

Survey of Optimal Production in Fed-Batch Yeast Reactors using Applications of Non-Linear Programming

Patel, Harshil

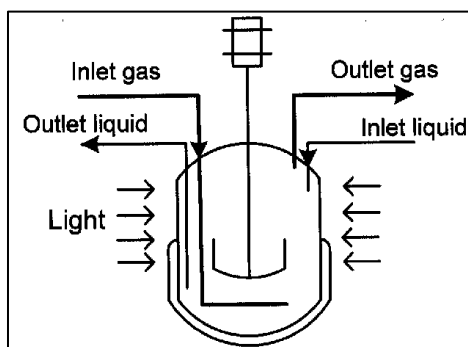
Preface

This survey relates to using nonlinear optimization to improve or optimize fed-batch reactor system regarding the discussion in the academic journal presented “*Optimization of fed-batch bioreactors using genetic algorithm: multiple control variables*” by Debasis Sarkar and Jayant M. Modak, published in Computers & Chemical Engineering. Furthermore, the original article, “*Optimal Production of Secreted Protein in Fed-Batch Reactors*” by Seujeung Park and W. Fred Ramirez, published in AIChE Journal, Volume 34. Additionally, it should be acknowledged this report was produced for ISE 3210 (Nonlinear and Dynamic Optimization) at Ohio State University and is not meant for reproduction or publication.

Introduction

Fed-batch reactors use an operational technique in which bio-technological processes where one or more nutrients are fed to the bioreactor during cultivation and in which the product remain in the bioreactor until the end of the cycle. Optimal control of fed-batch reactors or bioreactors need to maximize the biomass and/or the metabolite production at the end of the cycle time. The application of optimal fed-batch reactors can be applied in the ability of yeast cells to secrete or discharge proteins, including hormones, enzymes, toxins, and antimicrobial peptides. Yeast, or commonly called "baker's yeast" is used in baking bread and bakery products, serving as a leavening agent which causes the bread to rise by converting the fermentable sugars present in the dough into carbon dioxide and ethanol. Furthermore, with substantial post-translational modifications, there are many commercial applications of the proteins produced for mainly mass food baking and industrial food production. The secretion process and the macro-molecular activity taking place in secretion of yeast protein are further by Schekman and his associates (Novick et al., 1981). Figure 1 conveys a fed-batch reactor layout.

Figure 1: Fed-Batch Reactor



In simpler terms, the advantage or motivation in using optimal control policy is for the maximization of the harvest of baker's yeast protein bacteria in a continuous fed-batch reactor. This initial maximization is for the ease of the downstream processing later in the value chain.

Nonlinear-programming optimization techniques are used in the application for optimal control theory fed-batch bioreactors. The problem formulation consists of maximizing the harvest of the baker's yeast protein bacteria in a continuous fed-batch reactor with regard to important constraints which have been simplified for this survey. These constraints revolve around biomass, the glucose, the total protein and the secreted protein concentrations present in the initial feed of the fermentation process.

Optimal Production of Secreted Yeast Protein in a Fed-Batch Reactor (Optimization Model Formulation)

Optimal control theory is applied to optimize fed-batch production of secreted protein. We maximize an objective function that evaluates total harvest output of baker's yield protein in grams with respect to decision variables being around amounts of biomass, the glucose, the total protein and the secreted protein concentrations, and volume.

In this application x_1, x_2, x_3, x_4 are the biomass concentration (g/L), the glucose concentrations (g/L), the total initial protein (grams) and the secreted protein concentration (g/L), in that order. While x_5 is the culture volume (L). These are conveyed as the decision variables. The parameter t_f , is the final proposed cycle time which is 10 hours. The control $u(t)$ is bounded as $0 \leq u \leq 5$ and defines the specific growth rate or feed flow rate of the host cell, SEY2102 at 27 C° in the minimal media. The initial point to start the iterative solution method is $x^0 = [1, 5, 0, 0, 1, 10]$, this will find search and line methods the sequences after the initial point.

The yeast strain used to develop the material and system equations is SEY2102-s2I that contains the SUC2-s2 gene. Therefore, objective is defined is to maximize the total secreted SUC2-s2 product in the reactor at the end of each fed-batch running for a specified cycle time.

The yeast secretion harvest formulation is subject to or constrained by nine constraints. More A detailed description of the mathematical model and experimental system used is given in the original author's previous work (Park and Ramirez, 1988). The state variables, x_1, x_2, x_3, x_4, x_5 , as described above are the only state variables to be controlled. These variables are constrained by feed flow rate (given by differential equations) into the reactor given as state equations. The yield of biomass concentration is assumed to be constant and the glucose concentration of the feed is constant. The physically realizable feed flow rate is bounded. The lower limit of control action, u , or growth rate control of host cell, SEY2102 at 27 Celsius is constrained by zero and the upper limit is determined by the choice of feed system which will be 5. Also, the non-negative constraint of x_1, x_2, x_3, x_4, x_5 decision variables due to physical limitations needs to be considered. The variables, a_1, a_2, a_3 , are defined as system co-state variables to establish the problem formulation and are considered "dummy variables." They exist to supplement physical insight into the dynamic's bioreactor process during fermentation, describing the specific growth rate of the host cell, SEY2102 with given parameters that will be changed if another strain is used.

By comprising the systems and machine equations into a optimization formulation, the following mathematical model is build according to the synthesis of model parameter, decisions variables, and objective function.

$$\begin{aligned}
 & \max_{\substack{x_1, x_2, x_3, x_4, x_5, \\ u, \\ a_1, a_2, a_3}} & & (x_4(t_f)x_5(t_f)) \\
 & \text{subject to (s.t.)} \\
 & \frac{dx_1}{dt} = a_1x_1 - u \frac{x_1}{x_5} \\
 & \frac{dx_2}{dt} = -7.3a_1x_1 - u \frac{x_2 - 20}{x_5} \\
 & \frac{dx_3}{dt} = a_2x_1 - u \frac{x_3}{x_5} \\
 & \frac{dx_4}{dt} = a_3(x_3 - x_4) - u \frac{x_4}{x_5} \\
 & \frac{dx_5}{dt} = u \\
 & 0 \leq u \leq 5 \\
 & a_1 = \frac{21.87 x_2}{(x_2 + 0.4)(x_2 + 62.5)} \\
 & a_2 = \frac{x_2 e^{-5x_2}}{(x_2 + 0.1)} \\
 & a_3 = \frac{4.75a_1}{(a_1 + 0.12)} \\
 & x_i, \geq 0 \forall i = 1, 2, 3, 4, 5
 \end{aligned}$$

Problem Model with GAMs

Figure 2: GAMS Model

```

1 $Ontext
2 Survey of Optimal Production in Fed-Batch Yeast Reactors using Applications of Non-Linear Programming
3 Created by Harshil Patel on March 5, 2020
4
5 objective function is measured in grams
6 Volume is in liter (L)
7 mass in grams (g)
8 concentration is in grams per liter (g/L)
9
10 For this instance use tf(final cycle time) as 10 hours- this a data point as a parameter
11
12 Use iterative solution to solve when reduced gradient less than internal software tolerance.
13
14 $Offtext
15
16 *Set the number of iterations and save solutions for final evaluation (solution settings)
17 $if set n $set nh %nh%
18 $if not set nh $set nh 75
19
20
21
22 set nh Number of subintervals / 0x%nh% /;
23 alias (nh,k);
24
25
26 *System Setting
27 Scalar    tf    final time          / 10 /
28           x1_0  initial value for x1 / 1.0 /
29           x2_0  initial value for x2 / 5.0 /
30           x3_0  initial value for x3 / 0.0 /
31           x4_0  initial value for x4 / 0.0 /
32           x5_0  initial value for x5 / 1.0 /
33           h ;
34           h=tf/%nh%;
35
36 Variables x1(nh) biomass concentration
37           x2(nh) glucose concentration
38           x3(nh) total initial protein mass
39           x4(nh) secreted protein concentration
40           x5(nh) culture volume
41           u(nh) control variable
42           a1(nh) feedRate1
43           a2(nh) feedRate2
44           a3(nh) feedRate3
45           obj criterion ;
46
47 Equations eobj criterion definition
48           state1(nh) state equation 1
49           state2(nh) state equation 2
50           state3(nh) state equation 3
51           state4(nh) state equation 4
52           state5(nh) state equation 5
53           boundX1
54           boundX2
55           boundX3
56           boundX4
57           boundX5
58           ea1
59           ea2
60           ea3;
61

```

```

62 *Objective function
63 eobj..obj =e= x4['nh%']*x5['%nh%'];
64
65 *State Equations
66 state1(nh(k+1))..x1[k+1] =e= x1(k)+(h/2)*( a1(k)*x1(k) - u(k)*x1(k)/x5(k) +a1(k+1)*x1(k+1) - u(k+1)*x1(k+1)/x5(k+1) );
67
68 state2(nh(k+1))..x2[k+1] =e= x2(k)+(h/2)*( -7.3*a1(k)*x1(k) - u(k)*(x2(k)-20)/x5(k)-7.3*a1(k+1)*x1(k+1) - u(k+1)*(x2(k+1)-20)/x5(k+1) );
69
70 state3(nh(k+1))..x3[k+1] =e= x3(k)+(h/2)*( a2(k)*x1(k) - u(k)*x3(k)/x5(k) +a2(k+1)*x1(k+1) - u(k+1)*x3(k+1)/x5(k+1) );
71
72 state4(nh(k+1))..x4[k+1] =e= x4(k)+(h/2)*( a3(k)*(x3(k)-x4(k)) - u(k)*x4(k)/x5(k) +a3(k+1)*(x3(k+1)-x4(k+1)) - u(k+1)*x4(k+1)/x5(k+1) );
73
74 state5(nh(k+1))..x5[k+1] =e= x5(k) + (h/2)*( u(k) + u(k+1) );
75
76
77 *Descion Variable Bounds Equations
78 boundX1(nh(k))..x1(k) =g=0;
79 boundX2(nh(k))..x2(k) =g=0;
80 boundX3(nh(k))..x3(k) =g=0;
81 boundX4(nh(k))..x4(k) =g=0;
82 boundX5(nh(k))..x5(k) =g=0;
83
84
85 *Co-State Equations
86 ea1(nh(k)).. a1(k) =e= 21.87*x2(k)/((x2(k)+0.4)*(x2(k)+62.5));
87 ea2(nh(k)).. a2(k) =e= (x2(k)*exp(-5*x2(k)))/(0.1+x2(k));
88 ea3(nh(k)).. a3(k) =e= 4.75*a1(k)/(0.12*a1(k));
89
90
91
92 *Control Feed Rate Bounds for System (Found out how to does this on GAMS website)
93 u.lo(nh) = 0.0;
94 u.up(nh) = 5;
95
96
97 *Initial point
98 x1.l[nh]=1.0;
99 x2.l[nh]=5.0;
100 x3.l[nh]=0.0;
101 x4.l[nh]=0.0;
102 x5.l[nh]=1.0;
103 u.l(nh) =0.0;
104
105 x1.fx ['0'] = x1_0;
106 x2.fx ['0'] = x2_0;
107 x3.fx ['0'] = x3_0;
108 x4.fx ['0'] = x4_0;
109 x5.fx ['0'] = x5_0;
110
111 Model reactor /all/;
112 Solve reactor maximizing obj using nlp;

```

Solution

Figure 3: Optimal Solution

	LOWER	LEVEL	UPPER	MARGINAL
---- UAR obj	-INF	7.2819	+INF	.
obj criterion				
**** REPORT SUMMARY :	0	NONOPT		
	0	INFEASIBLE		
	0	UNBOUNDED		
	0	ERRORS		
	77	PROJECTED		
EXECUTION TIME	=	0.016 SECONDS	2 MB 30.2.0 r482c588 WEX-WEI	
USER: GAMS Demo license for Harshil Patel		G20030510002C0-GEN		
OHIO STATE UNIVERSITY, United States of America		DL003557		
**** FILE SUMMARY				
Input	C:\Users\Harshil Patel\OneDrive - Beam Technologies\Documents\ise 3210 hmk\reactor_fed_batch_NLP.gms			
Output	C:\Users\Harshil Patel\OneDrive - Beam Technologies\Documents\ise 3210 hmk\reactor_fed_batch_NLP.lst			

Figure 5 : Solution Time

```

** Optimal solution. Reduced gradient less than tolerance.
--- Restarting execution
--- reactor_fed_batch_NLP.gms(112) 2 Mb
--- Reading solution for model reactor
*** Status: Normal completion
--- Job reactor_fed_batch_NLP.gms Stop 03/05/20 15:41:47 elapsed 0:00:00.503

```

Table 1: Objective Function Value and Decision Values

Objective Function Value (g)	obj	7.2819
biomass concentration (g/L)	X1	2.111
glucose concentrations (g/L)	X2	2.643
total initial yeast yield(grams)	X3	1.989
total initial protein (grams)	X4	1.841
culture volume (L)	X5	3.956

The GAMS representation of the application presented in the previous section is displayed in Figure 2. The CONOPT solution algorithm was the only optimizer able to solve this application. The model was build using an iterative solution method that was able expand to 75 iterations without limiting computational power which was throttled due to only having a 7-day free license. The initial point was additionally set to start the iterative solution method at a certain objective function value. The optimal solution was identified by GAMS solvers when the reduced gradient (gradient loss) was less than the system tolerance. Additionally, the model was adapted to use differential equations through methodology found on GAMs website. The data point can be changed about the problem include changing yeast strand so the constraint parameters would change as a result. Also, initial point, termination criteria can also change.

Conclusion

The optimization of fed-batch bioreactors involving multiple singular control variables to maximize protein production from baker's yeast. The optimal yield from the reactor process or fermentation is 7.2819 grams. The optimal control variables should be set to biomass concentration (g/L), the glucose concentrations (g/L), the total initial protein (grams) and the secreted protein concentration (g/L), as shown in Table 1. While is the culture volume (L) should be set to. This solution was produced uniquely by applying an iterative solution principle. The solution time was 0.016 seconds; thus, computation time was not constrained. Although it should be noted, due to the constraint on the iterations form not having a full license a better optimal solution may exist. But it should be noted, in the last few iterations, the solution was conversing to the current optimal solution so are likely at an optimal solution which could be the global minimum or just a local minimum. Better solution algorithms may be needed. Overall a pumping up of the fresh media at the maximum rate close to the final time significantly increases the total amount of secreted protein in the culture for a short time. This, as stated before is very beneficial for mass food or baked goods production.

References

Sarkar, Debasis & Modak, Jayant. (2004). Optimization of fed-batch bioreactors using genetic algorithm: Multiple control variables. *Computers & Chemical Engineering*. 28. 789-798. 10.1016/j.compchemeng.2004.02.018.

Park, S. and Fred Ramirez, W. (1988), Optimal production of secreted protein in fed-batch reactors. *AIChE J.*, 34: 1550-1558. doi:[10.1002/aic.690340917](https://doi.org/10.1002/aic.690340917)

Park, S., and W. F. Ramirez, “Dynamics of Heterologous Protein Secretion from *Saccharomyces Cerevisiae*,” *Biotech. Bioeng.*, 33 (1988).

Schekman, R., P. Novick, S. Ferro-Novick, B. Esmond, W. Hansen, T. Etcheverry, and T. Stevens, “Protein Secretion and Organelle Assembly in Yeast,” *em Rec. Adv. Yeast Mol. Biol.*, 1, 143 (1982).

Novick, P., S. Ferro, and R. Schekman, “Order of Events in the Yeast Secretory Pathway,” *Cell*. 25,461 (1981).

Acknowledgement

The author would like to thank Dr. Ramteen Sioshani, and Professor of Integrated Systems Engineering at The Ohio State University. for teaching about nonlinear programming techniques and processes and how to work iterative solutions to complex NLPs formulations. Furthermore, the author would like to thank the honorable Hyeon Jun Kim, the Graduate Teaching Assistant who performs to a high caliber, exceeding expectations and a great instructor in learning operations research.

Furthermore, the gams model is attached to the submission for clarity and changing input data for scenario manager for different strains of yeasts.

About Author

Harshil Patel is third year Industrial and Systems Engineering student at The Ohio State University. He is currently taking Operations Research coursework. His interests include machine learning, blockchain, and integrated robotics, and cybersecurity. He can be contacted at patel.3001@osu.edu.