Dynamic Reoptimization of a Fed-Batch Fermentor Using Adaptive Critic Designs

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

***Abstract*—Traditionally, fed-batch biochemical process optimization and control uses complicated off-line optimizers, with no on-line model adaptation or reoptimization. This study demonstrates the applicability of a class of adaptive critic designs for on-line reoptimization and control of an aerobic fed-batch fermentor. Specifically, the performance of an entire class of adaptive critic designs, viz., heuristic dynamic programming (HDP), dual heuristic programming (DHP) and generalized dual heuristic programming (GDHP), was demonstrated to be superior to that of a heuristic random optimizer (HRO), on optimization of a fed-batch fermentor operation producing monoclonal antibodies.**

***Index Terms*—Dual heuristic programming (DHP), fed-batch fermentor, generalized dual heuristic programming (GDHP), heuristic dynamic programming (HDP), heuristic random optimizer (HRO), monoclonal antibodies.**

NOMENCLATURE

c_{A}	Concentration of amino acid in the reactor (mass/volume) at any time t.
$c_{\mathrm{A}}(0)$	Concentration of amino acid in the feed to the reactor as well as the concentration of amino acid in the initial charge (mass/volume).
c_{O}	Concentration of dissolved oxygen in the reactor broth (mass/volume) at any time t.
$c_{\mathrm{O}}(0)$	Concentration of dissolved oxygen in the reactor broth (mass/volume) at time $t = 0$.
$c_{\mathrm{O}}{}^{*}$	Equilibrium solubility of oxygen in the broth (mass/volume).
c_{L}	Concentration of the lactate in the reactor (mass/volume) at any time t.
K_{A}	Concentration of amino acid at which the specific product synthesis rate is half its maximum value (mass/volume) (A Michaelis–Menten constant).
K_d	Concentration of inhibitor at which the specific death rate of cells is half its maximum value (mass/volume) (A Michaelis–Menten constant).
K_I	Concentration of glucose at which the specific rate of inhibitor formation is half its maximum value (mass/volume) (A Michaelis–Menten constant).
K_{O2}	Concentration of dissolved oxygen at which the specific growth rate of cells has a factor of half contributed by the presence of dissolved oxygen (mass/volume) (A Michaelis–Menten constant).
K_{S}	Concentration of glucose at which the specific growth rate of cells has a factor of half contributed by the glucose(mass/volume) (A Michaelis–Menten constant).
$k_{d\max}$	Maximum value of the specific death rate of cells (mass of dead cells/mass of viable cells/time).
$k_{I\max}$	Maximum value of the specific rate of inhibitor formation (mass of inhibitor/mass of viable cells/time).
$k_{\mathrm{L}}a$	Volumetric liquid side mass transfer coefficient for oxygen (time^{-1}).
m	Factor that accounts for the growth yield factor not being equal to its maximum value (mass of glucose consumed per unit mass of viable cells per unit time).
c_P	Concentration of the product in the reactor (mass/volume) at any time t.
q_0	Volumetric flow rate of the feed into the reactor (volume/time).
c_{S}	Concentration of glucose in the reactor (mass/volume) at any time t.
$c_{\mathrm{S}}(0)$	Concentration of glucose in the feed to the reactor as well as the concentration of the glucose in the initial charge (mass/volume).
T_{b}	Batch time.
t	Time.
V	Volume of the contents in the reactor at any time.
$V(0)$	Volume of the initial charge in the reactor.
c_{Xd}	Concentration of dead cells in the reactor (mass/volume).
c_{Xv}	Concentration of viable cells in the reactor (mass/volume) at any time t.
$c_{\mathrm{Xv}}(0)$	Concentration of viable cells in the reactor (mass/volume) at time $t = 0$.
$Y^{*}_{\mathrm{Xv/S}}$	Maximum value of the growth yield factor (mass of viable cells/mass of glucose consumed).
$Y_{\mathrm{xv/A}}$	Yield of cell mass per unit mass of amino acid uptake (mass of cells formed/mass of amino acid consumed).
$Y_{\mathrm{xv/O2}}$	Yield of cell mass per unit mass of oxygen uptake (mass of cells formed/mass of oxygen consumed).
$Y_{\mathrm{xv/S}}$	Yield of cell mass per unit mass of substrate intake (mass of cells formed/mass of glucose consumed).
$\mu_{\max}$	Maximum specific rate of growth of cells (time^{-1}).
$\pi_{\max}$	Maximum value of specific product synthesis rate (mass of product/mass of viable cells/time).

Manuscript received March 9, 2000; revised May 30, 2001. This work was supported by the National Science Foundation under Grant 0096118.

M. S. Iyer is with Dresser-Rand Control Systems, Houston, TX 77043 USA.

D. C. Wunsch, II is with the Applied Computational Intelligence Laboratory, Department of Electrical and Computer Engineering, University of Missouri-Rolla, Rolla, MO 65409-0040 USA.

Publisher Item Identifier S 1045-9227(01)09518-2.

I. Introduction

BIOCHEMICAL processes, producing high value end products like vitamins, baker's yeast, and antibiotics [1], [2], are statistically nonstationary and, therefore, need continually adapting recipes for optimal performance. Fed-batch fermentations have been widely investigated for both optimization and control. Most of the optimization studies on fermentations have reported the use of detailed theoretical techniques [3]–[7], while very few [8], [9], have reported the use of direct search techniques. Advanced control of biochemical processes has also been simulated. These entail the use of perfect models of the process. Of current interest is the use of hybrid or single models involving neural networks, fuzzy logic, expert systems, and first principles [3], [4], [6], [7]. In fed-batch fermentations, the most important aspects to be considered are the changes in process dynamics during the operation of the batch and the consequent need to adjust the process model. Once the model is altered, there is also a need to reoptimize the process using the now available improved model. One published approach used a heuristic random optimizer (HRO) [9]–[12], for both off-line and on-line optimization.

We compare a variety of control schemes including off-line optimization, on-line model reparameterization and on-line reoptimization of a fed-batch fermentor using an entire class of adaptive critic designs (ACDs) [13]. 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 [14] was used for dynamic model parameter adjustment. Model dependent designs of heuristic dynamic programming (HDP), dual heuristic programming (DHP) and generalized dual heuristic programming (GDHP) [13] performed on-line reoptimization. The process performance obtained using each ACD was compared with that obtained using the HRO for both just off-line optimization, as well as combined off-line optimization and dynamic on-line reoptimization. Although the study was conducted for a specific case of cultivation of mammalian hybridoma cells to produce monoclonal antibodies [15]–[17], the development of all technologies used here is perfectly general and is easily applicable to any batch process that can be modeled.

II. The Biochemical Growth System

The system studied for optimization and control was the in vitro growth of hybridoma cells and the production of monclonal 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 fulfil the cells' requirement 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 [18]. 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 ions. Lactate and ammonium ions were the major cellular waste products, whose accumulation caused feedback inhibition of cellular metabolic processes.

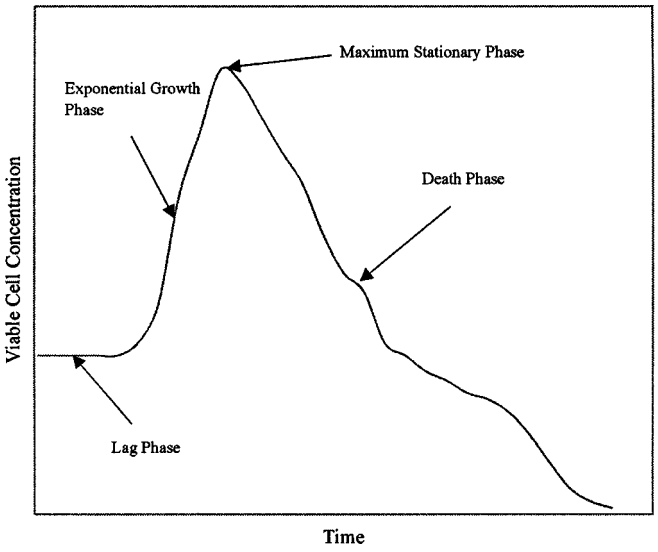

Fig. 1. Typical cell growth curve.

Fig. 1 shows the typical profile of live cell concentration as a function of fermentation batch time. The various phases of cell growth that are depicted in Fig. 1 are the lag phase, the exponential phase and the death phase. The lag phase is characterized by an insignificant increase in cell population. The exponential growth phase is characterized by a high rate of increase in cell population. The maximum stationary phase is the stage where the cell population has reached its maximum size and the death phase is the stage where the rate of death of cells far exceeds their rate of growth, resulting, thereby, in a significant reduction in cell population.

III. Model Development, Assumptions, and Sources of Process-Model Mismatch

Before delving into the details of model development, the concepts of a *model* and a *process simulator* must be elucidated. The process simulator refers to the simulation of an actual operating fermentor while the model refers to the mathematical description of the fermentor that a process engineer would formulate. For this study, both the model and the simulator were computer generated. However, for an actual operating fermentor, only the model needs to be developed, while the operating process would replace the simulator. In what follows, the term *process* will be used in lieu of *process simulator*.

The detailed phenomenological model, used in this study, is presented in Appendix A. Other results using a similar model are presented in [9]. The process had almost the same form as the model. However, differences were incorporated between the model and the process to account for the fact that a process engineer would never be able to perfectly model the process. The differences incorporated can be broadly classified into three categories: functional mismatch, differences in values of parameters

TABLE I
VALUES OF PARAMETERS USED IN THE MODEL AND PROCESS

Parameter	Value used in model	Value used in process
c_O^*	0.034 g/l	0.0352 g/l
K_A	0.009 g A/l	0.00836 g A/l
K_d	28 g I/l	28.22 g I/l
K_I	0.11 g I/l	0.1188 g I/l
k_La	72 h^{-1}	69.48 h^{-1}
K_{O2}	6.4 μg O_2/l	6.34 μg O_2/l
K_S	0.2 g/l	0.24 g/l
m	0.002	0.0021
$Y_{xv/A}$	1 g viable cells/g amino acid	0.832 g viable cells/g amino acid
$Y^*_{Xv/S}$	0.6 g viable cells/g glucose	0.44 g viable cells/g glucose
$Y_{xv/O2}$	1.0 g viable cells/g O_2	0.804 g viable cells/g O_2
μ_{max}	0.017 h^{-1}	0.01536 h^{-1}
π_{max}	0.0365 g P/g viable cells/h	0.0369 g P/g viable cells/h

TABLE II
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/h	0.16 g dead cells/ g viable cells/h	0.08 g dead cells/ g viable cells/h	0.0786 g dead cells/ g viable cells/h
$k_{I\,max}$	0.1675 g inhibitor/ g viable cells/h	0.1638 g inhibitor/ g viable cells/h	0.1675 g inhibitor/ g viable cells/h	0.3348 g inhibitor/ g viable cells/h

and measurement errors like noise and bias. Table I lists values of parameters that were used in the process and the model [16]. The value of $C_O{}^*$ was set to be the solubility of pure oxygen in water at 35°C [10], which is the typical culture condition [11]. It should be noted, that different values of $C_O{}^*$ have been quoted in literature.

Two case studies were formulated to investigate process-model mismatch due to errors in estimating parameters. Case 1 featured an erroneously low estimate of $k_{d\max}$ (specific death rate of cells), while Case 2 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 II.

IV. THE HRO

The HRO [12] uses three procedures in computing the optimum value of an objective function. The first procedure involves the search for a feasible starting point for the optimization to commence. This procedure is invoked only if the starting point provided to the optimizer is infeasible and continues until a feasible starting point has been found. The second procedure is a gross search, where the optimizer yields a local vicinity of the global optimum. The final procedure is a fine search, which hones in on the global optimum. The HRO has the advantages of constraint handling and scale independent stopping criteria.

The HRO has the following parameters that need to be selected.

A. Mean and Standard Deviation for the Search

The mean for all variables was selected to be zero.

To start with, the standard deviation was selected for each variable as a hundredth of the feasible range for the variable. This was further modified by offline simulation so as to attain the best optimizer performance.

B. Factor to Increase or Decrease the Mean Depending on the Outcome of a Search

These factors were determined by trial and error so as to attain the best optimizer performance during offline simulation.

C. Factor to Increase or Decrease the Standard Deviation Depending on the Outcome of a Search

These factors were determined again by trial and error so as to ensure best performance of the optimizer for offline simulation.

The HRO has been compared with a lot of different deterministic optimization routines like Broyden–Fletcher–Shanno (BFS), Fletcher–Reeves (FR), Davidson–Fletcher–Powell (DFP), Cauchy, and other basic random search techniques [11] on the Eason and Fenton function. The HRO has been found to be superior to the above mentioned techniques in terms of number of iterations taken to find the same optimum point.

Besides, the HRO was also found to be superior to the deterministic Levenberg–Marquardt method, the BFS method, the error backpropagation method, and the Nelder–Mead method for neural network training [11], both in terms of iter-

ation count as well as training error. Since the HRO has been demonstrated to be competitive with deterministic optimization schemes, there was no effort to compare it with sequential quadratic programming (SQP), which is a standard algorithm for nonlinear programming.

The comparison between ACD and HRO is, admittedly, not an exhaustive study. However, given the superiority of the HRO to several optimization schemes, it was chosen as a reasonable tool to benchmark the performance of the ACDs.

V. Off-Line Optimization

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

- $c_S(0)$ concentration of glucose in the continuous feed to the process, as well as in the process at the start of fermentation;
- $c_A(0)$ concentration of amino acid in the continuous feed to the process, as well as in the process at the start of fermentation;
- $V(0)$ volume of the reactor contents at the start of fermentation;
- $q_0(1)$ feed rate to the reactor in the first reaction stage where there is a net increase in the population of cells with time;
- $q_0(2)$ feed rate to the reactor in the second reaction stage where there is a net decrease in the population of cells with time;
- $c_{Xv}(0)$ initial inoculum of viable cells;
- T_b operating batch time;
- $c_O(0)$ concentration of dissolved oxygen at the start of fermentation C_{L0}.

As reported above, the feed rate to the reactor, q_0, was not considered to be a unique value. Instead, two stages of optimization were considered for q_0, where q_0 would be unchanged within a stage but would change from one stage to the other. The first stage corresponded to the phases of the growth-cycle that resulted in a net increase in total cell population with time {a positive value of $d(Vc_{X_v})/dt$}, while the second stage corresponded to the phases of the growth-cycle that resulted in a net decrease in total cell population with time {a negative value of $d(Vc_{X_v})/dt$}.

The concentration of dissolved oxygen at the start of fermentation was considered to be the decision variable, in lieu of the mixer speed or air flow rate, since either of the last two would effectively result in a certain dissolved concentration of oxygen. For a more thorough analysis and to allow the possibility of varying the mixer speed on-line, it is recommended that future studies use the mixer speed or air flow rate as both an off-line and on-line optimal decision variable.

Finally, instead of searching explicitly for the batch time, the same was determined as the time when the process hit the volume constraint (5 L in this case) or when the average production rate dropped, whichever happens earlier. The latter concept is applicable here since it has been observed [9] that the average production rate of antibodies is a unimodal function of the operating fermentation time.

TABLE III
Values of Decision Variables Obtained by Off-Line Optimization

Decision Variable	Optimal Value
$c_S(0)$	98.9 g/l
$c_A(0)$	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 h and 20 minutes
$c_{Xv}(0)$	30 mg/l
$c_O(0)$	29 mg/l

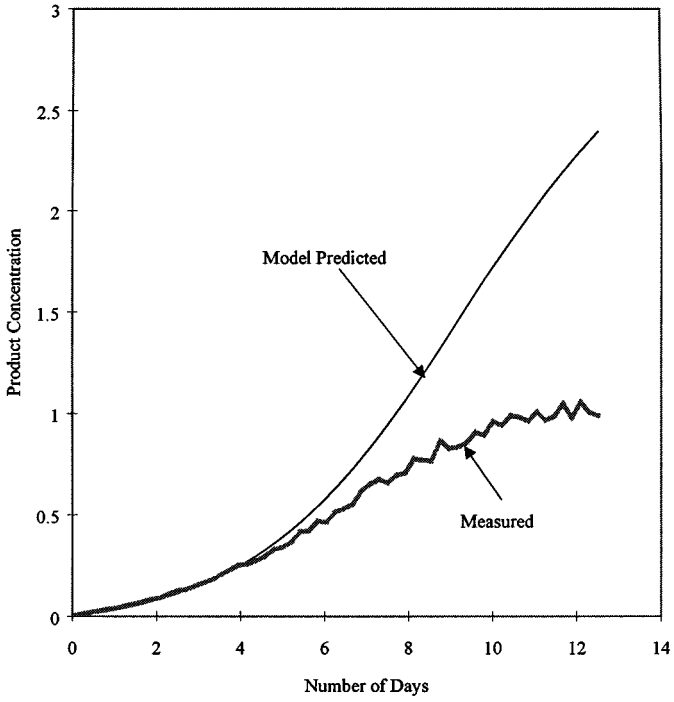

Fig. 2. Comparison of product concentration profiles along the off-line optimal recipe schedule for case (1).

The best off-line optimization results, obtained from multiple random starts, are given in Table III.

The model and the process were both integrated along the *recipe schedule* (the sequence of staged changes in the feed rate) specified by the off-line optimizer. Fig. 2 shows a comparison of trajectories for the product concentration, as predicted by the model and as measured in the process, under off-line optimal operating conditions (case 1 only). It is seen that the model predicted a much higher value of the final product concentration as compared to what would be attainable in the process. This was true for case (2) also (simulation not shown). This difference is a result of process-model mismatch.

VI. Development of ACDs

The application of the entire class of ACD [12], [22] was based on the block diagram in Fig. 3. The action network had, as inputs, the system state at any time t. Output from the action network, which constituted the feed rate of nutrients to the reactor, at the current system state, served as inputs to the process

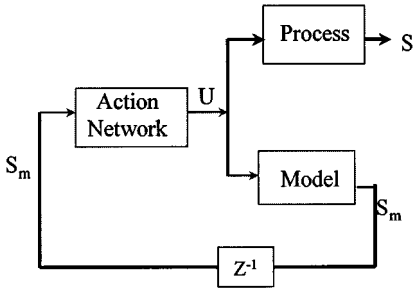

Fig. 3. Implementation details of action network.

and the model. The model predicted state at the next sampling was the next input to the action network and progressively this procedure was continued till the end of the batch. The end of the batch operation was determined by either a decrease in average production rate, or a measurement of process volume greater than 5 L, whichever occurred earlier.

It is important to note that the critic network is not depicted in Fig. 3, thus implying its absence in the implementation of ACD. However, the critic network was a very essential element involved in the training of the action network, as will be evident later. Outlined below is the procedure used to obtain the trained action and critic networks for each of the model dependent adaptive critic designs that was investigated.

VII. HDP

A. Training of Critic

The critic was a 9-10-1 self-organizing feedforward network, trained to estimate the Bellman cost function [19], $J(t)$, associated with each system state. There was no one-step penalty imposed on any state since a reference state was not known accurately. In other words, the critic was trained, using error backpropagation [20], to minimize the following error for all states.

$$e = \gamma J(t+1) - J(t) \tag{1}$$

Here $J(t+1)$ refers to the Bellman cost function associated with the system state at the next sampling. The inputs to the network were the system state (eight inputs that comprised the volume of reactor contents and concentrations of seven state variables) and the remaining time of operation (ninth input). The remaining time of operation was used an input to the critic since it is a variable factor that directly influences the Bellman cost of operating along a particular trajectory. To cite an example, if a particular trajectory of optimal decision variables is adapted for a bioreactor operation, the time for which that particular trajectory is used decides the cost of operating along that certain trajectory. The discount factor, γ, was determined by trial and error to have a value of 0.5.

B. Training of Action

The action network was a 9-5-1 feedforward network that was trained, using the node decoupled extended Kalman filter [21], 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, i.e., $\partial J(t+1)/\partial A(t)$. Here, $A(t)$ is the vector of manipulated variable moves predicted by the action network. The gradient, was obtained using (2) below.

$$\frac{\partial J(t+1)}{\partial A(t)} = \left[\frac{\partial J(t+1)}{\partial R(t+1)}\right]\left[\frac{\partial R(t+1)}{\partial A(t)}\right] \tag{2}$$

Here $\partial J(t+1)/\partial R(t+1)$ was obtained using the architecture of the critic network, corresponding to the predicted system state at the next sampling as the inputs, while $\partial R(t+1)/\partial A(t)$ was obtained using the architecture of 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, $d(Vc_{X_V})/dt$, 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.

VIII. DHP

A. Training of Critic

The critic was a 9-10-9 self-organizing feedforward network, trained to estimate the gradient of the Bellman cost function with respect to the system state. Again, there was no one-step penalty imposed on any state. The following error was minimized for all states:

$$e = \gamma\frac{\partial J(t+1)}{\partial R(t)} - \frac{\partial J(t)}{\partial R(t)} \tag{3}$$

The value of γ was again 0.5. The gradient, $\partial J(t+1)/\partial R(t)$, was evaluated as

$$\frac{\partial J(t+1)}{\partial R(t)} = \left[\frac{\partial J(t+1)}{\partial R(t+1)}\right]\left[\frac{dR(t+1)}{dR(t)}\right] \tag{4}$$

Here $\partial J(t+1)/\partial R(t+1)$ was obtained as outputs from the critic network, corresponding to the predicted system state at the next sampling as the network inputs, and $dR(t+1)/dR(t)$ was obtained using the dynamic model and the architecture of the action network as per (5)

$$\frac{dR(t+1)}{dR(t)} = \frac{\partial R(t+1)}{\partial R(t)} + \left[\frac{\partial R(t+1)}{\partial A(t)}\right]\left[\frac{dA(t)}{dR(t)}\right] \tag{5}$$

Values of $\partial J(t)/\partial R(t)$ were simply the outputs of the critic network, corresponding to inputs comprising the current system state.

B. Training of Action

The action network had exactly the same configuration and function as for HDP. There was no difference in the training procedure *vis-a-vis* HDP. The error vector for training was obtained using (2). The first factor on the right-hand side of (2) was obtained as outputs from the critic network, when the predicted system state the next sampling was the network input. The second factor on the right-hand side of (2) was obtained using the model.

IX. GDHP

A. Training of Critic

The critic was a 9-10-10 self-organizing feedforward network, trained to estimate both the Bellman cost function, as well its gradient with respect to the system state. Again, there was no one-step penalty imposed on any state. The error minimized was the sum of the errors minimized for HDP and DHP. The network inputs were the same as those for either of HDP and DHP. The various gradients were evaluated similar to DHP.

B. Training of Action

The action network had exactly the same configuration and function as for either of HDP and DHP. The evaluation of various gradients was similar to DHP, as was the training procedure.

It is worth noting that the critic in each of the three cases, viz. HDP, DHP, and GDHP had different configurations due to the different functions. While the critic estimated only the Bellman cost function in the case of the HDP, it estimated the gradient of the cost function in case of DHP. The functionality of critic in GDHP was, in effect, the summation of its functionality in HDP and DHP.

X. Model Reparameterization: The IMPOL Technique

The IMPOL technique [14] 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.

Consider a dynamic process at an operating point in time, where the state variable y (whose mismatch is to be eliminated) assumes a measured value $y(t)$. Further, let $y_{\mathrm{m}}(t)$ represent the value of the state variable predicted by a dynamic model used to represent the process. Owing to unavoidable sources of PMM described above, $y(t)$ and $y_{\mathrm{m}}(t)$ are never equal. The PMM is defined as

$$\mathrm{PMM} = y(t) - y_{\mathrm{m}}(t) \tag{6}$$

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

$$\phi(t) = \phi(t - \Delta t) - \frac{\mathrm{PMM}}{\left(\frac{\partial \mathrm{PMM}}{\partial \phi}\right)_t} \tag{7}$$

where Δt is the update time interval. Equation (7), being a one-step algorithm, could overestimate changes in the parameter ϕ due to numerical instabilities. Further, noise on measurements could contaminate $\phi(t)$. Hence a relaxation coefficient α, of the order of 0.1, is multiplied with the second term of (7), resulting in (8) below

$$\phi(t) = \phi(t - \Delta t) - \alpha \frac{\mathrm{PMM}}{\left(\frac{\partial \mathrm{PMM}}{\partial \phi}\right)_t} \tag{8}$$

The use of (8) is deemed sufficient for model adjustment insofar as control relevant issues are concerned. The adjustment of the model, at every sampling, in a one-step mode should suffice in keeping PMM to a desirably low value, obviating, thereby, the need for a complete iterative Newton's technique to eliminate the mismatch.

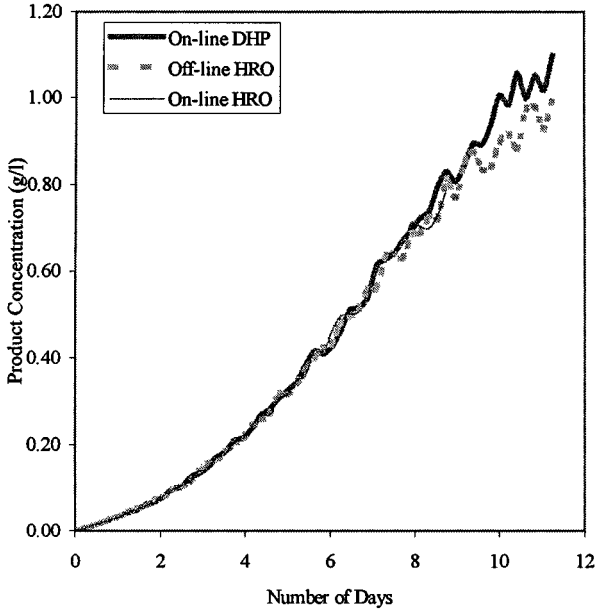

Fig. 4. Comparison of product concentration profiles along the various optimal operating schedules for case (1).

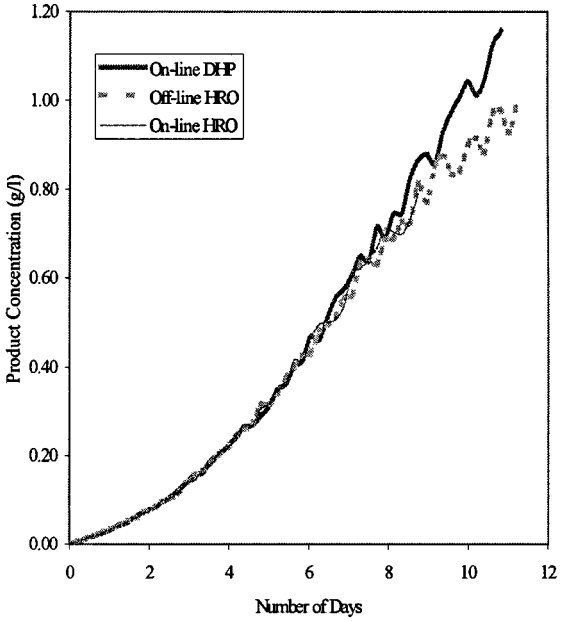

Fig. 5. Comparison of product concentration profiles along the various optimal operating schedules for case (2).

XI. Dynamic Model Reparameterization and On-Line Reoptimization Using HRO and ACD

The sequential strategy used for on-line reoptimization is as follows.

1) The product concentration in the process was measured. (Noise was incorporated in the measurement to account for random measurement errors.)
2) The extent of PMM was estimated using (6).

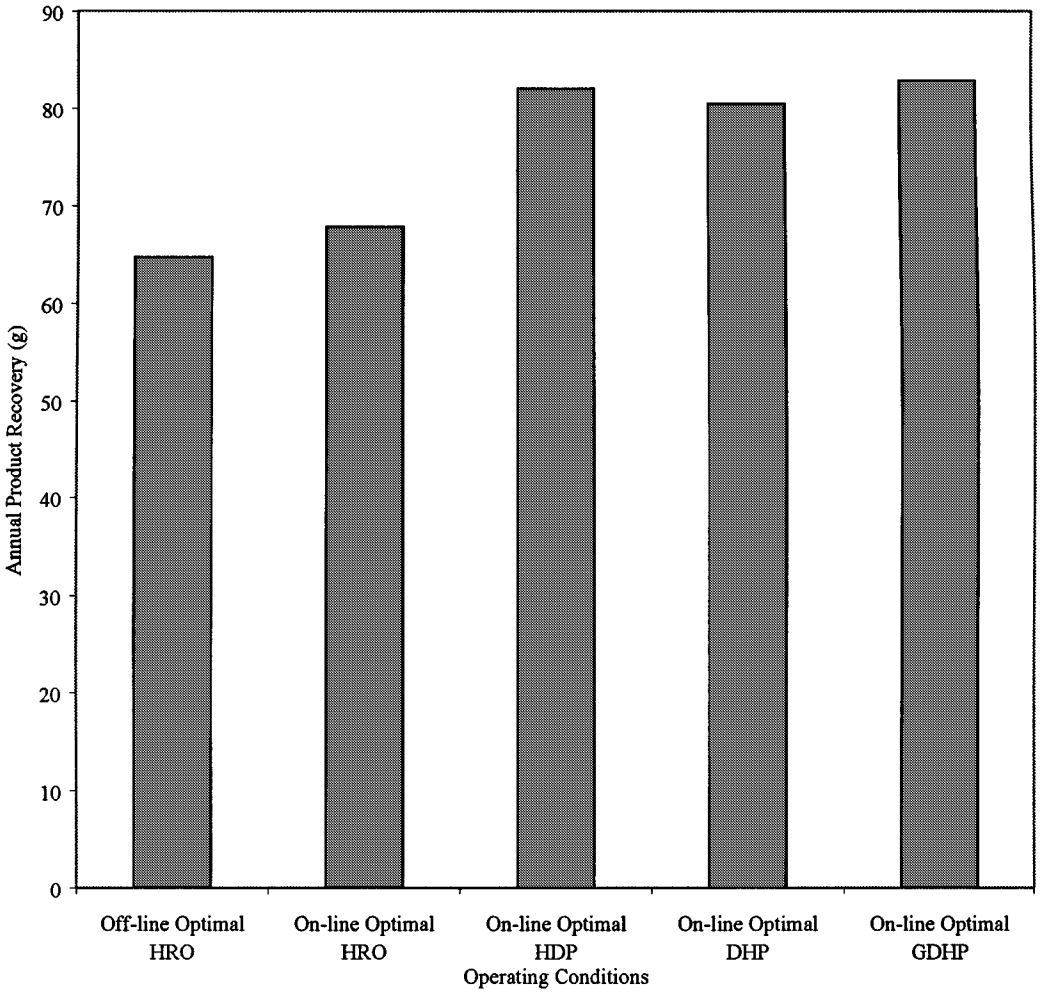

Fig. 6. Comparison of annual recovered production of antibodies per batch along various optimal operating schedules for case (1).

3) The PMM was eliminated using the IMPOL technique. The parameter $\pi_{\max}$, representing the maximum value of the specific product synthesis rate in (A13) was selected for adjustment since it is directly involved in the rate of product formation. [This can be seen by combining (6) and (A13) in Appendix A.] The realization of (8) that was used for adjustment of $\pi_{\max}$ was

$$\pi_{\max}(t) = \pi_{\max}(t - \Delta t) + \frac{\alpha(\text{PMM})}{\Delta t\left(\frac{AX_v}{K_A + A}\right)}. \tag{9}$$

4) Once model adjustment was performed, both HRO and adaptive critic designs were utilized for on-line reoptimization. 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 does not hit the volume constraint, while maintaining the highest possible average production rate of the desired product. While using adaptive critic designs for on-line reoptimization, there was no on-line retraining of either the action and critic networks. Any changes in the model (due to model adjustment to eliminate process-model mismatch) were reflected solely in the system state (that acted as an input vector to the networks). While on-line retraining of the networks is certainly realizable, it was found that the same was not warranted for this system to produce true optimal process performance.

At this juncture, it should be clained again that at no time was the biochemical system modeled using neural networks. The system model was a dynamic model based on first principles, with model adjustment performed using the IMPOL technique. Only static neural networks were employed for ACDs since both the critic and action networks need to just provide an estimate of the cost of operation and manipulate the flows into the reactor to minimize the cost of operation, respectively. Neither of the above functions requires information on process dynamics explicitly. What is required, fundamentally, though is on-line dynamic retraining of the critic and action networks. However, one of the strong points of ACDs is their robustness in the face of model uncertainties, since adaptive critic designs behave as a low-pass filter. Hence, on-line retraining of critic and action networks was eliminated from this study.

XII. Comparison of Results Using HRO and ACD

In order to maintain clarity, the dynamic product concentration profiles are shown for off-line optimal trajectory (using HRO) and on-line optimal trajectories using the HRO and only

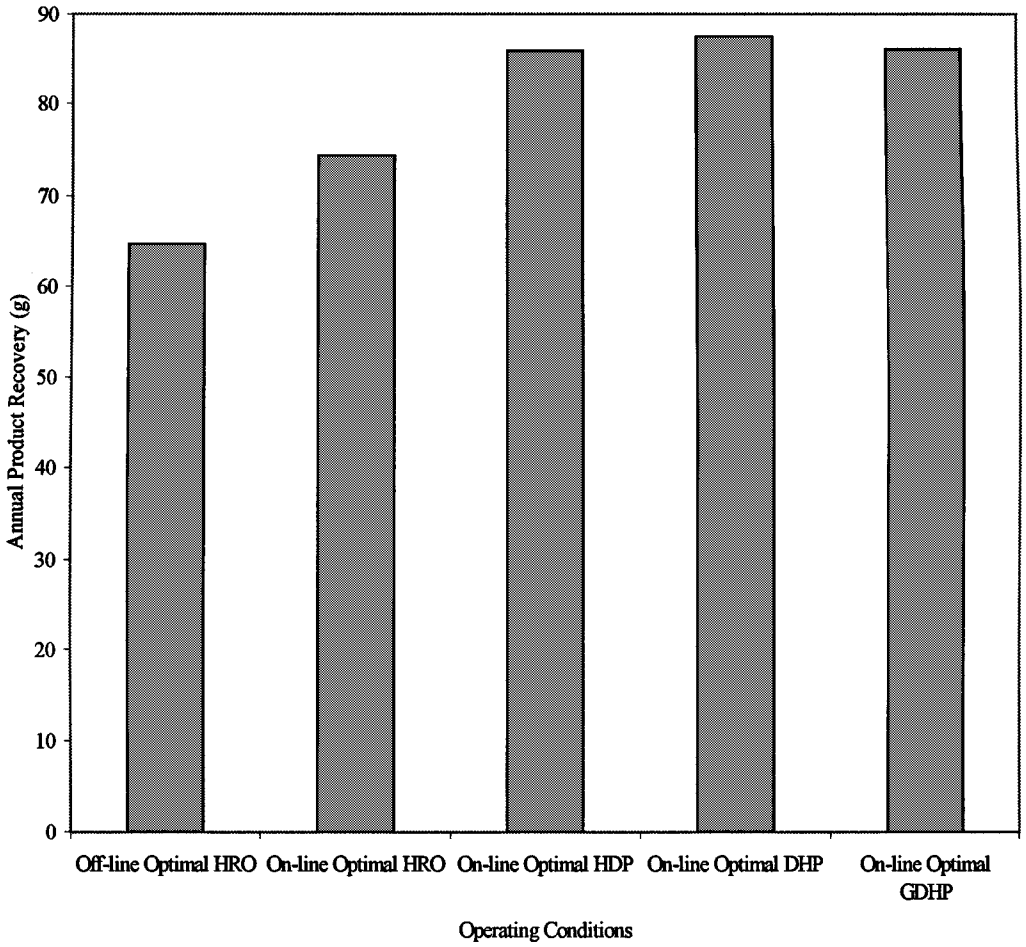

Fig. 7. Comparison of annual recovered production of antibodies per batch along various optimal operating schedules for case (2).

DHP from the class of ACDs, in Figs. 4 and 5, for Cases (1) and (2), respectively. Figs. 6 and 7 depict the annual product yields obtained using all the optimization strategies, for Cases (1) and (2) respectively. It is clearly seen that, in either case, the entire class of adaptive critic designs 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 reoptimization, using the HRO, resulted in an average production rate of 67.8 g/annum per batch. The use of HDP resulted in an average production rate of 82.06 g/annum, per batch while the corresponding figures were 80.5 and 82.9, respectively, for DHP and GDHP. A similar improvement in product yield was observed for Case (2) also, as noted from Fig. 5.

Figs. 8 and 9 are representative plots that depict the variation of the decision variables, viz. the feed rate and operating batch time, respectively, along various optimal trajectories obtained using HRO and DHP for Case (1). It should be noted here that the severe changes in feed rate shown in Fig. 8 were permitted only to test the efficacy of DHP. As experts in the field of biochemical engineering are aware, such severe changes are typically avoided on a real fermentation process. They could be included as part of a control study by limiting the rate of change of feed rate that is set by the optimizer.

The feed rate profiles depicted in Fig. 8 throw up an interesting analysis. The feed rate profile obtained by DHP (Fig. 8) shows a low feed flow rate at the start of the fermentation process. During this time, the cells are typically in their lag phase and hence, the DHP realizes that it would not be economical to increase the feed rate. However, once the cells enter their exponential growth phase, the feed rate is increased highly. This results in a very high concentration of the substrate (glucose), which consequently increases the cell growth rate even further since the growth rate is directly proportional to the substrate concentration [18]. This results in very high rate of product formation during this phase. However, later, the feed rate is cut back slightly so as to ensure that the production of lactate does not cause the death rate of cells to be higher than the growth rate, i.e., effectively to avoid the cell cultivation to go the death phase.

On the other hand, both the offline and online HRO schemes consistently maintain higher feed rate of nutrients during the lag phase (where there is no cell growth that could result in useful product) and less flow rate during the exponential growth phase (during which the bulk of the product is formed). Hence, the productivity using either offline or online HRO scheme is expected to be lower even theoretically and it is validated by experimental results.

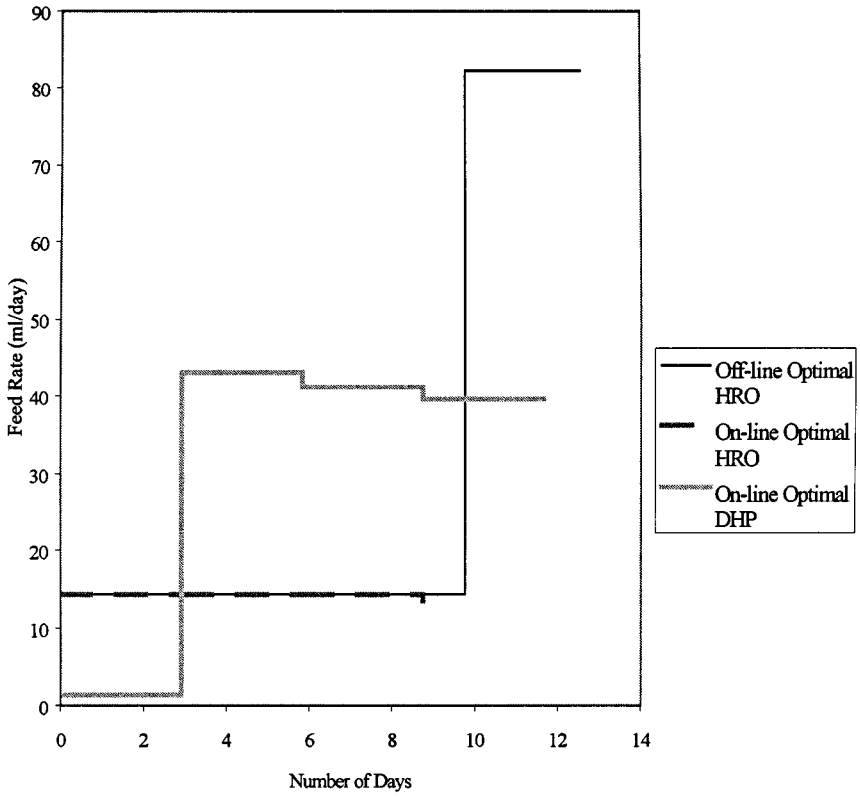

Fig. 8. Comparison of profile of feed rates along various optimal operating schedules.

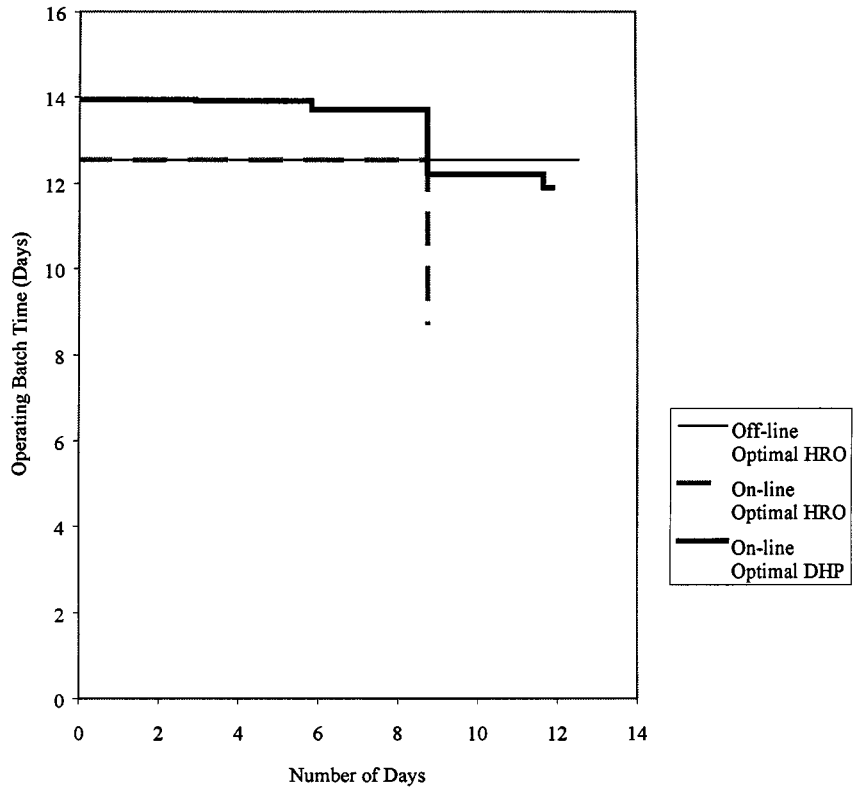

Fig. 9. Comparison of operating batch times along various optimal operating schedules.

In addition to improved productivity, the use of ACDs offers significant advantages over traditional direct search optimization routines like the HRO.

1) ACDs facilitate easy constraint handling via penalty functions and bounded activation functions in neural networks.

2) With traditional optimization routines, improvements in the model are translated into improved optimal operation only by dynamic reoptimization. However, with ACD, 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 ACDs are, in general, robust with respect to model uncertainties [22].

XIII. Conclusion

This study demonstrates the applicability of a simple scheme for off-line optimization and on-line model parameter adjustment and reoptimization using ACDs. The significant benefits in production, obtained by the use of ACD, is a pointer to possible avenues in exhaustive application of ACDs in the field of bioreactor optimization.

Appendix

The phenomenological model of a fed-batch fermentor is as follows.

A. The Overall Mass Balance on the Reactor Contents

$$\frac{dV}{dt} = q_0 \tag{A1}$$

The term on the left refers to the net rate of change in the volume of the reactor contents and the term on the right is the feed flow rate into the reactor. Liquid density is assumed constant.

B. The Mass Balance on Viable Cells

$$\frac{d(Vc_{\mathrm{Xv}})}{dt} = \mu(c_{\mathrm{S}}, c_{\mathrm{O}})\, c_{\mathrm{Xv}} V - K_d c_{\mathrm{Xv}} V \tag{A2}$$

The term on the left refers to the net rate of change of the total mass of viable cells in the reactor, the first term on the right gives the rate of increase in cell population due to growth and the second term on the right describes the rate of decline in cell population due to death.

C. The Mass Balance on Dead Cells

$$\frac{d(tVc_{\mathrm{Xd}})}{dt} = k_d c_{\mathrm{Xv}} V \tag{A3}$$

The term on the left refers to the net rate of change of dead cell population in the reactor and the term on the right is the rate of increase in dead cell population.

D. The Mass Balance on Dissolved Oxygen

$$\frac{d(Vc_{\mathrm{O}})}{dt} = K_{\mathrm{L}} a\,(c_{\mathrm{O}}{}^{*} - c_{\mathrm{O}})\, V - \frac{\mu(c_{\mathrm{s}}, c_{\mathrm{O}})\, c_{\mathrm{Xv}} V}{Y_{\mathrm{Xv/O_2}}} \tag{A4}$$

The term on the left refers to the net rate of change of dissolved oxygen content in the fermentation broth, the first term on the right describes the rate of transfer of oxygen from the gas phase while the second term on the right describes the rate of consumption of oxygen during cell growth.

E. The Mass Balance on Glucose

$$\frac{d(Vc_{\mathrm{S}})}{dt} = q_0 c_{\mathrm{s0}} - \frac{\mu(c_{\mathrm{s}}, c_{\mathrm{O}})\, V c_{\mathrm{Xv}}}{Y_{\mathrm{Xv/S}}} \tag{A5}$$

The term on the left refers to the net rate of change of total amount of glucose in the reactor, the first term on the right gives the rate of inflow of glucose to the reactor and the second term on the right gives the rate of consumption of glucose during cell growth.

F. The Mass Balance on the Antibody

$$\frac{d(Vc_P)}{dt} = \pi(c_{\mathrm{A}})\, V c_{\mathrm{Xv}} \tag{A6}$$

The term on the left describes the net rate of change of the total amount of antibody in the reactor and the term on the right describes the rate of antibody formation in the reactor.

G. The Mass Balance on the Lactate

$$\frac{d(Vc_{\mathrm{L}})}{dt} = K_{\mathrm{L}}(c_{\mathrm{s}})\, V c_{\mathrm{Xv}} \tag{A7}$$

The term on the left describes the net rate of change in the inhibitor concentration in the reactor and the term on the right describes the rate of increase of inhibitor concentration in the reactor.

H. The Mass Balance on the Amino Acid

$$\frac{d(Vc_{\mathrm{A}})}{dt} = q_0 c_{\mathrm{A0}} - \frac{\mu(c_{\mathrm{s}}, c_{\mathrm{O}})\, V c_{\mathrm{Xv}}}{Y_{\mathrm{Xv/A}}} \tag{A8}$$

The term on the left refers to the net rate of change of total amount of amino acid in the reactor, the first term on the right gives the rate of inflow of amino acid to the reactor and the second term on the right gives the rate of consumption of amino acid during cell growth.

The variation of some process parameters as functions of process states is described below.

1) The Specific Rate of Growth of Cells $\mu(c_{\mathrm{S}}, c_{\mathrm{O}})$: In the model

$$\mu(c_{\mathrm{s}}, c_{\mathrm{O}}) = \mu_{\max} \frac{c_{\mathrm{s}}}{K_{\mathrm{S}} + c_{\mathrm{s}}}. \tag{A9A}$$

In the process

$$\mu(c_{\mathrm{s}}, c_{\mathrm{O}}) = \mu_{\max} \frac{c_{\mathrm{s}}^{1.2}}{K_{\mathrm{S}} + c_{\mathrm{s}}^{1.2}} \frac{c_{\mathrm{o}}^{1.2}}{K_{\mathrm{o_2}} + c_{\mathrm{o}}^{1.2}}. \tag{A9B}$$

Here, it should be noted that if the effect of substrate concentration on the specific growth rate is negligible, then K_{O2} would be the concentration of dissolved oxygen that would yield a specific growth rate, which is half of the maximum possible. On the other hand, if the concentration of dissolved oxygen had a negligible effect of the specific growth rate, then the value of K_s is the concentration of substrate that yields a specific growth rate, which is half of the maximum possible.

Further, as experts in the field would be aware, the specific growth rate of cells could be represented by a factor that account for effects of amino acid concentration. However, since the substrate is typically the limiting nutrient and since the aim of this study was to test the efficacy of adaptive critic designs, this factor has been omitted from this study. It is recommended that more detailed equations be used for a more rigorous analysis.

2) The Specific Rate of Death of Cells $k_d(c_L)$: In the model

$$K_d(c_L) = K_{d\max}\left(\frac{c_L}{K_d + c_L}\right). \tag{A10A}$$

In the process

$$K_d(c_L) = K_{d\max}\left(\frac{c_L^{1.2}}{K_d + c_L^{1.2}}\right). \tag{A10B}$$

3) The Growth Yield Factor $Y_{Xv/S}$: In both the model and process

$$\frac{1}{Y_{Xv/S}} = \frac{1}{{Y_{Xv/S}}^*} + \frac{m}{\mu(c_s, c_O)}. \tag{A11}$$

In cell cultivation, the instantaneous yield of cell mass per unit consumption of glucose ($Y_{Xv/S}$) is a function of glucose and oxygen concentration in the fermentation broth. There is a certain concentration of the glucose and of oxygen when the instantaneous yield is a maximum ($Y^*_{Xv/S}$). For other concentrations of glucose and oxygen, the yield is less than the maximum. This effect is modeled in (A11) above by the introduction of a nonzero factor m. So long as $m/\mu(c_s, c_O)$ is nonzero, the value of $Y_{Xv/S}$ will always be less than $Y^*_{Xv/S}$. The value of m was selected so as to ensure that at a specific growth rate, equal to the maximum specific growth rate $\mu_{\max}$, the instantaneous yield of cell mass was about 95% of the maximum yield factor. However, the selection of this number will not have a bearing on the results of optimization. A different value for m would only result in a different trajectory for cell growth and, consequently, a different trajectory for product formation.

Further units on m are mass of glucose consumed per unit mass of cells per unit time while units on $\mu(c_s, c_O)$ are time^{-1}. Hence the ratio, $m/\mu(c_s, c_O)$, has the same units as those of either $1/Y_{Xv/S}$ or $1/Y^*_{Xv/S}$, viz. mass of glucose consumed per unit mass of viable cells produced.

4) The Specific Rate of Inhibitor Formation: In the model

$$K_I(c_s) = K_{I\max}\left(\frac{c_s}{K_I + c_s}\right) \tag{A12A}$$

In the process

$$K_I(c_s) = K_{I\max}\left(\frac{c_s^{1.2}}{K_I + c_s^{1.2}}\right) \tag{A12B}$$

5) The Specific Product Synthesis Rate: In the model

$$\pi(c_A) = \pi_{\max}\left(\frac{c_A}{K_A + c_A}\right). \tag{A13A}$$

In the process

$$\pi(c_A) = \pi_{\max}\left(\frac{c_A^{1.2}}{K_A + c_A^{1.2}}\right). \tag{A13B}$$

ACKNOWLEDGMENT

The authors would like to thank Prof. R. Russell Rhinehart, Head of the School of Chemical Engineering, Oklahoma State University, for his guidance in the initial stages of the study. In addition, the HRO developed by him with coworkers has been of great utility.

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] R. M. Dekkers, "Optimal control of a fed-batch fermentation," in *Innovations in Biotechnology*, E. H. Houwink and R. R. van der Meer, Eds. Amsterdam: Elsevier, 1984, pp. 313–330.

[4] A. Johnson, "The control of fed-batch fermentation processes-a survey," *Automatica*, vol. 23, no. 6, pp. 691–705, 1987.

[5] S. J. Parulekar, "Analytical optimization of some single-cycle and repeated fed-batch fermentations," *Chem. Eng. Sci.*, vol. 47, no. 15/16, pp. 4077–4097, 1992.

[6] J. Schubert, R. Simutis, M. Dors, I. Havlik, and A. Lubbert, "Bioprocess optimization and control: Application of hybrid modeling," *J. Biotechnol.*, vol. 35, pp. 51–68, 1994.

[7] K. Shimizu, "A tutorial review on bioprocess systems engineering," *Comput. Chem. Eng.*, vol. 20, no. 6/7, pp. 915–941, 1996.

[8] K. Dairaku, E. Izumoto, and H. Morikawa, "Optimal quality control of bakers' yeast fed-batch culture using population dynamics," *Biotechnol. Bioeng.*, vol. 24, pp. 2661–2674, 1982.

[9] M. S. Iyer, T. F. Wiesner, and R. R. Rhinehart, "Dynamic reoptimization of a fed-batch fermentor," *Biotechnol. Bioeng.*, vol. 63, no. 1, pp. 10–21, 1999.

[10] H. S. Harned and B. B. Owen, *The Physical Chemistry of Electrolytic Solutions (QC561.H33)*.

[11] S. Dhir, K. J. Morrow, R. R. Rhinehart, and T. F. Wiesner, "On-line dynamic optimization of a fed-batch fermentor," *Biotechnol. Bioeng.*, vol. 67, no. 2, pp. 197–205, 2000.

[12] J. Li and R. R. Rhinehart, "Heuristic random optimization," *Comput. Chem. Eng.*, vol. 22, no. 3, pp. 427–444, 1998.

[13] D. V. Prokhorov and D. C. Wunsch, "Adaptive critic designs," *IEEE Trans. Neural Networks*, vol. 8, pp. 997–1007, Sept. 1997.

[14] R. R. Rhinehart and J. B. Riggs, "Two simple methods for on-line incremental model parametrization," *Comput. Chem. Eng.*, vol. 15, no. 3, pp. 181–189, 1991.

[15] 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.

[16] J. G. Gaertner and P. Dhurjati, "Fractional factorial study of hybridoma behavior. 1. Kinetics of growth and antibody production," *Biotechnol. Progress*, vol. 9, pp. 298–308, 1993.

[17] 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 Eng.*, vol. 9, pp. 13–21, 1993.

[18] J. E. Bailey and D. F. Ollis, *Biochemical Engineering Fundamentals*, New York: McGraw-Hill, 1986.

[19] R. E. Bellman, *Dynamic Programming*. Princeton, NJ: Princeton Univ. Press, 1957.

[20] Y. Chauvin and D. E. Rumelhart, Eds., *Backpropagation: Theory, Architectures and Applications*. Hillsdale, NJ: Lawrence Erlbaum, 1995.

[21] G. Puskorius and L. Feldkamp, "Decoupled extended kalman filter training of feedforward layered neural networks," in *Proc. Int. Joint Conf. Neural Networks (IJCNN)*, Seattle, WA, 1991, pp. 771–777.

[22] Prokhorov and V. Danil, "Adaptive Critic Designs and Applications," Ph.D. dissertation, Elect. Eng. Dept., Texas Tech Univ., 1997.

Mahesh S. Iyer received the B.ChE. and M.ChE. degrees from the Department of Chemical Engineering, Bombay University, Bombay, India, in July 1993 and July 1995, respectively, and the Ph.D. degree from the Department of Chemical Engineering, Texas Tech University, Lubbock, in December 1997.

From February 1998 to August 1998, he worked as a Postdoctoral Research Associate in the Applied Computational Intelligence Laboratory, Texas Tech University. He was a Visiting Assistant Professor in the Department of Chemical Engineering at Texas Tech University from September 1998 to July 1999. Since July 1999, he has been working as a Senior Control Systems Engineer with Dresser-Rand Control Systems. Currently, his work activities involve modeling, advanced process control, and optimization.

Donald C. Wunsch, II (S'86–M'90–SM'94) received the B.S. degree in applied mathematics from the University of New Mexico, Albuquerque, in 1984, and completed a Humanities Honors Program at Seattle University in 1981, received the M.S. degree in applied mathematics, and the Ph.D. degree in electrical engineering from the University of Washington, Seattle, in 1987 and 1991, respectively.

He was Senior Principal Scientist at Boeing, where he invented the first optical implementation of the ART1 neural network, featured in the 1991 Boeing Annual Report and other optical neural networks and applied research contributions. He has also worked for International Laser Systems and Rockwell International and consulted for Sandia Labs, White Sands Missile Range, Texas Tech and Accurate Automation Corporation. He was Associate Professor of Electrical and Computer Engineering and Computer Science, at Texas Tech University. Since July 1999, he has been the Mary K. Finley Missouri Distinguished Professor of Computer Engineering in the Department of Electrical and Computer Engineering, University of Missouri, Rolla. He heads the Applied Computational Intelligence Laboratory and also has a joint appointment in Computer Science. His research activities include adaptive critic designs; neural-network pattern analysis, optimization, forecasting, and control; bioinformatics; financial engineering; fuzzy risk assessment for high consequence surety; intelligent agents; graph theory; quantum logic; and Go. He is heavily involved in research collaborations with former Soviet scientists. He has more than 100 research publications in various aspects of computational intelligence.

Dr. Wunsch is an Academician in the International Academy of Technological Cybernetics and in the International Informatization Academy; and is recipient of the Halliburton Award for excellence in teaching and research at Texas Tech and a National Science Foundation CAREER Award. He is a member of the International Neural Network Society, Association for Computing Machinery, Society of Photo Instrumentation Engineering, a life member of Phi Kappa Phi, a life member of the American Association of Artificial Intelligence, a life member of Sigma Xi and previously served as an Associate Editor of the IEEE TRANSACTIONS ON NEURAL NETWORKS and voting member of the IEEE Neural Network Council. He is Technical Program Coordinator for IJCNN 2002.