Optimization of Bioethanol Ethanol Production in Fed-batch Fermentation

Zheng Li *, Alim Dewan *
M Nazmul Karim *+

*Department of Chemical Engineering, Texas Tech University, Lubbock, TX 79409, USA.
Corresponding author; email:naz.karim@ttu.edu

Abstract: This study focuses on improving the productivity of a fed-batch ethanol fermentation process by developing and implementing in real time, an optimum feeding policy. A constraint-based stoichiometric model is developed using the systems biology approach, which can quantitatively predict the cellular behavior of the E. coli strain, KO11. However, the predictions from this model are accurate under low concentrations of substrate. In order to extend the usage of the model to higher substrate concentrations, we modified this model by introducing changes in glucose utilization rate to restrict the metabolic capacity of the cell. Using the proposed model, the fed-batch optimization problem becomes a constrained optimization problem, and a modified Iterative Dynamic Programming (IDP) algorithm was developed with an adaptive-stage updating methodology, which was applied to solve the global optimization problem. Experiments were carried out based on the optimal feed profile generated using the modified stoichiometric model and the Iterative Dynamic Programming algorithm. Approximately 90% of the theoretical ethanol yield was obtained; the fermentation consumed a total of 521.5 g of glucose and produced 237.5 g of ethanol.

Keywords: Iterative Dynamic Programming, Constraint-based Stoichiometric Model, Optimization, E. coli KO11, Fed-batch Fermentation, Bioethanol

1. INTRODUCTION

Biofuels provide a valuable route to resolving environmental issues from reliance on petroleum and keeping the price of energy affordable. Currently, most biofuel is in the form of bioethanol, which is produced from starch or sugar (Sticklen, 2008). To meet the increasing requirement of energy, one promising way is engineering more advanced microbes to utilize more abundant feeding stock, such as cellulosic biomass (Ryu & Karim, 2011; Sticklen, 2008). Another potential way is increasing the ethanol productivity through optimal control of fermentation process based on current industrial microbial process utilizing cellulosic feed stock.

This study focuses on improving the productivity from optimizing the fermentation process. To optimize the productivity of ethanol, the modeling of fermentation process becomes an essential task. Traditionally, empirical methods based on Monod equations (Huang, Shieh, & Wang, 2010; Mekarapiruk & Luus, 2000) have been adopted for the optimization of these processes. The major problem is that the model requires significant amount of data to determine the parameters (Roeva et al., 2007). However, current experimental facilities limit the number of data due to long-time interval of sampling and delay involves in biochemical analysis. Without enough data to determine the various parameters of the model, the nonlinear behavior and time- vary properties cannot accurately predicted. Besides Monod type of models, data-driven models, such as artificial neural networks (ANNs) (Kiran & Jana, 2009) are also used in fermentation modeling. However, large amount of experimental data still needed to ensure the fidelity of predictive values. Constraint-based model developed by Palsson and co-works (Varma & Palsson, 1994) is a genome-scale metabolic model. It was programmed into a software package (CORBRA) running in the MATLAB environment which allows examination of new knock out strategies and interventions of certain genes (Becker et al., 2007). This constraint-based model is a structured model based on stoichometry of biochemical reactions in the pathway of a given microorganism. With the development of System Biology, main pathway of common industrial microorganism can be identified which provide the basis of structured modeling. In this study, constraint-based model is used to model Escherichia coli (E. coli) KO11 for bioethanol production.

Fermentation can be carried out in several modes including continuous, batch, and fed batch. Fed-batch bioreactors have a number of well known advantages over batch. For example, undesired effects in fermentation such as substrate inhibition can be minimized. Besides, extension of process time and control of fermentation are also beneficial aspects of fed-batch operation (Kiran & Jana, 2009). In general, optimization of fed-batch fermentation can be described as an optimization problem of the nonlinear process. The feeding is usually taken as the main manipulate variable to optimize the objective function. There are three basic types of numerical methods for solving this kind of optimal control problems: indirect method, direct method (Barclay, Gill, & Ben Rosen, 1998; Pushpavanam, Rao, & Khan, 1999), and dynamic programming (Mekarapiruk & Luus, 2000; Rein, 1994). Direct method usually needs to numerically evaluate the
derivative of the cost function. However, this method is limited because continuous differentiabilities of state function are required. Indirect method transforms the original problem into boundary value problems (BVPs). However, if states are constrained, BVPs are difficult to solve. Iterative Dynamic Programming (IDP), developed from original dynamic programming, can iteratively find the global optimal conditions. The drawback of heavy computational burden is improved by the proposed adaptive-stage method, in this study.

2. MATERIAL AND METHODS

The paper illustrates the concept of constraint-based stoichiometric model approach to modeling of batch and fed-batch cultivation of E. coli K011 for bioethanol production. Optimization using modified IDP is carried out based on the proposed model, and compared with the Sequential Quadratic Programming algorithm. Experiments have been carried out to validate the proposed model and optimization methods.

2.1 Modified Constraint-based Stochiometric Model

Constraint-based stoichiometric model developed from System Biology perspective is a structured modeling approach which can quantitative predict cellular behavior, such as growth rate, product formation rate and substrate utilization rate (Becker, et al., 2007). This structured model utilizes stoichiometric relations of reactions in the pathway of microbial metabolism. Even in the main pathway, there is large number of reactions. In this study, 100 reactions in the main pathway (Ohta, Beall, Mejia, Shanmugam, & Ingram, 1991) and 97 corresponding metabolites are considered. So, a sparse matrix S_{max}, consisting of from stoichiometric reactions is formed where m is the number of metabolites, and n is number of fluxes. Metabolic flux v_{i+1}, defined as a vector, is flux in the pathway corresponding to various reactions. The product of sparse matrix S and metabolic flux v, is accumulation rates of internal metabolites. Because internal metabolites to get equilibrium very fast from either intercellular or extracellular oscillations (Varma & Palsson, 1994), a metabolic steady state is reached, which means accumulation rate of internal metabolites is zero. However, even with steady state assumption, the flux balance equations are still not solvable since the dimension of flux n is more than the number of metabolites, m. In order to solve these undetermined equations, an objective function is required to provide a unique solution. There are several ways to choose this objective function including minimizing ATP production, minimizing nutrient uptake, and maximizing metabolite production. (Becker, et al., 2007) In this study, a maximum cellular growth objective function is used to determine the solution of metabolic flux equations. Vector v_{n×1}, consisting of the weight and biomass is obtain from the experiment done by Varma and Palsson (Varma & Palsson, 1994).

Max \[ Z = \sum c \cdot v \]

subject to \[ S \cdot v = 0 \quad (1) \]

\[ \alpha \leq v \leq \beta \]

Where Z represents objective function to maximize the biomass growth rate, while S represents stoichiometric matrix and v is the metabolic flux vector within the upper limit vector \( \beta \) and lower limit vector \( \alpha \), respectively. Among them, 99 components in the vector \( \alpha \) and \( \beta \) are only used to constrain the reaction direction in anaerobic condition. The only one upper limit which needed to be specified is the upper limit for extracellular glucose. In other words, the flux balance model needs maximum glucose utilization for providing the upper bound on glucose uptake limit. In previous study (Varma & Palsson, 1994), maximum glucose utilization rate is determined as the ratio of the growth rate to the biomass yield in a batch reactor. However, for fermentation with high concentration of glucose, significant toxic substances are produced such as ethanol, acetic acid, and other acids. These toxic substances inhibit the ability of uptaking glucose of microbes, especially during fed-batch fermentation. Also, substrate inhibition regulates the glucose utilization. Consequently, maximum glucose utilization rate could be considered as a function of current ethanol concentration, acetic acid concentration and substrate concentration.

\[ \beta_{\text{glucose}} = \frac{k_1 s}{k_2 + s + k_1 + k_4 + p_2} \quad (2) \]

\( \beta_{\text{glucose}} \) is the maximum glucose utilization rate which is the upper limit of extracellular glucose uptake rate in vector \( \beta \) of (1). \( s \) represents concentration of glucose. \( p_1 \) and \( p_2 \) represent concentration of ethanol and acetic acid, respectively. \( k_1, k_2, k_3, k_4, \) are parameters required to be identified. To indentify these parameters, a nonlinear least squares method is to minimize the Sum of Square Error between the predicted value and experimental value. Since \( p_1 \) and \( p_2 \) are determined once \( s \) is measured, dataset of glucose concentration is only required to indentify these parameters in the model.

2.2 Mathematical Model for Fed-batch Process

In fed-batch fermentation, nutrients including growth media and substrate are continuously fed into the bioreactor during fermentation without draining any media. Based on the mass balance, the equations for the fed-batch operation are expressed as follows:

\[ \frac{dx}{dt} = \mu - u/V \]

\[ \frac{ds}{dt} = u/V \cdot (s_f - s) - Q_s x \]

\[ \frac{dp_1}{dt} = -u/V \cdot p_1 + Q_{p1} x \]

\[ \frac{dp_2}{dt} = -u/V \cdot p_2 + Q_{p2} x \]

\[ \frac{dV}{dt} = u \]

where \( x_0 \) is initial concentration of biomass \( x \), \( u \) is feed flow rate (L h\(^{-1}\)), \( s_f \) is glucose concentration in the feed (g L\(^{-1}\)), \( V \) is working volume of the bioreactor (L), \( s \) is substrate concentration in the bioreactor (g L\(^{-1}\)), \( p_1 \) is ethanol concentration in the bioreactor (g L\(^{-1}\)), \( p_2 \) is acetic acid concentration in the bioreactor (g L\(^{-1}\)), \( Q_s \) is substrate uptake rate obtained from extracellular glucose flux of constraint-
Based stoichiometric model, $Q_{p1}$ is the vector of ethanol secretion rates obtained from extracellular ethanol flux of constraint-based stoichiometric model, $Q_{p2}$ is the vector of acetic acid secretion rates obtained from extracellular acetic acid flux of constraint-based stoichiometric model, $\mu$ in the specific growth rate obtained from constraint-based stoichiometric model, it is calculated from the products of $c$ and $v$ in (1).

Based on the mathematical formulation, the aim of fermentation is finding optimal feed profile to maximize the profitability and regulate acetic acid concentration. The objective function is then,

$$\max_{u \in \mathbb{R}^P} J = p_1(P) - M \sum_{k=1}^{P} u(k) - N \sum_{k=1}^{P} \max(0, p_2(k) - p_{z2})$$

$$0 \leq u(k) \leq 0.15 \quad k = 1, ..., P$$

where $M$ is the ratio of the cost of the glucose to the sales price of ethanol. $N$ is the penalty factor for regulating acetic acid concentration. In this study, $M$ and $N$ are chosen as 0.1 and 1 based on performance of optimization. When concentration of acetic acid is higher than the threshold value $p_{z2}$, the third term of the right hand side of (4) will penalize the objective function. The threshold value $p_{z2}$ is select as 12 g L$^{-1}$ based on literature values (Takahashi, Takahashi, Carvalhal, & Alterthum, 1999).

In order to solve the optimization problem efficiently, a piecewise constant control policy is employed to maintain a certain feed rate within each time stage: $u(t) = u(k)$, $t_{k-1} < t < t_k$; $j = 1, ..., P$, so $P$ is corresponding to the final time $t_f$. And feed rate is constrained by the physical limits of the pump. In this study the maximum value 0.15 L h$^{-1}$ is selected. Therefore, the fed-batch optimization becomes a constrained optimization problem, and modified Iterative Dynamic Programming with adaptive-stage updating method is applied to solve such a global optimization problem.

### 2.3 Modified Iterative Dynamic Programming

Iterative Dynamic Programming (IDP) is an optimal control algorithm which developed from Bellman’s Dynamic Programming algorithm (Bellman, 1957). Compared to gradient method IDP is a robust stochastic method with fewer restrictions on the system (Mekarapipuk & Luus, 2000; Rein, 1994). Since IDP is a stochastic search the optimal points instead of extracting sensitivity information from derivatives and the use of costate, it make the algorithm unsusceptible to local optimal problem. For the same reasons, however, the long converge time and computation burden become the mainly drawback for the application of IDP in the real optimal control problem. With the development of high speed computer and parallel computation, this problem is minimized.

A dynamic model could be described by the vector ordinary differential equation

$$dx/ dt = f(x,u) \quad (5)$$

Where $x$ is a state vector, $u$ is the control vector, and initial condition of state is known. Control variables are regulated by upper and lower bound limits.

$$u_{lower,j} \leq u_j(t) \leq u_{upper,j}, \quad j = 1, 2, ..., P$$

In Iterative Dynamic Programming, more stages mean smoother control profile, but with a penalty of computational complexity. However, fewer stages could quickly approach the optimal value, but the controls are not smooth for a given number of stages. Multiple encouraging approaches have been developed in the past decade (Mekarapiruk & Luus, 2000; Rein, 1994; Thompson & Cluett, 2005), including simulated annealing filtering with a two-step method: control damping and pivot point test controls. These approaches make control profile smooth, but it adds extra computational burden to programming. The proposed method uses changing of stages as a trade off against the stage numbers and smoothing. Starting with limited number of stages, and increasing stage numbers during the iteration for which the cost function does not change much (smaller than 5%). The proposed approach for optimization is outlined as follows.

Step 1: Divide the fed-batch fermentation time into $P$ time stages, each of length $L$.

Step 2: An initial control policy is chosen at first iteration, and for subsequent iterations a previous best control policy is chose.

Step 3: By using the initial control policy or previous best control policy, solve equations (1) and (3) to generate state stages.

Step 4: Generate random control vector for each of the state stages. The random control vector is selected from allowable region.

Step 5: Starting from time, $t_f - L$ minimize the cost function of equation (4) based on states equation (1) to (3) by using the manipulated variable obtained from equations (4) to (6). The optimal value of manipulated variable at stage $P$ is stored in control stages. After the final stage is calculated, repeat the procedure of step 4 and step 5 to solve for the best control policy at previous stage at stage 1 is calculated.

Step 6: Reduce the region for allowable control. If the cost function does not change much (5%) compared with the previous iteration, increase the stage number. In this study, the stage number is doubled. After adjusting the number of stage, continue to Step 1. If the cost function still changes significantly, the best control policy will be refining by going to Step 2 directly.

### 2.4 Batch and fed-batch cultivation of E. coli KO11

E.coli KO11 was used as the organism in the experiments. Stock culture was inoculated into four individual 1 L shake flasks containing 200 mL with 25 g/L Luria-broth and 100 mg/L Chloramphenicol. The flasks were incubated in shaker at 300 rpm and 37 °C for 16 h. The pre-cultivated cells were spun down by centrifugation at 5000xg, 4°C for 10 min. Cells
were then transferred to a 7 L Applikon Technology autoclavable bioreactors with a 5 L maximum and 2 L minimum working volume. Fermentation experiments were conducted at 35°C, 300 rpm, anaerobically. Cultivation was maintained at pH 6.8 by automatic PI control using 4 mol/L NaOH or 1 mol/L HCl. The initial batch cultivation medium contained 25 g/L Luria-broth and 100 mg/L chloramphenicol. For fed-batch cultivation, the feeding solution consisted of glucose at a certain concentration measured in each experiment and other components were the same as the initial cultivation medium. After the batch-growth phase, the feeding started automatically by a pump controlled by the developed algorithm in MATLAB programming program. Ethanol and acetic acid were measured by HPLC (Varian ProStar 230, Agilent Technologies, Wilmington, DE, USA), 0.01 N H₂SO₄ was used as the mobile phase, refractive index detector was maintained at 60 °C column temperature. Glucose was measured by biochemistry analyzer (model 2700 select, YSI Incorporated, Yellow Springs, Ohio).

Table 1. Estimated model parameter values. Model I is model from Huang and coworkers’ paper, parameters are identified based on our batch experimental data. Model II is the proposed constraint-based stoichiometric model.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \mu_m )</td>
<td>0.74</td>
<td>( k_1 )</td>
<td>0.85</td>
</tr>
<tr>
<td>( v_{p_1} )</td>
<td>1.254</td>
<td>( k_2 )</td>
<td>103.32</td>
</tr>
<tr>
<td>( v_{p_2} )</td>
<td>1.052</td>
<td>( k_3 )</td>
<td>236.36</td>
</tr>
<tr>
<td>( Y_{p_1/s} )</td>
<td>0.024</td>
<td>( k_4 )</td>
<td>98.36</td>
</tr>
<tr>
<td>( Y_{p_2/s} )</td>
<td>0.0573</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( K_s )</td>
<td>94.56</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( K_{s*} )</td>
<td>486.25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( K_{p_1} )</td>
<td>85.67</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( K_{p_2} )</td>
<td>9.63</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( K_{p_1}' )</td>
<td>13.36</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( K_{p_1}'' )</td>
<td>106.32</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( K_{p_2}' )</td>
<td>56.32</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( K_{p_2}'' )</td>
<td>121.36</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3. RESULTS AND DISCUSSION

3.1 Model Comparison and Parameter Identification

Huang, et al (Huang, et al., 2010) has applied Monod type kinetic model to maximize ethanol production using *Saccharomyces diastaticus* LORRE 316. In order to compare their results with ours, the same model from Huang and coworkers’ paper was selected with ethanol and acetic acid as considered products. Batch experiment was carried out initiated with conditions for 0.27 g L⁻¹, 46.7 g L⁻¹, 0 g L⁻¹, 0 g L⁻¹ of cell biomass, glucose, ethanol, acetic acid concentrations respectively in the bioreactor, and other conditions were described in previous section. Results of experiment are shown in Figure 1. Values of parameters were estimated with the same procedure in Huang’s (Huang, et al., 2010). The estimated parameters are listed as Model I. All the calculations were carried out using MATLAB 2010 running on an Intel(R) Core(TM) Duo CPU E7500 personal computer. To compare, parameters of proposed constraint based were estimated based on dataset of glucose concentration. Compared with Monod type kinetic models, proposed model only have four parameters that need to be identified. These parameters are calculated from minimizing the scaled Sum of Squares for Error (SSE). The values of parameters are listed as Model II. The scaled SSE of both model are compared. The SSE of kinetic model is 0.0342, while SSE of the proposed model is 0.0312. Considering that less parameters and smaller dataset are required in proposed model, the performance of the proposed model is better than the Monod type kinetic model.

3.2 Simulation

To investigate the performance of the proposed technique, Iterative Dynamic Programming (IDP) and Sequential Quadratic Programming (SQP), are studied. The following parameters were used: initial stage P=10, a reduction factor \( \alpha = 0.7 \), initial control policy \( u_0 = 0.05 \), an initial control region size \( r = 0.2 \). The MATLAB command ODE15s was used for integration of the stiff differential equations in the constrained based stoichiometric model. For comparison, SQP, which uses gradient optimization method, was also computed. The details of SQP implemented are available from previous study (Barclay, et al., 1998). The simulations are based on proposed constraint-based stoichiometric model with initial fermentation conditions as described in previous section. The results are shown in Figure 2. The system is started in a batch mode. Simulation results for both optimization methods are the same because same model is used. However, the results of the prediction are initiated different when feeding calculated from IDP and SQP. The feeding is stared when growth moves from exponential phase to stationary phase. The movement of phase change is indicated by \( dx/dt=0 \). As Figure 2C shows, the initial feed rate based on IDP is faster than feed rate based on SQP. The faster feeding rate decreases the concentrations of ethanol and acetic acid, so the cell grows faster with less inhibition due to toxic products, as Figure 2B shows. Feed profile based on IDP is terminated after 65 hours, while feed profile based on SQP is terminated after 80 hours. To utilize the left substrate, fermentation is continued with batch mode until terminated around 105 hours. Figure 2A and Figure 2B show ethanol produced based on IDP is higher than SQP with
nearly the same acetic acid production. It can be inferred that the optimization of SQP may encounter a local minimum because the nonlinear characteristic of the model.

Figure 1 Comparison of kinetic model and stoichiometric model. Dashed lines: prediction of Model I (Monod type kinetic model in Huang’s paper). Continuous lines: prediction of proposed constrained based stoichiometric model. Symbols with bars: experimental data with 95% confidence interval. Blue: glucose concentration. Green: ethanol concentration. Red: acetic acid concentration. Black: biomass concentration.

3.3 Experiment Validation

From analysis of previous sections, it is found that the optimal feed profile can be obtained by using IDP based on constraint based stoichiometric model. Fed-batch experiment was carried out to validate propose optimization method and the modeling method. The fermentation is initiated with the same conditions as previous section and feed profile is calculated by a feed policy using from IDP, as solid line Figure 2C shows. In first 16 hours, without feeding added into system, the bioreactor ran as a batch. During the batch period, glucose concentration and the cell density were measured.

The model parameter values are given Table 1. Modified Iterative Dynamic Programming was performed to find optimal feed profile based on the model. After the feed was initiated, it was observed that the concentration of toxic products including ethanol and acetic acid decreased, while concentration of substrate eventually increased. After 75 hours, the fed-batch feeding was stopped based on the proposed feeding profile. To achieve high ethanol productivity, the fermentation system could be terminated, and the unused substrate could be recycled, since after 75 hours, as shown in Figure 3, the ethanol productivity decreased. However, in this study, we continued the fermentation until the unused substrate is consumed.
Approximately 90% of the theoretical ethanol yield was obtained; the fermentation consumed a total of 521.5 g of glucose and produced 237.5 g of ethanol.

Figure 3 Simulation and experimental validation of fermentation using the proposed feed profile. Symbols with bars: experimental data with 95% confidence interval. Blue: glucose concentration. Continuous lines: optimization results using IDP based on constraint-based stoichiometric model. Green: ethanol concentration. Red: acetic acid concentration. Black: biomass concentration. Conditions: initial glucose concentration: feed concentration: 149.05 g L⁻¹, final ethanol concentration: 47.5 g L⁻¹.

4. CONCLUSIONS

This article proposed a feeding strategy based on a metabolic model to optimize ethanol production using Escherichia coli KO11. Through modeling of substrate inhibition and product inhibition, proposed constraint-based stoichiometric model can predict microbial behavior including growth rate, product secretion rate, and substrate consumption rate in the fermentation process. Compared with Monod type model, the proposed model has fewer parameters which needed to be identified. Because of the uncertainty and noise in measurement of biosystems, future study may include the use of state estimation methods. An optimal feeding strategy and minimization of total fermentation time are accomplished by Iterative Dynamic Programming approach, which maximizes ethanol production and also minimizes toxic byproduct formation. Iterative Dynamic Programming is modified to increase computational efficiency. Experiments were carried out to valid proposed modeling and optimization methods.

REFERENCES


