

Article

A Two-Stage Method for Parameter Identification of a Nonlinear System in a Microbial Batch Process

Gongxian Xu , Dongxue Lv and Wenxin Tan

Department of Mathematics, Bohai University, No. 19, Keji Road, Jinzhou 121013, China; dutxgx@aliyun.com (D.L.); isopeisope@aliyun.com (W.T.)

* Correspondence: gxxu@bhu.edu.cn; Tel.: +86-0416-340-0145

Received: 8 December 2018; Accepted: 16 January 2019; Published: 18 January 2019



Abstract: This paper deals with the parameter identification of a microbial batch process of glycerol to 1,3-propanediol (1,3-PD). We first present a parameter identification model for the excess kinetics of a microbial batch process of glycerol to 1,3-PD. This model is a nonlinear dynamic optimization problem that minimizes the sum of the least-square and slope errors of biomass, glycerol, 1,3-PD, acetic acid, and ethanol. Then, a two-stage method is proposed to efficiently solve the presented dynamic optimization problem. In this method, two nonlinear programming problems are required to be solved by a genetic algorithm. To calculate the slope of the experimental concentration data, an integral equation of the first kind is solved by using the Tikhonov regularization. The proposed two-stage method could not only optimally identify the model parameters of the biological process, but could also yield a smaller error between the measured and computed concentrations than the single-stage method could, with a decrease of about 52.79%. A comparative study showed that the proposed two-stage method could obtain better identification results than the single-stage method could.

Keywords: microbial batch process; parameter identification; optimization problem; nonlinear programming; numerical differentiation; genetic algorithm

1. Introduction

There are widespread applications for 1,3-propanediol (1,3-PD) [1]. In the microbial production of 1,3-PD, the bio-dissimilation of glycerol to 1,3-PD has attracted the interest of researchers since the 1980s because it possesses a relatively high yield and productivity [2]. In recent years, much research has been conducted to improve the dissimilation process of glycerol from mathematical modeling, biochemical analysis, process optimization, etc. [2–24]. For example, to describe the cell growth kinetics of multiple-inhibitions and product formation in a continuous bio-dissimilation process, a quantitative description has been given by Xiu et al. [3], Zeng and Deckwer [4], and Zeng and Biebl [5]. Lama et al. [6] addressed the metabolic engineering of *Klebsiella pneumoniae* J2B for the production of 1,3-PD from glucose. Lee et al. [7] and Sun et al. [8] reviewed the advances in biological and chemical methods for the conversion of glycerol into 1,3-PD. Vivek et al. [9] carried out a comparative evaluation of the metabolite fluxes in 1,3-PD production of cell recycling, simple batch, and continuous fermentation processes by using the *Lactobacillus brevis* N1E9.3.3 strain. Rodriguez et al. [10] addressed the kinetic modeling of raw glycerol to 1,3-PD production by *Shimwellia blattae*. This kinetic model can describe the influence of the initial glycerol concentration on 1,3-PD production. Liu et al. [11] presented the bi-objective dynamic optimization of a nonlinear time-delay system to optimize 1,3-PD production in a microbial batch process. Yuan et al. [12] gave an optimal control strategy of a nonlinear batch system with a time delay to maximize 1,3-PD production. Hirokawa et al. [13] used the engineered cyanobacterium *Synechococcus elongatus* to improve the production of 1,3-PD by optimizing the gene

expression level of a metabolic pathway and operation conditions. Narisetty et al. [14] improved 1,3-PD production by maintaining physical conditions and optimizing media composition. To address the multi-objective optimization of the continuous culture process of glycerol to 1,3-PD, Xu et al. [15] proposed four optimization models and solved them with the normal-boundary intersection (NBI) and weighted-sum (WS) methods. Xu and Li [16] proposed a multilevel programming method to infer the common metabolic objective function for glycerol bio-dissimilation to 1,3-PD by *Klebsiella pneumoniae*. This approach is more feasible and has a better prediction performance than the existing method [16]. Xu and Wang [17] presented three parameter identification models of a microbial batch process of glycerol to 1,3-PD to identify the parameter values of the nonlinear biological system by considering three different error criteria of biomass, glycerol, 1,3-PD, acetic acid, and ethanol. These parameter identification models are dynamic optimization problems. They can be solved by transforming the original dynamic optimization problems into corresponding nonlinear programming problems. However, the transformed nonlinear programming problems are difficult to solve for global optimality. To deal with this difficulty, in this study, a two-stage method is proposed to efficiently handle the parameter identification of the microbial batch process of glycerol to 1,3-PD.

In the following, we first present a parameter identification model for the excess kinetics of the microbial batch process of glycerol to 1,3-PD by *Klebsiella pneumoniae*. Then, a two-stage method is proposed to efficiently solve the presented dynamic optimization problem. Section 4 presents the parameter identification results obtained by the proposed two-stage method. A comparative study is also given to show that the proposed two-stage method can obtain better identification results than the single-stage method. Finally, some brief conclusions of the present work are given in Section 5.

2. Parameter Identification Model for the Microbial Batch Process

2.1. Microbial Batch Process

Based on previous work [2], the material balance relationships of the microbial batch process of glycerol to 1,3-PD by *Klebsiella pneumoniae* can be expressed as follows:

$$\frac{dX}{dt} = \mu X, \quad (1)$$

$$\frac{dC_S}{dt} = -q_S X, \quad (2)$$

$$\frac{dC_{PD}}{dt} = q_{PD} X, \quad (3)$$

$$\frac{dC_{HAc}}{dt} = q_{HAc} X, \quad (4)$$

$$\frac{dC_{EtOH}}{dt} = q_{EtOH} X, \quad (5)$$

$$t \in [t_0, t_N], \quad (6)$$

$$X(t_0) = X_0, C_S(t_0) = C_{S0}, C_{PD}(t_0) = C_{PD0}, C_{HAc}(t_0) = C_{HAc0}, C_{EtOH}(t_0) = C_{EtOH0}, \quad (7)$$

$$\mu = 0.67 \frac{C_S}{0.28 + C_S} \left(1 - \frac{C_S}{2039}\right) \left(1 - \frac{C_{PD}}{939.5}\right) \left(1 - \frac{C_{HAc}}{1026}\right) \left(1 - \frac{C_{EtOH}}{360.9}\right), \quad (8)$$

$$q_S = m_S + \frac{\mu}{Y_S^m} + \Delta q_S^m \frac{C_S}{C_S + K_S^*}, \quad (9)$$

$$q_{PD} = m_{PD} + \mu Y_{PD}^m + \Delta q_{PD}^m \frac{C_S}{C_S + K_{PD}^*}, \quad (10)$$

$$q_{HAc} = m_{HAc} + \mu Y_{HAc}^m + \Delta q_{HAc}^m \frac{C_S}{C_S + K_{HAc}^*}, \quad (11)$$

$$q_{EtOH} = m_{EtOH} + \mu Y_{EtOH}^m + \Delta q_{EtOH}^m \frac{C_S}{C_S + K_{EtOH}^*}, \tag{12}$$

where t denotes the culture time (h); $t_0 \geq 0$ and $t_N > 0$ are the initial and terminal times of the culture process, respectively (h); X denotes the biomass (g/L); C_S denotes the substrate (glycerol) concentration (mmol/L); C_{PD} denotes the concentration of product 1,3-PD (mmol/L); C_{HAc} denotes the concentration of product acetic acid (mmol/L); C_{EtOH} denotes the concentration of product ethanol (mmol/L); $X_0, C_{S0}, C_{PD0}, C_{HAc0}$, and C_{EtOH0} are the initial values of X, C_S, C_{PD}, C_{HAc} , and C_{EtOH} , respectively; μ denotes the specific growth rate of cells (1/h); q_S denotes the specific consumption rate of substrate glycerol (mmol/(g·h)); $q_{PD}, q_{HAc}, q_{EtOH}$ denote the specific formation rates of products 1,3-PD, acetic acid, and ethanol, respectively (mmol/(g·h)); m_S is the maintenance term of glycerol consumption under substrate-limited conditions (mmol/(g·h)); m_{PD}, m_{HAc} , and m_{EtOH} denote the maintenance terms of corresponding product formation under substrate-limited conditions, respectively (mmol/(g·h)); Y_S^m denotes the maximum growth yield (g/mmol); Y_{PD}^m, Y_{HAc}^m and Y_{EtOH}^m are the product yield of the corresponding products (mmol/g); Δq_S^m is the maximum increment of the glycerol consumption rate under substrate-sufficient conditions (mmol/(g·h)); $\Delta q_{PD}^m, \Delta q_{HAc}^m$ and Δq_{EtOH}^m are the maximum increments of the corresponding product formation rate under substrate-sufficient conditions (mmol/(g·h)); and $K_S^*, K_{PD}^*, K_{HAc}^*$ and K_{EtOH}^* denote the saturation constants of glycerol and corresponding products in the excess terms, respectively (mmol/L).

By introducing Equations (8)–(12) into Equations (1)–(5), Equations (1)–(7) have the following formulations:

$$\begin{aligned} \frac{dX}{dt} &= 0.67 \frac{XC_S}{0.28+C_S} \left(1 - \frac{C_S}{2039}\right) \left(1 - \frac{C_{PD}}{939.5}\right) \left(1 - \frac{C_{HAc}}{1026}\right) \left(1 - \frac{C_{EtOH}}{360.9}\right), \\ \frac{dC_S}{dt} &= -m_S X - \frac{0.67}{Y_S^m} \frac{C_S X}{0.28+C_S} \left(1 - \frac{C_S}{2039}\right) \left(1 - \frac{C_{PD}}{939.5}\right) \left(1 - \frac{C_{HAc}}{1026}\right) \left(1 - \frac{C_{EtOH}}{360.9}\right) - \Delta q_S^m \frac{C_S X}{C_S + K_S^*}, \\ \frac{dC_{PD}}{dt} &= m_{PD} X + \frac{0.67 Y_{PD}^m C_S X}{0.28+C_S} \left(1 - \frac{C_S}{2039}\right) \left(1 - \frac{C_{PD}}{939.5}\right) \left(1 - \frac{C_{HAc}}{1026}\right) \left(1 - \frac{C_{EtOH}}{360.9}\right) + \Delta q_{PD}^m \frac{C_S X}{C_S + K_{PD}^*}, \\ \frac{dC_{HAc}}{dt} &= m_{HAc} X + \frac{0.67 Y_{HAc}^m C_S X}{0.28+C_S} \left(1 - \frac{C_S}{2039}\right) \left(1 - \frac{C_{PD}}{939.5}\right) \left(1 - \frac{C_{HAc}}{1026}\right) \left(1 - \frac{C_{EtOH}}{360.9}\right) + \Delta q_{HAc}^m \frac{C_S X}{C_S + K_{HAc}^*}, \\ \frac{dC_{EtOH}}{dt} &= m_{EtOH} X + \frac{0.67 Y_{EtOH}^m C_S X}{0.28+C_S} \left(1 - \frac{C_S}{2039}\right) \left(1 - \frac{C_{PD}}{939.5}\right) \left(1 - \frac{C_{HAc}}{1026}\right) \left(1 - \frac{C_{EtOH}}{360.9}\right) + \Delta q_{EtOH}^m \frac{C_S X}{C_S + K_{EtOH}^*}, \\ &t \in [t_0, t_N], \end{aligned}$$

$$X(t_0) = X_0, C_S(t_0) = C_{S0}, C_{PD}(t_0) = C_{PD0}, C_{HAc}(t_0) = C_{HAc0}, C_{EtOH}(t_0) = C_{EtOH0}.$$

2.2. Parameter Identification Model

To identify the values of the parameters $m_S, m_{PD}, m_{HAc}, m_{EtOH}, Y_S^m, Y_{PD}^m, Y_{HAc}^m, Y_{EtOH}^m, \Delta q_S^m, \Delta q_{PD}^m, \Delta q_{HAc}^m, \Delta q_{EtOH}^m, K_S^*, K_{PD}^*, K_{HAc}^*$ and K_{EtOH}^* in Equations (1)–(12), we propose the following parameter identification model for the microbial batch process of glycerol to 1,3-PD:

$$\begin{aligned} \min F &= \frac{1}{N} \sum_{j=1}^N \left[(X(t_j) - X_e(t_j))^2 / X_{\max}^2 \right] + \frac{1}{N} \sum_{j=1}^N \left[(\dot{X}(t_j) - \dot{X}_e(t_j))^2 / \dot{X}_{\max}^2 \right] \\ &+ \frac{1}{N} \sum_{j=1}^N \left[(C_S(t_j) - C_{Se}(t_j))^2 / C_{S\max}^2 \right] + \frac{1}{N} \sum_{j=1}^N \left[(\dot{C}_S(t_j) - \dot{C}_{Se}(t_j))^2 / \dot{C}_{S\max}^2 \right] \\ &+ \frac{1}{N} \sum_{j=1}^N \left[(C_{PD}(t_j) - C_{PD_e}(t_j))^2 / C_{PD\max}^2 \right] + \frac{1}{N} \sum_{j=1}^N \left[(\dot{C}_{PD}(t_j) - \dot{C}_{PD_e}(t_j))^2 / \dot{C}_{PD\max}^2 \right] \\ &+ \frac{1}{N} \sum_{j=1}^N \left[(C_{HAc}(t_j) - C_{HAc_e}(t_j))^2 / C_{HAc\max}^2 \right] + \frac{1}{N} \sum_{j=1}^N \left[(\dot{C}_{HAc}(t_j) - \dot{C}_{HAc_e}(t_j))^2 / \dot{C}_{HAc\max}^2 \right], \\ &+ \frac{1}{N} \sum_{j=1}^N \left[(C_{EtOH}(t_j) - C_{EtOH_e}(t_j))^2 / C_{EtOH\max}^2 \right] + \frac{1}{N} \sum_{j=1}^N \left[(\dot{C}_{EtOH}(t_j) - \dot{C}_{EtOH_e}(t_j))^2 / \dot{C}_{EtOH\max}^2 \right] \end{aligned} \tag{13}$$

subject to satisfying

$$\frac{dX}{dt} = 0.67 \frac{XC_S}{0.28 + C_S} \left(1 - \frac{C_S}{2039}\right) \left(1 - \frac{C_{PD}}{939.5}\right) \left(1 - \frac{C_{HAc}}{1026}\right) \left(1 - \frac{C_{EtOH}}{360.9}\right), \quad (14)$$

$$\frac{dC_S}{dt} = -m_S X - \frac{0.67}{Y_S^m} \frac{C_S X}{0.28 + C_S} \left(1 - \frac{C_S}{2039}\right) \left(1 - \frac{C_{PD}}{939.5}\right) \left(1 - \frac{C_{HAc}}{1026}\right) \left(1 - \frac{C_{EtOH}}{360.9}\right) - \Delta q_S^m \frac{C_S X}{C_S + K_S^*}, \quad (15)$$

$$\frac{dC_{PD}}{dt} = m_{PD} X + \frac{0.67 Y_{PD}^m}{0.28 + C_S} \frac{C_S X}{0.28 + C_S} \left(1 - \frac{C_S}{2039}\right) \left(1 - \frac{C_{PD}}{939.5}\right) \left(1 - \frac{C_{HAc}}{1026}\right) \left(1 - \frac{C_{EtOH}}{360.9}\right) + \Delta q_{PD}^m \frac{C_S X}{C_S + K_{PD}^*}, \quad (16)$$

$$\frac{dC_{HAc}}{dt} = m_{HAc} X + \frac{0.67 Y_{HAc}^m}{0.28 + C_S} \frac{C_S X}{0.28 + C_S} \left(1 - \frac{C_S}{2039}\right) \left(1 - \frac{C_{PD}}{939.5}\right) \left(1 - \frac{C_{HAc}}{1026}\right) \left(1 - \frac{C_{EtOH}}{360.9}\right) + \Delta q_{HAc}^m \frac{C_S X}{C_S + K_{HAc}^*}, \quad (17)$$

$$\frac{dC_{EtOH}}{dt} = m_{EtOH} X + \frac{0.67 Y_{EtOH}^m}{0.28 + C_S} \frac{C_S X}{0.28 + C_S} \left(1 - \frac{C_S}{2039}\right) \left(1 - \frac{C_{PD}}{939.5}\right) \left(1 - \frac{C_{HAc}}{1026}\right) \left(1 - \frac{C_{EtOH}}{360.9}\right) + \Delta q_{EtOH}^m \frac{C_S X}{C_S + K_{EtOH}^*}, \quad (18)$$

$$t \in [t_0, t_N], \quad (19)$$

$$X(t_0) = X_0, C_S(t_0) = C_{S0}, C_{PD}(t_0) = C_{PD0}, C_{HAc}(t_0) = C_{HAc0}, C_{EtOH}(t_0) = C_{EtOH0}, \quad (20)$$

$$\mu = 0.67 \frac{C_S}{0.28 + C_S} \left(1 - \frac{C_S}{2039}\right) \left(1 - \frac{C_{PD}}{939.5}\right) \left(1 - \frac{C_{HAc}}{1026}\right) \left(1 - \frac{C_{EtOH}}{360.9}\right) \geq 0, \quad (21)$$

$$q_S = m_S + \frac{0.67}{Y_S^m} \frac{C_S}{0.28 + C_S} \left(1 - \frac{C_S}{2039}\right) \left(1 - \frac{C_{PD}}{939.5}\right) \left(1 - \frac{C_{HAc}}{1026}\right) \left(1 - \frac{C_{EtOH}}{360.9}\right) + \Delta q_S^m \frac{C_S}{C_S + K_S^*} \geq 0, \quad (22)$$

$$q_{PD} = m_{PD} + \frac{0.67 Y_{PD}^m}{0.28 + C_S} \frac{C_S}{0.28 + C_S} \left(1 - \frac{C_S}{2039}\right) \left(1 - \frac{C_{PD}}{939.5}\right) \left(1 - \frac{C_{HAc}}{1026}\right) \left(1 - \frac{C_{EtOH}}{360.9}\right) + \Delta q_{PD}^m \frac{C_S}{C_S + K_{PD}^*} \geq 0, \quad (23)$$

$$q_{HAc} = m_{HAc} + \frac{0.67 Y_{HAc}^m}{0.28 + C_S} \frac{C_S}{0.28 + C_S} \left(1 - \frac{C_S}{2039}\right) \left(1 - \frac{C_{PD}}{939.5}\right) \left(1 - \frac{C_{HAc}}{1026}\right) \left(1 - \frac{C_{EtOH}}{360.9}\right) + \Delta q_{HAc}^m \frac{C_S}{C_S + K_{HAc}^*} \geq 0, \quad (24)$$

$$q_{EtOH} = m_{EtOH} X + \frac{0.67 Y_{EtOH}^m}{0.28 + C_S} \frac{C_S X}{0.28 + C_S} \left(1 - \frac{C_S}{2039}\right) \left(1 - \frac{C_{PD}}{939.5}\right) \left(1 - \frac{C_{HAc}}{1026}\right) \left(1 - \frac{C_{EtOH}}{360.9}\right) + \Delta q_{EtOH}^m \frac{C_S X}{C_S + K_{EtOH}^*} \geq 0, \quad (25)$$

$$\hat{m}_S \leq m_S \leq \bar{m}_S, \quad (26)$$

$$\hat{m}_{PD} \leq m_{PD} \leq \bar{m}_{PD}, \quad (27)$$

$$\hat{m}_{HAc} \leq m_{HAc} \leq \bar{m}_{HAc}, \quad (28)$$

$$\hat{m}_{EtOH} \leq m_{EtOH} \leq \bar{m}_{EtOH}, \quad (29)$$

$$\hat{Y}_S^m \leq Y_S^m \leq \bar{Y}_S^m, \quad (30)$$

$$\hat{Y}_{PD}^m \leq Y_{PD}^m \leq \bar{Y}_{PD}^m, \quad (31)$$

$$\hat{Y}_{HAc}^m \leq Y_{HAc}^m \leq \bar{Y}_{HAc}^m, \quad (32)$$

$$\hat{Y}_{EtOH}^m \leq Y_{EtOH}^m \leq \bar{Y}_{EtOH}^m, \quad (33)$$

$$\Delta \hat{q}_S^m \leq \Delta q_S^m \leq \Delta \bar{q}_S^m, \quad (34)$$

$$\Delta \hat{q}_{PD}^m \leq \Delta q_{PD}^m \leq \Delta \bar{q}_{PD}^m, \quad (35)$$

$$\Delta \hat{q}_{HAc}^m \leq \Delta q_{HAc}^m \leq \Delta \bar{q}_{HAc}^m, \quad (36)$$

$$\Delta \hat{q}_{EtOH}^m \leq \Delta q_{EtOH}^m \leq \Delta \bar{q}_{EtOH}^m, \quad (37)$$

$$\hat{K}_S^* \leq K_S^* \leq \bar{K}_S^*, \quad (38)$$

$$\hat{K}_{PD}^* \leq K_{PD}^* \leq \bar{K}_{PD}^*, \quad (39)$$

$$\hat{K}_{HAc}^* \leq K_{HAc}^* \leq \bar{K}_{HAc}^*, \quad (40)$$

$$\hat{K}_{EtOH}^* \leq K_{EtOH}^* \leq \bar{K}_{EtOH}^*. \quad (41)$$

In Equations (13)–(41), the objective function F is the sum of the least-square and slope errors of biomass (X), glycerol (C_S), 1,3-PD (C_{PD}), acetic acid (C_{HAc}), and ethanol (C_{EtOH}); $X_e(t_j)$, $C_{Se}(t_j)$,

$C_{PD_e}(t_j)$, $C_{HAc_e}(t_j)$, and $C_{EtOH_e}(t_j)$ are the measured data for biomass, glycerol, 1,3-PD, acetic acid, and ethanol at the sampling time t_j ($j = 1, 2, \dots, N$), respectively; $X(t_j)$, $C_S(t_j)$, $C_{PD}(t_j)$, $C_{HAc}(t_j)$, and $C_{EtOH}(t_j)$ denote the computed concentrations for biomass, glycerol, 1,3-PD, acetic acid, and ethanol at t_j , respectively; X_{max} , C_{Smax} , C_{PDmax} , C_{HAcmax} , and $C_{EtOHmax}$ are the maximum measured concentrations of biomass, glycerol, 1,3-PD, acetic acid, and ethanol in $[t_0, t_N]$, respectively; N is the number of sampled data points; $\dot{X}_e(t_j)$, $\dot{C}_{S_e}(t_j)$, $\dot{C}_{PD_e}(t_j)$, $\dot{C}_{HAc_e}(t_j)$, and $\dot{C}_{EtOH_e}(t_j)$ are the approximate experimental slopes for biomass, glycerol, 1,3-PD, acetic acid, and ethanol at t_j , respectively; $\dot{X}(t_j)$, $\dot{C}_S(t_j)$, $\dot{C}_{PD}(t_j)$, $\dot{C}_{HAc}(t_j)$, and $\dot{C}_{EtOH}(t_j)$ are the computed slopes for biomass, glycerol, 1,3-PD, acetic acid, and ethanol at t_j , respectively; \dot{X}_{max} , \dot{C}_{Smax} , \dot{C}_{PDmax} , \dot{C}_{HAcmax} , and $\dot{C}_{EtOHmax}$ are the maximum experimental slopes of biomass, glycerol, 1,3-PD, acetic acid, and ethanol in $[t_0, t_N]$, respectively; \hat{m}_S , \hat{m}_{PD} , \hat{m}_{HAc} , \hat{m}_{EtOH} , \hat{Y}_S^m , \hat{Y}_{PD}^m , \hat{Y}_{HAc}^m , \hat{Y}_{EtOH}^m , $\Delta\hat{q}_S^m$, $\Delta\hat{q}_{PD}^m$, $\Delta\hat{q}_{HAc}^m$, $\Delta\hat{q}_{EtOH}^m$, \hat{K}_S^* , \hat{K}_{PD}^* , \hat{K}_{HAc}^* , and \hat{K}_{EtOH}^* are the lower bounds of parameters m_S , m_{PD} , m_{HAc} , m_{EtOH} , Y_S^m , Y_{PD}^m , Y_{HAc}^m , Y_{EtOH}^m , Δq_S^m , Δq_{PD}^m , Δq_{HAc}^m , Δq_{EtOH}^m , K_S^* , K_{PD}^* , K_{HAc}^* , and K_{EtOH}^* , respectively; and \bar{m}_S , \bar{m}_{PD} , \bar{m}_{HAc} , \bar{m}_{EtOH} , \bar{Y}_S^m , \bar{Y}_{PD}^m , \bar{Y}_{HAc}^m , \bar{Y}_{EtOH}^m , $\Delta\bar{q}_S^m$, $\Delta\bar{q}_{PD}^m$, $\Delta\bar{q}_{HAc}^m$, $\Delta\bar{q}_{EtOH}^m$, \bar{K}_S^* , \bar{K}_{PD}^* , \bar{K}_{HAc}^* , and \bar{K}_{EtOH}^* are the upper bounds of parameters m_S , m_{PD} , m_{HAc} , m_{EtOH} , Y_S^m , Y_{PD}^m , Y_{HAc}^m , Y_{EtOH}^m , Δq_S^m , Δq_{PD}^m , Δq_{HAc}^m , Δq_{EtOH}^m , K_S^* , K_{PD}^* , K_{HAc}^* , and K_{EtOH}^* , respectively.

For the parameter identification model in Equations (13)–(41), we have the following remarks:

1. The inequality constraints in Equations (21)–(25) keep the specific growth rate μ , specific consumption rate q_S , and specific formation rates q_{PD} , q_{HAc} , and q_{EtOH} within certain physically and chemically feasible limits;
2. Obviously, the parameter identification model in Equations (13)–(41) is a nonlinear dynamic optimization problem with complex constraints. Therefore, it is difficult to solve for global optimality.

3. Two-Stage Method for the Parameter Identification Model

In this work, we propose a two-stage method to efficiently solve the presented parameter identification model in Equations (13)–(41). This method addresses the parameter identification problem in Equations (13)–(41) through a two-stage procedure instead of directly solving it.

3.1. Two-Stage Method

We first rewrite the parameter identification problem in Equations (13)–(41) as

$$\min F = \frac{1}{N} \sum_{j=1}^N \sum_{i=1}^5 \left[(x_i(t_j) - x_{ie}(t_j))^2 / x_{i\max}^2 \right] + \frac{1}{N} \sum_{j=1}^N \sum_{i=1}^5 \left[(\dot{x}_i(t_j) - \dot{x}_{ie}(t_j))^2 / \dot{x}_{i\max}^2 \right], \quad (42)$$

subject to satisfying

$$\frac{dx_1}{dt} = \mu(\mathbf{x}, \mathbf{p})x_1, \quad (43)$$

$$\frac{dx_2}{dt} = -q_S(\mathbf{x}, \mathbf{p})x_1, \quad (44)$$

$$\frac{dx_3}{dt} = q_{PD}(\mathbf{x}, \mathbf{p})x_1, \quad (45)$$

$$\frac{dx_4}{dt} = q_{HAc}(\mathbf{x}, \mathbf{p})x_1, \quad (46)$$

$$\frac{dx_5}{dt} = q_{EtOH}(\mathbf{x}, \mathbf{p})x_1, \quad (47)$$

$$t \in [t_0, t_N], \quad (48)$$

$$\mathbf{x}(t_0) = \mathbf{x}_0, \quad (49)$$

$$\mu(\mathbf{x}, \mathbf{p}) = \frac{0.67x_2}{0.28 + x_2} \left(1 - \frac{x_2}{2039}\right) \left(1 - \frac{x_3}{939.5}\right) \left(1 - \frac{x_4}{1026}\right) \left(1 - \frac{x_5}{360.9}\right) \geq 0, \quad (50)$$

$$q_S(\mathbf{x}, \mathbf{p}) = p_1 + \frac{0.67x_2}{p_5(0.28 + x_2)} \left(1 - \frac{x_2}{2039}\right) \left(1 - \frac{x_3}{939.5}\right) \left(1 - \frac{x_4}{1026}\right) \left(1 - \frac{x_5}{360.9}\right) + \frac{p_9x_2}{x_2 + p_{13}} \geq 0, \quad (51)$$

$$q_{PD}(\mathbf{x}, \mathbf{p}) = p_2 + \frac{0.67p_6x_2}{0.28 + x_2} \left(1 - \frac{x_2}{2039}\right) \left(1 - \frac{x_3}{939.5}\right) \left(1 - \frac{x_4}{1026}\right) \left(1 - \frac{x_5}{360.9}\right) + \frac{p_{10}x_2}{x_2 + p_{14}} \geq 0, \quad (52)$$

$$q_{HAc}(\mathbf{x}, \mathbf{p}) = p_3 + \frac{0.67p_7x_2}{0.28 + x_2} \left(1 - \frac{x_2}{2039}\right) \left(1 - \frac{x_3}{939.5}\right) \left(1 - \frac{x_4}{1026}\right) \left(1 - \frac{x_5}{360.9}\right) + \frac{p_{11}x_2}{x_2 + p_{15}} \geq 0, \quad (53)$$

$$q_{EtOH}(\mathbf{x}, \mathbf{p}) = p_4 + \frac{0.67p_8x_2}{0.28 + x_2} \left(1 - \frac{x_2}{2039}\right) \left(1 - \frac{x_3}{939.5}\right) \left(1 - \frac{x_4}{1026}\right) \left(1 - \frac{x_5}{360.9}\right) + \frac{p_{12}x_2}{x_2 + p_{16}} \geq 0, \quad (54)$$

$$\mathbf{p}^l \leq \mathbf{p} \leq \mathbf{p}^u, \quad (55)$$

where $\mathbf{x} \in \mathbb{R}^5, \mathbf{x}_0 \in \mathbb{R}^5, \mathbf{x}_e \in \mathbb{R}^5, \mathbf{x}_{\max} \in \mathbb{R}^5, \mathbf{p} \in \mathbb{R}^{16}, \mathbf{p}^l \in \mathbb{R}^{16}$, and $\mathbf{p}^u \in \mathbb{R}^{16}$ have the following formulations:

$$\mathbf{x} = (x_1, x_2, x_3, x_4, x_5)^T = (X, C_S, C_{PD}, C_{HAc}, C_{EtOH})^T,$$

$$\mathbf{x}_0 = (x_{10}, x_{20}, x_{30}, x_{40}, x_{50})^T = (X_0, C_{S0}, C_{PD0}, C_{HAc0}, C_{EtOH0})^T,$$

$$\mathbf{x}_e = (x_{1e}, x_{2e}, x_{3e}, x_{4e}, x_{5e})^T = (X_e, C_{Se}, C_{PDe}, C_{HAce}, C_{EtOHe})^T,$$

$$\mathbf{x}_{\max} = (x_{1\max}, x_{2\max}, x_{3\max}, x_{4\max}, x_{5\max})^T = (X_{\max}, C_{S\max}, C_{PD\max}, C_{HAc\max}, C_{EtOH\max})^T,$$

$$\mathbf{p} = (p_1, p_2, p_3, p_4, p_5, p_6, p_7, p_8, p_9, p_{10}, p_{11}, p_{12}, p_{13}, p_{14}, p_{15}, p_{16})^T \\ = (\bar{m}_S, \bar{Y}_S^m, \Delta \bar{q}_S^m, \bar{K}_S^*, \bar{m}_{PD}, \bar{Y}_{PD}^m, \Delta \bar{q}_{PD}^m, \bar{K}_{PD}^*, \bar{m}_{HAc}, \bar{Y}_{HAc}^m, \Delta \bar{q}_{HAc}^m, \bar{K}_{HAc}^*, \bar{m}_{EtOH}, \bar{Y}_{EtOH}^m, \Delta \bar{q}_{EtOH}^m, \bar{K}_{EtOH}^*)^T,$$

$$\mathbf{p}^l = (p_1^l, p_2^l, p_3^l, p_4^l, p_5^l, p_6^l, p_7^l, p_8^l, p_9^l, p_{10}^l, p_{11}^l, p_{12}^l, p_{13}^l, p_{14}^l, p_{15}^l, p_{16}^l)^T \\ = (\hat{m}_S, \hat{Y}_S^m, \Delta \hat{q}_S^m, \hat{K}_S^*, \hat{m}_{PD}, \hat{Y}_{PD}^m, \Delta \hat{q}_{PD}^m, \hat{K}_{PD}^*, \hat{m}_{HAc}, \hat{Y}_{HAc}^m, \Delta \hat{q}_{HAc}^m, \hat{K}_{HAc}^*, \hat{m}_{EtOH}, \hat{Y}_{EtOH}^m, \Delta \hat{q}_{EtOH}^m, \hat{K}_{EtOH}^*)^T,$$

$$\mathbf{p}^u = (p_1^u, p_2^u, p_3^u, p_4^u, p_5^u, p_6^u, p_7^u, p_8^u, p_9^u, p_{10}^u, p_{11}^u, p_{12}^u, p_{13}^u, p_{14}^u, p_{15}^u, p_{16}^u)^T \\ = (\bar{m}_S, \bar{Y}_S^m, \Delta \bar{q}_S^m, \bar{K}_S^*, \bar{m}_{PD}, \bar{Y}_{PD}^m, \Delta \bar{q}_{PD}^m, \bar{K}_{PD}^*, \bar{m}_{HAc}, \bar{Y}_{HAc}^m, \Delta \bar{q}_{HAc}^m, \bar{K}_{HAc}^*, \bar{m}_{EtOH}, \bar{Y}_{EtOH}^m, \Delta \bar{q}_{EtOH}^m, \bar{K}_{EtOH}^*)^T.$$

The general method for solving a dynamic optimization problem depends on the discretization of dynamic system equations within the framework of the direct transcription [25–27]. By using a certain discretization technique, one can convert an infinite-dimensional dynamic optimization problem into a finite-dimensional, large-scale, nonlinear programming problem. To discretize the dynamic system equations, a collocation method should be used. For example, we can use the first-order Runge–Kutta (RK) method and higher-order RK methods, such as the implicit trapezoidal method, to discretize the dynamic system equations.

The implicit trapezoidal method is applied here to discretize the dynamic system in Equations (43)–(47). The discretized equations are written as

$$x_1(t_j) = x_1(t_{j-1}) + 0.5\eta_j[\mu(\mathbf{x}(t_j), \mathbf{p})x_1(t_j) + \mu(\mathbf{x}(t_{j-1}), \mathbf{p})x_1(t_{j-1})], \quad (56)$$

$$x_2(t_j) = x_2(t_{j-1}) + 0.5\eta_j[-q_S(\mathbf{x}(t_j), \mathbf{p})x_1(t_j) - q_S(\mathbf{x}(t_{j-1}), \mathbf{p})x_1(t_{j-1})], \quad (57)$$

$$x_3(t_j) = x_3(t_{j-1}) + 0.5\eta_j[q_{PD}(\mathbf{x}(t_j), \mathbf{p})x_1(t_j) + q_{PD}(\mathbf{x}(t_{j-1}), \mathbf{p})x_1(t_{j-1})], \quad (58)$$

$$x_4(t_j) = x_4(t_{j-1}) + 0.5\eta_j[q_{HAc}(\mathbf{x}(t_j), \mathbf{p})x_1(t_j) + q_{HAc}(\mathbf{x}(t_{j-1}), \mathbf{p})x_1(t_{j-1})], \quad (59)$$

$$x_5(t_j) = x_5(t_{j-1}) + 0.5\eta_j[q_{EtOH}(\mathbf{x}(t_j), \mathbf{p})x_1(t_j) + q_{EtOH}(\mathbf{x}(t_{j-1}), \mathbf{p})x_1(t_{j-1})], \quad (60)$$

where $\eta_j = t_j - t_{j-1}, j = 1, 2, \dots, N$.

The implicit trapezoidal equations (56)–(60) help reduce the dimension of the dynamic optimization problem in Equations (42)–(55), but they still involve a large-scale nonlinear programming

problem. To reduce the computation time, we use the following modified formulations of the implicit trapezoidal method by applying experimental data to Equations (56)–(60):

$$x_1(t_j) = x_{1e}(t_{j-1}) + 0.5\eta_j[\mu(x_e(t_j), \mathbf{p})x_{1e}(t_j) + \mu(x_e(t_{j-1}), \mathbf{p})x_{1e}(t_{j-1})], \tag{61}$$

$$x_2(t_j) = x_{2e}(t_{j-1}) + 0.5\eta_j[-q_S(x_e(t_j), \mathbf{p})x_{1e}(t_j) - q_S(x_e(t_{j-1}), \mathbf{p})x_{1e}(t_{j-1})], \tag{62}$$

$$x_3(t_j) = x_{3e}(t_{j-1}) + 0.5\eta_j[q_{PD}(x_e(t_j), \mathbf{p})x_{1e}(t_j) + q_{PD}(x_e(t_{j-1}), \mathbf{p})x_{1e}(t_{j-1})], \tag{63}$$

$$x_4(t_j) = x_{4e}(t_{j-1}) + 0.5\eta_j[q_{HAc}(x_e(t_j), \mathbf{p})x_{1e}(t_j) + q_{HAc}(x_e(t_{j-1}), \mathbf{p})x_{1e}(t_{j-1})], \tag{64}$$

$$x_5(t_j) = x_{5e}(t_{j-1}) + 0.5\eta_j[q_{EtOH}(x_e(t_j), \mathbf{p})x_{1e}(t_j) + q_{EtOH}(x_e(t_{j-1}), \mathbf{p})x_{1e}(t_{j-1})], \tag{65}$$

where $j = 1, 2, \dots, N$.

Replacing Equations (43)–(47) with Equations (61)–(65), we can convert the dynamic optimization problem in Equations (42)–(55) into the following nonlinear programming problem:

$$\min F = \frac{1}{N} \sum_{j=1}^N \sum_{i=1}^5 [(x_i(t_j) - x_{ie}(t_j))^2 / x_{i\max}^2] + \frac{1}{N} \sum_{j=1}^N \sum_{i=1}^5 [(\dot{x}_i(t_j) - \dot{x}_{ie}(t_j))^2 / \dot{x}_{i\max}^2], \tag{66}$$

subject to satisfying

$$x_1(t_j) = x_{1e}(t_{j-1}) + 0.5\eta_j[\mu(x_e(t_j), \mathbf{p})x_{1e}(t_j) + \mu(x_e(t_{j-1}), \mathbf{p})x_{1e}(t_{j-1})], j = 1, 2, \dots, N, \tag{67}$$

$$x_2(t_j) = x_{2e}(t_{j-1}) + 0.5\eta_j[-q_S(x_e(t_j), \mathbf{p})x_{1e}(t_j) - q_S(x_e(t_{j-1}), \mathbf{p})x_{1e}(t_{j-1})], j = 1, 2, \dots, N, \tag{68}$$

$$x_3(t_j) = x_{3e}(t_{j-1}) + 0.5\eta_j[q_{PD}(x_e(t_j), \mathbf{p})x_{1e}(t_j) + q_{PD}(x_e(t_{j-1}), \mathbf{p})x_{1e}(t_{j-1})], j = 1, 2, \dots, N, \tag{69}$$

$$x_4(t_j) = x_{4e}(t_{j-1}) + 0.5\eta_j[q_{HAc}(x_e(t_j), \mathbf{p})x_{1e}(t_j) + q_{HAc}(x_e(t_{j-1}), \mathbf{p})x_{1e}(t_{j-1})], j = 1, 2, \dots, N, \tag{70}$$

$$x_5(t_j) = x_{5e}(t_{j-1}) + 0.5\eta_j[q_{EtOH}(x_e(t_j), \mathbf{p})x_{1e}(t_j) + q_{EtOH}(x_e(t_{j-1}), \mathbf{p})x_{1e}(t_{j-1})], j = 1, 2, \dots, N, \tag{71}$$

$$\mathbf{x}(t_0) = \mathbf{x}_0, \tag{72}$$

$$\mu(\mathbf{x}, \mathbf{p}) = \frac{0.67x_2}{0.28 + x_2} \left(1 - \frac{x_2}{2039}\right) \left(1 - \frac{x_3}{939.5}\right) \left(1 - \frac{x_4}{1026}\right) \left(1 - \frac{x_5}{360.9}\right) \geq 0, \tag{73}$$

$$q_S(\mathbf{x}, \mathbf{p}) = p_1 + \frac{0.67x_2}{p_5(0.28 + x_2)} \left(1 - \frac{x_2}{2039}\right) \left(1 - \frac{x_3}{939.5}\right) \left(1 - \frac{x_4}{1026}\right) \left(1 - \frac{x_5}{360.9}\right) + \frac{p_9x_2}{x_2 + p_{13}} \geq 0, \tag{74}$$

$$q_{PD}(\mathbf{x}, \mathbf{p}) = p_2 + \frac{0.67p_6x_2}{0.28 + x_2} \left(1 - \frac{x_2}{2039}\right) \left(1 - \frac{x_3}{939.5}\right) \left(1 - \frac{x_4}{1026}\right) \left(1 - \frac{x_5}{360.9}\right) + \frac{p_{10}x_2}{x_2 + p_{14}} \geq 0, \tag{75}$$

$$q_{HAc}(\mathbf{x}, \mathbf{p}) = p_3 + \frac{0.67p_7x_2}{0.28 + x_2} \left(1 - \frac{x_2}{2039}\right) \left(1 - \frac{x_3}{939.5}\right) \left(1 - \frac{x_4}{1026}\right) \left(1 - \frac{x_5}{360.9}\right) + \frac{p_{11}x_2}{x_2 + p_{15}} \geq 0, \tag{76}$$

$$q_{EtOH}(\mathbf{x}, \mathbf{p}) = p_4 + \frac{0.67p_8x_2}{0.28 + x_2} \left(1 - \frac{x_2}{2039}\right) \left(1 - \frac{x_3}{939.5}\right) \left(1 - \frac{x_4}{1026}\right) \left(1 - \frac{x_5}{360.9}\right) + \frac{p_{12}x_2}{x_2 + p_{16}} \geq 0, \tag{77}$$

$$\mathbf{p}^l \leq \mathbf{p} \leq \mathbf{p}^u. \tag{78}$$

To efficiently solve the nonlinear programming problem in Equations (66)–(78), we first regard $\mu(t_j)$ ($j = 0, 1, \dots, N$), $q_S(t_j)$, $q_{PD}(t_j)$, $q_{HAc}(t_j)$, and $q_{EtOH}(t_j)$ as the identified parameters and rewrite it as

$$\min F = \frac{1}{N} \sum_{j=1}^N \sum_{i=1}^5 [(x_i(t_j) - x_{ie}(t_j))^2 / x_{i\max}^2] + \frac{1}{N} \sum_{j=1}^N \sum_{i=1}^5 [(\dot{x}_i(t_j) - \dot{x}_{ie}(t_j))^2 / \dot{x}_{i\max}^2], \tag{79}$$

subject to satisfying

$$x_1(t_j) = x_{1e}(t_{j-1}) + 0.5\eta_j[\mu(t_j)x_{1e}(t_j) + \mu(t_{j-1})x_{1e}(t_{j-1})], j = 0, 1, \dots, N, \tag{80}$$

$$x_2(t_j) = x_{2e}(t_{j-1}) + 0.5\eta_j[-q_S(t_j)x_{1e}(t_j) - q_S(t_{j-1})x_{1e}(t_{j-1})], j = 0, 1, \dots, N, \tag{81}$$

$$x_3(t_j) = x_{3e}(t_{j-1}) + 0.5\eta_j[q_{PD}(t_j)x_{1e}(t_j) + q_{PD}(t_{j-1})x_{1e}(t_{j-1})], j = 0, 1, \dots, N, \tag{82}$$

$$x_4(t_j) = x_{4e}(t_{j-1}) + 0.5\eta_j[q_{HAc}(t_j)x_{1e}(t_j) + q_{HAc}(t_{j-1})x_{1e}(t_{j-1})], j = 0, 1, \dots, N, \tag{83}$$

$$x_5(t_j) = x_{5e}(t_{j-1}) + 0.5\eta_j[q_{EtOH}(t_j)x_{1e}(t_j) + q_{EtOH}(t_{j-1})x_{1e}(t_{j-1})], j = 0, 1, \dots, N, \tag{84}$$

$$\mathbf{x}(t_0) = \mathbf{x}_0, \tag{85}$$

$$\mu_j^l \leq \mu(t_j) \leq \mu_j^u, j = 0, 1, \dots, N, \tag{86}$$

$$q_{Sj}^l \leq q_S(t_j) \leq q_{Sj}^u, j = 0, 1, \dots, N, \tag{87}$$

$$q_{PDj}^l \leq q_{PD}(t_j) \leq q_{PDj}^u, j = 0, 1, \dots, N, \tag{88}$$

$$q_{HAcj}^l \leq q_{HAc}(t_j) \leq q_{HAcj}^u, j = 0, 1, \dots, N, \tag{89}$$

$$q_{EtOHj}^l \leq q_{EtOH}(t_j) \leq q_{EtOHj}^u, j = 0, 1, \dots, N, \tag{90}$$

where $\mu_j^l, q_{Sj}^l, q_{PDj}^l, q_{HAcj}^l$ and q_{EtOHj}^l are the lower bounds of $\mu(t_j), q_S(t_j), q_{PD}(t_j), q_{HAc}(t_j)$, and $q_{EtOH}(t_j)$, respectively; and $\mu_j^u, q_{Sj}^u, q_{PDj}^u, q_{HAcj}^u$ and q_{EtOHj}^u are the upper bounds of $\mu(t_j), q_S(t_j), q_{PD}(t_j), q_{HAc}(t_j)$, and $q_{EtOH}(t_j)$, respectively.

For the parameter identification problem in Equations (79)–(90), we have the following remarks:

1. In the computation of the optimization problem in Equations (79)–(90), the slopes $\dot{x}_i(t_j)$ ($i = 1, 2, \dots, 5$) in the objective function of Equation (79) can be computed by the following formulations:

$$\dot{x}_1(t_j) = \mu(t_j)x_{1e}(t_j),$$

$$\dot{x}_2(t_j) = -q_S(t_j)x_{1e}(t_j),$$

$$\dot{x}_3(t_j) = q_{PD}(t_j)x_{1e}(t_j),$$

$$\dot{x}_4(t_j) = q_{HAc}(t_j)x_{1e}(t_j),$$

$$\dot{x}_5(t_j) = q_{EtOH}(t_j)x_{1e}(t_j).$$

The slopes $\dot{x}_{ie}(t_j)$ ($i = 1, 2, \dots, 5$) of experimental data can be estimated by the method given in Section 3.2;

2. The optimization problem in Equations (79)–(90) is a relatively simple quadratic programming problem compared to the nonlinear programming problem in Equations (66)–(78). Thus, it is easy to obtain the globally optimal solution of the problem in Equations (79)–(90) with the available quadratic programming algorithms.

Let the optimal values of $\mu(t_j)$ ($j = 0, 1, \dots, N$), $q_S(t_j), q_{PD}(t_j), q_{HAc}(t_j)$, and $q_{EtOH}(t_j)$ be $\mu^*(t_j), q_S^*(t_j), q_{PD}^*(t_j), q_{HAc}^*(t_j)$, and $q_{EtOH}^*(t_j)$, respectively. Substituting $\mu^*(t_j), q_S^*(t_j), q_{PD}^*(t_j), q_{HAc}^*(t_j)$, and $q_{EtOH}^*(t_j)$ into $\mu(\mathbf{x}, \mathbf{p}), q_S(\mathbf{x}, \mathbf{p}), q_{PD}(\mathbf{x}, \mathbf{p}), q_{HAc}(\mathbf{x}, \mathbf{p})$, and $q_{EtOH}(\mathbf{x}, \mathbf{p})$, respectively, we have the following $4(N + 1)$ nonlinear equations:

$$p_1 + \frac{\mu^*(t_j)}{p_5} + \frac{p_9 x_{2e}(t_j)}{x_{2e}(t_j) + p_{13}} = q_S^*(t_j), j = 0, 1, \dots, N, \tag{91}$$

$$p_2 + p_6 \mu^*(t_j) + \frac{p_{10} x_{2e}(t_j)}{x_{2e}(t_j) + p_{14}} = q_{PD}^*(t_j), j = 0, 1, \dots, N, \tag{92}$$

$$p_3 + p_7 \mu^*(t_j) + \frac{p_{11} x_{2e}(t_j)}{x_{2e}(t_j) + p_{15}} = q_{HAc}^*(t_j), j = 0, 1, \dots, N, \tag{93}$$

$$p_4 + p_8\mu^*(t_j) + \frac{p_{12}x_{2e}(t_j)}{x_{2e}(t_j) + p_{16}} = q_{EtOH}^*(t_j), j = 0, 1, \dots, N. \tag{94}$$

After solving the nonlinear Equations (91)–(94), we can obtain the values of the identified parameters $p_k(j = 1, 2, \dots, 16)$. This goal can be achieved by solving the following optimization problem:

$$\min \bar{F} = \sum_{j=0}^N [F_1^2(t_j) + F_2^2(t_j) + F_3^2(t_j) + F_4^2(t_j)], \tag{95}$$

subject to satisfying

$$p^l \leq p \leq p^u. \tag{96}$$

In this problem, $F_1(t_j)$, $F_2(t_j)$, $F_3(t_j)$, and $F_4(t_j)$ have the following formulations

$$F_1(t_j) = p_1 + \frac{\mu^*(t_j)}{p_5} + \frac{p_9x_{2e}(t_j)}{x_{2e}(t_j) + p_{13}} - q_S^*(t_j), j = 0, 1, \dots, N,$$

$$F_2(t_j) = p_2 + p_6\mu^*(t_j) + \frac{p_{10}x_{2e}(t_j)}{x_{2e}(t_j) + p_{14}} - q_{PD}^*(t_j), j = 0, 1, \dots, N,$$

$$F_3(t_j) = p_3 + p_7\mu^*(t_j) + \frac{p_{11}x_{2e}(t_j)}{x_{2e}(t_j) + p_{15}} - q_{HAc}^*(t_j), j = 0, 1, \dots, N,$$

$$F_4(t_j) = p_4 + p_8\mu^*(t_j) + \frac{p_{12}x_{2e}(t_j)}{x_{2e}(t_j) + p_{16}} - q_{EtOH}^*(t_j), j = 0, 1, \dots, N.$$

The optimal values of the identified parameters $p_k(k = 1, 2, \dots, 16)$ can be obtained by solving the problem in Equations (95) and (96) with the available optimization solver.

3.2. Computing the Slopes of Experimental Data

As stated in the optimization problem in Equations (79)–(90), the objective of Equation (79) involves the slopes $\dot{x}_{ie}(t_j)$ of $5N$ experimental data. These experimental slopes can be estimated by some numerical methods, such as the artificial neural network method [28] and the cubic spline interpolation algorithm [29]. In this work, based on the Tikhonov regularization method [30], we present the following procedures to estimate the experimental slopes $\dot{x}_{ie}(t_j)$.

We first divide the time interval $[t_0, t_N]$ into $N_K - 1$ equidistant subintervals $[t_0, t_0 + l], \dots, [t_0 + rl, t_0 + (r + 1)l], \dots, [t_0 + (N_K - 2)l, t_N]$ ($l = (t_N - t_0)/(N_K - 1)$, $r = 0, 1, 2, \dots, N_K - 2$), such that $t_j \in \{t_0, t_0 + l, t_0 + 2l, \dots, t_0 + rl, t_0 + (r + 1)l, \dots, t_0 + (N_K - 2)l, t_N\}$ ($j = 0, 1, \dots, N$). Then, we compute the slopes (denoting them as d_{is} , $s = 1, 2, \dots, N_K$) at $t_0 + rl$ ($r = 0, 1, 2, \dots, N_K - 2$) by the following equations:

$$e_i = (B^T B + 0.48\beta^T \beta)^{-1} B^T y_i, i = 1, 2, \dots, 5,$$

where

$$e_i = (d_{i1}, d_{i2}, \dots, d_{iN_K}, x_i(t_0))^T,$$

$$y_i = (x_{ie}(t_0), x_{ie}(t_1), \dots, x_{ie}(t_N))^T,$$

$$\beta = \begin{bmatrix} 1 & -2 & 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 & \dots & 0 & 0 \\ \vdots & & \ddots & \ddots & \ddots & \ddots & \vdots & \vdots \\ 0 & \dots & 0 & 1 & -2 & 1 & 0 & 0 \end{bmatrix},$$

$$B = [A \quad E].$$

Here, both the matrix $A = (A_{js})_{(N+1) \times N_K}$ and vector $E = (1, 1, \dots, 1)^T \in \mathbb{R}^N$ are obtained by the numerical integral $x_{ij} = \sum_{s=1}^{N_K} A_{js}d_{is} + x_i(t_0)$, $i = 1, 2, \dots, 5$, $j = 0, 1, \dots, N$ of integral equations $x_i(t) = \int_{t'=t_0}^t d_i(t')dt' + x_i(t_0)$, $i = 1, 2, \dots, 5$.

4. Optimization Results and Discussion

In this work, we applied the Tikhonov regularization method given in Section 3.2 to estimate the experimental slopes $\dot{x}_{ie}(t_j)$ and used the genetic algorithm in MATLABTM software to solve both the optimization problems of Equations (79)–(90) and of Equations (95) and (96). In the implementation of the genetic algorithm, we set all the algorithm parameters to be the default values. Experimental concentration data $x_{ie}(t_j)$ ($i = 1, 2, \dots, 5$) were drawn from Feng and Xiu [31] and Xu and Wang [17]. The initial point x_0 of the studied system was selected as $(0.1905, 400.043, 0, 0, 0)^T$. The parameter N_K used in the Tikhonov regularization method was set to be 502. The lower and upper bounds for the optimized variables $\mu(t_j)$ ($j = 0, 1, \dots, N$), $q_S(t_j)$, $q_{PD}(t_j)$, $q_{HAc}(t_j)$, and $q_{EtOH}(t_j)$ and the identified parameters p_k ($k = 1, 2, \dots, 16$) are given in Tables 1 and 2, respectively.

Table 1. The lower and upper bounds in the optimization problem of Equations (79)–(90).

| Parameter | Lower Bound | Upper Bound |
|-----------------|-------------|-------------|
| $\mu(t_0)$ | 0 | 0.67 |
| $\mu(t_1)$ | 0 | 0.67 |
| $\mu(t_2)$ | 0 | 0.67 |
| $\mu(t_3)$ | 0 | 0.67 |
| $\mu(t_4)$ | 0 | 0.67 |
| $q_S(t_0)$ | 0 | 100 |
| $q_S(t_1)$ | 0 | 100 |
| $q_S(t_2)$ | 0 | 100 |
| $q_S(t_3)$ | 0 | 100 |
| $q_S(t_4)$ | 0 | 100 |
| $q_{PD}(t_0)$ | 0 | 100 |
| $q_{PD}(t_1)$ | 0 | 100 |
| $q_{PD}(t_2)$ | 0 | 100 |
| $q_{PD}(t_3)$ | 0 | 100 |
| $q_{PD}(t_4)$ | 0 | 100 |
| $q_{HAc}(t_0)$ | 0 | 100 |
| $q_{HAc}(t_1)$ | 0 | 100 |
| $q_{HAc}(t_2)$ | 0 | 100 |
| $q_{HAc}(t_3)$ | 0 | 100 |
| $q_{EtOH}(t_0)$ | 0 | 100 |
| $q_{EtOH}(t_1)$ | 0 | 100 |
| $q_{EtOH}(t_2)$ | 0 | 100 |
| $q_{EtOH}(t_3)$ | 0 | 100 |
| $q_{EtOH}(t_4)$ | 0 | 100 |

Figure 1 presents the optimal values of the variables $\mu(t_j)$ ($j = 0, 1, \dots, N$), $q_S(t_j)$, $q_{PD}(t_j)$, $q_{HAc}(t_j)$, and $q_{EtOH}(t_j)$, obtained by solving the optimization problem in Equations (79)–(90) in the first stage of the two-stage method. Applying these optimal values to the optimization problem in Equations (95) and (96) and then solving it in the second stage of the two-stage method, we could obtain the optimal values of 16 identified parameters p_k ($k = 1, 2, \dots, 16$). These results are shown in Table 2. Table 2 also gives the lower and upper bounds for identified parameters p_k ($k = 1, 2, \dots, 16$). Table 3 shows the comparison between the proposed two-stage method and the single-stage method applied in Reference [17]. As can be seen in the total error function F value, the proposed two-stage method in this work could yield a smaller error between the measured and computed concentrations than the single-stage method applied in Reference [17], with a decrease of about 52.79%. This conclusion clearly

shows the tractability and effectiveness of the proposed two-stage method in handling the parameter identification of the microbial batch process of glycerol to 1,3-PD.

Table 2. The lower bounds, upper bounds, and optimal values of the identified parameters.

| Parameter | Lower Bound | Optimal Value | Upper Bound |
|-----------|-------------|---------------|-------------|
| p_1 | 1 | 4.945 | 5 |
| p_2 | 0.0001 | 0.024 | 2 |
| p_3 | 10 | 10.075 | 50 |
| p_4 | 1 | 20.427 | 30 |
| p_5 | -5 | -4.823 | -1 |
| p_6 | 10 | 34.755 | 100 |
| p_7 | 1 | 4.246 | 50 |
| p_8 | 1 | 23.830 | 100 |
| p_9 | -2 | -0.285 | -0.01 |
| p_{10} | 10 | 19.235 | 50 |
| p_{11} | 1 | 1.113 | 10 |
| p_{12} | 10 | 86.994 | 100 |
| p_{13} | -10 | -4.554 | -0.01 |
| p_{14} | 2 | 20.192 | 50 |
| p_{15} | 1 | 2.370 | 20 |
| p_{16} | 10 | 95.515 | 100 |

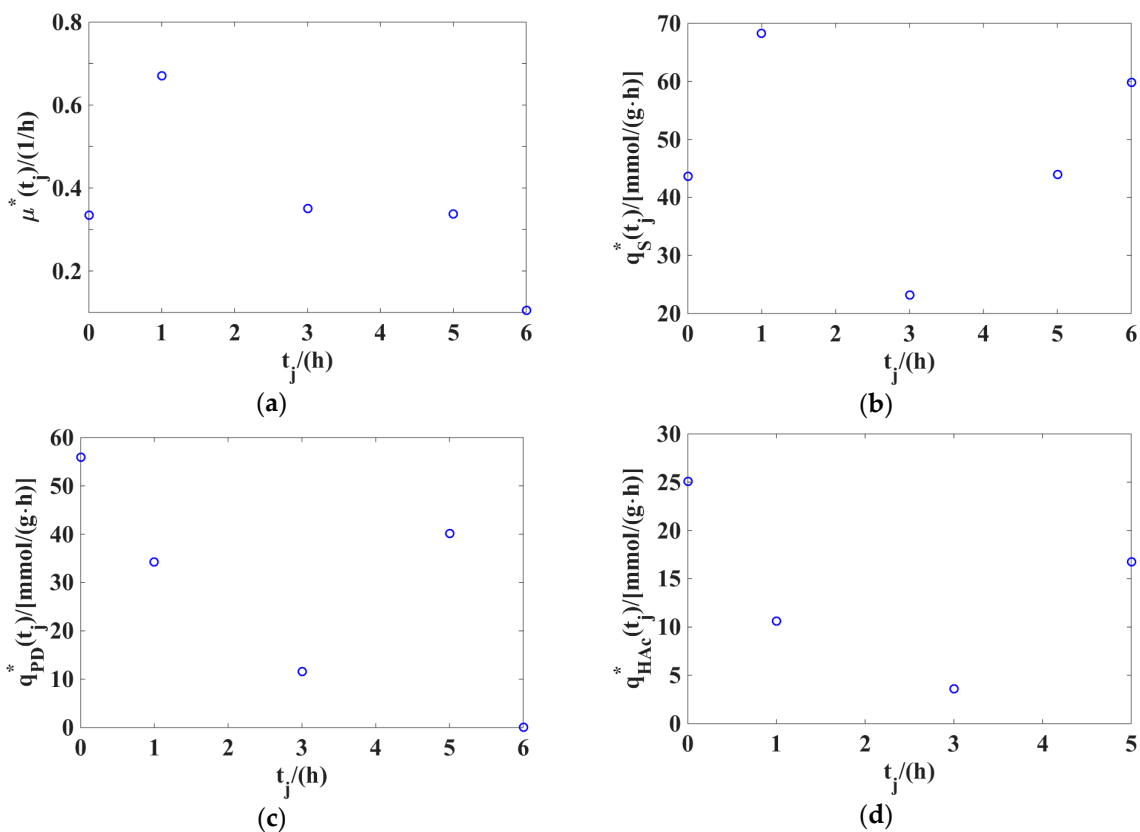


Figure 1. Cont.

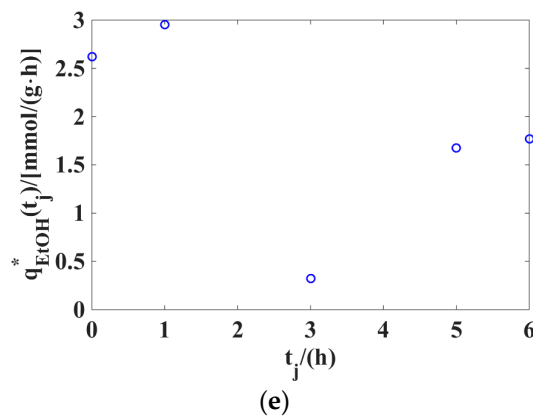


Figure 1. The optimal values $\mu^*(t_j)$, $q_s^*(t_j)$, $q_{PD}^*(t_j)$, $q_{HAc}^*(t_j)$, and $q_{EtOH}^*(t_j)$ of the variables $\mu(t_j)$, $q_s(t_j)$, $q_{PD}(t_j)$, $q_{HAc}(t_j)$, and $q_{EtOH}(t_j)$, obtained by solving the optimization problem in Equations (79)–(90) in the first stage of the two-stage method: (a) $\mu^*(t_j)$; (b) $q_s^*(t_j)$; (c) $q_{PD}^*(t_j)$; (d) $q_{HAc}^*(t_j)$; (e) $q_{EtOH}^*(t_j)$.

Table 3. Comparison between the proposed two-stage method and the single-stage approach from Reference [17].

| Method | Proposed Two-Stage Method | Single-Stage Method from Reference [17] |
|--------|---------------------------|---|
| F | 0.245 | 0.519 |

Figures 2 and 3 present the comparisons of biomass, glycerol, 1,3-PD, acetic acid, and ethanol between the experimental data and the computed values for the parameter identification of the microbial batch process. From Figures 2 and 3, we can see that the calculated concentrations with the proposed two-stage method were better fitted to the experimental data than with the single-stage method. To further measure the goodness-of-fit of the model, the coefficient of determination in statistical analysis was used as a metric. The coefficient of determination of the presented study was 0.9994 (0.9994 is closer to 1), and was larger than the coefficient of determination (0.9243) obtained by the single-stage method, with an improvement of about 8.13%. This concludes that the biological model obtained by the two-stage method provided better goodness-of-fit for the experimental concentration data than the single-stage method did.

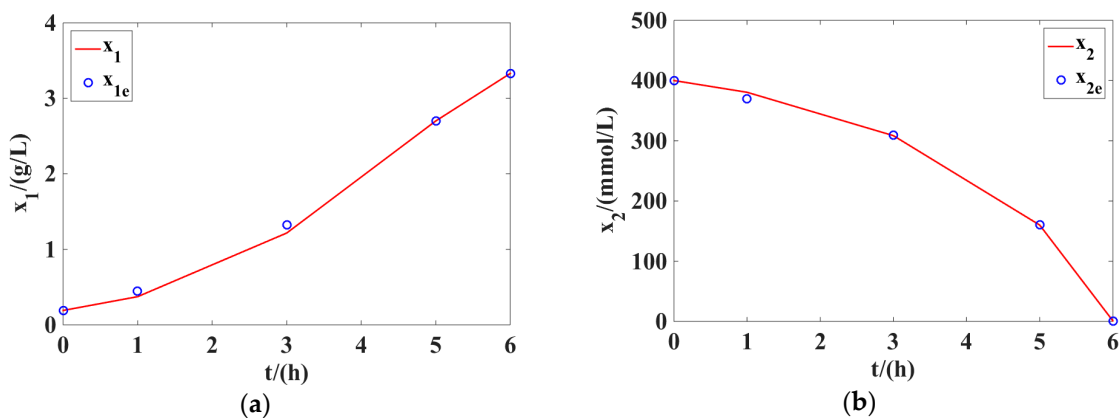


Figure 2. Cont.

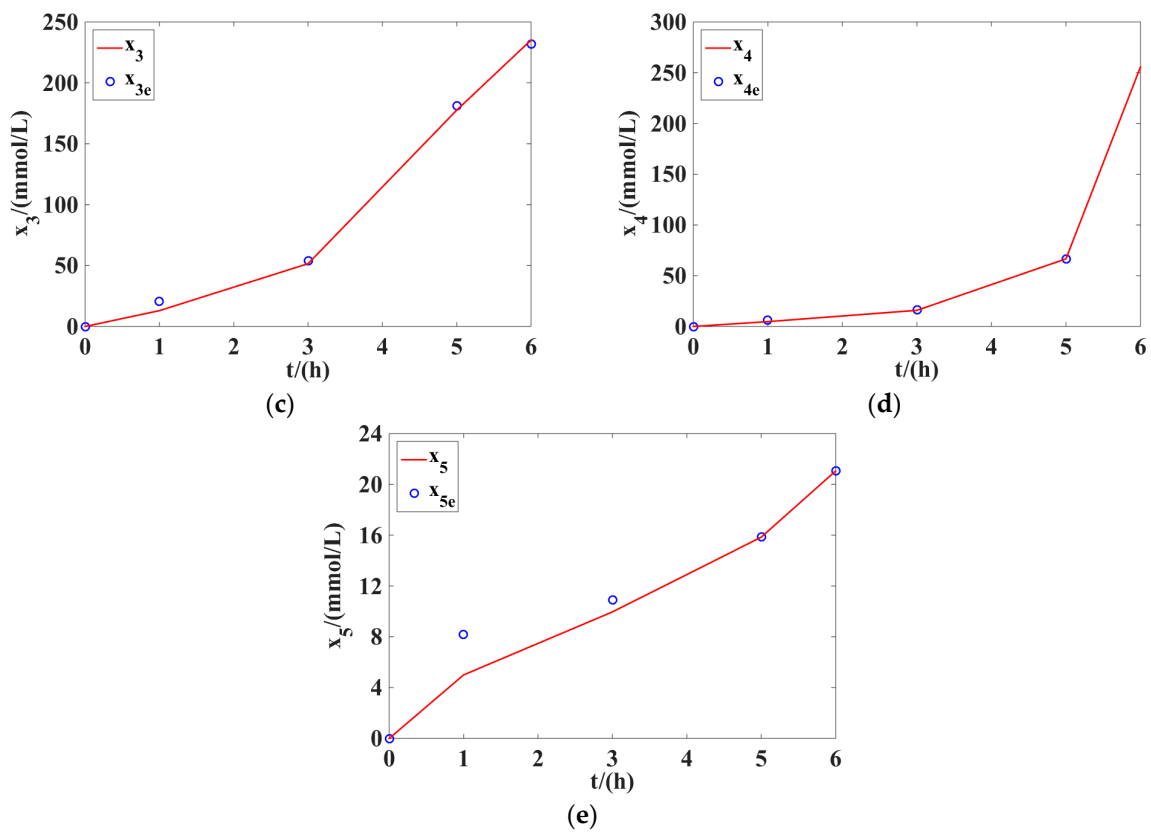


Figure 2. Comparison between the computed values and the experimental data in the proposed two-stage method: (a) biomass; (b) substrate glycerol; (c) 1,3-propanediol (1,3-PD); (d) acetic acid; (e) ethanol.

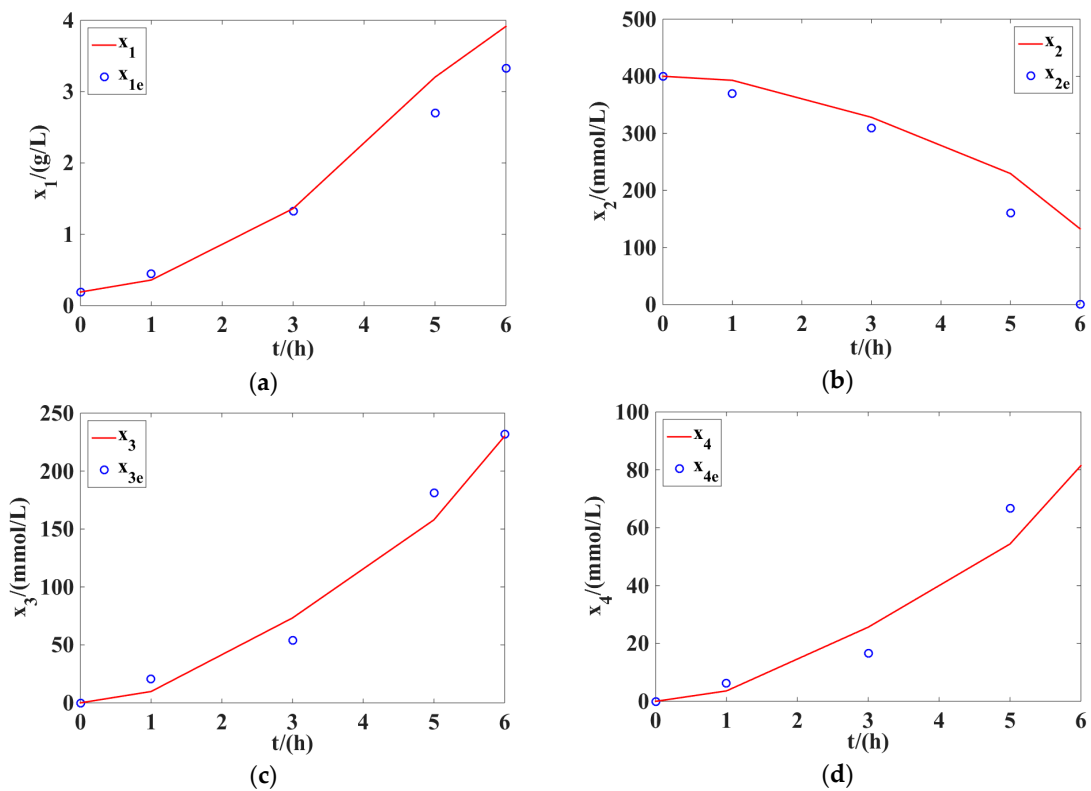


Figure 3. Cont.

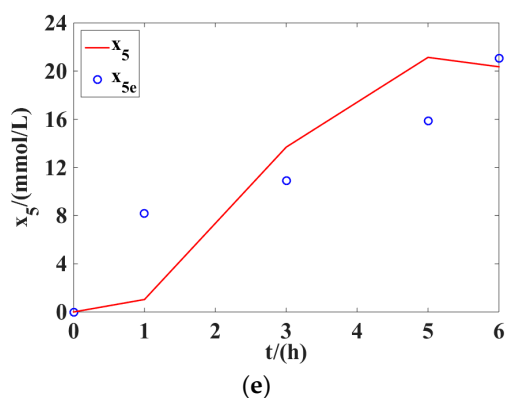


Figure 3. Comparison between the computed values and the experimental data in the single-stage method: (a) biomass; (b) substrate glycerol; (c) 1,3-PD; (d) acetic acid; (e) ethanol.

5. Conclusions

This paper addressed the problem of parameter identification for a microbial batch process of glycerol to 1,3-PD. A two-stage method was proposed to efficiently solve the presented parameter identification problem. In the first stage of this method, a simple quadratic programming problem is first required to be solved. The optimized variables of this quadratic programming problem are the specific growth rate μ , specific consumption rate q_s , and specific formation rates q_{PD} , q_{HAc} , and q_{EtOH} at time t_j ($j = 0, 1, \dots, N$). Applying the optimization results of the quadratic programming problem to the second stage of the proposed two-stage method, we can obtain the values of the 16 identified parameters. A comparative study was conducted and showed that the proposed two-stage method could obtain better identification results than the single-stage method could. Although the two-stage method was proposed here to identify the parameters of a microbial batch process of glycerol to 1,3-PD, it can also be applied to the parameter identification of other biological systems.

Author Contributions: G.X., D.L., and W.T. contributed to the design and implementation of the research, to the analysis of the results, and to the writing of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China (Grant number 11101051), the Liaoning Provincial Natural Science Foundation of China (Grant number 20180550839), and the Basic Research Fund of Liaoning Education Department (Grant number LF2017002).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Biebl, H.; Menzel, K.; Zeng, A.-P.; Deckwer, W.-D. Microbial production of 1,3-propanediol. *Appl. Microbiol. Biotechnol.* **1999**, *52*, 289–297. [[CrossRef](#)] [[PubMed](#)]
2. Xiu, Z.; Song, B.; Wang, Z.; Sun, L.; Feng, E.; Zeng, A.-P. Optimization of dissimilation of glycerol to 1,3-propanediol by *Klebsiella pneumoniae* in one- and two-stage anaerobic cultures. *Biochem. Eng. J.* **2004**, *19*, 189–197. [[CrossRef](#)]
3. Xiu, Z.; Zeng, A.-P.; Deckwer, W.-D. Multiplicity and stability analysis of microorganisms in continuous culture: Effects of metabolic overflow and growth inhibition. *Biotechnol. Bioeng.* **1998**, *57*, 251–261. [[CrossRef](#)]
4. Zeng, A.-P.; Deckwer, W.-D. A kinetic model for substrate and energy consumption of microbial growth under substrate-sufficient conditions. *Biotechnol. Prog.* **1995**, *11*, 71–79. [[CrossRef](#)] [[PubMed](#)]
5. Zeng, A.-P.; Biebl, H. Bulk-chemicals from biotechnology: The case of microbial production of 1,3-propanediol and the new trends. In *Tools and Applications of Biochemical Engineering Science*; Schüger, K., Zeng, A.-P., Eds.; Springer: Berlin/Heidelberg, Germany, 2002; Volume 74, pp. 239–259, ISBN 978-3-540-45736-7.
6. Lama, S.; Seol, E.; Park, S. Metabolic engineering of *Klebsiella pneumoniae* J2B for the production of 1,3-propanediol from glucose. *Bioresour. Technol.* **2017**, *245*, 1542–1550. [[CrossRef](#)]

7. Lee, C.S.; Aroua, M.K.; Daud, W.M.A.W.; Cognet, P.; Pérès-Lucchese, Y.; Fabre, P.-L.; Reynes, O.; Latapie, L. A review: Conversion of bioglycerol into 1,3-propanediol via biological and chemical method. *Renew. Sustain. Energy Rev.* **2015**, *42*, 963–972. [[CrossRef](#)]
8. Sun, Y.; Shen, J.; Yan, L.; Zhou, J.; Jiang, L.; Chen, Y.; Yuan, J.; Feng, E.; Xiu, Z. Advances in bioconversion of glycerol to 1,3-propanediol: Prospects and challenges. *Process Biochem.* **2018**, *71*, 134–146. [[CrossRef](#)]
9. Vivek, N.; Aswathi, T.V.; Sven, P.R.; Pandey, A.; Binod, P. Self-cycling fermentation for 1,3-propanediol production: Comparative evaluation of metabolite flux in cell recycling, simple batch and continuous processes using *Lactobacillus brevis* N1E9.3.3 strain. *J. Biotechnol.* **2017**, *259*, 110–119. [[CrossRef](#)]
10. Rodriguez, A.; Wojtusik, M.; Masca, F.; Santos, V.E.; Garcia-Ochoa, F. Kinetic modeling of 1,3-propanediol production from raw glycerol by *Shimwellia blattae*: Influence of the initial substrate concentration. *Biochem. Eng. J.* **2017**, *117*, 57–65. [[CrossRef](#)]
11. Liu, C.; Gong, Z.; Teo, K.L.; Loxton, R.; Feng, E. Bi-objective dynamic optimization of a nonlinear time-delay system in microbial batch process. *Optim. Lett.* **2018**, *12*, 1249–1264. [[CrossRef](#)]
12. Yuan, J.; Liu, C.; Zhang, X.; Xie, J.; Feng, E.; Yin, H.; Xiu, Z. Optimal control of a batch fermentation process with nonlinear time-delay and free terminal time and cost sensitivity constraint. *J. Process Control* **2016**, *44*, 41–52. [[CrossRef](#)]
13. Hirokawa, Y.; Maki, Y.; Hanai, T. Improvement of 1,3-propanediol production using an engineered cyanobacterium, *Synechococcus elongatus* by optimization of the gene expression level of a synthetic metabolic pathway and production conditions. *Metab. Eng.* **2017**, *39*, 192–199. [[CrossRef](#)] [[PubMed](#)]
14. Narisetty, V.; Astray, G.; Gullón, B.; Castro, E.; Parameswaran, B.; Pandey, A. Improved 1,3-propanediol production with maintained physical conditions and optimized media composition: Validation with statistical and neural approach. *Biochem. Eng. J.* **2017**, *126*, 109–117. [[CrossRef](#)]
15. Xu, G.; Liu, Y.; Gao, Q. Multi-objective optimization of a continuous bio-dissimilation process of glycerol to 1,3-propanediol. *J. Biotechnol.* **2016**, *219*, 59–71. [[CrossRef](#)] [[PubMed](#)]
16. Xu, G.; Li, C. Identifying the shared metabolic objectives of glycerol bioconversion in *Klebsiella pneumoniae* under different culture conditions. *J. Biotechnol.* **2017**, *248*, 59–68. [[CrossRef](#)] [[PubMed](#)]
17. Xu, G.; Wang, M. Parameter identification of a biological process: A comparative study. In Proceedings of the 2016 3rd International Conference on Information Science and Control Engineering (ICISCE), Beijing, China, 8–10 July 2016; IEEE Computer Society: Los Alamitos, CA, USA, 2016; pp. 1034–1038.
18. Xu, G.; Wang, D.; Li, C. Optimization of continuous bioconversion process of glycerol to 1,3-propanediol. *Int. J. Bioautomation* **2018**, *22*, 199–212. [[CrossRef](#)]
19. Wischral, D.; Zhang, J.; Cheng, C.; Lin, M.; De Souza, L.M.G.; Pessoa, F.L.P.; Pereira, N., Jr.; Yang, S. Production of 1,3-propanediol by *Clostridium beijerinckii* DSM 791 from crude glycerol and corn steep liquor: Process optimization and metabolic engineering. *Bioresour. Technol.* **2016**, *212*, 100–110. [[CrossRef](#)] [[PubMed](#)]
20. Silva, J.P.; Almeida, Y.B.; Pinheiro, I.O.; Knoelchermann, A.; Silva, J.M.F. Multiplicity of steady states in a bioreactor during the production of 1,3-propanediol by *Clostridium butyricum*. *Bioprocess Biosyst. Eng.* **2015**, *38*, 229–235. [[CrossRef](#)] [[PubMed](#)]
21. Wang, L.; Lin, Q.; Loxton, R.; Teo, K.L.; Cheng, G. Optimal 1,3-propanediol production: Exploring the trade-off between process yield and feeding rate variation. *J. Process Control* **2015**, *32*, 1–9. [[CrossRef](#)]
22. Kumar, V.; Durgapal, M.; Sankaranarayanan, M.; Somasundar, A.; Rathnasingh, C.; Song, H.; Seung, D.; Park, S. Effects of mutation of 2,3-butanediol formation pathway on glycerol metabolism and 1,3-propanediol production by *Klebsiella pneumoniae* J2B. *Bioresour. Technol.* **2016**, *214*, 432–440. [[CrossRef](#)]
23. Hirokawa, Y.; Maki, Y.; Tatsuke, T.; Hanai, T. Cyanobacterial production of 1,3-propanediol directly from carbon dioxide using a synthetic metabolic pathway. *Metab. Eng.* **2016**, *34*, 97–103. [[CrossRef](#)] [[PubMed](#)]
24. Celińska, E. *Klebsiella* spp. as a 1,3-propanediol producer—the metabolic engineering approach. *Crit. Rev. Biotechnol.* **2012**, *32*, 274–288. [[CrossRef](#)]
25. Biegler, L.T. *Nonlinear Programming: Concepts, Algorithm, and Applications to Chemical Processes*; SIAM: Philadelphia, PA, USA, 2010; pp. 287–324, ISBN 978-0-898717-02-0.
26. Betts, J.T. *Practical Methods for Optimal Control and Estimation Using Nonlinear Programming*; SIAM: Philadelphia, PA, USA, 2010; pp. 91–149, ISBN 978-0-898716-88-7.
27. Guo, T.; Allison, J.T. On the use of mathematical programs with complementarity constraints in combined topological and parametric design of biochemical enzyme networks. *Eng. Optim.* **2017**, *49*, 345–364. [[CrossRef](#)]

28. Almeida, J.S.; Voit, E.O. Neural-network-based parameter estimation in S-system models of biological networks. *Genome Inform.* **2003**, *14*, 114–123. [[CrossRef](#)] [[PubMed](#)]
29. Deng, Z.; Tian, T. A continuous optimization approach for inferring parameters in mathematical models of regulatory networks. *BMC Bioinform.* **2014**, *15*, 256. [[CrossRef](#)]
30. Lubansky, A.S.; Yeow, Y.L.; Leong, Y.-K.; Wickramasinghe, S.R.; Han, B. A general method of computing the derivative of experimental data. *AIChE J.* **2006**, *52*, 323–332. [[CrossRef](#)]
31. Feng, E.; Xiu, Z. *Nonlinear Fermentation Dynamic System: Identification, Control and Parallel Optimization*; Science Press: Beijing, China, 2012; ISBN 978-7-03-034959-0.



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).