Robust control of fed-batch cultures of Escherichia coli

Merouane Abadli

Thèse de doctorat

NNT: 2021UPASG072

Robust control of fed-batch cultures of *Escherichia coli*

Commande robuste de cultures fed-batch de Escherichia coli

Thèse de doctorat de l'Université Paris-Saclay

École doctorale n° 580, Sciences et Technologies de l'Information et de la Communication (STIC)
Spécialité de doctorat: Automatique
Unité de recherche : Université Paris-Saclay, CNRS, CentraleSupélec, Laboratoire des signaux et systèmes, 91190, Gif-sur-Yvette, France
Référent : Faculté des sciences d'Orsay

Thèse présentée et soutenue à Paris Saclay, le 20 Octobre 2021, par

Merouane ABADLI

Composition du Jury

Antoine CHAILLET Professeur, CentraleSupélec/L2S	Président
Rudibert KING Professeur, Technische Universität Berlin	Rapporteur & Examinateur
Jesus PICO Professeur, Universitat Politècnica de València	Rapporteur & Examinateur
Anne-Lise HANTSON Professeur, Université de Mons	Examinatrice

Direction de la thèse

Sihem TEBBANI Professeur, CentraleSupélec/L2S	Directrice de thèse
Alain VANDE WOUWER Professeur, Université de Mons/FPMs	Directeur de thèse
Didier DUMUR Professeur, CentraleSupélec/L2S	Co-encadrant & invité

Invités

Laurent DEWASME Docteur ingénieur, Université de Mons/FPMs	Invité

UMONS
University of Mons

université
PARIS-SACLAY

Robust control of fed-batch cultures of *Escherichia coli*

THÈSE DE DOCTORAT

soumise à

L'UNIVERSITÉ DE MONS

pour l'obtention du grade de

DOCTEUR EN SCIENCES DE L'INGÉNIEUR ET TECHNOLOGIE

Thèse présentée et soutenue à Gif-sur-Yvette, le 20 Octobre 2021, par:

Merouane ABADLI

Composition du Jury

Antoine CHAILLET Professeur, CentraleSupélec/L2S	President
Laurent DEWASME Docteur ingénieur, Université de Mons/FPMs	Secretary
Jesus PICO Professeur, Universitat Politècnica de València	Reviewer & Examiner
Rudibert KING Professeur, Technische Universität Berlin	Reviewer & Examiner
Anne-Lise HANTSON Professeur, Université de Mons	Examiner
Philippe BOGAERTS Professeur, Université Libre de Bruxelles	Examiner

Direction de la thèse

Sihem TEBBANI Professeur, CentraleSupélec/L2S	Supervisor
Alain VANDE WOUWER Professeur, Université de Mons/FPMs	Supervisor
Didier DUMUR Professeur, CentraleSupélec/L2S	Co-supervisor

Important information

This work is a joint PhD thesis of Université de Paris-Saclay (France) and Université de Mons (Belgium). The jury members are:

Antoine CHAILLET Professor, CentraleSupélec/L2S	President
Laurent DEWASME Senior Researcher and part-time Lecturer, Université de Mons	Secretary
Jesus PICO Professor, Universitat Politècnica de València	Reviewer
Rudibert KING Professor, Technische Universität Berlin	Reviewer
Anne-Lise HANTSON Professor, Université de Mons	Examiner
Philippe BOGAERTS Professor, Université Libre de Bruxelles	Examiner
Sihem TEBBANI Professor, CentraleSupélec/L2S	Supervisor
Alain VANDE WOUWER Professor, Université de Mons/FPMs	Supervisor
Didier DUMUR Professor, CentraleSupélec/L2S	Co-supervisor

The thesis was first presented in a private defense, held at Université of Mons on the 22nd of September 2021 in the presence of all jury members.

A public defense was held at Université Paris-Saclay on 20th of October 2021.

The roles of the jury members was adapted in the cover page of this manuscript in accordance with the regulation of each university. Consequently, they may differ for the two universities, but it is emphasized that the actual composition of the jury was the same.

"So the problem is ... not so much to see what nobody has yet seen, as to think what nobody has yet thought concerning that which everybody sees. "

Arthur Schopenhauer

"Parerga und Paralipomena" , 1851

"If I have seen further it is by standing on the shoulders of Giants "

Isaac Newton

letter to Robert Hooke, 1675

Acknowledgements

This thesis is the fruit of the collaboration between the L2S laboratory at CentraleSupélec, and the SECO laboratory of the polytechnic faculty of Mons.

I would like begin by thanking the jury members who were more than generous with their extreme patience and precious time. Their expertise and feedback on my work will surely contribute to my professional and personal development. I thank my committee chairman Pr. Antoine Chaillet for his support, patience, and good advice throughout the defense process. A special thanks to the reviewers Pr. Rudibert King and Pr. Jesus Pico for the countless hours of reflecting, reading, and examining my manuscript. I thank them for the constructive feedback on the different aspect of the thesis and their precious advice on the improvement of the presented work. Thank you Pr. Anne-Lise Hantson, Pr. Philippe Bogaerts and Dr. Laurent Dewasme for agreeing to examine my work and to be part of my jury and my committee. Your advice and comments helped me to advance on this research project and finalize my thesis.

I would like to express all my gratitude to my supervising team, Pr. Sihem Tebbani and Didier Dumur from CentraleSupélec, and Pr. Alain Vande Wouwer from the Polytechnic faculty of Mons. I thank them for guiding me throughout my thesis with professionalism, rigor, patience, availability, and generosity. I extend my sincere gratitude to them for their constant support, encouragement and the precious time they have given me during the thesis.

I would like to thank Sihem and Didier for their warm welcome to the lab and the automation department, for the trust they placed in me, for their kindness and care during difficult times, and for the precious opportunity they gave me to be able to participate in teaching activities at the department.

I thank Alain, for the warm welcome he always given me on each of my visits to Mons. I thank him for his constant support, good advice, regular encouragement and patience during my experiments. I also thank him for motivating me to learn and explore new concepts.

My deepest gratitude to Dr. Laurent Dewasme, without whom finishing this research project wouldn't have been possible. I thank him for his participation and guidance during the experiments, his implication in scientific writing process, and his discussions and feedback that were vital for finalizing this thesis.

I am also grateful to Pr. Anne-Lise Hantson for welcoming me to the department of Applied Chemistry and Biochemistry in order to carry out the fermentation experiments that were essential to the conclusion of this work. I would like to thank her for the availability, the advice, and the support during my stays at the biochemical lab.

I would like to thank my colleagues for creating a warm working atmosphere that allowed me to finish this doctoral PhD project. My sincere thanks to:

- The doctors and doctoral students of the Automation Department of CentraleSupélec for all the good times we shared.
- My colleagues in the automation service at the Polytechnic Faculty of Mons.

- The doctoral students and reaserchers of the Chimical and Biochimical of the Polytechnic Faculty of Mons.

I would like to thank my friends, without whom I would not have made it through the difficult moments of this research project. Mehdi, Abdelhak, Mohamed, Abdou, and Mourad for always being there for me when I needed them the most. Sarah and Fawzi for providing company, encouragement, and generosity. Amira, Katia, and Amina for their motivation, availability, and helpful advice and guidance. Mokrane and Aladdin for our daily interactions and discussions that helped motivate and lift the spirit of each one of us. Tiziri for providing constant support especially during my stays at Mons. And all my friends from ENP and CAP.

Most importantly, I am grateful for my family's unconditional, unequivocal, and loving support. I would like to pay homage to parents Kamel and Saïda for their love, wise counsel and sympathetic ear. Without your sacrifices I would never have had so many opportunities. My deepest gratitude to my brothers Sofiane and Badreddine and my sister Yasmine for the constant encouragement and support. I would like to thank all my family members, especially my grandmother who shaped the man I am today.

Finally, I dedicate this work to the memory of those we lost during the pandemic and in the recent years. Particularly: My grandparents, my uncle Hamid, and my teacher/friend/mentor Pr. Abdelkrim Bouchouata.

Abstract

Escherichia coli is a widespread cellular host for the industrial production of protein-based biopharmaceutical products considering its physiological and biological features. This production is mostly operated in fed-batch mode due to the scalable process, the low operational costs, and the relatively simple media conditions. One key challenge to maximize the bioprocess productivity is related to the production of acetate, a metabolic byproduct inhibiting the cell respiratory capacity and affecting the cells metabolic performance. This production must be maintained as low as possible.

In this thesis, model-based control strategies are considered to avoid acetate accumulation, thus maximizing biomass productivity, and to drive the culture near the optimal metabolic operating conditions. In addition, software sensors are developed to estimate the evolution of the non-measured key variables required to implement the control strategies.

To this end, three control strategies are proposed. First, the generic model control method is investigated within an adaptive framework in order to regulate the biomass growth rate to a desired reference. Secondly, a robust version of the generic model control strategy is developed to regulate the acetate concentration to a low value. Finally, the last part of the thesis focuses on the implementation of a nonlinear model predictive controller to limit acetate accumulation and compare the performance with the previously described control methods. Furthermore, an Unscented Kalman filter estimating the glucose and acetate concentrations based on the biomass measurements is implemented and coupled to the previously mentioned control schemes.

The bioprocess is a complex, nonlinear, uncertain, and time-varying system. Thereby, the developments in this study are focused on the robustness of the implemented methods towards model uncertainties and unpredicted dynamics.

The performance and robustness of the control and estimation strategies are tested and tuned by means of different scenarios of simulation runs. Fed-batch cultures of *E. coli* BL21(DE3) strain are successfully carried on a lab-scale bioreactor, highlighting the potential of the proposed strategies in real-time conditions. Theoretical developments and experimental results allow to assess the advantages of the different proposed approaches and show their tractability for further applications in an industrial framework.

The proposed control strategies presented in this thesis lead to an average gain of up to 20% in biomass productivity compared to the conventional operating mode.

Résumé

Escherichia coli est un hôte cellulaire très répandu pour la production industrielle de produits biopharmaceutiques à base de protéines, compte tenu de ses caractéristiques physiologiques et biologiques. Cette production est généralement réalisée en mode fed-batch en raison de l'évolutivité du procédé, des faibles coûts opérationnels et des conditions de milieu de culture relativement simples à mettre en œuvre. Un défi majeur pour maximiser la productivité du bioprocédé est lié à la production d'acétate, un produit métabolique inhibant la capacité respiratoire des cellules et affectant leur performance métabolique. Dès lors, il s'agit de limiter au maximum sa production par les microorganismes.

Dans cette thèse, des stratégies de commande à base de modèles sont envisagées pour éviter l'accumulation d'acétate, maximisant ainsi la productivité de la biomasse. Ces stratégies ont pour objectif d'opérer la culture en restant le plus proche possible des conditions opératoires optimales. En outre, des capteurs logiciels sont développés pour estimer l'évolution des variables clés non mesurées nécessaires à la mise en œuvre des lois de commande.

À cette fin, trois stratégies de commande sont développées. La méthode de commande par modèle générique est tout d'abord mise en œuvre dans un cadre adaptatif afin de réguler le taux de croissance de la biomasse à une référence désirée. Ensuite, une version robuste de la stratégie de commande par modèle générique est développée afin de réguler la concentration d'acétate à une valeur restant faible. Enfin, la dernière partie de la thèse s'intéresse à la mise en œuvre d'une structure de commande prédictive non linéaire pour limiter l'accumulation d'acétate et comparer les performances avec les méthodes de commande décrites précédemment. De plus, un filtre de Kalman « sans parfum » (Unscented Kalman Filter) estimant les concentrations de glucose et d'acétate à partir des mesures de biomasse est implémenté et couplé aux schémas de commande mentionnés précédemment. Enfin, la dernière partie de la thèse se concentre sur l'implémentation d'un contrôleur prédictif basé sur un modèle non-linéaire pour limiter l'accumulation d'acétate et comparer les performances avec les méthodes de contrôle décrites précédemment.

Le bioprocédé est un système complexe, non linéaire, incertain, variant dans le temps. Aussi, les développements présentés dans cette thèse se focalisent sur la robustesse des méthodes mises en œuvre vis-à-vis des incertitudes de modèle et des dynamiques non modélisées. La performance et la robustesse des schémas de commande et d'estimation sont testées et ajustées au travers de différents scénarios de simulation. Des cultures en mode Fed-batch de la souche *E. coli* BL21(DE3) sont réalisées avec succès sur un bioréacteur de laboratoire, mettant en évidence le potentiel des stratégies proposées dans un contexte de conditions opératoires en temps réel. Les développements théoriques et résultats expérimentaux permettent en outre de mettre en évidence les avantages des différentes approches proposées, et illustrent également la généralisation envisageable à des procédés industriels de plus grande échelle. Les stratégies de commande proposées dans cette thèse permettent un gain moyen jusqu'à 20% de la productivité de la biomasse par rapport au mode de fonctionnement conventionnel.

Contents

List of Figures

List of Tables

List of Abbreviations

TCA	Tricarboxylic Acid Cycle
PEP	Phosphoenolpyruvate
PDH	Pyruvate Dehydrogenase
PTA	Phosphotransacetylase
ACKA	Acetate Kinase
ACS	Acetyl-coa Synthetase
PoxB	Pyruvate Dehydrogenase
PFL	Pyruvate Formate Lyase
CSTR	Continuously Stirred Tank Reactor
FIA	Flow Injection Analysis
VISA	Virtual Instrument Software Architecture
DCU	Digital Control Unit
DO	Dissolved Oxygen
DLL	Dynamic Library Files
OD	Optical Density
DCW	Dry Cell Weight
MRAC	Model Reference Adaptive Control
OTR	Oxygen Transfer Rate
CTR	Carbon Dioxide Transfer Rate
MPC	Model Predictive Control
NMPC	Nonlinear Model Predictive Control
ANN	Artificial Neural Network
PCA	Principle Component Analysis
PLS	Partial Least Squares
GMC	Generic Model Control
ODE	Ordinary Differential Equation
FOC	First Order Linearizing Controller
PID	Proportional Integral Derivative Controller
EKF	Extended Kalman Filter
UKF	Unscented Kalman Filter

Preface

Context and motivation:

Escherichia coli is one of the most popular cellular hosts for the industrial production of protein-based drugs via non-microbial systems. Indeed, a vast array of biopharmaceutical products have been produced through *E. coli* fermentation, such as insulin, somatotropin, human parathyroid hormone, and others (Ferrer-Miralles et al., 2009). *Escherichia coli* is preferred for its physiological and biological features, such as flexible culture conditions, fast growth, and high production yields (Lee, 1996; Pontrelli et al., 2018).

Fed-batch cultivation of genetically modified strains of *E. coli* is the most common method that rapidly and efficiently produces high-quality proteins while maintaining industrial processes economic viability (Lee, 1996; Pontrelli et al., 2018). The advantage is the scalable process, the low operational costs, and the relatively simple media conditions.

However, some obstacles are still being faced to reach high cell densities in these bioprocesses, particularly stemming from host cells metabolic performance (Chou, 2007). The main challenge to ensure the process efficiency and productivity is the accumulation of acetate, a metabolic by-product inhibiting cell growth (Luli and Strohl, 1990). Acetate formation occurs when the capacity for energy generation within the cell is exceeded due to high flux into the main metabolic pathways caused by an excess in the carbon source (Han et al., 1992; Van De Walle and Shiloach, 1998). This mechanism is referred to as ”overflow metabolism” or ”Crabtree effect” (Crabtree, 1929).

Acetate presence in high concentration causes the inhibition of the cell respiratory capacity, leading to the decrease of biomass production yield and consequently the decrease of the recombinant protein production (Riesenberg et al., 1991; Rothen et al., 1998).

The goal of the work presented in this thesis is to propose and develop practical solutions to avoid overflow metabolism and maximize the biomass productivity in fed-batch *E. coli* cultures. This objective is achieved by closed-loop control and estimation strategies that drive the culture near the optimal metabolic operating conditions. In this manuscript, we attempt to answer the following questions:

- Is there a mathematical representation of the fed-batch *E. coli* process that accurately describes overflow metabolism and acetate production? Is there enough data to properly utilize this model to develop and test the control strategies?

- Which key components are available for on-line measurement? And is it possible to estimate the non-measured variables and parameters, based on these measurements?

- What are the available materials and hardware for this study? Is the reactor equipped for closed-loop operation? What software solutions are required to implement our control strategies?
- How can we translate the biological objectives into a defined control objective? What are the control variables? And which control method should be used depending on the available materials and measurements?
- What is the difference between the control methods in terms of metabolic performance? Difficulty of implementation? Control performance?

Outline:

This thesis is organized as follows:

An introduction to *Escherichia coli*, the bacterial system studied in this thesis, is given in chapter 1 . The physiological and metabolic aspects of the microorganism are presented and discussed. The chapter also discusses the overflow metabolism phenomenon and acetate excretion via the fermentation pathways. The chapter includes a general presentation of a bioprocess and its main components, along with the different operating modes used for cell cultivation.

In chapter 2, a macroscopic representation of the dynamics and the kinetics of bioprocesses is presented. These dynamic models use reaction schemes and mass balance principles to derive a state-space representation of the biological system. Additionally, kinetic models used to describe the different reaction rates are presented in this chapter. Lastly, the state-space model for fed-batch *E. coli* cultures is obtained using the general modeling approach. The model dynamics are illustrated in simulation runs showcasing the different metabolic regimes of the bioprocess.

Chapter 3 is dedicated to state estimation. The different software sensor configurations found in the literature are introduced. Then the estimation of the state variables and kinetic parameters in the studied bioprocess is discussed. For this task, the Kalman filtering methods are presented and then implemented to estimate the acetate and glucose concentrations based on the biomass concentration. The efficiency of these algorithms is illustrated in simulation runs.

Chapter 4 includes a presentation of the developed closed-loop system for the lab-scale bioreactor. This system comprises a real-time monitoring software solution, control and estimation blocks, and a peristaltic pump control interface. This program allows the testing and validation of the algorithms presented in this thesis. The chapter also includes a description of the bioprocess hardware, materials, methods, and protocols used during the experiments.

The different control strategies found in the literature for fed-batch bioprocesses are presented in chapter 5. This presentation highlights the difference between the control methods depending on the requirements, complexity, and control objectives in order to provide a guide for choosing the appropriate method for the studied bioprocess. The chapter then discusses the control objectives of the

current study on fed-batch *E. coli* cultures. The available hardware setup and the availability of a reliable process model guided the control strategies presented in the following chapters.

Chapter 6 introduces an adaptive biomass regulation strategy based on the Generic Model Control method (GMC). The GMC algorithm is presented and then applied to the *E. coli* model to track a predefined biomass concentration trajectory corresponding to a specific growth rate chosen to satisfy the control objectives and avoid acetate accumulation. A model order reduction is applied to avoid using the kinetic terms and ensure low feeding rates. Parameter adaptation is performed using the linear Kalman filter to estimate the unmeasured kinetic terms and adapt for unpredictable dynamics. The strategy is validated through simulation runs and experiments using the BL21(DE3) *E. coli* strain.

Chapter 7, a robust variation of the Generic Model Control strategy is presented and applied to regulate the acetate concentration at a defined low level. A robust design procedure using the $\mathcal{LMI}$ formalism is presented to account for model mismatch while ensuring the desired closed-loop transient response. The robust GMC controller is combined with the state estimation by the UKF, and the strategy is validated both in simulation runs and through Fed-batch experiments.

Chapter 8 discusses implementing the nonlinear model predictive control (NMPC) strategy to regulate the acetate concentration to a low level. A control vector parametrization (CVP) approach is used to reduce the complexity of the optimization problem and improve the calculation efficiency. The NMPC strategy is coupled to the UKF estimator and validated through simulations and fed-batch experiments on the lab-scale reactor. A comparison with the robust GMC structure is given in this chapter.

Finally, Chapter 9 draws the main conclusions and perspectives of this work.

Contributions:

The main contributions of this work are:

- The development of three control strategies for fed-batch *E. coli* cultures with the objective of avoiding acetate accumulation and driving the culture near the optimal operating conditions. The strategies varied from linearizing control to nonlinear predictive control.

- The transformation of the lab-scale bioreactor from an open-loop process to a reliable closed-loop system with flexible monitoring tools allowing the acquisition of several measurements from different manufacturers. The same interface includes tools that facilitate the integration of advanced control and estimation algorithms, and feeding flow rate manipulation.

- The experimental validation of the proposed regulation strategies on lab-scale fed-batch BL21(DE3) *E. coli* cultures. Providing a proof of concept for future implementations on higher scale reactors.

Publications:

The thesis work has resulted in the following publications to international journal and conferences:

Journal articles

Abadli, Merouane, Laurent Dewasme, Sihem Tebbani, Didier Dumur, and Alain Vande Wouwer (July 2020). "Generic model control applied to E. coli BL21(DE3) Fed-batch cultures". In: *Processes* 8.7, p. 772. ISSN: 22279717.

Abadli, Merouane, Laurent Dewasme, Sihem Tebbani, Didier Dumur, and Alain Vande Wouwer (2021). "An Experimental Assessment of Robust control and Estimation of Acetate Concentration in Escherichia coli BL21 (DE3) Fed-batch Cultures". In: *Biochemical Engineering Journal*, p. 108103.

Conferences with proceedings

Abadli, Merouane, Laurent Dewasme, Didier Dumur, Sihem Tebbani, and Alain Vande Wouwer (Oct. 2019). "Generic model control of an Escherichia coli fed-batch culture". In: *2019 23rd International Conference on System Theory, Control and Computing, ICSTCC 2019 - Proceedings*. Institute of Electrical and Electronics Engineers Inc., pp. 212–217. ISBN: 9781728106991.

Conferences with abstracts

Abadli, Merouane, Sihem Tebbani, Didier Dumur, Laurent Dewasme, and Alain Vande Wouwer (Mar. 2018a). "Nonlinear model predictive control of Escherichia coli culture". In: *37th Benelux Meeting on Systems and Control*. Soesterberg, Netherlands.

Presentations

Abadli, Merouane, Sihem Tebbani, Didier Dumur, Laurent Dewasme, and Alain Vande Wouwer (June 2018b). "Nonlinear model predictive control of Escherichia coli cultures". In: *Groupe de travail sur le thème de : La Commande Prédictive Non Linéaire*. Châtillon, France.

Résumé en français

Contexte et motivations

Escherichia coli est la cellule hôte la plus utilisée dans la production biopharmaceutique industrielle. Jusqu'à un tiers des protéines thérapeutiques approuvées sont produites par la fermentation fed-batch à haute densité cellulaire de souches génétiquement modifiées de *E. coli* (Baeshen et al., 2015). Cela découle des différentes propriétés biologiques de ce micro-organisme. La souplesse des conditions de culture, la rapidité de la croissance, le rendement élevé et la facilité de mise à l'échelle du procédé ont fait de *E. coli* l'hôte principal dans l'industrie biotechnologique (Lee, 1996; Pontrelli et al., 2018).

Escherichia coli est une bactérie hétérotrophe de la famille des entérobactéries. Elle peut effectuer des métabolismes complexes et survivre dans des conditions de stress et de culture difficiles. Elle peut se développer à différentes conditions de température de pH et se multiplier en utilisant diverses sources de carbone en présence d'une quantité élevée ou limitée d'oxygène. Le glucose est considéré comme la principale source de carbone dans le métabolisme de l'*E. coli*.

En cas de croissance aérobique sur le glucose, les cellules d'*E. coli* peuvent produire de l'acétate par la voie fermentative. Cependant, des complications peuvent survenir pendant la phase de croissance exponentielle. La sécrétion d'acétate dans le milieu de culture peut inhiber la croissance cellulaire à des concentrations élevées (Eiteman and Altman, 2006). L'inhibition provient de la diminution de l'efficacité respiratoire en cas d'excès de glucose. Ce phénomène est connu sous le nom de métabolisme de débordement ou effet Crabtree bactérien (Crabtree, 1929; De Deken, 1966; Doelle et al., 1982). Par conséquent, il est indispensable de déterminer une stratégie d'alimentation qui favorise la croissance des cellules et évite l'accumulation de l'acétate dans le milieu de culture.

Le but du travail présenté dans cette thèse est de proposer et de développer des solutions pratiques pour éviter le métabolisme de débordement et maximiser la productivité de la biomasse dans les cultures fed-batch de *E. coli*. Cet objectif est atteint à travers des stratégies de commande et d'estimation en boucle fermée en déterminant le taux d'alimentation approprié qui conduit la culture près des conditions métaboliques optimales.

Modèle dynamique des cultures fed-batch de *E. coli*

Le modèle macroscopique de la croissance de *Escherichia coli* est présenté ci-après. Le schéma réactionnel qui décrit la croissance cellulaire de *E. coli* sur le glucose

en conditions aérobiques est composé de trois voies cataboliques (Retamal et al., 2018; Rocha and Ferreira, 2002) :

$$S + k_{O1}O \xrightarrow{\mu_1\ X} k_{X1}X + k_{C1}C \tag{1a}$$

$$S + k_{O2}O \xrightarrow{\mu_2\ X} k_{X2}X + k_{A2}A + k_{C2}C \tag{1b}$$

$$A + k_{O3}O \xrightarrow{\mu_3\ X} k_{X3}X + k_{C3}C \tag{1c}$$

où

- S, O, X, C, et A représentent respectivement les concentrations de glucose (substrat), d'oxygène, de biomasse, de dioxyde de carbone et d'acétate.
- $k_{\xi i}$ $(\xi = [X\ S\ A\ O\ C]^\top;\ i = 1,2,3)$ sont les coefficients pseudo-stoechiométriques.
- μ_j $(j = 1,2,3)$ sont les taux de croissance spécifiques.

La croissance des cellules *E. coli* est modélisée suivant la théorie du goulot d'étranglement de Sonnleitner et Käppeli (Sonnleitner and Käppeli, 1986). La théorie du goulot d'étranglement suppose que les cellules sont susceptibles de changer leur métabolisme en raison de leur capacité respiratoire limitée, ce qui entraîne un métabolisme de débordement contrôlé par le niveau de substrat.

Si la concentration en substrat est supérieure au seuil critique correspondant à la capacité oxydative disponible ($S > S_{crit}$), l'acétate est produit par les cellules par la voie métabolique fermentaire. La culture est dite en régime oxydo-fermentaire (réactions (1a) et (1b)).

D'autre part, l'acétate (s'il est présent dans le milieu de culture) est consommé lorsque la concentration en substrat est inférieure au niveau critique ($S < S_{crit}$), et la culture est dite en régime oxydatif (réactions (1a) et (1c)).

Lorsque la concentration en substrat est au niveau critique et remplit exactement la capacité respiratoire, la culture est optimale, correspondant à la limite entre les deux régimes de fonctionnement, et l'acétate n'est ni produit ni consommé. Le modèle cinétique pour les taux spécifiques est basé sur ces régimes de fonctionnement :

$$\mu_1 = min(q_s, q_{s_{crit}}) \tag{2a}$$

$$\mu_2 = max(0, q_s - q_{s_{crit}}) \tag{2b}$$

$$\mu_3 = max(0, q_{AC}) \tag{2c}$$

où μ_1, μ_2, et μ_3 sont les taux spécifiques liés aux réactions cataboliques décrivant l'oxydation du substrat (1a), la production d'acétate (fermentation) (1b), et l'oxydation de l'acétate (1c) (Bastin and Dochain, 1990).

Les termes cinétiques liés aux taux de consommation q_j sont définis par :

$$q_s(S) = q_{s_{max}} \frac{S}{K_s + S} \tag{3a}$$

$$q_{s_{crit}}(A) = q_{O_{max}} \frac{K_{iA}}{K_{iA} + A} \tag{3b}$$

$$q_{AC}(S, A) = (q_{s_{crit}} - q_s) \frac{A}{K_A + A} \tag{3c}$$

où

- q_s et q_{AC} représentent respectivement les taux de consommation du substrat et de l'acétate.
- $q_{s_{crit}}$ représente le taux de consommation critique du substrat.
- $q_{S\,\mathrm{max}}$ représente le taux de consommation maximal de glucose.
- q_{Omax} représente la valeur maximale de la capacité respiratoire.

En analysant le bilan massique du schéma réactionnel (1), on obtient les équations différentielles suivantes (Retamal et al., 2018) :

$$\dot{X} = (k_{X1}\mu_1 + k_{X2}\mu_2 + k_{X3}\mu_3)X - D\,X \tag{4a}$$

$$\dot{S} = -(\mu_1 + \mu_2)X - D\,(S - S_{in}) \tag{4b}$$

$$\dot{A} = (k_{A2}\mu_2 - \mu_3)X - D\,A \tag{4c}$$

$$\dot{O} = -(k_{O1}\mu_1 + k_{O2}\mu_2 + k_{O3}\mu_3)X - D\,O + OTR \tag{4d}$$

$$\dot{C} = (k_{C1}\mu_1 + k_{C2}\mu_2 + k_{C3}\mu_3)X - D\,C - CTR \tag{4e}$$

$$\dot{V} = F_{in} \tag{4f}$$

où

- V est le volume du milieu de culture.
- F_{in} est le débit d'alimentation d'entrée.
- D est le taux de dilution ($D = \frac{F_{in}}{V}$).
- S_{in} est la concentration en glucose dans le milieu d'alimentation.
- $\mu_{\{1,2,3\}}$ sont les taux spécifiques donnés par equations (2) and (3c),.

Les taux de transfert de gaz OTR et CTR peuvent être modélisés par les équations suivantes :

$$OTR = k_L a_O\,(O_{sat} - O) \tag{5}$$

$$CTR = k_L a_C\,(C - C_{sat}) \tag{6}$$

où

- $k_L a_O$ et $k_L a_C$ sont respectivement les coefficients de transfert volumétrique des concentrations d'oxygène et de dioxyde de carbone dissous.

- O_{sat} et C_{sat} sont respectivement les concentrations d'oxygène et de dioxyde de carbone dissous à saturation.

Commande adaptative par modèle générique (GMC) de la concentration en biomasse

Le métabolisme de débordement et l'accumulation d'acétate conduisent à la diminution du rendement de la production de biomasse et par conséquent à la diminution de la production de protéines recombinantes (Riesenberg et al., 1991; Rothen et al., 1998).

Selon la théorie du goulot d'étranglement, afin de maximiser la productivité de la biomasse et d'éviter un métabolisme de débordement, la concentration en substrat doit être maintenue à un certain seuil critique correspondant à la capacité critique d'oxydation des cellules (Jana and Deb, 2005). Pour atteindre cet objectif, une stratégie d'alimentation en boucle fermée est nécessaire pour maintenir le bioprocédé près des conditions de fonctionnement optimales.

Une formulation simple du problème de commande consiste à réguler les concentrations de substrat ou d'acétate à de faibles valeurs. Cependant, le manque d'outils fiables de mesure en ligne des concentrations d'acétate et de glucose constitue un obstacle majeur à l'application de ces stratégies, puisque le niveau critique de la concentration de glucose dans les cultures d'*E. coli* est très faible par rapport à la sensibilité des sondes disponibles sur le marché.

Nous proposons une stratégie de commande adaptative basée sur la linéarisation des équations du modèle non linéaire, appelée Commande par Modèle Générique (GMC) (Lee and Sullivan, 1988). L'objectif est de bénéficier de la mesure en ligne de la concentration de la biomasse afin de développer et d'implémenter un algorithme GMC pour contrôler la productivité de la biomasse dans une fermentation fed-batch de *E. coli*.

Dans cette stratégie de commande que nous proposons, une trajectoire prédéfinie de biomasse correspondant à une production limitée d'acétate est imposée par le régulateur. Les avantages de cette approche sont l'inclusion du modèle non linéaire du bioprocédé dans la conception de loi de commande et la compensation des incertitudes du modèle par une adaptation en ligne utilisant un estimateur de paramètres.

Une mise en œuvre expérimentale de la stratégie de commande est effectuée sur un bioréacteur de laboratoire afin de tester ses performances et sa robustesse dans des conditions réelles d'exploitation.

Commande par modèle générique (GMC)

La commande par modèle générique est basée sur la linéarisation par retour de sortie, incluant les non-linéarités du système dans la conception de la loi de commande. L'objectif principal du schéma de commande est de suivre une trajectoire nominale de sortie désirée (Peter and Lee, 1993). Considérons le système non linéaire suivant :

$$\dot{x} = f(x) + g(x)u \tag{7}$$
$$\dot{y} = h(x) \tag{8}$$

où

- $x \in \mathbb{R}^n$ est le vecteur d'état
- $u \in \mathbb{R}$ est l'entrée de commande
- $y \in \mathbb{R}$ est la sortie du système.
- $f : \mathbb{R}^n \to \mathbb{R}^n$, $g : \mathbb{R}^n \to \mathbb{R}^n$ sont des fonctions non linéaires des états x,
- $h : \mathbb{R}^n \to \mathbb{R}$ est la fonction de sortie.

D'après Equation (8), la dynamique de la sortie est donnée par (Isidori et al., 1995) :

$$\dot{y} = \frac{\partial h}{\partial x}[f(x) + g(x)u] = L_f h(x) + L_g h(x)u \tag{9}$$

où

- $L_f h(x) = \frac{\partial h}{\partial x} f(x)$ est la dérivée de Lie de h le long de f.
- $L_g h(x) = \frac{\partial h}{\partial x} g(x)$ est la dérivée de Lie de h le long de g.

Dans la procédure de conception de la GMC, la sortie y est comparée à une trajectoire de référence prédéterminée y_{ref}. L'équation de sortie peut alors être définie à l'aide d'un régulateur proportionnel-intégral sous la forme :

$$\dot{y} = \hat{u} = G_1(y_{ref} - y) + G_2 \int_0^t (y_{ref} - y)\partial\tau \tag{10}$$

où G_1 et G_2 sont des gains de commande (constants par rapport au temps). Leur réglage est effectué en fonction du comportement dynamique souhaité. Si $L_g h(\mathrm{x}) \neq 0$ (c'est-à-dire que le système est de degré relatif 1), la commande satisfaisant equations (9) and (10) est dérivée de l'équation suivante :

$$u = \frac{1}{L_g h}\left(-L_f h + \hat{u}\right) \tag{11}$$

La réponse désirée de la boucle fermée est définie en fixant le coefficient d'amortissement ξ et la pulsation propre ω_0. G_1 et G_2 sont réglés de manière à conférer les ξ et ω_0 désirés.

Application de la stratégie GMC aux cultures de *E. coli*

L'objectif de la commande de la culture fed-batch de *E. coli* est de favoriser la production de biomasse, d'atteindre des densités cellulaires élevées et de maximiser la productivité de la biomasse.

Nous proposons de réguler le taux de croissance de la biomasse en suivant une trajectoire sous-optimale prédéterminée satisfaisant les objectifs de commande et maintenant la culture proche des conditions optimales. L'avantage de cette approche est son faible coût d'exploitation et sa praticité, puisqu'elle repose uniquement sur la mesure en ligne de la biomasse qui est fournie par la sonde turbidimétrique avec un faible bruit de mesure.

Le taux de croissance ciblé μ_{set} correspond à une concentration de substrat inférieure à la valeur critique ($S_{set} < S_{crit}$), et à une concentration initiale en acétate égale à zéro. Cette trajectoire de fonctionnement permet au procédé d'évoluer près de la limite entre les modes oxydatif et oxydo-fermentatif, avec une marge de sécurité pour éviter les commutations métaboliques et favoriser la croissance cellulaire.

L'application directe de la GMC au modèle macroscopique soulève quelques problèmes. La détermination précise des taux de croissance spécifiques est difficile, car la cinétique est basée sur le principe du métabolisme de débordement représenté par des commutations métaboliques entre les deux régimes. De plus, une trajectoire imposée de biomasse pourrait éventuellement conduire à des valeurs élevées du débit d'alimentation.

Conception de la commande GMC à l'aide d'un modèle réduit

Une commande basée sur un modèle réduit est développée en appliquant la technique de perturbation singulière : la dynamique du substrat, de l'oxygène et du dioxyde de carbone est considérée plus rapide que celle de la biomasse et de l'acétate. Ainsi, les variables rapides sont considérées comme étant en quasi-état d'équilibre et leur dynamique est mise à zéro.
Sous ces hypothèses, l'équation suivante est obtenue pour la concentration de la biomasse :

$$\dot{X} = -\bar{k_{11}}\,\frac{F_{in}}{V}\,S_{in} - \bar{k_{12}}\,OTR + \bar{k_{13}}\,CTR - \frac{F_{in}}{V}\,X \tag{12}$$

où les paramètres $\bar{k_{11}}$, $\bar{k_{12}}$, et $\bar{k_{13}}$ sont des fonctions des paramètres du modèle. En appliquant le schéma de la GMC, on obtient la loi de commande suivante :

$$F_{in} = \frac{-\bar{k_{12}}\,OTR + \bar{k_{13}}CTR - \hat{F}}{X + \bar{k_{11}}\,S_{in}}\,V \tag{13}$$

$$\hat{F} = G_1(X_{ref} - X) + G_2\int_0^t (X_{ref} - X)\partial\tau \tag{14}$$

En supposant que $X + \bar{k_{11}}\,S_{in} \neq 0$, ce qui est satisfait en général.

L'avantage de la réduction du modèle est que la condition de fonctionnement désirée (faible concentration de substrat) est directement intégrée dans l'algorithme de commande.

Étant donné que OTR et CTR ne sont pas disponibles pour une mesure en ligne dans notre montage expérimental et que la biomasse X est la seule variable mesurée, une stratégie adaptative de la GMC est développée. Les signaux non disponibles sont reconstruits à base de mesures, en adaptant également la loi de commande soumise à l'incertitude des paramètres.

GMC adaptative

La loi de commande de (14) comporte la variable non mesurable et incertaine suivante : $-\bar{k_{12}}\, OTR + \bar{k_{13}} CTR$. Le dispositif expérimental n'étant pas équipé d'analyseurs de gaz, un algorithme d'estimation des paramètres cinétiques est développé. L'équation de la dynamique de la biomasse (12) peut être réécrite comme suit :

$$\dot{X} = \gamma \; - D \; (X + \bar{k_{11}} S_{in}) \tag{15}$$

où $D = \frac{F_{in}}{V}$, et γ est le paramètre incertain et non mesurable variant dans le temps donné par :

$$\gamma = -\bar{k_{12}}\, OTR + \bar{k_{13}} CTR \tag{16}$$

À condition que $X^* = X + \bar{k_{11}} S_{in}$ soit disponible pour une mesure en ligne, γ peut être estimé à l'aide d'un filtre de Kalman linéaire de la même manière que celle présentée dans (Gonzalez et al., 2016).

Un filtre de Kalman discret (Welch and Bishop, 1995) peut être appliqué pour estimer l'évolution de X^* et de γ . La structure de commande mise à jour est décrite dans la Figure 1.

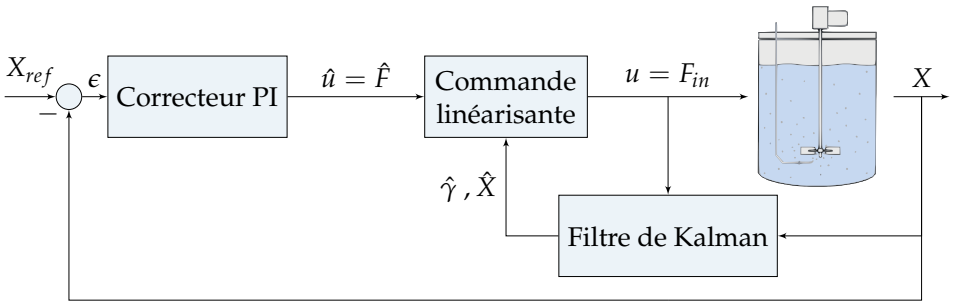


FIGURE 1: Commande GMC couplée au filtre de Kalman

La loi de commande, après inclusion du paramètre estimé $\hat{\gamma}$ devient :

$$F_{in} = \frac{\hat{\gamma} - \hat{F}}{X + \bar{k_{11}}\, S_{in}}\, V \tag{17}$$

Résultats expérimentaux

Des expériences fed-batch sont réalisées pour tester la stratégie de commande dans des conditions expérimentales réelles. Les paramètres de commande G_1 et G_2 ont été réglés en simulation, le temps de réponse choisi t_r est égal à 1 h ($\omega_0 = 3rad/h$), et le rapport d'amortissement est fixé à $\xi = 1$.

Les concentrations en biomasse, glucose, acétate, ainsi que le débit d'alimentation sont représentés dans la Figure 2. La concentration de biomasse commence à partir de 0.3 g/L et atteint 1.5-1.7 g/L à la fin de la phase batch, caractérisée par un épuisement du glucose. La phase de fed-batch commence à 6 h, et l'algorithme de commande est lancé.

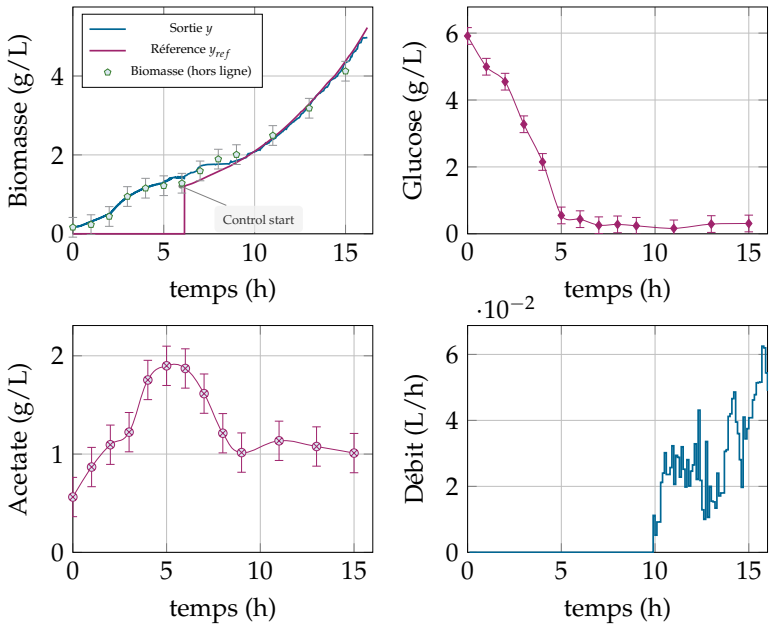


FIGURE 2: Evolution de la biomasse mesurée, du profil de référence, des concentrations de glucose et d'acétate, et du débit d'alimentation

La taux de croissance de référence imposé est $\mu_{set} = 0,18$ h $^{-1}$. Dans la Figure 2, la biomasse maintient une croissance exponentielle proche de la trajectoire de référence montrant que la régulation est efficace. La concentration en glucose reste proche de zéro et presque constante pendant la phase d'alimentation des deux expériences, confirmant l'hypothèse de dynamique rapide présentée dans la section de réduction du modèle ($\dot{S} = 0$).

La concentration d'acétate reste inférieure à 2 g/L pendant l'expérience. L'évolution indique un basculement métabolique entre les modes oxydatif et oxydo-fermentatif, et lorsque le glucose est presque épuisé à t = 5h, la concentration d'acétate commence à diminuer, c'est-à-dire que la culture est en régime oxydatif.

Les performances du régulateur à modèle générique, en termes de robustesse, sont satisfaisantes. Le contrôleur est capable de maintenir l'erreur de poursuite de la biomasse proche de zéro dans les deux expériences malgré l'inadéquation

du modèle résultant des incertitudes de modélisation. Le régulateur parvient à s'adapter aux variations du signal de la biomasse en agissant sur le débit d'alimentation.

Discussion

Les expériences ont montré que la combinaison de la commande par modèle générique et de l'adaptation des paramètres permet d'atteindre les objectifs de commande en temps réel. Le suivi de la trajectoire de la biomasse est effectué de manière adéquate malgré la présence de perturbations et d'incertitudes sur les paramètres du modèle.

La stratégie combinée GMC-Kalman présente plusieurs caractéristiques intéressantes. Tout d'abord, les coûts de développement et de conception sont assez minimes. Un autre avantage est le fait que la loi de commande est calculée à l'aide d'équations algébriques simples, et ne nécessite pas la résolution en temps réel d'équations différentielles non linéaires complexes. Cette caractéristique réduit la complexité de calcul du schéma de commande et le rend facilement intégrable dans la plupart des bioréacteurs.

La disponibilité des mesures de la biomasse rend le schéma de commande très pratique. L'estimation des variables d'état n'est pas nécessaire puisque la variable mesurée est la variable commandée, et l'estimation des paramètres est effectuée à l'aide d'un filtre de Kalman linéaire. L'adaptation en ligne des paramètres cinétiques renforce la robustesse du système en boucle fermée face aux dynamiques imprévisibles.

Cependant, malgré ses caractéristiques intéressantes, la stratégie de commande présente certaines limites concernant les performances métaboliques. La commande vise à réguler la concentration de biomasse avec un taux de croissance défini inférieur au taux critique pour éviter le métabolisme de débordement. Selon la théorie du goulot d'étranglement, ce taux de croissance sous-optimal peut correspondre aux régimes oxydatif ou oxydo-fermentatif. Cependant, le comportement observé est que les cellules fonctionnent principalement en régime oxydatif.

Bien que l'accumulation d'acétate soit évitée en fonctionnant dans ce mode, les écarts par rapport au taux de croissance de référence (dus à une forte inadéquation du modèle, à une variation des paramètres due aux conditions d'oxygénation ou à de fortes perturbations des mesures de la biomasse) peuvent entraîner une baisse du rendement de production de la biomasse et de la productivité de la biomasse par rapport aux valeurs théoriques attendues. De plus, cette déviation peut également provoquer une accumulation d'acétate si la culture passe en mode oxydo-fermentatif sans aucune indication en ligne pour l'utilisateur sur le signal de la biomasse.

Une solution pratique à ce problème est de réguler la concentration d'acétate à une faible valeur, car elle est directement liée au taux de croissance optimal. Cette approche nécessite une estimation robuste des variables d'état, ce qui augmente la complexité de la stratégie de commande, mais améliore en revanche la productivité du procédé.

Commande robuste par modèle générique (GMC) de la concentration d'acétate

Cibler un faible taux de croissance de biomasse offre une solution pour éviter l'accumulation d'acétate. Cependant, cela ne permet pas d'atteindre le potentiel du bioprocédé car ce choix entraîne une faible productivité et un temps de culture élevé (Srinivasan et al., 2001). Une alternative consiste à réguler la concentration de glucose ou d'acétate à des niveaux spécifiques (Dewasme et al., 2011a,b; Santos et al., 2012a). Le principal défi dans les cultures fed-batch de *E. coli* est la difficulté de mise en œuvre en raison de l'exigence de mesures précises à faibles concentrations d'acétate et/ou de glucose.

Commande GMC de la concentration d'acétate

Nous proposons une autre solution pour éviter le métabolisme de débordement en régulant la concentration d'acétate autour d'une valeur basse A_{ref}. Cette valeur doit être choisie aussi proche de zéro que possible afin de maintenir le bioprocédé près de la limite métabolique optimale. D'autre part, une marge de sécurité doit être prise pour éviter les commutations métaboliques entre les régimes opératoire (Dewasme et al., 2011a).

Comme la variable commandée (acétate) n'est pas disponible pour la mesure en ligne, un algorithme d'estimation d'état est nécessaire. Un filtre de Kalman non parfumé (UKF) est mis en œuvre pour estimer les concentrations d'acétate et de glucose à base du modèle du procédé et des mesures de biomasse.

Le schéma de la GMC présenté précédemment est appliqué pour réguler la concentration d'acétate dans des cultures fed-batch de *E. coli*. En considérant la concentration d' acétate comme la sortie contrôlée, et en supposant sa disponibilité pour la mesure $(y = A)$, la loi de commande suivante est obtenue :

$$F_{in} = V\frac{\hat{u} + (k_{A2}\mu_1 + \mu_3)\, X}{k_{A2}S_{in} - A} \tag{18}$$

$$\hat{u} = G_1(A_{ref} - A) + G_2\int_0^t (A_{ref} - A)\partial\tau \tag{19}$$

où $\theta = (k_{A2}\mu_1 + \mu_3)$ est un terme cinétique supposé incertain.

Conception d'une commande robuste

Les incertitudes structurelles et paramétriques ainsi que les erreurs d'estimation peuvent être regroupées dans une erreur paramétrique globale :

$$\delta = \bar{\theta} - \theta \tag{20}$$

où δ est une fonction non linéaire de (S, A, O) représentant les éventuelles déviations des termes non linéaires dues aux incertitudes du modèle, et $\bar{\theta}$

représente la valeur nominale exacte (inconnue). Suivant une approche similaire à celle développée dans (Dewasme et al., 2011a), le paramètre incertain δ est supposé borné et appartenant à l'ensemble Δ défini par :

$$\Delta := \{\delta : \underline{\delta} \leq \delta \leq \bar{\delta}\} \tag{21}$$

avec $\underline{\delta}$ et $\bar{\delta}$ représentant respectivement les valeurs minimale et maximale de l'ensemble polytopique borné Δ.

Les paramètres de commande G_1 et G_2 sont conçus pour assurer certaines caractéristiques de robustesse et performance du système global en boucle fermée.

Contraintes de robustesse

Le problème de conception de la commande robuste consiste à déterminer les paramètres du correcteur G_1 et G_2 de manière à minimiser la norme infinie de la fonction de transfert en boucle fermée (Chilali and Gahinet, 1996).

Le lemme borné réel (Chilali and Gahinet, 1996) pour les systèmes continus donne une formulation équivalente en LMI (inégalité matricielle linéaire) du problème de commande. La résolution de cette LMI permet de reconstruire le vecteur de retour d'état K qui stabilise le système en boucle fermée et compense la perturbation bornée δ.

Contraintes de performance

En plus d'assurer la robustesse de la boucle fermée, il est désiré d'obtenir certaines performances en termes de réponse transitoire (par exemple, amortissement, temps de réponse, etc.). En d'autres termes, des contraintes sont ajoutées sur l'emplacement des pôles de la boucle fermée.

Les contraintes de placement des pôles peuvent être exprimées à l'aide de régions LMI, qui sont connues pour avoir des propriétés géométriques intéressantes (convexité, symétrie, ...) (Chilali and Gahinet, 1996). Une région appropriée satisfaisant ce critère est l'intersection du demi-plan $s < -\rho < 0$, du disque de rayon r et du secteur conique défini par un angle Θ. La région correspondante $\mathcal{S}(\rho, r, \Theta)$ est définie comme suit :

$$\mathcal{S}(\rho, r, \Theta) = \{a < -\rho < 0, \quad |s = a + jb| < r, \quad a\tan(\Theta) < -|b|\} \tag{22}$$

De cette façon, il est possible de fixer un taux de convergence minimal ρ, un rapport d'amortissement minimal $\zeta = cos(\Theta)$, et une pulsation propre amortie maximale $\omega_d = rsin(\Theta)$ (Wood, 1972). Notre problème de conception de la commande consiste alors à trouver un gain de retour d'état K qui :

- garantit la performance H_∞.
- place les pôles en boucle fermée dans la région LMI $\mathcal{S}(\rho, r, \Theta)$ définie par (22).

Résultats expérimentaux

Des expériences de commande ont été réalisées pour tester les performances et la robustesse de la stratégie GMC robuste développée dans un environnement en temps réel. Chaque expérience comprenait une phase batch suivie d'une phase fed-batch (phase de commande). Les résultats d'une expérience sont présentés dans la Figure 3, qui illustre l'évolution de la biomasse mesurée (en ligne & hors ligne), des concentrations de glucose et d'acétate (hors ligne), et leurs estimations, ainsi que le débit d'alimentation (sortie du régulateur).

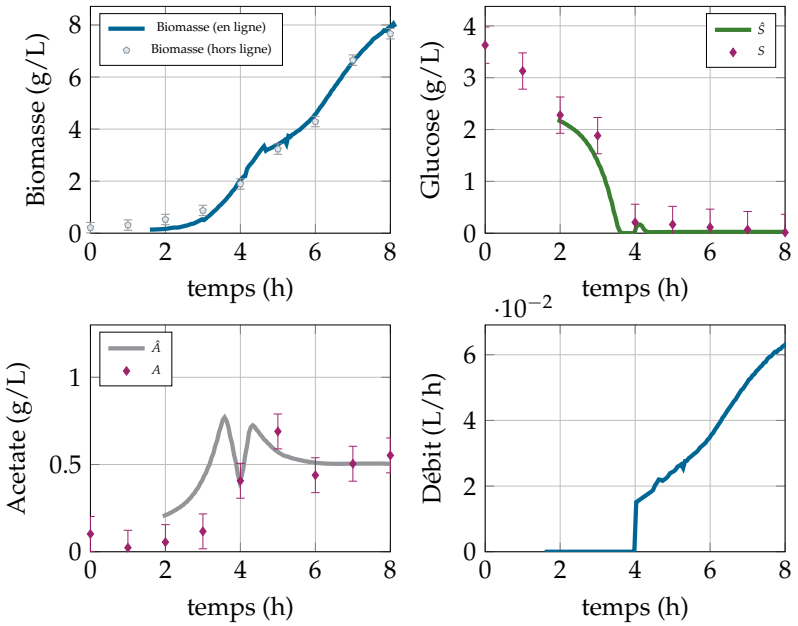


FIGURE 3: Evolution de la biomasse mesurée, des concentrations estimées et mesurées de glucose et d'acétate, et du débit d'alimentation

.

Après avoir atteint les conditions opératoires désirées, le réacteur est inoculé avec la préculture, et la phase batch commence. Comme le montre Figure 3, la biomasse suit une croissance exponentielle et atteint jusqu'à 2 g/L. Comme la consommation de glucose entraîne la production d'acétate, la culture est en mode respiro-fermentaire.

La phase batch a duré 4 heures après la consommation de la majorité du glucose dans le milieu. La phase de fed-batch a commencé juste après la mesure d'une faible concentration de glucose. Le contrôleur GMC est lancé après la mise en place de la référence acétate et des paramètres de commande. La solution d'alimentation est injectée par le contrôleur et les cellules reprennent leur croissance. La phase de fed-batch se poursuit jusqu'à atteindre la limite de saturation de la sonde turbidimétrique (autour d'une concentration de biomasse de 8 g/L).

La mesure en ligne de la concentration de la biomasse fournie par la sonde turbidimétrique, et le modèle cinétique avec les valeurs des paramètres identifiés du (Retamal et al., 2018) sont utilisés pour estimer les concentrations d'acétate et

de glucose en utilisant l'UKF. L'estimation est lancée pendant la phase batch après avoir atteint une concentration de biomasse mesurée supérieure à 0,2 g/L pour garantir une bonne précision d'estimation.

Les performances de l'UKF dans la phase fed-batch sont satisfaisantes, malgré les erreurs d'initialisation et les incertitudes du modèle. Les estimations du glucose et de l'acétate correspondent très bien aux mesures hors ligne pendant la période de commande, et la convergence est atteinte en moins d'une heure. Les valeurs de l'erreur quadratique moyenne de l'estimation pour chaque état estimé (le substrat et l'acétate) sont du même niveau de la sensibilité de la mesure (0,1 g/L).

Comme on peut le voir dans Figure 3, l'accumulation d'acétate est évitée, et la concentration est limitée à moins de (1 g/L) pendant la phase fed-batch. La concentration estimée en acétate converge vers la référence désirée, en respectant le temps de stabilisation choisi.

Discussion

La stratégie de commande présentée constitue une approche pratique pour éviter le métabolisme de débordement dans les cultures fed-batch de *E. coli*. Cependant, elle offre une solution sous-optimale, car la régulation de la concentration de substrat au niveau critique n'est pas pratique en raison du manque de mesures précises en ligne.

Afin d'évaluer l'efficacité de l'approche proposée, une comparaison est effectuée en simulation avec la stratégie de régulation du taux de croissance présentée dans la section précédente. À cette fin, nous avons réglé la régulation de la biomasse pour suivre un taux de croissance défini μ_{set} choisi à 90% de la valeur maximale théorique (μ_{max} = 0,26 L/h), correspondant à la concentration critique de substrat et à la capacité oxydative maximale. D'autre part, nous fixons la régulation de l'acétate pour suivre une référence de 0,5 g/L. Nous supposons que la concentration d'acétate est mesurée avec un bruit blanc additif d'une déviation standard de 0,05 g/L.

Tout d'abord, nous supposons que les paramètres du modèle et le taux de croissance maximal μ_{max} sont parfaitement connus. Ensuite, nous introduisons une variation fixe de la capacité oxydative maximale q_{Omax} qui est directement liée à la valeur du taux de croissance maximal.

TABLE 1: L'effet de la variation des paramètres sur la performance de la commande

Variation de q_{Omax}	$\frac{\mu_X}{\mu_{max}}$% (GMC-X)	$\frac{\mu_X}{\mu_{max}}$% (GMC-A)	S_{crit} (g/L)
0%	89%	81%	0.0375
10%	81%	85%	0.046
20%	75%	89%	0.0529
30%	70%	93%	0.0628

Dans le cas du modèle idéal (aucune variation paramétrique), la régulation de la croissance de la biomasse (GMC-X) présente une performance globale légèrement supérieure. En revanche, la régulation de la concentration en acétate (GMC-A) à 0.5 L/h conduit à un taux de croissance de la biomasse de 0,21 h^{-1} correspondant à 81% de la valeur maximale comme on peut le voir dans Table 1. Ce résultat montre que la présence d'acétate dans le milieu réduit le taux de croissance de la biomasse, en raison d'un taux de consommation de substrat plus faible causé par l'activation des voies de consommation de l'acétate selon la théorie du goulot d'étranglement. Cependant, le maintien de l'acétate à une faible concentration réduit son effet inhibiteur, et maintient la culture proche des conditions optimales. L'introduction d'une variation de 20 % de q_{Omax} entraîne une augmentation de la concentration critique de substrat S_{crit} et par conséquent du taux de croissance maximal μ_{max}.

Malgré l'inadéquation du modèle, la régulation du taux de croissance de la biomasse présente une bonne performance dans le suivi du taux de référence. Cependant, elle ne correspond qu'à 75% de la nouvelle valeur maximale, et la productivité de la biomasse est donc également inférieure à sa valeur optimale par rapport au cas nominal. Ceci est dû à l'augmentation de l'écart entre le taux de référence μ_{set} et le nouveau taux de croissance maximal μ_{max}.

La régulation de l'acétate, en revanche, offre une performance plus cohérente, et donne un meilleur rapport de taux de croissance (89%). De plus, le rapport de taux de croissance est plus élevé avec une variation croissante de la capacité oxydative maximale, comme on peut le voir dans Table 1.

Ce résultat met en évidence un problème lié au choix d'un taux de croissance spécifique comme objectif de commande, car il nécessite une détermination précise de la valeur maximale, puis de cibler un taux de croissance plus faible pour éviter l'accumulation d'acétate. C'est une tâche difficile en raison de la nature incertaine des bioprocédés, car la variation des paramètres dépend de plusieurs facteurs tels que la variation des conditions opératoires entre les cultures. Si le taux de croissance maximal est sous-estimé, la productivité sous-optimale de la biomasse qui en résulte est inférieure à celle désirée. Si le taux de croissance maximal est surestimé, une régulation à 90% de cette valeur pourrait conduire à une accumulation d'acétate et à des commutations métaboliques, et donc à une inhibition de la croissance.

D'autre part, la régulation de la concentration d'acétate et son maintien à une faible valeur offrent un meilleur compromis pratique, puisque l'accumulation est évitée, et le taux de croissance obtenu est cohérent dans le cas d'une inadéquation du modèle. Il s'agit d'un résultat intéressant car l'approche de régulation de l'acétate est robuste face aux changements des conditions d'exploitation et n'est pas spécifique à la souche bactérienne. La stratégie pourrait être appliquée à une souche différente tout en assurant le même niveau de performance sans avoir besoin d'estimer μ_{max} avec précision.

Commande prédictive non linéaire de la concentration d'acétate

La commande prédictive non linéaire (NMPC) est souvent envisagée pour les problèmes de régulation de procédés suite à ses nombreux avantages et à sa large utilisation dans les applications industrielles (Forbes et al., 2015; Qin and Badgwell, 2000). La NMPC est une stratégie de commande à base de modèle développée pour les procédés non linéaires. Elle consiste à résoudre un problème d'optimisation en ligne sous un ensemble de contraintes. La NMPC peut prédire, en utilisant un modèle dynamique non linéaire du procédé, l'effet des valeurs de commande sur les variables commandées sur un horizon fini. Une formulation générale du problème NMPC consiste à minimiser une fonction quadratique (fonction coût) sur un horizon fini.

Dans cette section, la NMPC est appliquée pour maximiser la productivité de la biomasse dans des cultures fed-batch de *E. coli*, en régulant la concentration d'acétate à une faible valeur. L'algorithme NMPC est implémenté dans un environnement en temps réel, et l'effort de calcul est réduit grâce à la technique de paramétrisation du vecteur de commande (CVP) (Banga et al., 2005).

Le filtre de Kalman non parfumé (UKF) est utilisé pour fournir des estimations des concentrations d'acétate et de glucose sur la base de la mesure de la biomasse.

Commande prédictive non linéaire appliquée aux cultures de *E. coli*

De la même manière que pour la stratégie de commande précédente, nous désirons maintenir la concentration d'acétate à une valeur de consigne faible A_{ref} pour éviter son accumulation durant la culture. Cette régulation est obtenue en agissant sur le débit d'alimentation F_{in}, également contraint de suivre une trajectoire de référence pré-calculée F_{ref} afin de lisser le comportement de la commande.

Conception de la commande

Dans ce qui suit, une expression équivalente en temps discret du modèle continu de *E. coli* est utilisée pour mettre en œuvre la stratégie de commande prédictive. En considérant un temps d'échantillonnage constant T_s, le modèle discret est défini par :

$$\begin{aligned} \xi_{k+1} &= \mathcal{F}\left(\xi_k, F_{in_k}\right) \\ z_k &= H\xi_k \end{aligned} \tag{23}$$

où

- $\xi_k = [X_k \; S_k \; A_k \; V_k]^T$ et $z_k = A_k$ sont respectivement l'état discret la sortie échantillonnés au temps kT_s.
- $\mathcal{F}$ est la fonction de transition non linéaire

- $H = \begin{bmatrix} 0 & 0 & 1 & 0 \end{bmatrix}$ est la matrice de mesure.
- F_{in} est l'entrée de commande, paramétrée à l'aide d'une approximation constante par morceaux.

Le suivi de trajectoire est réalisé à l'aide de la stratégie de commande prédictive non linéaire (NMPC). En se basant sur le modèle (23) et des objectifs de commande définis précédemment, la fonction coût de la NMPC peut être défini comme suit :

$$\Phi\left(\check{\xi}_{\star}, F_{in_{\bullet}}\right) = \sum_{i=1}^{N_p} \left(\check{A}_{k+i} - A_{ref_{k+i}}\right)^2 + \lambda \sum_{i=1}^{N_c} \left(F_{in_{k+i}} - F_{ref_{k+i-1}}\right)^2 \tag{24}$$

où $\check{A}$ est la concentration d'acétate prédite, et F_{ref} est le profil d'alimentation de référence pré-calculé. N_p et N_c sont respectivement les horizons de prédiction et de commande, et $\lambda > 0$ est le gain de pénalité de commande.

La formulation du problème NMPC consiste à minimiser la fonction coût $\Phi\left(\check{\xi}_{\star}, F_{in_{\bullet}}\right)$ à l'instant k comme suit :

$$\min_{\check{A}_k \ldots \check{A}_{k+N_p-1}, F_{in_k} \ldots F_{in_{k+N_c-1}}} \Phi\left(\check{\xi}_{\star}, F_{in_{\bullet}}\right) \tag{25}$$

sous des contraintes liées à la dynamique prédite du système :

$$\text{s.t.} \quad \check{\xi}_{k+i+1} = \mathcal{F}\left(\check{\xi}_{k+i}, F_{in_{k+i}}\right), \quad i = \overline{0, N_p - 1} \tag{26a}$$

$$\check{A}_{k+i+1} = H\mathcal{F}\left(\check{\xi}_{k+i}, F_{in_{k+i}}\right), \quad i = \overline{0, N_p - 1} \tag{26b}$$

$$0 \leqslant F_{in_{k+i}} \leqslant F_{max}, \quad i = \overline{0, N_c - 1} \tag{26c}$$

$$0 \leqslant \check{\xi}_{k+i} \leqslant \check{\xi}_{max}, \quad i = \overline{0, N_p - 1} \tag{26d}$$

Afin d'éviter de résoudre le problème d'optimisation sous contraintes, ce dernier est transformé en un problème de programmation non linéaire (NLP) à l'aide de la technique de paramétrisation des vecteurs de commande (CVP) présentée dans (Banga et al., 2005). De plus, un changement de variable $F_{in} = e^{v}$ vise à éliminer les contraintes de positivité sur F_{in} et améliore le conditionnement du problème d'optimisation. La fonction de coût correspondante s'écrit :

$$\Phi\left(\check{\xi}_{\star}, v_{\bullet}\right) = \sum_{i=1}^{N_p} \left(\check{A}_{k+i} - A_{ref_{k+i}}\right)^2 + \lambda \sum_{i=1}^{N_p} \left(e^{v^{ref}_{k+i-1}} - e^{v_{k+i-1}}\right)^2 \tag{27}$$

et la formulation du problème NMPC se réduit à résoudre à chaque instant kT_e :

$$\min_{e^{v_k} \ldots e^{v_{k+N_p-1}}} \Phi\left(\check{\xi}_{\star}, v_{\bullet}\right) \tag{28}$$

Résultats expérimentaux

Pour valider la stratégie de commande combinée du NMPC et l'estimation des concentrations d'acétate et de glucose par l'UKF, des expériences en fed-batch sont réalisées en utilisant la souche *E. coli*. BL21(DE3).

L'horizon de prédiction N_p et le paramètre de pénalité λ ont été réglés en simulation par essais et erreurs afin d'atteindre les objectifs de la commande tout en empêchant une déviation excessive de F_{in} par rapport au profil d'alimentation de référence F_{ref}. L'évolution de la biomasse, du glucose, des concentrations d'acétate et du débit d'alimentation dans l'une des expériences est illustrée dans la Figure 4.

La première expérience est réalisée en deux phases, un batch suivi d'un fed-batch. La phase batch est réalisée pour atteindre une concentration minimale de biomasse de 1 g/L et s'assurer que les cellules sont en phase exponentielle au moment de lancement de la régulation et d'injection du milieu d'alimentation. L'estimation par l'UKF est lancée pendant cette phase après l'initialisation du vecteur d'état.

La phase fed-batch commence après la consommation totale du glucose. Le régulateur NMPC est lancé après la configuration de la référence d'acétate et des paramètres de commande. Le régulateur génère un profil d'alimentation exponentiel, et la concentration d'acétate estimée converge vers la référence désirée en moins de 30 min et reste dans l'intervalle 0,33-0,5 L/h jusqu'à la fin de la culture. En même temps, la concentration de substrat reste dans un état quasi-stationnaire.

Après la fin de la première expérience, un rafraîchissement de la culture est effectué. Un volume de la culture est extrait à l'aide d'une pompe péristaltique, laissant 500 mL de culture de volume. Ensuite, 3L d'un milieu stérilisé en autoclave sont injectés par une pompe péristaltique, et la phase batch de la seconde expérience commence.

La deuxième expérience suit le même protocole que la première, la phase batch dure 2 h, et le glucose est rapidement consommé par les cellules. La consigne de concentration en acétate est fixée à 1,5 g/L, l'alimentation suit une courbe exponentielle, et la concentration en acétate converge vers la référence imposée. À t=12 h, un changement de point de consigne est introduit et A_{ref} est fixé à $0,7$ g/L. La commande NMPC adapte le débit d'alimentation en fonction de la nouvelle référence, et la concentration d'acétate estimée suit la nouvelle référence en 20 minutes.

Les performances de la stratégie NMPC-UKF sont très satisfaisantes. L'UKF estime avec précision les concentrations d'acétate et de glucose en présence de bruit de mesure sur la concentration de la biomasse. Les erreurs quadratiques moyennes des concentrations en substrat et en acétate sont respectivement de 0,089 g/L et 0,068 g/L, ce qui est cohérent avec la sensibilité des mesures et les niveaux de bruit (0,1 g/L).

La commande NMPC régule la concentration d'acétate estimée à la consigne fixée, et la convergence est atteinte en 20 minutes. La concentration d'acétate reste dans une plage acceptable pendant la durée de la fermentation, et les conditions de culture sont bien adaptées à la croissance de la biomasse.

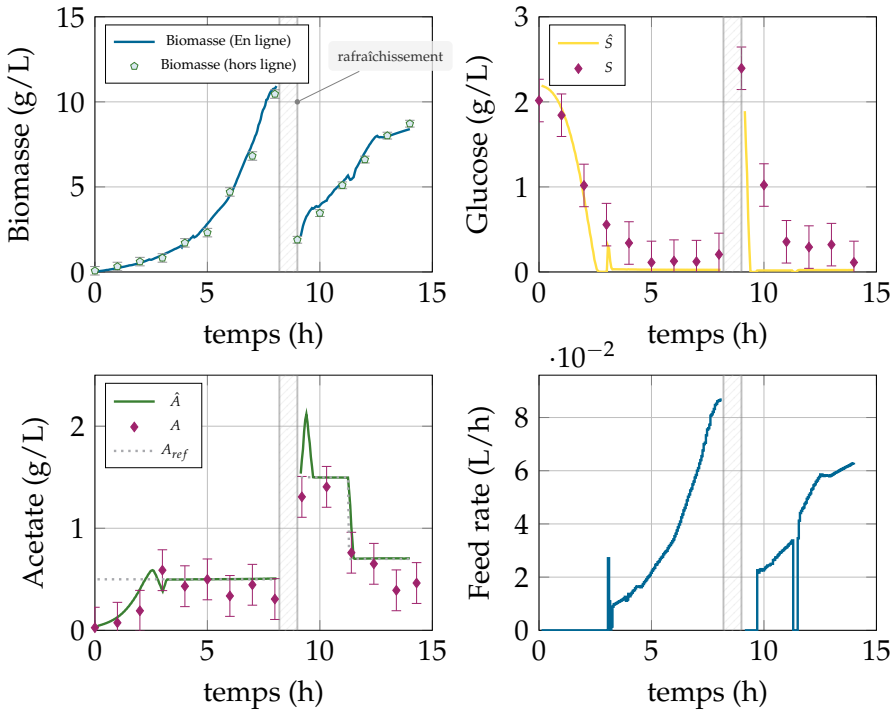


FIGURE 4: Évolution de la biomasse mesurée, des concentrations du glucose et d'acétate (mesurées hors ligne et estimées en ligne) et du débit d'alimentation.

Étude comparative

Afin d'évaluer les performances des stratégies de commande présentées dans cette thèse, une étude en simulation est effectuée et présentée dans cette section. Les trois méthodes de commande sont comparées dans les mêmes conditions de fonctionnement et avec un réglage de commande similaire.

Les commandes NMPC et GMC robuste (GMC-A) sont paramétrées pour suivre une référence d'acétate de $A_{ref} = 0,7$ g/L. D'autre part, la régulation du taux de croissance de la biomasse à l'aide de la stratégie GMC (GMC-X) est réglée pour suivre un profil de référence correspondant à cette valeur de concentration en acétate.

La simulation est réalisée sur 10 h en supposant que le vecteur d'état est disponible pour la mesure, avec un bruit de mesure ajouté de 0,05 g/L et 0,02 g/L sur les concentrations de biomasse et d'acétate.

Tout d'abord, les stratégies de commande sont comparées dans le cas du modèle idéal, où tous les paramètres du modèle sont supposés être parfaitement connus. Ensuite, 500 simulations de Monte Carlo (MC) sont effectuées en tenant compte des incertitudes du modèle. Les paramètres du modèle sont choisis aléatoirement avec une variation maximale de 30 % autour de leur valeur nominale suivant une distribution normale.

Les résultats des simulations de Monte Carlo donnent un aperçu des performances des méthodes de commande. La productivité de la biomasse sur 10 h pendant les 500 simulations est résumée dans le tableau 2.

On peut voir que le NMPC surpasse les commandes GMC en moyenne et donne une meilleure productivité. On peut également observer que la régulation d'acétate a une valeur moyenne et maximale plus élevée. Le régulateur de biomasse vise un taux de croissance spécifique, et ainsi, lorsque le procédé est capable d'atteindre des taux de croissance plus élevés, la régulation de la biomasse maintient le taux de croissance plus proche de la référence spécifiée, ce qui conduit à une productivité de biomasse plus faible. Comme prévu, toutes les commandes développées donnent de meilleurs résultats que la boucle ouverte, avec une amélioration moyenne de 20 % de la productivité de biomasse.

TABLE 2: Productivité de biomasse des méthodes de commande durant 500 simulations MC

Méthode	**Moyenne**	**Min**	**Max**	**Unitée**
Boucle ouverte	0,0161	0,0079	0,0183	g/(h · g de substrat)
GMC-X	0,0187	0,0176	0,0208	g/(h · g de substrat)
GMC-A	0,0190	0,0184	0,0206	g/(h · g de substrat)
NMPC	0,0192	0,0185	0,0206	g/(h · g de substrat)

Chapter 1

General aspects of *Escherichia coli* as a host cell

1.1 Introduction

Escherichia coli has become one of the best characterized cellular organisms on earth, and the predominant organism in research and production laboratories.

E. coli rise to fame in metabolic engineering and synthetic biology stems from its numerous traits. The simple culture conditions, the rapid growth, and the metabolic flexibility have made *E. coli* one of the most studied host organisms, leading to a significant biochemical and physiological knowledge of the cellular system and a broad set of developed genetic and genomic engineering tools. Hence, the different non-pathogenic strains of *E. coli* can be found in the productions of pharmaceuticals, food, chemicals, and fuels.

The rapid progress in metabolic engineering and synthetic biology allowed overcoming many limitations of the bacterial system. New engineered phenotypes in *E. coli* have proven to surpass traditional native producers. A demonstration of this can be found in (Gusyatiner et al., 2017), where a bio-engineered strain of *E. coli* has been utilized for the industrial production of amino acids, traditionally achieved using the *Corynebacterium glutamicum*. Another example is the production of n-butanol, where *E. coli* demonstrated a similar efficiency to that achieved with the natural producer *Clostridia* (Ohtake et al., 2017; Shen et al., 2011).

E. coli is also the preferred proof-of-concept model organism and among the top choices as a host when no natural producer exists. Successful examples of industrial production of lysine, 1,3-propanediol (PDO), and 1,4 butanediol using *E. coli* can be found in (Burgard et al., 2016; Kojima et al., 2000; Sabra et al., 2016; Sanford et al., 2016).

As such, *E. coli* is the most preferred micro-organism for research purposes and a key player in developing modern biological engineering and industrial microbiology.

This chapter introduces *E. coli*, the bacterial system chosen for this study, and we present the general aspects of bioprocesses. It is thereby divided into two main sections. We first start by describing the *E. coli* cell, its physiology, and the main strains used in laboratory applications. Additionally, we examine the aerobic and anaerobic metabolism of *E. coli*, with a focus on acetate production and overflow metabolism.

Finally, we present and discuss the general definition of a bioprocess, its main components and the different kinds of cultures used for cell cultivation.

1.2 *E. coli* as a host cell

The choice of *E. coli* for this study is motivated by its interesting traits as a host cell. *E. coli* strains can be easily cultured under a variety of growth conditions. In addition, *E. coli* can be easily enhanced genetically, allowing the extensive physiological analysis and the engineering of new phenotypes, thus significantly reducing industrial development costs (Meyer and Schmidhalter, 2012).

Various chemicals such as tryptophan, phenylalanine, threonine, lysine, and others have been produced on the industrial scale using *E. coli* strains. As for biopharmaceutical applications, recombinant proteins are the main product of *E. coli* fermentations. It was reported that 34% of recombinant proteins in US and European markets are expressed in *E. coli* (Meyer and Schmidhalter, 2012).

Nevertheless, using *E. coli* as a host has some limitations. *E. coli* is incapable of producing proteins that require complex assembly or proteins with high numbers of disulfide bonds (Meyer and Schmidhalter, 2012).

Furthermore, culture conditions that favor contamination resistance (high and low pH, high temperatures) are not suitable for *E. coli* growth (Bhalla et al., 2013; Hasunuma and Kondo, 2012; Tao et al., 2005; Wernick et al., 2016). Phage attack is also a significant concern to industrial production using bacterial strains in non-sterile conditions (Melo et al., 2018; Samson and Moineau, 2013).

Despite these drawbacks, *E. coli* maintains it's position as one the most versatile organisms used on the industrial scale. The large knowledge of its physiology, genetics and metabolism has enabled a big progress in metabolic engineering and synthetic biology using *E. coli*. This progress lead to overcoming many limitation and to the development of new phenotypes engineered in *E. coli* (Pontrelli et al., 2018).

1.3 *E. Coli* strains

To date, more than 700 serotypes of *E. coli* have been identified (Liu, 2019). Current biotechnology research is focused on the strains *E. coli* B (BL21) and *E. coli* K (MG109), and *E. coli* W (ATCC 9637).

The K strains (K-12) are used in recombinant protein production, in laboratory and industrial scale. K-12 strains are characterized by lower reducing power in the cytoplasm to form disulfide bonds better. They are frequently chosen as a host for plasmid DNA production (Phue et al., 2008).

Most common applications using *E. coli* are carried using K-12 strains. However, BL21 and its derivatives quickly became preferable for biopharmaceutical productions (Pontrelli et al., 2018). The strains B and K and their derivatives have been extensively studied. Their response under high glucose concentrations in their growth media is significantly different (Lee, 1996).

The difference in the metabolism resides within the glycolytic pathway and the tricarboxylic acid cycle (TCA) cycle. The B and K strains process glucose and pyruvate at different rates and thus generate more or less acetate production and consumption. The acetate production and consumption rates are crucial to the growth and recombinant protein production efficiency using *E. coli* (Shiloach and Rinas, 2009).

Hence, applications that require high protein expression are mostly achieved using BL21 as a host because of the lower sensitivity to high glucose concentrations (Meyer and Schmidhalter, 2012). The reduced sensitivity leads to reduced acetate production and high growth rates, and efficient glucose consumption (Phue et al., 2008).

W strains are known for the extensive substrate range, less acetate production, and higher product tolerance (Prieto et al., 1996). The W strain (ATCC 9637) is best known for ethanol production (Park et al., 2011) .

1.4 *E. coli* physiology

Escherichia coli is named after the German bacteriologist Theodor Escherich who isolated the type species in the human colon in 1885 (Feng et al., 2002). It is a non-spore-forming, Gram-negative heterotrophic bacterium from the Enterobacteria family.

The *E. coli* forms a rod-shaped cell composed of a fimbria and a cell wall. It includes a protective outer membrane, a periplasmic space with a peptidoglycan layer, and an inner cytoplasmatic cell membrane. The cell is about 2.0-6.0 μm in length and 0.25–1.0 μm in diameter, with a cell volume of 0.7 μm^3 (Nelson and Cox, 2017).

The cytoplasmatic membrane and the layers outside it constitute the cell envelope (Figure 1.1). The plasma membrane consists of a thin bilayer of lipid molecules and proteins. The cytoplasm contains about 15,000 ribosomes, thousands of different enzymes, numerous metabolites and cofactors, and various inorganic ions. The nucleoid contains a single, circular molecule of DNA, and the

cytoplasm contains one or smaller plasmids. The outer membrane provides a barrier to toxins and certain antibiotics (e.g., penicillin) (Nelson and Cox, 2017).

E. coli has an extremely simple cell structure, with only one chromosomal DNA and a plasmid. However, it can perform complicated metabolisms and survive under stress and difficult culture conditions. It can grow at temperatures from 8-48 °C and pH values ranging from 4.4 to 10. It is a facultative anaerobic bacterium since it can grow with various carbon sources in the presence or absence of oxygen.

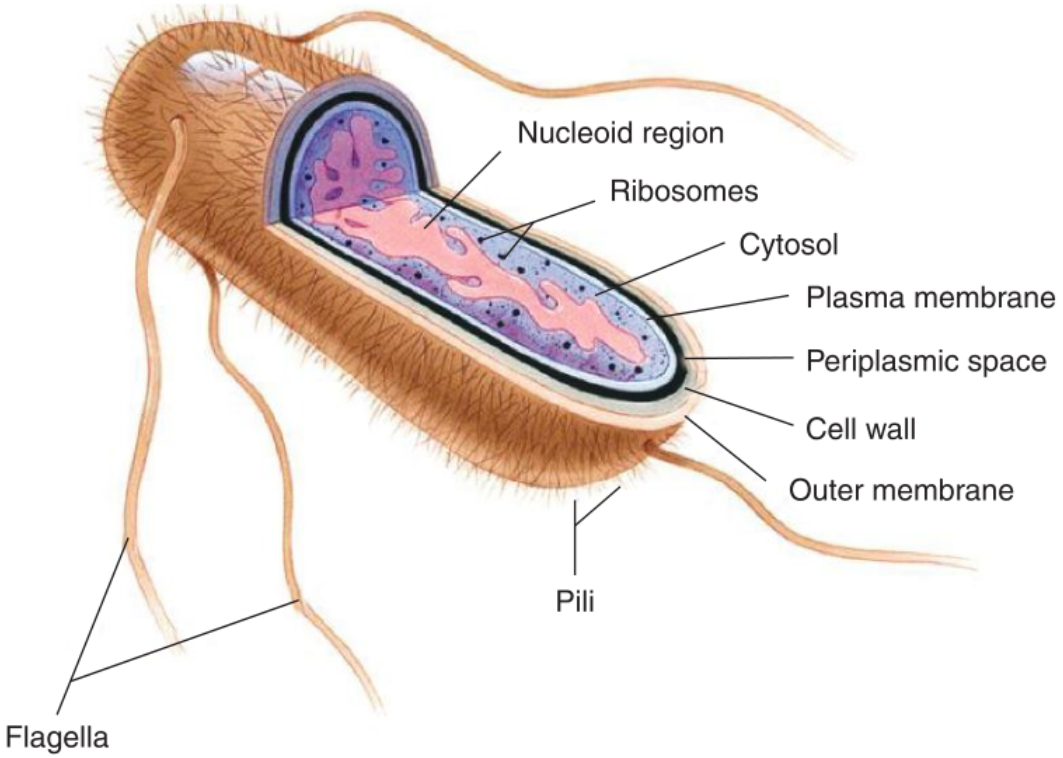


FIGURE 1.1: The *E. coli* cell structure (Moulton, 2014)

1.5 *E. coli* metabolism

The cell activities are performed through a set of biochemical reactions taking place inside the cell itself. The sum of these chemical transformations is called metabolism. These reactions are catalyzed by enzymes and organized in metabolic pathways and are regulated according to the cell's need and function. Metabolism provides energy for vital processes and the synthesis of organic material. The principal metabolism components are the carbon source, the metabolic products, the biomass constituents, the intercellular metabolites, and the enzymes.

This section presents a brief description of the most important metabolic pathways of *E. coli*, under aerobic and anaerobic conditions. Figures 1.2 and 1.3 represent the central metabolic pathways under aerobic and anaerobic conditions, respectively, considering only the carbon fluxes. The aerobic metabolic growth

is composed of two main pathways: glycolysis and the tricarboxylic acid cycle (TCA). During these transformations, part of the energy released is converted to the cofactors ATP and NADH (Nelson and Cox, 2017).

Glucose is considered as the primary carbon source in *E. coli* metabolism. Under aerobic conditions, glucose is catabolized through glycolysis. Glycolysis is the metabolic route where glucose is converted to phosphoenolpyruvate (PEP) and pyruvate through a series of enzymatic reactions, releasing some energy in the form of ATP and NADH.

Then, pyruvate is oxidized to acetyl-CoA and CO_2 by the pyruvate dehydrogenase (PDH) enzyme. This pathway involves five coenzymes and three enzymes and links the glycolysis to the TCA cycle. The TCA cycle (also known as the Krebs cycle) is a catabolic phase of aerobic respiration, where oxygen is used as the final electron acceptor to generate ATP and cofactors. The TCA cycle pathway has the role of energy production and conservation. Acetyl-CoA is metabolized in the TCA cycle, and a variety of biosynthesis products are generated (Nelson and Cox, 2017).

In aerobic growth, acetate is produced through two main pathways. The first pathway is the breakdown of acetyl-CoA by phosphotransacetylase (PTA) and acetate kinase (ACKA). This pathway is called the overflow pathway since it rapidly converts acetyl-CoA to acetate via acetyl phosphate. The overflow pathway is reversible, where the acetate can be converted back to acetyl-CoA via the acetyl-CoA synthetase (ACS) enzyme (Phue et al., 2008).

The second pathway for acetate production is the pyruvate oxidation by the enzyme pyruvate dehydrogenase (PoxB), generating acetate and CO_2. This pathway has lower energy conservation compared to the overflow pathway (Kirkpatrick et al., 2001).

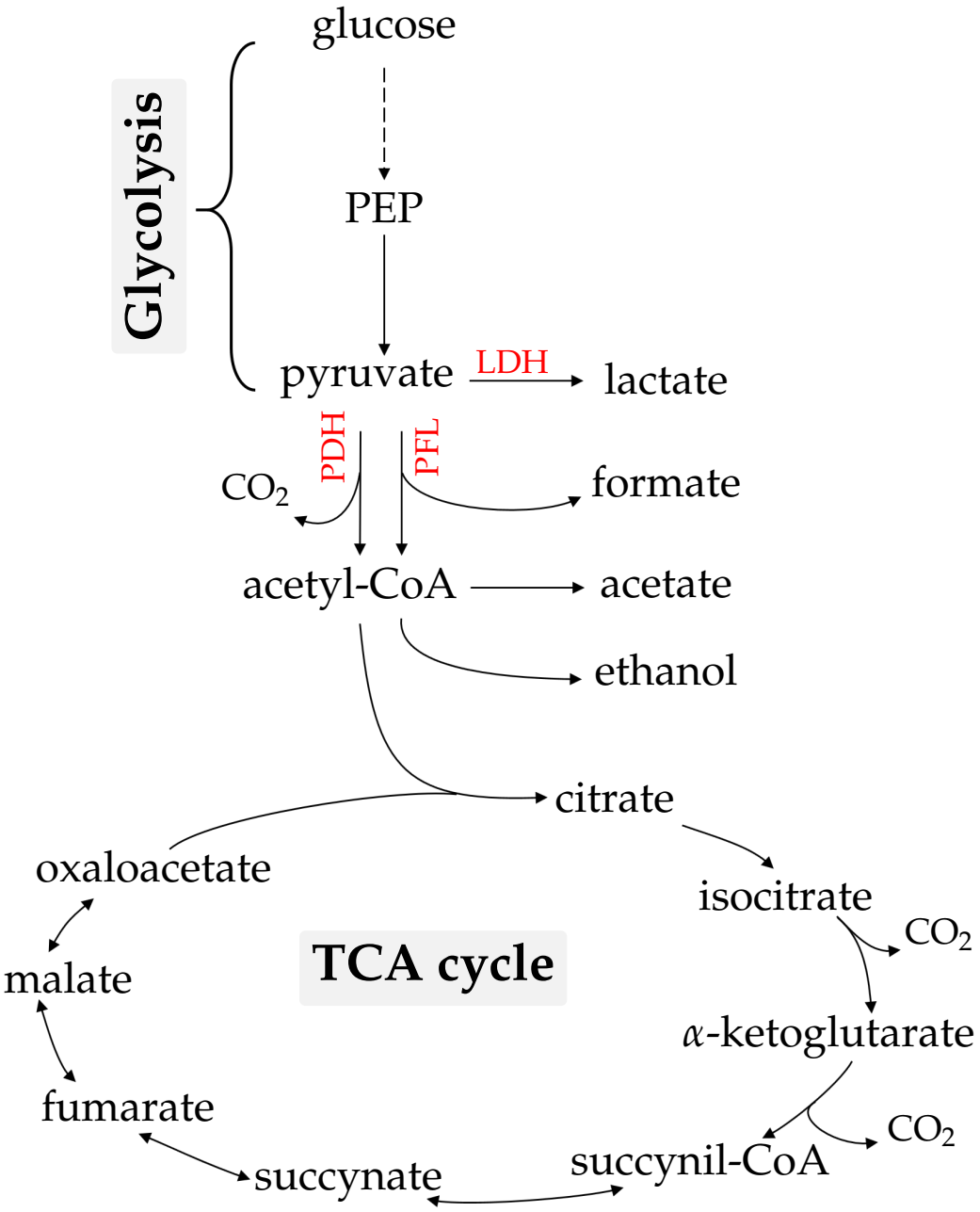


FIGURE 1.2: A simplified central aerobic pathway of *Escherichia coli* (Moulton, 2014)

The cells growth occurs also under limited oxygen and active respiratory pathways. Under these conditions, the growth is characterized by limited ATP production, leading to the downregulation of the TCA cycle and incomplete glucose oxidation (Koebmann et al., 2002).

Fermentation by-products are excreted during anaerobic growth. The list includes succinate, formate, acetate, lactate, and ethanol. These by-products are produced with ratios that allow the cell to regulate redox balance and ATP formation (Xu et al., 2014).

Acetate is produced from acetyl-COA, and the reaction generates one ATP. However, different enzymes are involved compared to aerobic growth. The pyruvate formate lyase (PFL) intervenes in the generation of acetyl-COA and formate from pyruvate (Xu et al., 2014).

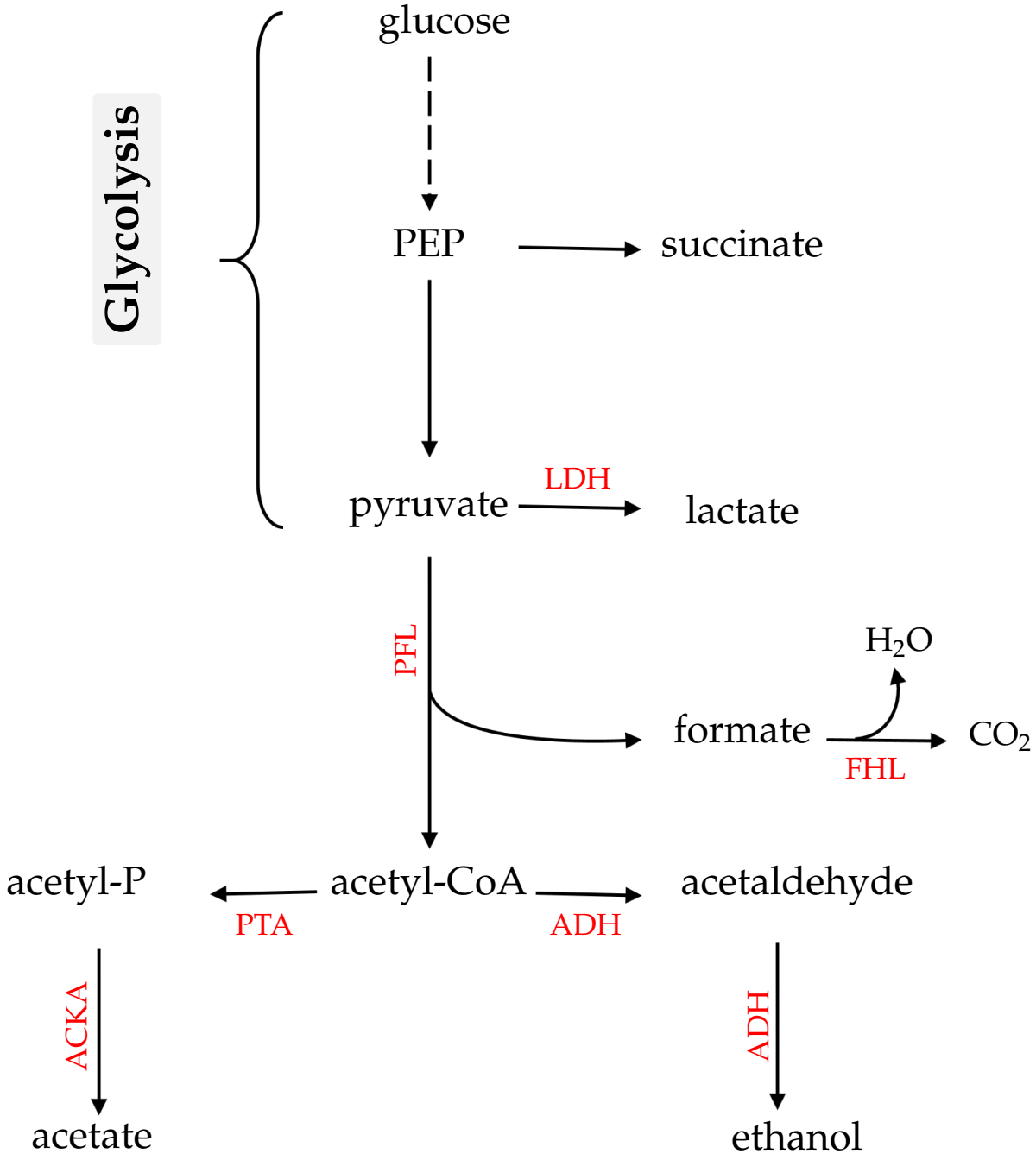


FIGURE 1.3: A simplified central pathway of *Escherichia coli* growth under limited oxygen conditions (Moulton, 2014)

1.6 Overflow metabolism

Under aerobic growth on glucose, acetate is excreted by *E. coli* cells through the fermentative pathway. However, complications can arise during the exponential growth phase. The secretion of acetate into the surrounding media can inhibit cell growth at higher concentrations (Eiteman and Altman, 2006).

The respiration efficiency in the case of excess glucose can decrease due to the production of acetate. This phenomenon is known as overflow metabolism or bacterial Crabtree effect. It was reported that up to 15% of glucose is converted to acetate under these conditions. (Crabtree, 1929; De Deken, 1966; Doelle et al., 1982; Wolfe, 2005).

Overflow metabolism can also cause the decoupling of transmembrane pH gradients, thus affecting osmotic pressure, intracellular pH, and amino acid synthesis. Adding to that, the excretion of acetate causes the acetogenesis of the *E. coli* cell.

Overflow metabolism is likely caused by an imbalance between fast glucose metabolism and the TCA cycle's limited capacity or respiration (Holms, 1986; Kadir et al., 2010; Shin et al., 2009). The repression of many TCA promoters and genes that encode enzymes leads to enzymatic limitations, and the lack of regulation in the maximum glucose uptake rate velocity (Luli and Strohl, 1990; Phue et al., 2008; Vemuri et al., 2006).

In the case of the aerobic excretion of acetate, the cell metabolism depends on the limited oxidative capacity (represented by a bottleneck) and the available glucose in the medium. Six possible scenarios presented in (Luli and Strohl, 1990; Xu et al., 2014) are summarized in Figure 1.4:

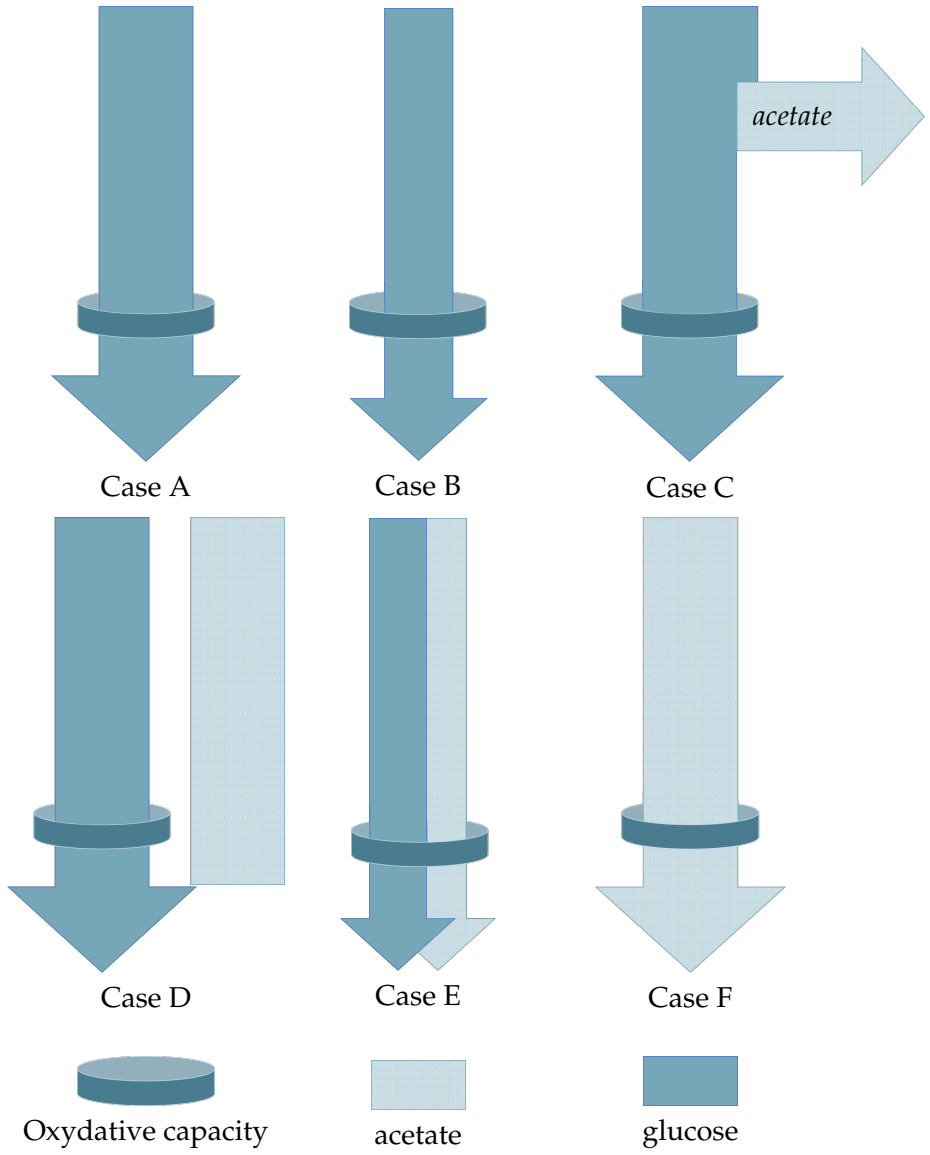


FIGURE 1.4: Representation of the bottleneck principle (Crabtree, 1929) in the case of *E. coli* cultures

- **Case A**: the oxidative capacity is completely filled by the glucose available in the culture medium.
- **Case B**: the available quantity of glucose is lower than the oxidative capacity.
- **Case C**: the available quantity of glucose is higher than the oxidative capacity. A portion of the glucose corresponding to the oxidative capacity is destined for cell growth, and the remaining amount is transformed into acetate through the fermentation metabolic pathway.
- **Case D**: the glucose completely fills the oxidative capacity despite the presence of acetate in the culture medium.

- **Case E**: the sum of the glucose and acetate quantities is lower than the oxidative capacity.
- **Case F**: the glucose is not present in the culture medium, so the acetate uses all the oxidative capacity.

Despite the presence of oxygen, the cells have a limited respiration capacity. The high glucose concentration causes an inhibition of the oxidative metabolic pathway. When the oxidative capacity is saturated, the cell can no longer oxidize the glucose, and the excess is directed to the fermentative catabolic pathway. Otherwise, two operating regimes that describe overflow metabolism can be distinguished:

- **The oxidative regime**: where the substrate concentration is lower than a critical threshold, corresponding to the maximal oxidative capacity of the cell. Acetate (if present in the medium) is also oxidized. (Cases B, E, and F).
- **The oxido-fermentative regime** where the substrate (glucose) concentration is higher than the critical threshold. The glucose excess is converted to acetate through the fermentative pathway. (Case C).

These two operating regimes are detailed in the chapter 2. They are the basis of the dynamic model of the fed-batch *E. coli* cultures considered in this work.

1.7 Bioprocess cultivation and operating modes

A bioprocess is defined as the structure developed to cultivate living micro-organisms (bacteria, yeast, fungi, etc.) or enzymes using a medium containing certain nutrients (carbon source, vitamins, minerals, etc.) under specific operating conditions (temperature, pH, oxygenation, etc.). The cultivation goal is to transform the substrates into desired products yielding from the cell's metabolism. Bioprocesses can be found in various applications, such as food processing, biomedical, pollution control, and energy production (Doran, 2013).

Cultivating micro-organisms in bioreactors depends on several factors: The product of interest, the micro-organism strain, the culture medium, and the available material means. Thereby, bioprocess design is a crucial step in the cultivation of micro-organisms. The design consists of choosing the suitable strain for the application, the bioreactor size and shape, the culture and the feeding medium compositions, and the appropriate operating conditions (temperature, pH, airflow, etc.). The cultivation mode is also an essential part of the design, as it depends mainly on the application objectives.

In this study, the continuously stirred tank reactor (CSTR) is considered. The CSTR is a tank reactor equipped with a mixing device to provide efficient agitation. This class of reactors is composed of several components (illustrated in Figure 1.5):

- A thermal **water jacket** composed of inner and outer walls, allowing temperature regulation at a constant value through hot/cold water circulation.

- An **agitator** that ensures the homogeneity of the cells and an adequate transport of nutrients and gases.
- A **sparger** used in aerobic cultivations to continuously supply the oxygen to the growing cells.
- **Probes** connections for on-line measurements of main variables (pH, temperature, dissolved gases, etc.), and the concentrations of vital components (e.g., biomass).

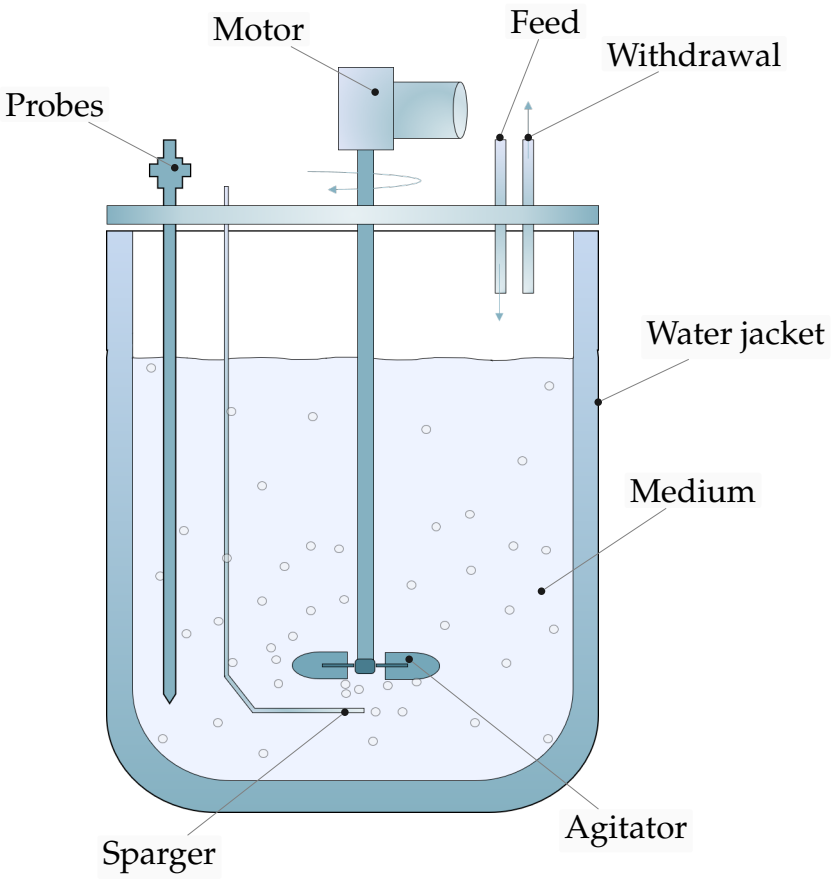


FIGURE 1.5: General scheme of a continuous stirred-tank bioreactor

Four main operating modes are distinguished in continuously stirred tank reactor (CSTR) cultures (Dochain, 2010), depending on the inlet (F_{in}) and outlet (F_{out}) feed-rate configuration, Figure 1.6 shows a diagram of these operating modes:

- **Batch mode:** In batch mode, all the nutrients and media are added at the beginning of the culture. The cells are inoculated, and the culture is operated without any addition or withdrawal of the culture medium. The batch process is a closed system, where the biomass and products are recovered at the end of the culture.

The inlet feed-rate and outlet feed rate in batch mode are equal to zero ($F_{in} = F_{out} = 0$). Therefore the culture volume is approximately constant.

Note that there is a slight variation of the volume throughout the cultivation due to base and acid (pH regulation), antifoam additions, . Sample withdrawal also slightly alters the culture volume. However, these variations are considered negligible compared to the total culture volume.

This batch mode is appropriate for fast experimentations such as strain characterization or optimization of culture medium. The advantages of cultures operated in batch mode are the short culture time, the low chance of contamination, and the ease of operation and management. This mode's disadvantages include the limited biomass and product yields, the short exponential growth periods, and the necessity of additional treatment to retrieve the products.

- **Fed-batch mode:** In fed-batch mode, the nutrients are supplied continuously during the cultivation to avoid the carbon source or other nutrients limitation. In this mode, a feed containing are injected until the end of the culture, with no medium withdrawal ($F_{in} > 0, F_{out} = 0$). Hence, the medium volume increases all along with the cell's growth.

  The culture ends when the objectives are met (quantity of biomass, quantity of products, culture time, etc.) or when the maximum volume is reached. As in batch mode, the biomass and products are harvested at the end of the culture.

  The advantage of the fed-batch mode is the higher obtainable productivity by reaching higher biomass and product yields. The cells are continuously doubling, and the exponential growth phase is extended. Other advantages are the possibility to change the substrate mid-culture and the variety of applications using different feeding strategies. The disadvantages of the fed-batch process include the build-up of inhibitory products and the higher risk of contamination.

  Fed-batch processes are widely used in biotechnological applications, particularly for the production of recombinant proteins and antibiotics.

- **Continuous mode:** The continuous mode describes the cultures operated with continuous addition and withdrawal of the culture medium. In this mode, the volume is kept constant by maintaining equal inlet and outlet rates ($F_{in} = F_{out} \neq 0$).

  Continuous processes reduce product inhibition and improve the space-time yield, leading to longer cultivations than the fed-batch mode. However, the contamination risk and long-term changes in the cultures are significantly higher.

Continuous cultivation is an ideal tool for a better understanding of the studied process since it is operated around a fixed operating point (called steady-state). The plant parameters remain constant for a considerable period of time in continuous mode.

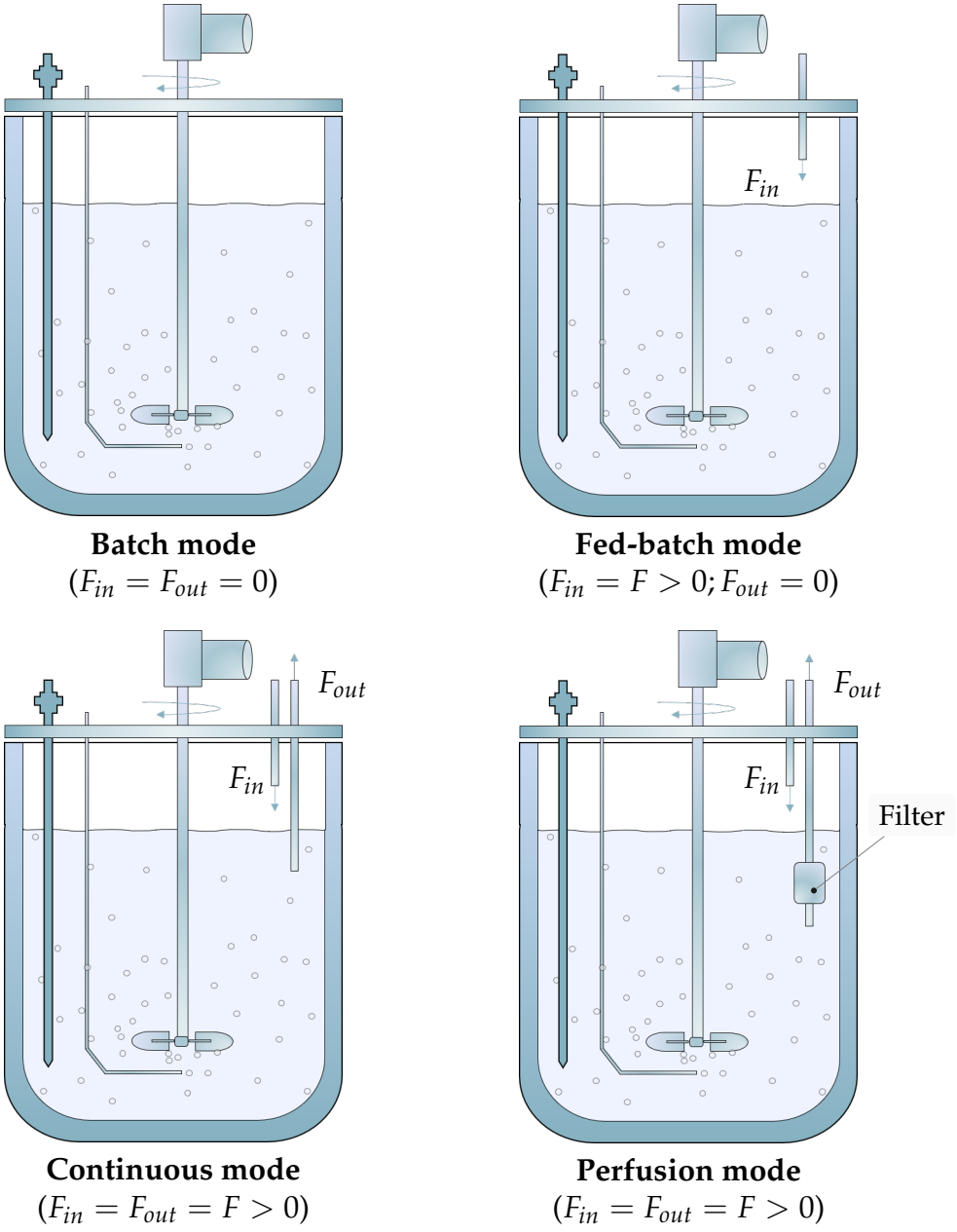
The challenge of continuous reactors is the determination of optimal operating conditions (Optimal operating points), that maximize culture efficiency and productivity.

- **Perfusion mode:** The perfusion mode is a particular variation of continuous mode, where the feed and substraction rates are non-zero. The main difference is that the outflow stream is filtered and the biomass is kept inside the reactor. Thereby, only the culture medium is renewed. This operating mode is used for animal cell cultures.

  The addition of the fresh nutrients and the withdrawal of the toxic and inhibiting products in perfusion mode provide a suitable environment for cell growth. As such, this mode leads to higher productivity and higher biomass and product yields.

  However, despite the advantages of the perfusion mode, it shares the same drawbacks of the continuous process. Fed-batch remains the most widely used mode for biotechnological production, mainly for the excellent characterization of the process due to the accumulated expertise within the industry. The experiments carried out in this work will consider the fed-batch mode for *E. coli* cultivation. The aim is to maximize the biomass productivity and cell growth since pharmaceutical and vaccine manufacturing requires maximal cell and protein production.

  A detailed description of the fed-batch bioreactor and the different materials and methods used in this work is given in chapter 4.



Batch mode
($F_{in} = F_{out} = 0$)

Fed-batch mode
($F_{in} = F > 0; F_{out} = 0$)

Continuous mode
($F_{in} = F_{out} = F > 0$)

Perfusion mode
($F_{in} = F_{out} = F > 0$)

FIGURE 1.6: The different operating modes of a continuous stirred-tank bioreactor

1.8 Conclusion

In this chapter, *Escherichia coli* is presented as a versatile micro-organism and a widely studied bacterial system in biotechnological applications. First, the various physiological features of *E. coli* are described and discussed. Then, the cell catabolism in aerobic and anaerobic conditions is summarized, considering the

central metabolic pathways. The focus is on the overflow metabolism and the production of acetate due to the byproduct's inhibitory effect, which will be of interest in the next chapters of this work.

The second part of this chapter discussed the biochemical processes and their basic composition and components. The different operating modes used in the cultivation of micro-organisms are presented, highlighting the fed-batch cultivation used in this study's experiments.

The following chapters will discuss the control and estimation of the *E. coli* fed-batch process, with the general objective of biomass productivity maximization. The closed-loop methods used throughout the work require a mathematical description of the system. Thereby, a macroscopic mathematical model for the *E. coli* fed-batch process is needed.

In the next chapter, a macroscopic dynamic model describing *E. coli* growth on glucose is presented. The model is based on overflow metabolism and the bottleneck assumption to describe the acetate dynamics during the fermentation.

Chapter 2

Dynamic modeling of *E. coli* fed-batch cultures

2.1 Introduction

Designing the bioprocess and choosing the appropriate cultivation mode is an essential step in culture optimization. The next challenge is to develop efficient online monitoring tools for the process.

For that purpose, the mathematical modeling of bioprocesses is a powerful tool that provides a comprehensive understanding of the evolution of living cells. Thereby, it can offer a good prediction of the process dynamics. The model is also the basis for adequate optimization and control applications.

The dynamic model's role is to describe the transfer phenomena of the main components, the metabolic reaction rates, and their stoichiometry based on the bioprocess operating mode and conditions.

Several steps are required to establish a dynamic model. First, a general reaction scheme describing the evolution of the main components is defined. The next

step is to select the appropriate kinetic model structure for the bioprocess. Finally, mass balance principles are applied to derive the differential equations. The resulting model quality depends on the available experimental data at hand for the parametric identification procedure.

Edwards and Wilke defined some essential properties that biological models must have (Schugerl and Bellgardt, 2014): (a) represent all the culture phases, (b) be flexible enough to use different data types (c) be easy to operate once the parameters are identified (d) have parameters with a physical significance.

Biosystems usually have complex attributes. Therefore the model must provide a compromise between a faithful and detailed representation of the process and a simple structure with the use of few parameters to simplify the identification procedure. In this chapter, we introduce a general structure of mathematical models of bioprocesses. This model is the backbone of the estimation and control strategies presented in the subsequent chapters.

The dynamic model structure is then utilized to derive the state space representation of the *E. coli* fed-batch culture.

2.2 General aspects of bioprocess modeling

Mathematical models can offer viable information about the dynamics of vital components of the culture and provide insight into the process evolution in various conditions.

These models are used to have a better understanding of the process and its optimal operating conditions. They can be used to implement monitoring tools and software sensors to predict non-measurable variables. Models are also essential to develop control strategies to reach the optimal conditions and design software sensors capable of estimating unmeasurable variables.

Modeling a process should balance a faithful representation of its dynamics and a structural simplicity to simplify the parameter identification procedure, control and estimation design phases, and provide a reasonable computational effort.

In this section, we introduce the dynamical models of cultivation in CSTR bioreactors. Thereby, the culture medium and components are supposed to be homogeneous.

2.2.1 Reaction schemes

The mechanistic model approach is based on macroscopic reactions, where only the significant reactants, catalysts, and products are considered. The reaction scheme is analogous to chemical reactions but not equivalent, as it is a qualitative representation of the macroscopic reactions.

Considering a bioprocess with a set of n_φ reactions involving R_r reactants and P_r products. The corresponding reaction scheme is described as follows (Bastin and Dochain, 1990):

$$\sum_{i \in R_r} k_{i,r} \xi_i \xrightarrow{\varphi_r} \sum_{j \in P_r} k_{j,r} \xi_j \quad r \in \left[1, \ldots, n_\varphi\right] \tag{2.1}$$

where

- $k_{i,r}$ and $k_{j,r}$ are the pseudo-stoichiometric coefficients or yield coefficients.
- φ_r is the reaction rate of reaction r.
- ξ_i is the i^{th} macroscopic component (reactant or product).
- R_r (P_r) is the set of reactants (products) in reaction r.
- n_φ is the number of reactions.

The reaction scheme is a qualitative representation of the relationship of the main components, and does not represent the chemical stoichiometry. The reactions do not necessarily satisfy elementary mass balances, thus the coefficient $k_{i,r}$ and $k_{j,r}$ are called the pseudo-stoichiometric coefficients. These coefficients (also called yield coefficients)

The components ξ_i are generally the population of micro-organisms or biomass (X), internal and external substrates (S), products (P), and enzymes. Other useful chemical components can be integrated into the reaction scheme.

The reaction scheme is not a detailed description of the process, and mass conservation is not always respected. However, it allows the combination of chemical and biochemical representations in one scheme and offers a valuable tool for deriving the process dynamical model.

2.2.2 General dynamic model

In this section, a general class of state-space models is presented. This model is derived from the reaction scheme and can describe a broad class of bioprocesses in a CSTR (Bastin and Dochain, 1990).

Applying mass balances to the reaction scheme yields the following set of nonlinear differential equations, written in a matrix form:

$$\frac{d\xi}{dt} = K\varphi - D\xi + F + Q \tag{2.2}$$

where

where $\xi_k \in \mathbb{R}^{n_\xi}$ and $z_k \in \mathbb{R}^{n_z}$ are the system state and output vectors at the time step k, respectively. $F_{in_k} \in \mathbb{R}^{n_F}$ the input. $\mathcal{F}$ is the nonlinear transition function, and H is the measurement matrix.

-
-

-
- feed, acid/base, and antifoam. V is the medium volume.
-
-

2.2.3 Kinetic models

The reaction rate vector φ describes the kinetics of biological reactions. In the case of microbial growth, the reaction rate vector is related to the biomass, and the specific rate of the reaction μ:

$$\varphi = \mu X \tag{2.3}$$

The kinetic laws are used to model these specific rates (μ). They describe in particular the phenomena of activation, limitation, and inhibition. These specific rate models are usually a function of main component concentrations ξ_i, and a set of kinetic parameters.

The models describe the activation or the inhibition by a substrate (S), the activation by the biomass (X) or the inhibition by a metabolic product (P). The most common kinetic models for the specific rates are presented below.

The Monod model

The Monod model is mostly used for growth rate modeling. This model describes the growth rate as a function of the substrate concentration:

$$\mu(S) = \mu_{\max} \frac{S}{K_S + S} \tag{2.4}$$

where

- K_S is the half saturation constant.
- $\mu_{\max}$ is the maximal specific growth rate.

The Monod law is an extension of the Michaelis-Menten model that describes the growth rate of a single substrate enzyme-catalyzed reaction (Michaelis, Menten, et al., 1913). It reflects the phenomenon of growth limitation due to lack of substrate S without considering its inhibitory effect. An illustration of the Monod law is given in Figure 2.1.

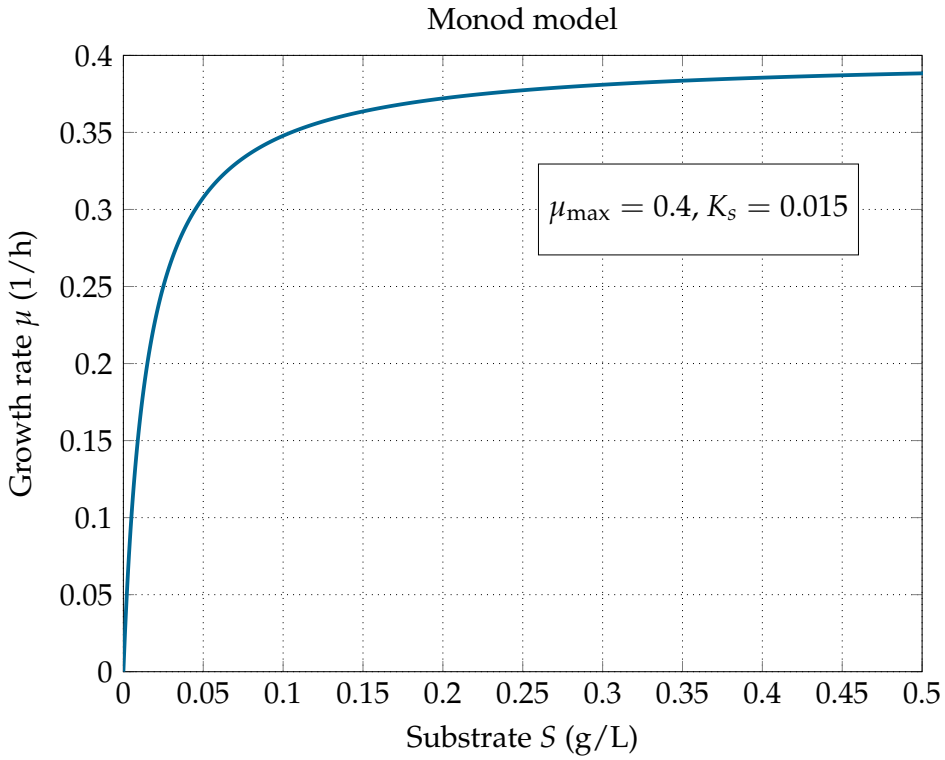


FIGURE 2.1: Evolution of the Monod model $\mu(S)$

The Haldane model (Haldane, 1930)

The Haldane model is used to describe the substrate inhibition on microbial growth:

$$\mu(S) = \mu_{\max} \frac{S}{S + K_S + S^2/K_I} \tag{2.5}$$

where

- K_S is the half saturation constant.
- $\mu_{\max}$ is the maximal specific growth rate.
- K_I is the substate inhibition constant.

The Haldane model is an empirical model first introduced for enzymatic reactions. It was extended by Andrews to model the growth of living organisms. An example of the Haldane law is illustrated in Figure 2.2.

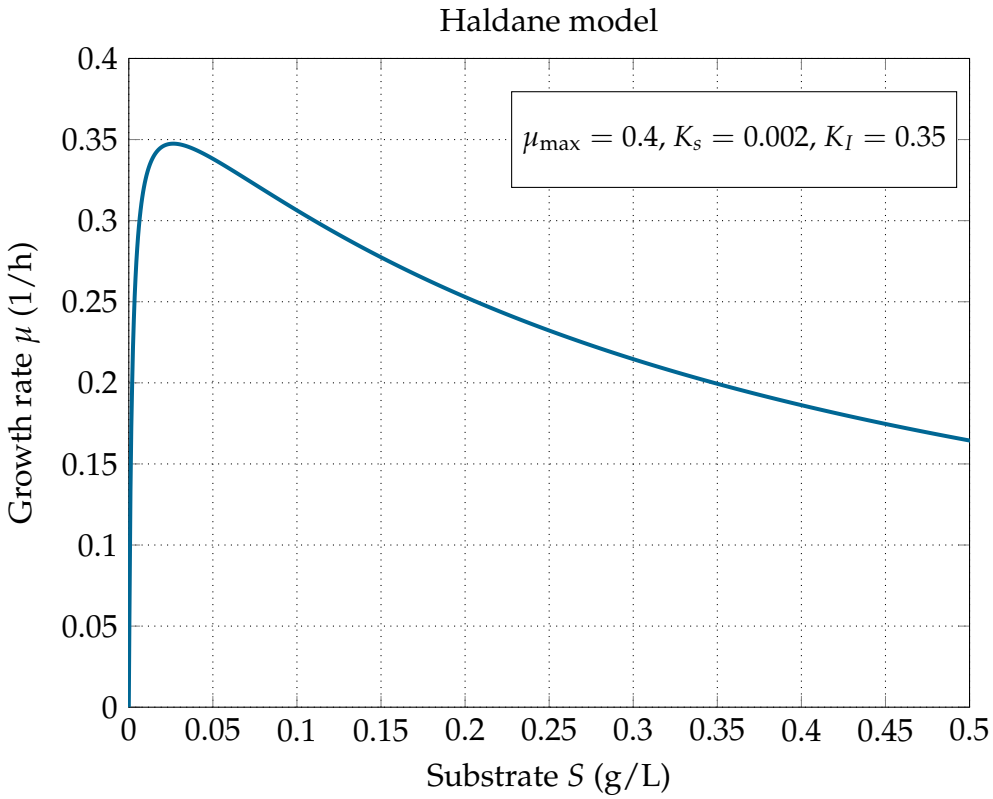


FIGURE 2.2: Evolution of the Haldane model $\mu(S)$

Contois model (Contois, 1959)

The Contois model describes the growth slowdown under high biomass X concentrations:

$$\mu(X, S) = \mu_{max} \frac{S}{k_c X + S} \tag{2.6}$$

where

- μ_{max} is the maximal specific growth rate.
- k_c is the biomass inhibition constant.

This law is similar to the Monod model, with a consideration of the possible inhibitory effect of X. Figure 2.3 shows a plot of the Contois model.

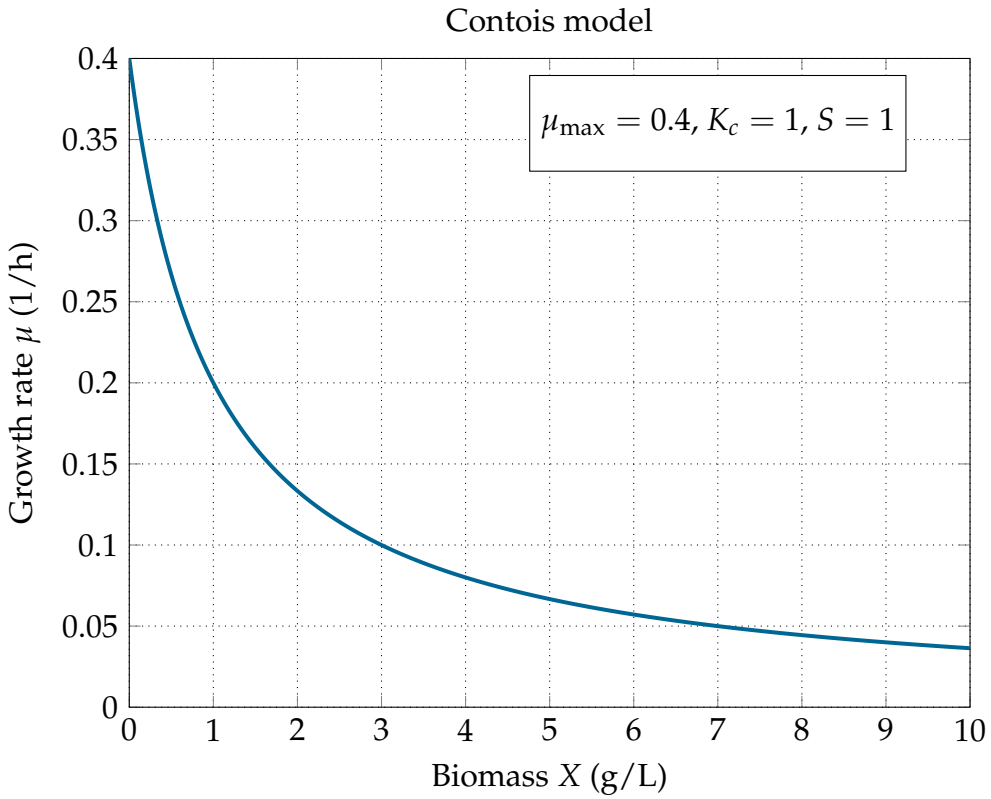


FIGURE 2.3: Evolution of the Contois model $\mu(X)$

Herbert model (Herbert, 1958)

$$\mu(S) = \mu_{\max} \frac{S}{K_S + S} - m \tag{2.7}$$

where

- K_S is the half saturation constant.
- $\mu_{\max}$ is the maximal specific growth rate.
- m is a maintenance/mortality coefficient.

An illustration of the Herbert law is given in Figure 2.4.

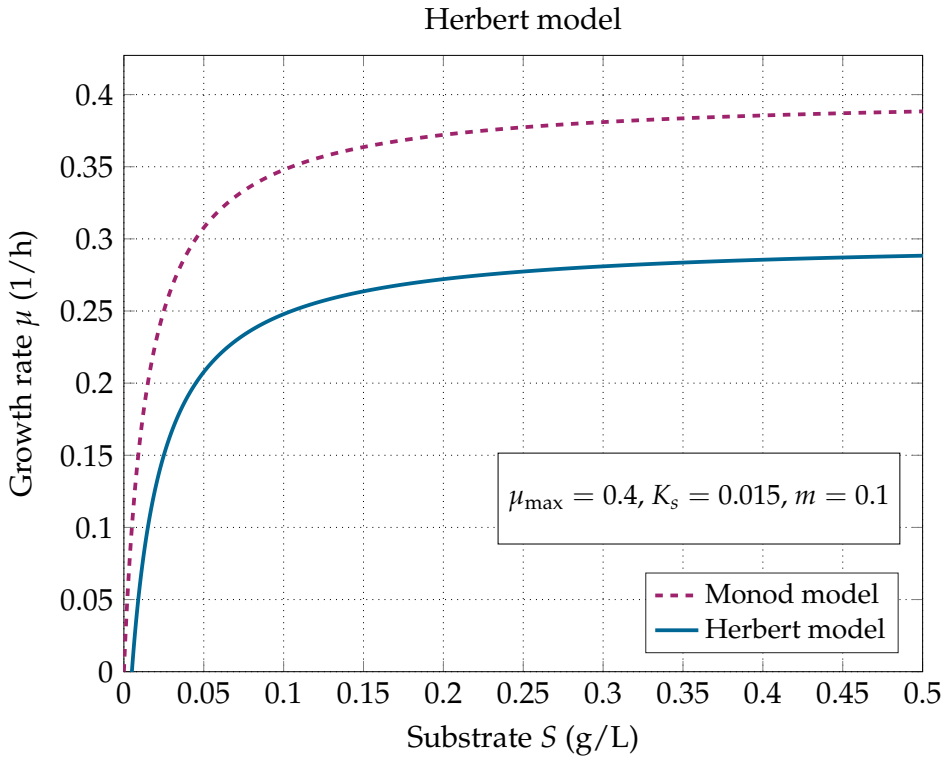


FIGURE 2.4: Evolution of the Herbert model $\mu(S)$

Product inhibition model

The growth inhibition by a by-product is a common phenomenon in alcoholic fermentations. To describe product inhibition, the following factor is added to the growth rate:

$$\mu(S,P) = \mu_{\max} \frac{S}{K_S + S} \frac{K_P}{K_P + P} \tag{2.8}$$

where

- K_S is the half saturation constant.
- $\mu_{\max}$ is the maximal specific growth rate.
- P is the product concentration.
- K_P is the product inhibition constant.

A plot illustrating the product inhibition is shown in Figure 2.5.

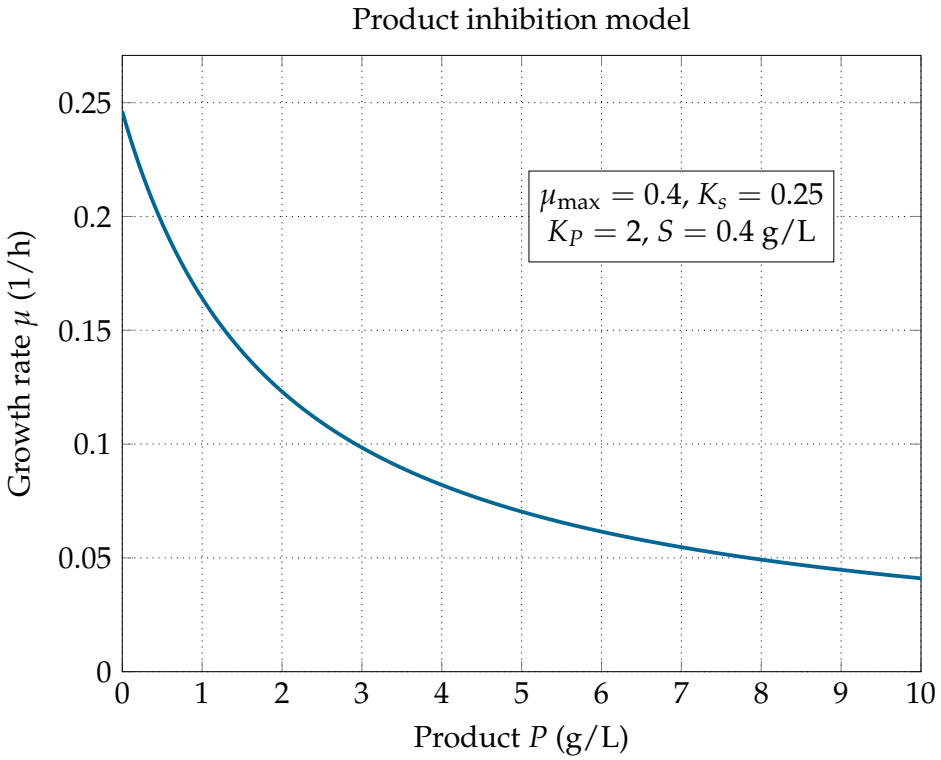


FIGURE 2.5:

In the case of fed-batch cultures of *E. coli*, the kinetics are described by Monod dynamics and product inhibition models.

2.2.4 Gas transfer models

The gas transfers in bioprocesses mainly concern oxygen, carbon dioxide, and nitrogen. In the case of aerobic cultures, such as bacteria and yeast, oxygen is a required substrate for cell growth. The accumulation of oxygen in a culture can be represented as follows:

$$\frac{dO}{dt} = OTR - OUR - D\,O \tag{2.9}$$

where

- OTR is the oxygen transfer rate from the gas phase to the liquid phase
- OUR is the oxygen uptake rate
- D is the dilution rate

The OTR classical model is based on Henry's law:

$$OTR = k_L a_O \,(O_{sat} - O) \tag{2.10}$$

where

- $k_L a_O$ is the volumetric oxygen transfer coefficient of the liquid phase
- O_{sat} is the oxygen saturation constant in the culture medium.

OUR depends on the growth of the micro-organism. It can be modeled as:

$$OUR = k_O \mu X \tag{2.11}$$

where

- k_O is the yield coefficient associated with the consumption of oxygen per biomass produced.

The carbon dioxide mass balance follows a similar approach. The transfer rate of carbon dioxide from the liquid to the gaseous phase is noted CTR and is modeled by:

$$CTR = k_L a_C \left(C - C_{sat}\right) \tag{2.12}$$

where

- $k_L a_C$ is the volumetric carbon dioxide transfer coefficient of the liquid phase.
- C_{sat} is the saturation constant of carbon dioxide in the medium of culture.

2.3 Macroscopic model of fed-batch *E. coli* cultures

The dynamic modeling concepts presented in the previous section will now be applied to a process of *Escherichia coli* fed-batch cultures. The first step is to define the specific reaction scheme related to this micro-organism. This reaction scheme will be used to establish the mechanistic model describing *E. coli* growth on glucose under aerobic conditions. Finally, the kinetic model based on the bottleneck theory is presented.

2.3.1 *E. coli* mechanistic models

Various mechanistic models in literature were used to describe overflow metabolism and acetate dynamics in fed-batch *E. coli* cultures.

Substrate consumption is the most described fermentation phenomenon in *E. coli* metabolism. The Monod model is the simplest representation of the substrate consumption rate. Despite being an empirical model, the Monod expression can adequately describe the substrate consumption dynamics. To model the inhibition by substrate or products, an inhibition term is often added to the Monod model (Cockshott and Bogle, 1999). The internal usage and partitioning of glucose as a substrate in *E. coli* cultures are studied in (Insel et al., 2007), an extensive mechanistic model describing *E. coli* metabolism is presented.

In fed-batch fermentation, the oxygenation conditions are mostly favorable for growth, and thereby the cells are assumed to be in exponential phase with a low death rate. The death rate factor is considered negligible in most mechanistic models (Horowitz et al., 2010).

Models describing acetate production and overflow metabolism can take various forms (Ko et al., 1994; Lin et al., 2001; Neubauer et al., 2003; Peebo et al., 2015; Rocha and Ferreira, 2002; Xu et al., 1999). In most approaches, overflow metabolism is divided into two distinct phases: The overflow phase, where acetate is produced, followed by the substrate limitation phase, where the acetate is consumed. (Xu et al., 1999) modeled acetate production and consumption using the bottleneck assumption. The culture switches between the two processes and the cells consume the accumulated acetate after substrate depletion. In (Insel et al., 2007), the acetate is also consumed after the substrate depletion, and the acetate production rate is modeled by a constant fraction of the substrate conversion efficiency. (Cockshott and Bogle, 1999) considered a constant parameter to model the specific acetate production rate. This parameter is estimated from experimental data. In (Ko et al., 1994), the acetate production rate is represented by an algebraic equation function of the specific growth rate, the biomass concentration, and the fraction of carbon flux to the Embden-Meyerhof-Parnas (EMP) pathway. The models based on the bottleneck theory are able to describe acetate evolution in *E. coli* fermentations adequately. However, the expressions used to model these dynamics are discrete conditional statements and discontinuous functions, presenting a hurdle for implementing advanced control laws and optimization algorithms.

Advanced proteomic analysis and system biology approaches were performed in (Peebo et al., 2015; Valgepea et al., 2010) to show that the intracellular production and consumption of acetate in *E. coli* metabolism is a continuous and simultaneous process called acetate cycling. As a result, the extracellular acetate excretion yields from an offset of the intracellular production and consumption equilibrium.

The transition from glucose to acetate consumption in *E. coli* batch cultures is instantaneous, suggesting that the acetate is already consumed in the background along with the glucose, and the acetate consumption pathways are already activated (Bernal et al., 2016; Enjalbert et al., 2015).

(Anane et al., 2017) presented a promising modeling approach of the *E. coli* fed-batch process based on the acetate cycling principle. The presented model is composed of continuous functions for the acetate and glucose kinetics. This development opens doors for the implementation of complex optimization and control algorithms in *E. coli* processes.

2.3.2 Reaction scheme

The reaction scheme that describes *E. coli* cell growth is based on the cellular metabolism presented in chapter 1. The scheme is composed of three catabolic

pathways detailed below (Retamal et al., 2018; Rocha and Ferreira, 2002):

$$S + k_{O1}O \xrightarrow{\mu_1\ X} k_{X1}X + k_{C1}C \tag{2.13a}$$

$$S + k_{O2}O \xrightarrow{\mu_2\ X} k_{X2}X + k_{A2}A + k_{C2}C \tag{2.13b}$$

$$A + k_{O3}O \xrightarrow{\mu_3\ X} k_{X3}X + k_{C3}C \tag{2.13c}$$

where

- S, O, X, C, and A represent the glucose (substrate), oxygen, biomass, carbon dioxide, and acetate concentrations, respectively.
- the parameters $k_{\xi i}$ $(\xi = [X\ S\ A\ O\ C]^\top;\ i = 1,2,3)$ are the pseudo-stoichiometric coefficients.
- μ_j $(j = 1,2,3)$ are the specific growth rates.

Each reaction describes a specific catabolic pathway. Reaction (2.13a) corresponds to the oxidation of glucose under aerobic conditions. This reaction is exergonic: it releases a form of energy that allows the cellular components biosynthesis.

Reaction (2.13b) describes the fermentation of glucose. It takes place either in anaerobic conditions, in the presence of excess glucose.

Reaction (2.13c) corresponds to acetate's oxidation, which is degraded to acetyl-CoA, and the latter allows the activation of the TCA cycle.

It is important to note that in reactions (2.13a and 2.13b), the yield coefficients are normalized with respect to substrate concentration, and in reaction (2.13c) with respect to the acetate concentration.

2.3.3 Kinetic model

As presented in chapter 1, the growth of *E. Coli* cells is modeled following the bottleneck assumption by Sonnleitner and Käppeli (Sonnleitner and Käppeli, 1986), applied to *Saccharomyces cerevisiae*.

The bottleneck theory assumes that the cells are likely to change their metabolism due to their limited oxidative capacity, leading to overflow metabolism controlled by the substrate level. Figure 2.6 illustrates the operating regimes depending on the substrate concentration.

If the substrate concentration is higher than the critical threshold corresponding to the available oxidative capacity ($S > S_{crit}$), acetate is produced by the cells through the fermentative metabolic pathway. The culture is said in the oxido-fermentative regime (reactions (2.13a) and (2.13b)).

On the other hand, acetate (if present in the culture medium) is consumed when the substrate concentration is lower than the critical level ($S < S_{crit}$), and the culture is said in the oxidative regime (reactions (2.13a) and (2.13c)).

When the substrate concentration is at the critical level and fills exactly the respirative capacity, the culture is optimal, corresponding to the edge between the two operating regimes, and acetate is neither produced nor consumed.

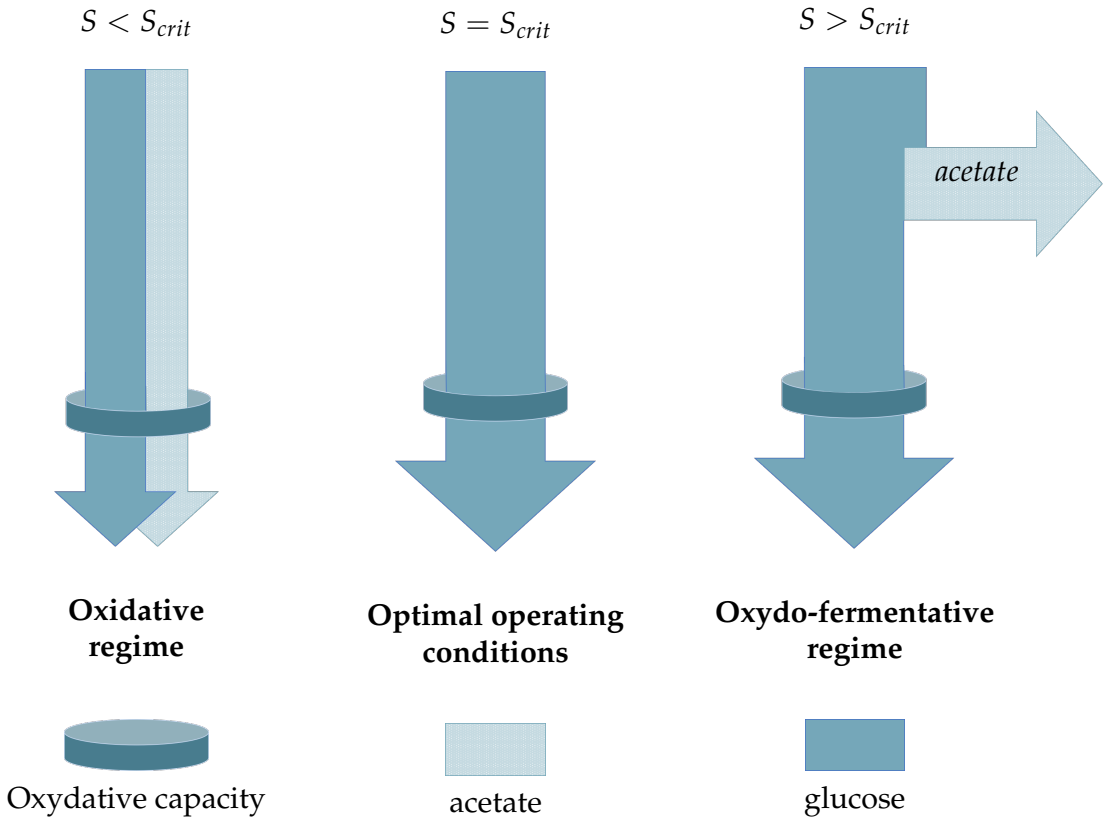


FIGURE 2.6: Operating regimes of the *E. coli* cell according to the Bottleneck assumption

The kinetic model for the specific rates is based on these operating regimes:

$$\mu_1 = min(q_s, q_{s_{crit}}) \tag{2.14a}$$
$$\mu_2 = max(0, q_s - q_{s_{crit}}) \tag{2.14b}$$
$$\mu_3 = (0, q_{AC}) \tag{2.14c}$$

where μ_1, μ_2, and μ_3 are the specific rates related to the catabolic reactions describing substrate oxidation (2.13a), acetate production (fermentation) (2.13b), and acetate oxidation (2.13c) Bastin and Dochain, 1990. Their proposed kinetic structures read (Retamal et al., 2018):

The kinetic terms related to consumption rates q_j are defined by:

$$q_s(S) = q_{s_{max}} \frac{S}{K_s + S} \tag{2.15a}$$

$$q_{s_{crit}}(A) = q_{O_{max}} \frac{K_{iA}}{K_{iA} + A} \tag{2.15b}$$

$$q_{AC}(S, A) = (q_{s_{crit}} - q_s) \frac{A}{K_A + A} \tag{2.15c}$$

where

- q_s and q_{AC} represent the substrate and acetate consumption rates respectively.
- $q_{s_{crit}}$ represents the substrate critical consumption rate.
- $q_{S\,\mathrm{max}}$ represents the maximal glucose consumption rate.
- q_{Omax} represents the maximal value of the respiratory capacity.

$$\mu_1 = q_s = q_{s_{crit}} \tag{2.16}$$

$$\mu_2 = 0 \tag{2.17}$$

$$\mu_3 = 0 \tag{2.18}$$

Equation (2.18) leads to:

$$\frac{q_{O_{max}}}{k_{OS}} = q_s = q_{s_{max}} \frac{S}{K_s + S} \tag{2.19}$$

$$S_{crit} = \frac{K_S q_{O_{max}}}{k_{OS} q_{s_{max}} - q_{O_{max}}} \tag{2.20}$$

2.3.4 Macroscopic model

Applying component-wise mass balances to the reaction scheme (2.13), the following differential equations are derived (Retamal et al., 2018):

$$\dot{X} = (k_{X1}\mu_1 + k_{X2}\mu_2 + k_{X3}\mu_3)X - D\,X \tag{2.21a}$$

$$\dot{S} = -(\mu_1 + \mu_2)X - D\,(S - S_{in}) \tag{2.21b}$$

$$\dot{A} = (k_{A2}\mu_2 - \mu_3)X - D\,A \tag{2.21c}$$

$$\dot{O} = -(k_{O1}\mu_1 + k_{O2}\mu_2 + k_{O3}\mu_3)X - D\,O + OTR \tag{2.21d}$$

$$\dot{C} = (k_{C1}\mu_1 + k_{C2}\mu_2 + k_{C3}\mu_3)X - D\,C - CTR \tag{2.21e}$$

$$\dot{V} = F_{in} \tag{2.21f}$$

where

- V is the culture medium volume.
- F_{in} is the inlet feed rate.
- D is the dilution rate ($D = \frac{F_{in}}{V}$).
- S_{in} is the glucose concentration in the feed medium.
- $\mu_{\{1,2,3\}}$ are the specific rates given by equations (2.14) and (2.15c),.

The state space model (2.21) can be written in the matrix form, following the structure of the general dynamic model (2.2):

$$\frac{d\xi}{dt} = K\varphi - D\xi + F + Q \tag{2.22}$$

$$\frac{dV}{dt} = DV = F_{in} \tag{2.23}$$

where:

$$\xi^\top = [X, S, A, O, C] \tag{2.24}$$

$$\varphi^\top = [\varphi_1\ \varphi_2\ \varphi_3] = [\mu_1\ X\ \mu_2\ X\ \mu_3\ X] \tag{2.25}$$

$$F^\top = \begin{bmatrix} 0 & DS_{in} & 0 & 0 & 0 \end{bmatrix} \tag{2.26}$$

$$Q^\top = \begin{bmatrix} 0 & 0 & 0 & OTR & -CTR \end{bmatrix} \tag{2.27}$$

$$K = \begin{bmatrix} k_{X1} & k_{X2} & k_{X3} \\ -1 & -1 & 0 \\ 0 & k_{A2} & -1 \\ -k_{O1} & -k_{O2} & -k_{O3} \\ k_{C1} & k_{C2} & k_{C3} \end{bmatrix} \tag{2.28}$$

The gas transfer rates OTR and CTR can be modeled with the classical equations:

$$OTR = k_L a_O \left(O_{sat} - O\right) \tag{2.29}$$

$$CTR = k_L a_C \left(C - C_{sat}\right) \tag{2.30}$$

where

- $k_L a_O$ and $k_L a_C$ are respectively the volumetric transfer coefficients of the dissolved oxygen and the carbon dioxide concentrations.
- O_{sat} and C_{sat} are respectively the dissolved oxygen and carbon dioxide concentrations at saturation.

2.4 Model Simulation

Two numerical simulations of the *E. coli* mechanistic model are presented hereafter. The model parameters used in these simulations and the rest of the work were estimated in a previous study (Retamal et al., 2018), performed on fed-batch cultures of *E. coli* BL21(DE3) strain. Tables 2.1 and 2.2 list the values of the kinetic and stoichiometric parameters.

The goal behind these simulations is to show the model ability to translate the different dynamics of the bioprocess and the cells behaviour under different substrate conditions. Furthermore, the simulation are designed to illustrate the metabolic switch between the operating regimes illustrated in Figure 2.6.

TABLE 2.1: Yield coefficients values of *E.coli* model (Retamal et al., 2018)

Yield coefficients	Values	Units
k_{X1}	0.184	g of X/g of S
k_{X2}	0.289	g of X/g of S
k_{X3}	0.041	g of X/g of A
k_{A2}	0.432	g of A/g of S
k_{O1}	0.737	g of O_2/g of S
k_{O2}	0.319	g of O_2/g of S
k_{O3}	1.341	g of O_2/g of A
k_{C1}	0.760	g of CO_2/g of S
k_{C2}	0.105	g of CO_2/g of S
k_{C3}	0.846	g of CO_2/g of A

TABLE 2.2: Kinetic coefficients values of *E.coli* model (Retamal et al., 2018)

Kinetic coefficients	Values	Units
q_{Omax}	1.403	h^{-1}
q_{Smax}	3.281	h^{-1}
K_S	0.050	g of S/L
K_A	0.392	g of A/L
K_{iA}	2.041	g of A/L

The profiles of the measured state variables and feed rate (X, S, A, F_{in}) with the corresponding model predictions are plotted in Figures 2.7 and 2.9. The specific growth rates are shown in Figures 2.8 and 2.10. The experiments were designed to trigger metabolic switches and observe the cells reaction to different operating conditions from substrate excess to starvation.

The first experiment is composed of a batch phase followed by a fed-batch one with exponential and constant feeding (Figure 2.7).

In the first part of the batch phase (A), we observe an exponential growth of the cells and the consumption of the main substrate (glucose). Acetate is produced during this phase, indicating that the cells are in oxido-fermentative regime (RF). This can be seen in the specific growth rates evolution (Figure 2.8). The specific rate μ_3 associated with acetate consumption is kept at zero, unlike its counterparts μ_1 and μ_2, indicating that only the first two reactions are active (equations (2.13a) and (2.13b)).

The second half of the batch phase (B) illustrates the cell's behavior in the case of lack of glucose. , and the acetate becomes the new substrate. The culture is now in the oxidative regime. The specific growth rates confirm this metabolic switch, as shown in Figure 2.8. μ_2 is kept at zero while μ_3 and μ_1 are active (equations (2.13a) and (2.13c)).

The fed-batch phase starts with an exponential feeding (C) with $\mu_{set} = 0.18$ h^{-1}, followed by a break of 1h (D). The feeding is performed in (E) with the same μ_{set}.

During (C), the cells reprise their growth, and the biomass increases exponentially. The glucose concentration remains constant and close to zero. Acetate is not produced during this phase, indicating that the culture is operating in the oxidative regime. This can be confirmed by the fact that the glucose concentration is lower than the critical threshold (S_{crit}) and the specific growth rate μ_2 is equal to zero. μ_3 is also equal to zero because the acetate is not available in the medium to be consumed by the cells. This shows that the growth rate reference μ_{set} is lower than the optimal growth rate.

The biomass concentration remains constant when the feeding is stopped (D). The cells start growing again after the feeding is reprised with the same set value (E). The culture operates in the same regime (oxidative), and the substrate concentration is lower than S_{crit}. A constant feeding is applied in (F). The cells continue growing with a lower growth rate. The specific growth rate μ_1 decreases to a lower value when the glucose supply falls to a value below 0.02 L/h. This indicates that the culture is still in suboptimal conditions in the oxidative regime.

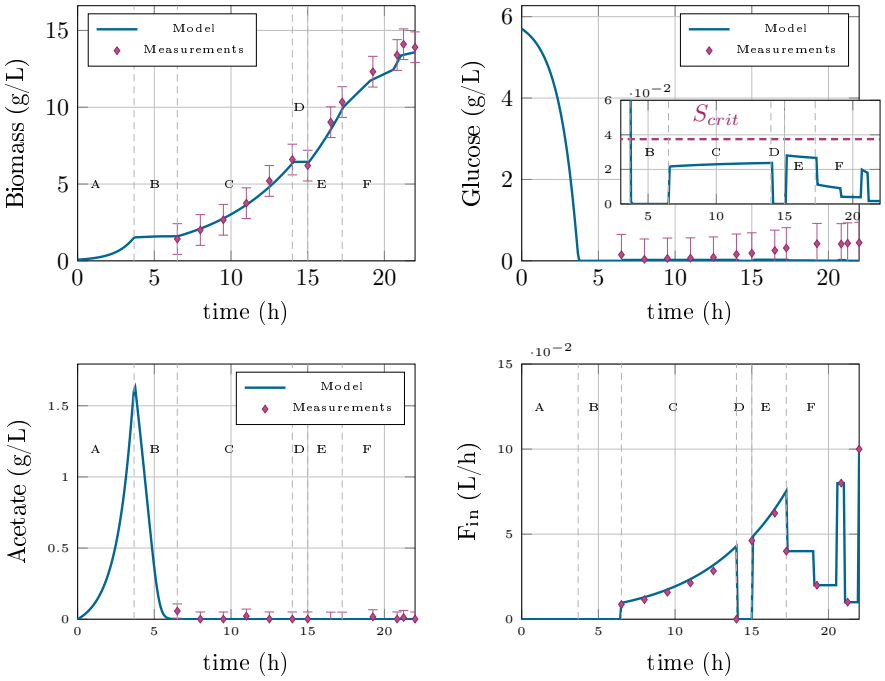


FIGURE 2.7: Experiment 1: Simulation of *E. coli* model with experimental data from (Retamal et al., 2018). Plot of the state variables ($\xi = [X\ S\ A]$) and the feed-rate (F_{in})

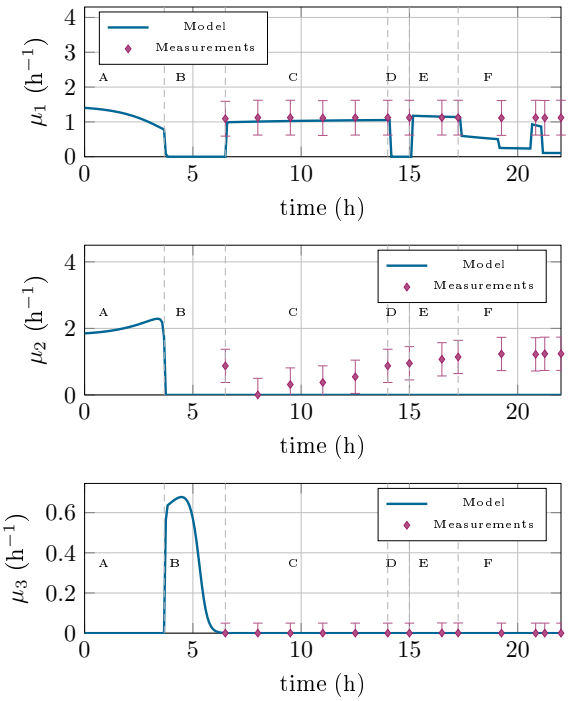


FIGURE 2.8: Experiment 1: Simulation of *E. coli* model with experimental data from (Retamal et al., 2018). Plot of the specific growth rates ($[\mu_1\ \mu_2\ \mu_3]$).

The second experiment follows a similar path to the first one. A batch phase followed by exponential feeding with different μ_{set} values and constant feeding.

During the batch phase, the cells consume the available substrate to multiply, producing acetate in the process due to the excess of glucose (A). The culture is in oxido-fermentative regime as can be seen in Figure 2.10 ($\mu_1, \mu_2 \neq 0 \quad \mu_3 = 0$). In the second half (B), biomass stabilization is observed due to the absence of glucose, and acetate is consumed. The culture switches to the oxidative regime ($\mu_1, \mu_3 \neq 0 \quad \mu_2 = 0$).

The fed-batch phase starts with exponential feeding with a higher $\mu_{set} = 0.22$ h^{-1}. The cells reprise their growth, and biomass concentration increases exponentially. The glucose concentration remains constant but higher than the critical value S_{crit} causing acetate accumulation. This triggers a metabolic switch to the oxido-fermentative regime ($\mu_1, \mu_2 \neq 0 \quad \mu_3 = 0$).

The feeding stops for 30 min (D), causing the drop of the growth rate, the glucose depletion, and consequently the switch to the oxidative regime. Reprising the feeding (E) with $\mu set = 0.18$ h^{-1} causes the cells to reprise their metabolism, but the low feeding rate puts the culture in oxidative mode ($\mu_1 \neq 0 \quad \mu_2, \mu_3 = 0$) and leads to the consumption of the acetate. The constant feeding (F) causes a drop in the growth rate, but it does not change the culture operating mode.

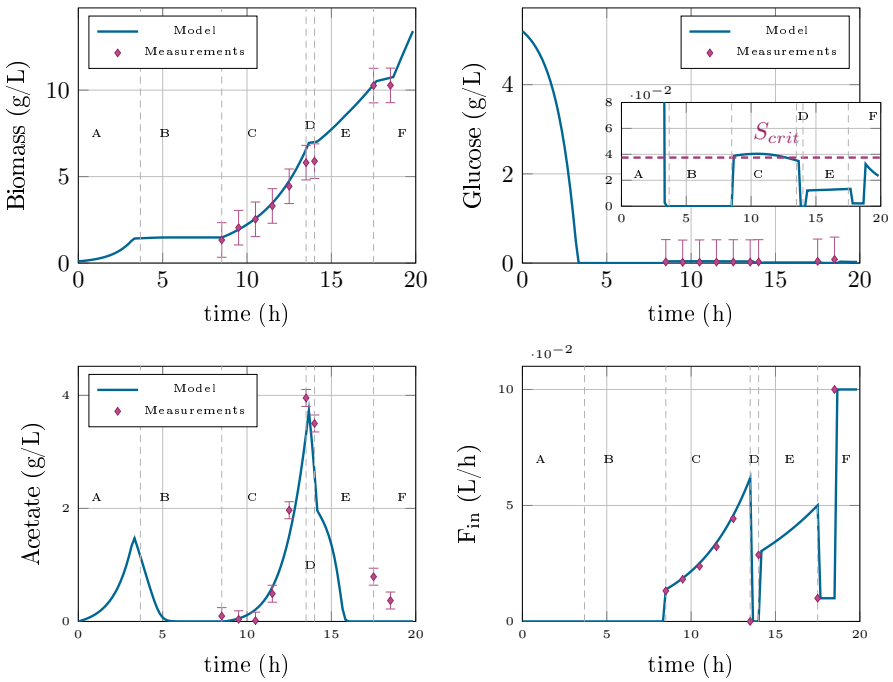


FIGURE 2.9: Experiment 2: Simulation of *E. coli* model with experimental data from (Retamal et al., 2018). Plot of the state variables ($\xi = [X\ S\ A]$) and the feed-rate (F_{in})

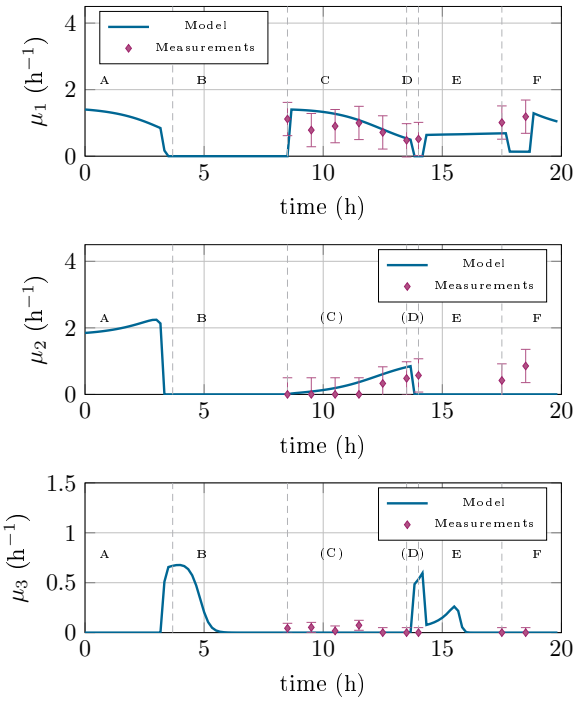


FIGURE 2.10: Experiment 2: Simulation of *E. coli* model with experimental data from (Retamal et al., 2018). Plot of the specific growth rates ($[\mu_1\ \mu_2\ \mu_3]$).

The simulations show an overall good fit between the model prediction and the experimental data. The model adequately describes the metabolic behaviour of the cells under various feeding conditions.

Nevertheless, the identified model presents some parametric uncertainty shown in Figure 2.8 indicating a possible mismatch in the parameters which is very common in bioprocess models. Therefore, the developed control and estimation strategies must consider the model mismatch in the design procedure.

2.5 Conclusion

In this chapter, we present a general representation of mathematical models used to describe the dynamics and the kinetics of bioprocesses. A dynamic model is a macroscopic approach based on chemical and biological reactions that illustrate the major components interactions.

Mass balance principles are then applied to the reaction scheme to obtain the state-space model. The model also includes a kinetic representation of the reaction rates depending on the empirical relations between the process components. A presentation of the main kinetic models used in bioprocesses models is given.

Lastly, the general dynamic model is applied to the case of fed-batch *E. coli* cultures, and a state space representation of the studied bioprocess is given. This model is then illustrated with a series of numerical simulations to highlight the different operating mode in *E. coli* metabolism.

The mechanistic model presented in this chapter is used to implement software sensors in order to estimate the non-measured variables. The next chapter will present the estimation scheme developed to predict the acetate concentration in fed-bacth *E. coli* cultures.

Chapter 3

State estimation and culture monitoring

3.1 Introduction

The main advances in biotechnological processes mainly include , bioprocess optimization, control, fault detection, and diagnosis. An essential common requirement in all these engineering tools is the necessity of reliable real-time information of viable process variables. Thereby, efficient bioprocess monitoring is an essential component for successful bioprocess operation.

Despite the recent advances in optical, ion-selective, and enzymatic sensors, reliable on-line measurement of many essential variables in bioprocesses is not possible. The available on-line measurement tools are either expensive or mainly concern pH, temperature, and dissolved gases. Nevertheless, many informative

variables such as the metabolic products and intracellular metabolite concentrations are not easily measured on-line. The absence of on-line measurements of the internal process variables represents a limitation for the effective operation of bioprocesses.

This information could provide a more comprehensive description of the process and improve the cultivation. Furthermore, incorporating this estimated information to develop novel control strategies can improve the process performance and productivity (Bastin and Dochain, 1990). In addition to control applications, variable and parameter estimation can be integrated into other engineering applications such as optimization, fault detection, and diagnosis.

In this chapter, a brief presentation of the estimation methods developed for bioprocess monitoring is described, focusing on the Kalman filtering algorithm since it is the estimation technique used in the various applications presented throughout the thesis.

Later in this chapter, the problem of acetate and glucose estimation in fed-batch *E. coli* cultures is discussed. The Kalman filter algorithm is applied to the mechanistic model presented in the previous chapter. A series of numerical simulations are performed to illustrate the developed estimation schemes.

3.2 State estimation methods used for bioprocess monitoring

On-line estimation of non-measured process variables and parameters provides a tool for better monitoring of bioreactors. The estimation methods are used to deliver a reliable real-time estimation of the internal process variables (called state variables) or parameters based on the knowledge of the process and the available measurements.

The incorporation of estimated signals of unknown process variables in the control strategies has significant advantages as it complements conventional sensor data and delivers new feedback signals for control and regulation purposes. The estimation can also provide viable information to compensate model mismatch and parametric uncertainties, thus improving the control performance and the process productivity (Bastin and Dochain, 1990).

In this section, the main state and parameter estimation methods used in bioprocess monitoring are presented and classified into different categories, including methods based on balance equations, observers, neural networks, and fuzzy logic-based estimation. The list of methods is not exhaustive but it is a presentation of the main estimation methods encountered in bioprocess monitoring

3.2.1 Balance Equation Based Methods

The balance equation methods are based on the theoretical and empirical relationships between the measurable and unmeasurable variables.

This empirical approach neglects the measurement noise and model uncertainties and only relies on the mathematical relationship between variables. Therefore, the estimation is performed through simple calculations, and the required computational cost is low.

Several bioprocess applications considered balance equation-based estimation. In (Zabriskie and Humphrey, 1978), the estimation of the biomass concentration and the biomass growth rate in aerobic batch fermentation processes were performed based on the dissolved oxygen measurements and mass balance equations. The method is validated on a bakers yeast culture.

(Grosz et al., 1984) introduced the relationship between the respiratory quotient and the product yield in fermentation processes. The respiratory quotient and the heat evolutions measurements are used to reconstruct the product yield for on-line identification and control purposes. The theoretical results are validated through simulations and experimental data of yeast and *E. coli* fermentations.

Another on-line estimation of *Streptomyces avermitilis* fermentation based on the oxygen measurements and mass balances is presented by (Gbewonyo et al., 1989). The biomass estimation is also investigated by (Beluhan et al., 1995) where maintenance equations are used to estimate the biomass concentration in baker's yeast fermentation. The on-line estimation is used in a feedback control scheme.

Mass balance estimation methods are based on input-output relations between the process state variables and parameters. They involve simple calculations and can provide important insight into the evolution of key variables.

However, the disadvantage of these techniques is that they do not consider system uncertainties or measurement noises, leading to significant estimation errors.

3.2.2 Observer-based Methods

State observers are model-based estimation schemes used to predict the evolution of the process parameters and variables. The mathematical model of the bioprocess and the available on-line measurements are used to reconstruct the non-measurable signals. The observer is provided with the system input u and output y to estimate the real process vector ξ and calculate the estimated state vector $\hat{\xi}$ (as presented in Figure 3.1).

The state observers are based on a mathematical model to provide the state and parameter estimation, and they can effectively deal with deterministic and stochastic, linear, and non-linear processes. Observers can also incorporate disturbances and uncertainties, making them a powerful tool for bioprocess monitoring. However, developing an accurate process model is a major task, and the quality of the estimation is linked to model structure and the identified parameters. Several types of observers have been developed and implemented for biotechnological applications; a detailed description of observers designed for bioprocesses can be found in (Bogaerts and Vande Wouwer, 2003).

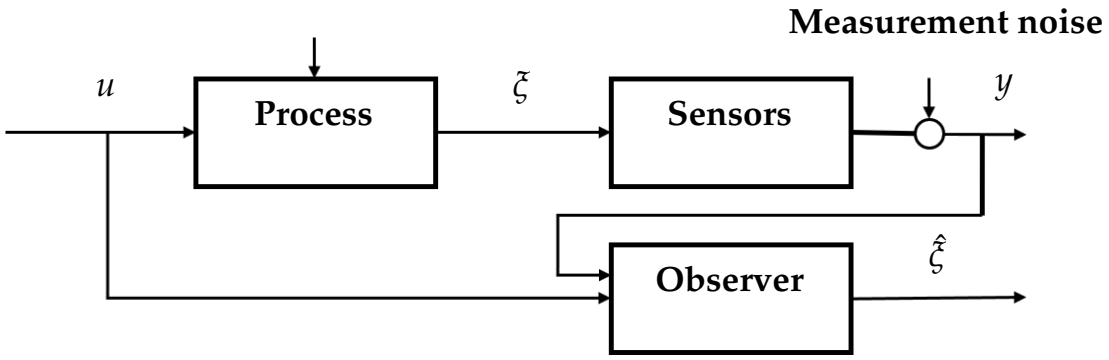


FIGURE 3.1: Observer block diagram

The observer type and design depend on the model structure (linear/nonlinear), the model uncertainty, the measurement nature (discrete/continuous), the sensor noise, and the estimation objectives.

Various observer-based estimation methods have been bioprocess monitoring. Each method has some advantages and drawbacks depending on the application. The main differences are:

- The ability to consider measurement noise and errors.
- The ability to consider model uncertainties.
- The aptitude to handle nonlinear dynamics or the need for local linearization.
- The convergence rate of the observer
- The observer convergence depending on the culture conditions (e.g. metabolic switches).

A brief description of the main observer types is given hereafter:

Exponential Observers

Exponential observers reconstruct the state vector based on a process model and on-line sensor measurements. In linear systems, the system observability condition must be guaranteed to design an exponential observer. As for nonlinear models, the local observability condition is the minimal requirement to implement this type of observers (Kwakernaak and Sivan, 1974).

The exponential observer uses the measurements in a closed-loop feedback structure. The estimation error is driven to zero, and the estimated state converges exponentially to the real process state. The observer parameters are tuned to adjust the convergence rate and improve the estimation quality. Examples of exponential observers are the Luenberger observer, the Kalman filter, the moving horizon observer, and the high-gain observer (Bogaerts and Vande Wouwer, 2003).

The specific case of Kalman filtering is detailed in section 3.3 since this filter is an essential part of the control strategies presented in this thesis.

Asymptotic Observers

Asymptotic observers have an open-loop structure, where only a part of the process model is used, and the missing part is compensated by the continuous and noise-free measurements. Unlike the exponential observer, the asymptotic observer does not include a correction term between the estimated state and the measurement. The advantage of the asymptotic observers is the possibility to estimate the state vector without the need to model the reaction kinetics. However, the estimation error convergence rate depends heavily on the operating conditions (Bastin and Dochain, 1990).

A recent example of asymptotic observers developed for bioprocesses can be found in (Dochain and Rapaport, 2018). The asymptotic observer is designed for systems that are unobservable on a subset domain boundary. The observer is applied to a batch microbial fermentation system with a single biogas measurement to estimate the biomass concentration. The results showed a guaranteed convergence and better overall performance compared to the Luenberger observer.

Hybrid observers

Hybrid observers provide a combination of the previous method features: the adjustable conversion rate of the exponential observers and the robustness towards the reaction kinetics of the asymptotic observers.

This technique consists in defining a confidence level in the mathematical model, and the observers switch between two extreme cases depending on this confidence level. When the model is assumed to be perfectly known, the exponential structure is used. Conversely, when the uncertainty level is high, the asymptotic structure is considered (Bogaerts, 1999).

Examples of hybrid observers applied to bioprocesses can be found in (Bogaerts and Coutinho, 2014; Hulhoven et al., 2006a). A hybridization parameter that reflects the kinetic model confidence is generally considered.

3.2.3 Neural Network Based Methods

Artificial neural networks (ANN) are computing systems composed of several layers of highly interconnected nodes. The network maps a set of input patterns to corresponding output patterns by learning from a series of a defined set of input-output examples from past data. Then, the neural network applies this learned information to new inputs to generate an appropriate predicted output.

In the case of state and parameter estimation, the inputs to the network estimator consist of the measured process variables and the process inputs. The process outputs act as a teacher signal that trains the neural network. The difference between the predicted output by the network and the process states is called the prediction error. The trained neural network performs iterations to minimize the prediction error so that the predicted states converges to the real process ones.

Artificial neural networks provide the possibility to estimate states and parameters of nonlinear processes and do not require developing and validating rigorous process models (black box models). The application of neural networks for state estimation and parameter identification for bioprocesses increased in recent years. A neural network-based estimation and prediction of bioprocess variables presented in (Karim and Rivera, 1992). The network was trained on available environmental and physiological data sets using the conjugate gradient method with unconstrained optimization.

In (Murugan and Natarajan, 2019), a Multiphase Artificial Neural Network (MANN) based estimator was developed to predict the biomass concentration in fed-batch Trichoderma cultures in the presence of insoluble substrates. The estimator comprises three nonlinear Auto Regressive with eXogenous input (NARX) models to capture the three phases of the microorganism. The proposed MANN-based estimator demonstrated good performance with acceptable deviation.

A three-layer feed-forward back propagation ANN was employed to estimate the biomass concentration in a microalgae cultivation with various nutrient sources (Ansari et al., 2020). The proposed ANN structure achieved high predictive performance for the waste-water treatment process.

In (Ahmad et al., 2016), an ANN model was proposed to estimate the biosurfactants yield, the surface tension reduction, and the emulsification index in *Klebsiella* sp. FKOD36 cultures. The ANN model used temperature, pH, incubation period, and gas measurements as inputs. The estimation scheme showed an efficient prediction of the variables compared to experimental data.

Neural networks are very adaptive and can provide efficient results in the presence of measurement noise and incomplete data. However, the training range restricts the generalization to different operating conditions, and new training data is usually required.

3.2.4 Fuzzy logic Based Methods

Fuzzy logic is a mathematical approach to handle uncertain, semi-qualitative, and linguistic information. The fuzzy set theory describes and manipulates imprecise and vague physical phenomena through a set of graded membership functions. A set of fuzzy if-then rules and an inference mechanism are used to determine the input-output mapping of the system (Zadeh, 1994).

The advantage of the fuzzy logic approach is the greater flexibility to capture incomplete and imperfect aspects of the process. Several bioprocess parameters and state estimation algorithms based on fuzzy logic were developed. In (Dohnal, 1985), a fuzzy-logic-based model was used to estimate the specific yield in a fermentation process, based on the dilution rate and the growth rate signals. A review of fuzzy logic systems application for estimation and control purposes in biological applications can be found in (Birle et al., 2013).

A stable linear matrix inequality-based fuzzy observer for fermentation processes is proposed in (Márquez-Vera et al., 2018). The fuzzy rules used are the

linear, mechanistic submodels. The observer is set to estimate the biomass concentration and is integrated into a state-feedback regulation of the product concentration. In (Boiocchi et al., 2015), a fuzzy logic-based estimation scheme has been implemented for the diagnosis and control of a Complete Autotrophic Nitrogen Removal (CANR) process. The estimation is based on the nitrogen concentration in the influent and in the effluent measurements.

The fuzzy logic-based estimation is a practical alternative for complex processes, where the development of an accurate model is not straightforward. However, they require a good understanding of the bioprocess to establish a comprehensive rule base.

3.3 Kalman Filtering

The classical Kalman filter has been used in various applications for more than 50 years due to its low computational requirement and status as the best estimator for linear systems with Gaussian error statistics (Anderson et al., 2008).

The Kalman filter provides optimal estimates of the state variables for linear systems, based on the noisy measurement and a mathematical model of the process. The estimate provided by the Kalman filter is the maximum likelihood estimate conditioned on all observations up to that instant of time.

There are different variations of the Kalman filter, depending on the nature of system dynamics (continuous or discrete, linear or not) and measurements (discrete or continuous) assuming gaussian distribution of the process and measurement noises. In the case of biochemical systems, the measurements are usually available at large sampling periods, leading to the consideration of discrete measurements. However, most bioprocess models are continuous, so considering Kalman filter for continuous dynamics is generally preferred. Nevertheless, in this work, since the system dynamics are slow compared to the sensor's characteristic times, it will be discretized, leading to an estimation problem for a discrete-time system with discrete-time measurements.

Furthermore, most bioprocesses are inherently nonlinear, therefore modeling their dynamics with a nonlinear mapping that reflects the essential structure of the process is more beneficial. Consequently, the estimation algorithms developed for bioprocesses should be based upon the nonlinear model structure or approximate the dynamics using a linearization of the model with some adaptation.

The Kalman filter algorithm has been appended for nonlinear systems in various forms. The Extended Kalman filter (EKF) is an estimation technique based on the linear Kalman filter equations used for the estimation of states and parameters in nonlinear processes. The EKF algorithm is based on linearizing the nonlinear equations around the current estimate and . This approximation may introduce estimation errors, especially in the case of strong nonlinearities.

An alternative to the EKF is the Unscented Kalman Filter (UKF) (Julier and Uhlmann, 1997). The UKF is a derivative-free estimation method based on a deterministic sampling approach called the Unscented transformation. This technique allows for better estimation in inherently nonlinear systems.

This section aims to present the Kalman filter equations for linear and nonlinear systems. The algorithms presented in the following are used throughout the thesis in both simulations and experiments for state and parameter estimation purposes.

In the case of nonlinear systems, the Extended Kalman filter (EKF) and the Unscented Kalman Filter (UKF) are tested and compared in simulation using the fed-batch *E. coli* process model presented in the previous chapter.

3.3.1 Linear Kalman filter (KF)

The Kalman filtering approach consists of minimizing the estimation error variance using an algorithm with two recursive steps. First, the process model is used to propagate the initial state estimates until a new measurement is available (prediction step). In the second step, the propagated model estimates are combined with the measurements to update or correct the estimates (Lewis et al., 2007).

Several studies considered the Kalman filter for bioprocess monitoring purposes. (Cha and Hitzmann, 2004) have employed the Kalman filter for noise filtering and prediction of the biomass, glucose, and ethanol concentrations in *S. cerevisiae* batch cultivation. In (Holmberg and Olsson, 1985), the Kalman filter was employed in an open aerator system for the on-line estimation of the oxygen transfer rate and the respiration rate, using air flow rate and dissolved oxygen measurements. Similarly, (Howell and Sodipo, 1985) presented a Kalman filter-based algorithm to estimate respiration and aeration rates in a sludge aeration basin using the measurement of dissolved oxygen concentration only. Other examples of state estimation in bacterial and microbial cell cultures using the Kalman filter can be found in (Chattaway and Stephanopoulos, 1989).

The classical Kalman filter addresses the general problem of state estimation in linear stochastic processes with Gaussian error statistics. The algorithm describing the prediction and update steps of the Kalman filter is presented in appendix A.

3.3.2 Extended Kalman Filter (EKF)

The classical Kalman filtering is applied to a variety of systems described by linear stochastic equations. However, most practical applications and processes (including bioprocesses) are described by nonlinear models.

The Extended Kalman filter (EKF) is the standard method for most nonlinear state and parameter estimation problems. In the EKF scheme, the nonlinear model equations are linearized around the current estimate, and the classical Kalman filter scheme is applied to the linearized model.

Various applications of the EKF have been derived for state and parameter estimation for biotechnological processes. (Lee et al., 1991) applied the EKF to a batch culture of *E. coli* to filter the noise on the dissolved oxygen measurements and improve the DO control strategy. (Cha and Hitzmann, 2004) presented an experimental study where the EKF was implemented to a fed-batch culture of S. cerevisiae

to filter the glucose measurements and estimate the biomass and glucose concentrations as well as a kinetic parameter (μ_{max}). The glucose measurements were provided by a special flow injection analysis (FIA) system, and the estimated variables were used in a feedback glucose control scheme. In another study, the EKF was implemented to filter the dilution rate and dissolved oxygen measurements in fed-batch cultivation of *S. cerevisiae* (Patnaik, 2005).

(Dewasme et al., 2013b) have studied the glucose and acetate concentrations estimation in fed-batch *E. coli* cultures using the EKF with various measurement configurations. The observability analysis on a mechanistic model and experimental data showed that the biomass measurement is sufficient for the states and growth rate estimations. (Rocha et al., 2006) have studied the parameter, and state estimation by the EKF on the same process using dissolved and exhaust oxygen and carbon dioxide measurements.

(Soons et al., 2007) have applied the EKF to a fed-batch cultivation of *Bordetella pertussis* to estimate the specific growth rate, biomass concentration, and oxygen mass transfer coefficient based on dissolved oxygen measurements. (Krämer and King, 2016) employed the EKF to a fed-batch culture of *S. cerevisiae* to predict the substrate and biomass concentrations using a mass balance model with Monod kinetics and the biomass measurements with NIR spectrometer. Another study on the same bioprocess considered the EKF to estimate the ethanol concentration based on temperature and dissolved oxygen measurements (Lisci et al., 2020).

An on-line estimation of the biomass, xylose, and ethanol concentrations in fed-batch *E. coli* cultures using the EKF is presented in (Hilaly et al., 1992). The estimation is based on the CO_2 evolution and alkali addition rate. The state estimation is combined with an optimization control scheme to maximize ethanol productivity.

In (Markana et al., 2018), the EKF was combined to an economic model predictive control strategy with multiple objectives (substrates and productivity) in a fed-batch bioreactor. The estimation of the state variables is based on the nutrient and foreign protein concentration measurements. The EKF was integrated in the shrinking horizon MPC algorithm, and results showed improved efficiency of the control scheme.

The EKF is one of the most used algorithm for state estimation problems in nonlinear stochastic and uncertain systems. The simplicity and the reasonable performance made the EKF a standard in several fields. A presentation of the EKF algorithm is given in appendix A.

Nevertheless, the EKF strategy have some drawbacks and limitations. The EKF in general is not an optimal estimator like the linear Kalman filter, because only the mean is propagated through the nonlinear model (Julier and Uhlmann, 1997). Furthermore, the linearization procedure in the EKF algorithm may introduce some estimation errors and can lead to the divergence of the algorithm. Especially in the case of bad initial estimate of the state, and incorrect process model.

An alternative to some of these drawbacks is the Unscented Kalman filter (UKF) (Julier and Uhlmann, 1997), presented in the next section.

3.3.3 Unscented Kalman Filter (UKF)

The Unscented Kalman filter (UKF) is introduced as an EKF alternative. It is a derivative free nonlinear variation of the Kalman filter that utilizes the unscented transformation to approximate the statistics of stochastic variables with Gaussian distribution (Julier and Uhlmann, 2004).

(Tebbani et al., 2013) employed the UKF in a *Chlamydomonas reinhardtii* microalgae culture to estimate the biomass, carbon dioxide, and oxygen concentrations based on a molar fraction in the output gas measurements. The estimation is validated through simulations and experiments on a lab-scale photobioreactor.

(Simutis and Lübbert, 2017) presented a hybrid version of the UKF applied to a recombinant therapeutic protein production process using *E. coli*. The UKF was employed to estimate the biomass growth rate and the state variables.

(Marafioti et al., 2009) applied the UKF for state and parameter estimation of a *Porphyridium purpureum* microalgae culture. The biomass concentration is reconstructed based on the total inorganic carbon measurement. The UKF produced better performance in comparison to the EKF when applied to the experimental data.

(Fernandes et al., 2015) presented an EKF and a UKF algorithm applied to a hybridoma cell culture to estimate the glutamine and glucose concentrations from the biomass, lactate, and ammonia measurements in fed-batch and continuous cultures. The UKF outperformed the EKF in terms of estimation accuracy.

(Dewasme et al., 2015) has applied the UKF as a part of a nonlinear model predictive control (NMPC) scheme in fed-batch cultures of hybridoma cells to produce monoclonal antibodies (MAb). The combined closed-loop scheme showed a satisfactory robust response.

(Wang et al., 2010a) have presented a robust version of the UKF, applied for estimating the biomass and substrate concentrations in fed-batch cultures of *S. cerevisiae* using the measurements of dissolved oxygen and carbon dioxide. The results show that the proposed approach presents better accuracy and stability on the state estimation than the strong tracking filter and classical UKF algorithm.

Contrarily to the EKF, the UKF algorithm (presented in appendix A) does not include any linearization procedure and the Jacobian matrices are not required. Thereby, the implementation of the UKF is more flexible, since the nonlinear map can be changed without altering the algorithm structure. The EKF and UKF performance is compared in the simulation in the following sections.

3.4 On-line estimation of the acetate and glucose concentrations in *E. coli* fed-batch cultures

In order to develop the control strategies presented in this thesis, it is necessary to determine the acetate concentration on-line. Since the available sensors in the market are either expensive or not precise in the working concentration range, the acetate concentration is estimated on-line from the biomass measurements.

3.4.1 Observability analysis

Before proceeding to the acetate and glucose concentrations estimation, the *E. coli* dynamic model observability has to be verified for the available measurements.

Observability is a structural system property that depends on the input signal in nonlinear systems, it relates to the possibility of estimating the state variables on the basis of the available measurement information.

The observability analysis of the *E. coli* model structure is detailed in (Dewasme et al., 2013b). The analysis is performed on the substrate and acetate concentrations, considering various measurement setups. The acetate observability analysis is summuarized in this section. First, the model is put in the following canonical form introduced in (Gauthier and Kupka, 1994; Zeitz, 1984):

$$\forall i \in \{1,\dots,q\}, \xi_i \in \mathbb{R}^{n_i}, n_1 \geqslant n_2 \geqslant \dots \geqslant n_q, \sum_{1\leqslant i\leqslant q} n_i = N = \dim \xi,$$

$$\begin{aligned} \dot{\xi} &= \begin{bmatrix} \dot{\xi}_1 \\ \dot{\xi}_2 \\ \vdots \\ \dot{\xi}_{q-1} \\ \dot{\xi}_q \end{bmatrix} = \begin{bmatrix} f_1\left(\xi_1,\xi_2\right) \\ f_2\left(\xi_1,\xi_2,\xi_3\right) \\ \\ f_{q-1}\left(\xi_1,\dots,\xi_{q-1},\xi_q\right) \\ f_q\left(\xi_1,\dots,\xi_{q-1},\xi_q\right) \end{bmatrix} \\ y &= \xi_1 \end{aligned} \tag{3.1}$$

where

- ξ is the state vector.
- y the vector of measurements.
- f_i is a partition of the nonlinear state equations
- q the number of partitions.

The system (3.1) (also called a *Lower Hessenberg System*) is said globally observable if the following condition is satisfied:

$$\operatorname{rank} \frac{\partial f_i}{\partial \xi_{i+1}} = n_{i+1} \;\; \forall i \in \{1,\dots,q-1\} \tag{3.2}$$

To assess if the *E. coli* process is observable, the dynamic model must be put in the canonical form of (3.1).

Under controlled operating conditions, several simplifications of the model are possible. The substrate concentration could be kept close to zero under controlled conditions ($S \approx 0, \dot{S} \approx 0$). In addition, ample oxygenation is provided ($\dot{O} = 0$) (Dewasme et al., 2013b). Under these assumptions, the system (2.21) can be written in the canonical form (3.1) using the following partition:

$$\begin{aligned} \dot{\xi} &= \begin{bmatrix} \dot{\xi}_1 \\ \dot{\xi}_2 \end{bmatrix} = \begin{bmatrix} \dot{X} \\ \dot{A} \end{bmatrix} = \begin{bmatrix} f_1(X, A) \\ f_2(X, A) \end{bmatrix} \\ y &= X \end{aligned} \tag{3.3}$$

Assuming that the biomass concentration is measured, the condition (3.2) reduces to:

$$\operatorname{rank} \frac{\partial f_1}{\partial \xi_2} \neq 0 \tag{3.4}$$

which is verified if the biomass concentration do not vanish ($X \neq 0$).

Similarly, assuming that the acetate concentration is regulated ($\dot{A} = 0$ and $K_{i_A} >> A$) and ample oxygenation is provided ($\dot{O} = 0$) the system (2.21) can be written in the following form:

$$\begin{aligned} \dot{\xi} &= \begin{bmatrix} \dot{\xi}_1 \\ \dot{\xi}_2 \end{bmatrix} = \begin{bmatrix} \dot{X} \\ \dot{S} \end{bmatrix} = \begin{bmatrix} f_1(X, S) \\ f_2(X, S) \end{bmatrix} \\ y &= X \end{aligned} \tag{3.5}$$

The condition (3.2) becomes:

$$\operatorname{rank} \frac{\partial f_1}{\partial \xi_2} \neq 0 \tag{3.6}$$

This condition is also verified if the biomass concentration do not vanish ($X \neq 0$).

Remark. *The observability analysis presented in the chapter considers the operating conditions of the fed-batch phase of the studied bioprocess (substrate in quasi-steady-state). More rigorous analysis should be performed under a higher substrate and acetate range.*

From equations (3.3) and (3.5), the bioprocess model is observable, and well adapted for the estimation of acetate and glucose concentrations based on the biomass measurement. The EKF and UKF strategies can then be applied to the system, following the structure illustrated in Figure 3.2.

3.5 Numerical simulations

In this section, several numerical simulations are achieved to test and compare the estimation of the acetate and glucose concentrations using the EKF and the UKF.

The process and measurement noise covariance matrices $\mathbf{Q}$ and $\mathbf{R}$ are chosen empirically by trial and error to ensure a good compromise between the filter stability, convergence rate, and estimation accuracy. In the following, they are denoted as follows:

$$\mathbf{Q} = diag(\sigma_X^2, \sigma_S^2, \sigma_A^2, \sigma_V^2) \quad \mathbf{R} = r_X^2$$

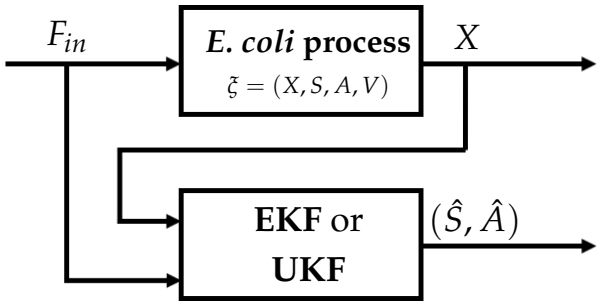


FIGURE 3.2: State estimation in *E. coli* process

where $\sigma_\star$ and r_X are respectively the state and measurement standard deviations.

The initial state covariance matrix $\mathbf{P}_0$ is chosen according to the deviation in the process initial conditions. Table 3.1 summarizes the filters parameters and matrices. The UKF specific tuning parameters (α,β, and κ) were determined from the literature (Fernandes et al., 2015; Julier and Uhlmann, 2004).

TABLE 3.1: EKF and UKF covariance matrices, tuning parameters, and initial conditions

Initial conditions		
X_0	0.1	g/L
S_0	5	g/L
A_0	0.1	g/L
V_0	3.15	L
μ_{set}	0.25	h^{-1}
S_{in}	500	g/L
Sampling time		
T_s	3	mn
UKF parameters		
σ_X	0.01	g/L
σ_S	0.1	g/L
σ_A	0.1	g/L
σ_V	0.001	L
r_X	0.1	g/L
$\mathbf{P}_0$	$10^{-4} \times \mathbb{I}_4$	g/L
α	1	-
β	2	-
κ	0	-
EKF parameters		
σ_X	0.05	g/L
σ_S	0.005	g/L
σ_A	0.02	g/L
σ_V	0.001	L
r_X	0.1	g/L
$\mathbf{P}_0$	$10^{-4} \times \mathbb{I}_4$	g/L

The EKF and UKF are first compared in the ideal model case, where no model uncertainties or mismatch are considered. A robustness study of the filters performance is carried out, considering uncertainties on the simulated process model parameters.

The convergence of the Kalman filters is first tested with erroneous initial conditions. Figure 3.3 shows the performance of the EKF and the UKF in estimating glucose and acetate concentrations based on the biomass concentration measurement affected by additive white noise with zero mean and a standard deviation of 0.1 g/L.

A predefined feeding profile is applied to the system, consisting of a batch phase of 6h followed by a fed-batch phase of 9h with an exponential feeding profile.

After a transient phase of 3 h, both states are well estimated, and the convergence is achieved. Both filters manage to estimate the glucose and acetate concentrations accurately based on the on-line biomass concentration measurement. Moreover, the filters are not affected by the change in the feeding-rate at 12 h, and the estimated glucose and acetate concentrations converge to the real values.

The EKF and UKF performances are similar in the nominal case. They both manage to reconstruct the acetate and glucose concentration signals adequately. Nevertheless, the UKF performance in terms of convergence rate and accuracy is slightly better, as shown in Table 3.2. The root mean square errors of the substrate ($\overline{e_S}$) and acetate ($\overline{e_A}$) estimates during this test include the initial transient phase and are coherent with the measurements' sensitivity and the noise levels (0.1 g/L).

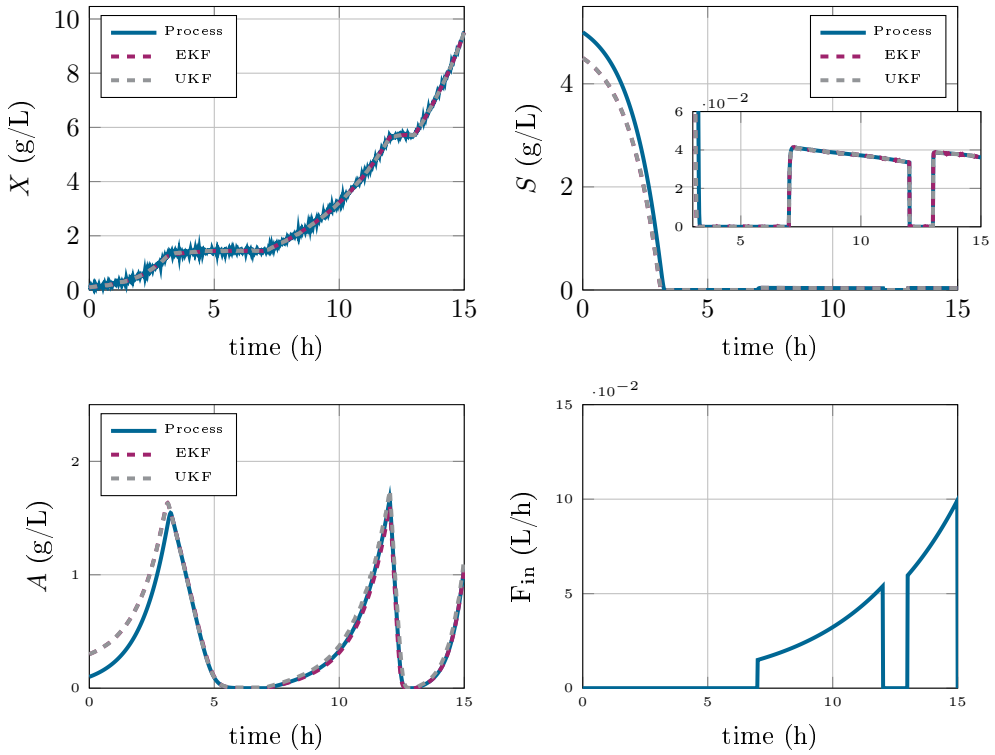


FIGURE 3.3: EKF and UKF applied to the *E. coli* fed-batch process. Estimation of the glucose and acetate concentrations in the ideal model case

TABLE 3.2: error of EKF and UKF in the ideal model case.

	EKF	**UKF**
$\overline{e_S}$ (g/L)	0.203	0.196
$\overline{e_A}$ (g/L)	0.153	0.107

The Kalman filters robustness with respect to model uncertainties is analyzed hereafter. The initial deviations, covariance matrices, and filter parameters are the same as those considered in the ideal model case.

We consider deviations of up to 15% around the nominal value in the plant kinetic and stoichiometric parameters (these deviations are randomly generated according to a Gaussian distribution). Figure 3.4 shows the evolution of the estimated and real process state variables during this test.

Both filters managed to predict the non-measured glucose and acetate concentrations. However, estimation errors are significantly higher than the perfect model case, especially at the end of the culture. This is an expected result with exponential observers, as they highly depend on the model quality. A practical solution is to increase the value of $\mathbf{Q}$ and lower the confidence in the model. However, it is not easy to guarantee the filter stability with high values of $\mathbf{Q}$.

It can be noticed that the UKF provides a closer estimated value to the process values compared to the EKF. This stems from the strong non-linearity in the process kinetic model and the linear approximation of the EKF algorithm. Leading to higher estimation errors, as shown in Table 3.3 when using the latter filter.

The UKF performance is better than the EKF in the model mismatch case. However, both approaches are highly dependent on the model quality. Therefore, any implementation of these estimation strategies must consider model mismatch in the design procedure and the parameters choice.

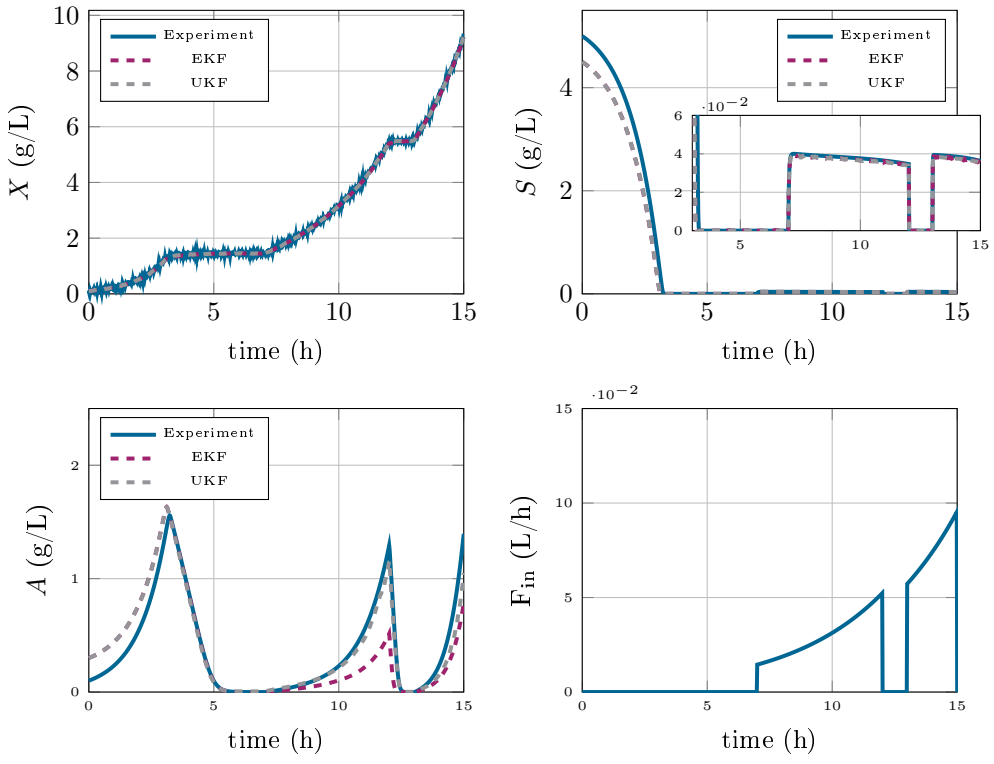


FIGURE 3.4: EKF and UKF applied to the *E. coli* fed-batch process. Estimation of the glucose and acetate concentrations in the model mismatch case

TABLE 3.3: error of EKF and UKF in the model mismatch case.

	EKF	**UKF**
$\overline{e_S}$ (g/L)	0.227	0.225
$\overline{e_A}$ (g/L)	0.183	0.157

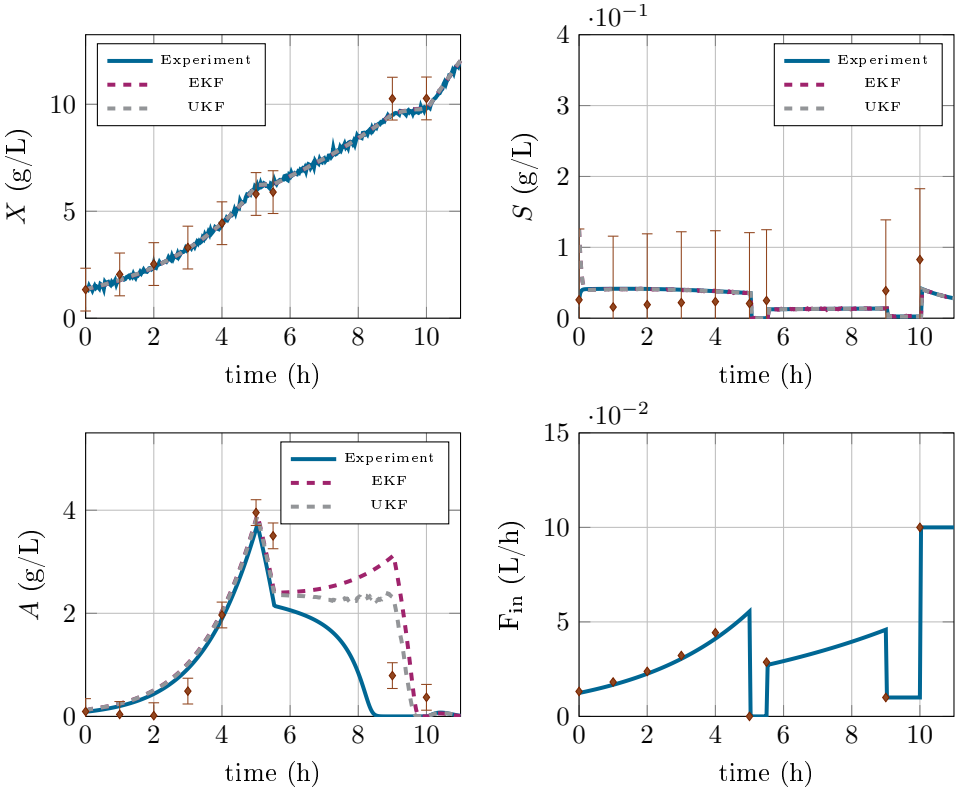


FIGURE 3.5: EKF and UKF applied to the *E. coli* fed-batch process. Comparison with experimental data from (Retamal et al., 2018)

Figure 3.5 illustrates the filters performance under real operating conditions, compared to experimental data from (Retamal et al., 2018). The model and the filters are initialized with the off-line measurements and the experimental feed-rate is applied as the model and estimators input. The Kalman filters showed a good performance in estimating the glucose and acetate concentrations, especially in the latter case where the model prediction is far from the measurements at 8h and 9h.

In the following chapters, the UKF is preferred for state estimation problems since it provides better performance than the EKF. An experimental validation of the UKF is performed using new data. Nevertheless, it still depends on the model accuracy, which should be addressed in control design using bioprocess models.

3.6 Conclusion

In this chapter, a brief overview of the estimation methods used for bioprocess monitoring is presented. The estimation method choice depends on several factors: The application in question, the required prediction accuracy, the convergence rate, the computational cost, the availability of previous data, and the availability of a detailed process model.

A focus is then placed on the Kalman filter due to its practicality and wide use for state and parameter estimation problems in biotechnological applications.

Due to the inherent nonlinear nature of bioprocesses, a nonlinear variation of the Kalman filter is considered for state estimation. The Extended and Unscented Kalman filters are applied to the *E. coli* fed-batch process to estimate the glucose and acetate concentrations based on the biomass measurements and the mechanistic model.

The filters performance is put to the test through numerical simulations. Both filters accurately predicted the glucose and acetate concentrations, with a slightly better performance of the UKF, especially when considering modelling errors and uncertainties. However, they both rely on the quality of the process model.

In the following chapters, the UKF is implemented on-lined and combined with the developed control strategies.

Chapter 4

Bioreactor harware and software setup

4.1 Introduction

Experimental validation of the developed control and estimation strategies requires efficient monitoring and utilizing the available hardware and software. One of the main obstacles to implement advanced optimization and control algorithms in bioprocesses is interacting efficiently with the hardware. The diversity of the sensor configurations and the closed environment of commercial monitoring software are significant problems.

A first step in the experimental validation of the algorithms presented in this work is implementing a closed-loop structure in one main program since the bioreactor is conceived to operate in open-loop mode. The developed program should contain all the available measurements from the reactor's control unit and additional measurements in one interface. This information should be available in real-time and can be manipulated by the control and estimation algorithms. The developed program should also be able to actuate the system input represented by the pump's feed-rate.

The monitoring and control of the bioreactor main variables during the fed-batch cultures is described in this chapter. The reactor setup and main components are presented, as well as the digital control unit and on-line sensors. Finally, the off-line measurements and analysis performed during the experiments are described.

4.2 Bioprocess monitoring software

The efficient operation of bioprocesses requires the availability of on-line measurements for the operator. This information should be easily manipulated, visualized, and recorded in real-time.

In addition to the lack of reliable sensors for several essential components, it is often challenging to incorporate the available measurements in on-line control and estimation schemes due to the diversity of sensor configurations and connectivity. Main variables like pH, temperature, dissolved oxygen, airflow, and agitation speed are measured and controlled by the bioreactor integrated units. These components are often specific to the manufacturer. Their communication protocols are not universal and require dedicated commercial software.

Some examples of commercial software for bioprocess monitoring are: UBICON (Universal Bioprocess Control System) by Software Electronic Systems Design, (Germany); AFS (Advanced Fermentation Software), by New Brunswick Scientific Inc. (USA); BioCommand FermExpert software by BioExpert Ltd. (Estonia).

Most programs allow data-acquisition and control for the main variables, data recording, visualization, and open-loop control of the feeding rate. However, several works reported the limited applications of these programs: (Shin et al., 2009; Soccol et al., 2012).

The commercial software's main disadvantage is the specificity to the manufacturer's hardware and the difficulty of manipulating the measurements outside the software defined environement.

A solution to these setbacks is developing dedicated software in a programing language (C++, C#, Python ...). This approach advantage is integrating advanced control and estimation algorithms, manipulating measured variables, and flexibility in terms of data manipulation. However, this approach is time-consuming and requires knowledge of different communication protocols. Examples of dedicated programs developed for data acquisition and control can be found in (Jaén et al., 2017; Pablos et al., 2014).

The MFCS program from B.Braun Biotech International (Germany) is an exclusive software for Biostat bioreactors. The MFCS has several features allowing data acquisition and control of the digital control unit of the bioreactor. However, it does not allow the implementation of user control algorithms and integration of external measurement devices. Examples of applying the MFCS for monitoring and control of *E. coli* fermentation can be found in (Huang et al., 1995; Tomson et al., 1995).

In most cases, the MFCS is used for data acquisition and saving, and additional measurements are implemented and visualized in a separate program. If the bioreactor control unit is not equipped with a specific measurement input (e.g. light intensity measurements), the operator is obliged to use separate software to visualize and record the new measurements. This increases the difficulty of implementing control and estimation schemes based on these measurements, especially when using multiple sensors from different manufacturers.

(Rocha, 2003) found the MFCS program inadequate for acetate regulation in fed-batch *E. coli* cultures due to the difficulty to integrate the additional measurements from a Flow Injection Analysis (FIA) system, and to implement adaptive control algorithms. A LabView program was developped to acquire on-line measurements of the acetate concentration from the FIA measurement system as well as the main reactor measurements. This information is then integrated into a closed-loop adaptive control scheme. The approach used to develop this program required setting up the communication protocols with the reactor measurements for every variable.

In this chapter, we present a data acquisition and control solution developed in the framework of this thesis in the LabView programming environment. In this program, communication is established directly between the MFCS software and the LabView program, allowing to port the measurements to the program in real-time using shared library nodes.

Labview is a visual object-oriented programing language dedicated to data acquisition and monitoring systems. It allows the development of virtual instruments and communication with several types of hardware with various communication protocols.

A LabView virtual instrument (VI) comprises three main parts: The front panel, the block diagram, and the connector. The front panel represents the user interface. It contains the main controls in the form of buttons, knobs, and sliders. It can display the data in the form of numerical indicators, graphs, charts, and other forms. The block diagram is the main code of the VI. It contains a vast library of dedicated blocks for various applications (Signal analysis, communication, mathematical expressions ...) connected by virtual wire. The connector is responsible for data flow between subroutines. Every VI can be integrated as a subVI in another VI by specifying the Input-Output map, making the applications modular and the workflow easy.

LabView has a dedicated serial communication library called VISA (Virtual Instrument Software Architecture) that offers a programming interface for serial port communication with most instruments' interfaces (RS232/485, Ethernet, USB).

LabView has been widely used for bioprocess data acquisition purposes. Examples of LabView applications for fermentation supervision can be found in (Johnsson et al., 2013; Stratz and Dittrich, 2015).

In the following sections, the data-acquisition solution implemented in LabView environment is described. The goal is to integrate the different on-line measurements from the bioreactor Digital Control Unit (DCU) and the biomass probe in one main interface.

4.3 Bioreactor monitoring and control

4.3.1 Bioreactor setup

The cultivations were performed in a Bioreactor consisting of a 5L jacketed glass vessel and a digital control unit or (DCU) (BIOSTAT B plus, Sartorius Stedim Biotech, Germany). The reactor is equipped with a water jacket and an agitation motor.

The monitoring of the cultures is possible thanks to a potentiometric pH sensor (Hamilton, Switzerland), dissolved oxygen (DO) probe (Hamilton, Switzerland), and a temperature sensor (Sartorius, Germany). These probes are directly connected to the DCU which allows the control and regulation of the main variables.

Also, biomass concentration is available on-line via an absorption-based photometric turbidity probe (Fundalux II, Sartorius, Germany). However, this probe cannot be connected to the DCU due to the lack of a suitable connection in the DCU configuration.

4.3.2 Monitoring interface

All the programs presented in this chapter were conducted in NI LabView 2016 environment. The computer used for the experiments has an Intel I3 3.2 Ghz processor with 8Gb of ram running on Windows 10.

As stated before, the main process variables (pH, temperature, pO2, ...) are controlled by the DCU and monitored on-line via the MFCS software. This real-time information can be ported to LabView using the shared-library nodes. These nodes allow to execute functions in dynamic library files (DLL), provided the function definition and configuration. Luckily, this information is provided by the MFCS, and the functions used to retrieve the values of the variables are well defined. Hence, all variables contained in the MFCS can be displayed, manipulated, and saved in LabView. An immediate advantage of this method is the flexibility of the sampling and recording rate since the MFCS rate is locked at 5 mn. The front panel representing the user interface of the LabView program is shown in Figure 4.1.

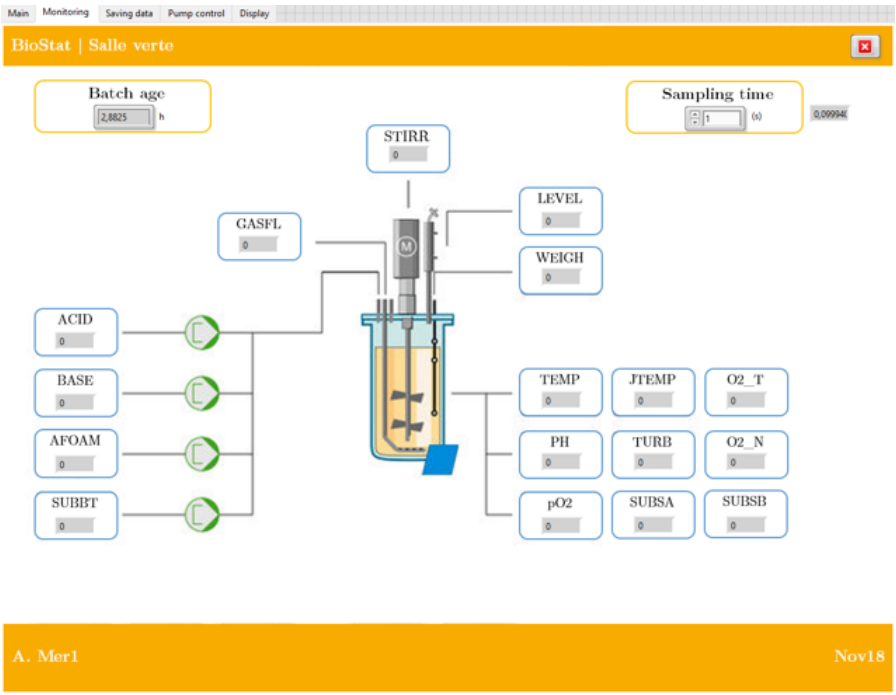


FIGURE 4.1: Front panel of the data aquisition interface

The biomass signal is measured separately by a data acquisition device from national instruments (NI USB-6000USB Multifunction DAQ Device, National Instruments, USA) and squired in the same LabView program using DAQMax library. The advantage of the programing in Labview is the combination of various measurements from different sources in the same environment and the ability to fix an appropriate sampling rate for these measurements and apply signal analysis tools (Filtering, FFT . . .).

All the individual measurements are integrated into one program that displays the variable values, plots chosen variables, and saves them in a spreadsheet with a specified recording time. Additionally, the program controls external peristaltic pumps and manipulates the feed-rate in real-time via the VISA module, thus creating a closed-loop structure. Furthermore, thanks to the network shared variables feature of LabView, these measurements are shared via the network and can be accessed in real-time by authorized computers or servers on the same network. This feature allows the remote supervision and control of the reactor and the implementation of complex algorithms on a separate powerful computer.

The control and simulation algorithms were coded in Matlab for simulations and implemented on-line in LabView using shared-library and MathScript nodes.

A diagram of the experimental setup showing all the devices described in this chapter is shown Figure 4.2, and a picture of the experimental setup is shown in Figure 4.3.

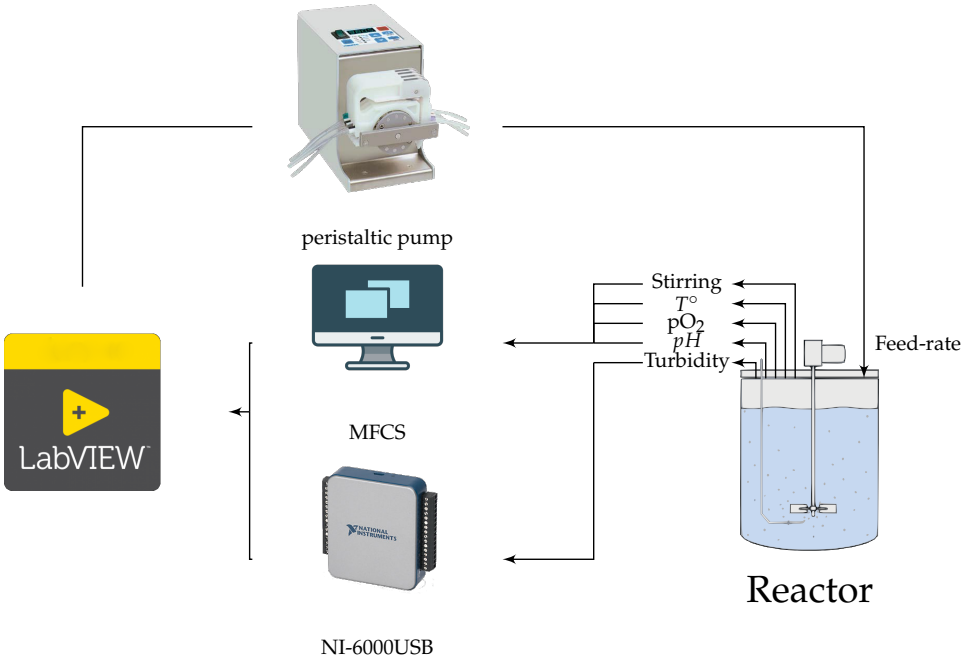


FIGURE 4.2: Real-time implementation diagram

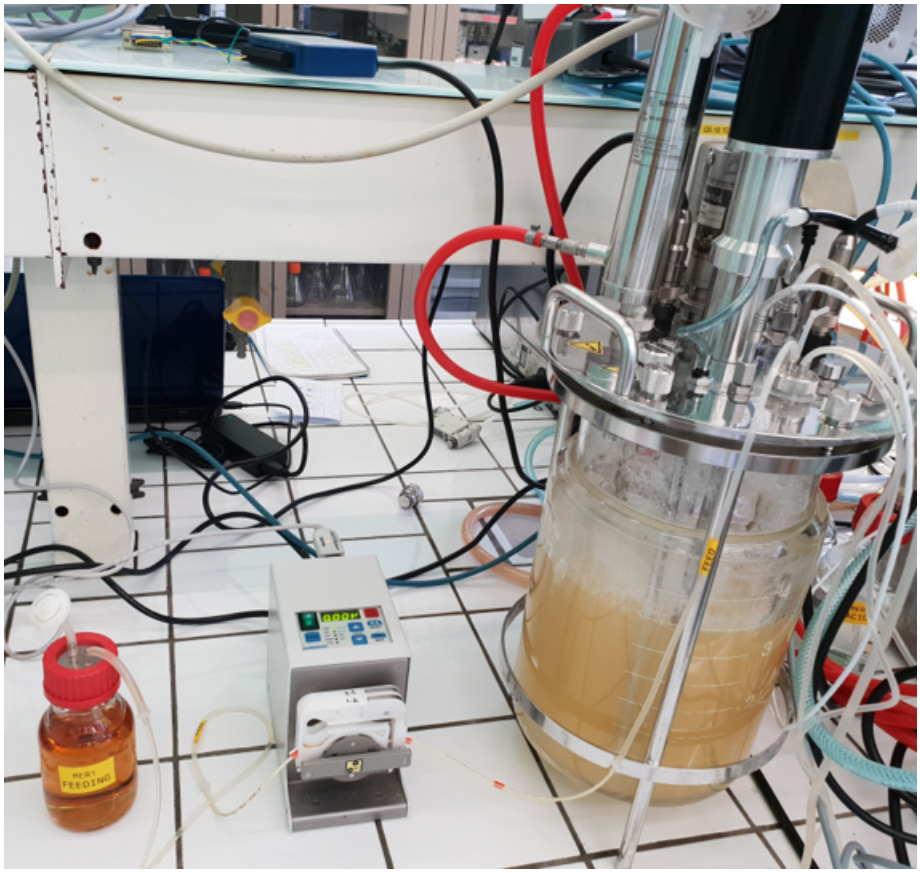


FIGURE 4.3: Picture of the bioreactor and the feeding system

4.4 Off-line measurements

In order to validate the on-line measurements of the biomass concentration, and the estimation quality of the glucose and acetate concentrations, several off-line analyses were performed on the samples. In the following sections, a brief description of the off-line analysis methods used during the cultures is provided.

4.4.1 Biomass measurements

Off-line biomass measurements were determined by measuring the optical density (OD) at 600 nm in a UV spectrophotometer (Shimadzu, Pharmacia Biotech, USA). Samples were diluted with deionized water to obtain OD in the linear range (0-0.3 absorbance units), and then correlated with dry cell weight (DCW) using a calibration curve. The calibration method and curve can be found in appendix B.

4.4.2 Glucose Measurements

Glucose concentration was determined using the di-nitrosalicylic acid (DNS) method. A reagent is prepared according to the procedure described in (Retamal et al., 2018; Rocha, 2003).

The glucose concentration is correlated using a calibration curve performed within the range 0.1-2 g/L. The measurement and calibration procedure are presented in appendix B. The sensitivity level of the measurements is considered around 0.1 gL^{-1}.

4.4.3 Acetate measurements

Acetate concentration measurements were performed with an enzymatic kit (Megazyme, Ireland) according to the manufacturer's instructions. The absorbance measurements were carried out in a Sunrise moduler absorbance microplate reader (Tecan, Austria). The calibration curve for acetate measurements is presented in appendix B. The sensitivity level of the measurements is considered around 0.1 gL^{-1}.

4.5 Cultures and experiments

The fermentations performed throughout this work were operated in batch and fed-batch modes. Initial batch cultures of $V = 3.5L$ were pre-equilibrated to the appropriate operating conditions (pH, temperature, DO) before inoculation with 5% v/v seed culture, where the initial OD_{600} in the fermenter reaches 0.3-0.6. Sterile filtered anti-foam was added via a peristaltic pump when necessary throughout the cultivations. The batch-phase is monitored during the day. The flag for the beginning of the feeding was the increase of the dissolved oxygen resulting from the exhaust of the glucose and acetate in the medium. Once the glucose is nearly depleted, the fed-batch phase starts, and the feeding solution is added with a rate

determined by the controller, and applied by a Reglo-digital peristaltic pump (Ismatec, Germany).

4.5.1 Microbial strain

The *E. coli* BL21 (DE3) strain was used for all fermentations. This strain is more suitable for high cell density cultivation due to low acetate production and lower sensitivity to varying growth conditions compared to other strains (Müller et al., 2018).

4.5.2 Operating conditions

The batch and fed-batch fermentations were conducted under controlled conditions summarized in Table 4.1. These set-points were determined according to literature (Retamal et al., 2018; Rocha, 2003). The pH is regulated by titration of 12.5% ammonium hydroxide (base) and phosphoric acid 0.5 M (acid). Dissolved oxygen was maintained above 30% air saturation by a two-level controller, increasing the agitation rate when the oxygen demand of the cells increases. When the maximal agitation rate is reached, the manipulated variable shifts to the airflow. Minimum values for airflow and agitation were imposed (1 L-min^{-1} and 200 rpm, respectively).

The temperature is controlled by the DCU at 37 °C using a heating water jacket.

TABLE 4.1: Set-points of the operating conditions

Variable	Setpoint	Controller
Temperature	37°C	PID
pH	7	PID
Dissolved Oxygen	30%	Two-level

4.5.3 Medium composition

The composition of the pre-cultures, batch, and feed medium is given in Tables 4.2 and 4.4. The mediums were prepared according to the protocol cited in (DeLisa et al., 2001; Rocha, 2003). During the preparation, solutions were filtered and sterilized in the autoclave to avoid contamination.

A minimal M9 medium with addition of trace metals and vitamins was used for the biomass regulation experiments. The M9 medium was applied to fed-batch control experiments of *E. coli* cultures by (Rocha, 2003). It has the advantage of being cheap and has a very low autofluorescence. During the acetate regulation experiments, a defined high-density fermentation medium (HDF) (DeLisa et al., 2001) was used. This medium allows high density and growth rates.

In all experiments, the salts and glucose solution of the batch medium was sterilized in the autoclave for 120 mn at 120-200 °C with a total volume of 2.9 L.

A 100 mL of the trace mineral solution was filter sterilized using acrodisc syringe filters (from Pall Corp., USA) and injected in the bioreactor using peristaltic pump prior to the batch phase.

4.5.4 Pre-cultures

Bacteria from frozen cryotubes were transferred into a petri dishes containing (per liter) 25 g of LB medium (Table 4.2) and 15 g of Agar, the dishes are incubated at 37°C for 24h. A single colony is transfered to 250 mL shaker flask containing 50 mL of LB medium, incubated at 37°C for 8h.

A volume of 10 mL of this culture was then inoculated into two 500 mL shake flasks containing 250mL of the application medium (Tables 4.3 and 4.4) , and incubated over night $(14 - 16h)$ at 37°C (DeLisa et al., 2001). This culture (500 mL) was used to innoculate the bioreactor to a final volume of 3.5 L, the minimal volume to ensure the complete immersion of the biomass probe.

TABLE 4.2: Composition of the LB media used during preparations

Component	**Concentration** (g/L)
Peptone	10
Yeast extract	5
NaCl	6

4.5.5 Experimental steps

In order to ensure reproducibility of the experiments, the following sequence of operating steps was performed during all the cultures:

Bioreactor preparation for autoclave

1. Prepare 2.9L of the salt media (M9 or HDF) and add it to the reactor vessel.
2. Turn on the stirrer and set it to 150 rpm to ensure the medium homogeneity.
3. Calibrate the dissolved oxygen probe 0 % with using a 2.0 M sodium sulfite (Na_2SO_3) solution.
4. Calibrate the dissolved oxygen probe 100 % by sparging the bioreactor with a maximal airflow (10 L-min^{-1}) and a maximal stirring (800 rpm).
5. Calibrate the pH probe with pH 4 and 7 standards and insert it into the bioreactor.
6. Prepare the tubing for the airflow, acid-base, feeding, and anti-foam.
7. Autoclave the bioreactor overnight.

TABLE 4.3: Composition of the M9 medium (Rocha, 2003)

Components	**Batch medium** (./L)	**Feeding solution** (./L)
Glucose	5–10 g	500 g
Salts solution		
Na_2HPO_4	6 g	-
KH_2PO_4	3 g	-
NH_4Cl	1 g	10g
NaCl	0.5 g	-
$MgSO_4 \cdot 7H_2O$	0.12 g	4g
Trace metals and vitamin solution		
$FeCl_3$	27 mg	-
$ZnCl_2$	2.0 mg	-
$CoCl_2$	2.0 mg	-
$NaMoO_4$	2.0 mg	-
$CaCl_2$	1.0 mg	-
$CuCl_2$	1.0 mg	-
H_3BO_3	0.50 mg	-
HCl	100 mg	-
Kanamicin	5 mg	-
Tiamine	3.4 mg	-

Launching the culture

1. After sterilization, connect the different tubing and probes, and prepare the feeding system (vessel, pump).
2. Sterile addition of 100 mL trace and vitamin solution.
3. Initialize temperature control and wait for stable readings.
4. Initialize pH control.
5. Initialize the cascade control of dissolved oxygen.
6. Launch MFCS and the LabView data acquisition and application (Control, Estimation) interfaces.
7. Inoculate for beginning the batch phase.
8. When the on-line flags for glucose depletion appear (sudden increase of pO2, pH increase), begin the fed-batch phase.

TABLE 4.4: Composition of the HDF media (DeLisa et al., 2001)

Components	**Batch medium** (./L)	**Feeding solution** (./L)
Glucose	5 g	500. 0g
Salts solution		
KH_2PO_4	13.3 g	–
$(NH_4)_2\,HPO_4$	4.0 g	–
$MgSO_4 \cdot 7\,H_2O$	1.2g	20.0g
Citric acid	1.7 g	–
Trace metals and vitamin solution		
EDTA	8.4 mg	13.0 mg
$CoCl_2 \cdot 6H_2O$	2.5 mg	4.0 mg
$MnCl_2 \cdot 4H_2O$	15.0 mg	23.5 mg
$CuCl_2 \cdot 4H_2O$	1.5 mg	2.5 mg
H_3BO_3	3.0 mg	5.0 mg
$Na_2MoO_4 \cdot 2H_2O$	2.5 mg	4.0 mg
$Zn\,(CH_3COO)_2 \cdot 2H_2O$	13.0 mg	16.0 mg
Fe^{III}Citrate	100.0 mg	40.0 mg
Thiamine ·HCl	4.5 mg	–

4.6 Conclusion

In order to efficiently operate the bioreactor, and implement the control and estimation methods presented throughout the thesis, a data-acquisition program is developed in the LabView environment.

This program provides the communication with the MFCS software, allowing real-time measurements in the LabView environment with more flexibility (sample rate, recording, exporting, real-time plotting of data ...).

The program also allows integrating several measurements in one environment and is flexible for a hardware upgrade. Network-shared variables feature in LabView offers the possibility of sharing the same information between computers on the same network. Meaning that the closed-loop system does not require a powerful computer to operate. The data acquisition interface can share the variables in real-time, and the complex calculation can be performed in another powerful computing machine or a server in another room.

Another advantage is the possibility to integrate modular control and estimation algorithms. The modular structure of the program allows modifying these algorithms quickly. The user can swap a component with another without the need to build a new program. The user-friendly interface facilitates the operation of the program.

The different materials and methods used in this work are presented in this chapter, including the bioreactor structure and hardware, the bacterial strain, the mediums, and the culture operation steps.

The following chapters describe the control and estimation strategies to maximize biomass production yield and promote cell growth in fed-batch *E. coli* cultures.

Chapter 5

Overview of control strategies for fed-batch cultures

5.1 Introduction

Fed-batch fermentation is the most common cultivation mode in industrial processes (Birol et al., 2002; Bodizs et al., 2007). This mode is preferred for large-scale production of several products such as chemicals, enzymes, pharmaceuticals and food products (Doran, 2013).

The progress in engineering tools developed for biotechnological applications focused on improving the productivity of the processes, improving the production yields, and reaching optimal operating conditions. These objectives can be

achieved by multiple approaches like genetic and genomic engineering of host strains, or by designing efficient control and optimization strategies (Pontrelli et al., 2018).

Several control strategies were developed throughout the years for bioprocesses. These types of systems have indeed specific features such as the strong non-linearities, the parametric uncertainties, and the multiple metabolic operating modes.

In this chapter, the main strategies based on feed rate manipulation in fed-batch cultures are presented and discussed. In all strategies, the goal is to determine and maintain an optimal feeding profile corresponding to the desired objectives behind the fermentation. The main difficulty is to define this optimal feed-rate, since it depends on several variables of the bioprocess, as well as the operating conditions.

Manipulating the feed-rate has a direct impact on the metabolism of the microorganisms, since the variation in the substrate concentration affects the growth rate, the products formation rate, the oxygen uptake rate, and other kinetic variables. Furthermore, a variation of the feed-rate directly changes the mass and volume dynamics, thereby affecting the bioprocess variables, and especially the oxygen dynamics due to the change in viscosity.

Closed-loop control schemes can maintain the optimal feed-rate corresponding to the desired objectives, and in most cases adapts to disturbances affecting the metabolism of the cells and measurement noise.

The list of the control strategies presented in this chapter is not exhaustive, but it may provide a guide to choose the appropriate control method to develop and implement, based on the available materials, process knowledge, or the amount of available data. The list considers the most widely used control strategies currently implemented either through industrial applications or lab-scale experiments.

5.2 General aspects of the control problem

5.2.1 Control objectives

The objective behind implementing a control scheme usually depends on economic drivers of the bioprocess. For example, the production of recombinant proteins and pharmaceutical products requires the maximization of the biomass concentration and the biomass production yield. Other objectives include: Maximizing/minimizing the product concentration, maximizing the process yield, maximizing the productivity, and maintaining an oxygen uptake rate profile.

Due to the exponential growth rates, metabolic shifts, varying volume dynamics, nonlinear kinetics, and feed disturbances, fed-batch processes dynamics are inherently nonlinear. Therefore, a major challenge to reach the defined control objective is the constantly changing nature of the optimal operating points in nonlinear dynamics.

5.2.2 Measured variables

A major obstacle for efficient and advanced control of fed-batch cultures is the limited number of on-line measurement devices. The standard measurements found in fed-batch cultivations include temperature, pH, dissolved oxygen, and carbon dioxide. Adding to these list, we find the on-line measurements of the agitation speed, and inlet gas rates (Alford, 2006; Sonnleitner, 2013).

Other advanced probes like spectroscopic sensors can be found on the market (Cervera et al., 2009), however they are not widely used on the industrial scale due to the lack of robust measurements at low concentrations of key variables, such as substrate or by-products concentrations.

5.2.3 Estimated variables

The lack of efficient sensors for key variables is a major challenge in industrial fermentations (Luttmann et al., 2012; Montague et al., 1989), since the process optimization and control requires robust on-line data. Thereby, software sensors can provide an alternative tool to include additional variables in the developed strategies based on the available measurements.

As described in chapter 3, software sensors use the available on-line data from the physical sensors and the process knowledge (in the form of dynamic models or historical data sets) to efficiently reconstruct and estimate the non-measured variables (Bogaerts and Vande Wouwer, 2003; Montague et al., 1989). Hence, estimation strategies are not only a powerful tool for on-line process monitoring, but also provide additional data to implement advanced control and optimization strategies (Luttmann et al., 2012; Sagmeister et al., 2013).

5.2.4 Control inputs and variables

In the case of control strategies designed for fed-batch cultures presented in this chapter, the control input is the inlet feeding flow rate. The controlled variable can be a measured variable (pH, oxygen, biomass ...), an estimated variable (by-product, substrate, ...), or a calculated term from the measured and estimated variables (growth rate, uptake rate, ...). The control strategy can also include multiple controlled variables, and multiple inputs. A general scheme of the main components of a control strategy is shown in Figure 5.1.

The control design procedure must take into consideration the parametric uncertainties and disturbances. Furthermore, the robustness of the developed strategies towards batch-to-batch variations, initial conditions errors, and measurement noise must be evaluated.

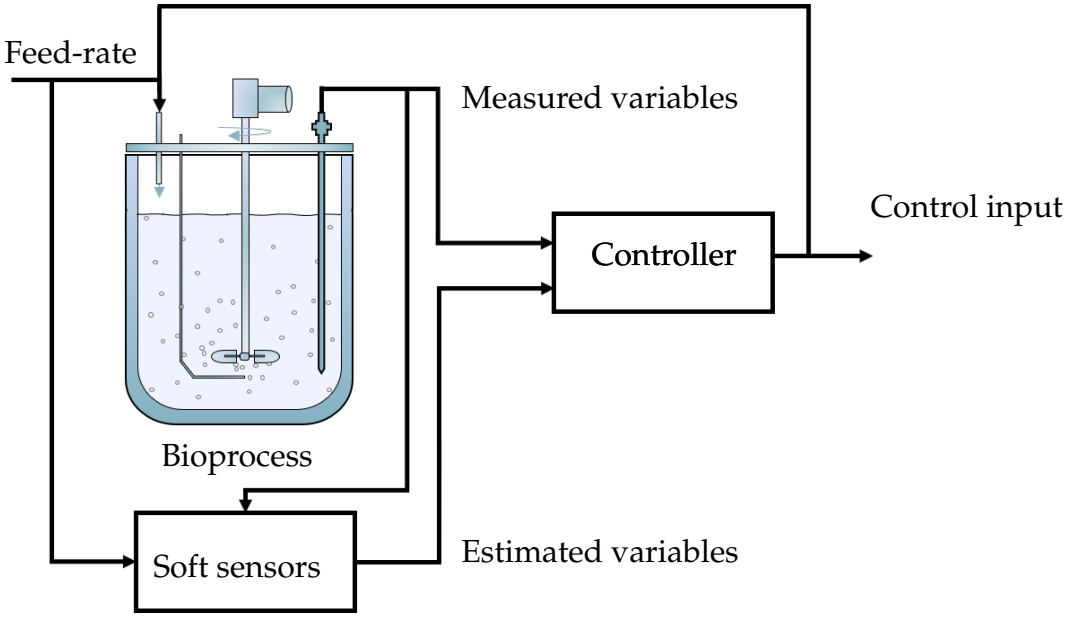


FIGURE 5.1: General scheme of the control strategy of a fed-batch process

5.3 Overview of control strategies for fed-batch cultivation

This section describes the most used control strategies developed for fed-batch processes. This presentation is not exhaustive but is intended to provide guidance.

5.3.1 Predetermined feeding control

Predetermined feeding control (also called open-loop control or non-feedback control) is the most used operating strategy in the fermentation industry (Oliveira et al., 2004). In this method, the control actions are independent from the process outputs and the manipulated variables. Thus, a feedback is not required to determine the controller actions. Non-feedback control is employed to apply a pre-calculated feeding profile to the process, based on the initial states and the operating conditions of the culture.

In fed-batch cultures, a pre-calculated exponential feed-rate profile is usually considered to achieve reproducibility of batches at the end of the exponential growth phase. Exponential feeding allows the cells to grow at a constant specific growth rate (Lee et al., 1999). This feeding method is usually applied to cells that exhibit overflow metabolism. In order to avoid cell growth inhibition and product formation, the growth rate is usually regulated below the maximal value.

The feeding profile is generated based on the initial conditions, and specific parameters such as the maximum specific growth rate (μ_{max}). An example of the

feeding profile is given by Equation (5.1) (Dochain, 2010; Lee et al., 1999), derived from mass balances with the assumption of a constant cell yield and constant growth rate.

$$F(t) = \frac{\mu_{set} \cdot X_0 V_0}{Y_{XS} \cdot S_{in}} e^{\mu_{set} t} \tag{5.1}$$

where $\mu_{set} < \mu_{max}$ is the set-point growth rate. Y_{XS} the yield of biomass on substrate, X_0 and V_0 are respectively the initial biomass concentration and culture volume.

The non-feedback control is used to reduce the variability between batches caused mainly by variations in initial biomass concentration (Jenzsch et al., 2006a). Examples of open-loop applications to fed-batch processes can be found in (Henes and Sonnleitner, 2007; Jenzsch et al., 2006a; Wechselberger et al., 2012).

The main advantage of this method is that it doesn't require any on-line measurements of the key variables and its implementation is fairly easy. However, as a direct consequence, it does not allow to regulate process variables, or reach most of the control objectives. Furthermore, process disturbances and system nonlinearity are major issues, especially when cell growth does not match the predetermined profile due to overfeeding or under feeding. Exponential growth of the cells leads to significant deviations in biomass profile from its reference (Lee et al., 1999).

5.3.2 Adaptive control

Adaptive control is a nonlinear control strategy, where the model parameters are estimated and adapted in real-time. Adaptive control covers a wide range of methods, but they all rely on changing a defined set of parameters in the control law in order to achieve a better response in presence of disturbances or uncertainties in the process. The adaptive controller types vary depending on the parameter adaptation method.

A simple scheme of adaptive control is the gain scheduling method. Gain scheduling techniques are also called "open-loop" adaptive strategies since the controller gain is adapted based on prior knowledge of the system to account for varying dynamics. A preprogrammed tuning is calculated and applied based on data from previous cultures. An example of the application of the gain scheduling method to fed-batch cultures can be found in (Hisbullah et al., 2001; Yuan et al., 2009). Gain scheduling strategies gained interest due to their simplicity to implement on specific manufacturers hardware. However, they lack robustness towards model uncertainties and disturbances.

Closed-loop adaptive control strategies can be divided into two main categories: direct or indirect, depending on the parameter tuning method. In direct adaptive control, the controller variables are adapted directly based on system measurements. In other words, the tracking error between the process and a reference model is used to directly modify a parameter in the control law.

In a variant of direct adaptive control, the optimal control action is designed to minimize the error between an ideal model and the process outputs. This method

refers to model reference adaptive control (MRAC), and has been applied in several fed-batch control applications by feed-rate manipulation (Oliveira et al., 2005, 2004; Soons et al., 2006a). The desired response of the system is defined by a model, that has optimal response to input disturbances. The minimization of the error between the actual response and the desired one is used to tune the controller parameters (Landau et al., 2011).

In indirect adaptive methods, the measurements are used to update the model parameters to reduce the error between the system response and the desired one. Then the controller parameter values are calculated based on the updated model parameters.

Nonlinear control methods such as feedback linearization and backstepping control are usually employed to bioprocesses with the addition of an adaptation of a set of parameters in the control scheme. These parameters are estimated online based on the measurements and the process model. The controller is then able to adapt to unpredictable system dynamics. The measurements are processed to predict and update a state variable or a system parameter, before their use in the adaptation algorithm. There is a big interest for adaptive control schemes based on traditional nonlinear controllers in fermentation applications due to the nonlinear and uncertain nature of bioprocesses (Bastin and Dochain, 1990; Gonzalez et al., 2015; Rocha and Ferreira, 2002; Smets et al., 2004).

In (Dewasme et al., 2011a), adaptive control is applied to control the by-product concentration P to a low predefined setpoint in fed-batch *E.coli* and *S. cerevisiae* cultures. A robust linearizing adaptive control law is implemented to estimate and adapt a kinetic parameter θ. The robustness of the closed-loop is ensured using the $\mathcal{LMI}$ formulation. The control scheme is compared to classic adaptive strategy developed in (Rocha, 2003), and the results showed a better performance in the presence of modeling errors and disturbances.

In (Oliveira et al., 2004), an indirect MRAC algorithm was applied to regulate the dissolved oxygen concentration and maximize the Oxygen transfer rate (OTR) in penicillin-G fermentation. First, a model identification is carried out from measured variables, then, the error from the reference model output is used to update the control parameters.

With the same control objective, the MRAC was used on an industrial 50L fed-batch fermentation of *P. pastoris* in (Oliveira et al., 2005). The OTR signal was estimated from the available state variable signals to reduce the measurement delays. The strategy allowed reaching a high cell density, with a biomass up to 670 g of wet weight/L.

(Soons et al., 2006b) applied a direct MRAC to regulate the growth rate μ to a defined set-point in a 5L *B. pertussis* fed-batch cultures. The method is based on a mechanistic model with two substrates compared to a reference model. The controller gains are calculated based on the updated state variables (μ, X and V).

Adaptive control is suitable to fermentation systems with high disturbance levels and unpredictable system dynamics (Landau et al., 2011). A general challenge for adaptive control is the determination of the Parameter Adaptation Algorithm (PAA) used to adapt the control parameters to reach optimal operating points and

ensure a stable system response. Another challenge is the need of sufficiently rich signals for the parameter and state estimation.

5.3.3 Model predictive control

Model predictive control (MPC) is an advanced control strategy used for a variety of control problems (Forbes et al., 2015), since it is based on on-line optimization and is known for handling complex multivariate dynamics and system constraints. These features allow the MPC strategies to be applied in a wide range of industrial fields (Qin and Badgwell, 2003).

The basic principle behind the MPC is to simulate the model up to a defined time in the future (called prediction horizon) in order to predict the current output and the future evolution of the system. The prediction is evaluated to define the action required at the current time, based on optimization of a cost function and a given defined prediction window (Stanke and Hitzmann, 2013).

The MPC scheme has been extended to nonlinear systems (NMPC), and used for a variety of control objectives. A cost function can be defined to maximize the production, minimize the costs, or to follow a trajectory for a certain variable (Seborg et al., 1989). The optimization is carried out up to a predetermined point in the future, and the algorithm is then repeated at the next time interval. The optimization of the cost function can also be subject to a set of operational constraints (Seborg et al., 1989).

The main difference between the different MPC implementations in fermentation applications is the choice of the cost function. Some applications aim to follow a defined trajectory for a specific variable, such as biomass concentration, X (Kuprijanov et al., 2013a; Zhang and Lennox, 2004) or substrate concentration S (Craven et al., 2014). In other examples, the objective is to maximize a certain process variable (Chang et al., 2016; Kovárová-Kovar et al., 2000a; Santos et al., 2010), rather than follow a determined trajectory.

The MPC is based on a mechanistic or data driven model of the system, therefore the efficiency of this control strategy depends heavily on the quality and accuracy of the model. Furthermore, the method is considered to have an expensive computational cost, especially when handling uncertain steep nonlinear dynamics (Laurí et al., 2014).

The MPC strategy is considered as a standard method in several industrial fields, however its application in industrial bioprocesses is not as wide, due to the lack of robust process models and the heavy computational cost.

In (Zhang and Lennox, 2004), the MPC strategy was employed to regulate the biomass trajectory in a penicillin fed-batch process. The biomass concentration was estimated using a Multivariate Partial Least Squares method based on 20 previous batches.

(Kuprijanov et al., 2013b) applied an adaptive MPC scheme to reproduce an optimal biomass profile obtained from a golden batch. The method relied on off-line measurements of the biomass and substrate concentrations inputs. An application

to an industrial 10 L culture of *E. coli* allowed reaching a high cell density and a final biomass concentration over 55 g/kg.

A nonlinear MPC (NMPC) strategy was applied to a mammalian fed batch culture in (Craven et al., 2014). The control objective is to maintain the substrate concentration (glucose) at a defined setpoint. The controller was tuned in simulation and then implemented on a 15L bioreactor. The glucose concentration was successfully regulated at 11 mM with a setpoint change to 15 mM after 100 h.

In (Santos et al., 2012b) an NMPC strategy was presented to control fed-batch *E. coli* cultures. The cost function is defined to maximize glucose oxidation rate and minimize the glucose fermentation rate, in order to avoid growth inhibition due to overflow metabolism. The advantage of this approach is the ability to track an unknown optimal operating point corresponding to the critical substrate concentration.

(Kovárová-Kovar et al., 2000b) presented a combination of the MPC with an artificial neural network (ANN) applied to *B. subtilis* fed-batch fermentation. The ANN is applied to predict the product formation, due to the lack of a reliable mechanistic model. The cost function aimed at maximizing the total product quantity and the product yield. The results showed a 10% increase of the product yield compared to previous batches.

(Chang et al., 2016) applied the MPC strategy to *S. cerevisiae* cultures, with a goal of regulating the dissolved oxygen to a defined set-point using a dynamic flux balance model.

The advantages of the MPC schemes are the full exploitation of the process dynamics, the simple control policy for complex systems, the consideration of operational constraints, and the on-line optimization procedure. On the other hand they require rigorous model identification, and come with a high computation load.

5.3.4 Fuzzy control

Fuzzy control is a method based on fuzzy logic theory (Zadeh, 1994), designed to bypass the use of complex models and to deal with the uncertainty aspect of processes. This property made fuzzy control suitable for fermentation processes known for their non-linear dynamics (Lee et al., 1999). Unlike model-based methods, fuzzy logic controllers do not rely on the initial knowledge of the system dynamics, since the control parameters are calculated relying on the user experience to evaluate the current state of the process. Fuzzy control is based on converting the quantitative data into qualitative parameters. The method relies on the following definitions:

- Fuzzy set: A linguistic term defining the properties of a variable.
- Membership function: A value between 0 and 1 translating the degree to which a variable belongs to the fuzzy set.

In fuzzy control, numerical data are converted based on their degree of fitting in a fuzzy set using the membership function. This operation is called fuzzification. The ensemble of the fuzzy sets is then used to determine the current state of

the process, which allows the definition of a control action. A set of conditional rules (If/else/then) is defined based on the a-priori knowledge of the system. These rules are applied to the interpreted condition or state of the process, and used to calculate a numerical output that represents the controller output. This step is called the defuzzification (Seborg et al., 1989).

Fuzzy control methods can provide some insight into complex nonlinear system dynamics without the need for model identification (Babuška and Verbruggen, 1996). However, despite its benefits, there are some limitations to fuzzy control applications as it depends entirely on the user knowledge and expertise. The fuzzy sets are often defined on a set of operating conditions, and the controller performance is affected outside these conditions.

Several works considered the fuzzy logic for control problems in fed-batch cultures. In (Zhang et al., 1994) two fuzzy sets were used in a baker's yeast fermentation for different phases of the culture. In the first phase (alcohol production phase), a set of rules was used to regulate the tracking errors. In the second phase, the growth rate μ_X became the new regulated variable and was controlled to reach the optimal growth rate μ_{max}.

(Horiuchi and Hiraga, 1999) applied the fuzzy control to a large-scale recombinant protein production process. The controller was employed to identify the current process phase in a four operating phases process. A trapezoid membership function was assigned to each state variable, and the manipulated variables were the pH and the feed-rate. The total product increased up to 16% after the implementation of the fuzzy controller compared to a manual control method.

In (Hisbullah et al., 2003), a hybrid fuzzy and PI control strategy with scheduled gain was implemented in a baker's yeast cultures. The objective is to regulate the oxygen and carbon dioxide evolution rates to predefined set-points based on biomass measurements. The controller managed to reduce the oscillations and the set point tracking offset compared to conventional controllers.

Fuzzy logic-based controllers combine the user knowledge of the process and historical data in a relatively simple structure. Adaptation for different processes and scales is also a big advantage since the method relies on linguistic rules rather than mathematical representation. On the other hand, the definition of the rules is not straightforward and requires a deep knowledge of the process.

5.3.5 Artificial neural networks

Artificial Neural Network (ANN) is a data driven control method that can describe a complex non-linear system without the need for explicit equations. ANN relies on past process data to predict the outcomes of the process inputs. Consequently, ANN can be found in various industrial applications (Glassey et al., 1994; Lübbert and Simutis, 1994).

In ANNs, the input data are processed by weighted functions to predict the outputs. These outputs are defined by basic functions that may be sigmoidal or step functions (Lee et al., 1999; Stanke and Hitzmann, 2013). ANN training is

commonly done using the back propagation algorithm (Rumelhart et al., 1986), as shown by (Bošković and Narendra, 1995; Peng et al., 2013).

To utilize ANN in a control strategy, an optimal value of a chosen process variable must be set, and then an optimization algorithm (such as the genetic algorithm (Holland, 1984)) is used to optimize ANN outputs for the given data set and solve the optimization problem.

Controllers based on ANNs have been applied successfully to fed-batch fermentation applications. In some set point tracking applications, ANN is used to predict a variable of interest, and then feed it to a closed loop control algorithm (Ferreira et al., 2001). Alternatively, ANN can be used directly in solving optimization problems, and determine an optimal solution (Chen et al., 2004; Peng et al., 2013).

The main drawback of ANN controllers is that the network cannot be used to get some understanding of the process or the relationship between variables. Furthermore, the performance depends heavily on the operating conditions of the training data (Babuška and Verbruggen, 1996).

Few examples of ANN control of industrial fed-batch processes exist. (Chen et al., 2004) employed a cascade recurrent neural network for fed batch *S. cerevisiae* process to maximize the biomass concentration. The controller is composed of two network blocks where the feed-rate and the volume were used to estimate the dissolved oxygen in the first block. The second block uses the DO as an input to predict the biomass concentration. A genetic algorithm is used to optimize a smooth feed profile.

Another ANN implementation can be found in (Peng et al., 2013) based on data from 6 batches of *B. casei* fed-batch cultures. An optimal trajectory was created using a genetic algorithm. The strategy showed an improvement in the product synthesis over 25 batches.

(Ferreira et al., 2001) presented a multilayer ANN used to control the substrate concentration in fed-batch fermentation. The network was used to interpret the glucose and sucrose measurements. The ANN was coupled to a PI controller to regulate the substrate concentration to a predefined set-point of 10 g/L.

5.3.6 Probing control

Probing control is a control method where perturbations are applied to the process inputs, in order to get a response in the controlled variable and base the control decisions on this response. Probing control is employed when the regulation set-points and the reference trajectories are unknown.

In fed-batch fermentation, a perturbation is applied to the feed-rate and the response of a kinetic variable signal is analyzed to calculate the feed rate for the next sampling time. Examples of applications of probing control in fed-batch fermentation are given in (Henes and Sonnleitner, 2007; Johnsson et al., 2013; Velut et al., 2007).

Self-optimizing control and extremum seeking control are the main forms of probing control used in bioprocesses (Dochain et al., 2011). They are used to solve

optimization problems. Self-optimizing control is based on maintaining control parameters at a constant value in order to reach near-optimal operating conditions with acceptable loss. Extremum seeking control is employed to find an unknown optimal operating point that maximizes or minimizes a cost function. This strategy is appropriate to situations where the optimum is unknown, but the optimal condition is known.

In most applications, the determination of the optimal feed rate is difficult, therefore probing control is an interesting alternative since it does not refer to set-point knowledge, but it finds the optimal point and adapts to its changes. This advantage makes probing control suitable for fed-batch fermentations, where the objective is to maximize a certain variable, while avoiding feeding excess (Dewasme et al., 2011b; Henes and Sonnleitner, 2007; Johnsson et al., 2013). Another advantage is the adaptation to disturbance rejection and different initial conditions.

In (Johnsson et al., 2013), probing control was applied to fed-batch *B. licheniform* cultures. The goal was to maximize the oxidative metabolism, whilst avoiding overflow metabolism. The frequency response of the DO signal was analyzed to interpret the system state. The biomass concentration at 10 h was 24% greater than the reference batch.

(Henes and Sonnleitner, 2007) employed the probing control strategy to *E. coli*, *S. cerevisiae*, and *P. pastoris* fed-batch cultures. The relative change in DO was used to determine the controller actions. An exponential feeding profile was used as a reference, and an adaptive strategy was used to reach the optimal feeding profile by manipulating the specific growth rate μ in the used model.

In (Velut et al., 2007), probing control was applied to *E. coli* cultures to regulate and limit the Oxygen Uptake Rate (OUR). The probing controller was combined to a temperature control loop in cascade, and the feed-rate is adapted in proportion to the response, and the temperature is varied to maintain the oxygen concentration to the defined set-point.

The main advantage of probing control is the tracking of unknown and varying optimal conditions. This requires a knowledge of the optimum nature and on-line access to signals affecting it. These signals are not always available, which limits the control objectives handled by probing control.

5.3.7 Statistical control

Statistical control approaches are based on evaluating the current process state compared to past performances using statistical multivariate methods. Hence, past cultures data is used to form an empirical model (Nomikos and MacGregor, 1994), rather than relying on process knowledge. It is applied to monitor the process and identify the deviation from a given optimal state.

Correlation between variables can be identified by multivariate analysis (Olkin and Sampson, 2001), two main methods are used: principle component analysis (PCA) (Nomikos and MacGregor, 1994) and partial least squares (PLS) (Nomikos and MacGregor, 1995).

Despite the challenge of the development of a separate model for each process and scale, statistical process control is applied to a variety of fed-batch control problems. Examples of multivariate analysis methods application for process modeling and monitoring, and control can be found in (Albert and Kinley, 2001; Doan and Srinivasan, 2008; Duran-Villalobos et al., 2016a; Ferreira et al., 2007; Glassey, 2013; Mears et al., 2016).

In (Albert and Kinley, 2001), an industrial application of multivariate analysis to provide on-line process control of fed-batch tysoline cultures. The PCA is applied to 144 batches, and 65 high yield batches are found desirable.

In (Duran-Villalobos et al., 2016b), a PLS model was developed and applied to maximize the biomass formation in fed-batch cultures. The PLS was combined with an optimization algorithm to determine the optimal manipulated variables for future batches based on their initial conditions.

5.3.8 Discussion

In the previous sections, various control schemes implemented on fed-batch processes are presented. Each method has some advantages and drawbacks depending on several factors. A classification of these factors may provide an insight into choosing the appropriate method for the desired application.

The presented strategies can be classified into three main categories: model-based, historical data based, or user experience-based. Table 5.1 shows a summary of the different aspects of each control method.

Before choosing a control strategy for the fed-batch culture, the user must consider several aspects related to the studied bioprocess. The number of available measurements is a first deciding factor. These measurements should be accurate, reliable, with minimal noise in order to be applicable in a control or estimation loop.

The available data prior to the application of any control method is another deciding factor. If a reliable process model is available, it should describe accurately the studied process in order to provide a basis for eventual model-based control scheme. In the case of the lack of a model and the difficulty to identify a new set of model parameters, the user may refer to historical data from previous fermentations, and choose a control method that doesn't rely on the explicit description of the system.

Another aspect to analyze is the complexity of the control method. Ideally the control strategy should be as simple as possible, in terms of parameter tuning, software requirements, and hardware implementation. Development and running costs should also be evaluated before choosing a control strategy.

Methods such as fuzzy control and probing control provide some interesting advantages, as they do not require a process model, and a big amount of data from previous cultivations. They also handle unpredictable dynamics and process disturbances. However, they rely heavily on the process knowledge and the choice of control objectives is constrained to this aspect.

TABLE 5.1: Summary of the requirements and benefits of the presented control strategies (Mears et al., 2017).

		PF	Adaptive	MPC	ANN	Fuzzy	Probing	SPC
Requirements	Historical data sets				■			■
	Process Model		■	■				
	User experience	■				■	■	
Advantages	Handles nonlinear dynamics		■	■	■	■	■	■
	Adaptation to unpredictible dynamics		■	■	■	■	■	■
	Provides process insight			■			■	
	Flexible objective definition		■	■				

In the case where a process model is available, it is valuable to opt for a model-based control strategy due to flexibility, versatility, and the relatively low development time. Model-based methods (especially the MPC) can be applied for a wide range of control objectives. Predictive control covers most of the advantages of other methods, and the flexibility to define objective functions and the possibility to include operational constraints. However, the implementation difficulty and the computational cost are higher than other methods.

If a reliable process model exists, model-based approaches provide a vast number of possibilities, . However, their main drawback is the quality of the model in terms of uncertainty and sensitivity. Therefore, the model development time, the uncertainty analysis, and the robust design of the control strategy must be considered.

Another aspect of the control strategies is the gained insight of the process, since it's very valuable from an industrial point of view. Model-based methods, fuzzy control, and probing control provide information for the user about the process state and behavior under different conditions.

The development and implementation costs of the control strategies are one of the deciding factors of the industrial application of the control strategies. The cost/benefit ratio is a standard quantifier of the bioprocess performance. The costs include the model development, the industrial software licenses, and the specific program development (Lübbert and Bay Jorgensen, 2001). In terms of benefits of

considering a control approach, the increase of the product yield or the productivity is usually the main focus area. Other benefits are the reduced variance in the variables and the reduced variability between batches, in order to stay as close as possible to the optimal operating conditions (Moulton, 2014).

Furthermore, the implementation costs should be minimized, meaning that the chosen method complexity should be as low as possible. Methods that require extensive user training and operation and constant operator presence lead to additional working hours and thereby increasing running costs.

5.4 Control of fed-batch *E. coli* cultures

In the context of this thesis, we consider the control of fed-batch *E. coli* cultures. The main goal behind the control strategies is to favor the biomass production, reach high cell densities, and maximize the biomass productivity, while avoiding acetate accumulation.

As mentioned in the previous chapters, in order to maintain the process in the optimal conditions, the culture must be driven to the boundary between the respiro-fermentative and the respirative modes, where the substrate concentration is neither limiting nor in excess ($S = S_{crit}$ and $q_s = q_{scrit}$).

A straightforward approach to reach this optimal condition is to control the substrate concentration accurately at the critical level corresponding to the critical oxidative capacity of the cells. Another suboptimal approach is to regulate the byproduct (acetate) concentration at a low value to stay close to the metabolic edge and avoid the growth inhibition due to byproduct accumulation.

The main obstacles facing these approaches are the metabolic switch between the operating modes in the neighborhood of the critical substrate concentration, which may disturb the cells growth and cause uncontrolled fluctuations and latencies. Another major obstacle is the resolution and sensitivity of the available acetate and substrate measurement devices at low levels.

Indeed, the glucose critical level is very low ($O\,(10^{-2})$ gL^{-1}) and a small measurement error of this vital variable can increase the metabolic switches or cause the accumulation of acetate throughout the culture. On the other hand, the acetate concentration at the optimum is equal to zero, and regulating it at this level is impractical since the culture can operate in respirative mode at this concentration causing a drop in the biomass production yield. Thereby, robust soft sensors with high estimation accuracy are required to implement these approaches.

A reliable process model was developed in a previous study (Retamal et al., 2018) on the same process. The model parameters were identified based on a set of fed-batch experiments designed to trigger metabolic switches between the operating modes.

Model-based methods are therefore considered since they provide flexibility in defining control objectives. The chosen control method should achieve the desired objectives while remaining simple to develop and easy to implement on a lab-scale reactor setup. The method should also take into consideration modeling errors and uncertainties.

The measurement setup of the studied process includes traditional probes (pH, pO_2, temperature ...) in addition to the agitation and airflow measurements. The biomass concentration can be measured on-line using the turbidimetric probe. This measurement is reliable and therefore can be used in an output-feedback control scheme.

Taking into account all the previous points, a biomass regulation strategy based on the direct measurements of the biomass concentration is presented in the following chapter. The controller is based on feedback linearization with the addition of parameter adaptation. The control objective is to track a predefined biomass trajectory corresponding to a defined growth rate chosen lower than the maximal one in order to avoid acetate accumulation.

The biomass concentration measurement can also be used to implement a state estimation strategy, provided the availability of sufficient data for parameter tuning. The state estimation allows the implementation of state-feedback control strategies which increases the number of possible control objectives.

5.5 Conclusion

The main differences between the control methods are the requirements, complexity, process operation, and the control objectives. The goal of this presentation is to provide a basis and a guide to formulate the control problems depending on the available materials, methods, and objectives.

In the context of this thesis, we consider the control of fed-batch *E. coli* cultures. The control objective is to maximize biomass concentration and biomass productivity while avoiding acetate accumulation.

Several control strategies were developed for this bioprocess. However, most of the research studies are performed in simulation. The control strategies presented in the following chapters are tested and validated in simulations and on the lab-scale bioreactor.

The availability of a reliable process model and on-line biomass measurements allows for the consideration of an adaptive model-based biomass growth rate control strategy. The advantage being the easiness of implementation on the lab-scale reactor.

The next chapter discusses implementing an adaptive generic model control (GMC) strategy to regulate the biomass concentration tracking the desired biomass trajectory.

Chapter 6

Generic Model Control of the biomass concentration

6.1 Introduction

As presented in chapter 1, a major challenge in reaching high cell densities in fed-batch *E. coli* fermentation is the overflow metabolism and the accumulation of acetate during the culture (Han et al., 1992; Van De Walle and Shiloach, 1998). This metabolic phenomenon leads to the decrease of the biomass production yield and consequently the decrease of the recombinant protein production (Riesenberg et al., 1991; Rothen et al., 1998).

According to the bottleneck theory, in order to maximize the biomass productivity and avoid overflow metabolism, the substrate concentration must be maintained at a certain critical threshold corresponding to the critical cell oxidation

capacity (Jana and Deb, 2005). To achieve this objective, a closed-loop feeding strategy is required to maintain the process near optimal operating conditions. Several possible formulations of the control objective and strategy have been developed throughout the years, examples of control schemes developed for similar fed-batch processes presenting overflow metabolism can be found in (Benattia et al., 2015; Dewasme et al., 2011a,b; Gonzalez et al., 2016; Santos et al., 2012a).

A straightforward formulation of the control problem is to regulate the substrate or acetate concentrations at low values. Achieving this control objectives puts the process closer to the metabolic edge between the respirative and the respiro-fermentative modes, and allocates the majority of the available substrate to biomass production. Some application of this control approach can be found in (Dewasme et al., 2010; Hafidi et al., 2008; Renard et al., 2006; Rocha and Ferreira, 2002).

A major practical hurdle facing the experimental application of these methods is the lack of reliable on-line monitoring tools of acetate and glucose concentrations. The critical level of the glucose concentration in *E. coli* cultures is very low compared to the sensitivity of the available probes on the market. Furthermore, the strong nonlinearities exhibited by the process and the need for an accurate determination of the stoichiometry and the kinetics present additional obstacles in the implementation of these control strategies. Therefore, developing an adaptive control scheme with minimum reliance on stoichiometry can be an attractive alternative, especially if it does not require on-line measurements of substrate and acetate concentrations.

In this chapter, we propose an adaptive control strategy based on feedback linearization of the nonlinear model equations called Generic Model Control (GMC) (Lee and Sullivan, 1988). The objective is to take advantage of the on-line measurement of the biomass concentration to develop and implement a GMC algorithm to control biomass productivity during the fed-batch fermentation of recombinant *E. coli*. In this control strategy, a pre-defined biomass trajectory corresponding to a limited acetate production is imposed by the controller. The advantages of this approach are the inclusion of the process nonlinear model in the control design and the compensation of the model inaccuracies by online adaptation using a parameter estimator. In addition, the on-line integration of the process model (i.e. the numerical solution of the mass balance ODEs) is not required, which leads to an easy experimental implementation. An experimental implementation of the control strategy is carried on a lab-scale bioreactor in order to test its performance and robustness under real operating conditions.

6.2 Generic Model Control

Generic Model Control (GMC) was developed by Lee and Sullivan (Lee and Sullivan, 1988) and is a control strategy based on feedback linearization, embedding the process nonlinearities in the design of the control law.

GMC was used in various process control applications, among which we can cite for instance (DeLisa et al., 2001), where GMC ensured the tracking of the foreign protein level reference trajectory in *E. coli* fermentations, (Costello et al., 1989) where it was applied to anaerobic digestion, and (Douglas et al., 1994) where GMC was used to control dual product composition in an industrial high purity distillation column. More specifically, in relation with the present study, the GMC strategy was applied in (Jenzsch et al., 2006b) to control the specific biomass growth rate in recombinant *E. coli* fed-batch cultures to an experimentally defined specific growth rate reference $\mu_{ref}(t)$. A model simplification was carried out during the control design and a state estimation using the Extended Kalman Filter was carried out. In the same spirit, a control strategy was developed in (Bastin and Dochain, 1990) and applied in (Rocha, 2003), based on an adaptive linearizing control law, regulating the acetate concentration with the dilution rate as manipulated variable.

Hereafter, the GMC is applied to regulate the biomass concentration in fed-batch *E. coli* cultures, and track a predefined biomass trajectory corresponding to the desired metabolic performance.

6.2.1 GMC principle

Generic Model Control (GMC) is based on the feedback linearization of the nonlinear dynamics of the system. The main objective of the control scheme is to track a desired output nominal trajectory (Lee and Sullivan, 1988; Peter and Lee, 1993). Consider the following nonlinear process:

$$\dot{x} = f(x) + g(x)u \tag{6.1}$$
$$y = h(x) \tag{6.2}$$

where

- $x \in \mathbb{R}^n$ is the state vector
- $u \in \mathbb{R}$ is the manipulated input
- $y \in \mathbb{R}$ is the system output.
- $f : \mathbb{R}^n \to \mathbb{R}^n$ $g : \mathbb{R}^n \to \mathbb{R}^n$ are nonlinear functions of the states x,
- $h : \mathbb{R}^n \to \mathbb{R}$ is the output map.

From Equation (6.2), the output dynamics is given by (Isidori et al., 1995):

$$\dot{y} = \frac{\partial h}{\partial x}[f(x) + g(x)u] = L_f h(x) + L_g h(x)u \tag{6.3}$$

where

- $L_f h(x) = \frac{\partial h}{\partial x} f(x)$ is the Lie derivative of h along f.
-

In the GMC strategy, the performance of the system is measured using the time derivative of the output $\dot{y}$ and an arbitrary specification function (Lee and Sullivan, 1988):

$$\dot{y} = \hat{u} \tag{6.4}$$

where $\hat{u}$ is an arbitrary function used to specify the desired closed-loop performance.

In the GMC design procedure, the output y is compared against a prescribed reference trajectory y_{ref}. When the process output is away from its desired reference y_{ref}, we would like the process to return towards steady-state with a defined convergence rate. In addition, we would like to have zero offset. Hence, the specification signal $\hat{u}$ can then be defined using a proportional-integral controller in the form:

$$\dot{y} = \hat{u} = G_1(y_{ref} - y) + G_2 \int_0^t (y_{ref} - y)\partial\tau \tag{6.5}$$

where G_1 and G_2 are tuning gains (constant with respect to time), whose values can be chosen to achieve a variety of responses. Their tuning is performed according to the desired dynamic behavior as detailed later. If $L_g h(\mathrm{x}) \neq 0$ (i.e., the system is of relative degree 1), the control input satisfying equations (6.3) and (6.5) is derived from the following equation:

$$u = \frac{1}{L_g h}\left(-L_f h + \hat{u}\right) \tag{6.6}$$

The resulting closed-loop transfer function (Figure 6.1) is given by:

$$\frac{Y(s)}{Y_{ref}(s)} = \tag{6.7}$$

where

- $Y(s)$ and $Y_{ref}(s)$ are respectively the Laplace transforms of y and y_{ref}.
- s is the Laplace variable.

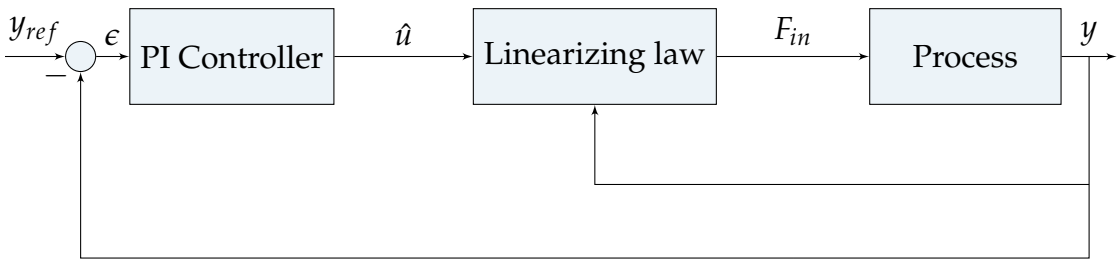


FIGURE 6.1: GMC structure

The desired closed-loop response is defined by setting the damping ratio ξ and the natural frequency ω_0, G_1 and G_2 are then chosen as follows:

$$\begin{cases} G_1 = & 2\xi\omega_0 \\ G_2 = & \omega_0^2 \end{cases} \tag{6.8}$$

Although the controlled system response is different from the classic second-order system due to uncertainties strong nonlinearities, similar plots of the closed-loop response for different values of ξ and ω_0 can be achieved. An example is given in Figures C.1 and C.2 in appendix C with multiple responses for the studied process. G_1 and G_2 can be tuned so as to confer the desired damping ratio ξ and a natural frequency ω_0 in the following steps:

- Choose ξ according to the desired response shape.
- Choose an appropriate rise time t_r and the corresponding $\omega_0 = \frac{3}{t_r}$.
- Calculate G_1 and G_2 using Equation (6.8).

The nonlinear closed-loop stability and the performance analysis of the GMC are detailed in (Zhou et al., 1992). Nominal stability is ensured for any positive values of G_1 and G_2. The proof is based on finding a strict Lyapunov function for the nominal process and applying a perturbation theorem. Another stability proof for a similar control structure with kinetic parameter estimation is given in (Gonzalez et al., 2016).

An anti-windup mechanism is also added to the integral term of $\hat{u}$ to avoid integration accumulation under input saturation:

$$\begin{aligned} \hat{u} &= G_1(X_{ref} - X) + G_2 \int_0^t [(X_{ref} - X) + u_\omega] d\tau \\ u_\omega &= K_\omega(u_{sat} - u) \end{aligned} \tag{6.9}$$

where K_ω is the anti-windup gain, . In the following, K_ω is determined by trial and error.

6.3 Application of the GMC scheme to *E. coli* Cultures

As presented in the previous chapters, our goal behind controlling fed-batch fermentation of *E. coli* is to favor the biomass production, reach high cell densities, and maximize the biomass productivity. Acetate accumulation and its growth inhibiting effect must be taken into consideration during the control design.

In this chapter, we propose regulating the biomass growth rate by tracking a predetermined suboptimal biomass trajectory. The idea is to reproduce a theoretical biomass profile reference satisfying the control objectives and maintaining the culture as close as possible to the optimal conditions. The advantage of this

approach is the low operating cost and practicality, since it relies only on the online biomass measurement which is provided by the turbidimetric probe with low measurement noise.

The reference biomass profile must satisfy the previously described control objectives. First, an exponential feeding reference $F_{in_{ref}}$ profile given in (Retamal et al., 2018) is applied to the system:

$$F_{in_{ref}} = \frac{\mu_{set}}{k_{XS}} \frac{X_0 V_0 e^{\mu_{set}\ t}}{S_{in}} \tag{6.10}$$

where μ_{set} is the setpoint specific growth rate, and k_{XS} is the yield coefficient defined as grams of produced biomass per grams of consumed substrate.

An important step of the control design is to choose the reference growth rate μ_{set}. We seek to track an exponential biomass reference trajectory corresponding to a constant growth rate, an initial substrate concentration lower than the critical value ($S^* < S_{crit}$), and an initial acetate concentration equal to zero. This operating trajectory allows the process to evolve close to the edge between the respirative and respiro-fermentative modes, with a safety net that avoids the metabolic switches and favor cell growth.

The reference growth rate μ_{set} should be carefully selected below the critical value $\mu_{set} \leq \mu_X^\star$, where $\mu_X^\star$ is the optimal growth rate. Figure 6.2 shows a simulation of the biomass productivity and acetate production for different values of μ_{set}. The biomass productivity is defined by the following equation:

$$P_X = \frac{1}{t_f} \frac{V\left(t_f\right) X\left(t_f\right) - V_0 X_0}{S_{in}\left(V\left(t_f\right) - V_0\right)} \tag{6.11}$$

Biomass productivity increases for increasing μ_{set} values until $\mu_{set} = 0.24\ \mathrm{h}^{-1}$. For the sake of security with respect to possible model uncertainties offsetting μ_X, μ_{set} will be selected in the range $[0.18, 0.22]$ in order to avoid acetate accumulation.

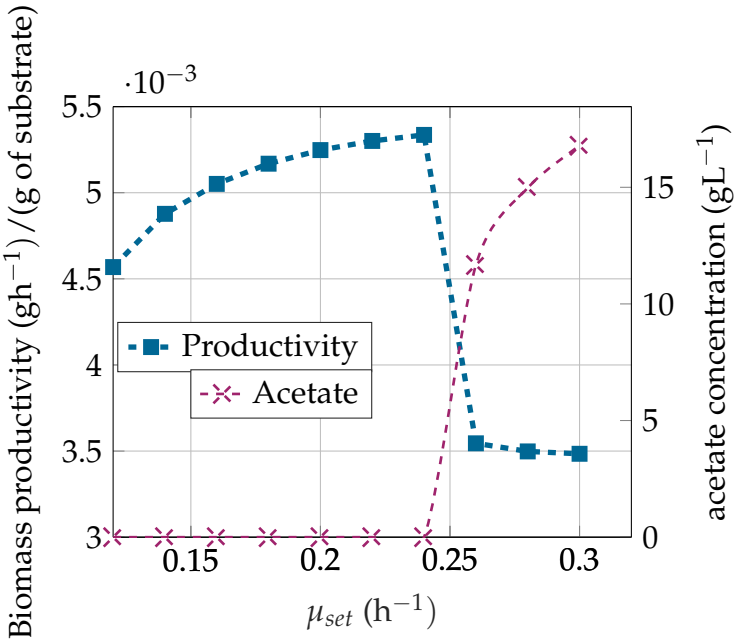


FIGURE 6.2: Biomass productivity (blue) and acetate production (magenta) for different μ_{set} values

GMC Design Using the Full-Order Model

The GMC scheme presented in the previous section is applied to the mechanistic model of the *E. coli* culture considering the biomass as the sole online measurement ($y = X$). Applying Equation (6.6), we obtain the following control law:

$$F_{in} = \frac{V}{X}\left((k_{X1}\mu_1 + k_{X2}\mu_2 + k_{X3}\mu_3)X - \hat{F}\right) \tag{6.12}$$

provided that $X \neq 0$. In this expression, $\hat{F}$ is given by the following:

$$\begin{aligned} \hat{F} &= G_1(X_{ref} - X) + G_2 \int_0^t (X_{ref} - X + F_\omega)d\tau \\ F_\omega &= K_\omega(F_{sat} - F_{in}) \end{aligned} \tag{6.13}$$

where K_ω is the anti-windup gain, and F_{sat} is the value of the saturation on the feed-rate F_{in}.

Unfortunately, the straightforward application of the GMC to the mechanistic model raises some problems. Accurate determination of the specific growth rates is difficult, since the kinetics are based on the overflow metabolism paradigm and are represented by metabolic switches. Moreover, an imposed biomass trajectory could possibly lead to high values of the flow-rate.

The biomass differential equation in respiro-fermentative mode is given by the following:

$$\dot{X} = \mu_X X - \frac{F_{in}}{V} X \tag{6.14}$$

where

$$\mu_X = (k_{X1} - k_{X2})q_{s_{crit}} + k_{X2}q_s \tag{6.15}$$

If the imposed biomass trajectory is exponential with a constant growth rate μ_{set}, the correspondent feeding profile is given by:

$$F_{in} = (\mu_X - \mu_{set})V|_{X=X_{set}} \tag{6.16}$$

The substrate dynamics is assumed to be fast and, therefore, in quasy-steady state ($S = S_{set}, \dot{S} = 0$) during the control period, implying the following:

$$\dot{S} = \mu_S X - \frac{F_{in}}{V}(S - S_{in}) = 0 \tag{6.17}$$

where

$$\mu_S = q_{s_{max}} \frac{S}{K_s + S}\bigg|_{S=S_{set}} \tag{6.18}$$

which leads to the following feed-rate profile:

$$F_{in} = \frac{\mu_S}{(S_{in} - S)} XV\bigg|_{S=S_{set}} \tag{6.19}$$

Assuming a low production of acetate, equations (6.16) and (6.19) show that the reference feeding trajectory depends on the pair (X_{set},S_{set}):

$$\frac{\mu_S}{(S_{in} - S)}\bigg|_{X_{set},S_{set}} = \left(\frac{\mu_X - \mu_{set}}{X}\right)\bigg|_{S_{set},X_{set}} \tag{6.20}$$

For each value of X_{set}, two possible solutions for S_{set} are obtained:

$$S_{set} = (\alpha_1 + \alpha_2 X_{set} \pm \sqrt{\alpha_3 + \alpha_4 X_{set} + \alpha_5 X_{set}^2})/\alpha_6 \tag{6.21}$$

The expressions of the α_j coefficients are given in appendix C. A numerical example with μ_{set}=0.18 h^{-1} and X_{set}=10 gL^{-1} is presented in Table 6.1. The first solution S_{set_1} is a low value of S, while the second solution S_{set_2} is rejected because it corresponds to a high value and physically non-achievable operating condition.

TABLE 6.1: α_j values and S_{set} solutions for μ_{set} = 0.18 h^{-1} and X_{set} = 10 gL^{-1}.

$\alpha_1 = -1.5602 \times 10^4$	$\alpha_2 = 164.0900$
$\alpha_3 = 2.4338 \times 10^8$	$\alpha_4 = -5.1201 \times 10^6$
$\alpha_5 = 2.6926 \times 10^4$	$\alpha_6 = -62.4058$
$S_{set_1} = 0.0294$ gL^{-1}	$S_{set_2} = 447.4088$ gL^{-1}

The GMC controller based on the mechanistic model can reproduce the correct

biomass profile corresponding to the reference growth rate μ_{set}. However, the system can converge to the undesired and unrealistic operating point corresponding to the high feed-rate. To avoid this behavior, a reduced model taking into account the low substrate concentration (during the control period) is presented hereafter.

GMC Design Using a Reduced Model

A control design based on a reduced model is developed by applying the singular perturbation technique (Rocha and Ferreira, 2002): the dynamics of substrate, oxygen, and carbon dioxide are considered much faster than the dynamics of biomass and acetate. Thus, the fast variables are considered to be in quasi steady-state and their dynamics are set to zero.
A fast-slow state partition is therefore proposed as follows:

$$\xi_f = \begin{pmatrix} S \\ O \\ C \end{pmatrix} \qquad \xi_s = \begin{pmatrix} X \\ A \end{pmatrix} \tag{6.22}$$

with indices f and s being respectively related to fast and slow dynamics. The dynamic system for the fast variables can be written from the system model:

$$\underbrace{\begin{pmatrix} \dot{S} \\ \dot{O} \\ \dot{C} \end{pmatrix}}_{\dot{\xi}_f} = \underbrace{\begin{pmatrix} -1 & -1 & 0 \\ -k_{O1} & -k_{O2} & -k_{O3} \\ k_{C1} & k_{C2} & k_{C3} \end{pmatrix}}_{K_f} \underbrace{\begin{pmatrix} \varphi_1 \\ \varphi_2 \\ \varphi_3 \end{pmatrix}}_{\varphi} - D \underbrace{\begin{pmatrix} S \\ O \\ C \end{pmatrix}}_{\xi_f} + \underbrace{\begin{pmatrix} DS_{in} \\ OTR \\ -CTR \end{pmatrix}}_{F_f+Q_f} \tag{6.23}$$

where

- D is the dilution factor ($D = \dfrac{F_{in}}{V}$).
- K_f is the stoichiometric matrix.
- φ is the reaction rate vector defined as $\varphi = \begin{pmatrix} \mu_1 X & \mu_2 X & \mu_3 X \end{pmatrix}^T$

Since the dynamics of these variables ($\dot{\xi}_f = 0$) as well as the dilution effects ($D\xi_f = 0$) are neglected, which are often several orders of magnitude smaller than the reaction terms, the following equation holds approximately:

$$K_f\, \varphi + F_f + Q_f = 0 \tag{6.24}$$

If K_f is full rank, Equation (6.24) can be rewritten as follows:

$$\varphi = K_f^{-1}(-Q_f - F_f) \tag{6.25}$$

The state space dynamics of the reduced model considering only slow variables is given by the following:

$$\underbrace{\begin{pmatrix}\dot{X}\\ \dot{A}\end{pmatrix}}_{\dot{\xi}_s} = \underbrace{\begin{pmatrix}k_{X1} & k_{X2} & k_{X3}\\ 0 & k_{A1} & -1\end{pmatrix}}_{K_s} \underbrace{\begin{pmatrix}\varphi_1\\ \varphi_2\\ \varphi_3\end{pmatrix}}_{\varphi} - \frac{F_{in}}{V}\underbrace{\begin{pmatrix}X\\ A\end{pmatrix}}_{\xi_s} \tag{6.26}$$

Replacing the expression of φ from (6.25) in (6.26) yields the following:

$$\dot{\xi}_s = \bar{K}(-Q_f - F_f) - \frac{F_{in}}{V}\xi_s \tag{6.27}$$

where

$$\bar{K} = K_s K_f^{-1} = \begin{pmatrix}\bar{k_{11}} & \bar{k_{12}} & \bar{k_{13}}\\ \bar{k_{21}} & \bar{k_{22}} & \bar{k_{23}}\end{pmatrix} \tag{6.28}$$

The following equation is for the biomass:

$$\dot{X} = -\bar{k_{11}}\,\frac{F_{in}}{V}\,S_{in} - \bar{k_{12}}\,OTR + \bar{k_{13}}\,CTR - \frac{F_{in}}{V}\,X \tag{6.29}$$

where the parameters $\bar{k_{11}}$, $\bar{k_{12}}$, and $\bar{k_{13}}$ are functions of the yield coefficients shown in Table 6.2.

TABLE 6.2: Theoritical dependency of $\bar{k_{ij}}$ parameters.

Parameter	Expression
$\bar{k_{11}}$	$\frac{1}{\bar{\delta}}(k_{X3}(k_{C1}k_{O2}) - k_{C2}k_{O1} - k_{X2}(k_{C1}k_{O3} - k_{C3}k_{O1}) + k_{C3}k_{O2}) + k_{X1}(k_{C2}k_{O3} - k_{c3}k_{O2})$
$\bar{k_{12}}$	$\frac{1}{\bar{\delta}}(k_{C3}k_{X1}) - (k_{C3}k_{X2}) - (k_{X3}(k_{C1} - k_{C2}))$
$\bar{k_{13}}$	$\frac{1}{\bar{\delta}}(k_{O3}k_{X1}) - (k_{O3}k_{X2}) - k_{X3}(k_{O1}k_{O2}))$
$\bar{\delta}$	$(k_{C1}k_{O3} - k_{C3}k_{O1} - k_{C2}k_{O3} + k_{C3}k_{O2})$

OTR and CTR can be calculated using the on-line measurements of the oxygen and carbon dioxide concentration using gas analyzers. The expressions of OTR and CTR under controlled conditions are given by (Retamal et al., 2018):

$$OTR = (k_{O1}\mu_1 + k_{O2}\mu_2 + k_{O3}\mu_3)X \tag{6.30}$$

$$CTR = (k_{C1}\mu_1 + k_{C2}\mu_2 + k_{C3}\mu_3)X \tag{6.31}$$

From equations (6.3), (6.13) and (6.29), the following control law is obtained:

$$F_{in} = \frac{-\bar{k_{12}}\,OTR + \bar{k_{13}}CTR - \hat{F}}{X + \bar{k_{11}}\,S_{in}}\,V \tag{6.32}$$

where $\hat{F}$ is still given by Equation (6.13) and assuming that $X + \bar{k_{11}}\,S_{in} \neq 0$, which is satisfied in general.

The advantage of the model reduction is that the desired operating condition (low substrate concentration) is directly embedded in the control algorithm.

Since OTR and CTR are not available for on-line measurement in our experimental setup and the biomass X is the sole measured variable, an adaptive GMC strategy is developed in the next section to reconstruct the unavailable signals, also adapting the control law subject to parameter uncertainty.

6.4 Adaptive GMC

The control law from Equation (6.32) contains the following unmeasurable and uncertain quantity: $-\bar{k_{12}}\ OTR + \bar{k_{13}}CTR$. Since the experimental setup is not equipped with gas analyzers, a kinetic parameter estimation scheme is developed and presented in this section. The biomass dynamics equation (6.29) can be rewritten as:

$$\dot{X} = \gamma\ - D\ (X + \bar{k_{11}}S_{in}) \tag{6.33}$$

where $D = \frac{F_{in}}{V}$, and γ is the uncertain and unmeasurable time-varying parameter given by:

$$\gamma = -\bar{k_{12}}\ OTR + \bar{k_{13}}CTR \tag{6.34}$$

We desire to model the dynamics of γ in order to apply a parameter estimation scheme. Let:

$$X^* = X + \bar{k_{11}}S_{in} \tag{6.35}$$

Hence, Equation (6.33) can be written in a compact form, assuming that S_{in} is constant:

$$\dot{X}^* = \gamma - D\ X^* \tag{6.36}$$

Provided $X^* = X + \bar{k_{11}}S_{in}$ is available for on-line measurement, γ can be estimated using a linear Kalman filter in the same way as presented in (Gonzalez et al., 2016), where the production rate of lactic acid is estimated in continuous mode.

Two estimation approaches are presented in the following, based on the classical discrete Kalman filter which offers simplicity and easy implementation.

6.4.1 Constant evolution of γ

First, γ is assumed to be constant, which is described by an exogeneous system:

$$\begin{cases} \dot{X}^* = & \gamma - DX^* \\ \dot{\gamma} = & 0 \end{cases} \tag{6.37}$$

As biomass measurements are collected at discrete time (sampling time T_s), it can be convenient to discretize Equation (6.37) using, for instance, the Euler scheme

(assuming a sufficiently small value of T_s):

$$\begin{bmatrix} X^*_{k+1} \\ \gamma_{k+1} \end{bmatrix} = \begin{bmatrix} -D_k T_s + 1 & T_s \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X^*_k \\ \gamma_k \end{bmatrix} + \begin{bmatrix} v_{1,k} \\ v_{2,k} \end{bmatrix} \tag{6.38}$$

$$X^*_k = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} X^*_k \\ \gamma_k \end{bmatrix} + w_k \tag{6.39}$$

where D_k is the dilution rate at the time instant t_k ($D_k = \frac{F_{in_k}}{V_k}$), assumed to be piecewise constant. v_k and w_k are respectively the process and measurement noises, assumed to be centered Gaussian white noises with covariance matrices Q_{const} and R_{const}, respectively. The covariance matrix of the estimation error on the state vector $[X_k \ \ \gamma_k]^T$ is denoted P_{const}.

6.4.2 Ramp evolution of γ

Considering a ramp evolution of γ, Equation (6.37) becomes:

$$\begin{cases} \dot{X}^* = & \gamma - DX^* \\ \Gamma = & \ddot{\gamma} = 0 \end{cases} \tag{6.40}$$

As in the previous case, the model is discretized using Euler scheme and additive noises are considered:

$$\begin{bmatrix} X^*_{k+1} \\ \gamma_{k+1} \\ \Gamma_{k+1} \end{bmatrix} = \begin{bmatrix} -D_k T_s + 1 & T_s & 0 \\ 0 & 1 & T_s \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X^*_k \\ \gamma_k \\ \Gamma_k \end{bmatrix} + \begin{bmatrix} v_{1,k} \\ v_{2,k} \\ v_{3,k} \end{bmatrix} \tag{6.41}$$

$$X^*_k = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} X^*_k \\ \gamma_k \\ \Gamma_k \end{bmatrix} + w_k \tag{6.42}$$

The covariance matrices of v_k and w_k, as well as the error on the state vector $[X_k \ \ \gamma_k \ \ \Gamma_k]^T$ are denoted $\mathbf{Q}_{Lin}$, $\mathbf{R}_{Lin}$, and $\mathbf{P}_{Lin}$ respectively.

6.4.3 Kalman filtering

After modeling the dynamics of γ, a discrete Kalman filter (Welch and Bishop, 1995) can be applied to estimate the evolution of X^* and γ . The updated control structure is summarized in Figure 6.3.

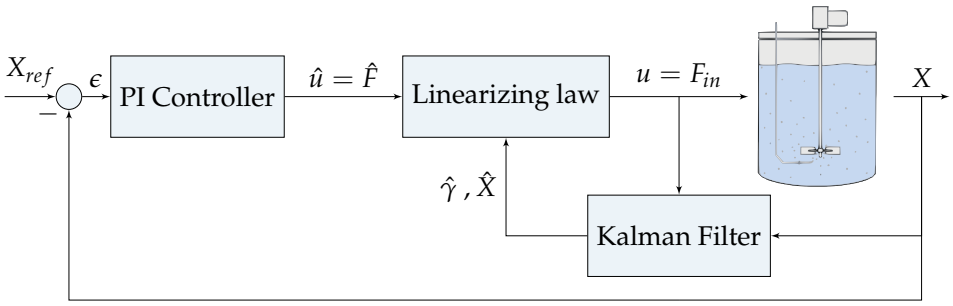


FIGURE 6.3: GMC combined with the Kalman filter

The updated control law, after including the estimated parameter $\hat{\gamma}$ becomes:

$$F_{in} = \frac{\hat{\gamma} - \hat{F}}{X + k_{11}^{-} S_{in}} V \tag{6.43}$$

The performances of both exogenous models will be compared in the next section.

6.5 Numerical simulations

This section highlights the control scheme performance using numerical simulations, based on the mechanistic model presented in chapter 2. The initial conditions and the different control and estimation parameters are given in Table 6.3.

TABLE 6.3: Control and estimation parameters

Initial conditions	$\begin{pmatrix} X_0 \\ S_0 \\ A_0 \\ V_0 \end{pmatrix} = \begin{pmatrix} 1.42\ \text{gL}^{-1} \\ 0.5\ \text{gL}^{-1} \\ 0.5\ \text{gL}^{-1} \\ 3.15\ L \end{pmatrix}$
Control parameters	$T_s = 0.05\ h$ G_1= 6, G_2= 9, $k_\omega = 0.2$ $\xi = 1$, $t_r = 1\ h$ $\omega_0 = \frac{3}{t_r} = 3\ rad/h$
Reference characteristics	$F_{in_{ref}} = \frac{\mu_{set}}{k_{XS}} \frac{V_0 X_0 e^{\mu_{set} t}}{S_{in}}$ $\mu_{set} = 0.18\ h^{-1}$, $k_{XS} = 0.2$
Estimator parameters	$\mathbf{P}_{Lin} = diag([0.1\ 1\ 1])$ $\mathbf{Q}_{Lin} = diag([0.01\ 0.1\ 0.1])$ $\mathbf{P}_{Const} = diag([0.1\ 1])$ $\mathbf{Q}_{Const} = diag([0.01\ 0.1])$ $\mathbf{R}_{Const} = R_{Lin} = 0.1$

6.5.1 Parameter estimation

First, the Kalman filter is tested with both the constant and ramp exogenous models for γ. Figure 6.4 shows the evolution of the variable γ constructed with the computation of OTR and CTR, and its estimate using both the constant and ramp exogenous models. The estimator convergence is achieved in 20 mn after a transient phase. Both exogenous models yield good performance of the Kalman filter. However, the ramp model presents a better performance regarding the initialization error as shown in Table 6.4. In the following (and in the experiments), the ramp model is applied to estimate the kinetic parameter γ.

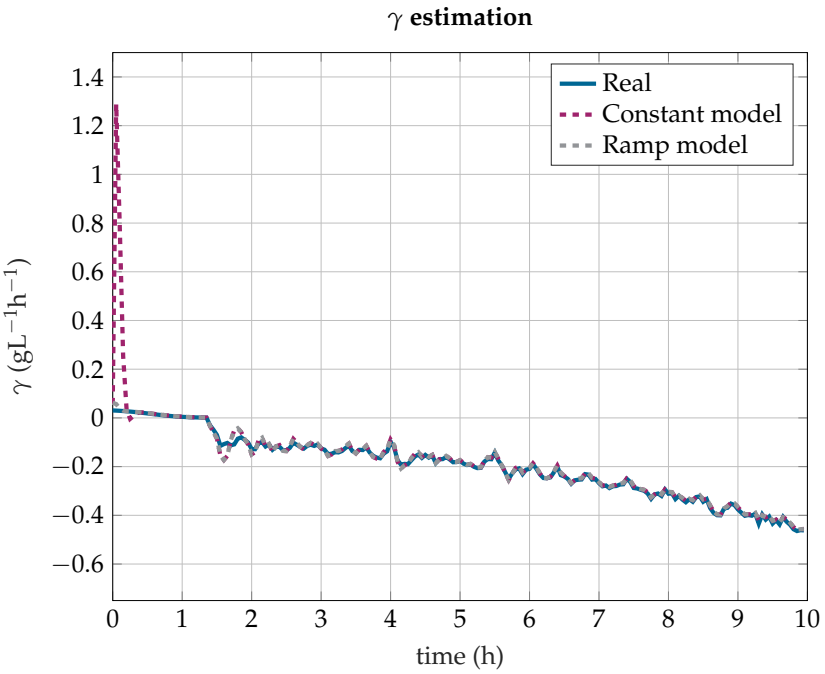


FIGURE 6.4: $\hat{\gamma}$ estimation based on biomass measurement using both constant and ramp exogenous models

TABLE 6.4: Root Mean Square error (RMSE) comparison between the constant and ramp exogenous models

Model	Constant	Ramp
RMSE ($gL^{-1}h^{-1}$)	0.0139	0.0030

6.5.2 GMC performance

Next, the control strategy is tested assuming that the variables are directly measurable (no parameter adaptation). Figures 6.5 to 6.7 show the simulation results of the GMC strategy based on the nominal and reduced models. The initial conditions and control parameters are the same in both scenarios, and all the state variables are assumed to be measurable. In both cases, the biomass is regulated

and the exponential reference trajectory is tracked after 0.8 h, the convergence time is identical for the two models as can be seen in Figure 6.6. The offset due to initialization and model mismatch is forced to zero by the integral term in the GMC control formulation.

However, in the case of the controller based on the full-order model, simulation shows that the feed-rate reaches unrealistic high values, as well as the substrate concentration which gets close to S_{in}. Indeed, as established in the previous sections, for every set biomass value X_{set}, two possible substrate concentrations can be obtained from the model, and only one corresponds to realistic growth conditions.

Conversely, in the case of the control law derived from the reduced model formulation, the system converges to the desired trajectory corresponding to low substrate concentration. The substrate and the acetate concentrations are in the expected practical range, and the calculated flow rate is more realistic.

While the full-order system may indeed converge to both substrate solutions, the reduced model, established under the assumption of low dilution rates, i.e. $D\xi_f = 0$ and therefore F_{in} will always converge to the expected realistic trajectory.

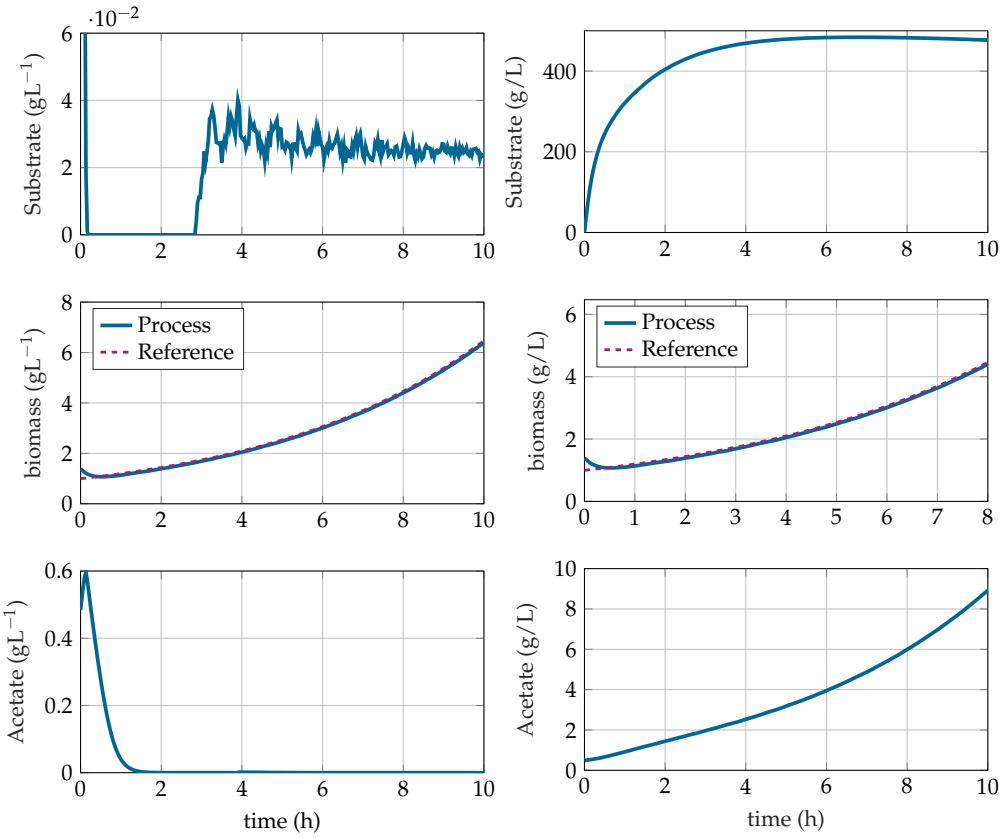


FIGURE 6.5: State variable evolutions with the full-order model (Right) and the reduced model (Left)

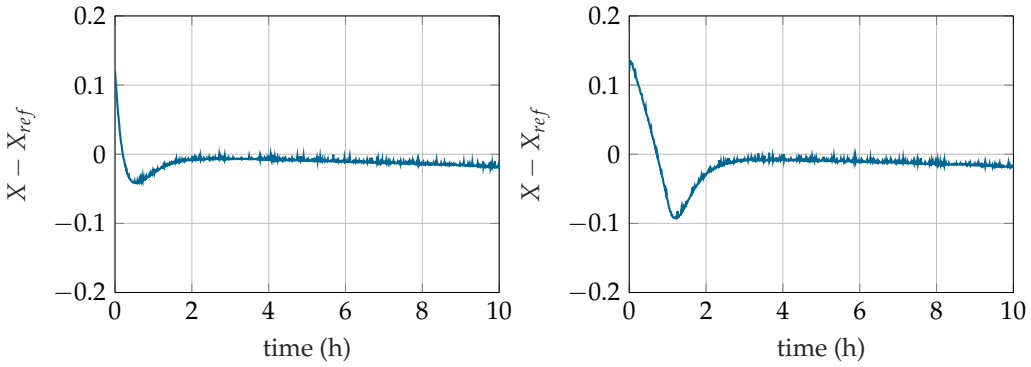


FIGURE 6.6: Biomass tracking error with the full-order model (Right) and the reduced model (Left)

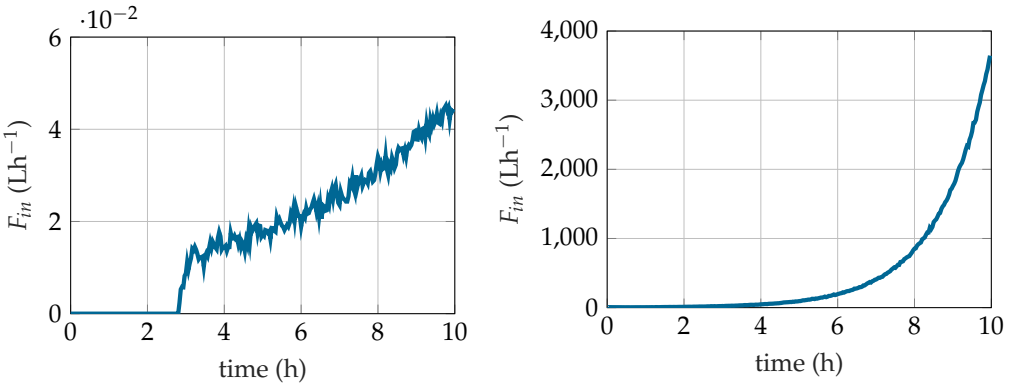


FIGURE 6.7: Flow rate evolution: full-order model (Right) and reduced model (Left)

6.5.3 Robustness of the control scheme

The GMC controller and the Kalman filter are now combined. To test further the robustness of the proposed controller, a sequence of Monte Carlo simulations is achieved, with 15% relative uncertainty on the plant parameters following a normal distribution, with the same control setting as the previous simulations. 500 Monte Carlo (MC) simulations were performed, and the histogram of the parameter $\bar{k}_{11}$ during the MC runs is shown in Figure 6.8. The biomass, substrate, acetate, and flow rate time evolutions are presented in Figure 6.9.

In all the runs, the corresponding reference substrate concentration S_{set} is reached, and the acetate concentration is equal to zero at the end of the culture, ending in respiratory mode. However, parameter variations imply a distribution of the final biomass concentration. Nevertheless, the goal of reaching high biomass concentrations while keeping the culture in good operating conditions is achieved in all experiments.

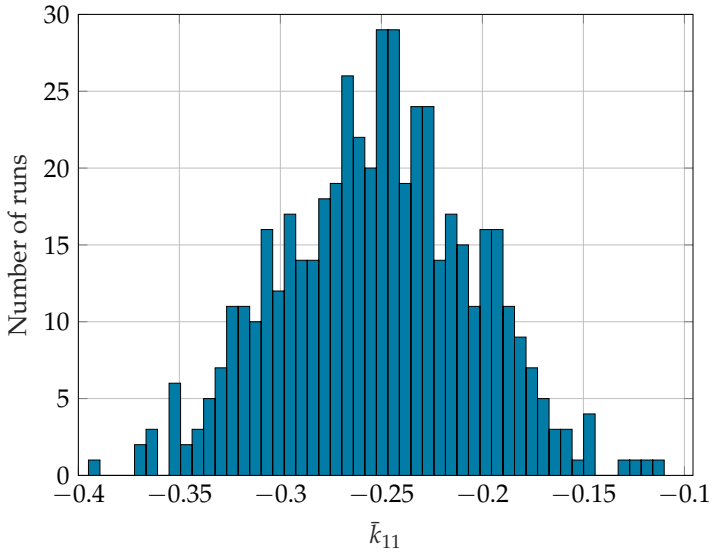


FIGURE 6.8: Histogram of the parameter $\bar{k}_{11}$ during 500 Monte Carlo runs

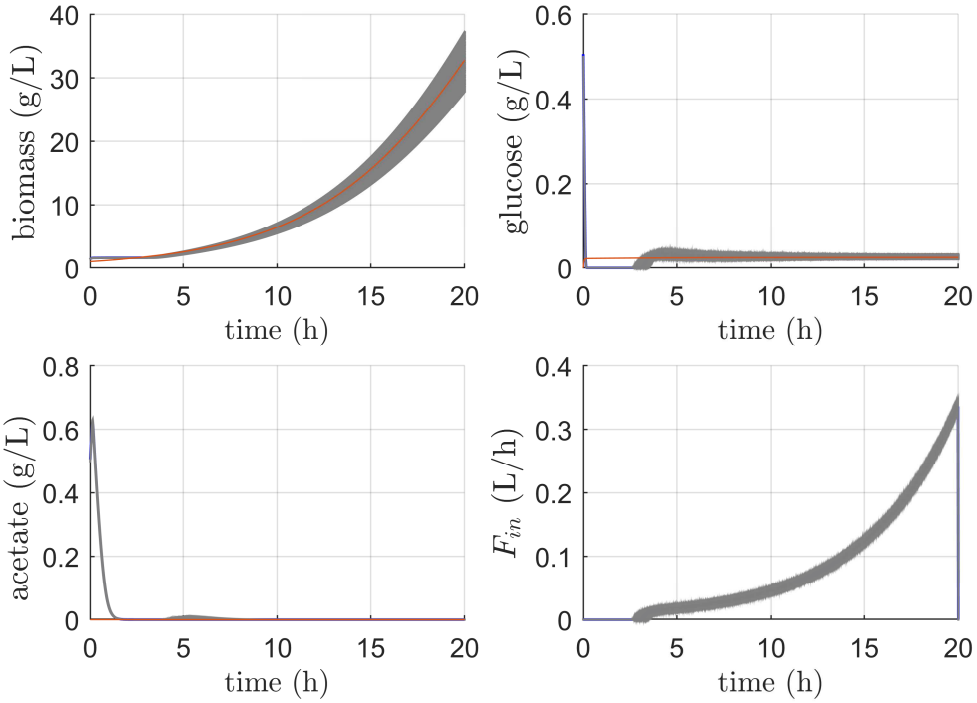


FIGURE 6.9: State variables and feed-rate evolution during 500 Monte Carlo runs

TABLE 6.5: Root mean square tracking errors for different values of $\bar{k}_{11}$

Deviation	Value	$min(\overline{e_X})$	$mean(\overline{e_X})$	$max(\overline{e_X})$
-30%	-0.1759	0.1444	0.0994	0.1591
-20%	-0.201	0.1456	0.1148	0.1661
-10%	-0.2261	0.1463	0.1337	0.1569
nominal	-0.2512	0.1481	0.1365	0.1636
+10%	-0.2764	0.1515	0.1381	0.1658
+20%	-0.3015	0.1533	0.139	0.1804
+30%	-0.3266	0.1575	0.144	0.2179

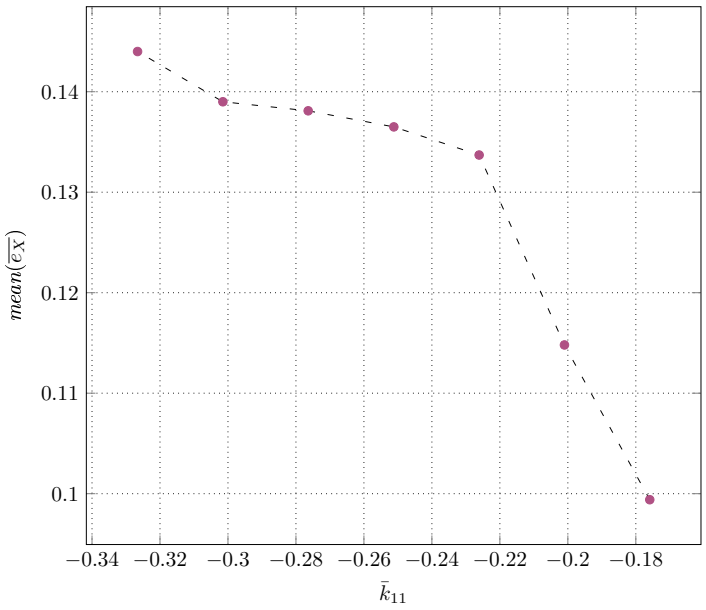


FIGURE 6.10: mean root mean square tracking errors for different values of $\bar{k}_{11}$

Comparably to the set-point change illustrated in Figure 6.12, the same simulation is achieved with a trajectory change from μ_{set}= 0.18 h^{-1} to μ_{set}= 0.22 h^{-1} at t= 3 h, before returning to μ_{set} = 0.18 h^{-1} at t= 6 h. Obviously, the controller is able to handle all of these types of disturbances, providing a fast and robust behavior.

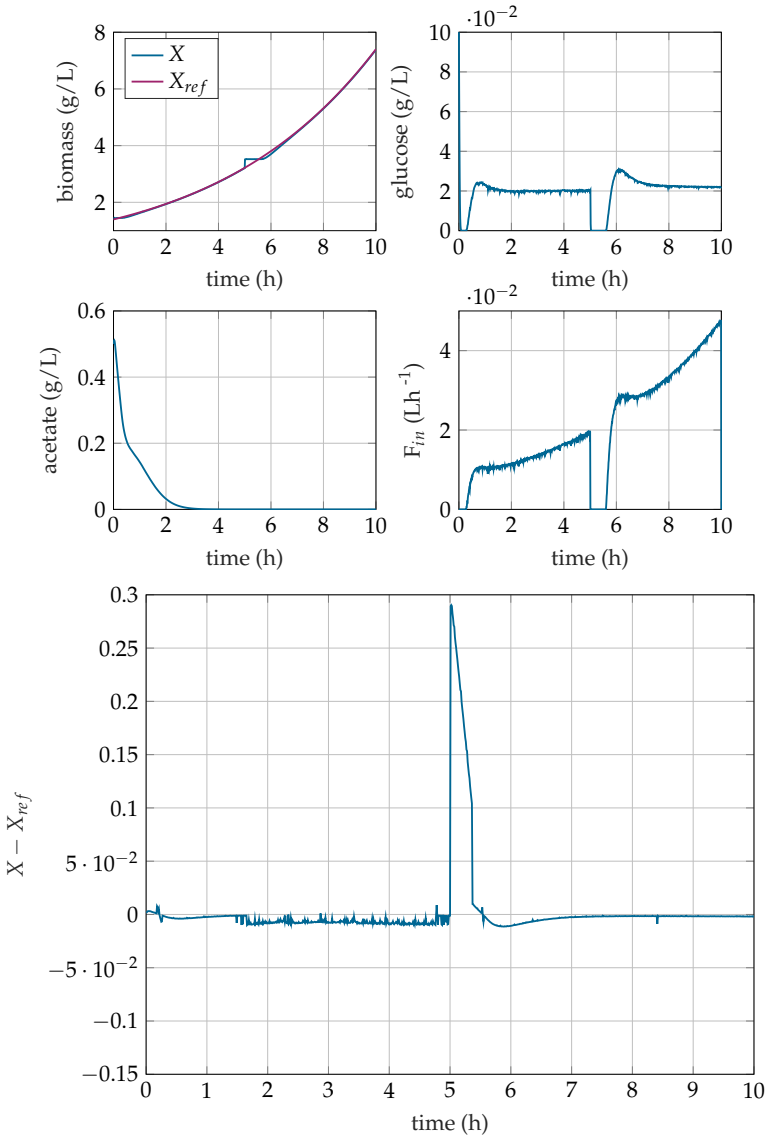


FIGURE 6.11: Closed loop response to a disturbance on the biomass signal at $t = 5\,h$

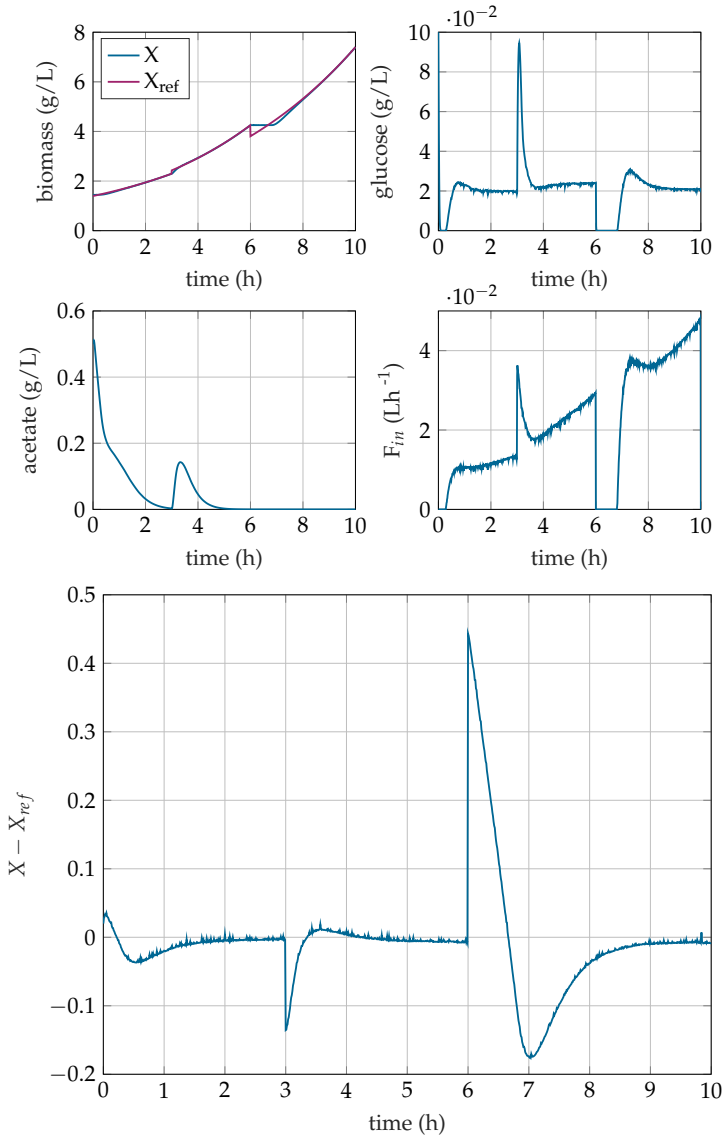


FIGURE 6.12: Closed-loop response to a set-point change μ_{set}= 0.18 h^{-1} and μ_{set}= 0.22 h^{-1}

6.5.4 Comparison with classical control strategies

In order to compare the performance of the proposed GMC strategy with other regular control schemes, a simulation is performed with a GMC controller, a first order linearizing controller (FOC) (Bastin and Dochain, 1990; Rocha, 2003) and a PID controller. A reference profile is imposed with μ_{set} = 0.18 h^{-1}, and the process parameter are considerer to have a 15 % variation around their nominal values. The first order linearizing controller has the same structure as the GMC but without the integral term as in:

$$\hat{F} = G_1 \left(X_{ref} - X \right) \tag{6.44}$$

The PID controller has the following transfer function:

$$C(s) = K_p + K_i \frac{1}{s} + K_d \frac{s}{1 + T_f s} \tag{6.45}$$

The PID parameters are chosen to have an overall dynamic behaviour close to the one obtained by the GMC, the parameters values are: $K_p = 0.07\ \mathrm{L^2 g^{-1} h^{-1}}$, $K_i = 0.01\ \mathrm{L^2 g^{-1} h^{-2}}$, $K_d = 0.01\ \mathrm{L^2 g^{-1}}$, $T_f = 0.5$ h. The first order controller gain G_1 is chosen equal to the GMC proportional gain.

It is noteworthy to point out that in the case of the PID controller, a metabolic switch from the respirative to the respiro-fermentative mode occurred at t=13 h, leading to acetate formation due to substrate excess ($S > S_{crit}$), and thus the control output (feed-rate) strayed from its initial exponential curve.

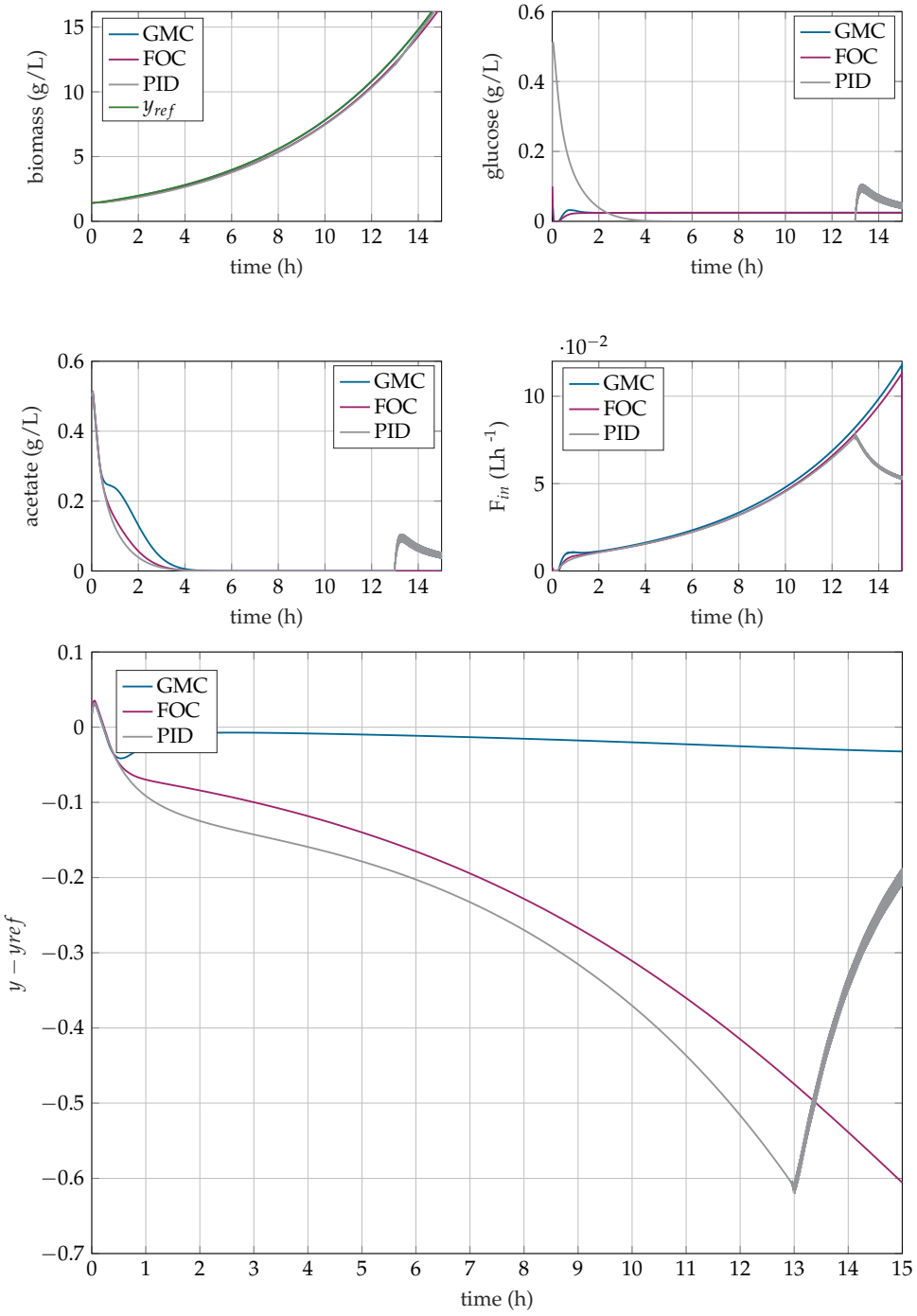


FIGURE 6.13: Comparison of the GMC performance with a first order controller (FOC) and a PID controller

6.6 Experimental results

Two fed-batch experiments were performed to challenge the controller under real experimental conditions. The control and estimation parameters are given in Table 7.4. The control parameters G_1 and G_2 were tuned in simulation, the chosen

response time t_r is equal to 1 h ($\omega_0 = 3rad/h$), and the damping ratio is fixed at $\xi = 1$. As far as the parameter γ estimation, the ramp model was used during the experiments.

Biomass, glucose, acetate concentrations, as well as the feed flow rate are shown in Figures 6.14 and 6.16. Operating conditions are also shown in Figures 6.15 and 6.17.

TABLE 6.6: Control and estimation parameters used in the experiments

Sampling time	$T_s = 0.05$ h
GMC parameters	G_1= 6, G_2= 9 $\xi = 1$, $t_r = 1\ h$, $\omega_0 = 3\ rad/h$
Reference	$F_{in_{ref}} = \frac{\mu_{set}}{k_{XS}} \frac{V_0 X_0 e^{\mu_{set} t}}{S_{in}}$ $\mu_{set} = 0.18\ \text{h}^{-1}$, $k_{XS} = 0.22$
Estimator covariance matrices	$\mathbf{P}_{Lin} = diag([0.1\ 1\ 1])$ $\mathbf{Q}_{Lin} = diag([0.01\ 0.1\ 0.1])$ $\mathbf{R}_{Lin} = 0.1$

During the fermentations, the initial biomass concentration ranged from 0.1-0.3 gL^{-1} and reached 1.5-1.7 gL^{-1} by the end of the batch phase, characterized by glucose depletion. The on-line flag for the end of the batch phase is the sudden increase of the pH, and consequently the decrease of base addition, as can be seen in Figures 6.15 and 6.17. The fed-batch phase starts around 6-8 h of culture time, and the control algorithm is launched. During this period, the RPM increases due to the important glucose oxidation inducing an increasing cell demand for oxygen. Base is added to compensate the pH decrease.

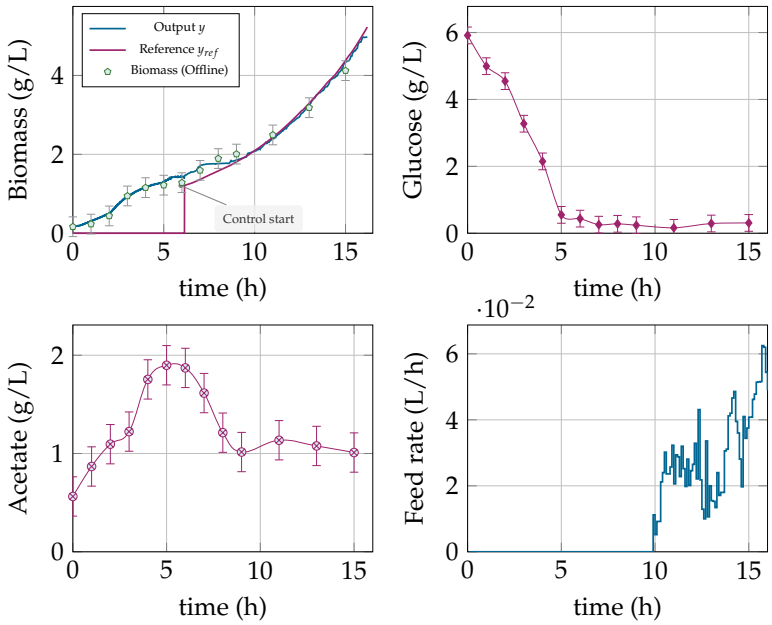


FIGURE 6.14: Experiment 1: Time evolution of the measured biomass, reference profile, glucose, acetate concentrations, and feed-rate

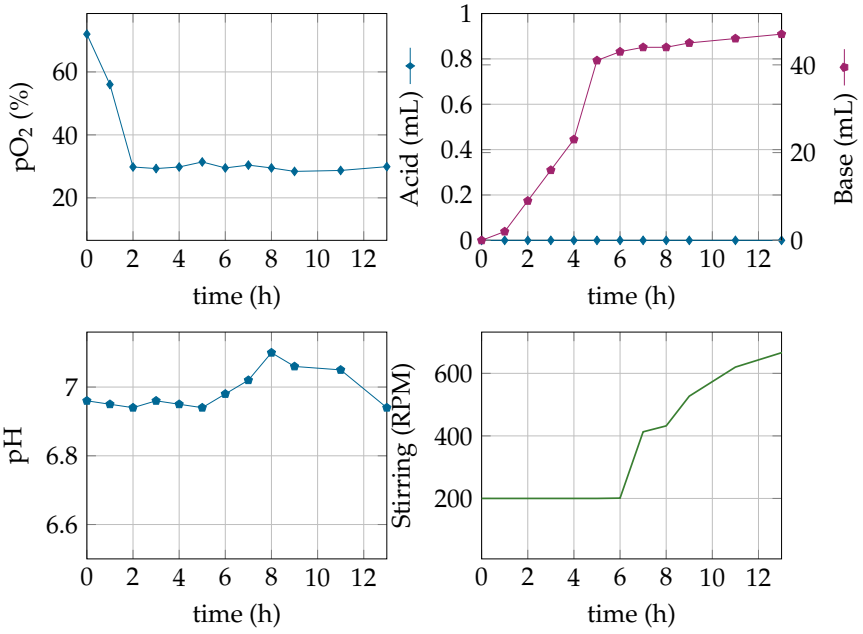


FIGURE 6.15: Experiment 1: Time evolution of pO_2, acid and base concentrations, pH, and Stirring

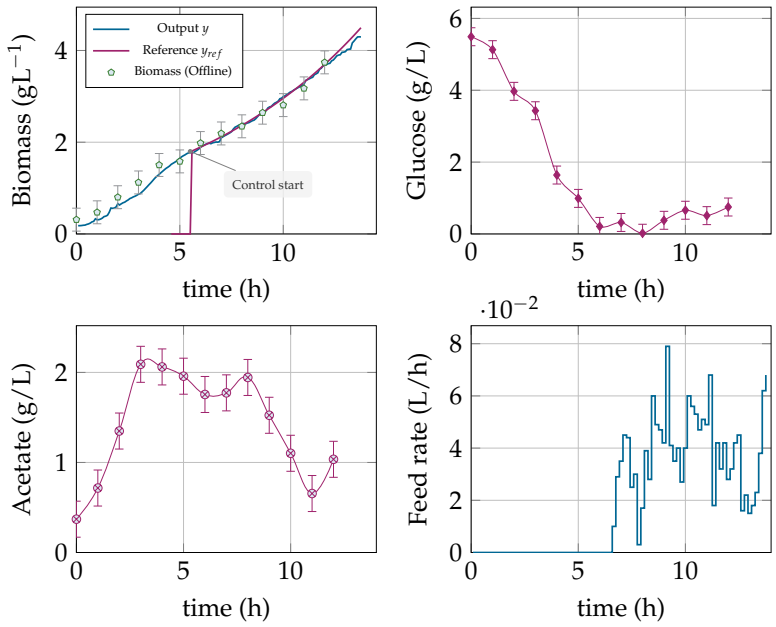


FIGURE 6.16: Experiment 2: Time evolution of the measured biomass, reference profile, glucose, acetate concentrations, and feed-rate

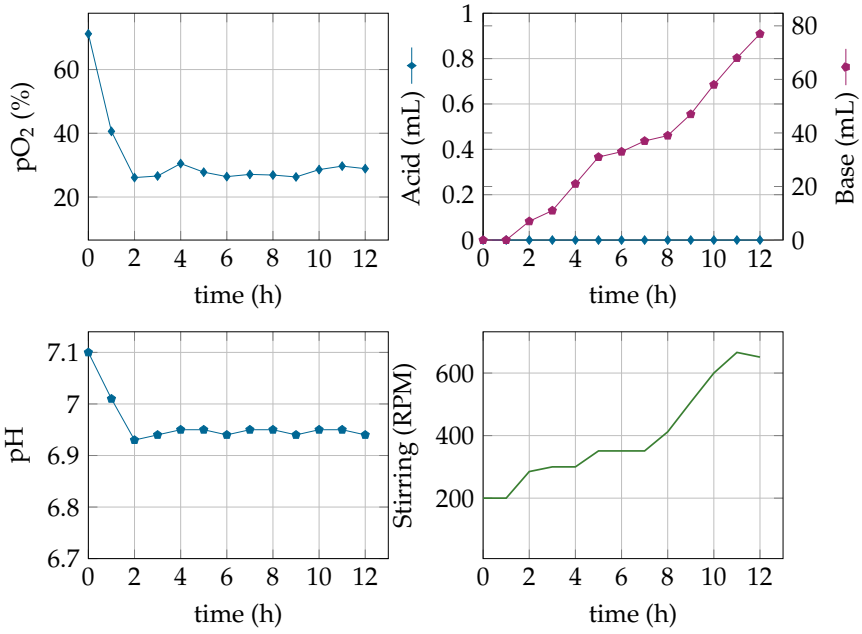


FIGURE 6.17: Experiment 2: Time evolution of pO_2, acid and base concentrations, pH, and Stirring

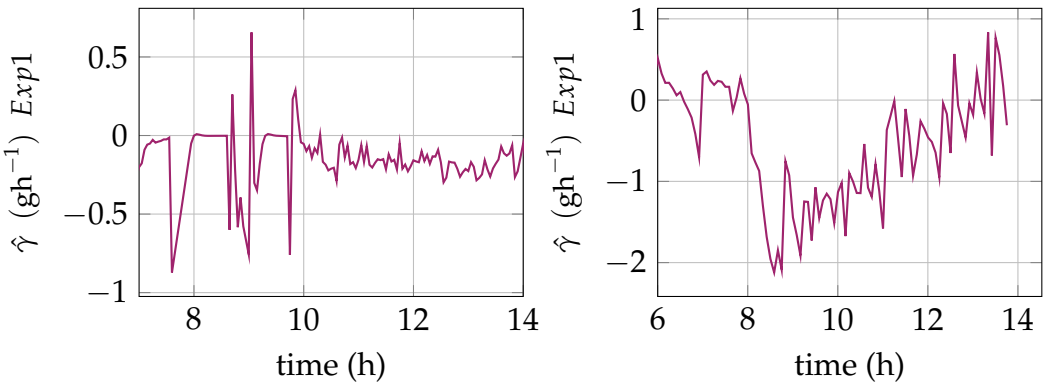


FIGURE 6.18: $\hat{\gamma}$ estimation during experiment 1 (left) and experiment 2 (right)

The imposed reference trajectory is $\mu_{set} = 0.18$ h $^{-1}$ and the algorithm is launched with an initial biomass concentration lower than the measured one to avoid excessive feeding at the start of the fed-batch. In Figures 6.14 and 6.16, the biomass maintains an exponential growth close to the reference trajectory showing that the regulation is efficient. The glucose concentration remains close to zero and almost constant during the fed-batch phase of both experiments, confirming the fast dynamics assumption presented in the model reduction section ($\dot{S} = 0$).

The acetate concentration remains under 2 gL^{-1} during both experiments. The evolution indicates a metabolic switch between respirative and respiro-fermentative modes, as can be seen in Figure 6.14 where the acetate concentration increases from 0.5 to 1.7 gL^{-1} due to overflow, and when the glucose is nearly depleted at t=4h, the acetate concentration starts to decrease, i.e the culture is in respirative mode. This behavior is reproduced during experiment 2 (Figure 6.16).

The Generic Model Controller performance, in terms of robustness, is quite satisfactory. The controller is able to maintain the biomass tracking error close to zero in both experiments despite the model mismatch resulting from modeling uncertainties and the use of basic minimal growth media (M9). The controller manages to adapt to the variations in the biomass signal by acting on the feed flow rate. We can see in experiment 1 (Figure 6.14) that the flow rate follows an exponential rate but is, however heavily distorted, due to the noisy biomass measurements. We can see in Figure 6.18 that measurement noise indeed affects γ estimation and, in turn, the calculated controller input. In the following experiments, a low pass filter is applied to the turbidimetric probe to reduce the measurement noise caused by airflow injection, leading to smoother profiles.

6.7 Conclusion

In this chapter, we presented a control strategy to regulate the biomass concentration in fed-batch *E. coli* BL21 (DE3) cultures. The considered control strategy is the Generic Model Control, based on input-output linearization of the nonlinear output equation and a proportional-integral control law.

Model reduction is applied to the mechanistic model presented in chapter 2, in order to obtain a control law which is independent of the kinetic terms (specific growth rates), and to avoid high feeding rates. A parameter estimation is coupled with the controller to predict the unmeasured kinetic terms in the control law, and adapt to model uncertainties and unpredicted dynamics.

Numerical simulations are carried out to test and tune the control and estimation algorithms. Results show that the accuracy and robustness of the proposed control strategy are quite satisfactory. The adaptation law was able to reconstruct the unmeasured dynamics, however, the closed-loop system is vulnerable to model mismatch and strong disturbances. A robust control design is required to take into account the effect of parametric uncertainties.

Fed-batch experiments of a BL21(DE3) *E. coli* strain are achieved with a lab-scale bioreactor and results show that the biomass concentration profile correctly tracks the given reference, and the controller manages to keep the culture in suitable growth conditions.

The experiments on the real lab-scale process showed that the combination of the Generic Model Control and the parameter adaptation is able to achieve the control objectives under real-time conditions. The biomass trajectory tracking is performed adequately despite the presence of disturbances and model mismatch.

The combined GMC-Kalman strategy has several interesting features. First, the development and design cost are fairly minimal, making its implementation on complex or old experimental setups simple and effective. The integration of the control scheme on an existing process does not require heavy modifications of the hardware and software configurations. In the present setup, only on-line biomass measurements are used, but the availability of a gas analyzer could provide missing information on OTR and CTR leading to a simplification of the control law. The proposed strategy can also be used to reproduce "golden batches" while adapting to disturbances due to the change in culture conditions.

Another advantage is the fact that the control law is calculated through simple algebraic equations, and doesn't require real-time solving of complex nonlinear differential equations. This feature lowers the computation complexity of the control scheme, and makes it easily integrable in most monitoring hardware.

The availability of the biomass measurements makes the control scheme very practical. The estimation of the state variables is not required since the measured variable is the controlled variable, and the parameter estimation is performed using a linear Kalman filter.

Finally, the addition of on-line kinetic parameter adaptation strengthens the robustness of the closed-loop system towards unpredictable dynamics.

However, despite its interesting features, the control strategy has some limitations concerning the metabolic performance. The controller aims at regulating the biomass concentration with a defined growth rate lower than the critical one to avoid overflow metabolism. According to the bottleneck theory, this suboptimal targeted growth rate can correspond to either the respirative or respiro-fermentative regimes. However, the observed behavior is that the cells operate mainly in respirative regime.

While the acetate accumulation is avoided by operating in this mode, deviations from the reference growth rate (due to strong model mismatch, parameter variation due to oxygenation conditions, or strong disturbances on the biomass measurements) can lead to a drop in the biomass production yield and biomass productivity compared to the expected theoretical values. Furthermore, this deviation can also cause an accumulation of the acetate if the culture switches to the respiro-fermentative mode without any informative on-line indication for the user on the biomass signal.

A practical solution to this problem is to regulate the acetate concentration at a low value, as it is directly linked to the growth rate. This approach requires robust state estimation, raising the complexity of the control strategy, but on the other hand improving the process productivity. The next chapter presents a robust GMC control strategy to regulate the acetate concentration at a defined low value.

Chapter 7

Robust Generic Model Control of the acetate concentration

7.1 Introduction

Acetate accumulation in fed-batch *Escherichia coli* cultivations affects the bioprocess efficiency and the biomass productivity (Han et al., 1992; Luli and Strohl, 1990; Van De Walle and Shiloach, 1998). Acetate presence in high concentrations causes the inhibition of the cell respiratory capacity, leading to the decrease of biomass production yield and consequently the decrease of the recombinant protein production (Riesenberg et al., 1991; Rothen et al., 1998).

Operating the process with a sufficiently low feed rate is a straightforward solution to avoid acetate accumulation. However, it does not allow reaching the full potential of the bioprocess since it leads to low productivity and high cultivation

time. It is therefore required to determine a closed-loop feeding strategy that maximizes biomass productivity while avoiding overflow metabolism (Srinivasan et al., 2001).

Several optimization schemes and process control architectures have been developed to reduce or avoid overflow metabolism (Dewasme et al., 2011a,b; Gonzalez et al., 2016; Hulhoven et al., 2006b; Jana and Deb, 2005; Rocha, 2003; Santos et al., 2012a; Valentinotti et al., 2003).

Two main control approaches can be categorized. The first approach, presented in the previous chapter, consists of regulating the specific biomass growth rate (Abadli et al., 2020; Jenzsch et al., 2006b) and imposing a reference biomass evolution profile. This type of control is made possible by the availability of reliable on-line biomass probes which allows convenient real-time implementation. However, the definition of a biomass reference profile is not straightforward as it relies on prior process knowledge (i.e., a growth model based on past experimental observations), and in practice, a suboptimal solution is often selected by targeting a reference slightly lower than the maximal growth rate, in order to ensure sufficient margin of security.

As seen in the previous chapter, limiting the specific growth rate presents some practical and metabolic limitations since its maximal level depends on the cells' oxidative capacity, which is by essence, uncertain. Therefore, targeting a growth rate close to its maximal value could lead to several uncontrolled metabolic switches provoking latencies. An example can be found in (Kleman and Strohl, 1994) where the glucose and oxygen consumption rates and CO_2 evolution rate suddenly and reproducibly decreased, causing a break of the metabolism for a period of 40 min and a drop in the biomass productivity.

The second approach consists of regulating either the substrate or the by-product concentration at specific levels (Dewasme et al., 2011a,b; Santos et al., 2012a). The substrate concentration should be close to a threshold corresponding to the critical oxidative capacity, while the by-production concentration should be close to zero. This approach was considered for several biological processes characterized with overflow metabolism such as *Saccharomyces cerevisiae* (Picó et al., 2009). The main challenge in *E. coli* fed-batch cultures is the difficulty of on-line implementation due to the requirement of accurate measurements of low-level concentrations of acetate and/or glucose.

This chapter proposes a control strategy combining the Generic Model Control with a software sensor to monitor and regulate the acetate concentration on-line. The approach is tested and validated through experimental runs on a lab-scale bioreactor.

An experimental implementation of linearizing control to fed-batch cultures of *Escherichia coli* cultures is published in (Rocha, 2003), where the acetate concentration is regulated to a pre-defined set-point. However, the control strategy relies on accurate knowledge of the model parameters, which is a significant drawback since a bioprocess model is always uncertain. Parameter adaptation strategies are usually applied to compensate the uncertainty in the kinetic terms of the process

model. However, stability is not guaranteed in the presence of unmodeled dynamics and high noise levels, this is why we propose a robust control design procedure, which will alleviate this difficulty.

Moreover, the control loop developed in (Rocha, 2003) is based on a flow injection analysis (FIA) device, whose market distribution has been disrupted, and no other similar device has been (re)developed in the meantime. The use of state estimation, or software sensors, seems therefore the most appropriate solution to avoid the burden of complex, unreliable sensing techniques. The results reported in (Rocha, 2003) show that exponential growth could not be sustained in the experimental studies, which might be an indication of the lack of accuracy of the FIA device. In this study, we propose an Unscented Kalman Filter (UKF) for the on-line reconstruction of the acetate concentration.

In this chapter, a robust version of the Generic Model Control (GMC) strategy is developed to control the acetate concentration to a low pre-defined value. $\mathcal{LMI}$s are considered in the control synthesis to derive the GMC control gains. The control design includes performance requirements using the regional pole placement technique. The approach ensures both the robust stability of the process in the presence of model uncertainties and process and measurement noise, and the desired transient performance of the closed-loop system.

Experimental implementation of the proposed strategy on a lab-scale reactor is performed to validate the control and estimation performance. Finally, a comparative discussion with the growth regulation strategy developed in chapter 6 is presented at the end of this chapter.

7.2 GMC control of the acetate concentration

7.2.1 Control objective

As described in previous chapters, the goal behind the control strategies is to drive the culture near the optimal operating conditions to maximize the biomass productivity. These conditions lie at the boundary of the respiro-fermentative and respirative modes (Dewasme et al., 2011a; Rocha, 2003), where all the available substrate is assumed to be allocated for biomass production, and the acetate is neither produced nor consumed.

Maintaining the culture at the edge between the respirative and respiro-fermentative modes requires controlling the substrate concentration to the critical value S_{crit}. An efficient on-line substrate measurement around this value is required, but the concentration level is below the resolution of currently available glucose probes (Dewasme et al., 2011a).

In the previous chapter, we proposed a solution to avoid acetate accumulation by regulating the biomass growth rate at a defined value, chosen slightly lower than the maximal growth rate. This choice was motivated by the availability of reliable biomass measurements, and the lack of on-line substrate and acetate measurements at low levels.

In this chapter, we propose another solution to avoid overflow metabolism by regulating the acetate concentration around a low value A_{ref}. This value should be chosen as close to zero as possible in order to maintain the process near the optimal metabolic edge. On the other hand, a safety margin should be taken to avoid metabolic switches between the operating regime and the accuracy of the estimation procedure (around 0.1 g/L) (Dewasme et al., 2011a).

It is reported in (Pinhal et al., 2019) that industrially relevant inhibitory levels of acetate concentration are on the order of 100 mM (6 g/L). The authors studied the effect of acetate presence in the culture medium on *E. coli* metabolism and showed that a concentration of 16.67 mM (1 g/L) corresponds to less than 20% drop in the growth rate compared to its maximal value.

Based on these elements, the control objective will be the tracking of a defined low acetate concentration reference A_{ref} chosen between 0.3-2 g/L.

7.2.2 Control design

The GMC paradigm presented in the previous chapter is applied to regulate the acetate concentration in fed-batch *E. coli* cultures. The control structure is illustrated in Figure 7.1. Considering acetate concentration as the controlled output, and assuming its availability for measurement ($y = A$).

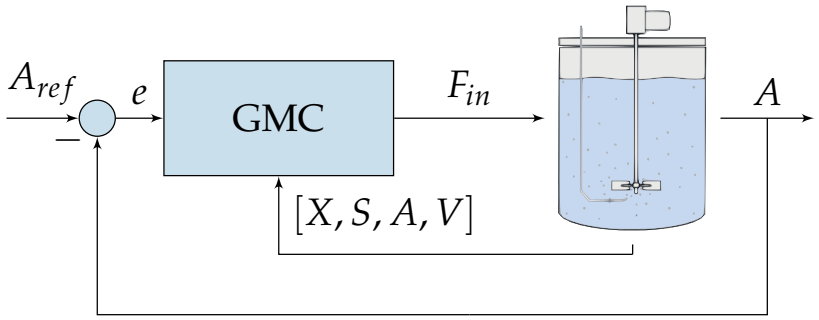


FIGURE 7.1: Generic model control applied to fed-batch *E. coli* cultures to regulate the acetate concentration

We recall the *E. coli* model equations presented in chapter 2:

$$\dot{X} = (k_{X1}\mu_1 + k_{X2}\mu_2 + k_{X3}\mu_3)X - D\,X \tag{7.1a}$$
$$\dot{S} = -(\mu_1 + \mu_2)X - D\,(S - S_{in}) \tag{7.1b}$$
$$\dot{A} = (k_{A2}\mu_2 - \mu_3)X - D\,A \tag{7.1c}$$
$$\dot{V} = F_{in} \tag{7.1d}$$

As the theoretical value of S_{crit} is very small (below 0.1 g/L) and assuming a quasi-steady state of S (i.e. no accumulation of glucose in the neighborhood of

the optimal operating conditions), the small quantity of substrate VS is almost instantaneously consumed by the cells ($\frac{d(VS)}{dt} \approx 0$ and $S \approx 0$), and Equation (7.1b) yields:

$$\mu_2 X = -\mu_1 X + DS_{in} \tag{7.2}$$

where μ_1 and μ_2 are nonlinear functions of S, A and O. Replacing $\mu_2 X$ by Equation (7.2), the mass balance equation of A (Equation (7.1c)) can be expressed as:

$$\dot{A} = -k_{A2}\mu_1 X - \mu_3 X - u\left(A - k_{A2}S_{in}\right) \tag{7.3}$$

where $u = D = \frac{F_{in}}{V}$ is the control input. Applying the GMC scheme introduced in chapter 6 yields:

$$\dot{A} = \hat{u} = G_1(A_{ref} - A) + G_2 \int_0^t (A_{ref} - A)\partial\tau \tag{7.4}$$

Equating (7.3) and (7.4), the following control law is obtained:

$$F_{in} = V\frac{\hat{u} + \left(k_{A2}\mu_1 + \mu_3\right) X}{k_{A2}S_{in} - A} \tag{7.5}$$

$$\hat{u} = G_1(A_{ref} - A) + G_2 \int_0^t (A_{ref} - A)\partial\tau \tag{7.6}$$

where $(k_{A2}\mu_1 + \mu_3)$ is an assumed uncertain kinetic term. The next section therefore explores a robust control design in order to compensate this uncertainty.

7.3 Robust control design

The linearizing control law obtained in the previous section can be written in the following form:

$$\begin{aligned} F_{in} &= V\frac{\hat{u} + \theta X}{k_{A2}S_{in} - A} \\ \hat{u} &= G_1(A_{ref} - A) + G_2 \int_0^t (A_{ref} - A)\partial\tau \end{aligned} \tag{7.7}$$

where θ is the kinetic term given by:

$$\theta = k_{A2}\mu_1 + \mu_3 \tag{7.8}$$

Structural and parametric uncertainties as well as estimation errors can be lumped into a global parametric error:

$$\delta = \bar{\theta} - \theta \tag{7.9}$$

where δ is a nonlinear function of (S, A, O) representing possible inexact cancellations of nonlinear terms due to model uncertainties, and $\bar{\theta}$ represents the hypothetical exact (unknown) value. Rewriting the control law in Equation (7.7) using the new expression of the kinetic term from Equation (7.9), we obtain:

$$F_{in} = V\frac{\hat{u} + \bar{\theta}X - \delta X}{k_{A2}S_{in} - A} \tag{7.10}$$

which corresponds to the perturbed reference system:

$$\dot{A} = \hat{u} - \delta X \tag{7.11}$$

Following a similar approach to the one developed in (Dewasme et al., 2011a), the time-varying parameter δ is assumed bounded and belonging to the set Δ defined by:

$$\Delta := \{\delta : \underline{\delta} \leq \delta \leq \bar{\delta}\} \tag{7.12}$$

with $\underline{\delta}$ and $\bar{\delta}$ respectively representing the minimal and maximal values of the assumed bounded polytopic set.

The control parameters G_1 and G_2 are designed to ensure some robustness and tracking performance to the overall closed-loop system. To this end, the acetate tracking error ($\tilde{A}_1 = A_{ref} - A$) dynamics can be modeled by the following augmented system, illustrated in Figure 7.2:

$$\begin{aligned} \dot{\tilde{A}}_1 &= \frac{d}{dt}(A_{ref} - A) = -\hat{u} + \delta X \\ \dot{\tilde{A}}_2 &= A_{ref} - A = \tilde{A}_1 \end{aligned} \tag{7.13}$$

Considering the state vector $x = \tilde{A} = \begin{bmatrix}\tilde{A}_1 & \tilde{A}_2\end{bmatrix}^T$, the performance output $e = \tilde{A}_1 = (A_{ref} - A)$ and the disturbance $w = [X\ A_{ref}]^T$, the control problem can be formulated as a state feedback controller ($\hat{u} = Kx, K = [G_1\ G_2]$) applied to the augmented system $\mathcal{M}$:

$$\mathcal{M} : \begin{cases} \dot{x} = A_{\mathcal{M}}x + B_w w + B_u \hat{u} \\ e = C_e x + D_{ew} w + D_{eu}\hat{u} \end{cases} \tag{7.14}$$

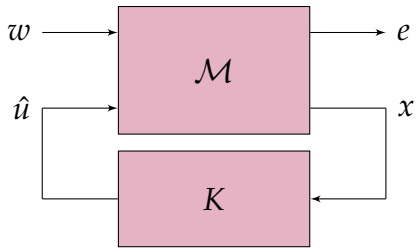


FIGURE 7.2: Robust control scheme

The state-space matrices are given by:

$$\begin{aligned} A_{\mathcal{M}} &= \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \quad B_w = \begin{bmatrix} \delta & 0 \\ 0 & 0 \end{bmatrix} \quad B_u = \begin{bmatrix} -1 \\ 0 \end{bmatrix} \\ C_e &= [1\ 0] \qquad D_{ew} = [0\ 0] \quad D_{eu} = 0 \end{aligned} \tag{7.15}$$

and the representation of the closed-loop system is therefore given by:

$$\left[\begin{array}{c} \dot{x} \\ e \end{array}\right] = \left[\begin{array}{c|c} A_f & B_f \\ \hline C_f & D_f \end{array}\right] \left[\begin{array}{c} x \\ \hline w \end{array}\right] = \left[\begin{array}{c|c} A_{\mathcal{M}} + B_u K & B_w \\ \hline C_e + D_{eu} K & D_{ew} \end{array}\right] \left[\begin{array}{c} x \\ \hline w \end{array}\right] \tag{7.16}$$

7.3.1 Robustness constraints

The control design problem consists in determining the controller parameters G_1 and G_2 so as to limit the infinity norm of the closed-loop transfer function within a predefined performance index (Chilali and Gahinet, 1996), $(\left\| \boldsymbol{T}(s) = D_f + C_f \left(s\mathbb{I}_n - A_f\right)^{-1} B_f \right\|_\infty < \gamma_\infty)$, where s is the Laplace variable. First, the following assumptions on the plant parameters are considered:

Assumption 1. *The pair $(A_{\mathcal{M}}, B_u)$ and $(A_{\mathcal{M}}, C_e)$ are respectively stabilizable and detectable*

Assumption 2. $D_{eu} = \mathbb{O}_{n_e, n_u}$

Under the previous assumptions, the Bounded Real Lemma (Chilali and Gahinet, 1996) for continuous-time systems gives an equivalent $\mathcal{LMI}$ formulation of the control problem:

Lemma 1. *The $\mathcal{H}_\infty$ norm of the continuous-time transfer function $\boldsymbol{T}(s)$ associated to the closed-loop system (7.16) is strictly smaller than γ_∞ if and only if there exists a symmetric positive definite matrix Q_∞ verifying:*

$$\begin{aligned} &Q_\infty > 0 \\ &\begin{bmatrix} A_f Q_\infty + Q_\infty A_f^T & B_f & Q_\infty C_f^T \\ B_f^T & -\gamma_\infty \mathbb{I}_{n_w} & D_f^T \\ C_f Q_\infty & D_f & -\gamma_\infty \mathbb{I}_{n_e} \end{bmatrix} < 0 \end{aligned} \tag{7.17}$$

According to the bounded real lemma, the closed-loop system (7.16) is stable if and only if there exists: $Q_\infty = Q_\infty^T > 0$ verifying:

$$\begin{bmatrix} AQ_\infty + B_u K Q_\infty + Q_\infty A^T + Q_\infty K^T B_u^T & B_w & Q_\infty C_e^T + Q_\infty K^T D_{eu}^T \\ B_w^T & -\gamma_\infty \mathbb{I}_{n_w} & D_{ew}^T \\ C_e Q_\infty + D_{eu} K Q_\infty & D_{ew} & -\gamma_\infty \mathbb{I}_{n_e} \end{bmatrix} < 0 \tag{7.18}$$

Considering $L = KQ_\infty$, the following $\mathcal{LMI}$ is obtained:

$$\begin{bmatrix} AQ_\infty + B_u L + Q_\infty A^T + L^T B_u^T & B_w & Q_\infty C_e^T + L^T D_{eu}^T \\ B_w^T & -\gamma_\infty \mathbb{I}_{n_w} & D_{ew}^T \\ C_e Q_\infty + D_{eu} L & D_{ew} & -\gamma_\infty \mathbb{I}_{n_e} \end{bmatrix} < 0 \tag{7.19}$$

and the controller given by $K = LQ_\infty^{-1}$ ensures a level of robustness w.r.t the bounded uncertainty δ. Next, the desired performance constraints are defined and added to the robustness condition (7.19).

7.3.2 Performance constraints

Besides ensuring the robustness of the closed-loop, it is desirable to achieve some performance in terms of the transient response (e.g. damping, response time, etc.). In other words, constraints are added to the location of closed-loop poles of system (7.16).

For a second-order system with poles $\lambda = -\zeta\omega_n \pm j\omega_d$, the step response is characterized by the undamped natural frequency $\omega_n = |\lambda|$, the damping ratio ζ, and the damped natural frequency ω_d. To ensure a desired transient response, specific bounds are imposed on these quantities, thus constraining the closed-loop poles λ in a prescribed region of the complex plane. Pole placement constraints can be expressed using $\mathcal{LMI}$ regions, which are known to have interesting geometric properties for control purposes (convexity, symmetry, ...) (Chilali and Gahinet, 1996). A suitable region satisfying this criterion is the intersection of the half-plane $s < -\rho < 0$, the disk of radius r and the conic sector defined by an angle Θ. The corresponding region $\mathcal{S}(\rho, r, \Theta)$ (Figure 7.3) is defined as follows:

$$\mathcal{S}(\rho, r, \Theta) = \{a < -\rho < 0, \quad |s = a + jb| < r, \quad a\tan(\Theta) < -|b|\} \tag{7.20}$$

In this way, it is possible to set a minimum decay rate ρ, a minimum damping ratio $\zeta = cos(\Theta)$, and a maximum damped natural frequency $\omega_d = rsin(\Theta)$ (Wood, 1972).

The poles of the closed-loop system (7.16) are contained in the region $\mathcal{S}(\rho, r, \Theta)$, if there exists a symmetric positive definite matrix $Q = Q^T$ verifying (Chilali and Gahinet, 1996):

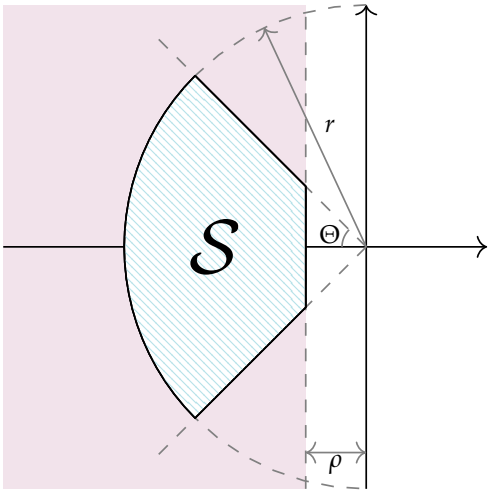


FIGURE 7.3: Representation of the region $\mathcal{S}(\rho, r, \Theta)$

$$
\begin{gathered}
A_f Q + Q A_f^T + 2\rho Q < 0 \\
\begin{bmatrix} -rQ & A_f Q \\ Q A_f^T & -rQ \end{bmatrix} < 0 \\
\begin{bmatrix} \sin\Theta\left(A_f Q + Q A_f^T\right) & \cos\Theta\left(A_f Q - Q A_f^T\right) \\ \cos\Theta\left(Q A_f^T - A_f Q\right) & \sin\Theta\left(A_f Q + Q A_f^T\right) \end{bmatrix} < 0
\end{gathered}
\tag{7.21}
$$

Our control design problem consists then in finding a state-feedback gain K that:

- guarantees the H_∞ performance $\|\boldsymbol{T}(s)\|_\infty < \gamma_\infty$.
- places the closed-loop poles in the $\mathcal{LMI}$ region $\mathcal{S}(\rho, r, \Theta)$ defined by Equation (7.20).

The first criterion (robustness) is ensured by solving Equation (7.19), and computing the matrix Q_∞. On the other hand, a sufficient condition to ensure the performance constraints given by Equation (7.21) is to take $Q = Q_\infty$ (Chilali and Gahinet, 1996), yielding:

$$
\begin{gathered}
A_f Q_\infty + Q_\infty A_f^T + 2\rho Q_\infty < 0 \\
\begin{bmatrix} -rQ_\infty & A_f Q_\infty \\ Q_\infty A_f^T & -rQ_\infty \end{bmatrix} < 0 \\
\begin{bmatrix} \sin\Theta\left(A_f Q_\infty + Q_\infty A_f^T\right) & \cos\Theta\left(A_f Q_\infty - Q_\infty A_f^T\right) \\ \cos\Theta\left(Q_\infty A_f^T - A_f Q_\infty\right) & \sin\Theta\left(A_f Q_\infty + Q_\infty A_f^T\right) \end{bmatrix} < 0
\end{gathered}
\tag{7.22}
$$

The robust GMC control design procedure based on $\mathcal{LMI}$s is summarized in the following steps:

- **Step1**: Select a suitable range for the uncertain variable δ.
- **Step2**: Determine the values of ρ, r, Θ in order to meet a suitable transient performance.
- **Step3**: Solve (off-line) the bounded real lemma (Equation (7.19)) and the performance $\mathcal{LMI}$ (Equation (7.22)) simultaneously, to compute the gain $K = [G_1\ G_2]$, and obtain the robust GMC controller parameters G_1 and G_2.

7.4 Numerical simulations

In this section, a set of numerical simulations is achieved to validate the performance and robustness of the GMC scheme before the experimental validation. The cultures are achieved in a 5-L bioreactor and the kinetic and stoichiometric parameters are those estimated in a previous work (Retamal et al., 2018). The GMC regulation is first tested assuming the acetate concentration is available for measurement, and then coupled to the UKF estimator presented in chapter 3.

The UKF tuning parameters (α, β, κ), the process and measurement noise covariance matrices $\mathbf{Q}$ and $\mathbf{R}$, and the initial state covariance matrix $\mathbf{P}_0$ are given in Table 7.1.

TABLE 7.1: UKF covariance matrices, sigma point tuning parameters, and initial conditions

Parameter	Value	Unit
σ_X	0.01	g/L
σ_S	0.1	g/L
σ_A	0.1	g/L
σ_V	0.001	g/L
r_X	0.01	g/L
$\mathbf{P}_0$	$10^{-4} \times \mathbb{I}_4$	g/L
$[\alpha, \beta, \kappa]$	$[1, 2, 0]$	-
X_0	0.1	g/L
S_0	5	g/L
A_0	0.1	g/L
V_0	3.5	L
S_{in}	500	g/L

First, the performance of the robust GMC design based on the $\mathcal{LMI}$ approach with the regional pole assignment is tested. The control objective is to regulate the acetate set-point A_{ref}, chosen sufficiently low to approach the neighborhood of the optimal trajectory but also sufficiently high to stay within the limit of the

observation sensitivity (0.1 g/L) and maintain the culture in respiro-fermentative mode. The acetate concentration is assumed available on-line for feedback, with the consideration of measurement noise.

The first step in our design approach is to define upper and lower bounds for the parametric uncertainty δ. The expression of the kinetic parameter θ is given by:

$$\theta = k_{A2}\mu_1 + \mu_3 \tag{7.23}$$

The expression of the uncertain term θ, and the kinetic terms μ_1 and μ_3 contain the parameters $k_{A2}, K_s, K_{iA}, k_{OS}, K_A$. Additionally the GMC control law contains the parameter k_{A2}. These parameters can deviate from their nominal values, thereby deviations of maximum 15% are considered. In addition, the parametric sensitivity study presented in appendix D shows that the parameters $q_{S\,\max}$, $q_{O\,\max}$, k_{X2}, have an influence on the acetate concentration. Consequently, the range Δ can be defined by $\underline{\delta} = 0$ and $\bar{\delta} = 0.1$.

Regarding the performance constraints, we desire to enforce a maximal settling time $T_s = \frac{4}{\zeta\omega_n} = \frac{4}{\rho}$ equal to 2 h, and to prevent fast controller dynamics.

To this end, we characterize the section $\mathcal{S}(\rho, r, \theta)$ as the intersection of the half-plane $x < -\rho = -\frac{4}{T_s}$ with the disk of radius $r = 4$ centered at the origin, and the conic section defined by $\Theta = \frac{\pi}{2}$ rad.

In light of these constraints, the $\mathcal{LMI}s$ (Equations (7.19) and (7.22)) are solved numerically using the solver *SeDuMi* (Sturm, 1999) and the following results are obtained:

$$Q_\infty = \begin{bmatrix} 0.143 & -0.034 \\ -0.034 & 0.015 \end{bmatrix} \qquad K^T = \begin{bmatrix} G_1 = 5.61 \\ G_2 = 9.55 \end{bmatrix} \tag{7.24}$$

corresponding to the following damping ratio and natural frequency:

$$\zeta = 0.91 \qquad \omega_n = 3.09 \ \text{rad/h} \tag{7.25}$$

satisfying the performance constraints regarding the settling time $T_s = \frac{4}{\zeta\omega_n} = 1.43$ h. It can be seen in Figure 7.4 that the obtained poles are located in the chosen region of the complex plane.

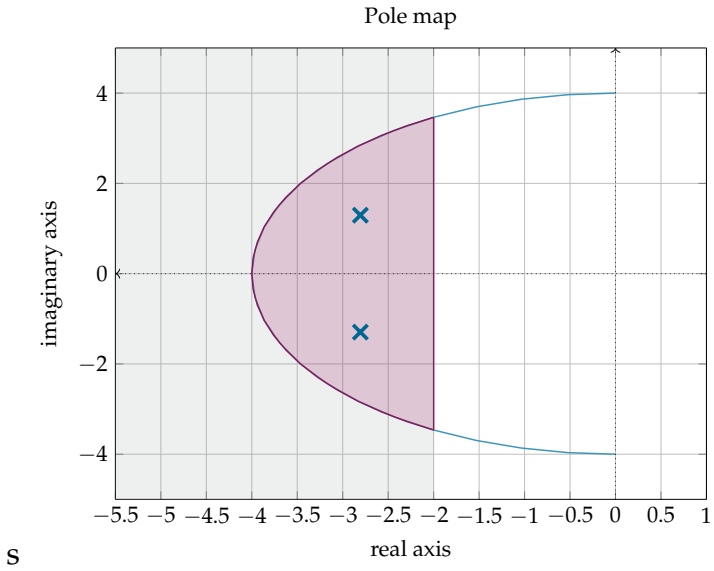


FIGURE 7.4: Plot of the pole location (blue) and the imposed region $\mathcal{S}$ (red)

Figures 7.5 and 7.6 show the closed-loop response of biomass, substrate, acetate concentrations, and the corresponding feed flow-rate in 50 different runs, with $A_{ref} = 0.5$ g/L and kinetic parameter deviations. A white noise is added to the acetate concentration measurement with zero mean and a standard deviation of 0.1 g/L. In all the runs, biomass follows a similar exponential growth in the first hours, while the model errors show their effect in the final hours. Nevertheless, the model uncertainties have a minor influence on the controller performance as can be observed in the acetate evolution, where the set-point is regulated and robust convergence is achieved by the controller. The noisy acetate signal has a mean value of 0.49 g/L and the tracking error ($A_{ref} - A$) has a RMSE of 0.0314 g/L which is lower than the measurement noise amplitude (0.1 g/L). Note that the maximal 2 h settling time condition is also satisfied.

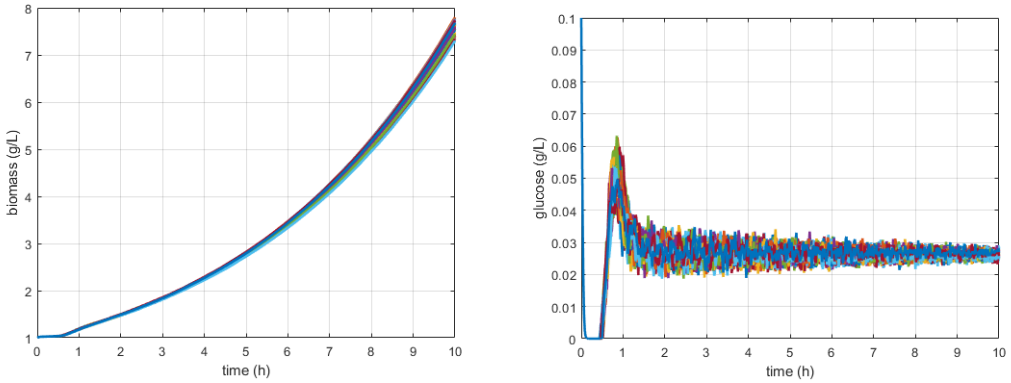


FIGURE 7.5: Biomass and substrate concentrations in 50 runs with kinetic parameter deviations (up to 15%) and a measurement noise standard deviation of 0.1 g/L using the robust GMC control strategy.

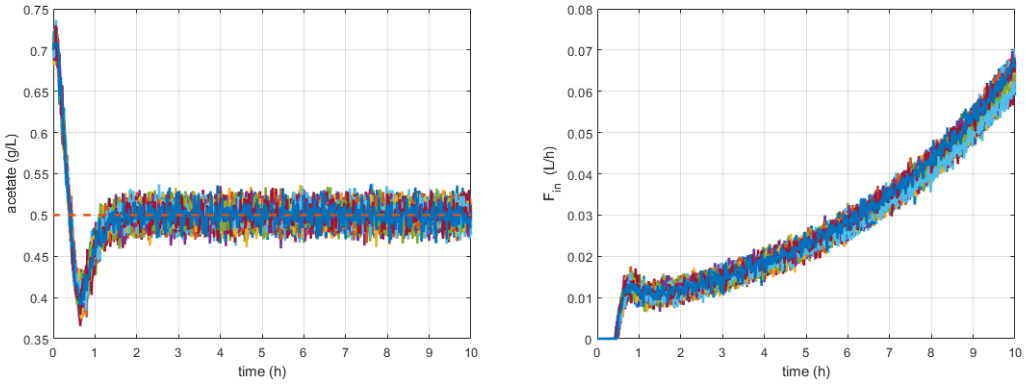


FIGURE 7.6: Acetate concentration and feed flow-rate in 50 runs with kinetic parameter deviations (up to 15%) and a measurement noise standard deviation of 0.1 g/L using the robust GMC control strategy.

Figure 7.7 illustrates an example of the kinetic parameter θ variation in the first run, compared to the theoretical real value $\bar{\theta}$ in the presence of measurement noise. The parametric uncertainties create a bias (δ) to the θ value. Nevertheless, the robust design of the controller allows to compensate for this error.

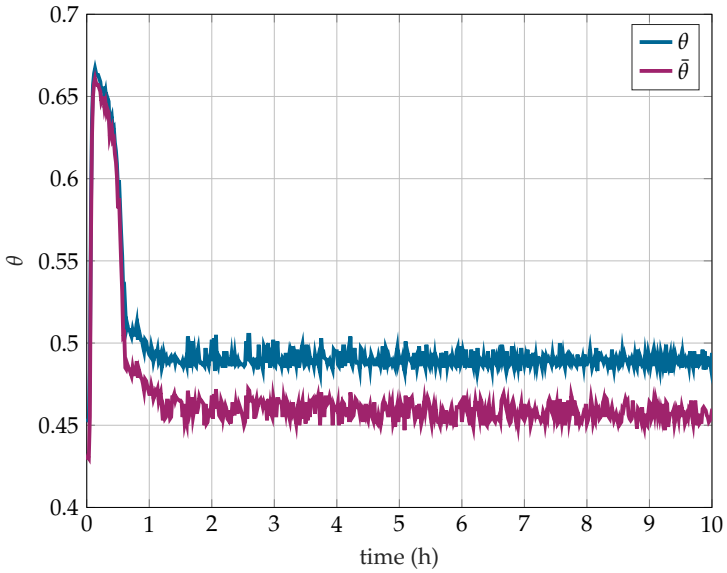


FIGURE 7.7: kinetic parameter θ evolution with random parameter variations and measurement noise (std = 0.1 g/L)

The different biomass productivity levels in all runs are shown in Figure 7.8, where the productivity remains higher than 90 % of the nominal value in 90% of the runs, which is satisfactory from an operational point of view.

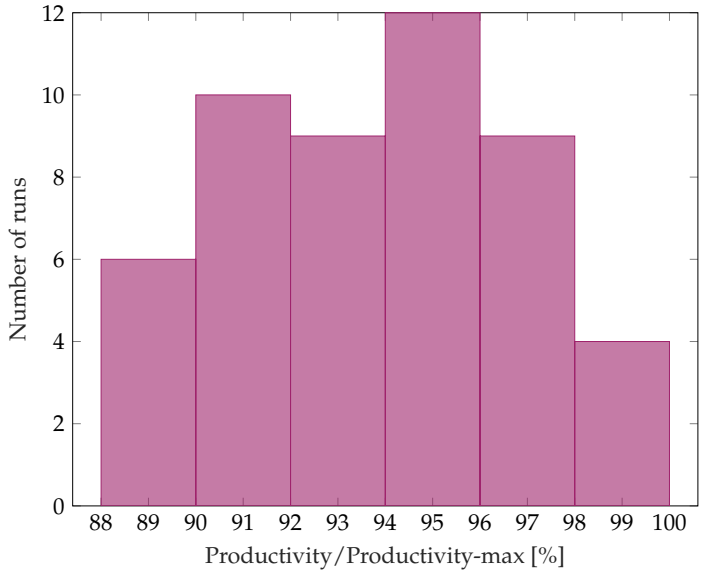


FIGURE 7.8: Productivity levels of the 50 runs with random parameter variations using the robust GMC strategy

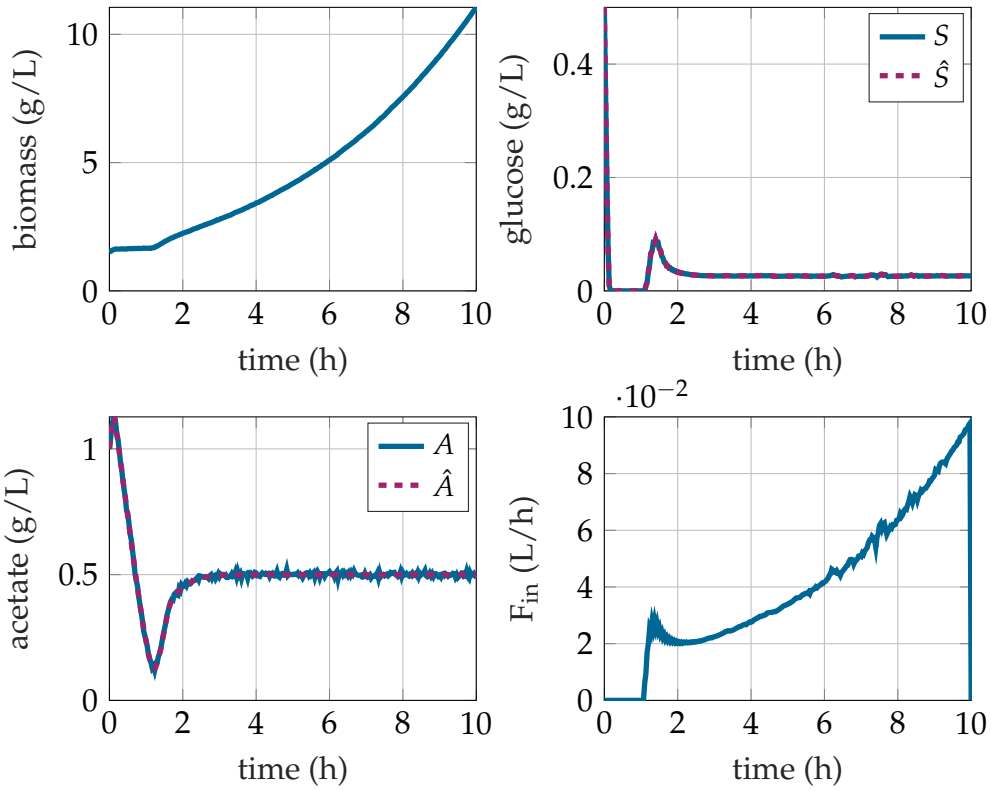


FIGURE 7.9: Coupled UKF-GMC with random parameter values (±15% variation) and a white measurement noise (std =0.1 g/L)).

Finally, the UKF and robust GMC are coupled, and their overall performance

assessed in a new set of numerical simulations. The UKF initial conditions are selected randomly with a maximum deviation of 20% from the real values. Kinetic parameter variations of $\pm 15\%$ of the nominal values and a white measurement noise with a standard deviation of 0.1 g/L are considered. As shown in Figure 7.9, the UKF behaves very well and converges in the first hours to the real state trajectories. We can observe small estimation errors with peaks and troughs around the real substrate value when the substrate (glucose) concentration reaches a critical level of S_{crit}. Fortunately, this is not too detrimental for the controller which is still able to track the acetate concentration reference set-point.

7.4.1 Comparison with the classical GMC

In order to test further the performance and robustness of the control approach, a comparison is achieved with the classical GMC algorithm presented in (Lee and Sullivan, 1988). The parameter tuning is performed by selecting a desired rise time t_r. In the presented simulations, the following parametrization for the classical approach is chosen:

$$\xi = 1, \;\; t_r = 2\,\text{h}, \;\; G_1 = 3, \;\; G_2 = 2.25$$

The classical and the robust controllers are tested in the ideal model case (no parameter variation), and in the case of a random variation in all model parameters up to $\pm 30\%$ of their nominal value. We assume that the acetate concentration is available for measurements, with additive centered white noise with a standard deviation of 0.05 g/L. A series of 100 Monte Carlo (MC) simulations is performed and the results are summarized in Table 7.2.

The results of one simulation are shown in Figure 7.10, where both approaches perform similarly in the ideal model case. However, we can see that with increasing levels of parameter variation, the robust GMC performs better in terms of reference tracking. The mean square errors ($\overline{e_A}$) and the mean acetate concentration ($\overline{A}$) summarized in Table 7.2 show that the robust tuning of the parameters allows the controller to achieve the control objective accurately. We can also observe that a $\pm 30\%$ variation is the breakpoint of both methods, with a slight advantage to the robust GMC design. Note that the robust GMC design was performed assuming a maximum of 15% variation of the plant parameters.

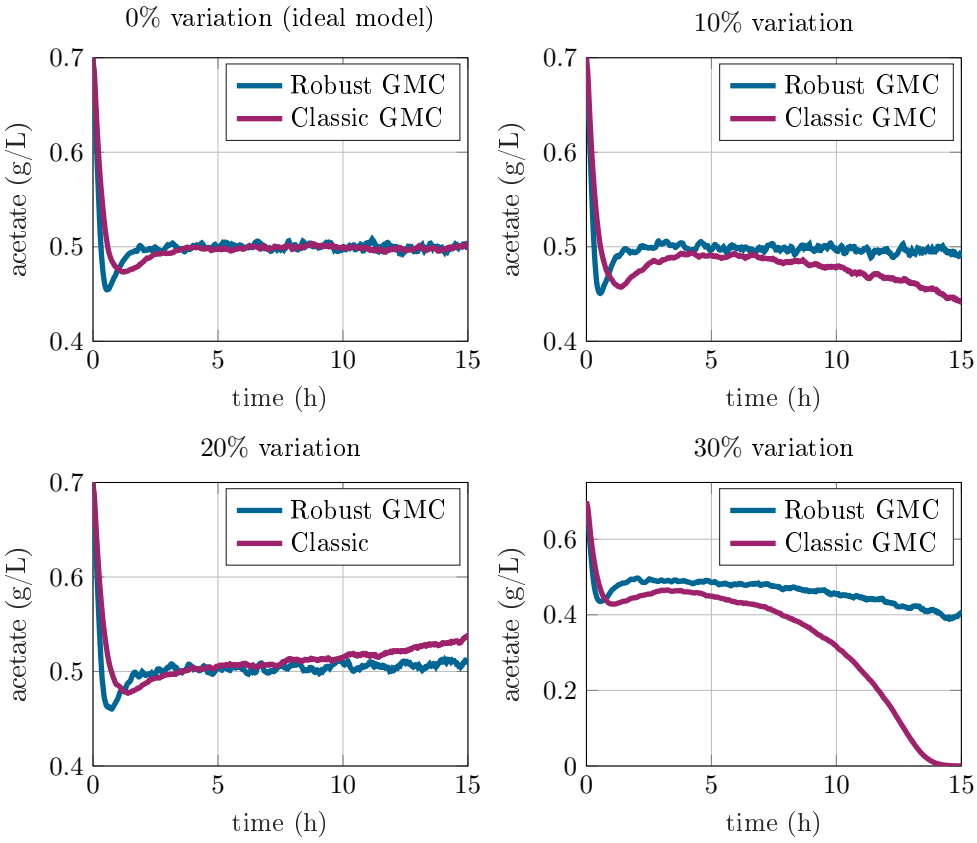


FIGURE 7.10: Comparison between the classical and robust tuning of the GMC strategy, with increasing levels of parameter variation.

TABLE 7.2: Results of 100 Monte Carlo simulations comparing the classical and robust GMC strategies

	$\overline{A}$ (Classic)	$\overline{A}$ (Robust)	$\overline{e_A}$ (Classic)	$\overline{e_A}$ (Robust)
0%	0.4996	0.4996	0.0226	0.0192
10%	0.4968	0.4989	0.0293	0.0200
15%	0.5009	0.5001	0.0343	0.0207
20%	0.4992	0.4998	0.0456	0.0224
25%	0.4885	0.4974	0.0643	0.0261
30%	0.4744	0.4944	0.0843	0.0294
35%	0.4741	0.4913	0.0889	0.0418

7.5 Experimental results and discussion

Two control experiments were performed to test the tracking performance and robustness of the developed UKF-GMC strategy in a real-time environment. Each experiment consisted of a batch phase followed by a fed-batch phase (control

phase). The evolution of the measured biomass (on-line & off-line), glucose, acetate concentrations (off-line), and their estimates, as well as the feed flow-rate (controller output), are shown in Figures 7.11 and 7.13. The operating conditions are also illustrated in Figures 7.12 and 7.14.

TABLE 7.3: Control & estimation parameters and initial conditions used in the experiments

	Experiment 1	**Experiment 2**
Sampling time	$T_s = 3\ min$	$T_s = 3\ min$
Acetate reference	$A_{ref} = 0.5$ g/L	$A_{ref} = 0.7$ g/L
Q	$diag[10^{-4}, 10^{-2}, 10^{-2}, 10^{-8}]$ g/L	$diag[10^{-4}, 10^{-2}, 10^{-2}, 10^{-8}]$ g/L
R	10^{-4} g/L	10^{-4} g/L
$\mathbf{P}_0$	$10^{-4} \times \mathbb{I}_4$ g/L	$10^{-4} \times \mathbb{I}_4$ g/L
$[\alpha, \beta, \kappa]$	[1, 2, 0]	[1, 2, 0]

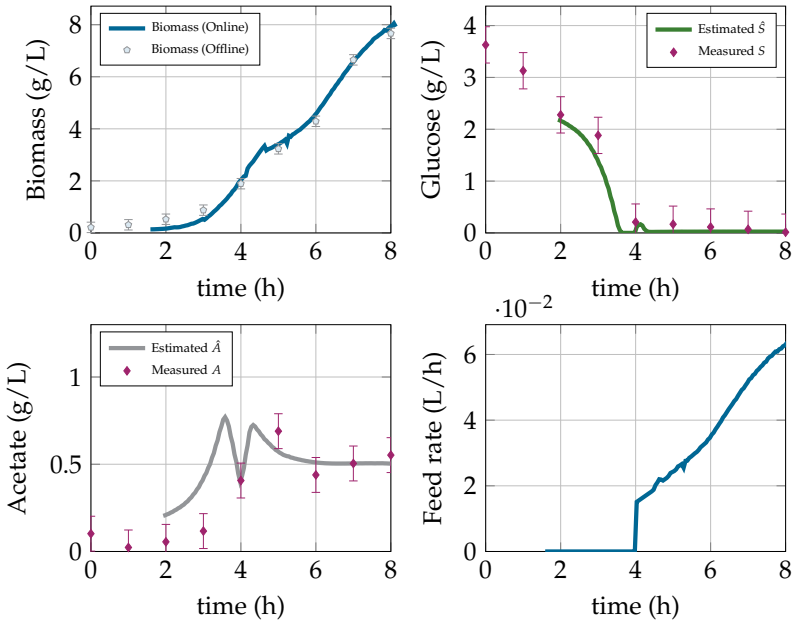


FIGURE 7.11: Experiment 1: Time evolution of the measured biomass, the estimated and measured glucose and acetate concentrations, and the feed-rate

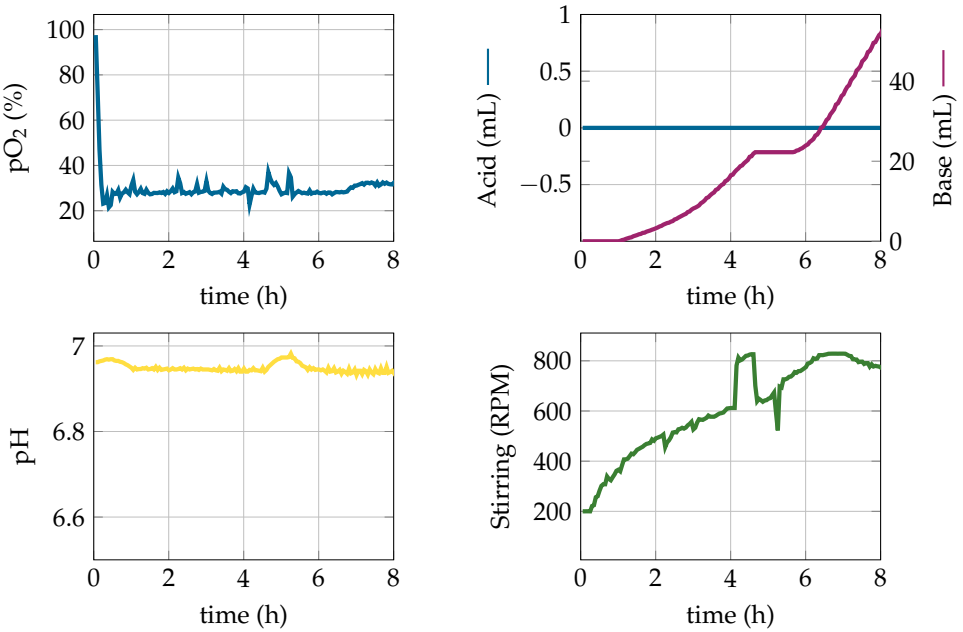


FIGURE 7.12: Experiment 1: Time evolution of the pO_2, acid and base concentrations, pH and stirring

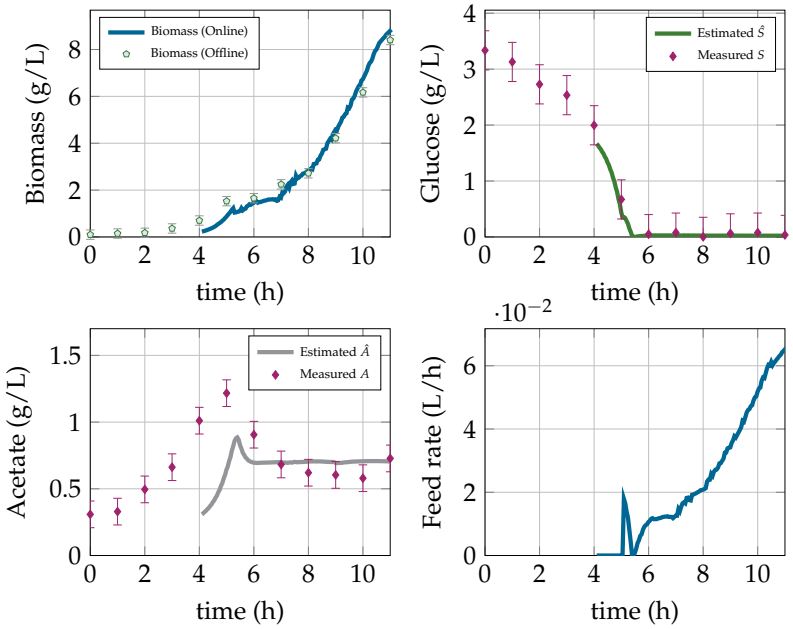


FIGURE 7.13: Experiment 2: Time evolution of the measured biomass, the estimated and measured glucose and acetate concentrations, and the feed-rate

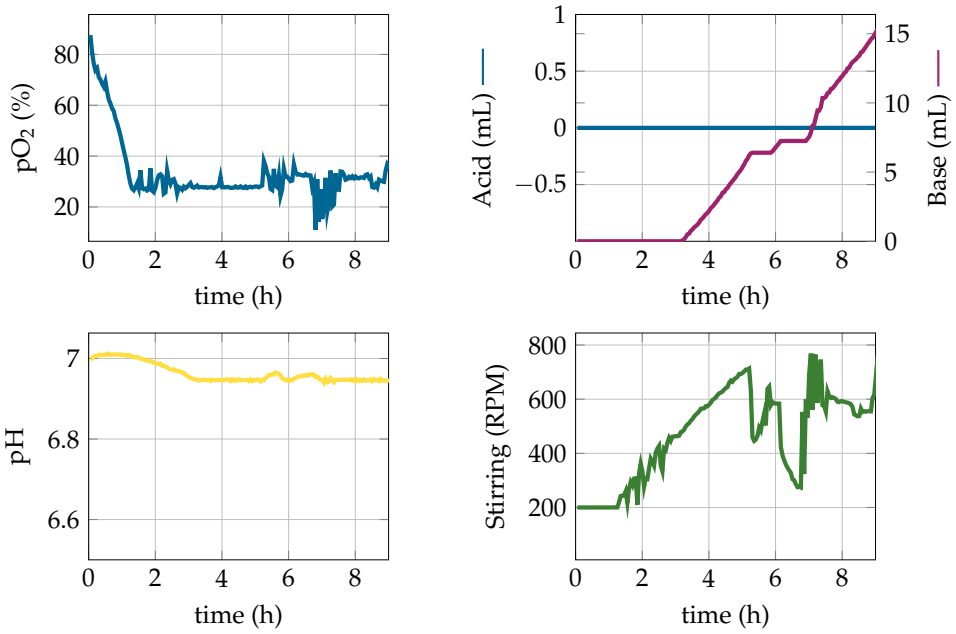


FIGURE 7.14: Experiment 2: Time evolution of the pO_2, acid and base concentrations, pH and stirring

7.5.1 Culture evolution

The control parameters and acetate references for each experiment, as well as the values for the measurement and process noise covariance matrices are given in Table 7.4.

TABLE 7.4: Control & estimation parameters and initial conditions used in the experiments

	Experiment 1	**Experiment 2**
T_s	3 min	3 min
A_{ref}	0.5 g/L	0.7 g/L
σ_X	0.01 g/L	//
σ_S	0.1 g/L	//
σ_A	0.1 g/L	//
σ_V	0.001 g/L	//
r_X	0.01 g/L	//
$\boldsymbol{P}_0$	$10^{-4} \times \mathbb{I}_4$ g/L	//

After reaching the desired operating conditions, the reactor is inoculated with the seed culture, and the batch phase begins. As shown in Figures 7.11 and 7.13, the initial biomass concentration in the reactor ranges from 0.1-0.2 g/L. During this phase, the biomass follows an exponential growth and reaches up to 2 g/L. Since glucose consumption leads to acetate production, the culture is in respiro-fermentative mode. Note that the estimation algorithm is launched during the batch phase.

The batch phase lasts for 4-5 h, and ends after the consumption of the glucose in the medium. On-line indicators of the glucose depletion are the sudden decrease of the stirring speed due to the decrease of cell demand for oxygen, the sudden increase of the pH combined with the stagnation of injected base volume. .

Practically, the feeding (control) should be launched slightly before the complete depletion of glucose, to prevent the interruption of cell growth, and avoid a possible metabolic switch (acetate consumption) as small delays in reprising the growth have been observed in previous experiments. These delays can affect the control performance at the beginning of the fed-batch phase, leading to undesired transitory behaviors. To this end, off-line glucose measurements are performed at every sampling point (1 h), and the control algorithm is launched when the glucose level is below 0.5 g/L.

In experiment 1 (Figure 7.12), the batch phase lasted for 4 h after the almost complete consumption of the glucose in the medium. The fed-batch phase started right after measurement of a low glucose concentration. The first on-line appeared around 5 h, as the pH increase, stirring speed decrease, and base stagnation indicate the consumption of acetate and the drop of the glucose concentration below the critical level. This occurrence of the metabolic switch during the control phase gives the controller the ability to adjust the feeding depending on the culture conditions.

Similarly in experiment 2 (Figure 7.14), the batch phase lasted 5h due to the lower initial biomass concentration, and higher initial acetate concentration. The feeding started while the glucose was not completely consumed and the acetate concentration was higher than the setpoint. A peak can be observed at 5 h indicating an adaptation of the feed by the controller to reduce the acetate and glucose concentrations. The online flags of glucose depletion and metabolic switch can be observed right after the start of the fed-batch phase.

The GMC controller is launched after setting up the acetate reference and the control parameters. The feed solution is injected by the controller and the cells resume their growth, resulting in an increase of the stirring speed due to the glucose oxidation, and the decrease of pH due to CO_2 emission which requires base addition to maintain the pH around its set-point.

The fed-batch phase continues until reaching the saturation limit of the turbidimetric probe (around a biomass concentration of 8 g/L). The maximum attainable cell density depends on the oxygenation limitation related to the bioreactor scale as can be observed in several studies (Dewasme et al., 2010; Retamal et al., 2018; Rocha, 2003). Therefore, the end of the fed-batch phase is forced by either an exhausted feed medium, or the limiting oxygenation conditions.

7.5.2 Acetate and glucose estimation

As presented in the simulation section, the on-line biomass concentration measurement provided by the turbidimetric probe, and the kinetic model with identified parameter values from (Retamal et al., 2018) are used to estimate the acetate and glucose concentrations using the UKF. The estimation is launched during the

batch phase after reaching a measured biomass concentration higher than 0.2 g/L to ensure a good estimation accuracy.

The measurement noise affecting the biomass concentration signal is considered as a centered white noise with a standard deviation of 0.1 g/L. On the other hand, the degree of confidence in the model regarding the substrate and acetate concentration signals is lower compared to the biomass concentration.

The initialization of the estimated state vector remains a challenge. The biomass and glucose concentrations can be measured every hour, and therefore accurate initial conditions for these variables is possible. The acetate concentration on the other hand is measured after the end of the culture and therefore initialization errors are significantly higher for this variable. Nevertheless, the UKF manages to correctly estimate the acetate concentration.

In both experiments, the UKF performance in the fed-batch phase is satisfactory, despite the initialization errors and the model uncertainties. The glucose and acetate estimations fit very well with the off-line measurements during the control period, and the convergence is achieved in less than 1 h. Table 7.5 shows the estimation mean square error values for each estimated state (i.e., substrate and acetate) during the fed-batch phase of both experiments, which are on par with than the measurement sensitivity (0.1 g/L).

TABLE 7.5: Experimental study - UKF estimation mean square errors (in g/L)

	$\overline{e_S}$(g/L)	$\overline{e_A}$(g/L)
Experiment 1	0.0885	0.0679
Experiment 2	0.0381	0.1132

7.5.3 GMC control performance

The control objective, as explained in previous sections, is to regulate the acetate concentration to a predefined set-point, and maintain the culture in the respiro-fermentative mode close to the optimal limit. As can be seen in Figures 7.11 and 7.13, acetate accumulation is avoided in both cultures, and the concentration is limited to less than (1 g/L) during the fed-batch phase.

In the first experiment (Figure 7.11), the estimated acetate concentration is regulated and converges to the desired reference, respecting the chosen settling time. The second experiment (Figure 7.13) presents the same performance regarding the GMC algorithm convergence, with a different set-point and a longer control time.

7.5.4 Discussion

The presented control method provides a practical approach to avoid overflow metabolism in *E. coli* fed-batch cultures. However, it offers a suboptimal solution, since regulating the substrate concentration at the critical level is impractical due to the lack of accurate on-line measurements.

In order to evaluate the efficiency of the proposed approach, a comparison is performed in simulation with the growth rate regulation strategy presented in the previous chapter. The growth rate is regulated to a set reference value μ_{set}, usually chosen slightly below the maximal growth rate in order to avoid acetate accumulation while maximizing the biomass productivity. This control objective is achieved by tracking a predefined biomass trajectory corresponding to the chosen reference growth rate (De Battista et al., 2012; Rocha et al., 2008).

For this purpose, we set the biomass regulation to track a defined growth rate μ_{set} chosen at 90% of the theoretical maximal value (μ_{max} = 0.26 L/h), corresponding to the critical substrate concentration and the maximal oxidative capacity. On the other hand, we set the acetate regulation to track a reference of 0.5 g/L. The acetate concentration is assumed to be measured with additive white noise of a 0.05 g/L standard deviation. The comparison is achieved with the initial conditions and control parameters presented in Table 7.6.

TABLE 7.6: Initial conditions and control parameters for the comparison simulations

Initial conditions		
X_0	1	g/L
S_0	0	g/L
A_0	0.2	g/L
V_0	3.5	L
GMC-X parameters		
G_1	6	-
G_2	9	-
GMC-A parameters		
G_1	5.61	-
G_2	9.55	-

First, we assume that the model parameters and maximal growth rate μ_{max} are perfectly known. Then, we introduce a fixed variation in the maximal oxidative capacity q_{Omax} which is directly linked to the maximal growth rate value.

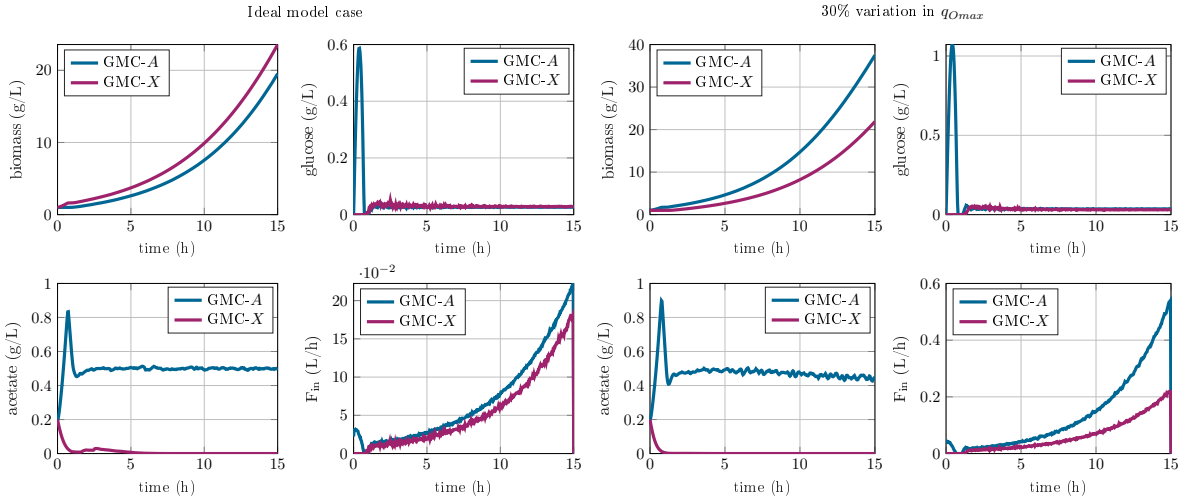


FIGURE 7.15: Comparison between the control approaches in the ideal model case, and in the presence of parametric variation. Plot of the state variables and the feed-rate.

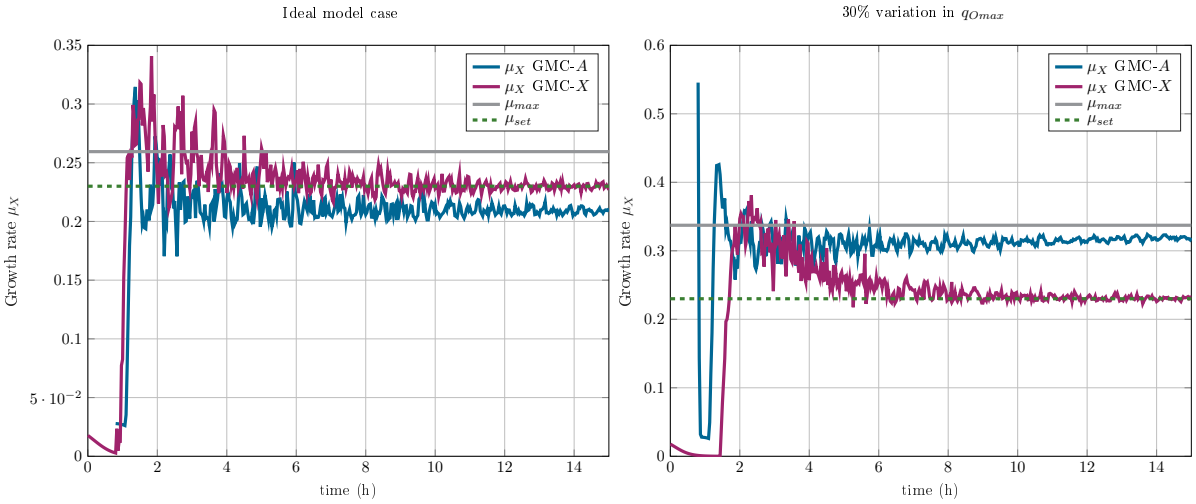


FIGURE 7.16: Comparison between the control approaches in the ideal model case, and in the presence of parametric variation. Plot of the specific biomass growth rates.

TABLE 7.7: The effect of parameter variation on the control performance

Variation in q_{Omax}	$\frac{\mu_X}{\mu_{max}}$% (GMC-X)	$\frac{\mu_X}{\mu_{max}}$% (GMC-A)	S_{crit}
0%	89%	81%	0.0375
10%	81%	85%	0.046
20%	75%	89%	0.0529
30%	70%	93%	0.0628

Simulation results are shown in Figures 7.15 and 7.16. In the ideal model case (no parameter variation), the biomass growth regulation (GMC-X) has a slightly better overall performance. The reference growth rate is tracked accurately at 0.23 h^{-1} corresponding to 89% of its maximal value. On the other hand, regulating the acetate concentration (GMC-A) at 0.5 L/h leads to a biomass growth rate of 0.21 h^{-1} corresponding to 81% of the maximal value as can be seen in Table 7.7. This result shows that the presence of acetate in the medium reduces the biomass growth rate, due to lower substrate consumption rate caused by the activation of the acetate consumption pathways according to the bottleneck theory. However, keeping the acetate at a low concentration reduces its inhibitory effect, and keeps the culture close to the optimal conditions.

The introduction of 20% variation in q_{Omax} leads to an increase of the critical substrate concentration S_{crit} and consequently the maximal growth rate μ_{max}. Despite the model mismatch, the biomass growth rate regulation presents a good performance in tracking the reference rate. However, it corresponds to only 75% of the new maximal value, and therefore the biomass productivity is also lower than its optimal value compared to the nominal case. This is due to the increase in the gap between the reference μ_{set} and new maximal growth rate μ_{max}.

The acetate regulation on the other hand offers a more consistent performance, and gives a better growth rate ratio (89%). Furthermore, the growth rate ratio is higher with increasing variation in the maximal oxidative capacity as can be seen in Table 7.7.

This result highlights a problem with targeting a specific growth rate as a control objective, as it requires accurate determination of the maximal value, and then target a lower growth rate to avoid acetate accumulation. This is a difficult task due to the uncertain nature of bioprocesses, as parameter variation depends on several factors such as the variation in operating conditions between batches. If the maximal growth rate is underestimated, the resulting suboptimal biomass productivity is lower than the desired one. If the maximal growth rate is overestimated, a regulation at 90% of this value could lead to acetate accumulation and metabolic switches, and thereby a growth inhibition.

On the other hand, regulating the acetate concentration and maintaining it at a low value offers a better practical trade-off, since the accumulation is avoided, and the obtained growth rate is consistent in the case of model mismatch. This is an interesting result since the acetate regulation approach is robust towards the change in operating conditions, and is not specific to the bacterial strain. The strategy could be applied to a different strain while ensuring the same level of performance without the need to estimate μ_{max} accurately.

7.6 Conclusion

In this chapter, a robust Generic Model Control strategy is presented and applied to drive fed-batch cultures of *E. coli* BL21 (DE3) near the optimal operating conditions.

The control objective of this proposed strategy is to regulate the acetate concentration at a defined low level. This approach has several advantages compared to the growth rate regulation presented in the previous chapter.

Although it provides a suboptimal solution, maintaining the acetate concentration is a practical alternative to avoid overflow metabolism. The regulation of the substrate at the critical level is impractical since accurate measurements at a low level are not possible via the available sensors on the market.

Furthermore, targeting the optimal operating conditions and the critical substrate concentration causes fluctuations and metabolic switches between the oxidative and oxido-fermentative regimes. Thereby, maintaining the acetate concentration at a low level provides a safety margin and a practical solution to avoid cell stress caused by the repetitive switches.

Due to the uncertain nature of the bioprocess model, a robust design procedure using the $\mathcal{LMI}$ formalism is carried out, to compensate the model mismatch, disturbances, and measurement noise. Performance constraints are also formulated with $\mathcal{LMI}$s to ensure desired properties of the closed-loop transient response.

Since the controlled variable (acetate) is not available for on-line measurement, a state estimation algorithm is required and an Unscented Kalman Filter (UKF) is implemented. The robust GMC controller and the state estimation by the UKF were validated both through simulation runs and in real-time experimental conditions.

Finally, fed-batch experiments with a lab-scale reactor were performed in order to validate the efficiency of the coupled GMC-UKF strategy in driving the cultures near the optimal operating conditions.

The results showed that the proposed control strategy is not restricted to the studied strain since accurate determination of the maximal growth rate is not required. It is also adaptable to different control objectives such as substrate regulation at high concentrations in order to promote the product formation.

An improvement of the proposed control scheme is tracking a successively decreasing set-point calculated by numerical on-line optimization based on the estimation of the maximal growth rate. An experimental validation of this approach in future works could improve the process productivity since it provides a good trade-off between practicality and best achievable sub-optimality.

A performance comparison of the robust GMC with advanced control strategies developed for the same objective would give an insight about the practical efficiency of the proposed solution. To this end, a nonlinear model predictive controller (NMPC) is presented and implemented to the studied process in the next chapter. The NMPC has several interesting features like the explicit use of the nonlinear model and the on-line optimization procedure.

Chapter 8

Nonlinear model predictive control of the acetate concentration

8.1 Introduction

As mentioned in previous chapters, several solutions to avoid growth inhibition caused by overflow metabolism have been developed, ranging from genetic modifications of the strains to closed-loop control and optimization of the bioprocess. The common objective behind the control strategies is to maintain the substrate concentration at a critical threshold, corresponding to the maximal oxidation capacity (Jana and Deb, 2005). Various nonlinear control methods of processes exhibiting overflow metabolism have been proposed in the literature (Abadli et al., 2020; Benattia et al., 2015; Dewasme et al., 2011a,b; Gonzalez et al., 2016; Santos et al., 2012a).

Among these studies, Nonlinear Model Predictive Control (NMPC) is often considered following its several advantages and wide use in industrial applications (Forbes et al., 2015; Qin and Badgwell, 2000). NMPC is an advanced model-based control strategy developed for nonlinear processes. It consists in solving an optimization problem on-line under a set of defined constraints. NMPC can predict, using a nonlinear dynamic model of the process, the effect of control steps on controlled variables over a finite horizon. A general formulation of the NMPC

problem consists in minimizing the quadratic objective function over a finite moving horizon.

Various structures of the NMPC strategy have been developed throughout the years for numerous control purposes (Findeisen et al., 2007). The difference between the NMPC schemes lies in the definition of the cost function and the constraints. An overview of the nonlinear model predictive methods can be found in (Allgöwer and Zheng, 2012). Hereafter, the main NMPC approaches are briefly presented.

Explicit NMPC (E-NMPC) is an offline NMPC variation based on the parametric programming technique, where the control law is pre-calculated offline (Bemporad et al., 2002). The E-NMPC does not require significant computational resources and is suited for control problems for systems with fast dynamics. Other structures focus on the stability of the closed-loop by approximating an infinite-horizon optimization problem. The list includes Finite-horizon NMPC with terminal equality constraint (Keerthi and Gilbert, 1988), Finite-horizon NMPC with terminal inequality constraint (Morari and Lee, 1999; Scokaert et al., 1999), Finite-horizon NMPC with terminal cost (Bitmead et al., 1990), and the Quasi-infinite horizon NMPC (Chen and Allgöwer, 1998). Another variation is the Economic NMPC (E-NMPC), where the objective function is an economically-oriented cost function of the state variables and the control inputs (Diehl et al., 2010).

Robust MPC schemes are designed to account for a set of bounded uncertainties and disturbances in the process model while satisfying the control objectives. There are several approaches to robust NMPC: In Min-max MPC, the optimization is performed with respect to all possible evolutions of the bounded model uncertainty (Scokaert and Mayne, 1998). The variations of the uncertainty related to the process model parameters is included in the cost function to be optimized.

Another robust NMPC approach is the Linear parameter varying NMPC (LVP), which is based on the reformulation of the process nonlinear model with a linear parameter-varying model (LPV), mostly using linear and bilinear matrix inequalities (LMI/BMI) in the control. H_∞-NMPC approaches consider the implementation of the H_∞ problem in a receding horizon framework, by considering a particular choice of the objective function (Gautam et al., 2013; Wang et al., 2009, 2010b).

Robust NMPC approaches are designed to handle model mismatch and parametric uncertainties, which make them appropriate for bioprocess models. However, they are known to carry a high computational burden as the size of the optimization problem increases with the number of uncertainties and the prediction horizon. This may explain the fewer experimental implementations of the MPC strategies for fed-batch bioprocesses.

The NMPC scheme can be found in various industrial fields such as chemical plants, oil refineries, power electronics, automotive industry, biochemistry, and many others (Qin and Badgwell, 2003). However, only a few industrial applications of NMPC exist in biotechnology (Stanke and Hitzmann, 2013). The reasons stem from the difficulties of developing reliable dynamic models of bioprocesses, the high computational costs, and the absence of efficient real-time monitoring solutions.

Among the experimental studies on lab-scale bioprocesses, the NMPC strategy have been applied to a variety of micro-organism cultures such as mammalian cell cultures (Aehle et al., 2012; Craven et al., 2014), yeast cultures (Cochrane and Alamir, 2018; Preuß et al., 2000), hybridoma cell cultures (Dewasme et al., 2013a, 2015), fungi (Ashoori et al., 2009), microalgae (Del Rio-Chanona et al., 2015; Gorrini et al., 2020) and bacterial cultures (Rio-Chanona et al., 2016; Santos et al., 2012b; Tebbani et al., 2010; Ulonska et al., 2018).

These studies are usually conducted in fed-batch mode, where an inlet feed-rate is the sole manipulated input used to regulate variables such as substrates or specific rates to defined set-points (Aehle et al., 2012; Craven et al., 2014; Dewasme et al., 2015). Other control objectives may also aim at maximizing substrate oxidation rates while minimizing product fermentation rates (Santos et al., 2012a), or regulating/limiting byproduct formation (Dewasme et al., 2010; Logist et al., 2011; Tebbani et al., 2010; Valentinotti et al., 2004). Multiple input (feed rates) MPC policies are also examined in (Amribt et al., 2014; Ashoori et al., 2009; Rio-Chanona et al., 2016).

This chapter focuses on the experimental implementation and validation of the NMPC to maximize the biomass productivity in fed-batch *E. coli* BL21(DE3) cultures, by regulating the acetate concentration at a low value and limiting the overflow metabolism effect. The NMPC algorithm is implemented in a real-time environment, and the computing effort is reduced through the control vector parametrization (CVP) technique (Banga et al., 2005).

As discussed in the previous chapter, the unscented Kalman filter (UKF) is used to provide estimates of the acetate and glucose concentrations based on the biomass measurement. This choice is motivated by the good results obtained in the previous experiments.

The chapter is organized as follows. A general description of the NMPC basic principles is presented. These principles are then applied to the specific case of acetate concentration regulation in fed-batch *E. coli* cultures. A set of numerical simulations is presented to showcase the performance of the proposed NMPC strategy in the nominal case (no model mismatch), and in presence of parametric uncertainty in the process model. Finally, the NMPC and UKF algorithms are implemented in real-time to a lab-scale bioreactor for experimental validations. The chapter ends with final remarks and future work directions.

8.2 Principles of Nonlinear Model Predictive Control

Nonlinear model predictive control (NMPC) is an optimization-based feedback control strategy applied for nonlinear systems. It is applied for a variety of control objectives including stabilization, trajectory tracking, optimum seeking, and many other control problems.

The NMPC strategy is based on the receding horizon principle, where the optimal control problem is solved over a finite horizon at each sampling point. The optimized control input is applied to the process, and the optimization problem is solved again at the next sampling point.

The NMPC shares the same basic philosophy as the linear MPC. However, it has some major differences such as the explicit use of nonlinear models for the prediction and the inclusion of state and input constraints. Furthermore, the NMPC allows the online minimization of the cost function, and requires the availability of the state variables for the prediction either through measurements or estimation.

The basic principles of the NMPC problem are described hereafter (Allgöwer and Zheng, 2012). We consider the following discrete-time state-space nonlinear system:

$$\begin{aligned} \xi_{k+1} &= \mathcal{F}\left(\xi_k, F_{in_k}\right) \\ z_k &= H\xi_k \end{aligned} \tag{8.1}$$

where $\xi_k \in \mathbb{R}^{n_\xi}$ and $z_k \in \mathbb{R}^{n_z}$ are the system state and output vectors at the time step k, respectively. $F_{in_k} \in \mathbb{R}^{n_F}$ the input. $\mathcal{F}$ is the nonlinear transition function, and H is the measurement matrix.

A general formulation of the NMPC problem consists in minimizing the quadratic cost function Φ at the instant k over a finite moving horizon:

$$\min_{\xi_k \ldots \xi_{k+N_p-1}, F_{in_k} \ldots F_{in_{k+N_c-1}}} \Phi\left(\check{\xi}_\star, F_{in_\bullet}\right) \tag{8.2}$$

$$\text{s.t} \begin{cases} \check{\xi}_{k+i+1} = \mathcal{F}\left(\check{\xi}_{k+i}, F_{in_{k+i}}\right), \text{ for } i = \overline{0, N_p - 1} \\ \check{\xi}_k = \xi_k \end{cases} \tag{8.3}$$

where

- Φ is the cost function evaluated at instant k.
- N_p and N_c represent the prediction horizon and the control horizon respectively. They are not necessarily equal (i.e. $N_p \geq N_c \geq 1$) in order to reduce the computational load.
- $\check{\xi}_{k+i}$ is the predicted state vector at time instant $k+i$, $i = \overline{0, N_p - 1}$.
- ξ_k is the initial condition which is the plant state at time instant k.

The classical NMPC strategy is implemented in a moving horizon framework. At current time step k, the optimization problem (equations (8.2) and (8.3)) is initialized by the current state vector ξ_k and solved over the prediction horizon N_p. The first element of the resulting input vector F_{in_k} is applied to the system (8.1). At the next sampling time $k+1$, the optimization problem is solved again with the updated state vector ξ_{k+1} as the initial condition.

The NMPC is a model-based control strategy. A mathematical representation of the system is required to predict the evolution of each state variable. Therefore, the efficiency of the control scheme is directly linked to the model used for the

predictions. The choice of the model structure and the parameter identification procedure have a decisive role in the controller performance.

In the case of trajectory and set-point tracking problems, the cost function Φ to be optimized is a quadratic nonnegative function measuring the distance between the predicted model state $\check{\xi}$ and a determined known reference sequence ξ^{ref} over the horizon window. Another quadratic function weighting the control effort is added to achieve the control objective. The general expression for the cost function in this case is given by the following:

$$\Phi\left(\check{\xi}_{\star}, F_{in_{\bullet}}\right) = \sum_{i=1}^{N_p} \left\| \check{\xi}_{k+i} - \xi_{k+i}^{ref} \right\|_{\alpha_i}^2 + \sum_{i=1}^{N_c} \left\| F_{in_{k+i-1}} - F_{ref_{k+i-1}} \right\|_{\beta_i}^2 \tag{8.4}$$

where

- ξ^{ref} represents the reference trajectory or setpoint chosen to specify a desired closed-loop behavior.
- F_{ref} represents the reference feeding trajectory corresponding to ξ^{ref}.
- $\alpha_i \geq 0 \in \mathbb{R}^{n_\xi \times n_\xi}$ and $\beta_i \geq 0 \in \mathbb{R}^{n_F \times n_F}$ are the weighting matrices for the state tracking error and the control inputs, respectively.

The tuning parameters are implemented to penalize the deviations of the predicted controlled state vector $\check{\xi}$ from the reference ξ^{ref}, as well as the control input deviations at every sampling point in the prediction horizon. These parameters are tuned to avoid instability and to reach the desired transient performance of the closed-loop.

In addition to the cost function, constraints on the states and inputs can be defined:

$$\xi_{min} \leq \xi_{k+i} \leq \xi_{max}, \quad i = \overline{0, N_p - 1}. \tag{8.5}$$

$$F_{min} \leq F_{in_{k+j}} \leq F_{max}, \quad j = \overline{0, N_c - 1} \tag{8.6}$$

$$\tag{8.7}$$

However, this addition increases the complexity of the optimization problem, and the solution cannot be obtained explicitly as in the unconstrained case. An illustration of the receding horizon principle in the SISO case is shown in Figure 8.1.

8.3 Nonlinear Model predictive control applied to *E. coli* cultures

Reminding the objective behind the control strategy, that is to drive the process closer to the boundary of the respiro-fermentative and respirative regimes. Similarly, to the previous chapter, we desire to maintain the acetate concentration at a

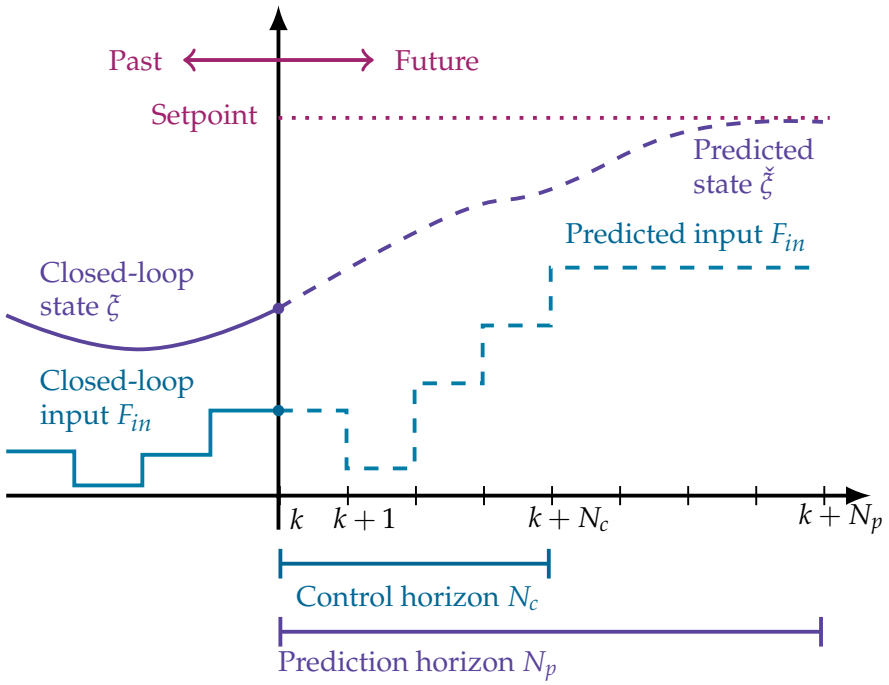


FIGURE 8.1: Illustration of the receding horizon principle

predefined low set-point to avoid its accumulation during the culture. To maintain the acetate concentration at a constant value, the bioprocess operates in the respiro-fermentative regime to compensate the dilution effect which tends to decrease the acetate concentration with time. Nevertheless, a low acetate concentration keeps the process close to the optimal metabolic edge (Dewasme et al., 2010; Rocha, 2003). In the proposed control strategy, the acetate concentration is regulated to the pre-defined set-point A_{ref}. This regulation is achieved by acting on the feed rate F_{in}, also constrained to track a pre-calculated reference trajectory F_{ref} in order to smooth the control behavior.

8.3.1 Determination of the reference feed-rate profile

In order to implement the NMPC strategy on the *E. coli* process, a reference feeding trajectory F_{ref} corresponding to the acetate reference must be pre-calculated based on the model equations defined by:

$$\dot{X} = (k_{X1}\mu_1 + k_{X2}\mu_2 + k_{X3}\mu_3)X - \frac{F_{in}}{V}X \tag{8.8a}$$

$$\dot{S} = -(\mu_1 + \mu_2)X - \frac{F_{in}}{V}(S - S_{in}) \tag{8.8b}$$

$$\dot{A} = (k_{A2}\mu_2 - \mu_3)X - \frac{F_{in}}{V}A \tag{8.8c}$$

$$\dot{V} = F_{in} \tag{8.8d}$$

Considering a constant acetate concentration ($A = A_{ref} \neq 0, \dot{A} = 0$), the differential equation (8.8c) leads to the following feed-rate:

$$F_{in} = \frac{\mu_A}{A} XV \Big|_{A=A_{ref}} \tag{8.9}$$

where $\mu_A = (k_{A2}\mu_2 - \mu_3) = k_{A2}(q_s – q_{scrit})$ in respiro-fermentative mode.

Generating the feeding profile using Equation (8.10) requires the availability of the substrate concentration. Assuming that there is no accumulation of the substrate concentration in the neighborhood of the optimal conditions, so that it is instantaneously consumed by the cells. The substrate is considered to be in quasi-steady state near the critical value ($S = S_{set}, \dot{S} = 0$).

From Equation (8.8b) we obtain the following equation:

$$F_{in} = \frac{\mu_S}{(S_{in} - S)} XV \Big|_{S=S_{set}} \tag{8.10}$$

where $\mu_S = -(\mu_1 + \mu 2) = q_S$ in respiro-fermentative mode. The substrate concentration S_{set} corresponding to the reference acetate concentration A_{ref} can then be calculated from the following equation:

$$\frac{q_S}{(S_{in} - S)} \Big|_{S=S_{set}} = \frac{\mu_A}{A} \Big|_{A=A_{ref}} \tag{8.11}$$

From equations (8.8a) and (8.8d) the following differential equation is obtained:

$$\frac{d(XV)}{dt} = \mu_X XV \tag{8.12}$$

hence:

$$XV = X_0 V_0 e^{\mu_X t} \tag{8.13}$$

where X_0 and V_0 are the initial conditions for the biomass concentration and the culture volume respectively, and $\mu_X = k_{X1}\mu_1 + k_{X2}\mu_2 + k_{X3}\mu_3$.

Finally, the expression of the reference trajectory F_{ref} for the pair (A_{ref}, S_{set}) can be generated by the following equation:

$$F_{ref} = \frac{\mu_A}{A} X_0 V_0 e^{\mu_X t} \Big|_{A=A_{ref}, S=S_{set}} = \frac{q_S}{(S_{\text{in}} - S)} X_0 V_0 e^{\mu_X t} \Big|_{A=A_{ref}, S=S_{set}} \tag{8.14}$$

8.3.2 Control design

In the following, a discrete-time equivalent expression of the continuous *E. coli* model (8.8) is used to implement the predictive control strategy. Considering a

constant sampling time T_s, the discrete model is defined as follows:

$$\begin{aligned} \xi_{k+1} &= \mathcal{F}\left(\xi_k, F_{in_k}\right) \\ z_k &= H\xi_k \end{aligned} \tag{8.15}$$

where

- $\xi_k = [X_k \; S_k \; A_k \; V_k]^T$ and $z_k = A_k$ are the discrete state and sampled output vectors at time kT_s, respectively.
- $\mathcal{F}$ is the nonlinear transition function
- $H = \begin{bmatrix} 0 & 0 & 1 & 0 \end{bmatrix}$ is the measurement matrix.
- F_{in} is the control input, parametrized using a piecewise constant approximation.

The trajectory tracking is achieved using the nonlinear model predictive control (NMPC) strategy. Based on the model (8.15) and the control objectives defined earlier, the NMPC objective cost can be defined as follows:

$$\Phi\left(\check{\xi}_\star, F_{in_\bullet}\right) = \sum_{i=1}^{N_p}\left(\check{A}_{k+i} - A_{ref_{k+i}}\right)^2 + \lambda \sum_{i=1}^{N_c}\left(F_{in_{k+i}} - F_{ref_{k+i-1}}\right)^2 \tag{8.16}$$

where $\check{A}$ is the predicted acetate concentration, and F_{ref} is the pre-calculated reference feeding profile. N_p and N_c are the prediction and control horizons respectively, and $\lambda > 0$ is the control penalty gain.

The NMPC problem formulation minimizes the quadratic cost function $\Phi\left(\check{\xi}_\star, F_{in_\bullet}\right)$ at the instant k and reads:

$$\min_{\check{A}_k \ldots \check{A}_{k+N_p-1}, F_{in_k} \ldots F_{in_{k+N_c-1}}} \Phi\left(\check{\xi}_\star, F_{in_\bullet}\right) \tag{8.17}$$

under constraints related to the predicted dynamics of the system:

$$\text{s.t.} \quad \check{\xi}_{k+i+1} = \mathcal{F}\left(\check{\xi}_{k+i}, F_{in_{k+i}}\right), \quad i = \overline{0, N_p - 1} \tag{8.18a}$$

$$\check{A}_{k+i+1} = H\mathcal{F}\left(\check{\xi}_{k+i}, F_{in_{k+i}}\right), \quad i = \overline{0, N_p - 1} \tag{8.18b}$$

$$0 \leqslant F_{in_{k+i}} \leqslant F_{max}, \quad i = \overline{0, N_c - 1} \tag{8.18c}$$

$$0 \leqslant \check{\xi}_{k+i} \leqslant, \qquad i = \overline{0, N_p - 1} \tag{8.18d}$$

A classical NMPC strategy follows the receding horizon principle, where the optimization problem (8.17 and 8.18) is solved online at instant k. The first element of the input vector is then applied to the system, and the optimization problem is

solved again with the updated initial state and input vectors at the next sampling point.

However, solving the constrained nonlinear optimization problem (8.17 and 8.18) raises two major difficulties. The first difficulty lies in the discretization of the continuous nonlinear system (8.8). The sampling period needs to be sufficiently small compared to the system's time constants such as the discrete model remains relevant compared to its the continuous counterpart. An appropriate choice of the sampling time in this case is in the order of 0.5-1 min, since the doubling time of *E. coli* is around 20 minutes and the response time is in the order of several tens of minutes. This leads to very long prediction horizons and an increase in the number of the decision variables of the optimization problem. Furthermore, it poses a practical problem for the on-line implementation of the control strategy due to the different operations executed between two sampling points (measurement, estimation, calculation of the control law).

The second difficulty is the nonlinear constraints (8.18), which increases the on-line computation time of the optimization procedure involved in the NMPC strategy.

In order to avoid to solve the constrained optimization problem, the optimization problem is transformed into a nonlinear programing problem (NLP) using the Control Vector Parametrization (CVP) technique presented in (Banga et al., 2005).

The CVP method is based on the discretization of the control variables only, and the control sequence is approximated by a piecewise constant function between two sampling instants. The advantage of this transformation is that it is no longer necessary to discretize the continuous model, and the predicted states are obtained by the integration of the continuous differential system (8.8).

Consequently, the sampling period required for the prediction can be much larger, and the constraints related to the state variables are removed from the optimization problem, since they are taken into account in an implicit way during the prediction of the states via the CVP method. Furthermore, the control horizon N_c can be chosen equal to the prediction horizon N_p, since the control vector is the only decision variable for the optimization problem. The CVP principle is illustrated in Figure 8.2.

The new formulation of the optimization problem is therefore given by:

$$\min_{v_k \dots v_{k+N_p-1}} \Phi\left(\check{\xi}_\star, F_{in_\bullet}\right) \tag{8.19}$$

Under the constraints:

$$0 \leqslant F_{in_k} \leqslant F_{max}, \quad k = \overline{1, N_p} \tag{8.20a}$$

In addition, a change of variable $F_{in} = e^v$ aims at removing the positivity constraints on F_{in} and improves the conditioning of the optimization problem. The

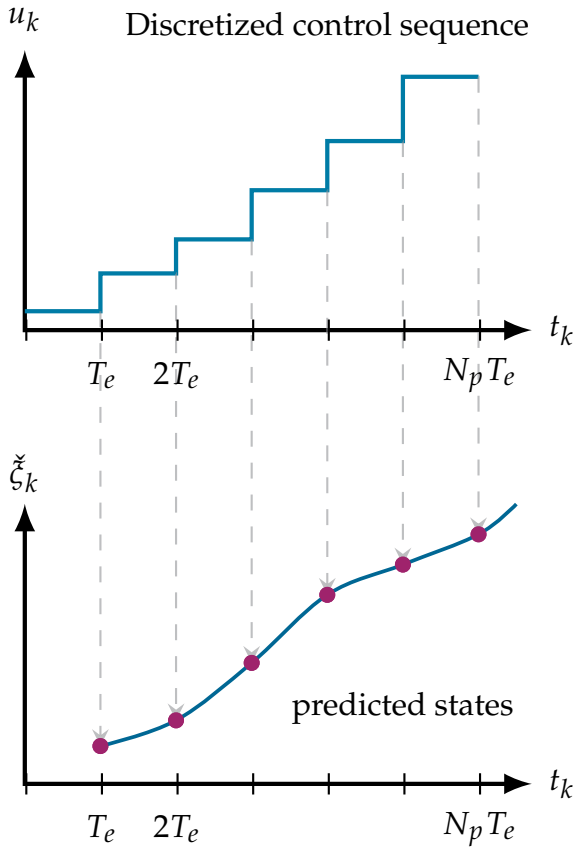


FIGURE 8.2: Principle of the CVP approach

corresponding cost function reads:

$$\Phi\left(\check{\xi}_{\star}, v_{\bullet}\right) = \sum_{i=1}^{N_p}\left(\check{A}_{k+i} - A_{ref_{k+i}}\right)^2 + \lambda \sum_{i=1}^{N_p}\left(e^{v^{ref}_{k+i-1}} - e^{v_{k+i-1}}\right)^2 \tag{8.21}$$

and the formulation of the NMPC problem is reduced to solving at each time kT_e:

$$\min_{e^{v_k} \ldots e^{v_{k+N_p-1}}} \Phi\left(\check{\xi}_{\star}, v_{\bullet}\right) \tag{8.22}$$

The constraint on the upper bound of the control input is removed from the optimization problem. The optimal input solution of the problem (8.18) is then limited to its upper bound before its application to the process.

This transformation of the initial optimization problem (8.17) to an unconstrained optimization problem is very beneficial for the implementation of the control strategy on the real process. A comparison between the classical and unconstrained NMPC formulations is presented in the simulations section.

In the following, the UKF estimated state $\hat{\xi}_k$ is used in the NMPC scheme.

8.4 Numerical simulations

The combined UKF-NMPC strategy is first assessed through numerical simulations, considering the fed-batch *E. coli* model presented in chapter 2. The initial plant conditions as well as UKF and controller parameters are given in Table 8.1.

Estimates initial conditions are randomly chosen using a normal distribution centred at the nominal state values with a standard deviation of 15%. The covariance matrix of measurement noise R depends on the quality of the turbidimetric probe signal. On the other hand, the covariance matrix of the model noise Q represents the confidence level in the model.

TABLE 8.1: UKF covariance matrices, sigma points tuning parameters, and initial conditions

Initial conditions		
X_0	1.0	g/L
S_0	0.1	g/L
A_0	1	g/L
V_0	3.5	L
S_{in}	500	g/L
UKF parameters		
σ_X	0.01	g/L
σ_S	0.1	g/L
σ_A	0.1	g/L
σ_V	0.001	g/L
r_X	0.01	g/L
P_0	$10^{-4} \times \mathbb{I}_4$	g/L
α	1	-
β	2	-
κ	0	-
Control parameters		
N_p	10	-
λ	0.05	-
Sampling time		
T_s	3 min	-

The acetate concentration regulation achieved by the NMPC is assessed, first considering a perfectly known model (no parametric uncertainty), and then a more realistic case with plant-model parameter mismatch. In both cases, the acetate concentration is supposed to be measured with a measurement noise level of $0.02 g/L$.

The optimization is performed by the LSQNONLIN optimizer under MATLAB R2018b environment on a 3.2 Ghz PC with 8Gb of ram. The optimization problem is solved using the '*trust-region-reflective*' algorithm .

Figure 8.3 shows the closed-loop profiles of the state and control variables for the ideal model case (no modeling errors or parameter uncertainties), with prediction horizon $N_p = 10$ and penalty parameter equal to $\lambda = 0.05$. λ was tuned by trial and error to prevent excessive deviation of F_{in} from the reference feeding profile F_{ref}.

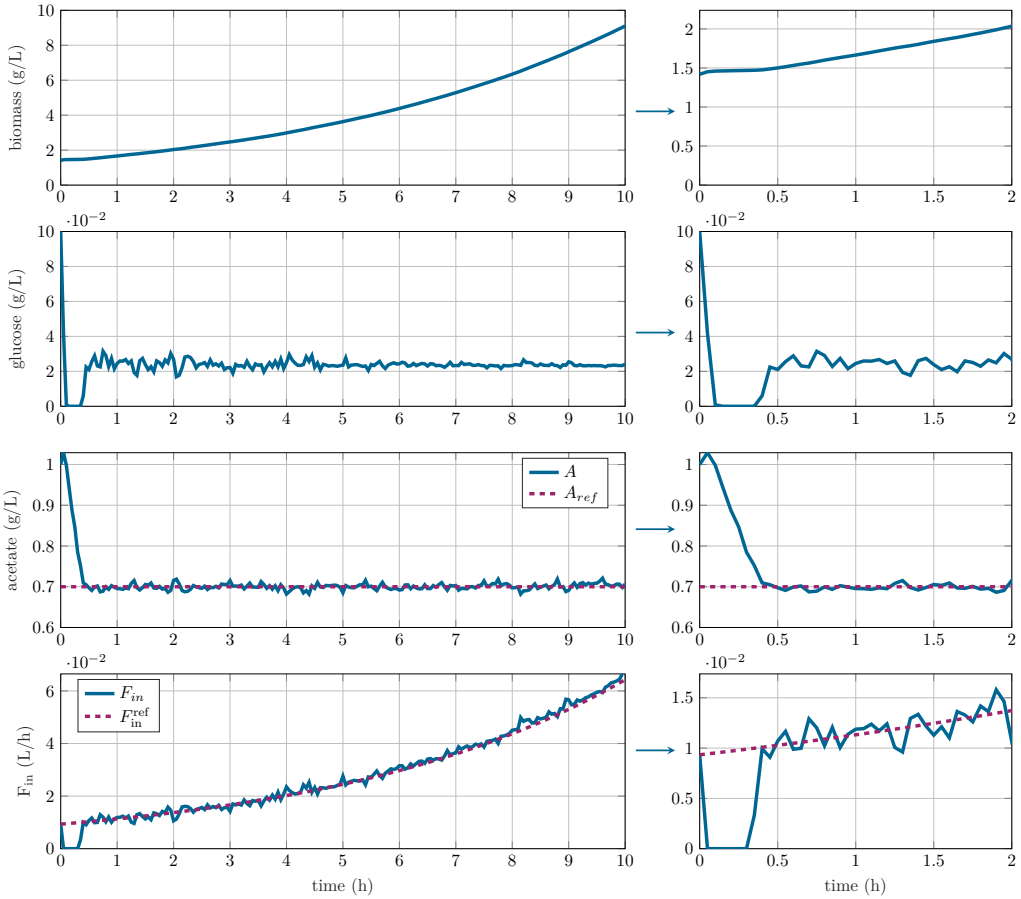


FIGURE 8.3: Ideal model case. Left: Closed loop evolution of the state variables, X, S, A, and feed flow rates F_{in} and F_{ref} with prediction horizon $N_p = 10$, $\lambda = 0.05$ and a noise standard deviation of 0.02 $[g/L]$. Right: Zoom over 2h.

It may be observed in Figure 8.3 that the regulation of the acetate concentration is achieved successfully. The biomass growth follows a typical exponential trajectory, and the substrate quickly reaches an a priori unknown and constant steady-state, assumed to be greater than the critical level since, to keep the acetate concentration constant, the cells must evolve in respiro-fermentative regime, compensating the dilution effects from the feed rate by a small acetate production.

The convergence to the steady-state is achieved in 25 minutes and the biomass at $t = 10$ h is 32.5 g.

In view of analyzing the NMPC robustness to plant-model mismatch, a simulation is performed by varying the most influential parameters on the acetate concentration. The parametric sensitivity study presented in appendix D shows that the most influential parameters on the acetate concentration A are $q_{S\,\max}$, $q_{O\,\max}$, k_{X2}, and k_{A2}.

During the simulation, the values of these parameters are randomly chosen following a normal distribution centred at their nominal values with a standard deviation of 15%. A set of 100 Monte Carlo (MC) simulations is performed using

the same control settings as previously for the ideal model (Table 8.1).

Figure 8.5 shows the histograms of the runs as functions of the randomly chosen values. The corresponding closed-loop profiles are represented in Figure 8.4. The acetate concentration is maintained at the reference in all runs, with a mean value of 0.7 g/L. The tracking error ($A_{ref} - A$) has a RMSE of 0.0314 g/L, which is still satisfactory. The convergence to the desired steady-state is still 25 minutes in average.

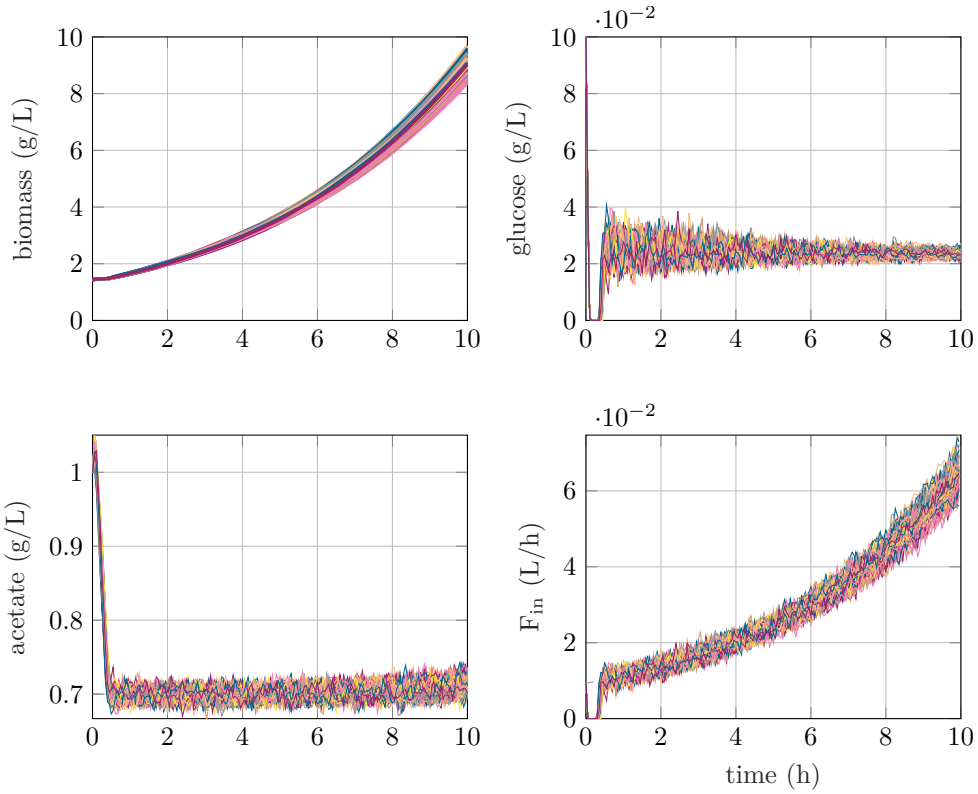


FIGURE 8.4: Plant/model mismatch case. Closed-loop evolution of the state variables X, S, and A with noise standard deviation of 0.02 g/L. Profiles of the 100 Monte Carlo experiments.

Histograms of the final biomass productions and productivities are also shown in Figure 8.5. Biomass production is, in average, slightly lower than in the ideal case (31.2g), nevertheless it stays above 30g in 96% of the MC experiments (Figure 8.5). In most of the cases, biomass productivity is higher than 90% of the maximal attained level.

To illustrate the advantages of the CVP method, a series of simulations is performed comparing the computation time required to solve the optimization problems before and after using the CVP technique.

The simulation consists in regulating the acetate concentration to the defined setpoint of $A_{ref} = 0.7$ g/L, with the same initial values for the state variables and control parameters described in Table 8.1.

Three methods were compared: The initial constrained NMPC problem (8.17), the NMPC problem with the CVP method (8.19), and the unconstrained NMPC

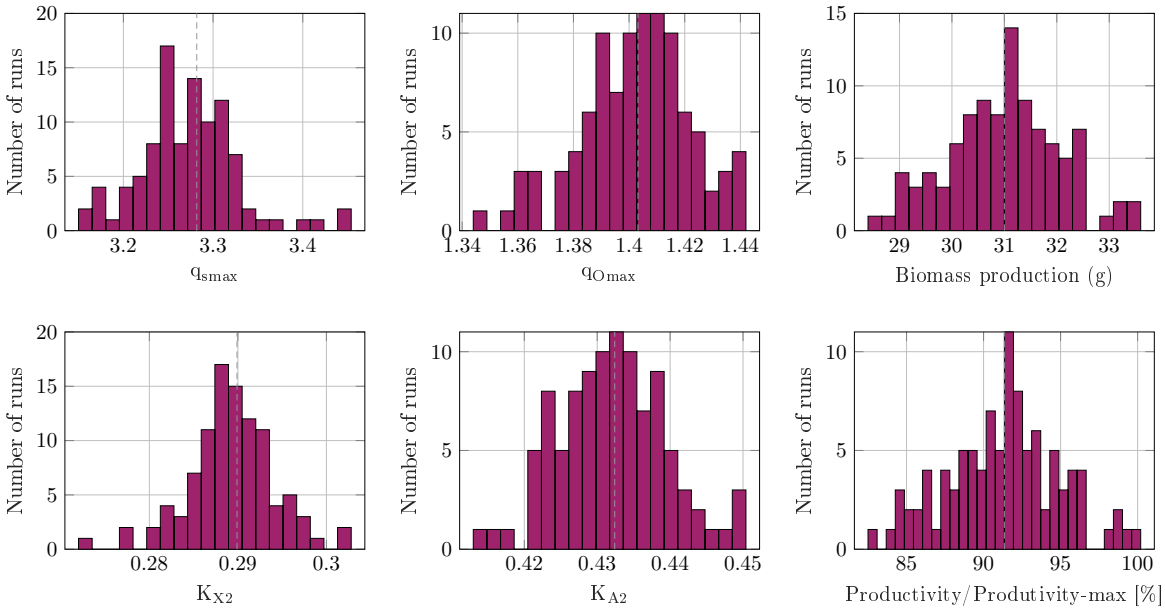


FIGURE 8.5: Histograms of the plant parameters $q_{S\,max}$, $q_{O\,max}$, k_{X2}, and k_{A2}, the biomass production and productivity during the MC simulations.

problem (8.22) with the variable change (VC) $F_{in} = e^{v}$. A total of 10 simulations is performed for each method, and the computation time δt required for solving the optimization problem for one time step (kT_e) is shown in Table 8.2.

The results show that the introduction of the CVP method offers a significant gain in the computation time, while achieving the same control problem objectives. Even though the variable change (VC) gives a close performance, it improves the conditioning of the optimization problem. This shows the advantage of the proposed method for the on-line implementation of the NMPC on bioprocesses. The low computation time offers the opportunity to consider lower sampling periods, and to couple the NMPC with other components in a more advanced control structure.

TABLE 8.2: Comparison between the computation time required for solving the different NMPC problems

	$\min(\delta t)$ (s)	$\max(\delta t)$ (s)	$\text{mean}(\delta t)$ (s)
Constrained NMPC	2.423	11.779	5.863
NMPC + CVP + VC	0.018	0.733	0.037
NMPC + CVP	0.019	0.793	0.038

The NMPC is now coupled to the UKF estimator, considering the estimated acetate concentration as the regulated variable. The same parameter uncertainty as well as biomass measurement noise level are considered. As shown in Figure 8.6, the UKF is able to correctly estimate the acetate concentration in the presence of

model and measurement noise, and the acetate regulation is performed accurately as both the estimated and actual states converge to the desired acetate value.

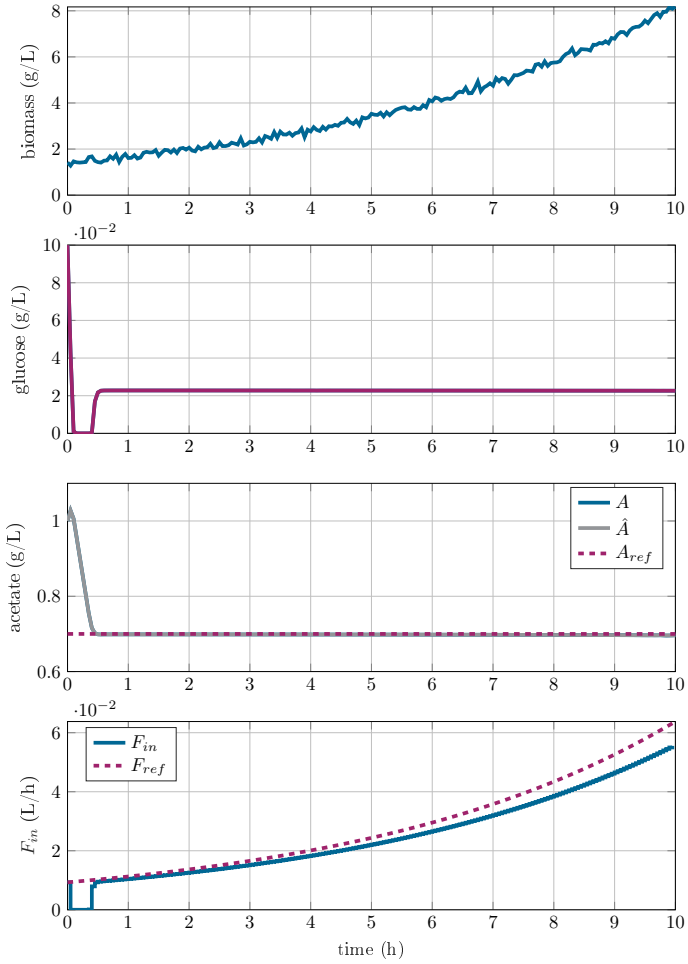


FIGURE 8.6: Coupled UKF-NMPC numerical simulations with measurement noise and parametric uncertainties.

In order to compare the performance of the NMPC and the GMC strategies, simulations are performed testing their performance in regulating the acetate concentration at $A_{ref} = 0.7$ g/L. The initial conditions for the simulation are the same of the previous simulations given in Table 8.1. The GMC parameters are the same used in the previous chapter ($G_1 = 5.61$, $G_2 = 9.55$) (Equation (7.24)). These parameters were determined using the robust design method, for a maximal parameter variation of 15%. Figure 8.7 shows the evolution of the state variables and the feed-rate using the two control strategies: NMPC and GMC, and Figure 8.8 shows a zoom over the first two hours.

The simulation shows that both control strategies achieved good regulation of the acetate concentration at the reference value $A_{ref} = 0.7$ g/L. As seen in the previous chapter, the GMC strategy is able to achieve the control objective. However, it induces a longer transient phase, a higher overshoot, and a longer convergence

time compared to the NMPC. Furthermore, the biomass quantity obtained with the NMPC strategy is marginally higher than that obtained using the GMC.

In conclusion, both the NMPC and the GMC strategies achieve the control objectives adequately, however the NMPC is better than the GMC regarding the transient performance, and this has a direct effect on the final biomass quantity.

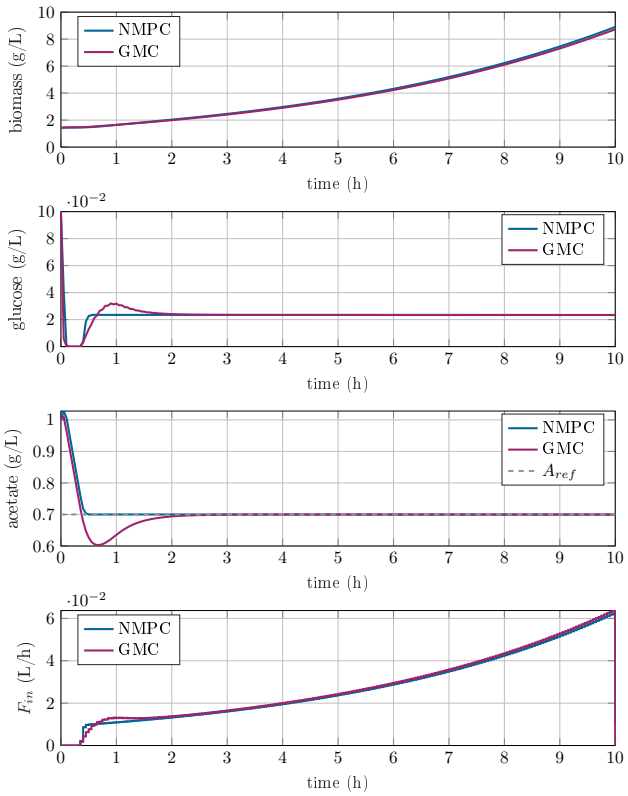


FIGURE 8.7: Comparison of the NMPC and GMC strategies. Plot of the state variables and the control inputs.

8.5 Experiment results and discussion

To validate the combined NMPC control strategy and the estimation of the acetate and glucose concentrations by the UKF, two fed-batch experiments are carried out using the strain *E. coli* BL21(DE3) according to the protocol described in chapter 4. The UKF parameters during the experiments are given in Table 8.3, and the UKF algorithm is presented in appendix A.

The NMPC parameters used in both experiments are given in Table 8.4. The prediction horizon and penalty parameter λ were tuned in simulation by trial and error in order to achieve the control objectives while preventing excessive deviation of F_{in} from the reference feeding profile F_{ref}.

The evolution of the biomass, glucose, acetate concentrations, and the feed flow rate in both experiments are illustrated in Figure 8.9, and the operating conditions are shown in Figure 8.10.

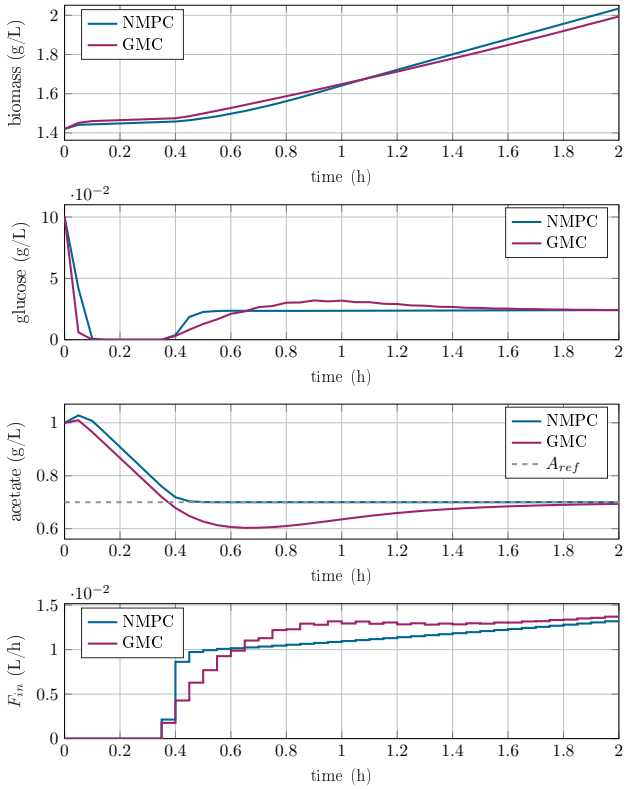


FIGURE 8.8: Comparison of the NMPC and GMC strategies. Plot of the state variables and the control inputs. Zoom over 2h.

TABLE 8.3: UKF parameters used in the experiments

UKF parameters		
σ_X	0.01	g/L
σ_S	0.1	g/L
σ_A	0.1	g/L
σ_V	0.001	g/L
r_X	0.01	g/L
$\mathbf{P}_0$	$10^{-4} \times \mathbb{I}_4$	g/L
α	1	-
β	2	-
κ	0	-

The first experiment is carried in two phases, a batch followed by a fed-batch. The batch phase is performed to reach a minimal biomass concentration of 1 g/L and ensure that the cells are in the exponential phase when initiating the regulation and injecting the feed media. The UKF estimation is launched during this phase after initializing the state vector.

The fed-batch phase starts after the near depletion of glucose. Even though the glucose concentration measurement is not available, on-line flags of the glucose depletion can be detected, consisting of a sudden drop in the stirring speed due to

TABLE 8.4: NMPC control parameters used in the experiments

NMPC parameters	
Sampling time	$T_s = 3$ min
Acetate reference	$A_{ref} = 0.5 - 1.5$ g/L
NMPC parameters	N_p= 10, λ= 0.05

the low cell demand for oxygen (Figure 8.10), coupled with an increase in the pH value due to the decrease of CO_2 released by the cells.

The NMPC controller is launched after setting up the acetate reference and the control parameters. The controller outputs an exponential feeding profile, and the estimated acetate concentration converges to the desired reference in less than 30 min and remains in the range 0.33-0.5 L/h until the end of the culture. At the same time, the substrate concentration remains in a quasi-steady state.

After the end of the first experiment, a culture refresh is performed. A volume of the culture is extracted using a peristaltic pump, leaving 500 mL of the culture of volume. Then, 3L of an autoclave sterilized HDF medium is injected by a peristaltic pump, and the batch phase of the second experiment starts.

The second experiment follows the same protocol as the first one, the batch phase lasts for 2 h, and the glucose is quickly consumed by the cells. The controller is launched as soon as the stirring drops significantly, indicating a drop in glucose oxidation. The acetate concentration set-point is fixed at 1.5 g/L, the feeding follows an exponential curve, and the acetate concentration converges to the imposed reference. At t=12 h, a set-point change is introduced and A_{ref} is set equal to 0.7 g/L.The NMPC controller adapts the feed-rate according to the new reference, and the estimated acetate concentration tracks the new reference in 20 min.

The NMPC-UKF strategy performance is highly satisfying. The UKF accurately estimates the acetate and glucose concentrations in the presence of measurement noise on the biomass concentration. The mean square errors of the substrate and acetate concentrations are $\overline{e_S} = 0.089$ g/L and $\overline{e_A} = 0.068$ g/L respectively, which is consistent regarding the sensitivity of the measurements and the noise levels (0.1 g/L). The NMPC controller regulates the estimated acetate concentration to the fixed set-point, and the convergence is achieved in 20 min. The acetate concentration remains at an acceptable range during the fermentation time, and the culture conditions are well suited for biomass growth.

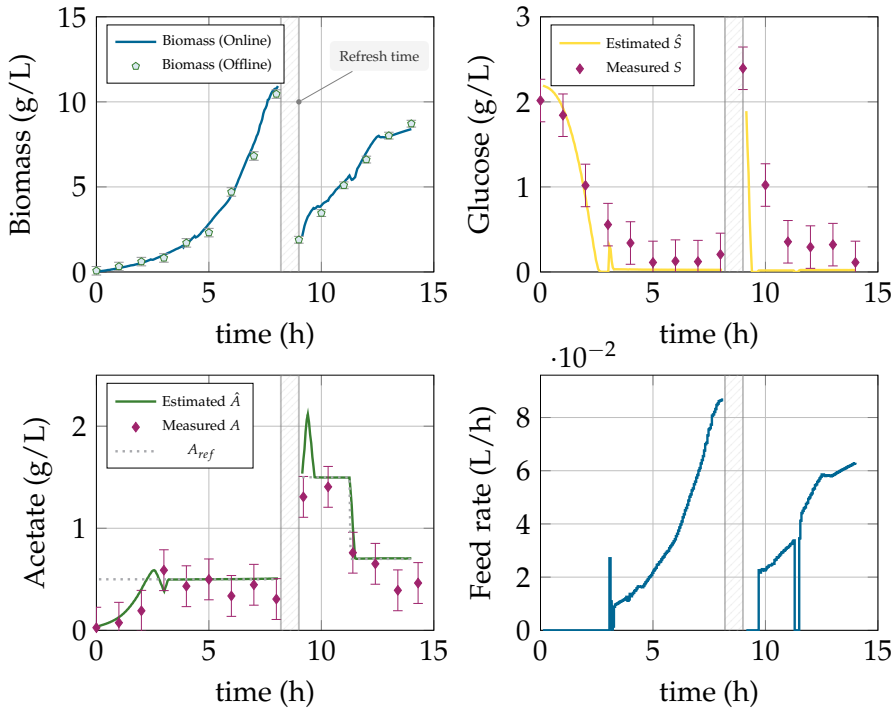


FIGURE 8.9: Experimental results: Time evolution of the measured biomass, glucose, acetate concentrations (offline and online estimation), and feed-rate

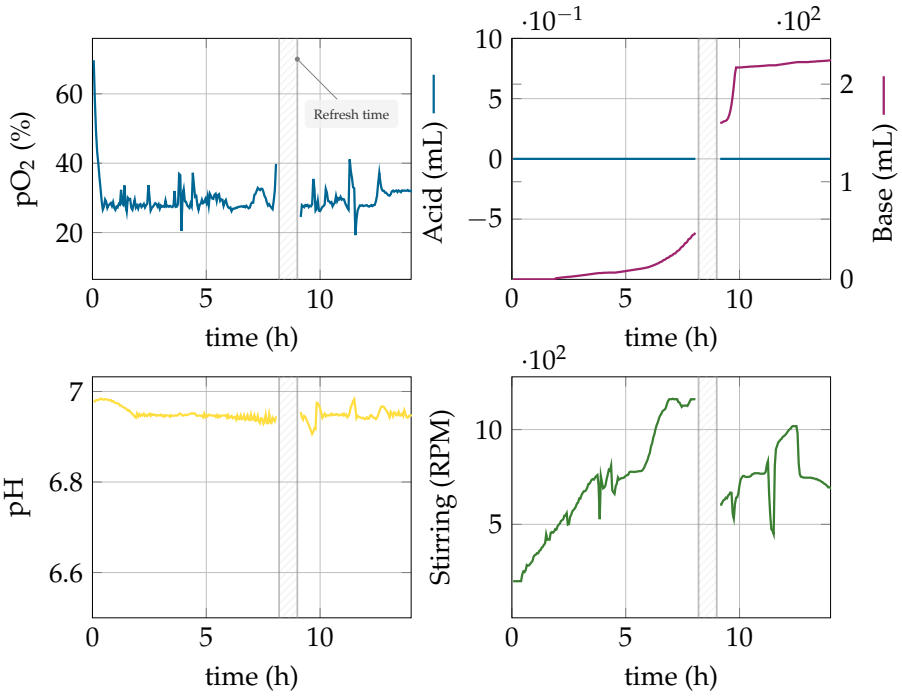


FIGURE 8.10: Experimental results: Time evolution of the pO_2, acid and base concentrations, pH, and stirring

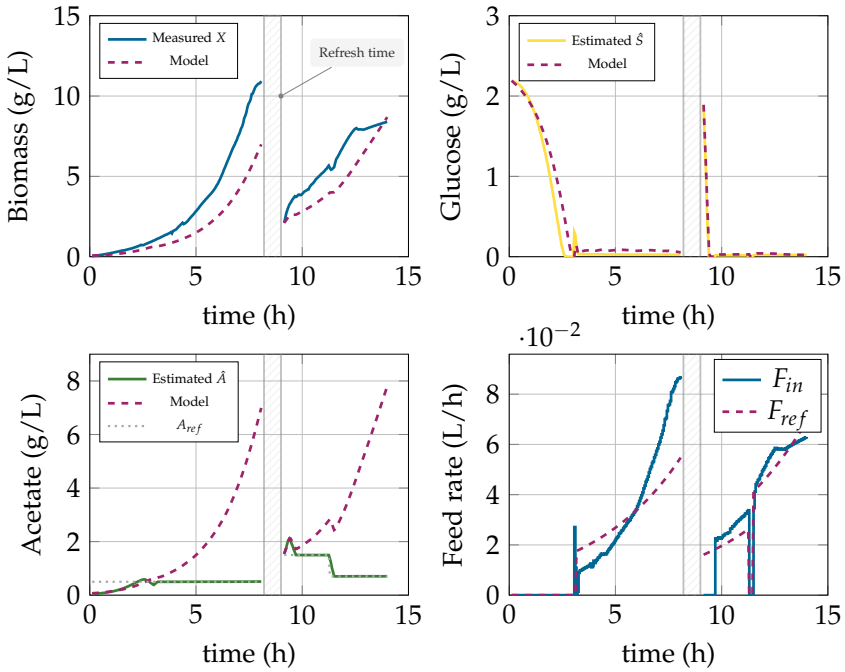


FIGURE 8.11: Comparison between the experimental results and the process model predictions

In order to compare the deviation of real process response from the model predictions, a simulation is performed using the experimental data and the process model used in the control and estimation algorithms.

The model is simulated in open-loop with the same operating conditions and the initial values for the state variables from the experiments. The obtained experimental feed-rate is applied to the model input.

The results presented in Figure 8.11 show a deviation between the model predictions and the measured and estimated state variables. The experimental feed-rate deviates from the reference trajectory, indicating the presence of a model mismatch. On the other hand, the acetate concentration is accumulated in the model simulation due the high glucose concentration resulting from the high feed-rate levels.

Consequently, the acetate presence leads to a lower growth rate than in the experiments and a low biomass concentration. The experimental and simulation results show the efficiency of the proposed predictive strategy to achieve the control objectives despite the model mismatch.

8.6 Comparative study

In order to evaluate the performance of the control strategies presented in this thesis, a simulation study is performed and presented in this section. The three control methods are compared under the same initial operating conditions and a similar control setting.

The NMPC and robust GMC (GMC-A) controllers are set to track an acetate reference of $A_{ref} = 0.7$ g/L. On the other hand, the biomass growth rate regulation using the GMC strategy (GMC-X) is set to track a reference profile corresponding to this acetate concentration value. To this end, the glucose concentration S_{set} corresponding to A_{ref} is calculated using Equation (8.11). Then, using the pair (S_{set}, A_{ref}), the reference biomass growth rate μ_{set} is calculated from kinetic model equations.

The simulation is performed over 10 h assuming that the state vector is available for measurement, with an added measurement noise of 0.05 g/L and 0.02 g/L on the biomass and acetate concentrations. The initial conditions, reference variables, and control parameters values are summarized in table Table 8.5.

TABLE 8.5: Initial conditions and control parameters

Initial conditions		
X_0	1.0	g/L
S_0	0.1	g/L
A_0	1	g/L
V_0	3.5	L
S_{in}	500	g/L
Reference variables		
A_{ref}	0.01	g/L
S_{set}	0.0234	g/L
μ_{set}	0.1925	h^{-1}
F_{max}	0.2	L/h

NMPC parameters		
N_p	10	-
λ	0.05	-
GMC-X parameters		
G_1	= 2.25	-
G_2	= 9	-
GMC-A parameters		
G_1	= 5.61	-
G_2	= 9.55	-
Sampling time		
T_e	3	min

First, the control strategies are compared in the ideal model case, where all the model parameters are assumed to be perfectly known. Then, 500 Monte Carlo simulations are performed considering model uncertainties. The model parameters are randomly chosen with a maximal variation of $\pm 30\%$ around their nominal value following a normal distribution.

In the ideal model case (Figure 8.12), the biomass trajectories are almost identical for the three methods. This behavior is expected since the simulation is designed to have the same setting for the three methods.

The acetate and glucose concentrations, on the other hand, have different dynamics. The NMPC and the GMC-A strategies regulate the acetate concentration at the reference A_{ref}. On the other hand, the biomass regulation (GMC-X) leads to a decrease in the acetate concentration to reach a zero value after 7.5 hours.

The feed rates have similar behavior, where the main difference between the controllers is the delays due to the transitory phase. The acetate dynamics affect the metabolic operating mode and the biomass evolution and, consequently, the required feed rate to reach the desired growth rate.

In the model mismatch case, the parameter variation leads to different responses of the controllers. An example of a mismatch simulation is shown in Figure 8.13. The biomass profiles in this case are different for each control strategy.

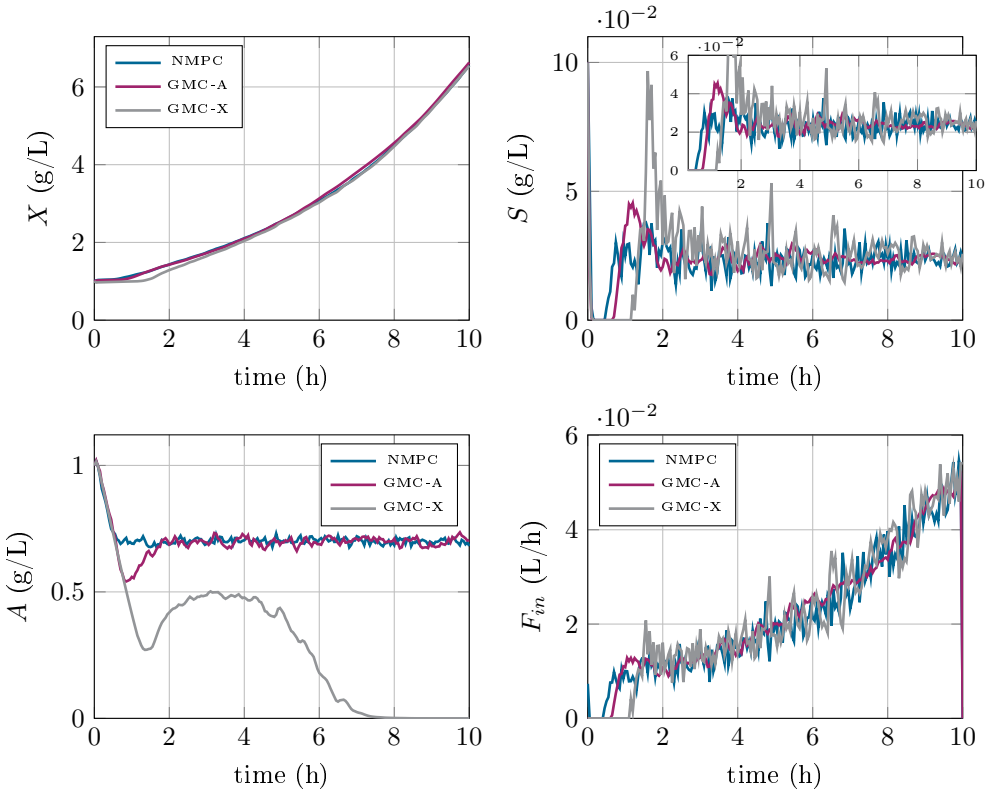


FIGURE 8.12: Evolution of the state variables in the ideal model case for the three control methods

The GMC-A and NMPC have higher final biomass concentration compared to the GMC-X with an advantage for the NMPC. This is due to the difference in the control objective definition.

The biomass regulation targets a specific growth rate and manages to reproduce the biomass profile in the case of model mismatch, leading to similar performance compared to the nominal case. However, since the optimal metabolic operating point is linked to the critical substrate concentration and the critical oxidative capacity, which is by nature uncertain.

The targeted growth rate is usually defined by setting it 5-10% lower than the maximal growth rate to avoid acetate accumulation. In the case where the maximal value changes due to the model mismatch, targeting μ_{set} leads to a lower metabolic performance due to underfeeding.

This strategy is useful for reproducing experiments and results with high yields and final product quality ("golden batches"), thanks to its consistent results. Furthermore, the operational advantage of the method is the availability of reliable biomass probes making its implementation fairly easy. However, this requires a knowledge of the bioprocess and a sufficient amount of historical data. The biomass growth regulation strategy is not the best fit for productivity maximization, especially in the case of uncertain optimal conditions or unreliable models.

On the other hand, acetate regulation offers a solution to this problem. Targeting an acetate concentration close to zero maintains the culture closer to the

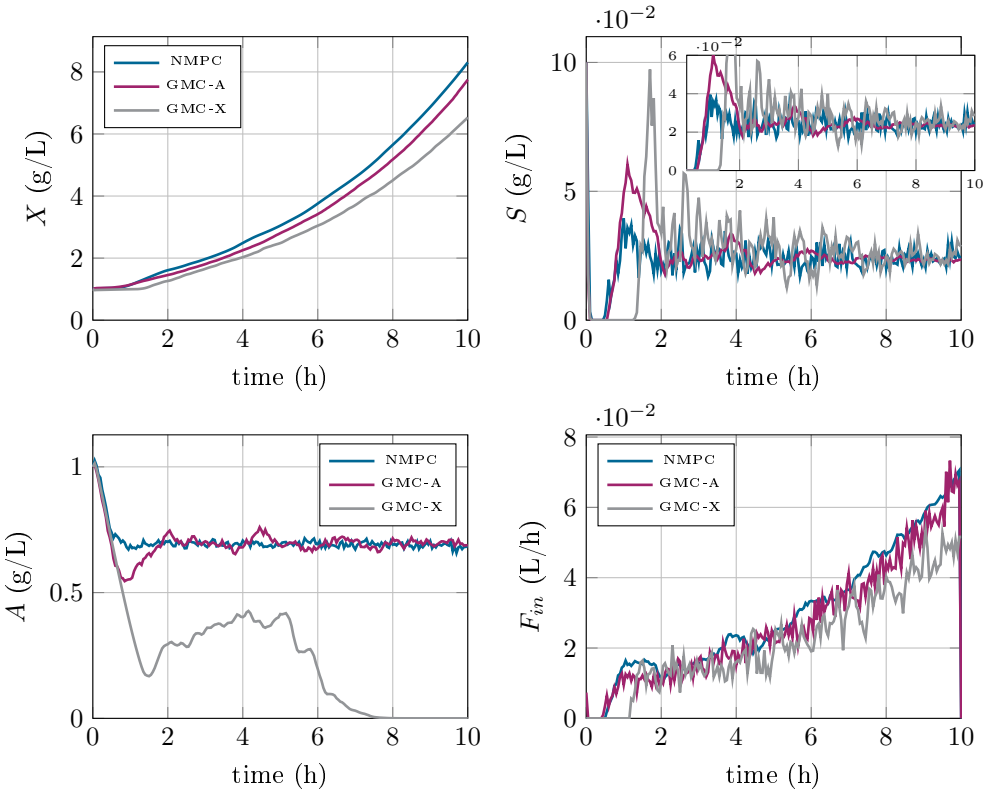


FIGURE 8.13: Evolution of the state variables in the model mismatch case for the three control methods

optimum with consistent metabolic performance. In the case where the optimum is unknown and uncertain, acetate fermentation is directly linked to the oxidative capacity, and therefore the metabolic performance is better thanks to the adaptation to the new maximal growth rate value.

The NMPC strategy leads to higher biomass quantity compared to the GMC-A. This is mainly due to the longer transitory phase in the GMC case and the more frequent fluctuations around the regulated acetate concentration value. This result can be explained by the fact that the NMPC uses the nonlinear model implicitly in the prediction phase, while the GMC algorithm includes linearization of the nonlinear dynamics leading to tracking errors and unpredicted dynamics.

The results of the Monte Carlo simulations give a performance overview of the control methods. The biomass productivity and productions over 10 h during the 500 runs are summarized in tables 8.6 and 8.7.

It can be seen that the NMPC outperforms the GMC controllers on average and gives higher productivity and biomass quantities. It can also be seen that acetate regulation has a higher average and maximal value. The biomass regulator targets a specific growth rate by design, and thereby, when the process is able to reach higher growth rates, the biomass regulation maintains the growth rate closer to the specified reference, leading to a lower biomass productivity. As expected, all the developed controllers lead to better results than the open-loop trajectory with an average improvement of 20% in the biomass productivity.

TABLE 8.6: Biomass productivity of the control methods during 500 MC simulations

Method	Mean	Min	Max	Unit
Open-loop	0.0161	0.0079	0.0183	g/(h · g of substrate)
GMC-X	0.0187	0.0176	0.0208	g/(h · g of substrate)
GMC-A	0.0190	0.0184	0.0206	g/(h · g of substrate)
NMPC	0.0192	0.0185	0.0206	g/(h · g of substrate)

TABLE 8.7: Biomass production of the control methods during 500 MC simulations

Method	Mean	Min	Max	Unit
Open-loop	22.864	22.0472	26.776	g
GMC-X	25.3944	18.4881	39.0349	g
GMC-A	25.7305	20.7225	56.4224	g
NMPC	26.6117	19.4120	57.1957	g

TABLE 8.8: Root mean square errors of the control methods during 500 MC simulations

Method	Mean	Min	Max	Unit
Open-loop	0.941	0.139	1.3821	g/L
GMC-X	0.1400	0.0744	0.8976	g/L
GMC-A	0.0623	0.0503	0.1531	g/L
NMPC	0.0560	0.0466	0.0888	g/L

TABLE 8.9: Computation time (between 2 sampling steps) of the control methods during 500 MC simulations

Method	Mean	Min	Max	Unit
Open-loop	0.0165	0.01413	0.0401	$(\times 10^{-3})$ s
GMC-X	0.0265	0.1413	0.0401	$(\times 10^{-3})$ s
GMC-A	0.0274	0.2854	0.0535	$(\times 10^{-3})$ s
NMPC	0.0210	0.8095	0.0968	s

Concerning the tracking performance, the root-mean-square errors of the controlled variable $(y–y_{ref})$ are given in Table 8.8. It can be seen that the NMPC has the lowest tracking error compared to the GMC-A regulation. The acetate control strategies provide better tracking performance than biomass regulation due to the exponential evolution of the biomass concentration.

As for the computation load (Table 8.9), The GMC methods have the lowest time required to calculate the control law. This result is expected since the GMC law is computed by a simple algebraic equation and does not require the integration of a nonlinear differential equation.

As for the NMPC, the computation time can sum up to 0.8 seconds between

sampling points, which is a good result compared to the sampling period ($T_s = 3$ min). However, this result is obtained by implementing a CVP method to simplify the optimization problem making it better conditioned. Considering constraints, a complex cost function or a robust variation of the NMPC (min-max NMPC, for example) can significantly raise the computation load. A computation time of the order of several seconds combined with the different delays in the LabView environment can sum up to several minutes. This may be a challenge, especially if a lower sampling time is required for the NMPC prediction to be accurate. Implementing time-saving strategies in the NMPC schemes is a necessary step before investigating advanced predictive variations.

8.7 Conclusions

In this chapter, an implementation of the nonlinear model predictive control strategy on the fed-batch *E. coli* process is presented. This strategy is experimentally implemented and validated on lab-scale fed-batch *E. coli* BL21(DE3) cultures. The goal behind the control scheme is to maximize the biomass production and favorize cell growth, by avoiding overflow metabolism and the accumulation of the acetate during the culture.

The control objective is to regulate the acetate concentration to a low value. This reference needs to be as close to zero as possible in order to maintain the process closer to the metabolic edge between the operating regimes. On the other hand, the reference value should be high enough to account for the measurement and estimation sensitivities, and to avoid cell stress caused by metabolic switches in the neighborhood of optimal operating conditions. Furthermore, the feed-rate is also set to track a specified feeding trajectory, precalculated offline depending on the couple (A_{ref}, S_{set}) in order to prevent excessive deviations from the optimal trajectory.

The NMPC strategy has several advantages, like the implicit use of the nonlinear model equations for the prediction, and the inclusion of state and control constraints. However, this comes with an additional computational cost, which is a major hurdle for experimental implementation of the NMPC.

To this end, the optimization problem was transformed into a nonlinear programming problem using the control vector parametrization (CVP) approach, and a variable change. This transformation allowed a significant reduction in the complexity of the optimization problem, and the obtention of an unconstrained nonlinear problem.

Simulations were achieved testing the proposed NMPC strategy to achieve the control objectives assuming an ideal model with no parameter uncertainty. The results showed a good performance to track the desired acetate concentration with relatively short transients. The feed-rate also follows the pre-calculated trajectory very well. The robustness of the strategy was tested using a set of Monte Carlo simulations, assuming a random variation of the model parameters around their nominal values. The results showed that the NMPC is able to achieve the control

objectives and regulate the acetate concentration to the low reference value, with a small deviation from the reference feeding trajectory.

Finally, the NMPC strategy was coupled to the UKF estimator. The acetate concentration is estimated by the pre-tuned UKF based on the biomass signal. The NMPC is then set to regulate the estimated acetate concentration. Results showed that the same level of performance is obtained by the combined control and estimation algorithms.

The performance of the NMPC was compared to that obtained with robust generic model control (GMC). The comparison was performed for identical scenarios, under nominal model assumption. The NMPC outperformed the GMC in achieving the control objectives of trajectory tracking. The NMPC strategy provide better transient response, and a higher biomass quantity. Furthermore, the NMPC design includes the explicit use of the nonlinear model for the prediction, while the GMC strategy requires a model reduction to avoid using the kinetic structure, and to prevent deviations of the feed-rates leading to excessive and high values.

The main contribution of this chapter is the experimental implementation of the proposed NMPC strategy. The NMPC algorithm was deployed on the control structure presented in the previous chapters, to regulate the estimated acetate concentration in a lab-scale fed-batch culture of *E. coli* BL21(DE3). The experiments were conducted according to the protocol presented in chapter 3. The results validated the combined NMPC-UKF strategy in regulating the acetate concentration despite the model mismatch. Furthermore, the experiments validated that regulating the acetate concentration to a low level leads to higher biomass concentrations and lower cultivation time, and consequently higher biomass productivity.

A possible improvement of the proposed NMPC strategy is to consider a cost function that includes a term accounting for bounded parameter uncertainty. This transformation may improve the NMPC robustness with respect to model mismatch especially when implemented on other strains with higher level of uncertainty. This may lead to an increase in the computational burden, but the proposed transformation using the CVP method provides a significantly low computation time compared to the sampling time, allowing a margin to include more objectives.

Chapter 9

General conclusions and perspectives

9.1 Conclusions

The work presented in this thesis aims at developing and implementing a practical solution to overflow metabolism in fed-batch *Escherichia coli* cultures through closed-loop control and estimation strategies. The objective behind the regulation of the bioprocess is to maximize bacterial growth and drive the culture near the optimal operating conditions. Another objective is to evaluate and test control and estimation algorithms developed for the bioprocess under realistic conditions in the lab-scale bioreactor.

Along with the developed strategies, the experimental setup is transformed into a reliable closed-loop system with a flexible monitoring and control software solution. This transformation allows to successfully implementing and validating the proposed control strategies on the real bioprocess, providing a framework for more validations of future works considering different control objectives and strategies. This thesis covered several steps to achieve these objectives:

- A presentation of the bacterial system used in this thesis: *Escherichia coli*, with a description of its physiological and metabolic features under aerobic and anaerobic conditions. The focus is then put on acetate production through the overflow metabolism mechanism, considering its inhibitory effect on bacterial growth. The different cultivation modes of bioprocesses are also presented, with a highlight on the fed-batch mode used in the framework of this thesis.

- A presentation of the macroscopic mathematical model of the fed-batch *E. coli* bioprocess. The model is based on the reaction scheme describing the metabolism of *E. coli*. The kinetic model is determined according to the bottleneck theory describing the overflow metabolism mechanism in the form of two metabolic regimes. The presented model is the basis of the estimation and control schemes implemented in this thesis.

- In order to implement the control and estimation strategies presented in this thesis, a closed-loop interactive system is developed for a lab-scale bioreactor. This system is composed of a data-acquisition LabView program gathering the real-time measurements from the various sensors with different acquisition software. Furthermore, the program has a modular structure for deploying the

estimation and control algorithms, making it flexible for software and hardware upgrade and changing the control structure and objectives. Finally, the control of the feed rate is made possible using peristaltic pumps, turning the bioreactor from an open-loop to a reliable closed-loop system.

– The developed control strategies require the on-line availability of the main state variables concentrations (Biomass, Glucose, Acetate). However, the biomass concentration is the only variable available for on-line measurement. To overcome this practical hurdle, software sensors are considered for state and kinetic parameter estimations. In the case of our process, the linear Kalman filter and the Unscented Kalman filter are implemented to estimate the glucose and acetate concentrations and other kinetic parameters based on the measured biomass concentration. These algorithms are tested and tuned in simulation, based on previous data from the bioprocess, and validated in the control experiments.

– The different control strategies developed and implemented in fed-batch fermentation by feed-rate manipulation are presented. The main differences between the control methods are the requirements, complexity, process operation, and control objectives. This presentation provides a basis and a guide to formulate the control problems depending on the available materials, methods, and objectives.

  This thesis focused on maximizing the biomass productivity and avoid acetate accumulation in fed-batch *E. coli* cultures. The availability of a reliable process model and on-line biomass measurements lead to the consideration of model-based methods to regulate the biomass growth rate and develop state estimators based on biomass measurement. Three strategies control are developed:

– First, an adaptive biomass regulation strategy is developed based on the Generic Model Control algorithm. A model order reduction is applied to obtain a control law independent from kinetic terms and avoid high flow rates. Parameter adaptation based on the linear Kalman filter is coupled to the controller to estimate the unmeasured terms of the control law and adapt to model mismatch.

  The controller is set to track a defined growth rate reference chosen slightly lower than the maximal growth rate to avoid acetate accumulation while maximizing biomass productivity. The performance of the proposed strategy is tested in simulation, and validated through fed-batch experiments of a BL21(DE3) *E. coli* strain were achieved on a lab-scale bioreactor.

– Secondly, a robust variation of the Generic Model Control strategy is developed and applied to regulate the acetate concentration at a defined low level in fed-batch cultures of *E. coli* BL21 (DE3). A robust design procedure using the $\mathcal{LMI}$

formalism is carried out to account for model mismatch, disturbances, and measurement noise. Performance constraints are also formulated with $\mathcal{LMI}$s to ensure desired properties of the closed-loop transient response.

The robust GMC controller combined with the state estimation by the UKF is validated both in simulation runs and in real-time experimental conditions.

– Finally, an implementation of the nonlinear model predictive control strategy on the fed-batch *E. coli* is considered to regulate the acetate concentration to a low value. The NMPC strategy is chosen for its several advantages, like the implicit use of the nonlinear model equations for the prediction and state and control constraints.

To improve the computation performance and reduce the complexity of the optimization problem, the latter is transformed into a nonlinear programming problem using the control vector parametrization (CVP) approach.

The NMPC strategy coupled to the UKF estimator is validated through simulations and experiments on the lab-scale reactor. The experiments also validated that regulating the acetate concentration to a low level leads to higher biomass concentrations, lower cultivation time, and higher biomass productivity.

This thesis presented practical control solutions to growth inhibition by acetate accumulation in fed-batch *E. coli* cultures. These control methods, along with monitoring and estimation solutions, were tested on real-time conditions in a lab-scale bioreactor. The obtained results highligthed the efficiency of the proposed strategies, and their benefits in comparison to open-loop strategy (increase of biomass productivity of about 20% over 10 h).

The difference between the proposed control strategies can be summarized in two main categories: Control objective and the method type and complexity.

Regarding the control objectives, two main approaches were investigated. Regulating the biomass concentration at a targeted reference growth rate and regulating the acetate concentration at a low level.

The biomass regulation is performed by tracking a generated biomass profile corresponding to a specified growth rate. The operational advantage of this approach is that it does not require state estimation, thanks to reliable biomass measurements. Another advantage is the reproducibility of results, even under model mismatch (assuming that the control method is robust).

The main issue lies in the definition of the control problem. The method relies on defining a distance from the maximal growth rate, which is an uncertain variable linked to the cells critical capacity. Underestimating the maximal value can lead to uncontrolled overflow and acetate accumulation due to overfeeding. This could be a problem at high cell densities where a slight increase in the substrate concentration can provoke acetate fermentation. Overestimating the maximal value, on the other hand, leads to underfeeding and long cultivation time. The resulting biomass yield is even lower than its maximal value.

The acetate concentration, on the other hand, is directly linked to the metabolic optimum. Acetate is produced when the cell's oxidative capacity is exceeded. Ideally, acetate should be equal to zero, but practically this can lead to uncontrolled metabolic switches causing cell stress and slowing their metabolism. Furthermore, an acetate concentration equal to zero can also be achieved by underfeeding, which is not ideal for a biomass yield maximization objective.

One could target the critical substrate concentration, but this value is very low (0.0375 g/L), and any small fluctuations around this value can cause a switch between respirative and respiro-fermentative regimes. In addition, no reliable measurement or estimation solutions exist for this concentration range.

Therefore, regulating the acetate concentration to a low value close to zero is more practical. This value is linked to the maximal oxidative capacity, and if the latter is uncertain, the metabolic performance is preserved. Another advantage is that acetate provides an on-line indicator of the culture metabolic state.

The challenge raised by this approach is the necessity of accurate acetate estimation, which requires a reliable process model, and experimental data. Regarding the control methods, this thesis investigated feed-back linearization and predictive methods.

The GMC offers a straightforward solution to the control problem with minimal computation burden. This feature makes the GMC a good candidate for real-time implementations of set-point and trajectory tracking problems. However, this comes with the cost of long transient phases and tracking errors caused by model uncertainties.

Considering a robust design and adaptive schemes can improve the control performance with uncertain models and unpredicted dynamics. The NMPC controller allows the inclusion of state and input constraints, flexibility in the definition of the control law, and implicit use of the nonlinear model. These features lead to better tracking performance and lower transient phases. Consequently, the biomass productivity and metabolic performance of the cells is higher when using the NMPC compared to the GMC.

However, the implementation of the NMPC comes with a high computational burden due to the strong nonlinearities in bioprocess models, especially when several operational constraints are considered. Furthermore, the robust design of the NMPC is not as straightforward as the GMC and raises the calculation complexity. Time-saving strategies can reduce the computation load while maintaining the benefits of the predictive strategies.

Finally, model-based strategies offer the opportunity to investigate several developed solutions from control theory. However, the performance is directly linked to the process model parameters. The uncertain nature of bioprocess and the complexity of the biological dynamics requires rigorous parameter estimation schemes.

9.2 Recommendations for future research

The work that has been undertaken during this thesis has underlined several areas for further research. These include further investigation and improvement in experiments, modeling, estimation, and control.

9.2.1 Experimental implementation

Hardware & software

Regarding implementing the closed-loop system, several improvements can be made regarding the hardware setup and the computation efficiency. Developing a dedicated bioprocess control library in LabView would be extremely useful to improve the efficiency of the current solution and easily implement other control structures and algorithms. Furthermore, it will provide a useful platform for experiments on different bioprocesses such as Microalgae and Hybridoma cells.

On this topic, using the shared libraries function in LabView allows the integration of Dynamic Link Libraries (DLL). This provides a significant gain in computation time since dynamic libraries contain compiled dynamic functions mainly developed in C or C++. Several efficient and open-source libraries dedicated to nonlinear integration and optimization can be found for these languages, such as `libIntegrate`, `OptimLib`, `ACADO`, and many others. Thereby, creating a dedicated bioprocess modeling, control, and estimation library based on shared libraries function can open doors for testing several algorithms.

Biological aspects

An area of improvement on the biological aspects of the experiments would be testing the same control algorithm on different strains of *E. coli*, and comparing the results to showcase the robustness towards the model kinetic and yield parameters.

Another perspective is testing and validating the strategies on higher bioreactor scales since the lab-scale reactor presented a limitation for testing the performance under higher cell densities.

The application of the developed algorithms and monitoring tools on other microorganisms exhibiting overflow metabolism would be an interesting area to explore. An example is the regulation of ethanol concentration in *Saccharomyces cerevisiae* cultures.

9.2.2 Model and Estimation

Regarding the process model, considering the differentiable *E. coli* model presented in (Anane et al., 2017) can lead to several improvements, especially regarding the state estimation. This model is based on the acetate cycling principle and contains continuous functions for the glucose and acetate kinetics. The consideration of this model requires carrying a new set of cultures using gas analyzers.

These experiments dedicated to model identification would be beneficial for developing new model-based strategies.

The soft sensors area can be improved by developing a joint parameter-state estimation strategy. For example, the on-line estimation of the maximal growth rate can improve the state estimation accuracy and the control efficiency. .

On the other hand, it would be interesting to develop and implement a moving horizon estimator (MHE) (Allgöwer et al., 1999) coupled with the NMPC controller. The MHE is an optimization-based state-estimation method that shares the same algorithmic structure as the MPC. Thereby, it allows the explicit use of nonlinear models leading to higher estimation quality than Kalman filtering methods. Another advantage of the MHE is the possibility to incorporate constraints on the estimated variables and the flexibility in the definition of the cost function. Furthermore, several robust variations of the MHE have been developed for uncertain systems, making it an appropriate choice for bioprocesses. The main challenge facing this implementation would be reducing the calculation complexity for reliable real-time results.

9.2.3 Control

Control objectives

Regarding the biological aspects of the control schemes, it would be interesting to consider the maximization of the glucose oxidation while minimizing the acetate formation. This concept is presented in a simulation study (Santos et al., 2012b), and achieved by optimizing a cost function in an NMPC framework. A recommendation for the real-time implementation is to add a safety margin from the optimum (metabolic edge) in order to avoid repetitive switches. A combination with the Receding Horizon Estimator with parameter and state estimation would be ideal for this type of control.

Control methods

The control schemes presented in this thesis managed to achieve the control objectives, but improvements on the strategies could be made. Implementing a robust predictive scheme using a min-max NMPC (Scokaert and Mayne, 1998) or a Tube NMPC (Langson et al., 2004) would improve the robustness, especially on higher cell densities. However, these strategies must be followed with computation time-saving methods (sensitivity analysis, efficient optimization algorithms, ...).

Another area is to consider probing methods such as Extremum seeking (ES) (Ariyur and Krstić, 2003), where a cost function is minimized, and the unknown optimum is tracked on-line without the need for a process model. A problem facing this implementation is the necessity of substrate measurements at low levels. This problem can be solved by robust substrate estimation or by combining the Extremum seeking algorithm with the probing control structure presented in (Kesson et al., 2001), where the dissolved oxygen measurements were used to detect and

avoid overflow metabolism. A robust formulation of the ES algorithm would also be beneficial for higher performance.

Appendix A

Kalman Filter algorithms

A.1 Linear Kalman Filter

The classical Kalman filter addresses the general problem of state estimation in linear stochastic processes. We consider the following discrete-time recursive stochastic system:

$$\xi_k = \mathbf{A}_{k-1}\xi_{k-1} + \mathbf{B}_{k-1}u_{k-1} + v_{k-1} \tag{A.1a}$$
$$y_k = \mathbf{C}_k\xi_k + w_k \tag{A.1b}$$

where $\xi \in \mathbb{R}^n$ is the state vector, $u \in \mathbb{R}^m$ is the system input, $y \in \mathbb{R}^p$ is the measurement vector.

The matrix $\mathbf{A}_k \in \mathbb{R}^{n\times n}$ in Equation (A.1a) relates the state ξ at instant k, and the matrix $\mathbf{B}_k \in \mathbb{R}^{n\times m}$ relates the control input u to the state ξ. The matrix $\mathbf{C}_k \in \mathbb{R}^{p\times n}$ in Equation (A.1b) relates the state ξ to the measurement y_k.

v_k and w_k represent the process and measurement noise, respectively. They are assumed to be non-correlated, have a zero mean and normal probability distributions:

$$\begin{aligned} v &\sim N(0, \mathbf{Q}) \\ w &\sim N(0, \mathbf{R}) \end{aligned} \tag{A.2}$$

where

- $\mathbf{Q}$ is covariance matrix of process noise v_k.
- $\mathbf{R}$ is covariance matrix of measurement noise w_k.

The process noise v_k is used to model the confidence in the system model, while the measurement noise w_k is used to model the measurement quality.

Given the knowledge of the process evolution prior to the instant k, we define the *a priori* state estimate at step k as $\hat{\xi}_k^- \in \mathbb{R}^n$. Similarly, given the measurement y_k at instant k, the *a posteriori* state estimate is refered to as $\hat{\xi}_k \in \mathbb{R}^n$.

Definition: The *a priori* and *a posteriori* estimation errors are defined as follows (Welch and Bishop, 1995):

$$\begin{aligned} e_k^- &\triangleq \xi_k - \hat{\xi}_k^- \\ e_k &\triangleq \xi_k - \hat{\xi}_k \end{aligned} \tag{A.3}$$

The *a priori* and *a posteriori* covariance matrices of the estimation errors are defined as:

$$\mathbf{P}_k^- = \mathbb{E}\left[e_k^- e_k^{-T}\right] \tag{A.4}$$

$$\mathbf{P}_k = \mathbb{E}\left[e_k e_k^T\right] \tag{A.5}$$

where $\mathbb{E}$ is the mathematical expectation.

A.2 Kalman filter algorithm

The Kalman filter algorithm consist in two major phases: a prediction phase between instants $k-1$ and k, and an update phase at instant k.

First, the initial estimated state vector and covariance matrix are initialized:

$$\begin{aligned} \hat{\xi}_0 &= \mathbb{E}\left[\xi_0\right] \\ \mathbf{P}_0 &= \mathbb{E}\left[e_0\, e_0^T\right] \end{aligned} \tag{A.6}$$

where $\mathbf{P}_0$ represents the initial covariance matrix of the estimation error, and ξ_0 and $\hat{\xi}_0$ represent the initial state and the initial state estimate, respectively.

Prediction:

In this step, the model and the covariance $\mathbf{P}_{k-1}^-$ are used to calculate the estimated *a priori* state $\hat{\xi}_k^-$.

$$\begin{aligned} \hat{\xi}_k^- &= \mathbf{A}_{k-1}\hat{\xi}_{k-1} + \mathbf{B}_{k-1}u_{k-1} \\ \mathbf{P}_k^- &= \mathbf{A}_{k-1}\mathbf{P}_{k-1}\mathbf{A}_{k-1}^T + \mathbf{Q} \end{aligned} \tag{A.7}$$

Update:

At instant k, the available measurement y_k is used to update the *a posteriori* estimated state $\hat{\xi}_k$ and the covariance matrix $\mathbf{P}_k$.

$$\mathbf{\Omega}_k = \mathbf{P}_k^- \mathbf{C}_k^T \left(\mathbf{C}_k \mathbf{P}_k^- \mathbf{C}_k^T + \mathbf{R} \right)^{-1} \tag{A.8}$$

$$\mathbf{P}_k = \left(I - \mathbf{\Omega}_k \mathbf{C}_k \right) \mathbf{P}_k^- \tag{A.9}$$

$$\hat{\xi}_k = \hat{\xi}_k^- + \mathbf{\Omega}_k \left(y_k - \mathbf{C}_k \hat{\xi}_k^- \right) \tag{A.10}$$

where $\mathbf{\Omega}$ is called the Kalman gain, it minimizes the trace of the covariance matrix of the *a posteriori estimation error* presented in equation Equation (A.9). The term $\left(y_k - \mathbf{C}_k \hat{\xi}_k^- \right)$ is called the residual or the measurement innovation term.

The prediction and update steps are repeated at the next sampling point. The algorithm is initialized with the previous *a posteriori* estimates and covariance matrix. These are then used as initial variables in a new prediction step to calculate the new *a priori* estimates.

Tuning the Kalman filter

The tuning of the filter parameters is performed prior to the initialization of the filter using the process noise $\mathbf{Q}$ and the measurement noise $\mathbf{R}$ covariance matrices.

The measurement noise covariance $\mathbf{R}$ is chosen depending on the measurement quality and the sensor characteristics. The process covariance matrix $\mathbf{Q}$ is usually tuned to model the uncertainty of the process model on each state variable. Adding sufficient uncertainty to the matrix $\mathbf{Q}$ improves the filter estimation quality in the case of uncertain systems.

Extended Kalman filter

The EKF is applied to systems represented by nonlinear differential equations, where a linearization is performed around the current state using the partial derivatives of the process and measurement functions to compute estimates. Let us consider the following stochastic continuous nonlinear system:

$$\begin{aligned} \dot{\xi}(t) &= f\left(\xi(t), u(t)\right) + v(t) \\ y(t) &= h(\xi(t)) + w(t) \end{aligned} \tag{A.11}$$

where $\xi \in \mathbb{R}^n$ is the state vector, $u \in \mathbb{R}^p$ is the input vector, $y \in \mathbb{R}^m$ is the measurement vector, $f(\cdot)$ and $h(\cdot)$ are nonlinear functions. $v \sim N\left(0, \mathbf{Q}\right)$ is the process noise vector, $w \sim N\left(0, \mathbf{R}\right)$ is the measurement noise vector. They are assumed to be non-correlated.

The extended Kalman filter applied to the system (A.11) is an extension of the classical Kalman filter in the linear case. It requires to linearize the nonlinear state and measurement equations around the estimated trajectories.

In the context of bioprocess monitoring, the concentration measurements are obtained at discrete times due to sampling or processing delays in the sensors. On the other hand, as presented in chapter 2, the dynamic models of bioprocesses are

continuous. Thus, a continuous-discrete of formulation of the Extended Kalman Filter is considered (Särkkä, 2007), where the state dynamics are modeled as continuous time process, and the measurements are obtained at discrete times.

The EKF algorithm follows the same steps as the linear Kalman filter: the prediction phase between instants $k-1$ and k, and the update phase at instant k. The initialization step also remains the same (Equation (A.6)).

The prediction phase uses the continuous model to obtain the *a priori* estimated state and covariance matrix, and the update phase uses the classical discrete Kalman filter equations.

Prediction:

The *a priori* prediction of the state, $\hat{\xi}_k^-$, and the *a priori* covariance matrix $\mathbf{P}_k^-$ are obtained by solving:

$$\begin{aligned} \dot{\hat{\xi}}(t) &= f(\hat{\xi}(t), u(t)) \\ \dot{\mathbf{P}}(t) &= \mathbf{A}\mathbf{P}(t) + \mathbf{P}(t)\mathbf{A}^\top + \mathbf{Q}(t) \end{aligned} \tag{A.12}$$

with

$$\begin{aligned} \hat{\xi}_k^- &= \xi(t_k) \\ \mathbf{P}_k^- &= \mathbf{P}(t_k) \end{aligned} \tag{A.13}$$

The matrix $\mathbf{A}$ is obtained by linearizing the dynamics around the current state estimate $\hat{\xi}_{k-1}$

$$\mathbf{A} = \left.\frac{\partial f(\xi, u)}{\partial \xi}\right|_{\xi=\hat{\xi}_{k-1}} \tag{A.14}$$

Update:

The measurement y_k is used to update the prediction and calculate the *a posteriori* estimate $\hat{\xi}_k$:

$$\mathbf{\Omega}_k = \mathbf{P}_k^- C_k^T \left(\mathbf{C}\mathbf{P}_k^-\mathbf{C}^T + R_k\right)^{-1} \tag{A.15}$$

$$\mathbf{P}_k = \left(I - \mathbf{\Omega}_k\mathbf{C}\right)\mathbf{P}_k^- \tag{A.16}$$

$$\hat{\xi}_k = \hat{\xi}_k^- + \mathbf{\Omega}_k\left(y_k - h\left(\hat{\xi}_k^-\right)\right) \tag{A.17}$$

where the matrix $\mathbf{C}$ is obtained by linearizing the measurement function around the current state estimate:

$$\mathbf{C} = \left.\frac{\partial h(\xi, u)}{\partial \xi}\right|_{\xi=\hat{\xi}_k^-} \tag{A.18}$$

The prediction and update steps are repeated for the next measurement at instant $k+1$.

A.3 The Unscented Kalman filter

The UKF algorithm is composed of three main steps. First, a set of deterministically chosen points (called sigma-points) around the current state estimate is generated via the unscented transform. Next, the classical prediction (between instants $k-1$ and k) and the update (at instant k) phases of the Kalman filter are performed. The estimated state and covariance matrix are initialized by Equation (A.6).

First, a discrete-time prediction model is obtained by the integration of the continuous time state-space model (A.11) using a Runge-Kutta method. Thus, the prediction step is performed using the following equations, considering a constant sampling time T_e :

$$\begin{aligned} \xi_k &= \Psi\left(\xi_{k-1}, F_{in_k}\right) \\ y_k &= H\xi_k \end{aligned} \tag{A.19}$$

where the index k is the discrete time, ξ_k and y_k are the discrete state vector and the sampled measurement at time kT_e, respectively. Ψ is the evolution function, and H is the measurement matrix. The control input F_{in_k} is parametrized using a piecewise constant approximation.

The unscented transform (UT):

The unscented transform (UT) is the first step in the UKF algorithm (Julier and Uhlmann, 2004), it is a mathematical transformation used to approximate the gaussian distribution of random variables.

The UT consists in generating a set of deterministically chosen sigma-points χ around the current state estimate. The UT role is to capture the mean of the stochastic variable $\hat{\xi}^-$ and covariance of the estimation error $\mathbf{P}$. The generated sigma-points are then propagated through the nonlinear function $\Psi[.]$ to estimate the mean and covariance of ξ.

First, a set of $2n+1$ sigma points is generated as follows:

$$\begin{cases} \chi_{k-1}^0 = \hat{\xi}_{k-1}^-, \;\; i = 0 \\ \chi_{k-1}^i = \hat{\xi}_{k-1}^- + (\sqrt{(n+\Lambda)\mathbf{P}})_i, \;\; i = 1, \ldots, 2n \\ \chi_{k-1}^i = \hat{\xi}_{k-1}^- - (\sqrt{(n+\Lambda)\mathbf{P}})_i, \;\; i = n+1, \ldots, 2n \end{cases} \tag{A.20}$$

where

- $\hat{\xi}_{k-1}^{-}\mathbb{R}^n$ the *a priori* estimated state at instant $k-1$.
- χ_{k-1}^{i} is the i^{th} generated sigma point at instant $k-1$.
- $\mathbf{P}$ is the covariance of the estimation error.
- Λ is a scaling parameter defined as $\Lambda = \alpha^2(n+\kappa) - n$.

The corresponding weights of the sigma points are computed as follows:

$$\begin{cases} \omega_m^0 = \frac{\Lambda}{n+\Lambda} \\ \omega_c^0 = \frac{\Lambda}{n+\Lambda} + \left(1 - \alpha^2 + \beta\right) \\ \omega_m^i = \omega_c^i = \frac{1}{2(n+\Lambda)}, \;\; i = 1, \ldots, 2n \end{cases} \tag{A.21}$$

where

- ω_{m}^i and ω_{c}^i represent respectively the weights of the mean and the covariance.
- α, κ, and β are positive constants tuned to set the appropriate spread of the sigma points.

Prediction

The prediction phase (between k and $k+1$) consists in computing the *a priori* predicted state $\hat{\xi}_k^-$ and the predicted *a priori* covariance matrix $\mathbf{P}_k^-$.

The sigma vectors are propagated through the nonlinear function Ψ, and the predicted value and covariance of the prediction error are computed using the weights ω^i.

$$\chi_k^{i-} = \Psi\left(\chi_{k-1}^{i}\right) \tag{A.22}$$

$$\hat{\xi}_k^- = \sum_{i=0}^{2n} \omega_m^i \chi_k^{i-} \tag{A.23}$$

$$\mathbf{P}_k^- = \sum_{i=0}^{2n} \omega_c^i \left[\chi_k^{i-} - \hat{\xi}_k^-\right]\left[\chi_k^{i-} - \hat{\xi}_k^-\right]^T + \mathbf{Q} \tag{A.24}$$

where

- χ_k^{i-} the *a priori* generated sigma point at instant k.

Update

The update step (at $k+1$) consists in using the measurements at time k to correct the estimated state $\hat{\xi}_k^-$ and covariance of the estimation error $\mathbf{P}_k^-$ obtained in the previous step.

The predicted measurement $\hat{y}_k$ is calculated by propagating the predicted sigma points through the measurement function $H[.]$:

$$\hat{Y}_k^i = H\left(\chi_{k-1}^{i-}\right) \tag{A.25}$$

$$\hat{y}_k = \sum_{i=0}^{2n} \omega_m^i \hat{Y}_k^i \tag{A.26}$$

The measurement covariance matrix $\mathbf{P}_{\tilde{y}_k\tilde{y}_k}$ and the cross covariance matrix $\mathbf{P}_{x_y y_k}$ are then computed:

$$\mathbf{P}_{\tilde{y}_k\tilde{y}_k} = \sum_{i=0}^{2n} \omega_c^i \left[\hat{Y}_k^i - \hat{y}_k\right] \left[\hat{Y}_k^i - \hat{y}_k\right]^T + \mathbf{R} \tag{A.27}$$

$$\mathbf{P}_{x_k y_k} = \sum_{i=0}^{2n} \omega_c^i \left[\chi_k^{i-} - \hat{\xi}_k^-\right] \left[\hat{Y}_{i(k)} - \hat{y}_k\right]^T \tag{A.28}$$

$$\tag{A.29}$$

Finally, the measurement update is performed using the Kalman filter equations:

$$\mathbf{\Omega}_k = \mathbf{P}_{x_k y_k} \mathbf{P}_{\tilde{y}_k\tilde{y}_k}^{-1} \tag{A.30}$$

$$\hat{\xi}_k = \hat{\xi}_k^- + \mathbf{\Omega}_k \left(y_k - \hat{y}_k\right) \tag{A.31}$$

$$\mathbf{P}_k = \mathbf{P}_k^- - \mathbf{\Omega}_k \mathbf{P}_{\tilde{y}_k\tilde{y}_k} \mathbf{\Omega}_k^T \tag{A.32}$$

where $\mathbf{\Omega}_k$ is the Kalman gain, and $\mathbf{P}_k$ is the *a posteriori* estimated state and covariance of the estimation error.

Appendix B

Culture monitoring

B.1 Biomass concentration measurements

The biomass concentration was obtained by determining the cell dry weight of a selected set of samples at different concentrations, then correlated with the optical density (OD) measurements. The procedure is summarized in the following steps:

- Place the filters $(45\ \mu\text{m})$ at 105°C overnight.
- Weight the filters using an analytical balance.
- Prepare standard solutions at different dilutions from a sample.
- Measure the OD of the prepared samples.
- Filter 10 mL of the samples and store them overnight at 105°C.
- Weight the filtered samples using an analytical balance.
- Calculate the biomass concentration X_{sample} using:

$$X_{\text{sample}} = \frac{\text{Average Cell Dry Weight}}{\text{Filtered Volume}}$$

- Build a chart of the biomass concentration X_{sample} versus the optical density (OD) in the linear range

The results of one of the dry weight and calibration experiments are given in Table B.1 and illustrated in Figure B.1.

B.1.1 Biomass probe calibration

The calibration is carried out using the off-line biomass measurements and the optical density obtained from the Turibidimetric probe (Fandalux, Germany) (Figure B.2). Several calibration methods were tested (linear, exponential) and the polynomial curve presented the best trade off between data fit in high densities, and stability under airflow disturbances.

TABLE B.1: Calibration of the biomass concentration from the optical density measurements

Filter	**Weight (g)**			**Optical Density**	X_{sample}
Sample	Filter	Filter + Biomass	Biomass	(abs)	(g/L)
0	0.301	0.301	0	0.098	0
1	0.280	0.300	0.0198	5.420	1.980
2	0.349	0.361	0.0170	3.097	1.170
2	0.302	0.311	0.008	2.587	0.800
4	0.301	0.306	0.005	1.805	0.540
5	0.366	0.369	0.003	1.234	0.370

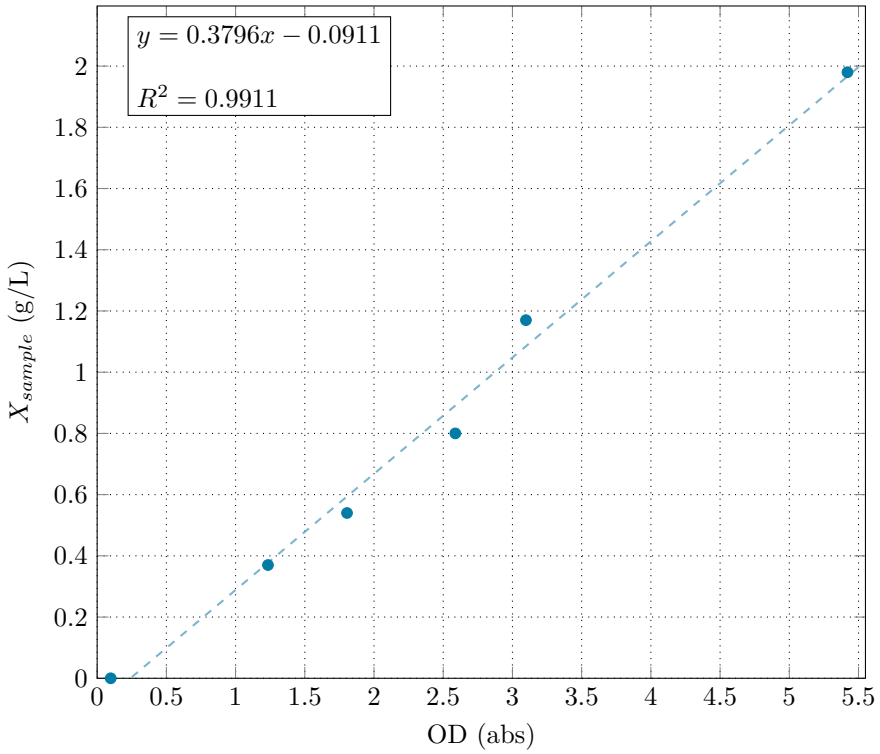


FIGURE B.1: Biomass concentration calibration curve from the optical density measurements

B.2 Glucose concentration measurements

The glucose measurements were performed using the dinitrosalysilic acid (DNS) method, described hereafter:

DNS reagent preparation

- Dissolve 1.00 g of DNS in 20 mL of NaOH 2 M
- Dissolve 30 g of Potassium sodium tartrate tetrahydrate (Rochelle salt) in 50mL of distilled water.

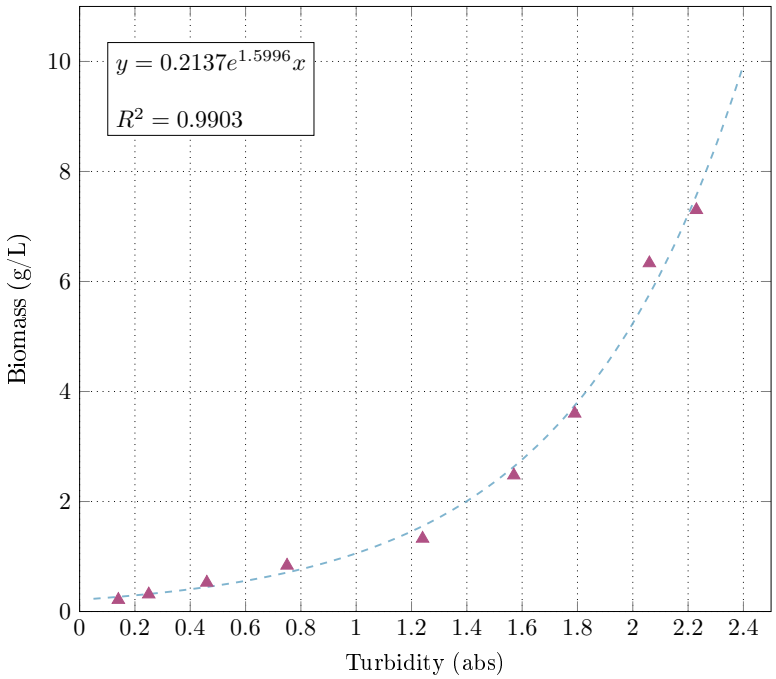


FIGURE B.2: Biomass concentration calibration curve from the optical density measurements

- Stir until complete dissolution then mix the prepared solutions.
- Heat and mix to homogenize.
- Fill with distilled water for a final volume of 100 mL.
- Store in amber bottle at 4°C.

Sample preparation and measurement procedure

- Add 1mL of culture supernatant and 1mL of DNS reagent.
- Boil the mix in a 100°C water bath for 5 min.
- Add 8 mL of distilled water after cooling to room temperature.
- Homogenize and read the absorbance at 540nm in the spectrophotometer.

The blank solution follows the same procedure as the samples, but the culture supernatant is replaced with distilled water.

Calibration

An example of glucose calibration results is given in Table B.2. Figure B.3 shows a calibration curve obtained during this experiment. The data fit is linear in the range of 0 to 1 g/L with a correlation coefficient of 0.996.

TABLE B.2: Calibration of the glucose measurements using the DNS procedure

Sample	**Concentration**	**OD 1**	**OD 2**	**OD 3**
	g/L	abs	abs	abs
1	0	0.044	0.046	0.042
2	0.2	0.165	0.168	0.166
3	0.4	0.260	0.276	0.266
4	0.5	0.341	0.341	0.356
5	0.8	0.487	0.491	0.480
6	1	0.637	0.629	0.635

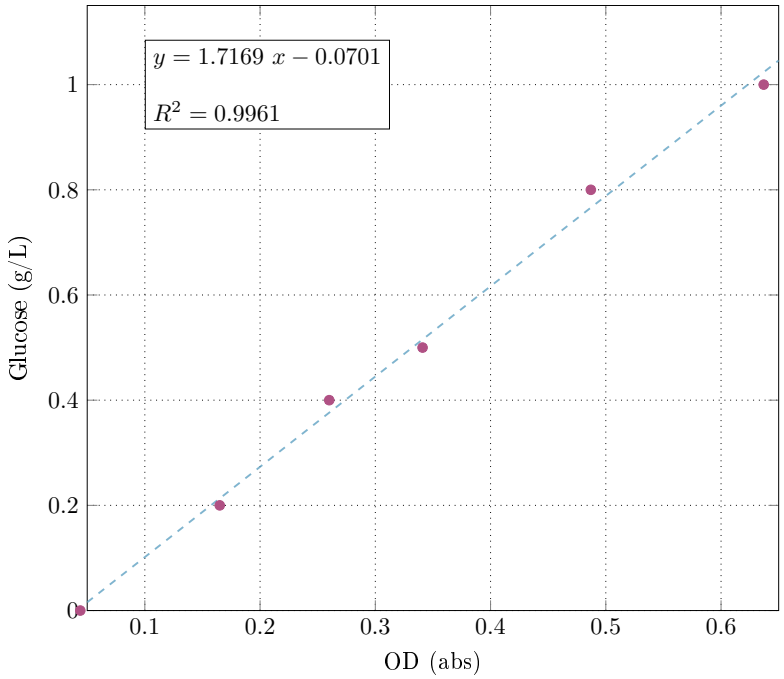


FIGURE B.3: Glucose concentration calibration curve from the optical density measurements

B.3 Acetate concentration measurements

Acetate concentration measurements were performed with an enzymatic kit (Megazyme, Ireland) according to the manufacturer's instructions. The calibration of the measurements was performed using the standard solution provided with the measurement kit. The concentrations are obtained by measuring the optical density (OD) of the samples and a calibration curve.

The acetate calibration is carried out by triplicate. Table B.3 and Figure B.4 show one of the calibration curves performed during the study. The method demonstrated linearity in a range of 0 to 0.25 g/L with a correlation coefficient of 0.98.

Calibration

TABLE B.3: Calibration of the acetate measurements using the enzymatic kit

Sample	OD	C(nM/μL)	C(g/L)
0	0.257	0.012714777	0.015016151
2	0.322	0.035051546	0.041395876
4	0.421	0.069072165	0.081574227
6	0.542	0.110652921	0.1306811
8	0.708	0.167697595	0.198050859
10	0.816	0.204810997	0.241881787

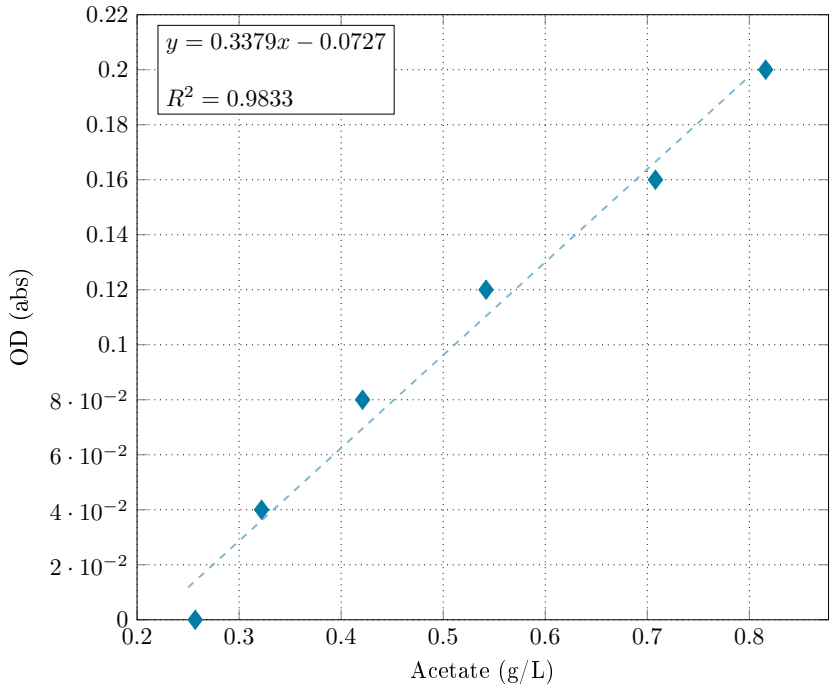


FIGURE B.4: Acetate concentration calibration curve from the Optical density measurements

Appendix C

GMC biomass regulation parameters

C.1 Closed-loop response and parameter tuning

A closed loop response showcasing the closed-loop behavior for different tuning values of the damping ratio ξ and the response time t_r are shown in Figures C.1 and C.2. These variables are used to set the GMC control parameters for the biomass regulation presented in chapter 6.

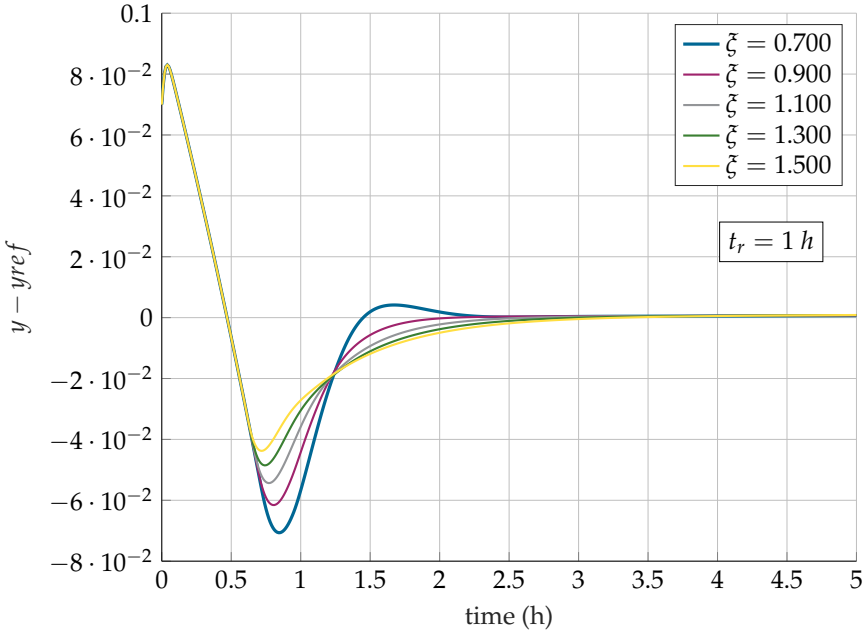


FIGURE C.1: GMC biomass tracking response specification with parameters for for different values of ξ for the *E. coli* BL21 model.

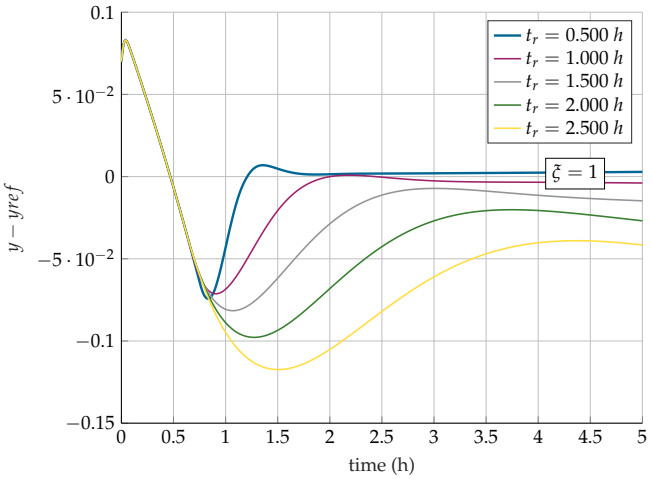


FIGURE C.2: GMC biomass tracking response specification for for different values of t_r for the *E. coli* BL21 model.

C.2 Expressions of the α_j coefficients

Theoretical expression of the α_j coefficients presented in chapter 5, calculated using Matlab Symbolic Toolbox:

$$
\begin{aligned}
\alpha_1 &= 9S_{in}K_{OS} - 9K_SK_{OS} + 50K_Sk_{X1}q_{O_{max}} - 50K_Sk_{X2}q_{O_{max}} - 50S_{in}k_{X1}q_{O_{max}} \\
&\quad + 50S_{in}k_{X2}q_{O_{max}} - 50S_{in}K_{OS}k_{X2}q_{S_{max}} \\
\alpha_2 &= +50K_{OS}q_{S_{max}} \\
\alpha_3 &= 81K_S^2K_{OS}^2 - 900K_S^2K_{OS}k_{X1}q_{O_{max}} + 900K_S^2K_{OS}k_{X2}q_{O_{max}} + 2500K_S^2k_{X1}^2q_{O_{max}}^2 \\
&\quad - 5000K_S^2k_{X1}k_{X2}q_{O_{max}}^2 + 2500K_S^2k_{X2}^2q_{O_{max}}^2 - 900K_SS_{in}K_{OS}^2k_{X2}q_{S_{max}} \\
&\quad + 162K_SS_{in}K_{OS}^2 + 5000K_SS_{in}K_{OS}k_{X1}k_{X2}q_{O_{max}}q_{S_{max}} - 1800K_SS_{in}K_{OS}k_{X1}q_{O_{max}} \\
&\quad - 5000K_SS_{in}K_{OS}k_{X2}^2q_{O_{max}}q_{S_{max}} + 1800K_SS_{in}K_{OS}k_{X2}q_{O_{max}} + 5000K_SS_{in}k_{X1}^2q_{O_{max}}^2 \\
&\quad - 10000K_SS_{in}k_{X1}k_{X2}q_{O_{max}}^2 + 5000K_SS_{in}k_{X2}^2q_{O_{max}}^2 - 900K_SXsetK_{OS}^2q_{S_{max}} \\
&\quad + 5000K_SXsetK_{OS}k_{X1}q_{O_{max}}q_{S_{max}} - 5000K_SXsetK_{OS}k_{X2}q_{O_{max}}q_{S_{max}} + 2500S_{in}^2K_{OS}^2k_{X2}^2q_{S_{max}}^2 \\
&\quad - 900S_{in}^2K_{OS}^2k_{X2}q_{S_{max}} + 81S_{in}^2K_{OS}^2 + 5000S_{in}^2K_{OS}k_{X1}k_{X2}q_{O_{max}}q_{S_{max}} - 900S_{in}^2K_{OS}k_{X1}q_{O_{max}} \\
&\quad - 5000S_{in}^2K_{OS}k_{X2}^2q_{O_{max}}q_{S_{max}} + 900S_{in}^2K_{OS}k_{X2}q_{O_{max}} + 2500S_{in}^2k_{X1}^2q_{O_{max}}^2 \\
&\quad - 5000S_{in}^2k_{X1}k_{X2}q_{O_{max}}^2 + 2500S_{in}^2k_{X2}^2q_{O_{max}}^2 \\
\alpha_4 &= 5000S_{in}K_{OS}^2k_{X2}q_{S_{max}}^2 + 900S_{in}K_{OS}^2q_{S_{max}} - 5000S_{in}K_{OS}k_{X1}q_{O_{max}}q_{S_{max}} \\
&\quad + 5000S_{in}K_{OS}k_{X2}q_{O_{max}}q_{S_{max}} \\
\alpha_5 &= 2500K_{OS}^2q_{S_{max}}^2 \\
\alpha_6 &= 2(9K_{OS} - 50k_{X1}q_{O_{max}} + 50k_{X2}q_{O_{max}} - 50K_{OS}k_{X2}q_{S_{max}})
\end{aligned}
$$

Appendix D

Sensitivity analysis

In view of analyzing the robustness of the control strategies to plant-model mismatch, the most influencing parameters are determined by a parametric sensitivity study. The influence of a parameter θ_j on the model output x_i can be quantified by:

$$\frac{d}{dt}\left(\frac{\partial x_i}{\partial \theta_j}\right) = \frac{\partial}{\partial \theta_j}\left(\frac{dx_i}{dt}\right) \tag{D.1}$$

The expressions of the parametric sensitivities for the *E. coli* model are therefore given by:

$$\frac{d}{dt}\left(\frac{\partial x_i}{\partial \theta_j}\right) = \frac{\partial \boldsymbol{F}_i}{\partial \theta_j} + \frac{\partial \boldsymbol{F}_i}{\partial X}\frac{\partial X}{\partial \theta_j} + \frac{\partial \boldsymbol{F}_i}{\partial S}\frac{\partial S}{\partial \theta_j} + \frac{\partial \boldsymbol{F}_i}{\partial A}\frac{\partial A}{\partial \theta_j} + \frac{\partial \boldsymbol{F}_i}{\partial V}\frac{\partial V}{\partial \theta_j} \tag{D.2}$$

where $\boldsymbol{F}_i$ represents the n differential equations of the state variables x_i, and θ_j represents the parameters ($j = \overline{1,9}$) from the following vector:

$$\theta = [k_{X1}\ k_{X2}\ k_{X3}\ k_{A2}\ q_{Smax}\ q_{Omax}\ K_s\ K_{iA}\ K_a] \tag{D.3}$$

Figure D.1 shows a simulation of model in open-loop over 2 h, with initial conditions : $[X_0\ S_0\ A_0\ V_0]^T = [0.1\ \text{g/L}\ 0\ \text{g/L}\ 0\ \text{g/L}\ 3.5\ \text{L}]^T$ and $F_{in} = 0.02\ \text{L/h}$. The growth rate $\mu_3 = 0$ confirms that the culture is maintained in oxydo-fermentative mode.

Figures D.2 to D.4 shows the normalized parameter sensitivity functions for the biomass, substrate, and acetate concentrations. Table D.1 summuarizes the degree of influence of the model parameters on each variable.

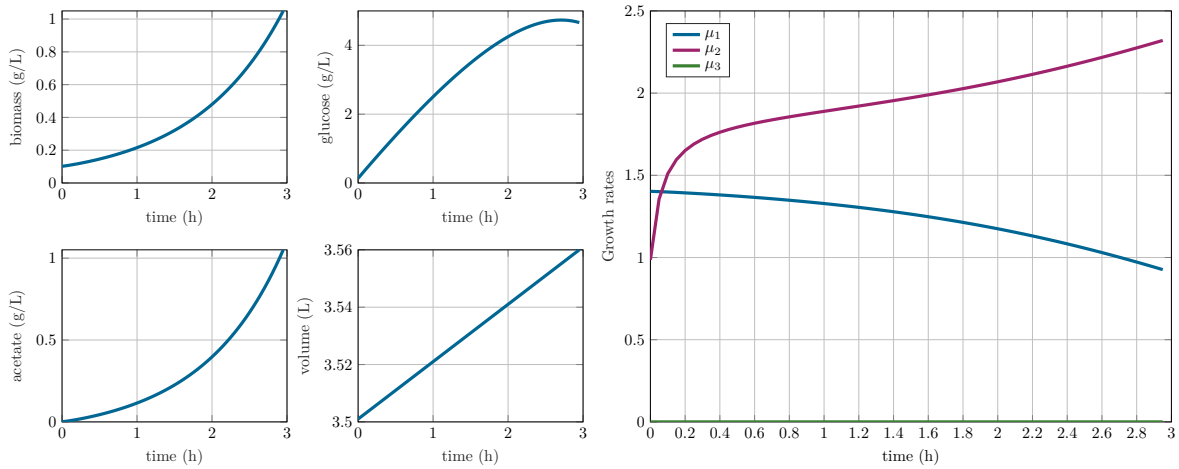


FIGURE D.1: Evolution of the state variables X, S, A, and V, and the growth rates μ_1, μ_2, and μ_3, in oxydo-fermentative mode.

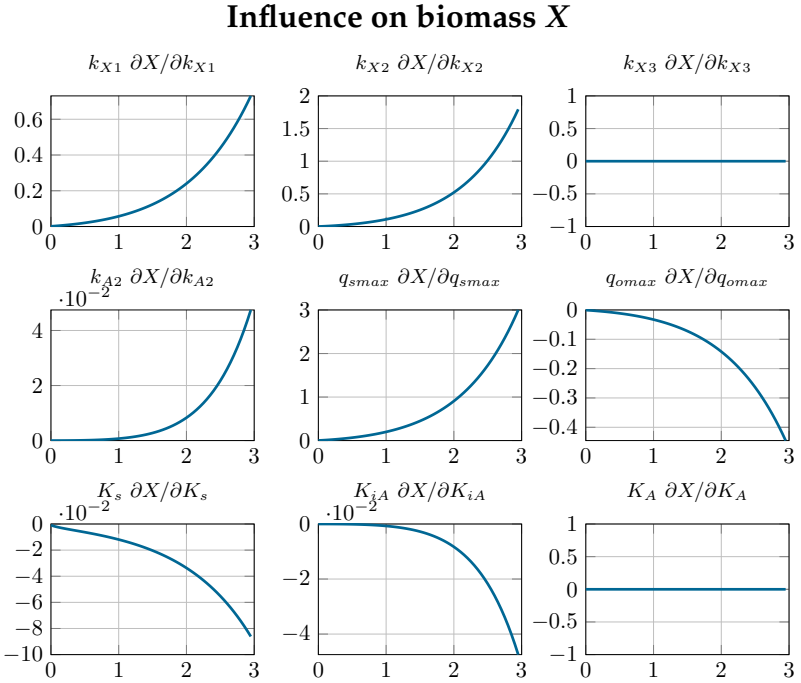


FIGURE D.2: Normalized sensitivity functions - influence on the biomass concentration X

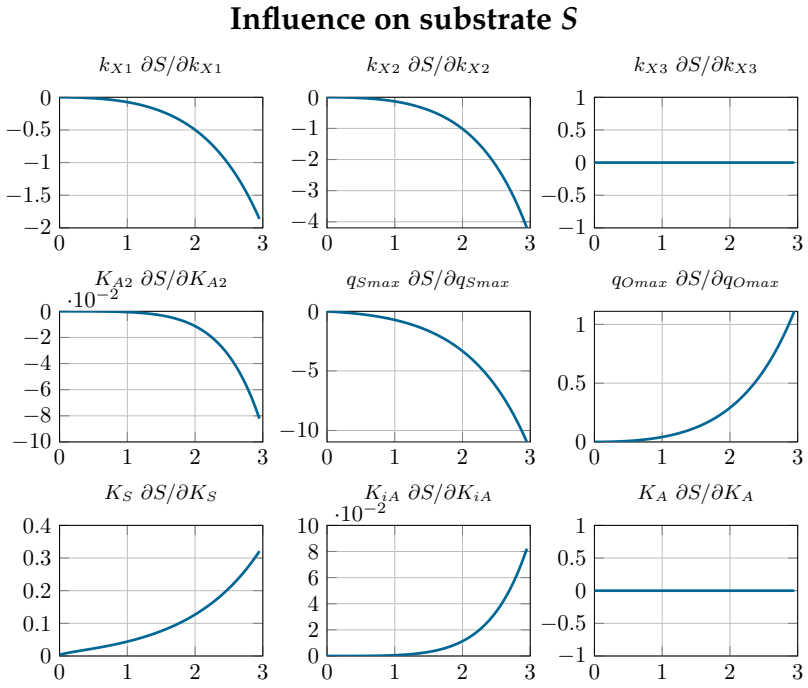


FIGURE D.3: Normalized sensitivity functions influence on the substrate concentration S

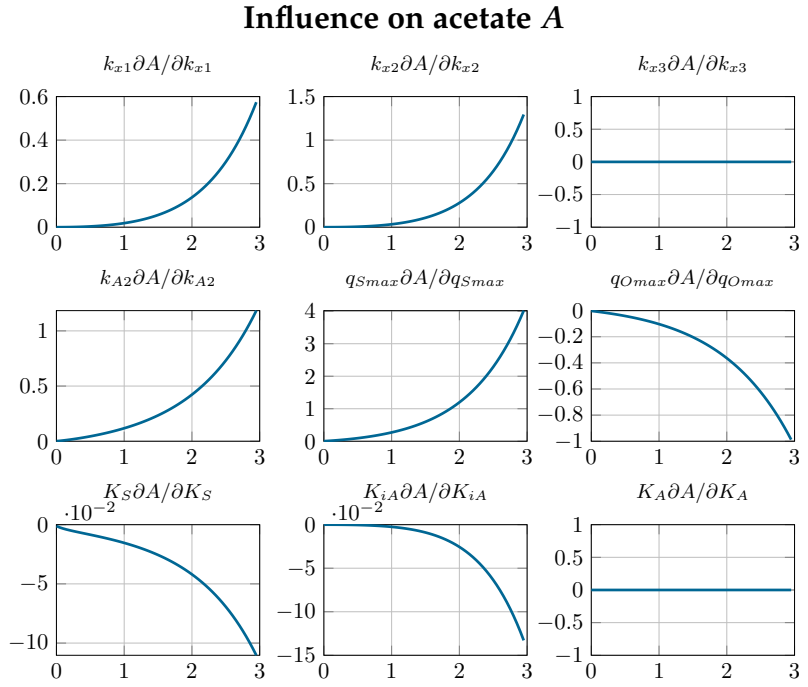


FIGURE D.4: Normalized sensitivity functions influence on the acetate concentration A

TABLE D.1: Ranking of the model parameters according to their respective influence on the state variables (From most to least influent)

Influence on X	**Influence on** S	**Influence on** A
k_{X2}	$q_{s\,\max}$	$q_{s\,\max}$
$q_{s\,\max}$	k_{X2}	k_{X2}
k_{X1}	k_{X1}	k_{A2}
K_S	$q_{O\,\max}$	$q_{O\,\max}$
K_{iA}	K_s	k_{X1}
k_{A2}	k_{A2}	K_{iA}
$q_{O\,\max}$	K_{iA}	K_S

Bibliography

Abadli, Merouane, Laurent Dewasme, Didier Dumur, Sihem Tebbani, and Alain Vande Wouwer (Oct. 2019). "Generic model control of an Escherichia coli fed-batch culture". In: *2019 23rd International Conference on System Theory, Control and Computing, ICSTCC 2019 - Proceedings*. Institute of Electrical and Electronics Engineers Inc., pp. 212–217. ISBN: 9781728106991.

Abadli, Merouane, Laurent Dewasme, Sihem Tebbani, Didier Dumur, and Alain Vande Wouwer (July 2020). "Generic model control applied to E. coli BL21(DE3) Fed-batch cultures". In: *Processes* 8.7, p. 772. ISSN: 22279717.

Abadli, Merouane, Laurent Dewasme, Sihem Tebbani, Didier Dumur, and Alain Vande Wouwer (2021). "An Experimental Assessment of Robust control and Estimation of Acetate Concentration in Escherichia coli BL21 (DE3) Fed-batch Cultures". In: *Biochemical Engineering Journal*, p. 108103.

Abadli, Merouane, Sihem Tebbani, Didier Dumur, Laurent Dewasme, and Alain Vande Wouwer (Mar. 2018a). "Nonlinear model predictive control of Escherichia coli culture". In: *37th Benelux Meeting on Systems and Control*. Soesterberg, Netherlands.

Abadli, Merouane, Sihem Tebbani, Didier Dumur, Laurent Dewasme, and Alain Vande Wouwer (June 2018b). "Nonlinear model predictive control of Escherichia coli cultures". In: *Groupe de travail sur le thème de : La Commande Prédictive Non Linéaire*. Châtillon, France.

Aehle, Mathias, Kaya Bork, Sebastian Schaepe, Artur Kuprijanov, Rüdiger Horstkorte, Rimvydas Simutis, and Andreas Lübbert (2012). "Increasing batch-to-batch reproducibility of CHO-cell cultures using a model predictive control approach". In: *Cytotechnology*. ISSN: 09209069.

Ahmad, Zulfiqar, David Crowley, Ninoslav Marina, and Sunil Kr Jha (2016). "Estimation of biosurfactant yield produced by Klebseilla sp. FKOD36 bacteria using artificial neural network approach". In: *Measurement: Journal of the International Measurement Confederation*. ISSN: 02632241.

Albert, S. and R.D. Kinley (2001). "Multivariate statistical monitoring of batch processes: An industrial case study of fermentation supervision". In: *Trends in Biotechnology* 19.2, pp. 53–62.

Alford, J.S. (2006). "Bioprocess control: Advances and challenges". In: *Computers and Chemical Engineering* 30.10-12, pp. 1464–1475.

Allgöwer, Frank, Thomas A Badgwell, Joe S Qin, James B Rawlings, and Steven J Wright (1999). "Nonlinear predictive control and moving horizon estimation—an introductory overview". In: *Advances in control*, pp. 391–449.

Allgöwer, Frank and Alex Zheng (2012). *Nonlinear model predictive control*. Vol. 26. Birkhäuser.

Amribt, Zakaria., Laurent. Dewasme, Alain. Vande Wouwer, and Philippe Bogaerts (2014). "Optimization and robustness analysis of hybridoma cell fed-batch cultures using the overflow metabolism model". In: *Bioprocess and Biosystems Engineering*. ISSN: 16157605.

Anane, Emmanuel, Diana C. López C, Peter Neubauer, and M. Nicolas Cruz Bournazou (2017). "Modelling overflow metabolism in Escherichia coli by acetate cycling". In: *Biochemical Engineering Journal*. ISSN: 1873295X.

Anderson, Brian D. O., John B. Moore, and Mansour Eslami (2008). "Optimal Filtering". In: *IEEE Transactions on Systems, Man, and Cybernetics*. ISSN: 0018-9472.

Ansari, Faiz Ahmad, Mahmoud Nasr, Ismail Rawat, and Faizal Bux (2020). "Artificial neural network and techno-economic estimation with algae-based tertiary wastewater treatment". In: *Journal of Water Process Engineering*. ISSN: 22147144.

Ariyur, Kartik B and Miroslav Krstić (2003). *Real time optimization by extremum seeking control*. Wiley Online Library.

Ashoori, Ahmad, Behzad Moshiri, Ali Khaki-Sedigh, and Mohammad Reza Bakhtiari (2009). "Optimal control of a nonlinear fed-batch fermentation process using model predictive approach". In: *Journal of Process Control*. ISSN: 09591524.

Babuška, R. and H.B. Verbruggen (1996). "An overview of fuzzy modeling for control". In: *Control Engineering Practice* 4.11, pp. 1593–1606.

Baeshen, Mohammed N., Ahmed M. Al-Hejin, Roop S. Bora, Mohamed M.M. Ahmed, Hassan A.I. Ramadan, Kulvinder S. Saini, Nabih A. Baeshen, and Elrashdy M. Redwan (2015). *Production of biopharmaceuticals in E. Coli: Current scenario and future perspectives*.

Banga, Julio R, Eva Balsa-Canto, Carmen G Moles, and Antonio A Alonso (June 2005). "Dynamic optimization of bioprocesses: efficient and robust numerical strategies". In: *Journal of biotechnology* 117.4, pp. 407–419. ISSN: 0168-1656.

Bastin, G. and D. Dochain (1990). "On-line estimation and adaptive control of bioreactors". In: *Process Measurement and Control* 1.

Beluhan, D., D. Gosak, N. Pavlović, and M. Vampola (1995). "Biomass estimation and optimal control of the baker's yeast fermentation process". In: *Computers and Chemical Engineering*. ISSN: 00981354.

Bemporad, Alberto, Manfred Morari, Vivek Dua, and Efstratios N Pistikopoulos (2002). "The explicit linear quadratic regulator for constrained systems". In: *Automatica* 38.1, pp. 3–20.

Benattia, S.E., S. Tebbani, and D. Dumur (2015). "Hierarchical Control Strategy based on Robust MPC and Integral Sliding mode - Application to a Continuous Photobioreactor". In: *IFAC-PapersOnLine* 48.23, pp. 212–217.

Bernal, Vicente, Sara Castaño-Cerezo, and Manuel Cánovas (2016). *Acetate metabolism regulation in Escherichia coli: carbon overflow, pathogenicity, and beyond*.

Bhalla, Aditya, Namita Bansal, Sudhir Kumar, Kenneth M. Bischoff, and Rajesh K. Sani (2013). *Improved lignocellulose conversion to biofuels with thermophilic bacteria and thermostable enzymes*.

Birle, S., M. A. Hussein, and T. Becker (2013). *Fuzzy logic control and soft sensing applications in food and beverage processes*.

Birol, G., C. Ündey, and A. Çinar (2002). "A modular simulation package for fed-batch fermentation: Penicillin production". In: *Computers and Chemical Engineering* 26.11, pp. 1553–1565.

Bitmead, Robert R, Michel Gevers, and Vincent Wertz (1990). "Adaptive optimal control the thinking man's GPC". In.

Bodizs, L., M. Titica, N. Faria, B. Srinivasan, D. Dochain, and D. Bonvin (2007). "Oxygen control for an industrial pilot-scale fed-batch filamentous fungal fermentation". In: *Journal of Process Control* 17.7, pp. 595–606.

Bogaerts, Ph. and D. Coutinho (2014). "Robust nonlinear state estimation of bioreactors based on H_∞ hybrid observers". In: *Computers and Chemical Engineering*. ISSN: 00981354.

Bogaerts, Ph. and A. Vande Wouwer (2003). "Software sensors for bioprocesses". In: *ISA Transactions* 42.4, pp. 547–558.

Bogaerts, Philippe (1999). "A hybrid asymptotic-Kalman observer for bioprocesses". In: *Bioprocess Engineering* 20.3, pp. 249–255.

Boiocchi, R., M. Mauricio-Iglesias, A. K. Vangsgaard, K. V. Gernaey, and G. Sin (2015). "Aeration control by monitoring the microbiological activity using fuzzy logic diagnosis and control. Application to a complete autotrophic nitrogen removal reactor". In: *Journal of Process Control*. ISSN: 09591524.

Bošković, J.D. and K.S. Narendra (1995). "Comparison of linear, nonlinear and neural-network-based adaptive controllers for a class of fed-batch fermentation processes". In: *Automatica* 31.6, pp. 817–840.

Burgard, Anthony, Mark J. Burk, Robin Osterhout, Stephen Van Dien, and Harry Yim (2016). *Development of a commercial scale process for production of 1,4-butanediol from sugar*.

Cervera, A.E., N. Petersen, A.E. Lantz, A. Larsen, and K.V. Gernaey (2009). "Application of near-infrared spectroscopy for monitoring and control of cell culture and fermentation". In: *Biotechnology Progress* 25.6, pp. 1561–1581.

Cha, Young-Lok and B. Hitzmann (2004). "Ultrasonic Measurements and its Evaluation for the Monitoring of Saccharomyces cerevisiae Cultivation". In.

Chang, L., X. Liu, and M.A. Henson (2016). "Nonlinear model predictive control of fed-batch fermentations using dynamic flux balance models". In: *Journal of Process Control* 42, pp. 137–149.

Chattaway, Thomas and Gregory Stephanopoulos (1989). "Adaptive estimation of bioreactors: monitoring plasmid instability". In: *Chemical Engineering Science*. ISSN: 00092509.

Chen, Hong and Frank Allgöwer (1998). "A quasi-infinite horizon nonlinear model predictive control scheme with guaranteed stability". In: *Automatica* 34.10, pp. 1205–1217.

Chen, L., S.K. Nguang, X.D. Chen, and X.M. Li (2004). "Modelling and optimization of fed-batch fermentation processes using dynamic neural networks and genetic algorithms". In: *Biochemical Engineering Journal* 22.1, pp. 51–61.

Chilali, Mahmoud and Pascal Gahinet (1996). "H_∞ design with pole placement constraints: An LMI approach". In: *IEEE Transactions on Automatic Control*. ISSN: 00189286.

Chou, C. Perry (2007). *Engineering cell physiology to enhance recombinant protein production in Escherichia coli*.

Cochrane, Jodie A. and Mazen Alamir (2018). "Model Predictive Control and Parameter Estimation applied to Gene Expression in Yeast". In: *IFAC-PapersOnLine*. ISSN: 24058963.

Cockshott, A. R. and I. D.L. Bogle (1999). "Modelling the effects of glucose feeding on a recombinant E. coli fermentation". In: *Bioprocess Engineering*. ISSN: 0178515X.

Contois, DE (1959). "Kinetics of bacterial growth: relationship between population density and specific growth rate of continuous cultures". In: *Microbiology* 21.1, pp. 40–50.

Costello, D. J., P. L. Lee, and P. F. Greenfield (1989). "Control of anaerobic digesters using Generic Model Control". In: *Bioprocess Engineering*. ISSN: 0178515X.

Crabtree, H.G. (1929). "Observations on the carbohydrate metabolism of tumours". In: *Biochem J* 23, pp. 536–545.

Craven, S., J. Whelan, and B. Glennon (2014). "Glucose concentration control of a fed-batch mammalian cell bioprocess using a nonlinear model predictive controller". In: *Journal of Process Control* 24.4, pp. 344–357.

De Battista, Hernán, Jesús Picó, and Enric Picó-Marco (2012). "Nonlinear PI control of fed-batch processes for growth rate regulation". In: *Journal of Process Control*. ISSN: 09591524.

De Deken, R.H. (1966). "The Crabtree effect: a regulatory system in yeast." In: *Journal of general microbiology* 44.2, pp. 149–156.

Del Rio-Chanona, Ehecatl Antonio, Pongsathorn Dechatiwongse, Dongda Zhang, Geoffrey C. Maitland, Klaus Hellgardt, Harvey Arellano-Garcia, and Vassilios S. Vassiliadis (2015). "Optimal Operation Strategy for Biohydrogen Production". In: *Industrial and Engineering Chemistry Research*. ISSN: 15205045.

DeLisa, M., H. Chae, W. Weigand, J. Valdes, G. Rao, and W. Bentley (Sept. 2001). "Generic model control of induced protein expression in high cell density cultivation of Escherichia coli using on-line GFP-fusion monitoring". In: *Bioprocess and Biosystems Engineering* 24.2, pp. 83–91.

Dewasme, L., Z. Amribt, L. O. Santos, A. L. Hantson, Ph Bogaerts, and A. Vande Wouwer (2013a). "Hybridoma cell culture optimization using nonlinear model predictive control". In: *IFAC Proceedings Volumes (IFAC-PapersOnline)*. ISBN: 9783902823588.

Dewasme, L., D. Coutinho, and A. Vande Wouwer (2011a). "Adaptive and robust linearizing control strategies for fed-batch cultures of microorganisms exhibiting overflow metabolism". In: *Lecture Notes in Electrical Engineering*. Vol. 89 LNEE. Springer, Berlin, Heidelberg, pp. 283–305. ISBN: 9783642195389.

Dewasme, L., S. Fernandes, Z. Amribt, L. O. Santos, Ph Bogaerts, and A. Vande Wouwer (2015). "State estimation and predictive control of fed-batch cultures of hybridoma cells". In: *Journal of Process Control*. ISSN: 09591524.

Dewasme, L., G. Goffaux, A. L. Hantson, and A. Vande Vande Wouwer (Feb. 2013b). "Experimental validation of an Extended Kalman Filter estimating acetate concentration in E. coli cultures". In: *Journal of Process Control*. Vol. 23. 2, pp. 148–157.

Dewasme, L., A. Richelle, P. Dehottay, P. Georges, M. Remy, Ph. Bogaerts, and A. Vande Wouwer (2010). "Linear robust control of S. cerevisiae fed-batch cultures at different scales". In: *Biochemical Engineering Journal* 53.1, pp. 26–37.

Dewasme, L., B. Srinivasan, M. Perrier, and A. Vande Wouwer (2011b). "Extremum-seeking algorithm design for fed-batch cultures of microorganisms with overflow metabolism". In: *Journal of Process Control* 21.7, pp. 1092–1104.

Diehl, Moritz, Rishi Amrit, and James B Rawlings (2010). "A Lyapunov function for economic optimizing model predictive control". In: *IEEE Transactions on Automatic Control* 56.3, pp. 703–707.

Doan, X.-T. and R. Srinivasan (2008). "Online monitoring of multi-phase batch processes using phase-based multivariate statistical process control". In: *Computers and Chemical Engineering* 32.1-2, pp. 230–243.

Dochain, D. and A. Rapaport (2018). "An asymptotic observer for batch processes with single biogas measurement". In: *IFAC-PapersOnLine*. ISSN: 24058963.

Dochain, Denis (2010). *Bioprocess Control*. London: Wiley. ISBN: 9780470611128.

Dochain, Denis, Michel Perrier, and Martin Guay (2011). "Extremum seeking control and its application to process and reaction systems: A survey". In: *Mathematics and Computers in Simulation* 82.3, pp. 369–380. ISSN: 0378-4754.

Doelle, Horst W., Ken N. Ewings, and Neil W. Hollywood (1982). "Regulation of glucose metabolism in bacterial systems". In.

Dohnal, Mirko (1985). "Fuzzy bioengineering models". In: *Biotechnology and Bioengineering*. ISSN: 10970290.

Doran, P.M. (2013). "Bioprocess development: an interdisciplinary challenge". In: *Bioprocess Engineering Principles*, pp. 3–11.

Douglas, P. L., P. S. Fountain, G. R. Sullivan, and W. Zhou (1994). "Model based control of a high purity distillation column". In: *The Canadian Journal of Chemical Engineering*. ISSN: 1939019X.

Duran-Villalobos, C.A., B. Lennox, and D. Lauri (2016a). "Multivariate batch to batch optimisation of fermentation processes incorporating validity constraints". In: *Journal of Process Control* 46, pp. 34–42.

Duran-Villalobos, Carlos A., Barry Lennox, and David Lauri (2016b). "Multivariate batch to batch optimisation of fermentation processes incorporating validity constraints". In: *Journal of Process Control* 46, pp. 34–42. ISSN: 0959-1524.

Eiteman, Mark A. and Elliot Altman (2006). *Overcoming acetate in Escherichia coli recombinant protein fermentations*.

Enjalbert, Brice, Muriel Cocaign-Bousquet, Jean Charles Portais, and Fabien Letisse (2015). "Acetate exposure determines the diauxic behavior of Escherichia coli during the glucose-acetate transition". In: *Journal of Bacteriology*. ISSN: 10985530.

Feng, Peter, Stephen Weagant, Michael Grant, and William Burkhardt (2002). "Bacteriological Analytical Manual : Enumeration of Escherichia coli and the Coliform Bacteria". In: *Bacteriological Analytical Manual*.

Fernandes, Sofia, Anne Richelle, Zakaria Amribt, Laurent Dewasme, Philippe Bogaerts, and Alain Vande Wouwer (2015). "Extended and Unscented Kalman Filter design for hybridoma cell fed-batch and continuous cultures". In: *IFAC-PapersOnLine*.

Ferreira, A.P., J.A. Lopes, and J.C. Menezes (2007). "Study of the application of multiway multivariate techniques to model data from an industrial fermentation process". In: *Analytica Chimica Acta* 595.1-2 SPEC. ISS. Pp. 120–127.

Ferreira, L.S., M.B. De Souza Jr, and R.O.M. Folly (2001). "Development of an alcohol fermentation control system based on biosensor measurements interpreted by neural networks". In: *Sensors and Actuators, B: Chemical* 75.3, pp. 166–171.

Ferrer-Miralles, Neus, Joan Domingo-Espín, José Corchero, Esther Vázquez, and Antonio Villaverde (2009). *Microbial factories for recombinant pharmaceuticals*.

Findeisen, Rolf, Frank Allgöwer, and Lorenz T Biegler (2007). *Assessment and future directions of nonlinear model predictive control*. Vol. 358. 7. Springer.

Forbes, M.G., R.S. Patwardhan, H. Hamadah, and R.B. Gopaluni (2015). "Model predictive controlin industry: Challenges and opportunities". In: *IFAC-PapersOnLine* 28.8, pp. 531–538.

Gautam, Ajay, Yun-Chung Chu, and Yeng Chai Soh (2013). "Robust H_∞ receding horizon control for a class of coordinated control problems involving dynamically decoupled subsystems". In: *IEEE Transactions on Automatic Control* 59.1, pp. 134–149.

Gauthier, J.P. and I.A.K. Kupka (1994). "Observability and observers for nonlinear systems". In: *SIAM Journal on Control and Optimization* 32.4, pp. 975–994.

Gbewonyo, K., D. Jain, G. Hunt, S. W. Drew, and B. C. Buckland (1989). "On-line analysis of avermectin fermentation cell growth kinetics in an industrial pilot plant". In: *Biotechnology and Bioengineering*. ISSN: 10970290.

Glassey, J. (2013). "Multivariate data analysis for advancing the interpretation of bioprocess measurement and monitoring data". In: *Advances in Biochemical Engineering/Biotechnology* 132, pp. 167–191.

Glassey, J., G.A. Montague, A.C. Ward, and B.V. Kara (1994). "Enhanced supervision of recombinant E. coli fermentation via artificial neural networks". In: *Process Biochemistry* 29.5, pp. 387–398.

Gonzalez, Karen, Sihem Tebbani, Filipa Lopes, Aurore Thorigné, Sébastien Givry, Didier Dumur, and Dominique Pareau (2016). "Regulation of lactic acid concentration in its bioproduction from wheat flour". In: *Control Engineering Practice* 54, pp. 202–213.

Gonzalez, Karen Vanessa, Sihem Tebbani, Didier Dumur, Filipa Lopes, Dominique Pareau, Aurore Thorigné, and Sébastien Givry (2015). "Adaptive Control of Lactic Acid Production Process from Wheat Flour". In: *IFAC-PapersOnLine* 48.8, pp. 1087–1092.

Gorrini, Federico Alberto, Jesús Miguel Zamudio Lara, Silvina Inés Biagiola, José Luis Figueroa, Héctor Hernández Escoto, Anne-Lise Hantson, and Alain Vande

Wouwer (2020). "Experimental Study of Substrate Limitation and Light Acclimation in Cultures of the Microalgae Scenedesmus obliquus—Parameter Identification and Model Predictive Control". In: *Processes* 8.12, p. 1551.

Grosz, Ron, Gregory Stephanopoulos, and K. Y. San (1984). "Studies on on-line bioreactor identification. III. Sensitivity problems with respiratory and heat evolution measurements". In: *Biotechnology and Bioengineering*. ISSN: 10970290.

Gusyatiner, M. M., Y. G. Rostova, M. Y. Kiryukhin, and A. Y. Romkina (2017). "Method for producing an L-amino acid using a bacterium of the family Enterobacteriaceae having a disrupted putrescine degradation pathway". In: *U.S. Patent 9840725B2*.

Hafidi, G., S. Tebbani, D. Dumur, and A. Vande Wouwer (2008). "Nonlinear Model Predictive Control applied to E. Coli Cultures". In: *IFAC Proceedings Volumes* 41.2, pp. 14570–14575.

Haldane, JBS (1930). "Union of enzymes with substrate". In: *Enzymes, 1st Edition. Longmans, Green & Co., London*, pp. 28–53.

Han, Keehyun, Henry C. Lim, and Juan Hong (1992). "Acetic acid formation in escherichia coli fermentation". In: *Biotechnology and Bioengineering* 39.6, pp. 663–671. ISSN: 10970290.

Hasunuma, Tomohisa and Akihiko Kondo (2012). *Consolidated bioprocessing and simultaneous saccharification and fermentation of lignocellulose to ethanol with thermotolerant yeast strains*.

Henes, B. and B. Sonnleitner (2007). "Controlled fed-batch by tracking the maximal culture capacity". In: *Journal of Biotechnology* 132.2, pp. 118–126.

Herbert, DENIS (1958). "Some principles of continuous culture". In: *Recent progress in microbiology* 1.

Hilaly, A. K., M. N. Karim, and J. C. Linden (1992). "Real time application of extended Kalman filtering in estimation and optimization of a recombinant Escherichia coli fermentation". In: *IFAC Symposia Series*. ISBN: 0080417108.

Hisbullah, M.A. Hussain, and K.B. Ramachandran (2001). "Comparative evaluation of various control schemes for fed-batch fermentation". In: *Bioprocess and Biosystems Engineering* 24.5, pp. 309–318.

Hisbullah, M.A. Hussain, and K.B. Ramachandran (2003). "Design of a fuzzy logic controller for regulating substrate feed to fed-batch fermentation". In: *Food and Bioproducts Processing: Transactions of the Institution of of Chemical Engineers, Part C* 81.2, pp. 138–146.

Holland, J.H. (1984). "Genetic algorithms and adaptation". In: *Adaptive Control of Ill-Defined Systems*, pp. 317–333.

Holmberg, Ulf and Gustaf Olsson (1985). "Simultaneous On-Line Estimation of Oxygen Transfer Rate and Respiration Rate". In: *IFAC Proceedings Volumes*. ISSN: 14746670.

Holms, W. H. (1986). "The Central Metabolic Pathways of Escherichia coli: Relationship between Flux and Control at a Branch Point, Efficiency of Conversion to Biomass, and Excretion of Acetate". In: *Current Topics in Cellular Regulation*.

Horiuchi, J.-I. and K. Hiraga (1999). "Industrial application of fuzzy control to large-scale recombinant vitamin B2 production". In: *Journal of Bioscience and Bioengineering* 87.3, pp. 365–371.

Horowitz, Joseph, Mark D. Normand, Maria G. Corradini, and Micha Peleg (2010). "Probabilistic model of microbial cell growth, division, and mortality". In: *Applied and Environmental Microbiology*. ISSN: 00992240.

Howell, J.A. and B.O. Sodipo (1985). "On-line Respirometry and Estimation of Aeration Efficiencies in an Activated Sludge Aeration Basin from Dissolved Oxygen Measurements". In: *IFAC Proceedings Volumes*. ISSN: 14746670.

Huang, Yu Liang, Titus J. Foellmer, Koon Chye Ang, Soo Beng Khoo, and Miranda G.S. Yap (1995). "Characterization and application of an on-line flow injection analysis/wall-jet electrode system for glucose monitoring during fermentation". In: *Analytica Chimica Acta*. ISSN: 00032670.

Hulhoven, X., A. Vande Vande Wouwer, and Ph Bogaerts (2006a). "Hybrid extended Luenberger-asymptotic observer for bioprocess state estimation". In: *Chemical Engineering Science*. ISSN: 00092509.

Hulhoven, Xavier, Frédéric Renard, Sandrine Dessoy, Ph Dehottay, Ph Bogaerts, and A Vande Wouwer (2006b). "Monitoring and control of a bioprocess for malaria vaccine production". In: *IFAC Proceedings Volumes* 39.9, pp. 143–148.

Insel, G., G. Celikyilmaz, E. Ucisik-Akkaya, K. Yesiladali, Z. P. Cakar, C. Tamerler, and D. Orhon (2007). "Respirometric evaluation and modeling of glucose utilization by Escherichia coli under aerobic and mesophilic cultivation conditions". In: *Biotechnology and Bioengineering*. ISSN: 00063592.

Isidori, Alberto, M. Thoma, E. D. Sontag, B. W. Dickinson, A. Fettweis, J. L. Massey, and J. W. Modestino (1995). *Nonlinear Control Systems*. 3rd. Berlin, Heidelberg: Springer-Verlag. ISBN: 3540199160.

Jaén, Karim E, Juan-Carlos Sigala, Roberto Olivares-Hernández, Karsten Niehaus, and Alvaro R Lara (2017). "Heterogeneous oxygen availability affects the titer and topology but not the fidelity of plasmid DNA produced by Escherichia coli". In: *BMC biotechnology* 17.1, pp. 1–12.

Jana, S. and J. K. Deb (2005). *Strategies for efficient production of heterologous proteins in Escherichia coli*.

Jenzsch, M., S. Gnoth, M. Beck, M. Kleinschmidt, R. Simutis, and A. Lübbert (2006a). "Open-loop control of the biomass concentration within the growth phase of recombinant protein production processes". In: *Journal of Biotechnology* 127.1, pp. 84–94.

Jenzsch, Marco, Rimvydas Simutis, and Andreas Luebbert (2006b). "Generic model control of the specific growth rate in recombinant Escherichia coli cultivations". In: *Journal of Biotechnology* 122.4, pp. 483–493.

Johnsson, O., J. Andersson, G. Lidén, C. Johnsson, and T. Hägglund (2013). "Feed rate control in fed-batch fermentations based on frequency content analysis". In: *Biotechnology Progress* 29.3, pp. 817–824.

Julier, Simon J. and Jeffrey K. Uhlmann (1997). "New extension of the Kalman filter to nonlinear systems". In: *Signal Processing, Sensor Fusion, and Target Recognition VI*. ISBN: 0819424838.

Julier, Simon J. and Jeffrey K. Uhlmann (2004). "Unscented filtering and nonlinear estimation". In: *Proceedings of the IEEE*.

Kadir, Tuty A.A., Ahmad A. Mannan, Andrzej M. Kierzek, Johnjoe McFadden, and Kazuyuki Shimizu (2010). "Modeling and simulation of the main metabolism in Escherichia coli and its several single-gene knockout mutants with experimental verification". In: *Microbial Cell Factories*. ISSN: 14752859.

Karim, M. N. and S. L. Rivera (1992). *Artificial neural networks in bioprocess state estimation.*

Keerthi, SS a and Elmer G Gilbert (1988). "Optimal infinite-horizon feedback laws for a general class of constrained discrete-time systems: Stability and moving-horizon approximations". In: *Journal of optimization theory and applications* 57.2, pp. 265–293.

Kesson, M., P. Hagander, and J.P. Axelsson (2001). "Avoiding acetate accumulation in Escherichia coli cultures using feedback control of glucose feeding". In: *Biotechnology and Bioengineering* 73.3, pp. 223–230.

Kirkpatrick, Christopher, Lisa M. Maurer, Nikki E. Oyelakin, Yuliya N. Yoncheva, Russell Maurer, and Joan L. Slonczewski (2001). "Acetate and formate stress: Opposite responses in the proteome of Escherichia coli". In: *Journal of Bacteriology*. ISSN: 00219193.

Kleman, G. L. and W. R. Strohl (1994). "Acetate metabolism by Escherichia coli in high-cell-density fermentation". In: *Applied and Environmental Microbiology*. ISSN: 00992240.

Ko, Yun-Fei -F, William E. Bentley, and William A. Weigand (1994). "A metabolic model of cellular energetics and carbon flux during aerobic Escherichia coli fermentation". In: *Biotechnology and Bioengineering*. ISSN: 10970290.

Koebmann, Brian J., Hans V. Westerhoff, Jacky L. Snoep, Dan Nilsson, and Peter R. Jensen (2002). "The glycolytic flux in Escherichia coli is controlled by the demand for ATP". In: *Journal of Bacteriology*. ISSN: 00219193.

Kojima, Hiroyuki, Yuri Ogawa, Kazue Kawamura, and Konosuke Sano (2000). "Method of producing L-lysine by fermentation". In: *U.S. Patent 6040160A*.

Kovárová-Kovar, K., S. Gehlen, A. Kunze, T. Keller, R. Von Däniken, M. Kolb, and A.P.G.M. Van Loon (2000a). "Application of model-predictive control based on artificial neural networks to optimize the fed-batch process for riboflavin production". In: *Journal of Biotechnology* 79.1, pp. 39–52.

Kovárová-Kovar, Karin, Stefan Gehlen, Axel Kunze, Thoralf Keller, Ralph von Däniken, Michael Kolb, and Adolphus P.G.M. van Loon (2000b). "Application of model-predictive control based on artificial neural networks to optimize the fed-batch process for riboflavin production". In: *Journal of Biotechnology* 79.1, pp. 39–52. ISSN: 0168-1656.

Krämer, Dominik and Rudibert King (2016). "On-line monitoring of substrates and biomass using near-infrared spectroscopy and model-based state estimation for enzyme production by S. cerevisiae". In: *IFAC-PapersOnLine*. ISSN: 24058963.

Kuprijanov, A., S. Schaepe, R. Simutis, and A. Lübbert (2013a). "Model predictive control made accessible to professional automation systems in fermentation technology". In: *Biosyst. Inf. Technol.* 2, pp. 26–31.

Kuprijanov, A., S. Schaepe, R. Simutis, and A. Lübbert (2013b). "Model predictive control made accessible to professional automation systems in fermentation technology". In: *Biosyst. Inf. Technol* 2.1, pp. 26–31.

Kwakernaak, H. and R. Sivan (1974). "Linear Optimal Control Systems". In: *IEEE Transactions on Automatic Control*. ISSN: 15582523.

Landau, I.D., R. Lozano, M. M'Saad, and A. Karimi (2011). "Adaptive regulation—Rejection of unknown disturbances". In: *Communications and Control Engineering* 9780857296634, pp. 477–497.

Langson, Wilbur, Ioannis Chryssochoos, SV Raković, and David Q Mayne (2004). "Robust model predictive control using tubes". In: *Automatica* 40.1, pp. 125–133.

Laurí, D., B. Lennox, and J. Camacho (2014). "Model predictive control for batch processes: Ensuring validity of predictions". In: *Journal of Process Control* 24.1, pp. 239–249.

Lee, J., S.Y. Lee, S. Park, and A.P.J. Middelberg (1999). "Control of fed-batch fermentations". In: *Biotechnology Advances* 17.1, pp. 29–48.

Lee, P.L. and G.R. Sullivan (1988). "Generic model control (GMC)". In: *Computers & Chemical Engineering* 12.6, pp. 573–580.

Lee, Sang Yup (1996). "High cell-density culture of Escherichia coli". In: *Trends in Biotechnology* 14.3, pp. 98–105.

Lee, Seung Chul, Young Bo Hwang, Ho Nam Chang, and Yong Keun Chang (1991). "Adaptive control of dissolved oxygen concentration in a bioreactor". In: *Biotechnology and Bioengineering*. ISSN: 10970290.

Lewis, Frank l., Lihua Xie, and Dan Popa (2007). *Optimal and Robust Estimation: With an Introduction to Stochastic Control Theory*. ISBN: 9781420008296.

Lin, H. Y., B. Mathiszik, B. Xu, S. O. Enfors, and P. Neubauer (2001). "Determination of the maximum specific uptake capacities for glucose and oxygen in glucose-limited fed-batch cultivations of Escherichia coli". In: *Biotechnology and Bioengineering*. ISSN: 00063592.

Lisci, Silvia, Massimiliano Grosso, and Stefania Tronci (2020). "A geometric observer-assisted approach to tailor state estimation in a bioreactor for ethanol production". In: *Processes*. ISSN: 22279717.

Liu, D. (2019). "Escherichia coli". In: *Encyclopedia of Microbiology*. ISBN: 9780128117378.

Logist, Filip, Boris Houska, Moritz Diehl, and Jan F. Van Impe (2011). "Robust multi-objective optimal control of uncertain (bio)chemical processes". In: *Chemical Engineering Science*. ISSN: 00092509.

Lübbert, A. and S. Bay Jorgensen (2001). "Bioreactor performance: A more scientific approach for practice". In: *Journal of Biotechnology* 85.2, pp. 187–212.

Lübbert, A. and R. Simutis (1994). "Using measurement data in bioprocess modelling and control". In: *Trends in Biotechnology* 12.8, pp. 304–311.

Luli, G. W. and W. R. Strohl (1990). "Comparison of growth, acetate production, and acetate inhibition of Escherichia coli strains in batch and fed-batch fermentations". In: *Applied and Environmental Microbiology* 56.4, pp. 1004–1011. ISSN: 00992240.

Luttmann, R. et al. (2012). "Soft sensors in bioprocessing: A status report and recommendations". In: *Biotechnology Journal* 7.8, pp. 1040–1048.

Marafioti, Giancarlo, Sihem Tebbani, Dominique Beauvois, Giuliana Becerra, Arsene Isambert, and Morten Hovd (2009). "Unscented Kalman filter state and parameter estimation in a photobioreactor for microalgae production". In: *IFAC Proceedings Volumes (IFAC-PapersOnline)*. ISBN: 9783902661548.

Markana, Anilkumar, Nitin Padhiyar, and Kannan Moudgalya (2018). "Multicriterion control of a bioprocess in fed-batch reactor using EKF based economic model predictive control". In: *Chemical Engineering Research and Design*. ISSN: 02638762.

Márquez-Vera, M. A., L. E. Ramos-Velasco, and B. D. Balderrama-Hernández (2018). "Stable fuzzy control and observer via LMIs in a fermentation process". In: *Journal of Computational Science*. ISSN: 18777503.

Mears, L., R. Nørregård, G. Sin, K.V. Gernaey, S.M. Stocks, M.O. Albaek, and K. Villez (2016). "Functional unfold principal component regression methodology for analysis of industrial batch process data". In: *AIChE Journal* 62.6, pp. 1986–1994.

Mears, Lisa, Stuart M. Stocks, Gürkan Sin, and Krist V. Gernaey (2017). "A review of control strategies for manipulating the feed rate in fed-batch fermentation processes". In: *Journal of Biotechnology* 245, pp. 34–46. ISSN: 0168-1656.

Melo, Alessandra Gonçalves de, Sébastien Levesque, and Sylvain Moineau (2018). *Phages as friends and enemies in food processing*.

Meyer, Hans-Peter and Diego R Schmidhalter (2012). "Microbial Expression Systems and Manufacturing from a Market and Economic Perspective". In: *Innovations in Biotechnology*.

Michaelis, Leonor, Maud L Menten, et al. (1913). "Die kinetik der invertinwirkung". In: *Biochem. z* 49.333-369, p. 352.

Montague, G.A., A.J. Morris, and A.C. Ward (1989). "Fermentation monitoring and control: A perspective". In: *Biotechnology and Genetic Engineering Reviews* 7.1, pp. 147–188.

Morari, Manfred and Jay H Lee (1999). "Model predictive control: past, present and future". In: *Computers & Chemical Engineering* 23.4-5, pp. 667–682.

Moulton, Gus G. (2014). *Fed-Batch Fermentation: A Practical Guide to Scalable Recombinant Protein Production in Escherichia Coli*. ISBN: 9781908818331.

Müller, Matthias, Wolfram Meusel, Ute Husemann, Gerhard Greller, and Matthias Kraume (Jan. 2018). "Application of heat compensation calorimetry to an E. coli fed-batch process". In: *Journal of Biotechnology* 266, pp. 133–143. ISSN: 18734863.

Murugan, Chitra and Pappa Natarajan (2019). "Estimation of fungal biomass using multiphase artificial neural network based dynamic soft sensor". In: *Journal of Microbiological Methods*. ISSN: 18728359.

Nelson, David L. and Michael M. Cox (2017). "Lehninger Principles of Biochemistry 7th". In: *W.H. Freeman and Company*. ISSN: 00063568.

Neubauer, P., H. Y. Lin, and B. Mathiszik (2003). "Metabolic load of recombinant protein production: Inhibition of cellular capacities for glucose uptake and respiration after induction of a heterologous gene in Escherichia coli". In: *Biotechnology and Bioengineering*. ISSN: 00063592.

Nomikos, P. and J.F. MacGregor (1994). "Monitoring batch processes using multiway principal component analysis". In: *AIChE Journal* 40.8, pp. 1361–1375.

Nomikos, P. and J.F. MacGregor (1995). "Multi-way partial least squares in monitoring batch processes". In: *Chemometrics and Intelligent Laboratory Systems* 30.1, pp. 97–108.

Ohtake, Toshiyuki, Sammy Pontrelli, Walter A. Laviña, James C. Liao, Sastia P. Putri, and Eiichiro Fukusaki (2017). "Metabolomics-driven approach to solving a CoA imbalance for improved 1-butanol production in Escherichia coli". In: *Metabolic Engineering*. ISSN: 10967184.

Oliveira, R., J.J. Clemente, A.E. Cunha, and M.J.T. Carrondo (2005). "Adaptive dissolved oxygen control through the glycerol feeding in a recombinant Pichia pastoris cultivation in conditions of oxygen transfer limitation". In: *Journal of Biotechnology* 116.1, pp. 35–50.

Oliveira, R., R. Simutis, and S. Feyo De Azevedo (2004). "Design of a stable adaptive controller for driving aerobic fermentation processes near maximum oxygen transfer capacity". In: *Journal of Process Control* 14.6, pp. 617–626.

Olkin, Ingram and Allan R Sampson (2001). "Multivariate analysis: overview". In.

Pablos, Tania E, Juan Carlos Sigala, Sylvie Le Borgne, and Alvaro R Lara (2014). "Aerobic expression of Vitreoscilla hemoglobin efficiently reduces overflow metabolism in Escherichia coli". In: *Biotechnology journal* 9.6, pp. 791–799.

Park, Jin Hwan, Yu Sin Jang, Jeong Wook Lee, and Sang Yup Lee (2011). "Escherichia coli W as a new platform strain for the enhanced production of L-Valine by systems metabolic engineering". In: *Biotechnology and Bioengineering*. ISSN: 00063592.

Patnaik, Pratap R. (2005). "The extended Kalman filter as a noise modulator for continuous yeast cultures under monotonic, oscillating and chaotic conditions". In: *Chemical Engineering Journal*. ISSN: 13858947.

Peebo, Karl, Kaspar Valgepea, Andres Maser, Ranno Nahku, Kaarel Adamberg, and Raivo Vilu (2015). "Proteome reallocation in Escherichia coli with increasing specific growth rate". In: *Molecular BioSystems*. ISSN: 17422051.

Peng, J., F. Meng, and Y. Ai (2013). "Time-dependent fermentation control strategies for enhancing synthesis of marine bacteriocin 1701 using artificial neural network and genetic algorithm". In: *Bioresource Technology* 138, pp. 345–352.

Peter, Ed. and L. Lee (1993). *Nonlinear Process Control: Applications of Generic Model Control*. London Ltd. UK: Springer-Verlag.

Phue, Je Nie, Jun Lee Sang, Loc Trinh, and Joseph Shiloach (2008). "Modified Escherichia coli B (BL21), a superior producer of plasmid DNA compared with Escherichia coli K (DH5α)". In: *Biotechnology and Bioengineering*. ISSN: 00063592.

Picó, J, F Garelli, H De Battista, and RJ Mantz (2009). "Geometric invariance and reference conditioning ideas for control of overflow metabolism". In: *Journal of Process Control* 19.10, pp. 1617–1626.

Pinhal, Stéphane, Delphine Ropers, Johannes Geiselmann, and Hidde De Jong (2019). "Acetate metabolism and the inhibition of bacterial growth by acetate". In: *Journal of Bacteriology*. ISSN: 10985530.

Pontrelli, Sammy, Tsan-Yu Chiu, Ethan I. Lan, Frederic Y.-H. Chen, Peiching Chang, and James C. Liao (2018). "Escherichia coli as a host for metabolic engineering". In: *Metabolic Engineering* 50, pp. 16–46.

Preuß, Karlheinz, Marie Véronique Le Lann, Jacques Proth, and Hervé Pingaud (2000). "Modelling and predictive control of fed-batch yeast growth on industrial pilot plant scale". In: *Chemical Engineering Journal*. ISSN: 13858947.

Prieto, María A., Eduardo Díaz, and José L. García (1996). "Molecular characterization of the 4-hydroxyphenylacetate catabolic pathway of Escherichia coli W: Engineering a mobile aromatic degradative cluster". In: *Journal of Bacteriology*. ISSN: 00219193.

Qin, S. Joe and Thomas A. Badgwell (2000). "An Overview of Nonlinear Model Predictive Control Applications". In: *Nonlinear Model Predictive Control*.

Qin, S.J. and T.A. Badgwell (2003). "A survey of industrial model predictive control technology". In: *Control Engineering Practice* 11.7, pp. 733–764.

Renard, F., A. Vande Wouwer, S. Valentinotti, and D. Dumur (2006). "A practical robust control scheme for yeast fed-batch cultures – An experimental validation". In: *Journal of Process Control* 16.8, pp. 855–864.

Retamal, C., L. Dewasme, A. L. Hantson, and A. Vande Wouwer (July 2018). "Parameter estimation of a dynamic model of Escherichia coli fed-batch cultures". In: *Biochemical Engineering Journal* 135, pp. 22–35. ISSN: 1873295X.

Riesenberg, D., V. Schulz, W.A. Knorre, H.-D. Pohl, D. Korz, E.A. Sanders, A. Roß, and W.-D. Deckwer (1991). "High cell density cultivation of Escherichia coli at controlled specific growth rate". In: *Journal of Biotechnology* 20.1, pp. 17–27.

Rio-Chanona, Ehecatl Antonio del, Dongda Zhang, and Vassilios S. Vassiliadis (2016). "Model-based real-time optimisation of a fed-batch cyanobacterial hydrogen production process using economic model predictive control strategy". In: *Chemical Engineering Science*. ISSN: 00092509.

Rocha, I. and E.C. Ferreira (2002). "Model-based Adaptive Control Of Acetate Concentration During The Production Of Recombinant Proteins With E. Coli". In: *IFAC Proceedings Volumes* 35.1, pp. 461–466.

Rocha, I., Ana Veloso, and Eugénio Ferreira (Feb. 2006). "Design of estimators for specific growth rate control in a fed-batch E. coli fermentation". In.

Rocha, Isabel (2003). *Model-based strategies for computer-aided operation of recombinant E. coli fermentation*.

Rocha, Isabel, Ana Veloso, Sonia Carneiro, Eugenio Ferreira, and Rafael Costa (2008). "Implementation of a specific rate controller in a fed-batch E. coli fermentation". In: *IFAC Proceedings Volumes (IFAC-PapersOnline)*. ISBN: 9783902661005.

Rothen, S. A., M. Sauer, B. Sonnleitner, and B. Witholt (1998). "Growth characteristics of Escherichia coli HB101[pGEc47] on defined medium". In: *Biotechnology and Bioengineering* 58.1, pp. 92–100.

Rumelhart, D.E., G.E. Hinton, and R.J. Williams (1986). "Learning representations by back-propagating errors". In: *Nature* 323.6088, pp. 533–536.

Sabra, W, C Groeger, An-Ping Zeng, Antonino Baez, Kwang Myung Cho, and James C Liao (2016). "Microbial Cell Factories for Diol Production". In: *Bioreactor Engineering Research and Industrial Applications I: Cell Factories*. Ed. by Qin Ye, Jie Bao, and Jian-Jiang Zhong. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 165–197. ISBN: 978-3-662-49161-4.

Sagmeister, P., P. Wechselberger, M. Jazini, A. Meitz, T. Langemann, and C. Herwig (2013). "Soft sensor assisted dynamic bioprocess control: Efficient tools for bioprocess development". In: *Chemical Engineering Science* 96, pp. 190–198.

Samson, Julie E. and Sylvain Moineau (2013). *Bacteriophages in food fermentations: New frontiers in a continuous arms race*.

Sanford, Karl, Gopal Chotani, Nathan Danielson, and James A. Zahn (2016). *Scaling up of renewable chemicals*.

Santos, L.O., L. Dewasme, D. Coutinho, and A. Vande Wouwer (2012a). "Nonlinear model predictive control of fed-batch cultures of micro-organisms exhibiting overflow metabolism: Assessment and robustness". In: *Computers & Chemical Engineering* 39, pp. 143–151. ISSN: 0098-1354.

Santos, L.O., L. Dewasme, D. Coutinho, and A. Vande Wouwer (2012b). "Nonlinear model predictive control of fed-batch cultures of micro-organisms exhibiting overflow metabolism: Assessment and robustness". In: *Computers and Chemical Engineering* 39, pp. 143–151.

Santos, L.O., L. Dewasme, and A. Vande Wouwer (2010). "Nonlinear model predictive control of fed-batch cultures of E. coli: performance and robustness analysis". In: *IFAC Proceedings Volumes* 43.14, pp. 1046–1051.

Särkkä, Simo (2007). "On unscented Kalman filtering for state estimation of continuous-time nonlinear systems". In: *IEEE Transactions on Automatic Control*. ISSN: 00189286.

Schugerl, K. and K.H. Bellgardt (2014). *Bioreaction Engineering: Modeling and Control*. ISBN: 9780874216561.

Scokaert, Pierre OM and David Q Mayne (1998). "Min-max feedback model predictive control for constrained linear systems". In: *IEEE Transactions on Automatic control* 43.8, pp. 1136–1142.

Scokaert, Pierre OM, David Q Mayne, and James B Rawlings (1999). "Suboptimal model predictive control (feasibility implies stability)". In: *IEEE Transactions on Automatic Control* 44.3, pp. 648–654.

Seborg, D.E., T.F. Edgar, and D.A. Mellichamp (1989). In: *Process Dynamics and Control*.

Shen, Claire R, Ethan I Lan, Yasumasa Dekishima, Antonino Baez, Kwang Myung Cho, and James C Liao (2011). "Driving Forces Enable High-Titer Anaerobic 1-Butanol Synthesis in ¡span class="named-content genus-species" id="named-content-1"¿Escherichia coli¡/span¿". In: *Applied and Environmental Microbiology*.

Shiloach, Joseph and Ursula Rinas (2009). "Glucose and Acetate Metabolism in E. coli – System Level Analysis and Biotechnological Applications in Protein Production Processes". In: *Systems Biology and Biotechnology of Escherichia coli*. Ed. by Sang Yup Lee. Dordrecht: Springer Netherlands, pp. 377–400. ISBN: 978-1-4020-9394-4.

Shin, Sooan, Dong Eun Chang, and Jae Gu Pan (2009). "Acetate consumption activity directly determines the level of acetate accumulation during Escherichia coli W3110 growth". In: *Journal of Microbiology and Biotechnology*. ISSN: 10177825.

Simutis, Rimvydas. and Andreas. Lübbert (2017). "Hybrid approach to state estimation for bioprocess control". In: *Bioengineering*. ISSN: 23065354.

Smets, I.Y., J.E. Claes, E.J. November, G.P. Bastin, and J.F. Van Impe (2004). "Optimal adaptive control of (bio)chemical reactors: Past, present and future". In: *Journal of Process Control* 14.7, pp. 795–805.

Soccol, Carlos Ricardo, Michele Rigon Spier, Luciana Porto de Souza Vandenberghe, Adriane Bianchi Pedroni Medeiros, Luiz Alberto Junior Letti, and Wilerson Sturm (2012). "Data acquisition systems in bioprocesses". In: *Data acquisition applications. Rijeka (Croatia): InTech*, pp. 79–106.

Sonnleitner, B. (2013). "Automated measurement and monitoring of bioprocesses: Key elements of the M3C strategy". In: *Advances in Biochemical Engineering/Biotechnology* 132, pp. 1–33.

Sonnleitner, B. and O. Käppeli (1986). "Growth of Saccharomyces cerevisiae is controlled by its limited respiratory capacity: Formulation and verification of a hypothesis". In: *Biotechnology and Bioengineering* 28.6, pp. 927–937.

Soons, Z. I.T.A., J. Shi, L. A. Van Der Pol, G. Van Straten, and A. J.B. Van Boxtel (2007). "Biomass growth and kLa estimation using online and offline measurements". In: *IFAC Proceedings Volumes (IFAC-PapersOnline)*. ISBN: 9783902661616.

Soons, Z.I.T.A., J.A. Voogt, G. van Straten, and A.J.B. van Boxtel (2006a). "Constant specific growth rate in fed-batch cultivation of Bordetella pertussis using adaptive control". In: *Journal of Biotechnology* 125.2, pp. 252–268.

Soons, Z.I.T.A., J.A. Voogt, G. van Straten, and A.J.B. van Boxtel (2006b). "Constant specific growth rate in fed-batch cultivation of Bordetella pertussis using adaptive control". In: *Journal of Biotechnology* 125.2, pp. 252–268. ISSN: 0168-1656.

Srinivasan, B. W., B. Sedghi, and D Bonvin (2001). "Terminal-cost optimization of a class of hybrid systems". In: *EEC*.

Stanke, M. and B. Hitzmann (2013). "Automatic control of bioprocesses". In: *Advances in Biochemical Engineering/Biotechnology* 132, pp. 35–63.

Stratz, Simone and Petra S Dittrich (2015). "A Microfluidic Device for Immunoassay-Based Protein Analysis of Single E. coli Bacteria". In: *Single Cell Protein Analysis*. Springer, pp. 11–25.

Sturm, Jos F (1999). "Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones". In: *Optimization methods and software* 11.1-4, pp. 625–653.

Tao, F., J. Y. Miao, G. Y. Shi, and K. C. Zhang (2005). "Ethanol fermentation by an acid-tolerant Zymomonas mobilis under non-sterilized condition". In: *Process Biochemistry*. ISSN: 00329592.

Tebbani, S., D. Dumur, G. Hafidi, and A. Vande Wouwer (2010). "Nonlinear Predictive Control of Fed-Batch Cultures of Escherichia coli". In: *Chemical Engineering & Technology* 33.7, pp. 1112–1124.

Tebbani, Sihem., Mariana. Titica, Sergiu Caraman, and Lionel Boillereaux (2013). "Estimation of chlamydomonas reinhardtii growth in a torus photobioreactor". In: *IFAC Proceedings Volumes (IFAC-PapersOnline)*.

Tomson, K., T. Paalme, P. S. Laakso, and R. Vilu (1995). "Automatic laboratory-scale fed-batch procedure for production of recombinant proteins using inducible expression systems of Escherichia coli". In: *Biotechnology Techniques*. ISSN: 0951208X.

Ulonska, Sophia, Daniel Waldschitz, Julian Kager, and Christoph Herwig (2018). "Model predictive control in comparison to elemental balance control in an E. coli fed-batch". In: *Chemical Engineering Science*. ISSN: 00092509.

Valentinotti, S., C. Cannizzaro, B. Srinivasan, and D. Bonvin (2004). "An Optimal Operating Strategy for Fed-Batch Fermentations by Feeding the Overflow Metabolite". In: *IFAC Proceedings Volumes* 37.1, pp. 365–371. ISSN: 1474-6670.

Valentinotti, S., B. Srinivasan, U. Holmberg, D. Bonvin, C. Cannizzaro, M. Rhiel, and U. von Stockar (2003). "Optimal operation of fed-batch fermentations via adaptive control of overflow metabolite". In: *Control Engineering Practice*. ISSN: 09670661.

Valgepea, Kaspar, Kaarel Adamberg, Ranno Nahku, Petri Jaan Lahtvee, Liisa Arike, and Raivo Vilu (2010). "Systems biology approach reveals that overflow metabolism of acetate in Escherichia coli is triggered by carbon catabolite repression of acetyl-CoA synthetase". In: *BMC Systems Biology*. ISSN: 17520509.

Van De Walle, Michele and Joseph Shiloach (1998). "Proposed mechanism of acetate accumulation in two recombinant Escherichia coli strains during high density fermentation". In: *Biotechnology and Bioengineering*.

Velut, S., L. de Maré, and P. Hagander (2007). "Bioreactor control using a probing feeding strategy and mid-ranging control". In: *Control Engineering Practice* 15.2, pp. 135–147.

Vemuri, G. N., E. Altman, D. P. Sangurdekar, A. B. Khodursky, and M. A. Eiteman (2006). "Overflow metabolism in Escherichia coli during steady-state growth: Transcriptional regulation and effect of the redox ratio". In: *Applied and Environmental Microbiology*. ISSN: 00992240.

Wang, Jianlin, Xuying Feng, Liqiang Zhao, and Tao Yu (2010a). "Unscented transformation based Robust Kalman filter and its applications in fermentation process". In: *Chinese Journal of Chemical Engineering*. ISSN: 10049541.

Wang, Rui, Guo-Ping Liu, Wei Wang, David Rees, and Yun-Bo Zhao (2009). "H_∞ Control for Networked Predictive Control Systems Based on the Switched Lyapunov Function Method". In: *IEEE transactions on industrial electronics* 57.10, pp. 3565–3571.

Wang, Rui, Bo Wang, Guo-Ping Liu, Wei Wang, and David Rees (2010b). "H_∞ Controller Design for Networked Predictive Control Systems Based on the Average Dwell-Time Approach". In: *IEEE Transactions on Circuits and Systems II: Express Briefs* 57.4, pp. 310–314.

Wechselberger, P., P. Sagmeister, H. Engelking, T. Schmidt, J. Wenger, and C. Herwig (2012). "Efficient feeding profile optimization for recombinant protein production using physiological information". In: *Bioprocess and Biosystems Engineering* 35.9, pp. 1637–1649.

Welch, Greg and Gary Bishop (1995). *An Introduction to the Kalman Filter*. Chapel Hill, NC, USA: University of North Carolina at Chapel Hill.

Wernick, David G., Sammy P. Pontrelli, Alexander W. Pollock, and James C. Liao (2016). "Sustainable biorefining in wastewater by engineered extreme alkaliphile Bacillus marmarensis". In: *Scientific Reports*. ISSN: 20452322.

Wolfe, Alan J. (2005). "The Acetate Switch". In: *Microbiology and Molecular Biology Reviews*. ISSN: 1092-2172.

Wood, Richard (1972). "Automatic control systems". In: *Water Research*. ISSN: 00431354.

Xu, Bo, Mehmedalija Jahic, and Sven Olof Enfors (1999). "Modeling of overflow metabolism in batch and fed-batch cultures of Escherichia coli". In: *Biotechnology Progress*. ISSN: 87567938.

Xu, Youqiang, Haipei Chu, Chao Gao, Fei Tao, Zikang Zhou, Kun Li, Lixiang Li, Cuiqing Ma, and Ping Xu (2014). "Systematic metabolic engineering of Escherichia coli for high-yield production of fuel bio-chemical 2,3-butanediol". In: *Metabolic Engineering*. ISSN: 10967176.

Yuan, J.Q., Y.F. Xue, K.L. Hu, H.T. Wu, and Q. Jia (2009). "On-line application oriented optimal scheduling for penicillin fed-batch fermentation". In: *Chemical Engineering and Processing: Process Intensification* 48.2, pp. 651–658. ISSN: 0255-2701.

Zabriskie, Dane W. and Arthur E. Humphrey (1978). "Real-time estimation of aerobic batch fermentation biomass concentration by component balancing". In: *AIChE Journal*. ISSN: 15475905.

Zadeh, Lotfi A. (1994). "Soft Computing and Fuzzy Logic". In: *IEEE Software*. ISSN: 07407459.

Zeitz, M. (1984). "Observability canonical (phase-variable) form for non-linear time-variable systems". In: *International Journal of Systems Science* 15.9, pp. 949–958.

Zhang, H. and B. Lennox (2004). "Integrated condition monitoring and control of fed-batch fermentation processes". In: *Journal of Process Control* 14.1, pp. 41–50.

Zhang, X.-C., A. Visala, A. Halme, and P. Linko (1994). "Functional state modeling and fuzzy control of fed-batch aerobic baker's yeast process". In: *Journal of Biotechnology* 37.1, pp. 1–10.

Zhou, Weibiao, Peter L. Lee, and Gerald R. Sullivan (1992). "Robust stability analysis of Generic Model Control". In: *Chemical Engineering Communications* 117.1, pp. 41–72.

université PARIS-SACLAY

ÉCOLE DOCTORALE
Sciences et technologies de l'information et de la communication (STIC)

Titre: Commande robuste de cultures fed-batch de Escherichia coli

Mots clés: Escherichia coli, Commande robuste, Biotechnologie, Fed-batch

Résumé: Escherichia coli est un hôte cellulaire très répandu pour la production industrielle de produits bio-pharmaceutiques à base de protéines. Cette production, principalement opérée en mode fed-batch, vise à maximiser la productivité de la biomasse. Cependant, l'accumulation d'acétate durant la culture inhibe la capacité respiratoire des cellules et diminue leur performance métabolique. Dans cette thèse, des stratégies de commande de l'alimentation sont envisagées pour éviter l'accumulation d'acétate et maximiser la productivité de la biomasse. A cette fin, des schémas de commande et d'estimation à base de modèle sont développés pour réguler le taux de croissance de la biomasse et la concentration en acétate. Les méthodes de commande vont du contrôle par modèle générique au contrôle prédictif par modèle non linéaire, et les variables non mesurées sont estimées à l'aide du filtre de Kalman « sans parfum ». Les développements ont porté sur la robustesse des méthodes proposées en raison de la nature incertaine du bioprocédé. La performance et la robustesse des schémas de commande et d'estimation sont testées et ajustées au travers de différents scénarios de simulation. Des cultures en mode Fed-batch de la souche E. coli BL21(DE3) sont réalisées avec succès sur un bioréacteur de laboratoire, mettant en évidence le potentiel des stratégies proposées dans un contexte de conditions opératoires en temps réel. Les stratégies de commande proposées dans cette thèse permettent un gain moyen jusqu'à 20% de la productivité de la biomasse par rapport au mode de fonctionnement conventionnel.

Title: Robust control of fed-batch cultures of Escherichia coli

Keywords: Escherichia coli, Robust Control, Biotechnology, Fed-batch

Abstract: Escherichia coli is a widespread cellular host for the industrial production of protein-based biopharmaceuticals. This production, mainly operated in fed-batch mode, aims to maximize biomass productivity. However, the accumulation of acetate during the culture inhibits the cells respiratory capacity and lowers their metabolic performance. In this thesis, closed-loop feeding control strategies are considered to avoid acetate accumulation and maximize biomass productivity. To this end, model-based control and estimation schemes are developed to regulate the biomass growth rate and the acetate concentration. The control methods ranged from the Generic Model Control and Nonlinear Model Predictive Control, and the non-measured variables are estimated using the Unscented Kalman Filter. The developments focused on the robustness of the proposed methods due to the uncertain nature of the bioprocess. The performance and robustness of the control and estimation strategies are tested and tuned by means of different scenarios of simulation runs. Fed-batch cultures of E. coli BL21(DE3) strain are successfully carried on a lab-scale bioreactor, highlighting the potential of the proposed strategies in real-time conditions. The proposed control strategies presented in this thesis lead to an average gain of up to 20% in biomass productivity compared to the conventional operating mode.

Université Paris-Saclay
Espace Technologique / Immeuble Discovery
Route de l'Orme aux Merisiers RD 128 / 91190 Saint-Aubin, France