Optimization of fed-batch bioreactor using neural network model

B. Chaudhuri, J.M. Modak

Abstract An algorithm using feedforward neural network model for determining optimal substrate feeding policies for fed-batch fermentation process is presented in this work. The algorithm involves developing the neural network model of the process using the sampled data. The trained neural network model in turn is used for optimization purposes. The advantages of this technique is that optimization can be achieved without detailed kinetic model of the process and the computation of gradient of objective function with respect to control variables is straightforward. The application of the technique is demonstrated with two examples, namely, production of secreted protein and invertase. The simulation results show that the discrete-time dynamics of fed-batch bioreactor can be satisfactorily approximated using a feedforward sigmoidal neural network. The optimal policies obtained with the neural network model agree reasonably well with the previously reported results.

1 Introduction

In recent years many bioprocess engineering problems have been studied using artificial neural networks [1]. Process identification and control are two areas where the ability of neural network models to "learn" relationships between input and output of a complex, nonlinear bioprocess have been utilized. In absence of reliable sensors for measurement of biomass, substrate and product concentrations, neural network model has been shown to be an effective "software sensor" to estimate the bioprocess parameters using available measurements [2]. It has also been demonstrated that the neural network models can also be used for dynamic modelling of the fermentation processes [3–8]. Neural network have been found suitable for online control of chemical and bio-processes [9–14]. These studies indicate that neural network models are more efficient than conventional linear input-output models for on-line identification and control of highly nonlinear dynamic systems. In these applications, the objective was to control the bioprocess variables such as pH, substrate concentration or specific growth rate at predetermined desired values. Very few studies on optimization of bioreactors, that is, determining optimal values of control variables, using neural network model have been reported [15–16] and the potential of neural networks in optimization of fermentation processes remains relatively unexplored. In this work, application of neural network model for optimization of fed-batch bioreactor is illustrated.

Many industrially important fermentation processes involving production of antibiotics, enzymes and organic acids are carried out in fed-batch mode of operation in which the substrate(s) is added continuously to an otherwise batch operation. Fed-batch processes are particularly useful when the growth and/or metabolite production by microorganisms is inhibited at high substrate concentrations due to phenomena such as substrate inhibition, end-product inhibition or catabolite repression. The underfeeding of the substrate can lead to cell starvation while overfeeding can result in lower rates as well as formation of undesirable products. Therefore, the controlled addition of the substrate is essential to achieve maximum production of desired product for such type of fermentation processes. The problem of determining the optimal substrate feed rate profile is a *singular control* problem, so called because the control variable, substrate feed rate, appears linearly in the mass balance equations describing the process. In recent years, numerous studies dealing with determination of optimal substrate feed rate have appeared in literature [17–21]. In these reports, optimization problem is formulated in the framework of Pontryagin's maximum principle and control variable iteration using gradient of the Hamiltonian is employed to determine the optimal control. A deterministic mathematical model of the fermentation process kinetics is essential for this approach. The fermentation process often do not easily lend themselves to such quantitative description of the process. Often one encounters a situation where lot of data is available through experimental runs. However, the lack of understanding of highly nonlinear and complex interactions between key variables of the fermentation process hinder model structure specification and parameter estimation. In absence of the model, the use of optimal control theory to develop optimal strategies for fed-batch bioreactor is not possible. Therefore, alternative strategies for optimizing bioprocesses need to be developed. One method of overcoming these limitations is to quantify the knowledge gained from experimental trials using techniques which are capable of describing the nonlinearity of

B. Chaudhuri, J.M. Modak
Department of Chemical Engineering
Indian Institute of Science, Bangalore 560012, India

Correspondence to: J.M. Modak

the process. Artificial neural networks are ideally suited for such a quantification because of their excellent pattern recognition ability.

The objective of present study is to demonstrate the use of neural network, in particular feedforward neural network model, for determining the optimal substrate feeding strategies for fed-batch bioreactors. The development of algorithm is discussed first. The application of the algorithm is illustrated with two examples. In the first example, the discrete time data required for neural network training is generated using deterministic model of the process, while previously reported experimental data is used in the second example. The optimal substrate feed rate profiles obtained with neural network model are compared with previously reported optimal strategies.

2 Development of algorithm

2.1 Optimization problem

Consider a typical fed-batch operation for metabolite production involving growth of cells (X), consumption of substrate (S) and production of metabolite (P). The growth limiting substrate at concentration S_F is fed to the bioreactor at the rate of $F(t)$ and the contents of the fermentor are withdrawn only at the end of the operation ($t = t_f$). Since substrate is continually added without any withdrawal, the volume of the bioreactor contents (V) increases with time and there is a upper limit on V as determined by the bioreactor size. Similarly the feed rate is also constrained between upper ($F_{\max}$) and lower ($F_{\min} = 0$) limits. The optimization problem is to determine feed rate $F(t)$ during entire operation ($0 \le t \le t_f$), which maximize the objective function defined in terms of status of the fermentor at the end of operation, for example, maximize the amount of metabolite at the end of operation:

$$\max_{F(t)} IP = (PV)_{t_f} \quad , \tag{1}$$

subject to:

$$V(t) \le V_{\max} \text{ and } 0 \le F(t) \le F_{\max}\ 0 \le t \le t_f \quad , \tag{2}$$

where P is the metabolite concentration and final time t_f is assumed to be fixed *a priori*.

In order to develop neural network model of the fed-batch fermentation process based on the data available at fixed time intervals, the discrete time domain description of the fermentation process is required. Let us consider that fermentation data is available at N sampling intervals, $t_0, t_1, \ldots, t_j, \ldots, t_N$. Data at $t = t_0 (= 0)$ denotes the initial conditions while $t = t_N (= t_f)$ represents the final outcome of the fermentation process. The fed-batch operation can now be viewed as N discrete processes, each of duration Δt_j, taking place in series. This is schematically shown in Fig. 1. As N approaches infinity the process becomes continuous in time domain. The state variables for the process are concentrations of cells (X), substrate (S) and metabolite (P) and the bioreactor volume (V). Let X_j, S_j, P_j and V_j represent the status of the fermentation at time $t = t_j$. The outputs of the fermentation process at $t = t_j$ (X_j, S_j, P_j and V_j) depend on the input to the process, that is, X, S, P and V at the time $t = t_{j-1}$, feed rate during the interval (t_{j-1}, t_j) and the duration of the sampling interval $\Delta t_j = t_j - t_{j-1}$. The feed substrate concentration S_F is assumed to be constant during entire fermentation operation during the period (t_0, t_f) and therefore, is not considered as a separate variable.

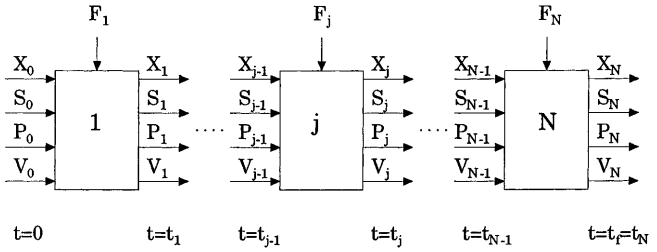

Fig. 1. Schematic representation of fed-batch fermentation process in discrete time domain

The serial nature of the N processes (Fig. 1) imply that the output of $j-1^{\text{th}}$ process is input for the j^{th} process. The feed rate F_j and sampling interval Δt_j are external inputs to the j^{th} process. The nonlinear relationship between the input and output of j^{th} process can be described by a neural network model. The same model can be used to describe fermentation process for all N intervals, if the model is trained by using the pair of input (X, S, P, and V at t_{j-1}) and output (X, S, P and V at t_j) for all $j = 1, 2, \ldots, N$. Thus, same neural network model can be used recursively to describe fermentation process during the entire operation $(0, t_f)$ once the initial conditions (X, S, P, and V at $t = 0$), feed rates during each interval $(F_1, F_2, F_3, \ldots, F_N)$ and the duration of sampling intervals (Δt_j) are specified.

2.2 Neural network model

A simple three layer feedforward neural network (denoted henceforth by FNN) is used to model fed-batch fermentation process in the j^{th} interval. For FNN with n_I, n_H and n_O neurons in input, hidden and output layers, respectively, the following set of equations apply [22]:

$$\text{Hidden layer} \quad \text{net}_i = \sum_{j=1}^{n_I} W_{ij} x_j + t_i$$

$$x_i = s(\text{net}_i)\ n_I + 1 < i < n_I + n_H \quad ,$$

$$\text{Output layer} \quad \text{net}_i = \sum_{j=n_I+1}^{n_I+n_H} W_{ij} x_j + t_i$$

$$x_i = s(\text{net}_i)\ n_{IH} + 1 < i < n_{IH} + n_0 \quad ,$$

$$\text{Activation function } s(z) = \frac{1}{1 + \mathrm{e}^{-z}} \quad , \tag{3}$$

where x_j, $j = 1, 2, \ldots, n_I$ are inputs to FNN and $n_{IH} = n_I + n_H$. The selection of input and output for neural network model to be developed for the process

during time interval (t_{j-1}, t_j) is now straightforward. Let x_i^{j-1} denote input to NN and x_i^j be the output:

$$\begin{aligned}\text{Input } & x_1^{j-1} = F_j,\ x_2^{j-1} = \Delta t_j,\ x_3^{j-1} = X_{j-1},\\ & x_4^{j-1} = S_{j-1},\ x_5^{j-1} = P_{j-1},\ x_6^{j-1} = V_{j-1},\\ \text{Output } & x_3^j = X_j,\ x_4^j = S_j,\ x_5^j = P_j\ . \end{aligned} \tag{4}$$

The bioreactor volume V_j is not included in the output of FNN model for the following reason. For the j^{th} process, the bioreactor volume increases because of addition of the substrate at the rate F_j for the duration Δt_j. Since F_j is constant during the interval, V_j and V_{j-1} are related by a simple relation:

$$V_j = V_{j-1} + F_j \Delta t_j\ . \tag{5}$$

Since F_j and Δt_j are the external inputs to the model, bioreactor volume can be eliminated from the output of the FNN. In general, if there are NS state variables describing the fed-batch fermentation process (X, S, P, V etc.) then the total number of inputs to FNN are $n_I = NS + 2$. The total number of outputs are $n_O = NS - 1$. Thus, FNN model of fermentation process can be summarized with following recursive equations:

$$\begin{aligned} x_i^j &= f\left(x_1^j, x_2^j, x_3^{j-1}, \ldots, x_k^{j-1}, \ldots, x_{NS+2}^{j-1}\right)\\ &\qquad i = 3, \ldots, n_I - 1;\ j = 1, 2, \ldots, N\ ,\\ x_{n_I}^j &= x_{n_I}^{j-1} + x_1^j x_2^j \quad j = 1, 2, \ldots, N\ , \end{aligned} \tag{6}$$

where x_1^j (feed rate) and x_2^j (sampling interval) are the external inputs to be specified for each interval and $x_i^o,\ i = 3, .., n_I$ specified as initial conditions. The nonlinear function f is defined by FNN equations given above, Eq. (3).

The neural network model of the fermentation process is trained with the discrete time data available. The input and output is scaled so that in each set the smallest magnitude element is 0.15 and the largest is 0.85, the latter because sigmoidal nodal transfer function saturates as scaled variables approach unity. The weights (W's) associated with the connections of neurons from one layer to another and the bias elements (t's) are the unknown parameters of the model. The training of the network consists of determining these parameters which was achieved using a modified backpropagation learning algorithm proposed by Werbos [22]. In the modified algorithm, the learning rate, which is assigned a fixed value in algorithm by Werbos [22], is determined through a one dimensional search in order to improve the rate of convergence of the algorithm. The weights (W's) and bias elements (b's) are initialized with random values between 0 and 1.

Apart from the ability of FNN to map complex relationship between inputs and outputs of the process, an important requirement in an optimization problem is the computation of the gradient of the objective function with respect to the control variables. The feedforward FNN has a convenient structure which makes the computation of the gradient, that is, $\frac{\partial(\text{input})}{\partial(\text{output})}$, straightforward. Using the chain rule of derivatives applied to the FNN model represented by Eq. (3), the gradient $\frac{\partial x_i}{\partial x_j}$ is given by:

$$\begin{aligned} \frac{\partial x_i}{\partial x_j} &= x_i(1 - x_i)\left(\sum_{k=n_I+1}^{n_{IH}} W_{ik} x_k (1 - x_k) W_{kj}\right)\\ &\qquad \text{for } j = 1, 2, \ldots, n_I \quad i = n_{IH} + 1, \ldots, n_{IH} + n_o\ , \end{aligned} \tag{7}$$

The details of the gradient calculation are presented in the following section.

2.3 Optimization with neural network model

The optimization problem defined by Eqs. (1) and (2) can now be expressed in the input/output notation for FNN introduced earlier, Eq. (4), as follows:

$$\max_{x_1^1, \ldots, x_1^j, \ldots, x_1^N} IP = \left(x_5^N x_6^N\right)\ , \tag{8}$$

subject to:

$$x_6^j \le V_{\max} \text{ and } 0 \le x_1^j \le F_{\max} \text{ for } j = 1, 2, \ldots, N\ . \tag{9}$$

The FNN model, Eq. (3), which is used for simulating the fermentation process is set of algebraic equations. Thus, the problem of determining the feed rates for N intervals is a constrained optimization problem in ordinary calculus for which a number of standard methods can be utilized. In this study, the generalized reduced gradient method (GRG) is used. The details of this method are available elsewhere [23] and therefore are repeated here. GRG requires calculation of gradient objective function with respect to control variables, that is, $\frac{\partial IP}{\partial x_1^j}, j = 1, 2, \ldots, N$. For the objective function given by Eq. (8), the gradient can be expressed as:

$$\frac{\partial IP}{\partial x_1^j} = x_6^N \frac{\partial x_5^N}{\partial x_1^j} + x_5^N \frac{\partial x_6^N}{\partial x_1^j}\ . \tag{10}$$

x_5^N, output of the N^{th} interval is directly related to x_k^{N-1} $k = 1, 2, \ldots, n_I$. Thus, feed rate in N^{th} interval x_1^N affects x_5^N explicitly. The feed rate in $N-1^{\text{th}}$ interval x_1^{N-1} determines x_k^{N-1} and thus x_1^{N-1} implicitly affects x_1^N. Extending the argument for all intervals, $x_1^1, x_1^2, \ldots, x_3^N$ (feed rates in all intervals) affect x_5^N implicitly. Therefore, the first term on the LHS of Eq. (10) can be evaluated as follows:

$$\frac{\partial x_k^N}{\partial x_1^j} = \sum_{l=1}^{n_I} \frac{\partial x_k^N}{\partial x_l^{N-1}} \frac{\partial x_l^{N-1}}{\partial x_1^j}\ . \tag{11}$$

The first term in above expression can be evaluated by using the gradient expression from FNN computations, Eq. (7). The second term can be evaluated using Eq. (11) regressively by substituting N by $N-1$ and so on. Actual computations proceed in forward direction, that is, evaluating $\frac{\partial x_k^i}{\partial x_1^j}$ starting with $i = 2$ and ending with $i = N$. The same procedure is repeated for all $j = 1, \ldots, N$.

The second term on the LHS of Eq. (10) represents partial derivative of bioreactor volume with respect to feed rate. In view of recursive Eq. (5), bioreactor volume is given by

$$V_N = V_0 + \sum_{j=1}^{N} F_j \Delta t_j \text{ or } x_6^N = x_6^0 + \sum_{j=1}^{N} x_1^j x_2^j \ , \qquad (12)$$

which yields $\partial x_6^N / \partial x_1^j = x_2^j$.

Some comments are in order for simplifying the gradient computations keeping in mind the nature of the fed-batch bioreactor operation.

1. The feed rate x_1^j and duration of sampling interval x_2^j for the j^{th} interval are external inputs which are not affected by the outputs of the previous interval, that is, $\frac{\partial x_i^j}{\partial x_k^{j-1}} = 0$ for $i = 1, 2$ and $k = 1, .., n_I$
2. As evident from Eq. (12), volume of the bioreactor is not affected by other state variables of the process, that is, $\frac{\partial x_6^j}{\partial x_k^{j-1}} = 0$ for $k = 1, .., n_I - 1$ and $\frac{\partial x_6^j}{\partial x_6^{j-1}} = 1$.
3. The feed rate in j^{th} interval affects only the processes following this interval and does not influence any of the preceeding interval processes, that is, $\frac{\partial x_k^i}{\partial x_1^j} = 0$ for all $i < j$.

Using these simplifications, the gradient of objective function with respect to all N feed rates can be computed. The optimal feeding strategy is then determined as a sequence of N feed rate sequences using GRG method. It should be pointed out that the objective function considered here is defined in terms of product concentration and bioreactor volume at the end of the operation. However, the computations can be easily extended to cases in which objective function contains other state variables as well. For example, objective function is defined as a profit function accounting for the substrate consumed, that is $IP = (PV)t_f - \alpha[(V_{t_f} - V_0)S_F - V_{t_f} S_{t_f}]$, where α is selling price of product relative to cost of the substrate. In such cases, the gradient calculation has to be suitably modified. It should also be pointed out that concentrations of various species and bioreactor volume alone were considered as state variables. However, other variables such as dissolved oxygen concentration, pH etc. can easily be incorporated, the only condition being that these are measurable quantities.

One limitation of using neural network approach for optimization purposes needs to be mentioned here. The approach presented here aims at determining the optimal substrate feed rate using the experimental data generated with predetermined feeding strategies. One can expect that the optimal strategy will yield metabolite quantities which are higher than that of the experimental data. Thus, optimization problem involves extrapolation of the data beyond the normal range of training data. It is well known that neural networks have excellent intrapolation ability, but their capability to extrapolate is rather poor. This problem can be overcome if the range of training data is extended and the optimal strategy is obtained as an intrapolation rather than extrapolation. This can be achieved by treating two constraints of the problem, namely, $V_{\max}$ and t_f as soft constraints for obtaining the training data. However, these are treated as hard constraints for the optimization problem. This idea is further illustrated with an example in the next section.

The procedure for determining the optimal substrate feeding strategies using neural network model can be summarized as follows:

Step 1: Experimental data for fed-batch fermentation process is collected at suitable time intervals using various different substrate feeding strategies. It is essential that the data is generated using strategies in which feed rate is constant during one sampling interval.

Step 2: The data is suitably scaled and backpropogation algorithm [22] is used for training the network model. The trained model is tested for its ability to predict the dynamics of fed-batch fermentation process.

Step 3: A sequence of feed rates is assumed as initial guess of the optimal feed rate. The trained dynamic neural network model is used to predict the state variables for all N intervals. It is also used for calculating the gradient of the objective function.

Step 4: The generalized reduced gradient (GRG) algorithm [23] is used to determine the optimal substrate feed rate.

3 Applications

The following two examples are provided to demonstrate the applicability of neural network model for determining optimal substrate feeding strategies for fed-batch bioreactor.

3.1 Example I: Secreted protein production

The fermentation process considered in this example consists of production of secreted protein by recombinant yeast strain growing on glucose. The first reason for choice of this example is that the detailed kinetic model and the optimization of fed-batch bioreactor for this process has been reported by Park and Ramirez [20]. The authors formulated the optimization problem in the framework of singular control theory and determined optimal substrate feeding strategies using an iterative numerical scheme. Therefore, the optimal results obtained with neural network model can easily be compared with the earlier results. The second reason is that the kinetics of the process is fairly complex, namely, the specific growth and metabolite production rates both exhibit inhibition kinetics. Furthermore, the specific growth rate of microorganism is maximum at the substrate concentration of 5 g/l while protein production is maximum at 0.1 g/l. Thus, this example provides a good test for demonstrating the applicability of the neural network approach. The model equations and required operating conditions are given in Appendix A. There are 5 state variables, namely, concentrations of microorganism (X), substrate (S), total protein (P_T), secreted protein (P_M) and the bioreactor volume (V). The optimization problem is to determine substrate feed rate $F(t)$ which maximized the amount of secreted protein ($(P_M V)t_f$) at the end of the operation.

The sampled data is generated by solving a set of ODE's describing the dynamics of the process and this is treated as "experimental data". Six different constant feed rates, namely, 0.1, 0.3, 0.5, 1.0, 2.0 and 5.0 l/h; which cover the

operating range of feed rates, were used for generating the data. As pointed out earlier, the constraints on bioreactor volume and operating time were treated as "soft constraints" for generating the training data. In other words, V_{max} of 20 l and t_f of 20 h was used for generating the data while original values (14.35 and 15 h, respectively, Appendix A) were used during the optimization phase of the algorithm. The duration of the sampling interval Δt was chosen as 0.5 h and was kept constant for all the intervals. Since 5 state variables describe the process ($NS = 5$), there are 7 inputs ($F_j, \Delta t_j, X_j, S_j, P_{M_j}, P_{Tj}$ and V_j) and 4 outputs (X_j, S_j, P_{M_j}, and P_{T_j}) for the neural network model. Since Δt_j (0.5 h)was same for all intervals, sampling interval was not used as one of the inputs for the neural network model thereby reducing the number of inputs to 6. By using a trial and error method, the number of hidden neurons was chosen as 5. Total of 240 pairs of input/output data was used to train the neural network model. The training of the network was continued till the RMS error was below 10^{-3}.

The trained neural network model is then tested for its ability to predict the dynamics of the process. The dynamic simulation is carried out by specifying the initial conditions and feed rates for each sampling interval and using trained FNN recursively for all sampling intervals. The results of FNN model predictions of cell mass, substrate, secreted protein and total protein are shown in Figs. 2 and 3. In these figures, solid lines represent profiles obtained with detailed kinetic model (Appendix A) while points denote the FNN model predictions. Figure 2 shows the results obtained with a constant feed rate ($F(t) = 1$ l/h) which was one of the feed rates used for generating the training data. It can be seen that the FNN model predictions of cell density and glucose concentration are in good agreement with the kinetic model simulation. The protein concentrations are over predicted by the model. It may be argued that the good agreement between FNN and

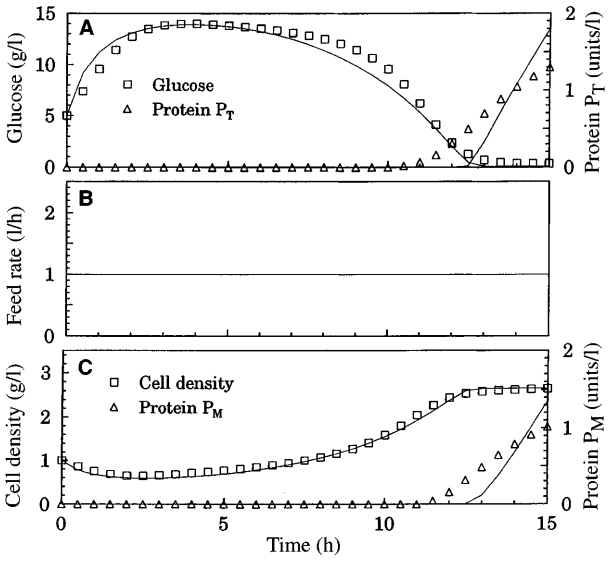

Fig. 2A-C. Dynamic simulation of secreted protein production in fed-batch bioreactor with substrate feed sequence used in generation of network training data. Solid lines – detailed kinetic model, points – neural network model

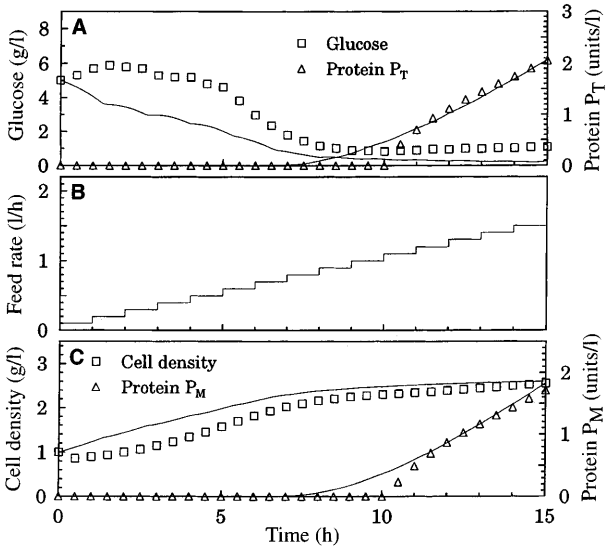

Fig. 3A-C. Dynamic simulation of secreted protein production in fed-batch bioreactor with substrate feed sequence not used in generation of network training data. Solid lines – detailed kinetic model, points – neural network model

kinetic model simulations is not surprising since the feed rate from the training data set is used for simulation. However, a basic difference between training the network and its use in dynamic simulation needs to noted. In training the network, each input/output data pair is individually presented to the network and the sum of squares of error between network predictions of the output and experimental value is minimized. As against this, dynamic simulation involves only specifying the initial conditions and feed rates for all intervals. The values of state variables at the end of one interval are calculated by the network model and presented as input for the next interval. In order to further demonstrate the ability of FNN model to predict the dynamics, the simulations are conducted with a linearly increasing feed rate (Fig. 3B) which is not used earlier in training the network. The simulation results in Fig. 3 clearly shows that the FNN model predictions are in good agreement with the detailed kinetic model simulation results even for this case. It is interesting to note in Figs. 2 and 3 that FNN model predicts no protein accumulation when glucose concentrations are very high (till about 10–12 h). Protein accumulation occurs only when glucose concentrations fall to a very low value. The relationship between the specific metabolite production rates and glucose concentration is explicitly contained in detailed kinetic model. On the other hand FNN model has no *a priori* knowledge about this relationship and yet it extracts this relationship through the training data that it sees. Thus, the pattern recognition ability of FNN model can be effectively used for predicting the dynamics of fed-batch bioreactor operation. This is a prerequisite for successful application of FNN model for determining the optimal substrate feed rates.

Having established the dynamic simulation ability of FNN model, the next step in the algorithm is determining the optimal substrate feed rates. The problem of maximizing the amount of secreted protein is formulated as

explained in previous section. In this phase of the algorithm, the constraints are imposed on the maximum bioreactor volume $V_{max} = 14.34$ l and final time $t_f = 15$ h. The constant feed rate of 1 l/h is used as an initial guess of the optimal feed rate and GRG algorithm is used to determine the optimal substrate feed rate. Figure 4 shows the optimal substrate feed rate and resulting state variable profiles. The dotted lines represent the optimization results reported by Park and Ramirez [20] (referred to henceforth as "true" results) while points denote the FNN model results. FNN model optimization gives the value of objective function of 31.3 which is very close to the true optimal (32.4). The true optimal substrate feeding policy consisted of four phases: singular feed rate which maintains the glucose concentration constant at 5 g/l till about 9 h in order to maximize growth of cells, a batch period till concentration decreases to 0.1 g/l, singular feed rate which maintains glucose concentration at 0.1 g/l till about 14.4 h in order to maximize protein production and finally feeding substrate at maximum rate. The FNN predicted optimal feed rate (Fig. 4B), although not numerically matching with the true optimal, does show the key features of the later. Three phases are seen in FNN optimal policy, initial singular feed rate keeping glucose at high levels, a batch period till concentration decrease to a low value and finally another singular interval for keeping glucose at constant level. The cell density and protein concentration profiles predicted by FNN model match reasonable well with true optimal profiles. There is a discrepancy in glucose concentration profiles in the initial stages till about 8 h. The examination of the specific rates in the model (Appendix A) reveals that the specific growth rate $C(S)$ is maximum at 5 g/l. However, $C(S)$ is insensitive to variations in glucose concentration in the range 5–10 g/l, that is growth rate decreases by less than 0.1% when glucose concentration is increased from 5 to 10 g/l. This insensitivity of $C(S)$ leads to FNN model predicting higher glucose concentrations in the first phase. It should be pointed out that the detailed kinetic model contains more physical information about the process than the FNN model. Furthermore, true optimal results are obtained with formulating the problem with optimal control theory and careful analysis of optimality conditions. Inspite of these shortcomings in the FNN model approach, the optimal results are remarkably similar to the true results.

3.2
Example 2: Invertase production

The fermentation process considered in this example consists of production of invertase by recombinant yeast cells utilizing glucose for their growth. Patkar and Seo [24] reported the fermentation kinetics of invertase production in fed-batch cultures. Authors reported experimental data for cell density (expressed as optimal density OD), glucose (G) and ethanol (E) concentrations and specific invertase activity (I) obtained with six different glucose feeding strategies. The sampling interval for these experiments varied between 0.2 to 2 h with majority of samples taken at about 1 h interval. In each of these experiments, the glucose feed rates was kept constant during sampling interval. Therefore, the data presented in the paper is ideally suited for demonstrating the applicability of the FNN approach of optimization of the fed-batch bioreactor. The number of state variables for the process is 5 (including the bioreactor volume V), and therefore, there are 7 inputs $(F_j, \Delta t_j, \mathrm{OD}_{j-1}, G_{j-1}, E_{j-1}, I_{j-1})$ and 4 outputs $(\mathrm{OD}_j, G_j, E_j, I_j)$ for the FNN model. The FNN model with 5 neurons in hidden layer is trained with 64 pairs of input/output till the error between model predictions and experimental data reduced to below 10^{-3}. Figure 5 shows the results of dynamic simulation using trained FNN model for one of the glucose feeding sequence labelled F1 [24] and shown here in Fig. 5B. The solid symbols are the experimental results while open symbols are the FNN predictions. The agree-

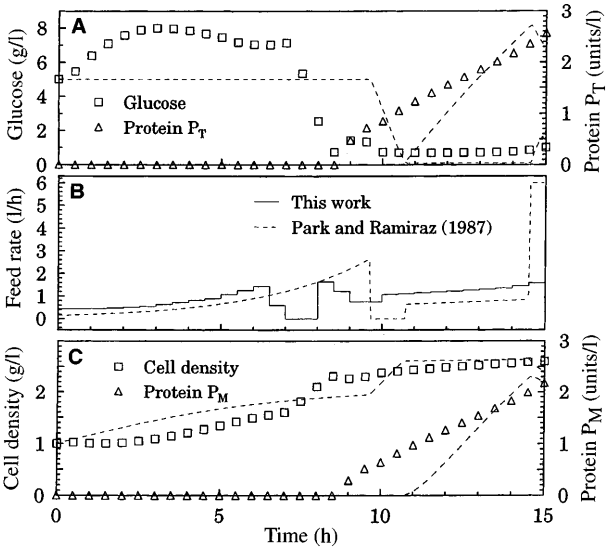

Fig. 4A-C. Optimal profiles for secreted protein production in fed batch bioreactor. Dashed lines – optimal results of Park and Ramirez [20]. Solid lines and points – optimal results using neural network model

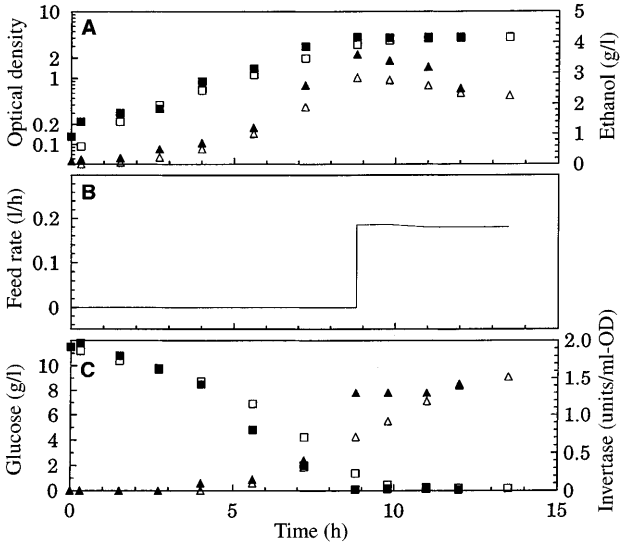

Fig. 5A-C. Dynamic simulation of invertase production in fed-batch bioreactor with substrate feed sequence used in generation of network training data. Filled symbols – experimental data, open symbols – neural network model prediction

ment between experimental data and FNN predictions in this example is not as good as that for the *Example 1* discussed above. This is to be expected since the data in this example is real experimental data while kinetic model simulated data was treated as experimental data in the *Example 1*. However, FNN model does correctly predict several trends observed experimentally. Several noteworthy features among these are: (a) initial exponential increase in OD followed by stationary phase (Fig. 5A), (b) initial accumulation of ethanol followed by its consumption (Fig. 5A) and (c) no invertase production at high glucose concentration followed by rapid accumulation of invertase when glucose concentrations are very low (Fig. 5C). The dynamic simulation with a feed rate (Fig. 6B) not used in the training data is shown in Fig. 6. Since experimental data for this feed rate is not available, only qualitative observations are possible. The feed rate (Fig. 6B) is similar to that shown in Fig. 5B except the feed rate is set at 0.28 l/h after 8.2 h. The experimental results of Fig. 5 show that maintaining feed rate of 0.18 beyond 8.2 h maintains glucose concentration at a very low value. It can be expected that when the feed rate is increased from 0.18 to 0.28 l/h (Fig. 6B), the glucose concentrations would increase. Indeed, FNN model does predict the increase in glucose levels upto 2 g/l (Fig. 6C). The invertase activity is more or less constant (Fig. 6C) once the glucose feeding at 0.28 l/h starts and the activity in general is lower compared to that obtained when glucose feed rate is lower at 0.18 g/l (Fig. 5C). These FNN model predictions are in concurrence with experimental observations of Patkar and Seo [24], namely, the glucose concentration tightly regulates the invertase activity and the activity remains more or less constant when glucose concentration exceeds about 2 g/l. Patkar et al. [21] reported a kinetic model for the invertase production which was developed using experimental data of the earlier paper [24]. The solid lines in Fig. 6 represent the simulation of this model. It can be seen that the FNN model results agree reasonably well with those of dynamic model simulation.

The optimal glucose feeding policies were also determined by Patkar et al. [21] using the optimal control theory and conjugate gradient algorithm. Thus, it is possible to compare the theoretical optimal policies with those determined using FNN approach. The details of the kinetic model and operating conditions are given in Appendix A. The objective function considered is maximization of amount of invertase, $(PVX)_{t_f}$. The optimal glucose feeding policy is determined using the trained network model and the gradients calculated using it. As pointed out earlier, the sampling interval for majority of the data was about 1 h, and therefore, the optimization with FNN model is done with fixed $\Delta t_j = 1$ h. The results of optimization are shown in Fig. 7. The dotted lines represent the optimal results reported by Patkar et al. [21] and points denote FNN model results. The maximum amount of invertase, 7.1 units, predicted by FNN model compares well with 7.3 units reported by Patkar et al. [21]. As in the previous example, the optimal feeding policy obtained with FNN model shows the same pattern as that of the theoretical optimal. The feeding policy (Fig. 7A) consists of a batch period till glucose concentrations decrease to a low level followed by feeding glucose to maintain low glucose concentrations. Eventhough glucose and invertase profiles (Fig. 7B) show some discrepancy, the general trends are correctly predicted by FNN model.

The above two examples clearly demonstrate the ability of neural network model in determining the optimal substrate feed rates for fed-batch bioreactor. The key factor contributing to the success of FNN model is its ability to predict the dynamics of the process correctly. The pattern recognition ability of FNN can easily be used for other applications as well. One such application is

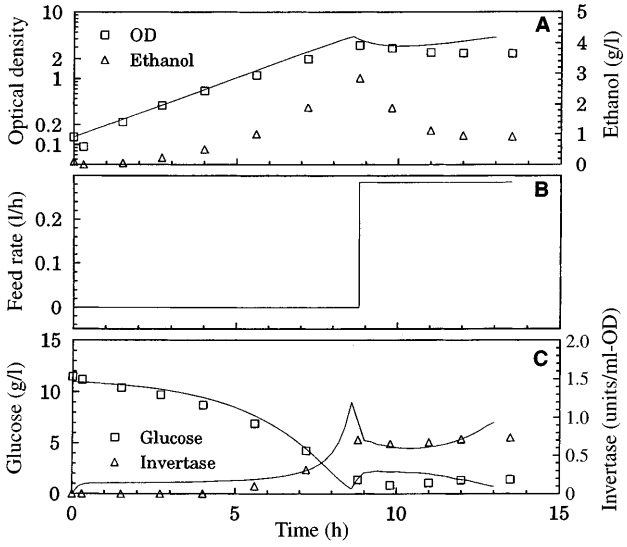

Fig. 6A-C. Dynamic simulation of invertase production in fed-batch bioreactor with substrate feed sequence not used in generation of network training data. Open symbols neural network model prediction

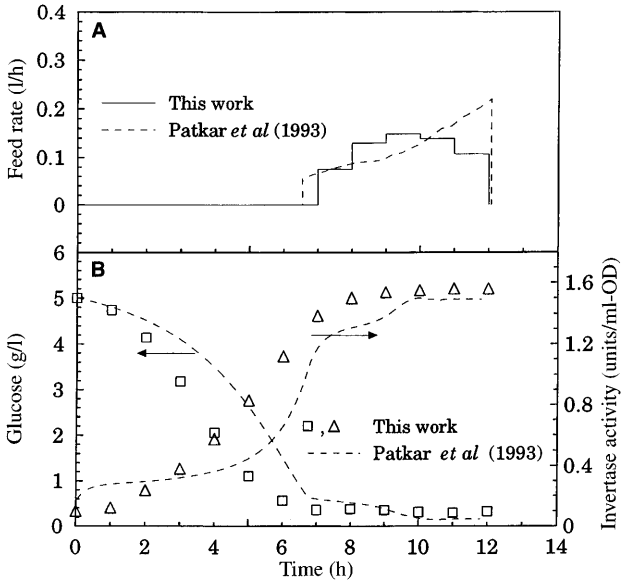

Fig. 7A,B. Optimal profiles for invertase production in fed batch bioreactor. Dashed lines – optimal results of Patkar et al. [21]. Solid lines and points – optimal results using neural network model

discussed here. Often, a large amount of data, for examples, concentrations of various species, temperature pH, dissolved oxygen concentration etc., is collected during the course of the fermentation. In order to use such a data set for optimization purposes or even for developing the detailed kinetic model of the process, it is necessary to identify the key variables which affect the performance of the process. Input-output sensitivity calculations performed with a trained neural network of the model using entire data set can be used for this purpose. In Example 2, there are 7 inputs, namely, feed rate, duration of sampling interval, cell, ethanol, glucose and invertase concentrations and the bioreactor volume and 4 outputs, concentrations of cell, ethanol, glucose and invertase. A simple test is conducted with the trained neural network model to determine what is contribution of each of these inputs in determining the output. For a given data set, each one of the 7 scaled inputs is assigned 100 different values (between 0–1 range) generated using random number generator while maintaining all other inputs at their nominal values. The output response for all 4 outputs is calculated using trained FNN model and percentage change in each of the outputs resulting from changes in all the inputs is computed. The procedure is repeated for all 64 data points and the average contribution of each input is then calculated for all four outputs. Table 1 shows the results of such a computation for the invertase production process. It should be kept in mind that the concentration of the species appearing in the input column are those at the beginning of the interval while concentrations appearing in the output are those at the end of the interval. Several interesting observations regarding the operation of fed-batch bioreactor can be noted from Table 1. The most dominant contribution for concentration of four species is the concentration of respective species themselves. For example, contribution of cell density as input is 55.2% for the cell density as the output. This is to be expected since the concentration of any species at the end of the sampling interval is strongly dependent on its concentration at the beginning of the interval. The highest contribution (19.7%) of glucose feed rate as the input is for the glucose concentration among all the outputs indicating stronger influence of feed rate on the substrate concentration. Some information regarding the kinetics of the process can also be extracted from Table 1. Ethanol concentration as a input has very little influence ($\leq$8.3%) on the concentrations all the species except ethanol. Invertase as a input also contributes very little ($\leq$3%) to the concentrations of other species. These observations indicate that ethanol and invertase have very little or nil effect on the kinetics of cell growth and substrate utilization. On the other hand, the invertase activity at the end of the interval is influenced by glucose concentration in almost equal measure as the invertase activity at the beginning of the interval. It is interesting to note that these features can be seen in the kinetic model of Patkar et al. [21]. Ethanol was not considered for the modelling purposes by the authors as seen from the mass balances given in Appendix A. Similarly, Invertase activity (P) does not affect the specific rates of cell growth, glucose utilization or invertase production. Thus, trained FNN can be utilized for extracting useful information regarding the process kinetics apart from its use in optimization.

Table 1. Percent contribution of each of the input for outputs of invertase production process

Inputs	Outputs			
	Cell density	Ethanol	Glucose	Invertase
Feed rate	12.4	13.4	19.7	14.7
Sampling interval	7.9	7.6	3.4	6.1
Cell density	55.2	29.5	7.8	14.3
Ethanol	5.9	30.7	7.4	8.3
Glucose	8.4	11.0	54.3	21.0
Invertase	2.6	2.8	2.8	29.1
Bioreactor volume	7.6	5.0	4.6	6.5

4 Conclusions

The application of feedforward neural network for determining the optimal substrate feed rate for fed batch fermentation processes is discussed in this paper. The optimization problem which falls under realm of calculus of variations is transformed into a problem in ordinary calculus by considering the process in discrete time domain, that is, dividing the entire operating period into N intervals and determining the feed rate for each of the N intervals. A FNN model of the process is developed using process data available at different sampling intervals. The gradient of the objective function with respect to control variables (N feed rates) can easily be computed using FNN model. The simulation results presented for secreted protein and invertase production clearly show that the FNN model captures the essential features of the process kinetics, and thereby, the model can be used for dynamic simulation of the process. The optimal feeding policies obtained with trained FNN agrees reasonably well with the previously reported results. The main advantage of the approach lies in the fact that optimization can be achieved without the knowledge of the detailed kinetic model of the process. When the kinetic model of the process is available, optimization problem can be formulated in the framework of optimal control theory. It can be expected that the discrete feed profile obtained with FNN will not satisfy all the optimality conditions. Even then FNN approach can be utilized for generating an initial estimate of the optimal which is required in any iterative scheme. Better results can perhaps be obtained in cycle-to-cycle optimization scheme in which FNN model is used for design of experiments. Such an approach would involve developing crude FNN model from preliminary data, generating more data using FNN optimal feed profile and refining the model using new data till no further improvement is achieved. Finally, it should be pointed out that the approach can be easily adopted for other dynamic optimization problems such as determining temperature profiles in batch reactor and multivariable optimization problems.

Appendix A: Models for fed-batch bioreactor

Example 1: Secreted protein production (Park and Ramirez, 1988)

Mass balance equations:

$$\left(\dot{XV}\right) = C(S)XV$$

$$\left(\dot{SV}\right) = FS_F - YC(S)XV$$

$$\left(\dot{P_M V}\right) = A(S)(P_T - P_M)V$$

$$\left(\dot{P_T V}\right) = B(S)XV$$

$$\left(\dot{V}\right) = F$$

$$A(S) = \frac{4.75C(S)}{0.12 + C(S)} \quad B(S) = \frac{Se^{-5.0S}}{0.1 + S}$$

$$C(S) = \frac{21.87S}{(S + 0.4)(S + 62.5)} \quad Y = 7.3$$

where the X (g/l), S (g/l), P_M (units/l), and P_T (units/l) represent the concentrations of cell, substrate, secreted protein and total protein, respectively.

Operating conditions: $X(0) = 1.0$, $S(0) = 5.0$, $P_M(0) = 0.0$, $P_T(0) = 0.0$, $V(0) = 1.0$, $S_F = 20$, $V_{\max} = 14.35$, $F_{\max} = 10.0$, $t_f = 15.0$

Objective functions: $\max_{F(t)} IP = (P_M V)t_f$

Example 2: Invertase production (Patkar et al. 1993)

Mass balance equations:

$$\left(\dot{XV}\right) = \left(R_r Y_{xr} + R_f Y_{xf}\right)XV$$

$$\left(\dot{SV}\right) = FS_F - R_t XV$$

$$\left(\dot{PXV}\right) = (\pi - k_d P)XV$$

$$\left(\dot{V}\right) = F$$

$$R_r = \frac{0.55S}{0.05 + S} \quad R_t = \max\left(\frac{1.25S}{0.95 + S}, R_r\right)$$

$$\pi = \frac{6.25S}{0.1 + S + 2.0S^2} \quad R_f = R_t - R_r$$

$$Y_{xr} = 0.6 \; Y_{xf} = 0.15, \; k_d = 1.85$$

where the X (OD), S (g/L) and I (units/OD/L) represent the concentrations of cell, substrate, and invertase, respectively.

Operating conditions: $X(0) = 0.15$, $S(0) = 5.0$, $P(0) = 0.1$, $V(0) = 0.6$, $S_F = 10$, $V_{\max} = 1.2$, $F_{\max} = 0.6$, $t_f = 12.0$

Objective functions: $\max_{F(t)} IP = (PXV)t_f$

References

1. **Baughman, D.R., Liu, Y.A.:** Neural Networks in Bioprocessing and Chemical Engineering, p. 5, San Diego: Academic Press 1995
2. **Karim, M.N., Rivera, S.L.:** Artificial neural networks in bioprocess state estimation. Advances in Biochemical Engineering, 46 (1992) 1–33
3. **Thibault, J., Van Breusegem, V., Cheruy, A.:** On-line prediction of fermentation variables using neural networks, Biotechnol. Bioeng., 36 (1990) 1041–1048
4. **Di Massimo, C., Montague, G.A., Willis, M.J., Tham, M.T., Morris, A.J.:** Towards improved penicillin fermentation via artificial neural networks. Computers Chem. Engng., 16 (1992) 283–291
5. **Psicos, D.C., Ungar, L.H.:** A hybrid neural network – First principles approach to process modelling, AIChE J., 38 (1992) 1499–1511
6. **Syu, M.J., Tsao, G.T.:** Neural network modeling of batch cell growth pattern. Biotechnol. Bioeng., 42 (1993) 376–380
7. **Glassey, J., Montague, G.A., Ward, A.C., Kara, B.V.:** Artificial neural network based experimental design procedures for enhancing fermentation development. Biotechnol. Bioengng., 44 (1994) 397–405
8. **Latrille, E., Corrieu, G., Thiabault, J.:** Neural network models for final process time determination in fermented milk production. Computers Chem. Engng., 18 (1994) 1171–1181
9. **Bhat, N., McAvoy, T.:** Use of neural nets for dynamic modeling and control of chemical process systems. Computers Chem. Engng., 14 (1990) 573–583
10. **Narendra, K.S., Parthasarathy, K.:** Identification and control of dynamical systems using neural networks. IEEE Trans. Neural Networks, 1 (1990) 4–27
11. **Hoskins, J.C., Himmelblau, D.M.:** Process control via artificial neural networks and reinforcement learning. Computers Chem. Engng., 16 (1992) 241
12. **Lee, M., Park, S.:** A new scheme combining neural feedforward control with model-predictive control. AIChE J., 38 (1992) 193–200
13. **Shi, Z., Shimizu, K.:** Neuro-fuzzy control of bioreactor systems with pattern recognition. J. Ferment. Bioeng., 74 (1992) 39–51
14. **Ye, K.S., Jin, S., Shimizu, K.:** Fuzzy neural networks for the control of high density cell cultivation of recombinant *Escherchia coli*, J. Ferm. Bioeng., 77 (1994) 663–669
15. **Chen, Q., Weigand, W.A.:** Dynamic optimization of nonlinear processes by combining neural net model with UDMC. AIChE J., 40 (1994) 1488–1497
16. **Sarkar, D., Modak, J.M.:** Adaptive optimization of continuous bioreactor using neural network model. Chem. Eng. Comm., 143 (1996) 99–116
17. **Lim, H.C., Tayeb, Y.J., Modak, J.M. Bonte, P.:** Computational algorithms for optimal feed rates for a class of fed-batch fermentation: Numerical results for penicillin and cell mass production, Biotech. Bioengng., 28 (1986) 1408–1420
18. **Menawat, A., Muthurasan, R., Coughanowr, D.R.:** Singular optimal control strategy for a fed-batch bioreactor: Numerical Approach. AIChE J., 33 (1987) 776–783
19. **Modak, J.M., Lim, H.C.:** Simple nonsingular control approach to fed-batch fermentation optimization. Biotech. Bioengng., 33 (1989) 11–15
20. **Park, S., Ramirez, W.F.:** Optimal production of secreted protein in fed-batch reactors. AIChE J., 34 (1988) 1550–1558
21. **Patkar A., Seo, J.H., Lim, H.C.:** Modeling and optimization of cloned invertase expression in *Saccharomyces cerevisiae*. Biotech. Bioengng., 41 (1993) 1066–1074
22. **Werbos, P.J.:** Backpropogation through time: What it does and how to do it. Proc. of the IEEE., 78 (1990) 1550–1560
23. **Reklatis, G.V., Ravindran, A., Ragsell, K.:** Engineering Optimization, pp. 381–390, New York: Wiley 1983
24. **Patkar, A., Seo, J.H.:** Fermentation kinetics of recombinant yeast in batch and fed-batch cultures. Biotech. Bioengng., 40 (1992) 103–109