OPTIMISATION OF FED-BATCH FERMENTATION PROCESS USING DEEP REINFORCEMENT LEARNING

CHAI WAN YING

FACULTY OF ENGINEERING
UNIVERSITI MALAYSIA SABAH
2023

OPTIMISATION OF FED-BATCH FERMENTATION PROCESS USING DEEP REINFORCEMENT LEARNING

CHAI WAN YING

THESIS SUBMITTED IN FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF ENGINEERING

FACULTY OF ENGINEERING
UNIVERSITI MALAYSIA SABAH
2023

UNIVERSITI MALAYSIA SABAH

BORANG PENGESAHAN STATUS TESIS

JUDUL : **OPTIMISATION OF FED-BATCH FERMENTATION PROCESS USING DEEP REINFORCEMENT LEARNING**

IJAZAH : **SARJANA KEJURUTERAAN**

BIDANG : **KEJURUTERAAN KIMIA**

Saya **CHAI WAN YING**, Sesi **2020-2023**, mengaku membenarkan tesis Sarjana ini disimpan di Perpustakaan Universiti Malaysia Sabah dengan syarat-syarat kegunaan seperti berikut:-

1. Tesis ini adalah hak milik Universiti Malaysia Sabah
2. Perpustakaan Universiti Malaysia Sabah dibenarkan membuat salinan untuk tujuan pengajian sahaja.
3. Perpustakaan dibenarkan membuat salinan tesis ini sebagai bahan pertukaran antara institusi pengajian tinggi.
4. Sila tandakan (/):

☐	SULIT	(Mengandungi maklumat yang berdarjah keselamatan atau kepentingan Malaysia seperti yang termaktub di dalam AKTA RAHSIA 1972)
☐	TERHAD	(Mengandungi maklumat TERHAD yang telah ditentukan oleh organisasi/badan di mana penyelidikan dijalankan)
☐	TIDAK TERHAD	

Disahkan Oleh,

_____ _____

CHAI WAN YING
MK2011037T

(Tandatangan Pustakawan)

Tarikh : 23 Ogos 2023

(Dr. Tham Heng Jin)
Penyelia Utama

DECLARATION

I hereby declare that this report is from my own work except for the information from articles, books and journals that have been properly cited and acknowledged in this paper.

18 May 2023

_____

CHAI WAN YING
MK2011037T

CERTIFICATION

NAME	:	**CHAI WAN YING**
MATRIC NO.	:	**MK2011037T**
TITLE	:	**OPTIMISATION OF FED-BATCH FERMENTATION PROCESS USING DEEP REINFORCEMENT LEARNING**
DEGREE	:	**MASTER OF ENGINEERING**
FIELD	:	**CHEMICAL ENGINEERING**
VIVA DATE	:	**18 MAY 2023**

CERTIFIED BY;

Signature

1. **MAIN SUPERVISOR**
   Dr. Tham Heng Jin

   _____

2. **CO-SUPERVISOR**
   Assoc. Prof. Dr. Kenneth Teo Tze Kin

   _____

3. **CO-SUPERVISOR**
   Dr. Tan Min Keng

   _____

ACKNOWLEDGEMENT

I would like to express my sincere gratitude to all those who have supported me throughout my journey in completing this thesis.

First and foremost, I would like to express my deepest appreciation to my supervisors, Dr. Tham Heng Jin, Dr. Tan Min Keng and Assoc. Prof. Dr. Kenneth Teo Tze Kin, who have provided invaluable guidance, support, and encouragement throughout my research. Your expertise and patience have been a constant source of inspiration and motivation to complete the research.

Besides, I would like to extend my gratitude to my family and friends who have supported me both emotionally and intellectually. Your unwavering support, motivation and encouragement have been a source of strength during the most challenging moments of my journey.

I would also like to thank the staff of Faculty of Engineering and Universiti Malaysia Sabah for their support in providing the necessary resources and facilities that were essential for the successful completion of my research. Furthermore, I would like to thank Ministry of Higher Education Malaysia for the Fundamental Research Grant Scheme (FRGS), grant no. FRGS/1/2020/TK0/UMS/02/10 for providing funds for this research.

In conclusion, I would like to express my sincere gratitude to everyone who has helped me in one way or another in completing this thesis. Your support and encouragement have been invaluable.

Chai Wan Ying

18 May 2023

ABSTRACT

Fed-batch fermentation process has always been a challenge for optimisation because it is highly non-linear and complex. Deep reinforcement learning is a self-learning algorithm through trial and error and experience, without any prior knowledge. This research aimed to determine the optimal feeding strategy for fed-batch baker's yeast fermentation process using the deep reinforcement learning algorithm in maximising the final production of yeast, while minimising the undesired ethanol formation. The kinetic and dynamic behaviour of the fed-batch baker's yeast fermentation was simulated and modelled using MATLAB, with no experimental work carried out. The proposed deep reinforcement learning algorithm, which integrates an artificial neural network with traditional reinforcement learning, was formulated based on the optimisation objective by manipulating only the substrate feeding rate. The performance of the proposed algorithm was compared with a pre-determined exponential feeding profile and a genetic algorithm. Results for the nominal condition show that the proposed algorithm produced final yeast concentration 33.42 g/l and 6.02 g/l higher than the exponential feeding and genetic algorithm, respectively. At the same time, the total ethanol formation is 0.19 g/l and 0.03 g/l lower than the exponential feeding and genetic algorithm, respectively. In other cases of different initial yeast and substrate concentrations, the proposed algorithm in general outperforms the exponential feeding profile while produces comparable results to the genetic algorithm. When dealing with model mismatch (±15% parameter variation in critical growth and maximum glucose uptake rate) and process disturbance (±20% deviation in substrate feeding concentration), the proposed algorithm was able to handle the changes with a minor effect on the yeast yield up to 13.78% and 2.52%, respectively, across all different initial condition cases. In conclusion, a deep reinforcement learning algorithm was successfully developed for the substrate feeding rate optimisation in the fed-batch baker's yeast fermentation process. The proposed algorithm improves the productivity of yeast while limiting ethanol formation and shows satisfactory performance in dealing with model mismatch and process disturbance.

ABSTRAK

OPTIMISATION OF FED-BATCH FERMENTATION PROCESS USING DEEP REINFORCEMENT LEARNING

Proses penapaian secara kumpulan makan sentiasa menjadi cabaran untuk pengoptimuman oleh sebab ketidaklineariti dan kerumitannya yang tinggi. Pembelajaran peneguhan mendalam ialah algoritma pembelajaran kendiri melalui percubaan dan kesilapan dan pengalaman tanpa sebarang pengetahuan terdahulu. Penyelidikan ini bertujuan untuk menentukan strategi pemakanan yang optimum bagi proses penapaian yis secara kumpulan makan dengan menggunakan algoritma pembelajaran peneguhan mendalam dalam memaksimumkan pengeluaran akhir yis sambil meminimumkan pembentukan etanol yang tidak diingini. Tingkah laku kinetik dan dinamik penapaian yis secara kumpulan makan telah disimulasikan dan dimodulkan menggunakan MATLAB, tanpa kerja eksperimen dijalankan. Algotitma pembelajaran peneguhan mendalam yang dicadangkan, yang mengintegrasikan rangkaian saraf tiruan dengan pembelajaran peneguhan tradisional, telah dirumuskan berdasarkan objektif pengoptimuman dengan hanya memanipulasi kadar pemakanan substrat. Prestasi algorithma yang dicadangkan telah dibandingkan dengan profil pemakanan secara eksponen dan algoritma genetik. Keputusan untuk keadaan nominal menunjukkan bahawa algoritma yang dicadangkan menghasilkan kepekatan yis akhir masing-masing dengan 33.42 g/l dan 6.02 g/l lebih tinggi daripada pemakanan secara eksponen dan algoritma genetik. Pada masa yang sama, jumlah pembentukan etanol masing-masing adalah 0.19 g/l dan 0.03 g/l lebih rendah daripada pemakanan secara eksponen dan algoritma genetik. Dalam kes lain kepekatan yis dan substrat awal yang berbeza, algoritma yang dicadangkan secara umum mengatasi profil pemakanan secara eksponen sambil menghasilkan keputusan yang setanding dengan algoritma genetik. Apabila menangani ketidakpadanan model (±15% variasi parameter dalam pertumbuhan kritikal dan kadar pengambilan glukosa maksimum) dan gangguan proses (±20% sisihan dalam kepekatan pemakanan substrat), algoritma yang dicadangkan dapat mengendalikan perubahan dengan kesan kecil pada penghasilan yis sehingga 13.78% dan 2.52%, masing-masing, merentas semua kes keadaan awal yang berbeza. Kesimpulannya, algoritma pembelajaran peneguhan mendalam telah berjaya direkakan untuk pengoptimuman kadar penyusuan substrat dalam process penapaian yis secara kumpulan. Algoritma pembelajaran peneguhan mendalam yang dicadangkan meingkatkan penghasilan yis sambil mengehadkan hasil etanol dan menunjukkan prestasi yang memuaskan dalam menangani ketidakpadanan model dan gangguan proses.

LIST OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS

C_e	-	Ethanol concentration	g/l
C_j	-	Output concentration of component j	g/l
C_{jo}	-	Input concentration of component j	g/l
C_o	-	Oxygen concentration	g/l
C_o^*	-	Oxygen saturation coefficient	g/l
C_s	-	Glucose (substrate) concentration	g/l
C_x	-	Yeast concentration	g/l
F_j	-	Substrate output flow rate	l/h
F_{jo}	-	Substrate input flow rate	l/h
K_e	-	Saturation constant for ethanol	g/l
K_i	-	Inhibition constant	g/l
K_o	-	Saturation constant for oxygen	g/l
K_s	-	Glucose (substrate) saturation constant	g/l
M_w	-	Weight-average molecular weight	g/mol
N_j	-	Mass of component j	g
Q_c	-	Carbon dioxide evolution rate	$g/g.h$
$Q_{e,max}$	-	Maximum specific ethanol uptake rate	$g/g.h$
$Q_{e,ox}$	-	Respiratory ethanol consumption	$g/g.h$
$Q_{e,prod}$	-	Ethanol production rate	$g/g.h$
$Q_{e,up}$	-	Ethanol uptake rate	$g/g.h$
Q_m	-	Cell maintenance rate	$g/g.h$
$Q_{o,lim}$	-	Oxidation capacity of yeast	$g/g.h$
$Q_{o,max}$	-	Maximum specific oxygen uptake rate	$g/g.h$
Q_o	-	Oxygen consumption rate	$g/g.h$
$Q_{s,lim}$	-	Limiting substrate flux	$g/g.h$
$Q_{s,max}$	-	Maximum specific glucose uptake rate	$g/g.h$
$Q_{s,ox}$	-	Oxidative glucose metabolism	$g/g.h$
$Q_{s,red}$	-	Reductive glucose metabolism	$g/g.h$
Q_s	-	Glucose (substrate) uptake rate	$g/g.h$
S_o	-	Concentration of glucose input feed	g/l

$Y_{c/e}$	-	Yield coefficient of carbon dioxide on ethanol	g/g
$Y_{c/s}^{ox}$	-	Oxidative yield coefficient of carbon dioxide on glucose	g/g
$Y_{c/s}^{red}$	-	Reductive yield coefficient of carbon dioxide on glucose	g/g
$Y_{e/o}$	-	Yield coefficient of ethanol on oxygen	g/g
$Y_{e/s}$	-	Yield coefficient of ethanol on glucose	g/g
$Y_{o/e}$	-	Consumption coefficient of oxygen on ethanol	g/g
$Y_{o/s}$	-	Consumption coefficient of oxygen on glucose	g/g
$Y_{x/e}$	-	Yield coefficient of yeast on ethanol	g/g
$Y_{x/s}^{ox}$	-	Oxidative yield coefficient of yeast on glucose	g/g
$Y_{x/s}^{red}$	-	Reductive yield coefficient of yeast on glucose	g/g
$k_L a_o$	-	Total volumetric mass transfer coefficient	$1/h$
r_j	-	Reaction rate change of component j	$g/l.h$
t_d	-	Time delay in substrate consumption	h
μ_{cr}	-	Critical specific growth rate	$1/h$
μ_{max}	-	Maximum specific growth rate	$1/h$
π^*	-	Optimal policy	
Δe	-	Error change	
$\int e$	-	Integral of error	
$\mathbf{V}$	-	Total reactor volume	l
A	-	Action space	
F	-	Glucose (substrate) feed flow rate	l/h
S	-	State space	
e	-	Error	
r	-	Reward function	
t	-	Time	h
μ	-	Total specific growth rate	$1/h$

LIST OF ABBREVIATION

A3C	-	Asynchronous Advantage Actor-Critic Algorithm
ADE	-	Adaptive Differential Evolution
AI	-	Artificial Intelligence
ANN	-	Artificial Neural Network
ATP	-	Adenine Triphosphate
BSA	-	Backtracking Search Optimisation Algorithm
CER	-	Carbon Dioxide Evolution Rate
CHO	-	Chinese Hamster Ovary
CTR	-	Carbon Dioxide Transfer Rate
DDP	-	Differential Dynamic Programming
DDPG	-	Deep Deterministic Policy Gradient
DE	-	Differential Evolution
DNN	-	Deep Neural Network
DO	-	Dissolved Oxygen
DPG	-	Deterministic Policy Gradient
DQN	-	Deep Q-learning Network
DRL	-	Deep Reinforcement Learning
EA	-	Evolutionary Algorithm
EBC	-	Elemental Balance Control
EF	-	Pre-determined Exponential Feeding
EMPC	-	Economic Model Predictive Control
FPA	-	Flower Pollination Algorithm
GA	-	Genetic Algorithm
HVAC	-	Heating, Ventilation and Air Conditioning
LM	-	Levenberg-Marquardt
MDP	-	Markov Decision Process
MFA	-	Model Free Adaptive
MIMO	-	Multi-inputs Multi-outputs
MPC	-	Model Predictive Control

MSA	-	Multi-step Action
NMPC	-	Non-linear Model Predictive Control
OADE	-	Adaptive Opposition-Based Differential Evolution
ODE	-	Opposition-based Differential Evolution
OTR	-	Oxygen Transfer Rate
OUR	-	Oxygen Uptake Rate
PI	-	Proportional Integral
PID	-	Proportional–Integral–Derivative
PSO	-	Particle Swarm Optimisation
PSRL	-	Partially Supervised Reinforcement Learning
ReLU	-	Rectified Linear Unit
RL	-	Reinforcement Learning
RQ	-	Respiratory Quotient
SIR	-	Substrate Intake Rate
SISO	-	Single-input Single-output

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

1.1 Background Study

Fermentation is a metabolic process carried out by microorganisms in the presence of carbon source, for their survival and growth while producing value-added products at the same time. Fermentation processes are widely applied in food processing industries to preserve food and improve the taste, flavour and texture. There are also applied in the manufacture of bioenergy, industrial chemicals and pharmaceutical products (Tullio, 2022). The biotechnology market worldwide was worth $369.62 billion in 2016 and is projected to reach $727.1 billion by 2025 (Ade *et al.*, 2018). In comparison with conventional chemical processes, bioprocesses able to produce valuable products with less energy consumption and less negative environmental impacts (Cheng *et al.*, 2018; Queiroz *et al.*, 2022). Bioprocesses typically operate under milder condition (e.g., lower temperature and atmospheric pressure) compared to conventional chemical processes, which reduces the energy requirement for heating, cooling, and maintaining high pressure. Besides, the by-products (or wastes) generated from bioprocesses are biodegradable and environmentally friendly. Conventional chemical processes, in contrast, produce persistent wastes that require extensive treatment. In addition, chemical synthesis typically involves multiple steps, whereas bioprocesses can obtain a complex compound formation in a single step. As a result, operation cost for chemical processes increases for equipment, energy and waste treatment.

Microorganisms are highly sensitive to the environmental conditions such as temperature, nutrient availability, pH and oxygen capacity. These factors influence the metabolic pathways and behaviour of microorganisms. A drawback of bioprocesses is that even a slight change in the broth environment can trigger

different metabolism mechanisms, resulting in the formation of multiple products. The presence of these multiple products can affect the quality and quantity of the desired product. Therefore, it is crucial to optimise and control bioprocess, in favour of the formation of the desired products. Hence, the goal of bioprocess optimisation in the industry is generally to maximise the productivity of targeted products while minimising any possible yield of undesired products in the most economical way.

Industrial fermentation processes are typically carried out in one of the following modes: batch, fed-batch or continuous mode (Yang & Sha, 2019). Table 1.1 below shows the advantage and disadvantage of these three different operation modes. Batch fermentation operates in a partially closed system where there is no feeding and discharge, and products are only removed at the end of the process. In batch operation, oxygen and pH solutions are allowed to regulate in the batch bioreactor, where oxygen concentration, pH and temperature are normally the observed parameters for production and quality improvement. Batch operation is relatively straightforward with minimal contamination risk. However, it has low product yields and requires long downtime for cleaning, sterilization, and preparation for new processing batches. In continuous fermentation process, the feeding and discharge occur at the same rate to maintain constant culture volume. Turbidostat and chemostat are the two primary control methods for continuous fermentation. Turbidostat regulates the inflow and outflow rate to control the turbidity in the bioreactor and to maintain constant cell density. Chemostat maintains the feeding and discharge rate at a constant level by adjusting a single limiting nutrient in the feeding medium to control the cell growth rate. In fed-batch operation, the substrate feeding input can be varied and regulated to control the reaction rate in the bioreactor with no output is discharged during the fermentation process. This becomes the advantage of fed-batch over the batch and continuous operation modes, as fed batch can overcome substrate inhibition on cell growth, improve productivity and control unwanted by-product formation (Yang & Sha, 2019). Therefore, the fed-batch mode of operation is still preferred by many biomanufacturing industries (Lindskog, 2018).

Table 1.1: The Advantage and Disadvantage of Different Operation Mode

Operation mode	Advantage	Disadvantage
Batch	Simple operation, low contamination risk	Low product yields, long downtime for new batch
Fed-batch	Higher product yield, avoid substrate inhibition effect, control by-product formation	Feeding regulation may be complicated, genetic instability
Continuous	High product yield, reduce downtime, avoid inhibition effect	Complex downstream processing, high contamination risk, genetic instability

However, the determination of optimal feeding profile for fed-batch bioprocess remains a challenging issue due to the high non-linearity and complexity of biological mechanism. As a result, the process optimisation of fed-batch bioprocess has drawn attention from many researchers. Many strategies have been studied such as pre-determined profile, model predictive control, fuzzy logic, artificial neural network and evolutionary algorithm. Reinforcement learning is a class of machine learning to solve problem and achieve goal through self-exploration and self-learning from experience without any guidance and training data. With the integration of artificial neural network (deep learning) as the function approximator, so called deep reinforcement learning, it has recently garnered interest in the control field. This is because deep learning can increase the computational efficiency of reinforcement learning in continuous and high dimensional problems. This type of machine learning has significant success in computer games, board games, traffic control, smart grids and robotics. Recently, it has been studied for its application in the chemical industries (Nian *et al.*, 2020). Fermentation process is a continuous process with infinite number of possible state variable values. Hence, in this work, reinforcement learning algorithm augmented with neural network is proposed to study its capability and potential for optimising the substrate feeding profile of fed-batch fermentation process.

1.2 Problem Statement

The high non-linearity and complexity of fermentation process (Zhang & Liu, 2019) is caused by the non-linear growth behaviour of microorganisms which is very sensitive to the culture environment. Plus, the metabolism mechanism of microorganisms is very complicated. A slight change in environment can cause different metabolic pathways. Substrate concentration is one of the important factors that influence the growth and metabolic process of microorganisms. Excessive substrate can lead to other undesired by-products formation which inhibit cell growth. Also, the multi-products formed will contaminate the broth and increase the post-processing cost. Therefore, the control and optimisation of the substrate feeding rate in fed-batch bioreactor is important to maximise the desired product yield while minimising by-product formation at the same time. These are the common requirements in various bioindustries.

In the search for the optimal substrate feeding profile, exponential feeding is one of the commonly used pre-determined profiles in the industry (Teworte *et al.*, 2022) because of its simplicity and ease of implementation. However, it lacks the ability to predict and overcome unexpected disturbances during production (Bolmanis *et al.*, 2023). Previous strategies have mostly focused on controlling and maintaining the process at a given set point, with limited studies on determining the optimal feeding profile. Given the abilities of deep reinforcement learning in self-learning through trial and error and experience without any prior knowledge, improvement computational efficiency and eliminate the need for human operators to monitor the process, it is of interest this research to investigate its capability in fed-batch fermentation process optimisation.