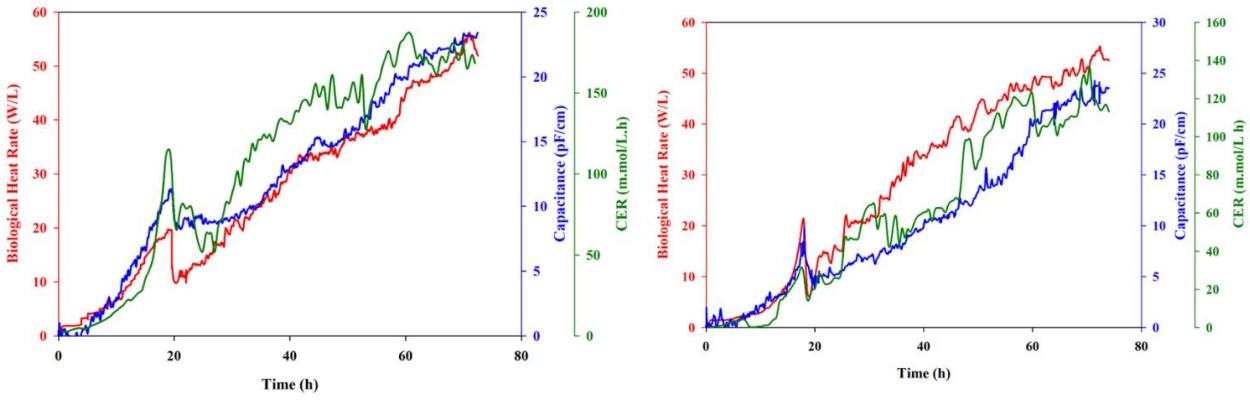


Figure 5.2: Real-time dynamic profiles from three process analysers i.e., fermentation calorimeter, capacitance probe and exhaust gas analyser. Metabolic heat rate (red continuous), capacitance (blue continuous), and CER (green continuous) A) NN-based MPC, B) GP-based MPC.

The heightened CER production in the induction phase suggests that a substantial portion of assimilated methanol is primarily directed towards energy metabolism, with less allocation to biomass production compared to glycerol batch phase. An excessive or abrupt increase in CER values serves as an indicator of extensive CO_2 oxidation, leading to the accumulation of formaldehyde in the culture broth and resulting in cytotoxic conditions [177]. Application of process analysers in conjunction with offline measurements facilitated the calculation of

thermochemical energetic yields. The metabolic heat generation arising from the balance between substrate uptake rate and its allocation to diverse metabolic reactions is an inherent characteristic of the organism, signifying the coupling between catabolic and anabolic processes. In the context of black box microbial stoichiometries, real-time heat rate measurements coupled with auxiliary process analysers provides a distinct advantage of predicting various yield coefficients. The heat yield coefficient attributed to the biomass generation ($Y_{Q/X}$) is obtained from the slope of the plot between the cumulative heat (Q) and biomass generated. The values of $Y_{Q/X}$ under conditions of balanced growth exhibit a high degree of consistency and are dependent on the specific metabolic pathways in operation. During the induction phase $Y_{Q/X}$ values for all experimental runs were in the range of 18 to 19.5 kJ/g [179]. Thus, the real-time measurements obtained from the process analysers were successful in deciphering the intricate relationship between the biomass production, metabolic heat generation and CO_2 production. The integration of real-time data from process analysers enabled to train the data-driven models for the prediction of biochemical state variable μ which facilitated the optimised feeding of methanol for enhanced production of huIFN α2b.

5.4.2 Experimental verification of ANN MPC

The architecture of the Artificial Neural Network (ANN) investigated here was denoted as 4-N-1, where "4" signifies the number of input nodes in the input layer, "N" represents the variable number of hidden nodes in the hidden layer, and "1" designates the single output node in the final layer. Each node corresponds to a neuron that computes a weighted sum of the outputs from the preceding layer. The identification of the optimal-performing network entailed the careful selection of an appropriate training algorithm and determining the number of neurons in the hidden layer. In accordance with established practices, the number of neurons was systematically chosen through an iterative process of trial and error, with the overarching objective of minimizing the prediction error between the actual target values and the corresponding values predicted by the network [198]. The optimal number of neurons was 12 with a lower RMSE being registered while training the network. The network's efficacy experienced a decline when the variable "N" was constrained to values exceeding 12 neurons (Table 5.1). A higher 'N' value tends to induce overfitting during the training phase, amplifying network complexity and computational burden. This, in turn, leads to a suboptimal generalization when applied to test data. Conversely, a lower count of neurons in the hidden layer contributes to underfitting, where the network inadequately captures the underlying

patterns in the data, further compromising its predictive performance. To assess and compare the performance of the developed model predictive control, real-time datasets that were not utilized during the training phase were employed. The ANN MPC consistently demonstrated the capability to maintain the specific growth rate within a range of 0.02 – 0.05 h^{-1} during both testing and validation datasets. Noteworthy deviations, ranging within ± 0.015 h^{-1} from the optimal specific growth rate of 0.035 h^{-1}, were observed at various intervals throughout the induction phase (Fig. 5.3). The manifestation of oscillatory behaviour in the feeding profiles as depicted in Fig. (5.4) is attributed to the intricate tuning of higher number of hyperparameters involved in development of ANN. The optimal feeding rate of methanol was determined as per Eq. (5.5) to maintain the desired setpoint values and F_o was determined from the batch-phase data as shown in Eq. (5.6).

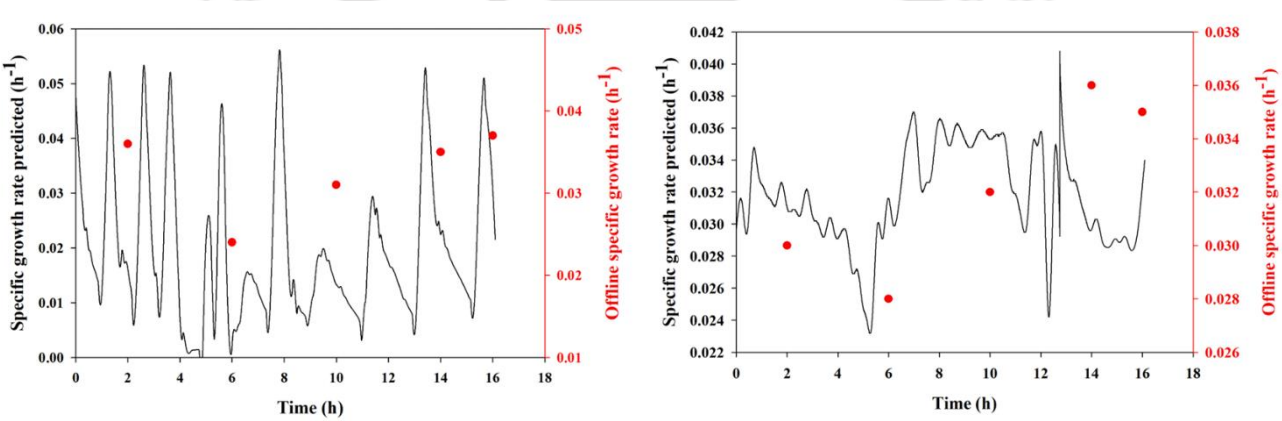


Figure 5.3: Comparison of data-driven predicted specific growth rate (black continuous) with the experimentally obtained data (red circles). A) ANN-based MPC, B) GP-based MPC.

Table 5.1: RMSE over training data of ANN for the prediction of specific growth rate

Number of neurons	RMSE
5	0.0088
7	0.0056
10	0.0024
12	0.00065
14	0.0034

The implementation of exponential feeding rates emerged as a highly effective strategy in maintaining the cellular metabolic demands. This approach facilitated concurrent substrate utilization, thereby maintaining the residual methanol concentration in close proximity to zero for a duration exceeding 15 hours. A comprehensive elucidation of the sustained and balanced utilization of substrate across diverse cellular activities is illustrated in Fig (5.5). Upon closer

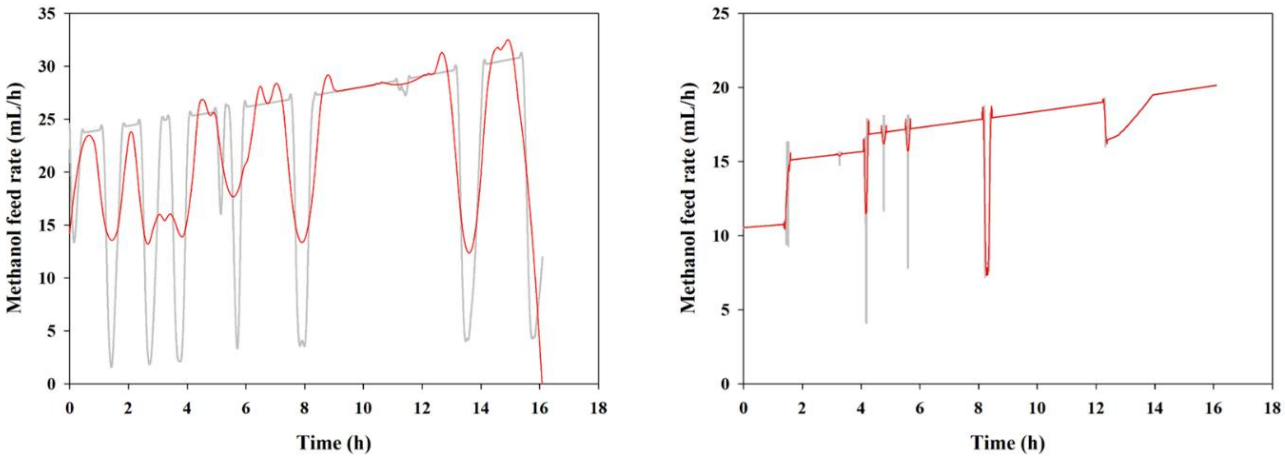


Figure 5.4: Influence of controller performance on methanol feeding rates. Pre-calculated reference feeding rate (grey continuous) compared with experimental feeding rate (red continuous). A) ANN-based MPC, B) GP-based MPC.

examination of the results, it was observed that, at the onset of huIFN α2b production, the cells demonstrated a notable maximum methanol consumption($q_s = 0.074\, g/g.h$). As the induction phase progressed, there was a discernible decline in the methanol consumption rate. This temporal evolution was accompanied by a significant reduction in the growth rate, transitioning from an initial value of $\mu_{ip} = 0.042\, h^{-1}$ to a subsequent value of $\mu_{ip} = 0.028\, h^{-1}$ at the end of the induction phase and this observed trend was mirrored with huIFN

α2b productivity. Owing to the weighted optimization of multiple control tasks within the Model Predictive Control (MPC) framework, a marginal disparity is observed at the commencement of the control process. Subsequently, a consistent decrement in the actual value ensues, attributed to the implementation of active soft constraints [199]. A similar phenomenon was observed for other recombinant proteins produced under Mut+ phenotype strains [176]. These findings align with established research suggesting that the induction of heterologous gene expression triggers an augmentation in intracellular product levels, accompanied by an elevation in binding protein levels and a concurrent reduction in μ_{ip} values. The plausible involvement of cellular stress response in the context of heterologous protein production, particularly relevant to complex proteins like huIFN α2b characterized by two disulfide bonds and a N-glycosylation site, further underscores the intricacies of the observed phenomena [21,172].

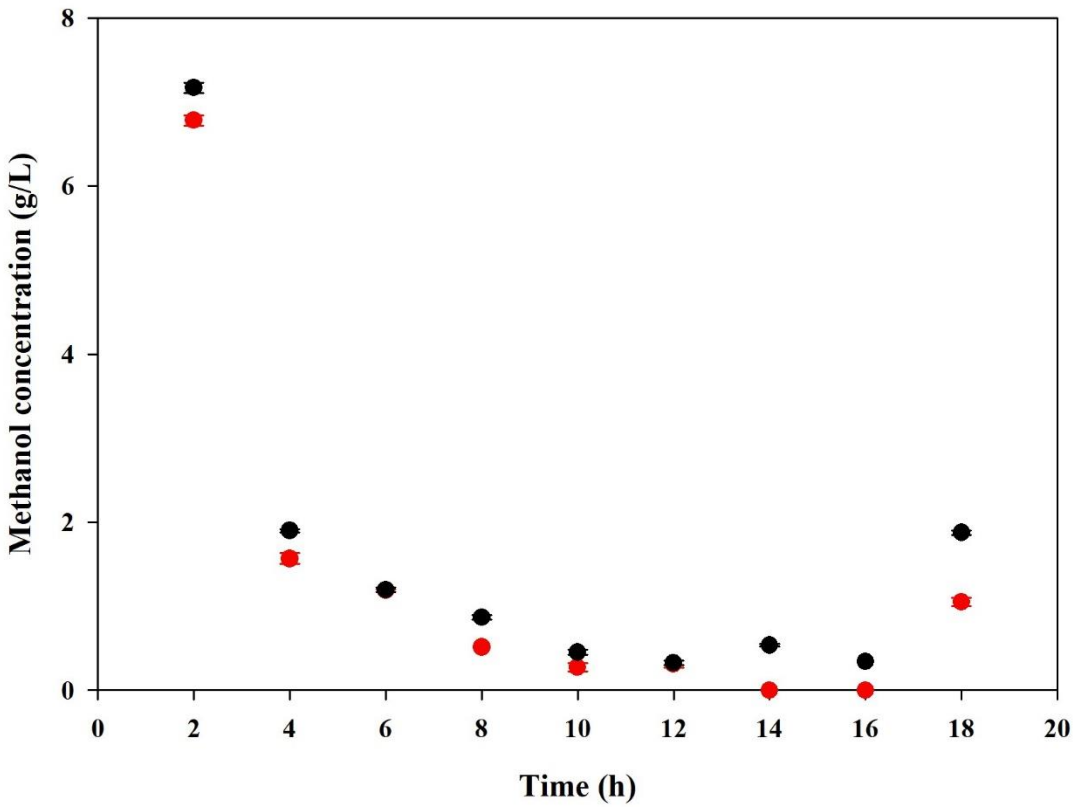


Figure 5.5: Impact of optimal methanol feeding strategy on residual methanol concentration. ANN-based MPC (black circles), GP-based MPC (red circles)

5.4.3 Experimental verification of GP MPC

GP prediction model assumes the form of black box system and the characterization arises from its capacity to furnish predictions of the output variable based on the input variables without manifesting the explicit functional relationship between them. The probabilistic approach for the regression was based on prior distribution of output variable (μ) and is characterized by a kernel function. The tuning of hyperparameter i.e., length scale was optimized to length scale = 8.57 during the training process to maximize the likelihood of the observed data and RMSE obtained at this optimal length scale was very low (0.00045). The length scale lesser than 8 resulted in wiggly function, as the model complexity increased to capture the finer details in the data. Conversely a higher length scale (> 8.57) resulted in overfitting of the data due to the evolution of smoother model.

GA MPC consistently exhibited the ability to sustain the specific growth rate within a range of 0.025 – 0.045 h^{-1} during both testing and validation datasets. Noteworthy deviations, falling within the range of ± 0.01 h^{-1} from the optimal specific growth rate of 0.035 h^{-1}, were discerned at diverse intervals throughout the induction phase (Fig. 5.3). The objective of attaining a stable output, reflected through optimal feeding of methanol was successfully realized through the implementation of GP MPC which resulted in enhanced production of huIFN α2b (Table 5.2). As illustrated in Fig (5.4), the feed rate was subjected to lesser oscillatory behaviour in GP MPC as compared to ANN MPC. Due to the straightforward approach of GP (lesser tuning of hyperparameters) as compared to ANN also contributed to the efficiency of GP MPC. The residual methanol concentration was maintained under methanol limiting conditions for longer durations and this tightly regulated methanol feeding reinstated the housekeeping activities more effectively which resulted in enhanced production of huIFN α2b. A 1.1-fold increment in huIFN α2b productivity was observed through GP MPC approach and also 14% decrease in methanol utilizing was observed due to the optimization of feeding rate.

Table 5.2: Kinetic parameters for the two adapted control methodologies for the production of huIFN α2b

Control method	Biomass concentration (g/L)	huIFN α2b (mg/L)	q_P (mg/g h)
NN-based MPC	108.2	252.32	0.55±0.03
GP-based MPC	116.65	283.56	0.61±0.03

5.4.4 Comparative assessment of various process strategies on huIFN α2b productivity

Deployment of process analysers viz., fermentation calorimeter, dielectric spectroscopy, exhaust gas analyser, and methanol sensor enabled to monitor and estimate various bioprocess state variables in real-time. Based on the type of process analyser utilized different process strategies were attempted in this thesis with the aim of enhancing huIFN α2b productivity. From chapter 3, it was observed that tightly controlling μ has a significant impact on huIFN α2b productivity. As per FDA regulation the amount of methanol present in the final drug should be lesser than 300 PPM and for this purpose a direct control on residual methanol concentration could be beneficial. In chapter 4, with the aid of methanol sensor, residual methanol concentration was controlled but a lower productivity was registered as compared with that of controlling μ. A combinatorial approach of controlling μ and optimally regulating the methanol feed aided by data-driven MPC was successful in enhancing huIFN α2b productivity (Table 5.3). Optimally regulated methanol feeding by MPC has channelized majority of methanol towards protein production and oxidation of methanol was limited which resulted in lower CER production compared to other process strategies. Moreover, data-driven based MPC resulted in lower methanol consumption which is of paramount importance at industrial scale in terms of safety and economy perspective. This underscores the importance of employing advanced control strategies that not only optimize productivity but also enhance process safety and cost-effectiveness on an industrial scale.

Table 5.3: Comparative assessment of various process strategies employed in this thesis on huIFN α2b productivity

Process parameter	Controller type	huIFN α2b (mg/L)	q_P (mg/g h)
Specific growth rate	Gain-scheduling PID (model free)	243.65	0.56±0.01
Residual methanol concentration	Model-based Adaptive PI	244.34	0.31±0.01
Both variables	Data-driven MPC	283.56	0.61±0.02

5.5 Conclusions

In this chapter, we explored the potential of controlling critical process parameters in huIFN α2b production. The utilization of data-driven models proves advantageous by providing reliable estimates of specific growth rate and thereby providing accurate input to the MPC. To effectively manage the inherent nonlinearities and discontinuities in the process MPC was leveraged to fulfil the multiple objectives and predict the process events. This proactive approach facilitated continuous product formation with a 1.03 and 1.9-fold enhanced productivity compared to μ controlling and residual methanol concentration controlling strategies. Furthermore, the optimal use of raw material (carbon substrate – methanol) with the aid of proposed MPC is of high relevance for industry as it provides the possibility of to significantly enhance the product concentration without any additional burden of substrate consumption. The obtained results highlight the potential for leveraging nonlinear approaches to achieve more robust and efficient control strategies in bioprocess applications.

Chapter 6 Conclusions and Future Perspectives

6.1 Conclusions

Research work carried out in this thesis is aligned with PAT framework for the real time monitoring and control of huIFN α2b production in glycoengineered *P. pastoris*. Deployment of process analyser viz., fermentation calorimeter, dielectric spectroscopy, exhaust gas analyser, and methanol sensor were found to be robust in elucidating the metabolic activity of *P. pastoris* growth. Implementation of various control strategies proved to be efficient in enhancing the productivity of huIFN α2b.

Chapter 2 enumerated the right choice of μ estimator model that needs to be implemented for real-time bioprocess monitoring from a calorimetric perspective. The estimator model based on the cumulative heat accurately estimated the μ values for high and low exothermic process. Data smoothing of estimator values did not influence the instantaneous heat rate-based estimator for the high exothermic system, leading to a more than 100-fold increase in performance indices. The versatility and accuracy of cumulative heat-based estimator could be deciphered from the thermodynamic perspective, with the focal point being biomass heat yield coefficient which accounts for energy dissipation from anabolic payload (i.e., formation of biomass) activities. The research outcome of this objective is highly crucial for a careful selection of a specific growth rate estimator model depending on the exothermicity and the scale of cultivation, which in turn serves as a reliable input to implement advanced control strategies in fed-batch fermentation and ultimately enhancing the yield, productivity, and quality of the product.

With an appropriate μ estimated values a stringent control of μ values in the range of 0.03 to 0.04 h^{-1} was successfully accomplished by implementing adaptive control (gain-scheduling) strategy in Chapter 3. This approach was the first to illustrate adaptive PID control with optimal tuning parameters, effectively controlling the specific growth rate at 0.035 $h^{-1}$with minor errors. Stringent control ensured that the methanol was efficiently uptaken by cells and channelized for growth and protein production, resulting in specific huIFN α2b productivity of 0.56 ± 0.01 mg/g.h. Process control approaches grants a deeper understanding of the nonlinear nature of cellular metabolism, and advances in control program may further influence the precision in quantifying process variables.

In Chapter 4, a comprehensive investigation about screening optimal methanol concentration for enhanced production of huIFN α2b. The residual methanol concentration was controlled stringently at the desired setpoint with the aid of a model-based adaptive control. This approach

was first to demonstrate the model-based adaptive PI control of substrate concentration for methanol limited fed-batch fermentation of *P. pastoris*. Tight control of methanol resulted in enhanced productivity compared to simple on/off control.

In Chapter 5, a data-driven approach to estimate the biochemical state variable was employed. The utilization of data-driven models proves advantageous by providing reliable estimates of specific growth rate and thereby providing accurate input to the MPC. To effectively manage the inherent nonlinearities and discontinuities in the process MPC was leveraged to fulfil the multiple objectives and predict the process events. This proactive approach facilitated continuous product formation with a 1.03 and 1.9-fold enhanced productivity compared to μ controlling and residual methanol concentration controlling strategies. Furthermore, the optimal use of raw material (carbon substrate – methanol) with the aid of proposed MPC is of high relevance for industry as it provides the possibility of to significantly enhance the product concentration without excess substrate consumption and thereby minimising the byproduct formation.

6.2 Future perspectives

Mechanistic modelling has been widely employed for process monitoring and control purposes. However, in the specific realm of real-time monitoring and control concerning critical quality attributes, such as the degree of glycosylation, a scarcity of a priori knowledge exists. Consequently, the development of mechanistic models for predicting the extent and pattern of glycosylation becomes a cumbersome endeavour. In the process of identifying CPPs and their intricate connections with product yield and Critical Quality Attributes CQAs, industries commonly rely on design of experiments and multivariate statistical methods. While these methodologies effectively predict the impact of CPPs on product yield, there remains an opportunity to enhance CQA prediction accuracy by capturing the intricate relationships within high-dimensional datasets.

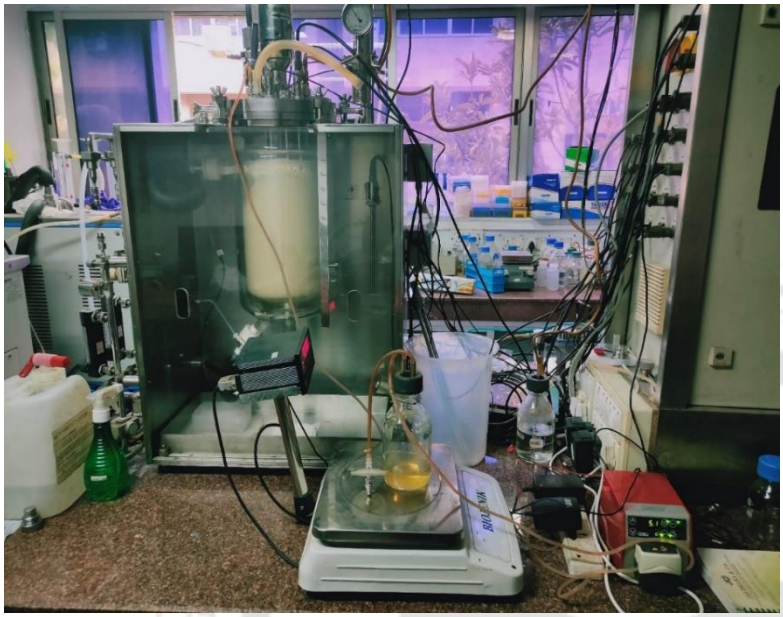
In this context, machine learning (ML) approaches present substantial potential due to their adeptness in handling nonlinear datasets. ML techniques have the capability to uncover new CPPs that could significantly contribute to predicting CQAs. Moreover, ML methods can be seamlessly integrated with mechanistic models, creating a 'hybrid ML' or 'white box ML' approach. This integration enables the elucidation of how CPPs mechanistically influence product yield and quality, thereby facilitating the rational design and control of bioprocesses. With the presence of an ample quantity of data, an effective approach involves the integration

of black-box statistical models with mechanistic modelling (hybrid model) to directly monitor glycosylation patterns and the degree of glycosylation for glycosylated protein production.

Moreover, the influence of precursor molecules on N-glycosylation necessitates attention. The optimization of the concentration and feeding rate of these precursor molecules holds the potential to enhance the homogeneity of glycan moieties. Strategic development for monitoring various stages of downstream processing for huIFN α2b and assessing real-time purity levels represents a pivotal step in establishing a comprehensive PAT enabled platform for recombinant protein production. Such an approach ensures a holistic understanding and control over the glycosylation process, contributing to the advancement of robust and high-quality bioprocessing methodologies.

Annexure

A1



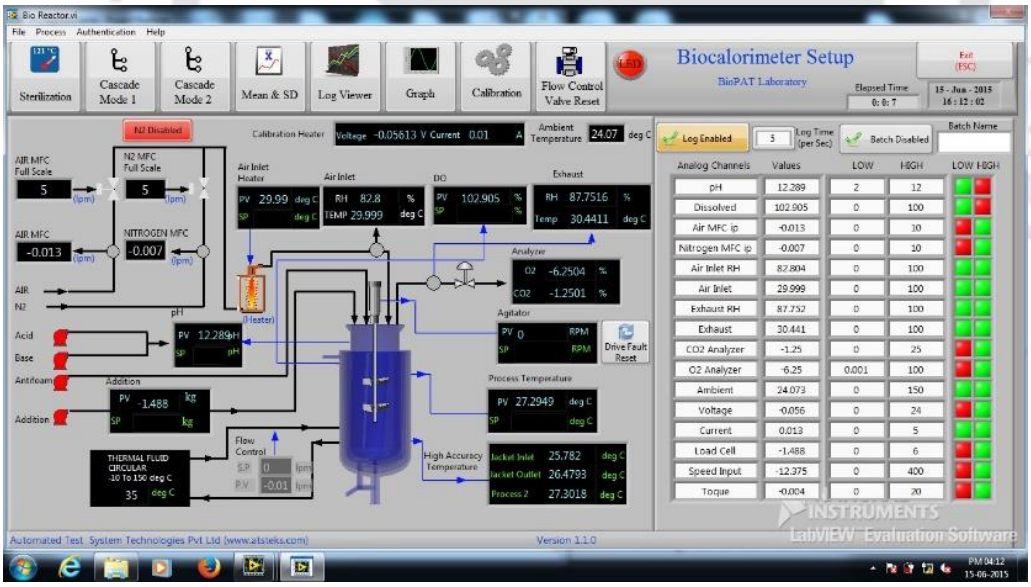


Figure A1: Biocalorimeter setup with data acquisition and SCADA at BioPAT Lab, IIT Guwahati.

A2

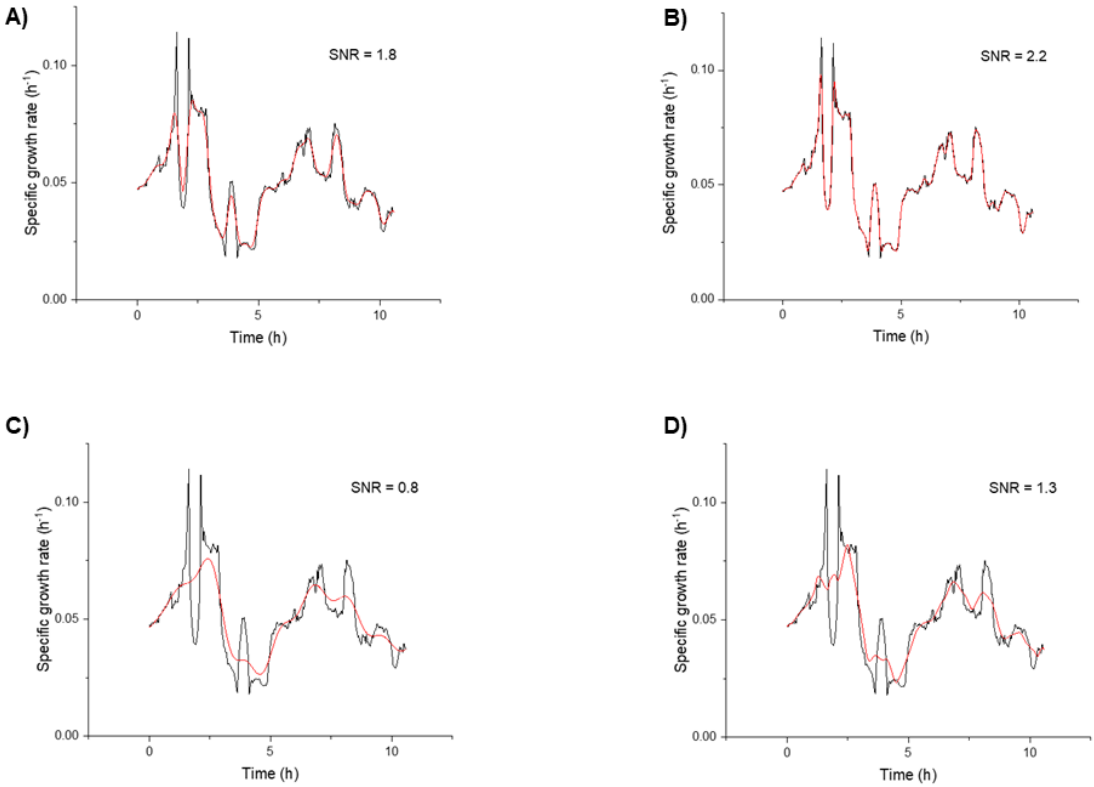


Figure A2: Effect of signal smoothing using filters, the raw signal (black continuous) and the filtered signal (red continuous) A) Moving average B) Savtizky Golay, C) Butterworth and D) Chebyshev

A3

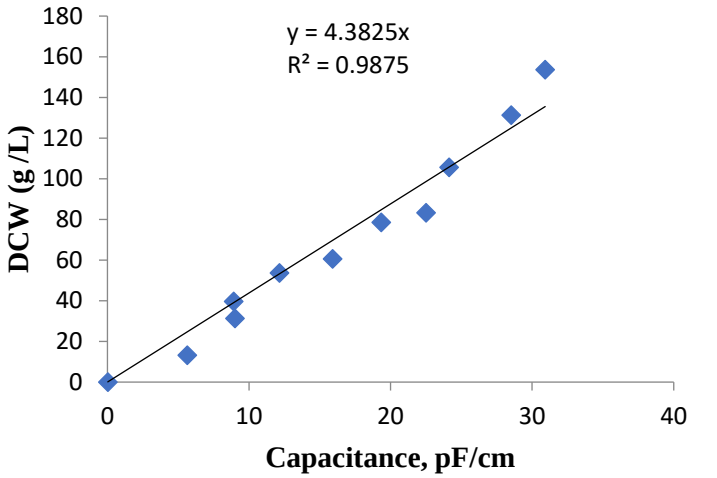


Figure A3: Correlation plot for the online capacitance (ΔC) and DCW

A4

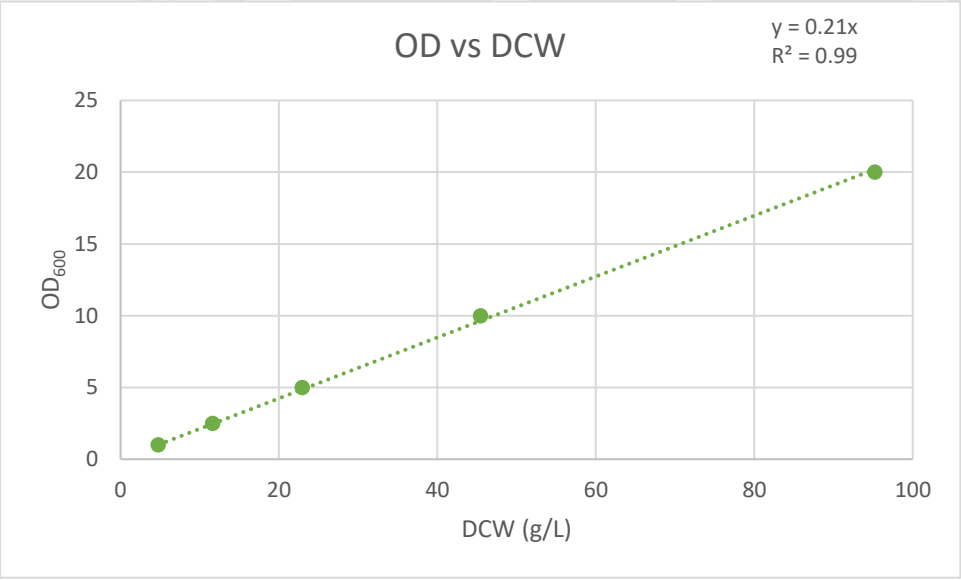


Figure A4: Standard plot for *P. pastoris* OD vs DCW

A5

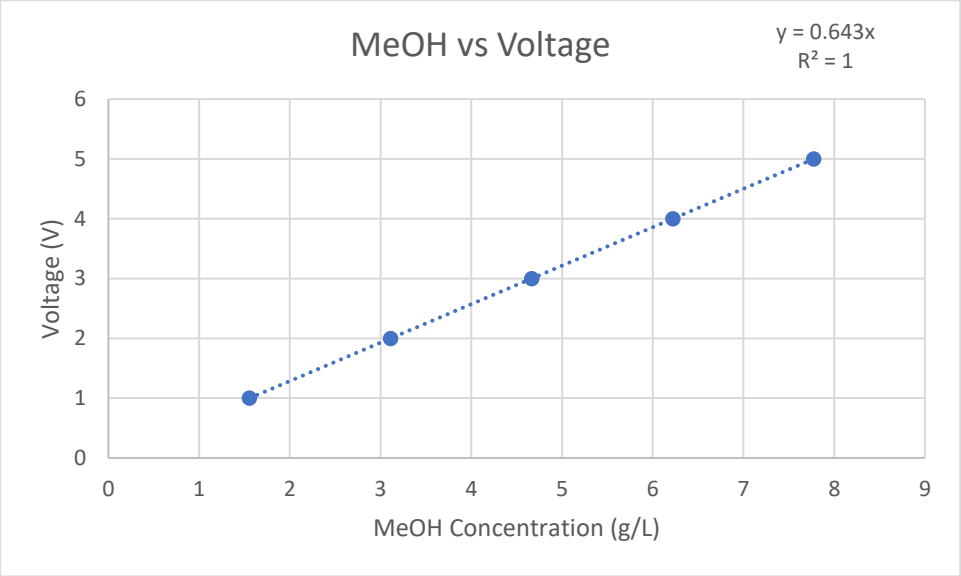


Figure A5: Calibration plot for methanol sensor.

A6. Sample MATLAB codes used for data-driven predictions

ANN based model

```
% Load your dataset into variables X (input features) and y (target labels)
% X = ... % Your input features, size: [26000, num_features]
% y = ... % Your target labels, size: [26000, 1]
% Define the architecture of the neural network
inputSize = size (X, 4); % Number of input features
hiddenSize = N; % Number of neurons in the hidden layer
outputSize = 1; % Number of output units
% Create and configure the neural network
net = feedforwardnet(hiddenSize); % Create a feedforward neural network
net.layers{1}.transferFcn = 'tanh'; % Set the transfer function for the hidden layer
net.layers{2}.transferFcn = 'purelin'; % Set the transfer function for the output layer
  % Setup Division of Data for Training, Validation, Testing
net.divideParam.trainRatio = 70/100;
net.divideParam.valRatio = 15/100;
net.divideParam.testRatio = 15/100;
% Train the neural network
net = train (net, X', y'); % Transpose X and y to match MATLAB's format
% Make predictions on new data
% Assuming newX is your new data, size: [num_samples, num_features]
newPredictions = net(newX'); % Transpose newX to match MATLAB's format
% testPredictions = net(testX');
% Calculate performance metrics (mean squared error)
```

```
% mse = mean((testPredictions - testy').^2);
```

GP based model

```
function [mu, sigma] = gaussian_process_rbf(X_train, y_train, X_test, sigma_n, sigma_f, l)


    % Kernel function (Radial Basis Function)
    kernel = @(x1, x2) sigma_f^2 * exp(-norm(x1 - x2)^2 / (2 * l^2));
    % Compute covariance matrices
    K_train = compute_kernel_matrix(X_train, kernel);
    K_test_train = compute_kernel_matrix(X_test, X_train, kernel);
    K_test = compute_kernel_matrix(X_test, kernel);
    % Add noise to the diagonal of the training covariance matrix
    K_train = K_train + sigma_n^2 * eye(size(K_train));
    % Compute mean and variance predictions
    mu = K_test_train * ((K_train \ y_train));
    sigma = diag(K_test) - sum((K_test_train / K_train) .* K_test_train, 2);
end
function K = compute_kernel_matrix(X1, X2, kernel)
    n1 = size(X1, 1);
    n2 = size(X2, 1);
    K = zeros(n1, n2);
    for i = 1:n1
        for j = 1:n2
            if nargin > 2
                K(i,j) = kernel(X1(i,:), X2(j,:));
```

```
            else
                K(i,j) = norm(X1(i,:) - X2(j,:))^2;
            end
        end
    end
end
```

Research Outcome

Publications from thesis work

1. **Pavan A**, Sivakumar R, Naresh M, Senthilkumar S*. Deployment of metabolic heat rate based soft sensor for estimation and control of specific growth rate in glycoengineered *Pichia pastoris* for human interferon alpha 2b production. ***Journal of Biotechnology*** DOI: https://doi.org/10.1016/j.jbiotec.2022.10.006 **[Impact factor 4.1]**

2. **Pavan A**, Sandhya S, Senthilkumar S*. Enhanced production of human interferon alpha 2b in glycoengineered *Pichia pastoris* by robust control of methanol feeding and implications of various control strategies. **Biochemical Engineering Journal** DOI: https://doi.org/10.1016/j.bej.2023.109152 **[Impact factor 3.9]**

3. **Pavan A**, Shika S, Senthilkumar S*. Metabolic heat based specific growth rate estimators: Does the choice of estimation model influence the state of bioprocesses? (Under Review)

4. **Pavan A,** Shikha S, Senthilkumar S*. Development of data-driven based model predictive control to design efficient feeding strategies for enhanced productivity (Under Review)

5. **Pavan A**, Senthilkumar S*. Unveiling the potential of specific growth rate control in fed-batch fermentation: Bridging the gap between product quantity and quality (Under Review)

Conferences

1. Delivered oral presentation on the topic "Metabolic heat rate-based specific growth rate estimators" at The 15th Asian Congress on Biotechnology in conjunction with The 7th International Symposium on Biomedical Engineering held on 2nd - 6th October 2022 at Bali, Indonesia (Presented Online)

2. Delivered oral presentation on the topic "Real-time estimation of specific growth rate and implications of control strategies on huIFN α2b production" at 8th International Bioprocessing India Conference - Recent Advancements & Applications in Bioprocessing for Biosimilars, Vaccines, and Bioenergy held on 16th - 18th December 2022 at National Chemical Laboratory (NCL), Pune.

Publications from collaborative work

1. Sivakumar R, **Pavan A**, Senthilkumar S*. Hybrid model based framework for soft sensing and forecasting key process variables in the production of hyaluronic acid by *Streptococcus zooepidemicus.* ***Biotechnology and Bioprocess Engineering*** doi.org/10.1007/s12257-022-0247-x **[Impact factor 3.2]**

2. Naresh M, **Pavan A**, Anjali J, Sivakumar S, Senthilkumar S*. Real-time metabolic heat-based specific growth rate sensor for monitoring and control of high molecular weight hyaluronic acid production by *Streptococcus zooepidemicus*. ***Applied Microbiology and Biotechnology*** doi.org/10.1007/s00253-022-11760-1 **[Impact factor 5.0]**

3. Subbi RRT, Ganesh N, **Pavan A**, Senthilkumar S*. Engineering precursor and co-factor supply to enhance D-pantothenic acid production in *Bacillus Megaterium.* ***Bioprocess and Biosystems Engineering*** doi.org/10.1007/s00449-022-02701-3 **[Impact factor 3.8]**

4. Srikanth Katla, Naresh Mohan, **Satya Sai Pavan**, Uttariya Pal, Senthilkumar Sivaprakasam*. Control of specific growth rate for the enhanced production of human interferon α2b in glycoengineered *Pichia pastoris*: process analytical technology guided approach, ***Journal of Chemical Technology and Biotechnology*** https://doi.org/10.1002/jctb.6118 **[Impact factor 3.4]**

5. Rengesh Balakrishnan, Subbi Rami Reddy Tadi, **Satya Sai Pavan**, Senthilkumar Sivaprakasam, Shyamkumar Rajaram*. Effect of nitrogen sources and neutralizing agents on D-lactic acid production from Kodo millet bran hydrolysate: comparative study and kinetic analysis, ***Journal of Food Science and Technology*** doi.org/10.1007/s13197-019-04124-7 **[Impact factor 3.1]**

6. Naresh Mohan, **Satya Sai Pavan**, Akshay Achar, Nivedhitha Swaminathan, Senthilkumar Sivaprakasam*. Calorespirometric investigation of *Streptococcus zooepidemicus* metabolism: Thermodynamics of anabolic payload contribution by growth and hyaluronic acid synthesis, ***Biochemical Engineering Journal***, doi.org/10.1016/j.bej.2019.107367 **[Impact factor 3.9]**

7. Srikanth Katla, **Satya Sai Pavan**, Naresh Mohan, Senthilkumar Sivaprakasam* Biocalorimetric monitoring of glycoengineered *P. pastoris* cultivation for the

production of recombinant huIFNα2b: A quantitative study based on mixed feeding strategies. ***Biotechnology Progress*** https://doi.org/10.1002/btpr.2971 **[Impact factor 2.9]**

8. Naresh Mohan, Subbi Rami Reddy Tadi, **Satya Sai Pavan**, Senthilkumar Sivaprakasam*. Deciphering the role of Critical Process Parameters governing High Molecular Weight Hyaluronic Acid synthesis in *Streptococcus zooepidemicus* cell factory. ***Applied Microbiology and Biotechnology*** doi.org/10.1007/s00253-020-10445-x **[Impact factor 5.0]**

Book chapters

1. **Pavan A,** Sandhya S, Sachin K, Senthilkumar S*. Bioprocess strategies for efficient production of pDNA vaccine production – (Book Chapter under review)
2. Naresh Mohan, **Satya Sai Pavan** & Senthilkumar Sivaprakasam* Process dynamics of Hyaluronic Acid biosynthesis through microbial route and various production strategies. Book title: Hyaluronic acid, History, Uses and Health effects. Nova Publishers Inc, New York.

References

[1] Virus interference. I. The interferon, Proc. R. Soc. Lond. B 147 (1957) 258–267. https://doi.org/10.1098/rspb.1957.0048.

[2] From interference to interferon: a brief historical introduction, Phil. Trans. R. Soc. Lond. B 299 (1982) 3–6. https://doi.org/10.1098/rstb.1982.0101.

[3] R. Bordens, S.E. Grossberg, P.P. Trotta, T.L. Nagabhushan, Molecular and biologic characterization of recombinant interferon-alpha2b, Semin Oncol 24 (1997) S9-41-S9-51.

[4] A.T. Nyman, H. Tölö, J. Parkkinen, N. Kalkkinen, Identification of nine interferon-α subtypes produced by Sendai virus-induced human peripheral blood leucocytes, Biochemical Journal 329 (1998) 295–302. https://doi.org/10.1042/bj3290295.

[5] P. Lengyel, Biochemistry of Interferons and Their Actions, Annu. Rev. Biochem. 51 (1982) 251–282. https://doi.org/10.1146/annurev.bi.51.070182.001343.

[6] S.R. Mohanty, S.S. Kupfer, V. Khiani, Treatment of chronic hepatitis B, Nat Rev Gastroenterol Hepatol 3 (2006) 446–458. https://doi.org/10.1038/ncpgasthep0550.

[7] W. Vogel, Treatment of acute hepatitis C virus infection, Journal of Hepatology 31 (1999) 189–192. https://doi.org/10.1016/S0168-8278(99)80399-7.

[8] E. Bajetta, M. Del Vecchio, P. Nova, A. Fusi, A. Daponte, M.R. Sertoli, P. Queirolo, P. Taveggia, M.G. Bernengo, S.S. Legha, B. Formisano, N. Cascinelli, Multicenter phase III randomized trial of polychemotherapy (CVD regimen) versus the same chemotherapy (CT) plus subcutaneous interleukin-2 and interferon-α2b in metastatic melanoma, Annals of Oncology 17 (2006) 571–577. https://doi.org/10.1093/annonc/mdl007.

[9] S. Li, X. Wu, P. Chen, Y. Pei, K. Zheng, W. Wang, E. Qiu, X. Zhang, Interferon-α versus interleukin-2 in Chinese patients with malignant melanoma: a randomized, controlled, trial, Anti-Cancer Drugs 30 (2019) 402–409. https://doi.org/10.1097/CAD.0000000000000741.

[10] B.J. Dezube, New therapies for the treatment of AIDS-related Kaposi sarcoma:, Current Opinion in Oncology 12 (2000) 445–449. https://doi.org/10.1097/00001622-200009000-00010.

[11] G.R. Angstreich, B.D. Smith, R.J. Jones, Treatment options for chronic myeloid leukemia: imatinib versus interferon versus allogeneic transplant:, Current Opinion in Oncology 16 (2004) 95–99. https://doi.org/10.1097/00001622-200403000-00002.

[12] J.J. Marler, J.B. Rubin, N.S. Trede, S. Connors, H. Grier, J. Upton, J.B. Mulliken, J. Folkman, Successful Antiangiogenic Therapy of Giant Cell Angioblastoma With Interferon Alfa 2b: Report of 2 Cases, Pediatrics 109 (2002) e37–e37. https://doi.org/10.1542/peds.109.2.e37.

[13] Structure and expression of human IFN-α genes, Phil. Trans. R. Soc. Lond. B 299 (1982) 7–28. https://doi.org/10.1098/rstb.1982.0102.

[14] E. Baron, S. Narula, From Cloning to a Commercial Realization: Human Alpha Interferon, Critical Reviews in Biotechnology 10 (1990) 179–190. https://doi.org/10.3109/07388559009038206.

[15] D.C. Andersen, L. Krummen, Recombinant protein expression for therapeutic applications, Current Opinion in Biotechnology 13 (2002) 117–123. https://doi.org/10.1016/S0958-1669(02)00300-2.

[16] D. Mattanovich, P. Branduardi, L. Dato, B. Gasser, M. Sauer, D. Porro, Recombinant Protein Production in Yeasts, in: A. Lorence (Ed.), Recombinant Gene Expression, Humana Press, Totowa, NJ, 2012: pp. 329–358. https://doi.org/10.1007/978-1-61779-433-9_17.

[17] P. De Brabander, E. Uitterhaegen, T. Delmulle, K. De Winter, W. Soetaert, Challenges and progress towards industrial recombinant protein production in yeasts: A review,

Biotechnology Advances 64 (2023) 108121. https://doi.org/10.1016/j.biotechadv.2023.108121.
[18] C. Bustos, J. Quezada, R. Veas, C. Altamirano, S. Braun-Galleani, P. Fickers, J. Berrios, Advances in Cell Engineering of the *Komagataella phaffii* Platform for Recombinant Protein Production, Metabolites 12 (2022) 346. https://doi.org/10.3390/metabo12040346.
[19] Z.E. Duman-Özdamar, B. Binay, Production of Industrial Enzymes via *Pichia pastoris* as a Cell Factory in Bioreactor: Current Status and Future Aspects, Protein J 40 (2021) 367–376. https://doi.org/10.1007/s10930-021-09968-7.
[20] B. Ma, X. Guan, Y. Li, S. Shang, J. Li, Z. Tan, Protein Glycoengineering: An Approach for Improving Protein Properties, Front. Chem. 8 (2020) 622. https://doi.org/10.3389/fchem.2020.00622.
[21] S. Katla, K.N.R. Yoganand, S. Hingane, C.T. Ranjith Kumar, B. Anand, S. Sivaprakasam, Novel glycosylated human interferon alpha 2b expressed in glycoengineered *Pichia pastoris* and its biological activity: N-linked glycoengineering approach, Enzyme and Microbial Technology 128 (2019) 49–58. https://doi.org/10.1016/j.enzmictec.2019.05.007.
[22] A.S. Rathore, R. Bhambure, V. Ghare, Process analytical technology (PAT) for biopharmaceutical products, Anal Bioanal Chem 398 (2010) 137–154. https://doi.org/10.1007/s00216-010-3781-x.
[23] A.S. Rathore, QbD/PAT for bioprocessing: moving from theory to implementation, Current Opinion in Chemical Engineering 6 (2014) 1–8. https://doi.org/10.1016/j.coche.2014.05.006.
[24] M. Streefland, D.E. Martens, E.C. Beuvery, R.H. Wijffels, Process analytical technology (PAT) tools for the cultivation step in biopharmaceutical production, Engineering in Life Sciences 13 (2013) 212–223. https://doi.org/10.1002/elsc.201200025.
[25] A. Guerra, M. Von Stosch, J. Glassey, Toward biotherapeutic product real-time quality monitoring, Critical Reviews in Biotechnology 39 (2019) 289–305. https://doi.org/10.1080/07388551.2018.1524362.
[26] P. Kroll, P. Sagmeister, W. Reichelt, L. Neutsch, T. Klein, C. Herwig, *Ex situ* online monitoring: application, challenges and opportunities for biopharmaceuticals processes, Pharmaceutical Bioprocessing 2 (2014) 285–300. https://doi.org/10.4155/pbp.14.22.
[27] V. Brunner, M. Siegl, D. Geier, T. Becker, Challenges in the Development of Soft Sensors for Bioprocesses: A Critical Review, Front. Bioeng. Biotechnol. 9 (2021) 722202. https://doi.org/10.3389/fbioe.2021.722202.
[28] J. Randek, C.-F. Mandenius, On-line soft sensing in upstream bioprocessing, Critical Reviews in Biotechnology 38 (2018) 106–121. https://doi.org/10.1080/07388551.2017.1312271.
[29] R. Luttmann, D.G. Bracewell, G. Cornelissen, K.V. Gernaey, J. Glassey, V.C. Hass, C. Kaiser, C. Preusse, G. Striedner, C.-F. Mandenius, Soft sensors in bioprocessing: A status report and recommendations, Biotechnology Journal 7 (2012) 1040–1048. https://doi.org/10.1002/biot.201100506.
[30] C. Mandenius, R. Gustavsson, Mini-review: soft sensors as means for PAT in the manufacture of bio-therapeutics, J of Chemical Tech & Biotech 90 (2015) 215–227. https://doi.org/10.1002/jctb.4477.
[31] A.S. Rathore, H. Winkle, Quality by design for biopharmaceuticals, Nat Biotechnol 27 (2009) 26–34. https://doi.org/10.1038/nbt0109-26.
[32] M. Mahmoodi, E. Nassireslami, Control algorithms and strategies of feeding for fed-batch fermentation of *Escherichia coli*: a review of 40 years of experience, Preparative Biochemistry & Biotechnology 52 (2022) 823–834. https://doi.org/10.1080/10826068.2021.1998112.

[33] M.M. Schuler, I.W. Marison, Real-time monitoring and control of microbial bioprocesses with focus on the specific growth rate: current state and perspectives, Appl Microbiol Biotechnol 94 (2012) 1469–1482. https://doi.org/10.1007/s00253-012-4095-z.
[34] S. Gnoth, M. Jenzsch, R. Simutis, A. Lübbert, Control of cultivation processes for recombinant protein production: a review, Bioprocess Biosyst Eng 31 (2008) 21–39. https://doi.org/10.1007/s00449-007-0163-7.
[35] Galvanauskas, Simutis, Levišauskas, Urniežius, Practical Solutions for Specific Growth Rate Control Systems in Industrial Bioreactors, Processes 7 (2019) 693. https://doi.org/10.3390/pr7100693.
[36] S. Mitra, G.S. Murthy, Bioreactor control systems in the biopharmaceutical industry: a critical perspective, Syst Microbiol and Biomanuf 2 (2022) 91–112. https://doi.org/10.1007/s43393-021-00048-6.
[37] H. Narayanan, M.F. Luna, M. Von Stosch, M.N. Cruz Bournazou, G. Polotti, M. Morbidelli, A. Butté, M. Sokolov, Bioprocessing in the Digital Age: The Role of Process Models, Biotechnology Journal 15 (2020) 1900172. https://doi.org/10.1002/biot.201900172.
[38] W.N. Reichelt, P. Thurrold, M. Brillmann, J. Kager, J. Fricke, C. Herwig, Generic biomass estimation methods targeting physiologic process control in induced bacterial cultures, Engineering in Life Sciences 16 (2016) 720–730. https://doi.org/10.1002/elsc.201500182.
[39] B. Henes, B. Sonnleitner, Controlled fed-batch by tracking the maximal culture capacity, Journal of Biotechnology 132 (2007) 118–126. https://doi.org/10.1016/j.jbiotec.2007.04.021.
[40] M. Jenzsch, S. Gnoth, M. Beck, M. Kleinschmidt, R. Simutis, A. Lübbert, Open-loop control of the biomass concentration within the growth phase of recombinant protein production processes, Journal of Biotechnology 127 (2006) 84–94. https://doi.org/10.1016/j.jbiotec.2006.06.004.
[41] M. Jenzsch, S. Gnoth, M. Kleinschmidt, R. Simutis, A. Lübbert, Improving the batch-to-batch reproducibility in microbial cultures during recombinant protein production by guiding the process along a predefined total biomass profile, Bioprocess Biosyst Eng 29 (2006) 315–321. https://doi.org/10.1007/s00449-006-0080-1.
[42] M. Aehle, A. Kuprijanov, S. Schaepe, R. Simutis, A. Lübbert, Increasing batch-to-batch reproducibility of CHO cultures by robust open-loop control, Cytotechnology 63 (2011) 41–47. https://doi.org/10.1007/s10616-010-9320-y.
[43] S. Schaepe, A. Kuprijanov, R. Simutis, A. Lübbert, Avoiding overfeeding in high cell density fed-batch cultures of *E. coli* during the production of heterologous proteins, Journal of Biotechnology 192 (2014) 146–153. https://doi.org/10.1016/j.jbiotec.2014.09.002.
[44] S. Chenikher, J.S. Guez, F. Coutte, M. Pekpe, P. Jacques, J.P. Cassar, Control of the specific growth rate of *Bacillus subtilis* for the production of biosurfactant lipopeptides in bioreactors with foam overflow, Process Biochemistry 45 (2010) 1800–1807. https://doi.org/10.1016/j.procbio.2010.06.001.
[45] U. Yüzgeç, M. Türker, A. Hocalar, On-line evolutionary optimization of an industrial fed-batch yeast fermentation process, ISA Transactions 48 (2009) 79–92. https://doi.org/10.1016/j.isatra.2008.09.001.
[46] A.S. Rathore, S. Mishra, S. Nikita, P. Priyanka, Bioprocess Control: Current Progress and Future Perspectives, Life 11 (2021) 557. https://doi.org/10.3390/life11060557.
[47] J. Lee, S.Y. Lee, S. Park, A.P.J. Middelberg, Control of fed-batch fermentations, Biotechnology Advances 17 (1999) 29–48. https://doi.org/10.1016/S0734-9750(98)00015-9.

[48] S.K. Yoon, W.K. Kang, T.H. Park, Fed-batch operation of recombinant *Escherichia coli* containing trp promoter with controlled specific growth rate, Biotechnology and Bioengineering 43 (1994) 995–999.
[49] F. Wang, G. Du, Y. Li, J. Chen, Regulation of CCR in the γ-CGTase production from *Bacillus macorous* by the specific cell growth rate control, Enzyme and Microbial Technology 39 (2006) 1279–1285. https://doi.org/10.1016/j.enzmictec.2006.03.014.
[50] L. Jae-Ho, Y.-H. Choi, S.-K. Kang, H.-H. Park, I.-B. Kwon, Production of human leukocyte interferon in *Escherichia coli* by control of growth rate in fed-batch fermentation, Biotechnology Letters 11 (1989) 695–698.
[51] M. Jenzsch, R. Simutis, A. Luebbert, Generic model control of the specific growth rate in recombinant Escherichia coli cultivations, Journal of Biotechnology 122 (2006) 483–493.
[52] H. De Battista, J. Picó, E. Picó-Marco, Nonlinear PI control of fed-batch processes for growth rate regulation, Journal of Process Control 22 (2012) 789–797. https://doi.org/10.1016/j.jprocont.2012.02.011.
[53] D.E. Seborg, T.F. Edgar, D.A. Mellichamp, F.J. Doyle III, Process dynamics and control, John Wiley & Sons, 2016.
[54] Fonseca, R. R., Franco, I. C., & Da Silva, F. V. (2016, August). Bioreactor temperature control using a generic fuzzy feedforward control system. In Proceedings of the 15th IASTED International Conference Intelligent Systems and Control (ISC 2016), Campinas, Brazil (pp. 16-18)., (n.d.).
[55] R.R. Fonseca, R.R. Sencio, I.C. Franco, F.V. Da Silva, An Adaptive Fuzzy Feedforward-Feedback Control System Applied to a Saccharification Process, Chemical Product and Process Modeling 13 (2018). https://doi.org/10.1515/cppm-2018-0014.
[56] A. Escalante-Sánchez, J. Barrera-Cortés, H.M. Poggi-Varaldo, T. Ponce-Noyola, I.S. Baruch, A soft sensor based on online biomass measurements for the glucose estimation and control of fed-batch cultures of *Bacillus thuringiensis*, Bioprocess Biosyst Eng 41 (2018) 1471–1484. https://doi.org/10.1007/s00449-018-1975-3.
[57] X.-C. Zhang, A. Visala, A. Halme, P. Linko, Functional state modeling and fuzzy control of fed-batch aerobic baker's yeast process, Journal of Biotechnology 37 (1994) 1–10. https://doi.org/10.1016/0168-1656(94)90196-1.
[58] K. Ye, S. Jin, K. Shimizu, Fuzzy neural network for the control of high cell density cultivation of recombinant *Escherichia coli*, Journal of Fermentation and Bioengineering 77 (1994) 663–673. https://doi.org/10.1016/0922-338X(94)90151-1.
[59] M. Butkus, J. Repšytė, V. Galvanauskas, Fuzzy Logic-Based Adaptive Control of Specific Growth Rate in Fed-Batch Biotechnological Processes. A Simulation Study, Applied Sciences 10 (2020) 6818. https://doi.org/10.3390/app10196818.
[60] Hisbullah, H. M., R. K., Comparative evaluation of various control schemes for fed-batch fermentation, Bioprocess and Biosystems Engineering 24 (2002) 309–318. https://doi.org/10.1007/s00449-001-0272-7.
[61] S. Duan, Z. Shi, H. Feng, Z. Duan, Z. Mao, An on-line adaptive control based on DO/pH measurements and ANN pattern recognition model for fed-batch cultivation, Biochemical Engineering Journal 30 (2006) 88–96. https://doi.org/10.1016/j.bej.2006.02.007.
[62] R. Oliveira, R. Simutis, S. Feyo De Azevedo, Design of a stable adaptive controller for driving aerobic fermentation processes near maximum oxygen transfer capacity, Journal of Process Control 14 (2004) 617–626. https://doi.org/10.1016/j.jprocont.2004.01.003.
[63] Z.I.T.A. Soons, J.A. Voogt, G. Van Straten, A.J.B. Van Boxtel, Constant specific growth rate in fed-batch cultivation of *Bordetella pertussis* using adaptive control, Journal of Biotechnology 125 (2006) 252–268. https://doi.org/10.1016/j.jbiotec.2006.03.005.

[64] V. Galvanauskas, R. Simutis, V. Vaitkus, Adaptive Control of Biomass Specific Growth Rate in Fed-Batch Biotechnological Processes. A Comparative Study, Processes 7 (2019) 810. https://doi.org/10.3390/pr7110810.
[65] I.D. Landau, R. Lozano, M. M'Saad, A. Karimi, Adaptive Control: Algorithms, Analysis and Applications, Springer London, London, 2011. https://doi.org/10.1007/978-0-85729-664-1.
[66] M.G. Forbes, R.S. Patwardhan, H. Hamadah, R.B. Gopaluni, Model Predictive Control in Industry: Challenges and Opportunities, IFAC-PapersOnLine 48 (2015) 531–538. https://doi.org/10.1016/j.ifacol.2015.09.022.
[67] A. Kuprijanov, S. Schaepe, R. Simutis, A. Lübbert, Model predictive control made accessible to professional automation systems in fermentation technology, Biosyst. Inf. Technol 2 (2013) 26–31.
[68] M. Aehle, K. Bork, S. Schaepe, A. Kuprijanov, R. Horstkorte, R. Simutis, A. Lübbert, Increasing batch-to-batch reproducibility of CHO-cell cultures using a model predictive control approach, Cytotechnology 64 (2012) 623–634.
[69] S. Ulonska, D. Waldschitz, J. Kager, C. Herwig, Model predictive control in comparison to elemental balance control in an *E. coli* fed-batch, Chemical Engineering Science 191 (2018) 459–467. https://doi.org/10.1016/j.ces.2018.06.074.
[70] J. Kager, A. Tuveri, S. Ulonska, P. Kroll, C. Herwig, Experimental verification and comparison of model predictive, PID and model inversion control in a *Penicillium chrysogenum* fed-batch process, Process Biochemistry 90 (2020) 1–11. https://doi.org/10.1016/j.procbio.2019.11.023.
[71] O.I. Abiodun, A. Jantan, A.E. Omolara, K.V. Dada, N.A. Mohamed, H. Arshad, State-of-the-art in artificial neural network applications: A survey, Heliyon 4 (2018) e00938. https://doi.org/10.1016/j.heliyon.2018.e00938.
[72] S. Rómoli, M. Serrano, F. Rossomando, J. Vega, O. Ortiz, G. Scaglia, Neural Network-Based State Estimation for a Closed-Loop Control Strategy Applied to a Fed-Batch Bioreactor, Complexity 2017 (2017) 1–16. https://doi.org/10.1155/2017/9391879.
[73] C. Murugan, P. Natarajan, Estimation of fungal biomass using multiphase artificial neural network based dynamic soft sensor, Journal of Microbiological Methods 159 (2019) 5–11. https://doi.org/10.1016/j.mimet.2019.02.002.
[74] T. Tavasoli, S. Arjmand, S.O. Ranaei Siadat, S.A. Shojaosadati, A. Sahebghadam Lotfi, A robust feeding control strategy adjusted and optimized by a neural network for enhancing of alpha 1-antitrypsin production in *Pichia pastoris*, Biochemical Engineering Journal 144 (2019) 18–27. https://doi.org/10.1016/j.bej.2019.01.005.
[75] A. Beiroti, S.N. Hosseini, M.R. Aghasadeghi, D. Norouzian, Comparative study of μ -stat methanol feeding control in fed-batch fermentation of *Pichia pastoris* producing HBsAg: an open-loop control *versus* recurrent artificial neural network-based feedback control, J. Chem. Technol. Biotechnol. 94 (2019) 3924–3931. https://doi.org/10.1002/jctb.6192.
[76] J. Peng, F. Meng, Y. Ai, Time-dependent fermentation control strategies for enhancing synthesis of marine bacteriocin 1701 using artificial neural network and genetic algorithm, Bioresource Technology 138 (2013) 345–352.
[77] Y. Cheng, X. Bi, Y. Xu, Y. Liu, J. Li, G. Du, X. Lv, L. Liu, Artificial intelligence technologies in bioprocess: Opportunities and challenges, Bioresource Technology 369 (2023) 128451. https://doi.org/10.1016/j.biortech.2022.128451.
[78] M. Aehle, S. Schaepe, A. Kuprijanov, R. Simutis, A. Lübbert, Simple and efficient control of CHO cell cultures, Journal of Biotechnology 153 (2011) 56–61. https://doi.org/10.1016/j.jbiotec.2011.03.006.
[79] A. Gautam, V. Sahai, S. Mishra, Development of a dual specific growth rate-based fed-batch process for production of recombinant human granulocyte colony-stimulating

factor in *Pichia pastoris*, Bioprocess Biosyst Eng 44 (2021) 103–112. https://doi.org/10.1007/s00449-020-02427-0.
[80] R. Hu, R. Cui, Q. Xu, D. Lan, Y. Wang, Controlling Specific Growth Rate for Recombinant Protein Production by *Pichia pastoris* Under Oxidation Stress in Fed-batch Fermentation, Appl Biochem Biotechnol 194 (2022) 6179–6193. https://doi.org/10.1007/s12010-022-04022-3.
[81] L. Jia, S. Rao, H. Li, C. Wu, Q. Wang, T. Li, A. Huang, Enhancing HSA-GCSFm fusion protein production by *Pichia pastoris* with an on-line model-based exponential and DO-stat control modes, Biochemical Engineering Journal 177 (2022) 108262. https://doi.org/10.1016/j.bej.2021.108262.
[82] P.P. Jacobs, M. Inan, N. Festjens, J. Haustraete, A. Van Hecke, M.M. Meagher, R. Contreras, N. Callewaert, Fed-batch fermentation of GM-CSF-producing glycoengineered *Pichia pastoris* under controlled specific growth rate, Microb Cell Fact 9 (2010) 93. https://doi.org/10.1186/1475-2859-9-93.
[83] W. Liu, H. Xiang, T. Zhang, X. Pang, J. Su, H. Liu, B. Ma, L. Yu, Development of a New High-Cell Density Fermentation Strategy for Enhanced Production of a Fungus β-Glucosidase in *Pichia pastoris*, Front. Microbiol. 11 (2020) 1988. https://doi.org/10.3389/fmicb.2020.01988.
[84] L. Mears, S.M. Stocks, G. Sin, K.V. Gernaey, A review of control strategies for manipulating the feed rate in fed-batch fermentation processes, Journal of Biotechnology 245 (2017) 34–46. https://doi.org/10.1016/j.jbiotec.2017.01.008.
[85] F. Ibáñez, P.A. Saa, L. Bárzaga, M.A. Duarte-Mermoud, M. Fernández-Fernández, E. Agosin, J.R. Pérez-Correa, Robust control of fed-batch high-cell density cultures: a simulation-based assessment, Computers & Chemical Engineering 155 (2021) 107545. https://doi.org/10.1016/j.compchemeng.2021.107545.
[86] L. Habegger, K. Rodrigues Crespo, M. Dabros, Preventing Overflow Metabolism in Crabtree-Positive Microorganisms through On-Line Monitoring and Control of Fed-Batch Fermentations, Fermentation 4 (2018) 79. https://doi.org/10.3390/fermentation4030079.
[87] J. Pinsach, C. De Mas, J. López-Santín, A simple feedback control of *Escherichia coli* growth for recombinant aldolase production in fed-batch mode, Biochemical Engineering Journal 29 (2006) 235–242. https://doi.org/10.1016/j.bej.2006.01.001.
[88] H. De Battista, J. Picó, E. Picó-Marco, V. Mazzone, ADAPTIVE SLIDING MODE CONTROL OF FED-BATCH PROCESSES USING SPECIFIC GROWTH RATE ESTIMATION FEEDBACK, IFAC Proceedings Volumes 40 (2007) 127–132. https://doi.org/10.3182/20070604-3-MX-2914.00023.
[89] M. Butkus, J. Repšytė, V. Galvanauskas, Fuzzy Logic-Based Adaptive Control of Specific Growth Rate in Fed-Batch Biotechnological Processes. A Simulation Study, Applied Sciences 10 (2020) 6818. https://doi.org/10.3390/app10196818.
[90] O. Johnsson, J. Andersson, G. Lidén, C. Johnsson, T. Hägglund, Feed rate control in fed-batch fermentations based on frequency content analysis, Biotechnology Progress 29 (2013) 817–824. https://doi.org/10.1002/btpr.1727.
[91] W. Sommeregger, B. Sissolak, K. Kandra, M. Von Stosch, M. Mayer, G. Striedner, Quality by control: Towards model predictive control of mammalian cell culture bioprocesses, Biotechnology Journal 12 (2017) 1600546. https://doi.org/10.1002/biot.201600546.
[92] L. Mears, S.M. Stocks, M.O. Albaek, G. Sin, K.V. Gernaey, Mechanistic Fermentation Models for Process Design, Monitoring, and Control, Trends in Biotechnology 35 (2017) 914–924. https://doi.org/10.1016/j.tibtech.2017.07.002.

[93] J. Glassey, Multivariate data analysis for advancing the interpretation of bioprocess measurement and monitoring data, Measurement, Monitoring, Modelling and Control of Bioprocesses (2013) 167–191.
[94] D. Levisauskas, Inferential control of the specific growth rate in fed-batch cultivation processes, Biotechnology Letters 23 (2001) 1189–1195. https://doi.org/10.1023/A:1010528915228.
[95] A. Survyla, D. Levisauskas, R. Urniezius, R. Simutis, An oxygen-uptake-rate-based estimator of the specific growth rate in *Escherichia coli* BL21 strains cultivation processes, Computational and Structural Biotechnology Journal 19 (2021) 5856–5863. https://doi.org/10.1016/j.csbj.2021.10.015.
[96] P. Wechselberger, P. Sagmeister, C. Herwig, Real-time estimation of biomass and specific growth rate in physiologically variable recombinant fed-batch processes, Bioprocess Biosyst Eng 36 (2013) 1205–1218. https://doi.org/10.1007/s00449-012-0848-4.
[97] S. Katla, N. Mohan, S.S. Pavan, U. Pal, S. Sivaprakasam, Control of specific growth rate for the enhanced production of human interferon α2b in glycoengineered *Pichia pastoris* : process analytical technology guided approach, J of Chemical Tech & Biotech 94 (2019) 3111–3123. https://doi.org/10.1002/jctb.6118.
[98] M. Dabros, M.M. Schuler, I.W. Marison, Simple control of specific growth rate in biotechnological fed-batch processes based on enhanced online measurements of biomass, Bioprocess Biosyst Eng 33 (2010) 1109–1118. https://doi.org/10.1007/s00449-010-0438-2.
[99] D. Ehgartner, T. Hartmann, S. Heinzl, M. Frank, L. Veiter, J. Kager, C. Herwig, J. Fricke, Controlling the specific growth rate via biomass trend regulation in filamentous fungi bioprocesses, Chemical Engineering Science 172 (2017) 32–41. https://doi.org/10.1016/j.ces.2017.06.020.
[100] J. Zitzmann, T. Weidner, G. Eichner, D. Salzig, P. Czermak, Dielectric Spectroscopy and Optical Density Measurement for the Online Monitoring and Control of Recombinant Protein Production in Stably Transformed *Drosophila melanogaster* S2 Cells, Sensors 18 (2018) 900. https://doi.org/10.3390/s18030900.
[101] B. Warth, G. Rajkai, C.-F. Mandenius, Evaluation of software sensors for on-line estimation of culture conditions in an *Escherichia coli* cultivation expressing a recombinant protein, Journal of Biotechnology 147 (2010) 37–45. https://doi.org/10.1016/j.jbiotec.2010.02.023.
[102] Z.I.T.A. Soons, M. Streefland, G. Van Straten, A.J.B. Van Boxtel, Assessment of near infrared and “software sensor” for biomass monitoring and control, Chemometrics and Intelligent Laboratory Systems 94 (2008) 166–174. https://doi.org/10.1016/j.chemolab.2008.07.009.
[103] R. Biener, A. Steinkämper, J. Hofmann, Calorimetric control for high cell density cultivation of a recombinant *Escherichia coli* strain, Journal of Biotechnology 146 (2010) 45–53. https://doi.org/10.1016/j.jbiotec.2010.01.004.
[104] R. Biener, A. Steinkämper, T. Horn, Calorimetric control of the specific growth rate during fed-batch cultures of *Saccharomyces cerevisiae*, Journal of Biotechnology 160 (2012) 195–201. https://doi.org/10.1016/j.jbiotec.2012.03.006.
[105] D. Paulsson, R. Gustavsson, C.-F. Mandenius, A Soft Sensor for Bioprocess Control Based on Sequential Filtering of Metabolic Heat Signals, Sensors 14 (2014) 17864–17882. https://doi.org/10.3390/s141017864.
[106] J. Kottelat, B. Freeland, M. Dabros, Novel Strategy for the Calorimetry-Based Control of Fed-Batch Cultivations of *Saccharomyces cerevisiae*, Processes 9 (2021) 723. https://doi.org/10.3390/pr9040723.

[107] N. Mohan, S.S. Pavan, A. Jayakumar, S. Rathinavelu, S. Sivaprakasam, Real-time metabolic heat-based specific growth rate soft sensor for monitoring and control of high molecular weight hyaluronic acid production by *Streptococcus zooepidemicus*, Appl Microbiol Biotechnol 106 (2022) 1079–1095. https://doi.org/10.1007/s00253-022-11760-1.
[108] P. Allampalli, S. Rathinavelu, N. Mohan, S. Sivaprakasam, Deployment of metabolic heat rate based soft sensor for estimation and control of specific growth rate in glycoengineered *Pichia pastoris* for human interferon alpha 2b production, Journal of Biotechnology 359 (2022) 194–206. https://doi.org/10.1016/j.jbiotec.2022.10.006.
[109] M.B. Haack, A.E. Lantz, P.P. Mortensen, L. Olsson, Chemometric analysis of in-line multi-wavelength fluorescence measurements obtained during cultivations with a lipase producing*Aspergillus oryzae* strain, Biotechnol. Bioeng. 96 (2007) 904–913. https://doi.org/10.1002/bit.21170.
[110] P. Ödman, C.L. Johansen, L. Olsson, K.V. Gernaey, A.E. Lantz, On-line estimation of biomass, glucose and ethanol in *Saccharomyces cerevisiae* cultivations using in-situ multi-wavelength fluorescence and software sensors, Journal of Biotechnology 144 (2009) 102–112.
[111] K. Nielsen, D. Gall, M. Jolley, G. Leishman, S. Balsevicius, P. Smith, P. Nicoletti, F. Thomas, A homogeneous fluorescence polarization assay for detection of antibody to Brucella abortus, Journal of Immunological Methods 195 (1996) 161–168.
[112] C. Justice, A. Brix, D. Freimark, M. Kraume, P. Pfromm, B. Eichenmueller, P. Czermak, Process control in cell culture technology using dielectric spectroscopy, Biotechnology Advances 29 (2011) 391–401. https://doi.org/10.1016/j.biotechadv.2011.03.002.
[113] B. Moore, R. Sanford, A. Zhang, Case study: The characterization and implementation of dielectric spectroscopy (biocapacitance) for process control in a commercial GMP CHO manufacturing process, Biotechnology Progress 35 (2019) e2782. https://doi.org/10.1002/btpr.2782.
[114] M. Li, B. Ebel, F. Blanchard, C. Paris, E. Guedon, A. Marc, Control of IgG glycosylation by in situ and real-time estimation of specific growth rate of CHO cells cultured in bioreactor, Biotechnology and Bioengineering 116 (2019) 985–993. https://doi.org/10.1002/bit.26914.
[115] U. Von Stockar, T. Maskow, J. Liu, I.W. Marison, R. Patiño, Thermodynamics of microbial growth and metabolism: An analysis of the current situation, Journal of Biotechnology 121 (2006) 517–533. https://doi.org/10.1016/j.jbiotec.2005.08.012.
[116] R.N. Landau, Expanding the role of reaction calorimetry, Thermochimica Acta 289 (1996) 101–126. https://doi.org/10.1016/S0040-6031(96)03081-X.
[117] T. Maskow, R. Kemp, F. Buchholz, T. Schubert, B. Kiesel, H. Harms, What heat is telling us about microbial conversions in nature and technology: from chip- to megacalorimetry: Calorimetry controls microbial activities, Microbial Biotechnology 3 (2010) 269–284. https://doi.org/10.1111/j.1751-7915.2009.00121.x.
[118] T. Maskow, H. Harms, Real Time Insights into Bioprocesses Using Calorimetry: State of the Art and Potential, Engineering in Life Sciences 6 (2006) 266–277. https://doi.org/10.1002/elsc.200520123.
[119] C. Larsson, G. Lidn, C. Niklasson, L. Gustafsson, Calorimetric control of fed-batch cultures of *Saccharomyces cerevisiae*, Bioprocess Engineering 7 (1991) 151–155. https://doi.org/10.1007/BF00387410.
[120] D. Voisard, U. Von Stockar, I.W. Marison, Quantitative calorimetric investigation of fed-batch cultures of *Bacillus sphaericus* 1593M, Thermochimica Acta 394 (2002) 99–111. https://doi.org/10.1016/S0040-6031(02)00243-5.

[121] M. Rohde, S. Paufler, H. Harms, T. Maskow, Calorespirometric feeding control enhances bioproduction from toxic feedstocks—Demonstration for biopolymer production out of methanol, Biotech & Bioengineering 113 (2016) 2113–2121. https://doi.org/10.1002/bit.25986.
[122] L. Sanchez-Garcia, L. Martín, R. Mangues, N. Ferrer-Miralles, E. Vázquez, A. Villaverde, Recombinant pharmaceuticals from microbial cells: a 2015 update, Microb Cell Fact 15 (2016) 33. https://doi.org/10.1186/s12934-016-0437-3.
[123] J.-M. Puertas, J. Ruiz, M.R. de la Vega, J. Lorenzo, G. Caminal, G. González, Influence of specific growth rate over the secretory expression of recombinant potato carboxypeptidase inhibitor in fed-batch cultures of *Escherichia coli*, Process Biochemistry 45 (2010) 1334–1341. https://doi.org/10.1016/j.procbio.2010.04.024.
[124] J. Schenk, K. Balazs, C. Jungo, J. Urfer, C. Wegmann, A. Zocchi, I.W. Marison, U. von Stockar, Influence of specific growth rate on specific productivity and glycosylation of a recombinant avidin produced by a *Pichia pastoris* Mut+ strain, Biotechnol. Bioeng. 99 (2008) 368–377. https://doi.org/10.1002/bit.21565.
[125] L. Olsson, J. Nielsen, On-line and in situ monitoring of biomass in submerged cultivations, Trends in Biotechnology 15 (1997) 517–522. https://doi.org/10.1016/S0167-7799(97)01136-0.
[126] S. Katla, N. Mohan, S.S. Pavan, U. Pal, S. Sivaprakasam, Control of specific growth rate for the enhanced production of human interferon α2b in glycoengineered *Pichia pastoris* : process analytical technology guided approach, J. Chem. Technol. Biotechnol. 94 (2019) 3111–3123. https://doi.org/10.1002/jctb.6118.
[127] P. Hrnčiřík, T. Moucha, J. Mareš, J. Náhlík, D. Janáčová, Software Sensors for Biomass Concentration Estimation in Filamentous Microorganism Cultivation Process, Chem. Biochem. Eng. q. (Online) 33 (2019) 141–151. https://doi.org/10.15255/CABEQ.2018.1387.
[128] U. Von Stockar, J.-S. Liu, Does microbial life always feed on negative entropy? Thermodynamic analysis of microbial growth, Biochimica et Biophysica Acta (BBA) - Bioenergetics 1412 (1999) 191–211. https://doi.org/10.1016/S0005-2728(99)00065-1.
[129] J.J. Hoijnen, M.C.M. van Loosdrecht, L. Tijhuis, A black box mathematical model to calculate auto- and heterotrophic biomass yields based on Gibbs energy dissipation, Biotechnol. Bioeng. 40 (1992) 1139–1154. https://doi.org/10.1002/bit.260401003.
[130] H.V. Westerhoff, K.J. Hellingwerf, K. Van Dam, Thermodynamic efficiency of microbial growth is low but optimal for maximal growth rate, Proc. Natl. Acad. Sci. U.S.A. 80 (1983) 305–309. https://doi.org/10.1073/pnas.80.1.305.
[131] J. Kottelat, B. Freeland, M. Dabros, Novel Strategy for the Calorimetry-Based Control of Fed-Batch Cultivations of *Saccharomyces cerevisiae*, Processes 9 (2021) 723. https://doi.org/10.3390/pr9040723.
[132] D. Voisard, C. Claivaz, L. Menoud, I.W. Marison, U. von Stockar, Use of reaction calorimetry to monitor and control microbial cultures producing industrially relevant secondary metabolites, Thermochimica Acta 309 (1998) 87–96. https://doi.org/10.1016/S0040-6031(97)00370-5.
[133] C. Larsson, G. Lidn, C. Niklasson, L. Gustafsson, Calorimetric control of fed-batch cultures of Saccharomyces cerevisiae, Bioprocess Engineering 7 (1991) 151–155. https://doi.org/10.1007/BF00387410.
[134] R. Biener, A. Steinkämper, J. Hofmann, Calorimetric control for high cell density cultivation of a recombinant *Escherichia coli* strain, Journal of Biotechnology 146 (2010) 45–53. https://doi.org/10.1016/j.jbiotec.2010.01.004.

[135] R. Biener, A. Steinkämper, T. Horn, Calorimetric control of the specific growth rate during fed-batch cultures of *Saccharomyces cerevisiae*, Journal of Biotechnology 160 (2012) 195–201. https://doi.org/10.1016/j.jbiotec.2012.03.006.
[136] M.M. Schuler, S. Sivaprakasam, B. Freeland, A. Hama, K.-M. Hughes, I.W. Marison, Investigation of the potential of biocalorimetry as a process analytical technology (PAT) tool for monitoring and control of Crabtree-negative yeast cultures, Appl Microbiol Biotechnol 93 (2012) 575–584. https://doi.org/10.1007/s00253-011-3507-9.
[137] N. Mohan, S. Sivaprakasam, Heat Compensation Calorimeter as a Process Analytical Tool To Monitor and Control Bioprocess Systems, Ind. Eng. Chem. Res. 56 (2017) 8416–8427. https://doi.org/10.1021/acs.iecr.7b01367.
[138] L. Mendes, N. Ogink, N. Edouard, H. Van Dooren, I. Tinôco, J. Mosquera, NDIR Gas Sensor for Spatial Monitoring of Carbon Dioxide Concentrations in Naturally Ventilated Livestock Buildings, Sensors 15 (2015) 11239–11257. https://doi.org/10.3390/s150511239.
[139] D. Paulsson, R. Gustavsson, C.-F. Mandenius, A Soft Sensor for Bioprocess Control Based on Sequential Filtering of Metabolic Heat Signals, Sensors 14 (2014) 17864–17882. https://doi.org/10.3390/s141017864.
[140] U. Von Stockar, V. Vojinović, T. Maskow, J. Liu, Can microbial growth yield be estimated using simple thermodynamic analogies to technical processes?, Chemical Engineering and Processing: Process Intensification 47 (2008) 980–990. https://doi.org/10.1016/j.cep.2007.02.016.
[141] N. Mohan, S.S. Pavan, A. Jayakumar, S. Rathinavelu, S. Sivaprakasam, Real-time metabolic heat-based specific growth rate soft sensor for monitoring and control of high molecular weight hyaluronic acid production by *Streptococcus zooepidemicus*, Appl Microbiol Biotechnol 106 (2022) 1079–1095. https://doi.org/10.1007/s00253-022-11760-1.
[142] P. Allampalli, S. Rathinavelu, N. Mohan, S. Sivaprakasam, Deployment of metabolic heat rate based soft sensor for estimation and control of specific growth rate in glycoengineered *Pichia pastoris* for human interferon alpha 2b production, Journal of Biotechnology 359 (2022) 194–206. https://doi.org/10.1016/j.jbiotec.2022.10.006.
[143] N. Mohan, S.S. Pavan, A. Achar, N. Swaminathan, S. Sivaprakasam, Calorespirometric investigation of *Streptococcus zooepidemicus* metabolism: Thermodynamics of anabolic payload contribution by growth and hyaluronic acid synthesis, Biochemical Engineering Journal 152 (2019) 107367. https://doi.org/10.1016/j.bej.2019.107367.
[144] A. Kawala-Sterniuk, M. Podpora, M. Pelc, M. Blaszczyszyn, E.J. Gorzelanczyk, R. Martinek, S. Ozana, Comparison of Smoothing Filters in Analysis of EEG Data for the Medical Diagnostics Purposes, Sensors 20 (2020) 807. https://doi.org/10.3390/s20030807.
[145] J.-S. Liu, V. Vojinović, R. Patiño, Th. Maskow, U. von Stockar, A comparison of various Gibbs energy dissipation correlations for predicting microbial growth yields, Thermochimica Acta 458 (2007) 38–46. https://doi.org/10.1016/j.tca.2007.01.016.
[146] I. Marison, J.-S. Liu, S. Ampuero, U. Von Stockar, B. Schenker, Biological reaction calorimetry: Development of high sensitivity bio-calorimeters, Thermochimica Acta 309 (1998) 157–173. https://doi.org/10.1016/S0040-6031(97)00424-3.
[147] S. Katla, S.S. Pavan, N. Mohan, S. Sivaprakasam, Biocalorimetric monitoring of glycoengineered *P. PASTORIS* cultivation for the production of recombinant huIFNα2b: A quantitative study based on mixed feeding strategies, Biotechnol Progress 36 (2020). https://doi.org/10.1002/btpr.2971.
[148] C. Jungo, J. Schenk, M. Pasquier, I.W. Marison, U. Von Stockar, A quantitative analysis of the benefits of mixed feeds of sorbitol and methanol for the production of recombinant

avidin with *Pichia pastoris*, Journal of Biotechnology 131 (2007) 57–66. https://doi.org/10.1016/j.jbiotec.2007.05.019.
[149] S. Paufler, M.-T. Weichler, H. Harms, T. Maskow, Simple improvement of the sensitivity of a heat flux reaction calorimeter to monitor bioprocesses with weak heat production, Thermochimica Acta 569 (2013) 71–77. https://doi.org/10.1016/j.tca.2013.07.001.
[150] B. Birou, I.W. Marison, U.V. Stockar, Calorimetric investigation of aerobic fermentations, Biotechnol. Bioeng. 30 (1987) 650–660. https://doi.org/10.1002/bit.260300509.
[151] Q. Zhou, V. Chen, C.P. Shannon, X.-S. Wei, X. Xiang, X. Wang, Z.-H. Wang, S.J. Tebbutt, T.R. Kollmann, E.N. Fish, Interferon-α2b Treatment for COVID-19, Front. Immunol. 11 (2020) 1061. https://doi.org/10.3389/fimmu.2020.01061.
[152] M. Vachher, A. Sen, R. Kapila, A. Nigam, Microbial therapeutic enzymes: A promising area of biopharmaceuticals, Current Research in Biotechnology 3 (2021) 195–208. https://doi.org/10.1016/j.crbiot.2021.05.006.
[153] Z. Yang, Z. Zhang, Engineering strategies for enhanced production of protein and bio-products in *Pichia pastoris*: A review, Biotechnology Advances 36 (2018) 182–195. https://doi.org/10.1016/j.biotechadv.2017.11.002.
[154] M. Vandermies, P. Fickers, Bioreactor-Scale Strategies for the Production of Recombinant Protein in the Yeast *Yarrowia lipolytica*, Microorganisms 7 (2019) 40. https://doi.org/10.3390/microorganisms7020040.
[155] T.I. Potgieter, S.D. Kersey, M.R. Mallem, A.C. Nylen, M. d'Anjou, Antibody expression kinetics in glycoengineered *Pichia pastoris*, Biotech & Bioengineering 106 (2010) 918–927. https://doi.org/10.1002/bit.22756.
[156] A.E. Cunha, J.J. Clemente, R. Gomes, F. Pinto, M. Thomaz, S. Miranda, R. Pinto, D. Moosmayer, P. Donner, M.J.T. Carrondo, Methanol induction optimization for scFv antibody fragment production in *Pichia pastoris*, Biotech & Bioengineering 86 (2004) 458–467. https://doi.org/10.1002/bit.20051.
[157] T. Ohya, M. Ohyama, K. Kobayashi, Optimization of human serum albumin production in methylotrophic yeast *Pichia pastoris* by repeated fed-batch fermentation, Biotech & Bioengineering 90 (2005) 876–887. https://doi.org/10.1002/bit.20507.
[158] Y. Brignoli, B. Freeland, D. Cunningham, M. Dabros, Control of Specific Growth Rate in Fed-Batch Bioprocesses: Novel Controller Design for Improved Noise Management, Processes 8 (2020) 679. https://doi.org/10.3390/pr8060679.
[159] S. Katla, N. Mohan, S.S. Pavan, U. Pal, S. Sivaprakasam, Control of specific growth rate for the enhanced production of human interferon α2b in glycoengineered *Pichia pastoris* : process analytical technology guided approach, J. Chem. Technol. Biotechnol. 94 (2019) 3111–3123. https://doi.org/10.1002/jctb.6118.
[160] R. Luttmann, D.G. Bracewell, G. Cornelissen, K.V. Gernaey, J. Glassey, V.C. Hass, C. Kaiser, C. Preusse, G. Striedner, C.-F. Mandenius, Soft sensors in bioprocessing: A status report and recommendations, Biotechnology Journal 7 (2012) 1040–1048. https://doi.org/10.1002/biot.201100506.
[161] P. Sagmeister, P. Wechselberger, M. Jazini, A. Meitz, T. Langemann, C. Herwig, Soft sensor assisted dynamic bioprocess control: Efficient tools for bioprocess development, Chemical Engineering Science 96 (2013) 190–198. https://doi.org/10.1016/j.ces.2013.02.069.
[162] R. Biener, A. Steinkämper, J. Hofmann, Calorimetric control for high cell density cultivation of a recombinant *Escherichia coli* strain, Journal of Biotechnology 146 (2010) 45–53. https://doi.org/10.1016/j.jbiotec.2010.01.004.

[163] W. Zhang, L.A. Smith, B.A. Plantz, V.L. Schlegel, M.M. Meagher, Design of Methanol Feed Control in *Pichia pastoris* Fermentations Based upon a Growth Model, Biotechnol. Prog. 18 (2002) 1392–1399. https://doi.org/10.1021/bp025516w.
[164] S. Craven, J. Whelan, B. Glennon, Glucose concentration control of a fed-batch mammalian cell bioprocess using a nonlinear model predictive controller, Journal of Process Control 24 (2014) 344–357. https://doi.org/10.1016/j.jprocont.2014.02.007.
[165] Hisbullah, M.A. Hussain, K.B. Ramachandran, Design of a Fuzzy Logic Controller for Regulating Substrate Feed to Fed-Batch Fermentation, Food and Bioproducts Processing 81 (2003) 138–146. https://doi.org/10.1205/096030803322088279.
[166] A. Kuprijanov, S. Gnoth, R. Simutis, A. Lübbert, Advanced control of dissolved oxygen concentration in fed batch cultures during recombinant protein production, Appl Microbiol Biotechnol 82 (2009) 221–229. https://doi.org/10.1007/s00253-008-1765-y.
[167] S. Gnoth, A. Kuprijanov, R. Simutis, A. Lübbert, Simple adaptive pH control in bioreactors using gain-scheduling methods, Appl Microbiol Biotechnol 85 (2010) 955–964. https://doi.org/10.1007/s00253-009-2114-5.
[168] S. Katla, B. Karmakar, S.R.R. Tadi, N. Mohan, B. Anand, U. Pal, S. Sivaprakasam, High level extracellular production of recombinant human interferon alpha 2b in glycoengineered *Pichia pastoris:* culture medium optimization, high cell density cultivation and biological characterization, J Appl Microbiol 126 (2019) 1438–1453. https://doi.org/10.1111/jam.14227.
[169] G. Flores-Cosío, E.J. Herrera-López, M. Arellano-Plaza, A. Gschaedler-Mathis, M. Kirchmayr, L. Amaya-Delgado, Application of dielectric spectroscopy to unravel the physiological state of microorganisms: current state, prospects and limits, Appl Microbiol Biotechnol 104 (2020) 6101–6113. https://doi.org/10.1007/s00253-020-10677-x.
[170] S. Sivaprakasam, M.M. Schuler, A. Hama, K.-M. Hughes, I.W. Marison, Biocalorimetry as a process analytical technology process analyser; robust in-line monitoring and control of aerobic fed-batch cultures of crabtree-negative yeast cells, J Therm Anal Calorim 104 (2011) 75–85. https://doi.org/10.1007/s10973-010-1259-x.
[171] J.M. Barrigón, R. Ramon, I. Rocha, F. Valero, E.C. Ferreira, J.L. Montesinos, State and specific growth estimation in heterologous protein production by *Pichia pastoris*, AIChE Journal 58 (2012) 2966–2979. https://doi.org/10.1002/aic.12810.
[172] J.L. Cereghino, J.M. Cregg, Heterologous protein expression in the methylotrophic yeast *Pichia pastoris*, FEMS Microbiol Rev 24 (2000) 45–66. https://doi.org/10.1111/j.1574-6976.2000.tb00532.x.
[173] A. Vieira Gomes, T. Souza Carmo, L. Silva Carvalho, F. Mendonça Bahia, N. Parachin, Comparison of Yeasts as Hosts for Recombinant Protein Production, Microorganisms 6 (2018) 38. https://doi.org/10.3390/microorganisms6020038.
[174] S. Macauley-Patrick, M.L. Fazenda, B. McNeil, L.M. Harvey, Heterologous protein production using the*Pichia pastoris* expression system, Yeast 22 (2005) 249–270. https://doi.org/10.1002/yea.1208.
[175] B. Wang, D. Li, T. Liu, H. Wang, F. Luo, Y. Liu, Subcutaneous injection of IFN alpha-2b for COVID-19: an observational study, BMC Infect Dis 20 (2020) 723. https://doi.org/10.1186/s12879-020-05425-5.
[176] J.M. Cregg, J.L. Cereghino, J. Shi, D.R. Higgins, Recombinant protein expression in *Pichia pastoris*, Mol Biotechnol 16 (2000) 23–52. https://doi.org/10.1385/MB:16:1:23.
[177] M. Gao, Z. Shi, Process Control and Optimization for Heterologous Protein Production by Methylotrophic *Pichia pastoris*, Chinese Journal of Chemical Engineering 21 (2013) 216–226. https://doi.org/10.1016/S1004-9541(13)60461-9.
[178] P. Çalık, Ö. Ata, H. Güneş, A. Massahi, E. Boy, A. Keskin, S. Öztürk, G.H. Zerze, T.H. Özdamar, Recombinant protein production in Pichia pastoris under glyceraldehyde-3-

phosphate dehydrogenase promoter: From carbon source metabolism to bioreactor operation parameters, Biochemical Engineering Journal 95 (2015) 20–36. https://doi.org/10.1016/j.bej.2014.12.003.
[179] S. Katla, S.S. Pavan, N. Mohan, S. Sivaprakasam, Biocalorimetric monitoring of glycoengineered *P . PASTORIS* cultivation for the production of recombinant huIFNα2b: A quantitative study based on mixed feeding strategies, Biotechnol Progress 36 (2020). https://doi.org/10.1002/btpr.2971.
[180] L.B. Trinh, J.N. Phue, J. Shiloach, Effect of methanol feeding strategies on production and yield of recombinant mouse endostatin from*Pichia pastoris*, Biotechnol. Bioeng. 82 (2003) 438–444. https://doi.org/10.1002/bit.10587.
[181] N.K. Khatri, F. Hoffmann, Impact of methanol concentration on secreted protein production in oxygen-limited cultures of recombinant*Pichia pastoris*, Biotechnol. Bioeng. 93 (2006) 871–879. https://doi.org/10.1002/bit.20773.
[182] H. Narayanan, M. Sokolov, M. Morbidelli, A. Butté, A new generation of predictive models: The added value of hybrid models for manufacturing processes of therapeutic proteins, Biotechnology and Bioengineering 116 (2019) 2540–2549. https://doi.org/10.1002/bit.27097.
[183] O. Cos, R. Ramon, J.L. Montesinos, F. Valero, A simple model-based control for*Pichia pastoris* allows a more efficient heterologous protein production bioprocess, Biotechnol. Bioeng. 95 (2006) 145–154. https://doi.org/10.1002/bit.21005.
[184] O. Grigs, E. Bolmanis, A. Kazaks, HBsAg Production in Methanol Controlled *P. pastoris* GS115 Mut^S Bioreactor Process, KEM 903 (2021) 40–45. https://doi.org/10.4028/www.scientific.net/KEM.903.40.
[185] National Instruments, Analog Input:Scaling:Offset Property, n.d.
[186] E. Bolmanis, O. Grigs, A. Kazaks, V. Galvanauskas, High-level production of recombinant HBcAg virus-like particles in a mathematically modelled *P. pastoris* GS115 Mut+ bioreactor process under controlled residual methanol concentration, Bioprocess Biosyst Eng 45 (2022) 1447–1463. https://doi.org/10.1007/s00449-022-02754-4.
[187] C. Jungo, C. Rérat, I.W. Marison, U. Von Stockar, Quantitative characterization of the regulation of the synthesis of alcohol oxidase and of the expression of recombinant avidin in a *Pichia pastoris* Mut+ strain, Enzyme and Microbial Technology 39 (2006) 936–944. https://doi.org/10.1016/j.enzmictec.2006.01.027.
[188] W.-C. Liu, S. Inwood, T. Gong, A. Sharma, L.-Y. Yu, P. Zhu, Fed-batch high-cell-density fermentation strategies for *Pichia pastoris* growth and production, Critical Reviews in Biotechnology 39 (2019) 258–271. https://doi.org/10.1080/07388551.2018.1554620.
[189] A. Surribas, R. Stahn, J.L. Montesinos, S.-O. Enfors, F. Valero, M. Jahic, Production of a Rhizopus oryzae lipase from *Pichia pastoris* using alternative operational strategies, Journal of Biotechnology 130 (2007) 291–299. https://doi.org/10.1016/j.jbiotec.2007.04.009.
[190] D. Wu, X.W. Yu, T.C. Wang, R. Wang, Y. Xu, High yield Rhizopus chinenisis prolipase production in *Pichia pastoris*: Impact of methanol concentration, Biotechnol Bioproc E 16 (2011) 305–311. https://doi.org/10.1007/s12257-009-3021-4.
[191] H. Hohenblum, B. Gasser, M. Maurer, N. Borth, D. Mattanovich, Effects of gene dosage, promoters, and substrates on unfolded protein stress of recombinant*Pichia pastoris*, Biotechnol. Bioeng. 85 (2004) 367–375. https://doi.org/10.1002/bit.10904.
[192] I.A. Pla, L.M. Damasceno, T. Vannelli, G. Ritter, C.A. Batt, M.L. Shuler, Evaluation of Mut+ and MutS *Pichia pastoris* Phenotypes for High Level Extracellular scFv Expression under Feedback Control of the Methanol Concentration, Biotechnol. Prog. 22 (2006) 881–888. https://doi.org/10.1021/bp060012+.

[193] E. Bolmanis, K. Dubencovs, A. Suleiko, J. Vanags, Model Predictive Control—A Stand Out among Competitors for Fed-Batch Fermentation Improvement, Fermentation 9 (2023) 206. https://doi.org/10.3390/fermentation9030206.
[194] M. Ławryńczuk, Modelling and nonlinear predictive control of a yeast fermentation biochemical reactor using neural networks, Chemical Engineering Journal 145 (2008) 290–307. https://doi.org/10.1016/j.cej.2008.08.005.
[195] K. Dubencovs, A. Suleiko, E. Sile, I. Petrovskis, I. Akopjana, A. Suleiko, V. Galvanauskas, K. Tars, J. Vanags, The Application of Adaptive Model Predictive Control for Fed-Batch *Escherichia coli* BL21 (DE3) Cultivation and Biosynthesis of Recombinant Proteins, Fermentation 9 (2023) 1015. https://doi.org/10.3390/fermentation9121015.
[196] M. Ebden, Gaussian Processes: A Quick Introduction, (2015). http://arxiv.org/abs/1505.02965 (accessed February 13, 2024).
[197] M. Rashedi, M. Rafiei, M. Demers, H. Khodabandehlou, T. Wang, A. Tulsyan, C. Undey, C. Garvin, Machine learning-based model predictive controller design for cell culture processes, Biotech & Bioengineering 120 (2023) 2144–2159. https://doi.org/10.1002/bit.28486.
[198] W. Wong, E. Chee, J. Li, X. Wang, Recurrent Neural Network-Based Model Predictive Control for Continuous Pharmaceutical Manufacturing, Mathematics 6 (2018) 242. https://doi.org/10.3390/math6110242.
[199] S. Tebbani, D. Dumur, G. Hafidi, A. Vande Wouwer, Nonlinear Predictive Control of Fed-Batch Cultures of *Escherichia coli*, Chem Eng & Technol 33 (2010) 1112–1124. https://doi.org/10.1002/ceat.201000029.