

Dynamic Re-optimization of a Fed-Batch Fermentor using Heuristic Dynamic Programming

Mahesh S. Iyer¹ and Donald C. Wunsch, II, *Senior Member, IEEE*

Applied Computational Intelligence Laboratory
Department of Electrical Engineering
Texas Tech University
Lubbock, TX 79409-3102
<http://www.acil.ttu.edu>

¹Address: Department of Chemical Engineering
Texas Tech University
Lubbock, TX 79409-3121

¹ Author to whom correspondence should be addressed

Abstract

Traditionally, fed-batch biochemical process optimization and control uses complicated theoretical off-line optimizers, with no on-line model adaptation or re-optimization. This study demonstrates the applicability, effectiveness, and economic potential of a simple phenomenological model for modeling, and an Adaptive Critic Design, Heuristic Dynamic Programming, for on-line re-optimization and control of an aerobic fed-batch fermentor. The results are compared with those obtained using a Heuristic Random Optimizer.

1 Introduction

Biochemical processes provide a good opportunity for optimization and control because they produce high value end products like vitamins, baker's yeast, and antibiotics [1], [2]. In addition, fermentation processes are often non-stationary and, therefore, need continually adapting recipes for optimal performance. Fed-batch fermentations have been widely investigated for both optimization and control. The most important aspects to be considered are the changes in process parameters and/or dynamics during the operation of the batch. This requires dynamically adjusting the process model, and re-optimization using the improved model. Previous research demonstrated this [3], [4], using a Heuristic Random Optimizer [5] for both off-line and on-line optimization.

This study explores a variety of control schemes including off-line optimization, on-line model re-parametrization, and on-line re-optimization of a fed-batch fermentor, using an Adaptive Critic Design, Heuristic Dynamic Programming [6]. Specifically, a rigorous phenomenological model was used to represent the fermentation process, with an intentionally different model for the optimizer (to account for the process-model mismatch that exists in an industrial setting). Off-line optimization was performed using the HRO. The one-step IMPOL technique [7] was used for dynamic model parameter adjustment. Heuristic Dynamic Programming (HDP) was utilized for on-line re-optimization, and the process performance obtained using the same was compared with that obtained using the HRO for both off-line and on-line optimization. Although the study was

conducted for a specific case of cultivation of mammalian hybridoma cells (animal cells) to produce monoclonal antibodies [8]-[10], the overall development is perfectly general, and is easily applicable to any batch process that can be modeled.

2 The Biochemical Growth System

The system studied for optimization and control was the *in vitro* growth of hybridoma cells and the production of monoclonal antibodies by these cell lines. The cell culture medium was complex, containing glucose as the main energy source. In addition, about 15 amino acids were added to fulfill the requirement of cells for protein synthesis.

Glucose was converted to lactate through the glycolytic pathway, and thence broken down to carbon dioxide and water in the Krebs cycle. High energy phosphates in the form of ATP were generated by the removal of electrons, and their tunneling through the electron transport system. Amino acids could also be interconverted into fats and carbohydrates, and subsequently used to generate additional energy by entry into the Krebs cycle.

The breakdown of amino acids into two carbon fragments, which is required for introduction into the Krebs cycle, resulted in the formation of ammonium ion. Lactate and ammonium ion were the major cellular waste products, whose accumulation caused feedback inhibition of cellular metabolic processes.

3 Model development, assumptions and sources of process-model mismatch

The detailed phenomenological model can be found elsewhere [3]. Basically, the model comprised the overall mass balance as well as balances on individual constituents like viable and dead cells, the substrates, glucose and amino acid (chiefly glutamine), dissolved oxygen, lactate (the inhibitor) and monoclonal antibodies (product). The process simulator (henceforth referred to as the process) had almost the same form as the model. The Process-model mismatch introduced can be classified into three categories, viz., functional mismatch, differences in values of parameters, and measurement

errors. It should be noted that the measurement error considered here was purely random error. Effects of any outliers or gross error were ignored.

Two case studies were formulated to investigate process-model mismatch due to errors in estimating parameters. The first study featured an erroneously low estimate of $k_{d\max}$ (specific death rate of cells) while the second study featured an erroneously low estimate of $k_{i\max}$ (specific rate of inhibitor formation). The values assumed by both the parameters, in the model and the process, are presented in Table I. The values assumed by all other parameters can be found elsewhere [3]. The model and process were formulated in such a way that the degree of process-model mismatch would be realistic by engineering standards.

4 The Heuristic Random Optimizer (HRO)

The HRO is a powerful optimization routine that has been demonstrated [5] to be superior or equivalent to a variety of optimization algorithms including Broyden-Fletcher-Shanno, Fletcher-Reeves, Cauchy, gradient descent, etc. It has the advantages of constraint handling and scale independent stopping criteria. Hence the HRO was chosen as both the off-line optimization algorithm, and a comparative non-neural network based optimization scheme to benchmark the performance of HDP.

5 Off-line Optimization

The generic approach used, for off-line optimization, was to determine the values of the following variables, so as to maximize the average production rate per batch.

- a) S_0 , the concentration of glucose in the continuous feed to the process as well as in the process at the start of fermentation,
- b) A_0 , the concentration of amino acid in the continuous feed to the process as well as in the process at the start of fermentation,
- c) V_0 , the volume of the reactor contents at the start of fermentation,
- d) $q_0(1)$, the feed rate to the reactor in the first reaction stage where there is a net increase in the population of cells with time,
- e) $q_0(2)$, the feed rate to the reactor in the second reaction stage where there is a net decrease in the population of cells with time,
- e) X_{v0} , the initial inoculum of viable cells,
- g) C_{Lo} , the concentration of dissolved oxygen at the start of fermentation.

The batch time was determined as the time when the process hit the volume constraint (5 liters in this case) or when the average production rate dropped, whichever came earlier. The latter concept is applicable here since it has been observed [3] that the average production rate is a unimodal function of the operating time of fermentation. The constraints, under which the optimization was performed [3], were based on solubility and process design considerations. The best off-line optimization results,

obtained from multiple random starts, are given in Table II.

6 Development of Heuristic Dynamic Programming

6.1 Training of Critic

The critic was a 9-10-1 self-organizing feedforward network, trained to estimate the Bellman Cost Function [11] associated with each system state. There was no one-step penalty imposed on any state, since a reference state was unknown. In other words, the critic was trained, using error backpropagation [12], to minimize the following error for all states.

$$e = \gamma J(t+1) - J(t) \quad (2)$$

The inputs to the network were the system state (eight inputs that comprised the volume of reactor contents and concentrations of 7 state variables) and the remaining time of operation (9th input). The discount factor, γ , was assigned a value of 0.5.

For this study, the Bellman Cost function (also the objective function) was the negative of the average production rate, per batch, of monoclonal antibodies.

6.2 Training of Action

The action network was a 9-5-1 feedforward network that was trained, using the Node Decoupled Extended Kalman Filter [13], to predict the feed rate to the reactor that would minimize the cost function predicted by the critic network. In other words, the error, which the action network was trained to minimize, was the gradient of the cost function relative to the control action given by the action network.

Eight of the nine inputs to the action network were the system state, while the ninth input was the sign of the

quantity, $\frac{d(vX_v)}{dt}$, i.e. sign of the rate of change of total

viable cell mass with time. This was included to ensure that comparisons of performance with the HRO (which utilized the above information while arriving at the feed rate) were meaningful.

The detailed methods of training are not being presented here. However, it should be noted that both the critic and action networks were trained as per the general techniques developed by Prokhorov and Wunsch [6].

7 Model Re-parametrization: The IMPOL Technique

During process operation, the true process parameters drift as per underlying relationships not exactly known to the engineer. Hence, dynamically, there is a need to adjust model parameters to ensure compliance with the process behavior. The IMPOL technique [7] is a one-step application of Newton's method, per control interval, to update a model parameter using the actual process-model mismatch (PMM) and the model sensitivity to the

parameter. For a dynamic process, the process-model mismatch is defined as

$$PMM = y(t) - y_m(t) \quad (4)$$

where $y(t)$ and $y_m(t)$ refer to measured and model predicted values of the state variable being considered. If the mismatch is to be eliminated by adjusting the value of a particular model parameter ϕ , then a one-step application of Newton's method would yield

$$\phi(t) = \phi(t - \Delta t) - \frac{PMM}{\left(\frac{\partial PMM}{\partial \phi}\right)_t} \quad (5)$$

where Δt is the update time interval. In order to eliminate overestimation of the parameter ϕ , and to avoid contamination effects of noise, a relaxation coefficient α , of the order of 0.1, is multiplied with the second term of (5). The resulting equation is

$$\phi(t) = \phi(t - \Delta t) - \alpha \frac{PMM}{\left(\frac{\partial PMM}{\partial \phi}\right)_t} \quad (6)$$

The use of (6) is deemed sufficient for model adjustment insofar as control relevant issues are concerned. While there is no a priori method to ascertain convergence, the adjustment of the model, at every sampling, in a one-step mode should suffice in keeping process-model mismatch to a desirably low value.

For this particular study, the parameter, π_{\max} , denoting the maximum value of the specific product synthesis rate, was adjusted using Equation (6). Evaluation of the gradient in Equation (6) was performed numerically.

It should be noted that, in this study, model re-parametrization and model parameter adjustment mean the same, and are being used interchangeably.

8 Dynamic Model Re-parametrization and On-line Re-optimization using HRO and HDP

The sequential strategy, used for on-line re-optimization, is as follows

- a) The product concentration in the process was measured (Noise was incorporated in the measurement).
- b) The extent of process-model mismatch, PMM, was estimated using (4).
- c) The process-model mismatch was eliminated using the IMPOL technique. The parameter π_{\max} , representing the maximum value of the specific product synthesis rate, was selected for adjustment, since it was directly involved in the rate of product formation.

- d) Once model adjustment was performed, both HRO and HDP were utilized for on-line re-optimization. Both were utilized to determine only the feed rate to the reactor. The remaining time of operation was determined as described previously, i.e., to ensure that the system doesn't hit the volume constraint while maintaining the highest possible average rate of production of the desired product. While using HDP for on-line re-optimization, there was no on-line retraining of either the action and critic networks. Any changes in the model were reflected solely in the system state, that acted as an input vector to the networks. It should be noted that the system state was that predicted by the model and not obtained from the process since quantities from a differentiable model are needed for HDP critic and action network training.

9 Comparison of Results using HRO and HDP

The comparison of actual measured product concentration profiles along off-line optimal (using HRO) and on-line optimal (using both HRO and HDP) trajectories is depicted in Fig. 1 for Case (1). Fig. 2 depicts the annual product yields for Case (1). It is clearly seen that HDP outperformed both off-line and on-line HRO insofar as average production rate was concerned. Specifically, for Case (1), the average off-line optimal production rate was 64.5 g/annum per batch. On-line re-optimization, using the HRO, resulted in an average production rate of 67.8 g/annum per batch. The use of HDP, for on-line optimization, resulted in an average production rate of 89.1 g/annum per batch. For Case (2), the corresponding figures were 68.47 g/annum per batch and 78.4 g/annum per batch respectively, along off-line and on-line optimal operations using the HRO, and 85.0 g/annum per batch along on-line optimal operation using HDP.

If the market demand for monoclonal antibodies is considered to be 5 kg/annum of recovered product, as is often the case [14]-[16], a detailed economic analysis for Case (1) indicated that the use of HDP resulted in an increase in the annual net profit by \$ 9.3 million and \$ 8.2 million respectively, over off-line and on-line optimal operations using the HRO. For Case (2), the corresponding figures were \$ 6.03 million and \$ 2.33 million respectively.

In addition to improved productivity and better economics, the use of Adaptive Critic Designs offers significant advantages over traditional direct search optimization routines like the HRO. These are

- a) Adaptive Critic Designs facilitate easy constraint handling via penalty functions and bounded activation functions in Neural Networks.
- b) Neural networks compute rapidly, thereby facilitating a much reduced control interval relative to optimizers like HRO. This advantage of reduction in control interval would be highly significant in processes with fast dynamics like chemical reactions (as opposed to

biochemical reactions). Another area where this advantage would be clearly observed is massive systems like refineries, where optimization involves determination of several decision variables, and computational time is an important aspect of process economics.

- c) With traditional optimization routines, improvements in the model are translated into improved optimal operation only by dynamic re-optimization. However, with Adaptive Critic Designs, even no on-line retraining results in significant improvements as opposed to both off-line and on-line optimal operation using conventional optimizers like HRO. This is due partly to the fact that the system state (that reflects changes in the model) is explicitly used while computing the control action, and also due to the fact that Adaptive Critic Designs do not, in general, require a perfect model for true optimal process performance [17].

10 Conclusions

This study demonstrates the applicability and economic potential of a simple scheme for off-line optimization and on-line model parameter adjustment and re-optimization using Heuristic Dynamic Programming. In general, Heuristic Dynamic Programming is robust towards model uncertainties, and tracks the global optimum closely. Besides, the significant economic benefits and increased computational power, obtained by the use of HDP, is a pointer to possible avenues in exhaustive application of Adaptive Critic Designs in the field of bioreactor control.

Acknowledgments

The authors would like to thank Prof. R. Russell Rhinehart, Head of the School of Chemical Engineering, at Oklahoma State University, for his guidance in the initial stages of the study. In addition, the HRO developed by him with co-workers has been of great utility. Many thanks are due to Prof. Ted Wiesner, from the Department of Chemical Engineering at Texas Tech University, for his immense help in developing the model.

References

- [1] W. Bamberger, and R. Iserman, "Adaptive On-Line Steady-State Optimization of Slow Dynamic Processes," *Automatica*, vol. 14, pp. 223-230, 1978.
- [2] I. W. D. Hackh, *Grant and Hackh's Chemical Dictionary*. New York: McGraw-Hill, 1987.
- [3] M. S. Iyer, T. F. Wiesner, and R. R. Rhinehart, "Dynamic Re-optimization of a Fed-Batch Fermentor," Accepted by *Biotechnology and Bioengineering*, 1998.
- [4] S. Dhir, K. J. Morrow, R. R. Rhinehart, and T. F. Wiesner, "On-line Dynamic Optimization of a Fed-Batch Fermentor," Submitted to *Biotechnology and Bioengineering*, 1998.
- [5] J. Li, R. R. Rhinehart, "Heuristic Random Optimization," *Computers and Chemical Engineering*, vol. 22, no. 3, pp. 427-444, 1998.
- [6] D. V. Prokhorov, D. C. Wunsch, "Adaptive Critic Designs," *IEEE Transactions on Neural Networks*, vol. 8, no. 5, pp. 997-1007, 1997.
- [7] R. R. Rhinehart, and J. B. Riggs, "Two Simple Methods for On-line Incremental Model Parametrization," *Computers and Chemical Engineering*, vol. 15, no. 3, pp. 181-189, 1991.
- [8] B. C. Batt, and D. S. Kompala, "A Structured Kinetic Modeling Framework for the Dynamics of Hybridoma Growth and Monoclonal Antibody Production in Continuous Suspension Cultures," *Biotechnol. Bioeng.*, vol. 34, pp. 515-531, 1989.
- [9] J. G. Gaertner, and P. Dhurjati, "Fractional Factorial Study of Hybridoma Behavior. 1. Kinetics of Growth and Antibody Production," *Biotechnology Progress*, vol. 9, pp. 298-308, 1993.
- [10] M. de Tremblay, M. Perrier, C. Chavarie, and J. Archambault, "Fed-batch culture of hybridoma cells: Comparison of optimal control approach and closed loop strategies," *Bioprocess Engineering*, vol. 9, pp. 13-21, 1993.
- [11] R. E. Bellman, *Dynamic Programming*. New Jersey: Princeton Univ. Press, 1957.
- [12] Y. Chauvin, and D. E. Rumelhart (Eds.), *Backpropagation : Theory, architectures, and applications*, Hillsdale, N. J., Lawrence Erlbaum Associates, 1995.
- [13] L. Singhal, and L. Wu, "Training Multilayer Perceptrons with the Extended Kalman Algorithm," in *Advances in Neural Information Processing Systems 1*, Denver 1988, D. S. Touretzky, Ed. San Mateo, CA; Morgan Kaufmann, 1989, pp. 133-140.
- [14] B. Atkinson, and F. Mavituna, *Biochemical Engineering and Biotechnology Handbook. Chapter 8*. Stockton Press (MacMillan Publishers), 1991.
- [15] *Monoclonal Antibody Purification Handbook (including MAbAssistant)*. 1993. Pharmacia.

[16] G. L. Smith, "Large Scale Animal Cell Culture," in *Bioprocess Engineering*, B.K. Lydersen, N.A. D'Elia, and K.L. Nelson, Eds., Wiley Interscience, 1994.

[17] E. W. Saad, Personal Communications, *Department of Electrical Engineering, Texas Tech University*, 1998.

Tables

Table I
Delineation of Cases (1) and (2)

Parameter	Case (1) Erroneously low $k_{d\max}$		Case (2) Erroneously low $k_{i\max}$	
	Value used in Model	Value used in Process	Value used in Model	Value used in Process
$k_{d\max}$	0.08 g dead cells/ g viable cells/hr	0.16 g dead cells/ g viable cells/hr	0.08 g dead cells/ g viable cells/hr	0.0786 g dead cells/ g viable cells/hr
$k_{i\max}$	0.1675 g inhibitor/ g viable cells/hr	0.1638 g inhibitor/ g viable cells/hr	0.1675 g inhibitor/ g viable cells/hr	0.3348 g inhibitor/ g viable cells/hr

Table II
Values of Decision Variables obtained by Off-line Optimization

Decision Variable	Optimal Value
S_0	98.9 g/l
A_0	11.4 g/l
V_0	4.64 l
$q_0(1)$	14.4 ml/day
$q_0(2)$	82.2 ml/day
T_b	12 days, 13 hr and 20 minutes
X_{v0}	30 mg/l
C_{L0}	29 mg/l

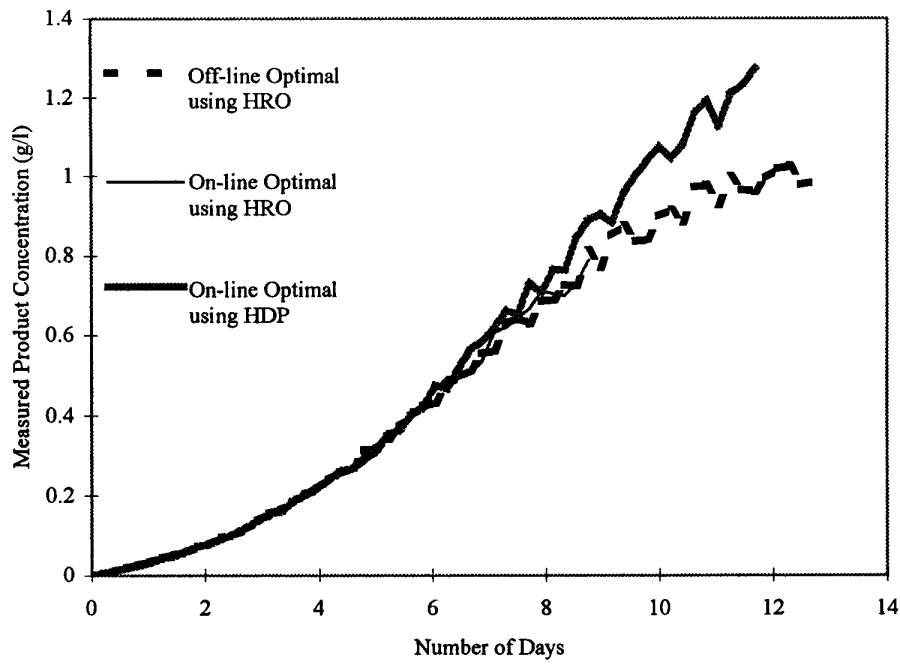


Fig. 1. Comparison of Product Concentration Profiles for Case (1) along various Optimal Recipe Schedules.

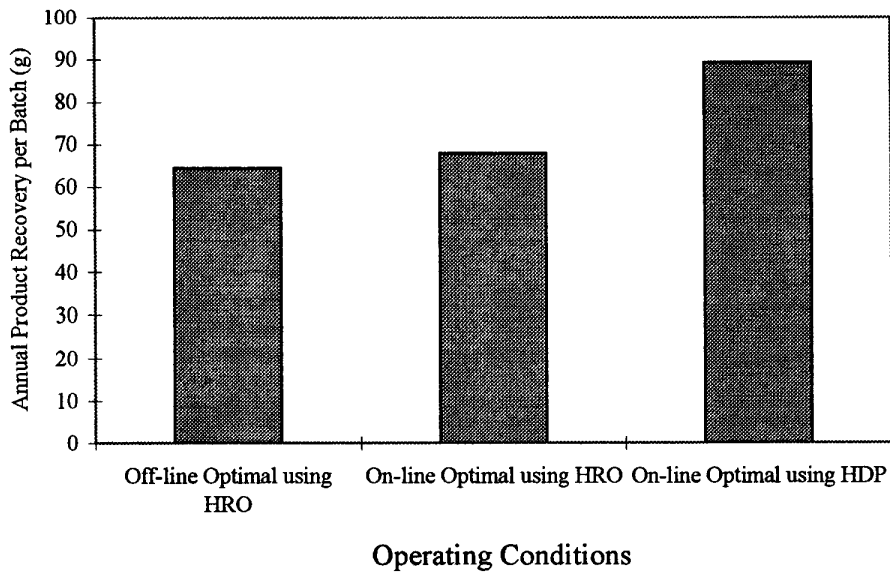


Fig. 2. Comparison of Annual Product Recovery per Batch for Case (1) along various Optimal Operating Schedules