0273-2289/01/91–93/0341/$13.00

A Hybrid Neural Network Algorithm for On-Line State Inference That Accounts for Differences in Inoculum of *Cephalosporium acremonium* in Fed-Batch Fermentors

Rosineide G. Silva, Antonio J. G. Cruz, Carlos O. Hokka, Raquel L. C. Giordano, and Roberto C. Giordano*

Departamento de Engenharia Química, Universidade Federal de São Carlos, C.P. 676, CEP 13565-905, São Carlos, SP, Brazil, E-mail: roberto@deq.ufscar.br

Abstract

One serious difficulty in modeling a fermentative process is the forecasting of the duration of the lag phase. The usual approach to model biochemical reactors relies on first-principles, unstructured mathematical models. These models are not able to take into account changes in the process response caused by different incubation times or by repeated fedbatches. To overcome this problem, we have proposed a hybrid neural network algorithm. Feedforward neural networks were used to estimate rates of cell growth, substrate consumption, and product formation from on-line measurements during cephalosporin C production. These rates were included in the mass balance equations to estimate key process variables: concentrations of cells, substrate, and product. Data from fed-batch fermentation runs in a stirred aerated bioreactor employing the microorganism *Cephalosporium acremonium* ATCC 48272 were used. On-line measurements strongly related to the mass and activity of the cells used. They include carbon dioxide and oxygen concentrations in the exhausted gas. Good results were obtained using this approach.

Index Entries: Neural networks; hybrid model; cephalosporin C production; state inference.

Introduction

Cephalosporin C is a natural β-lactam antibiotic produced by strains of a strictly aerobic fungus, *Cephalosporium acremonium*. Industrial production

*Author to whom all correspondence and reprint requests should be addressed.

of this antibiotic is carried out in aerated, stirred bioreactors using submerged cultures of mutant strains. This antibiotic may be hydrolyzed to manufacture 7-amino cephalosporanic acid (7-ACA), which is employed to produce the commercially available, semisynthetic cephalosporins *(1)*.

Many researchers have proposed mechanistic models to simulate the cephalosporin C production process *(2–5)*. These models were based on the overall biosynthetic mechanisms, and although simplifying assumptions were made, the problem of estimating a large number of parameters still remained. Furthermore, changes in strain and small differences in the reactivation of the microorganism may cause appreciable modifications on the system response.

In biochemical processes, it is often difficult to establish quantitative relationships between microbial growth and production rates. This difficulty is even more pronounced when the product is a secondary metabolite. On the other hand, the lack of reliable sensors to measure key process variables makes their inference a necessity for the on-line monitoring and control of the bioprocess.

Artificial neural networks (ANNs) emerged as a very useful tool for process modeling. The ability of ANNs to represent complex nonlinear relationships without prior knowledge of any model structure makes them a promising alternative in many practical situations. One way to use ANNs, called the black box approach, is to replace the complete first-principles model with the artificial net. Applications of this approach to bioprocesses have been reported in the literature *(6–11)*.

An alternative to this technique is called gray box modeling strategy, or hybrid modeling *(12)*. In this case, the mass and energy balances are used in conjunction with ANNs, which serve as an estimator of unmeasured parameters or variables. Psichogios and Ungar *(12)* reported the use of this hybrid model to study the dynamic behavior of a fed-batch stirred bioreactor. Following this approach, mass and energy conservation would not be violated by the model, which is not ensured when the ANN is a black box. According to van Can et al. *(13)*, another advantage of the hybrid model might be its better extrapolation properties.

In this article, a hybrid neural network model is proposed to infer state variables of the cephalosporin C production process. The hybrid model includes three neural networks, combined with mass balance equations. The first network estimates the specific rate of cell growth from selected on-line measurements and initial conditions. The second and third ones predict glucose consumption and specific antibiotic production, respectively, using the specific growth rate predicted by the previous network. The solution of the mass balances provides the concentrations of cells, substrate, and product, using the output of the three ANNs. Most important, the proposed model was able to forecast the main features of the system response when the inoculation followed a nonstandard procedure.

Modeling Cephalosporin C Production

Production of cephalosporin C by *C. acremonium* has usually been modeled using variations on Monod kinetics. Morphologic differentiation of the microorganism and catabolic repression by glucose are also studied *(2–5)*. Matsumura et al. *(2)* analyzed the induction of cephalosporin C by endogenous methionine, while Chu and Constantinides *(3)* and Basak et al. *(4)* focused on the role of enzymes synthesized by the cells in a complex enzymatic pathway. The complexity of the internal cellular relations makes it very difficult to quantify the production of cephalosporin from the metabolic reactions. All the available information does allow a grasp of the main features of the process, useful for modeling and optimization purposes. Nevertheless, it is virtually impossible that a pure, first-principles ("white box") model will be generic enough to take into account all the consequences of the performance of a biochemical reactor when some random alteration occurs in the inoculum. In this case, all kinetic parameters previously estimated from experimental data might be inaccurate, and systematic deviations from the predicted reactor trajectory would occur.

ANNs in Biochemical Process Simulation

ANNs may be used to circumvent the difficulties described beforehand. ANNs consist of interconnected units, called nodes or neurons, that are organized in layers. A very common type of neural net is the three-layer (input, hidden, and output) feedforward neural (FNN) network. The number of nodes in the input and output layers is equal to the number of input and output variables in the process under investigation. The activity of the input units reflects the raw information flux through the ANN. The optimal number of nodes in the hidden layers depends on the type and complexity of the task the ANN is designed to perform. Usually, this number is determined by trial and error. Each interneuron connection has a weight associated with it, emulating the synapses between actual neurons. These weights are the internal parameters of the network, whose fitting allows the ANN to "learn" information about the system to be modeled. This procedure is called the ANN training.

There are several training algorithms in the literature. Among them, one of the most popular is the backpropagation routine *(14)*, a modified gradient descendent method that minimizes an objective function by redistributing the output error back through the network, appropriately modifying the weights of the nodes' connections. The behavior of the ANNs depends mainly on the weights and the transfer function (typically nonlinear) that is specified for the nodes. The most popular architecture uses sigmoidal functions *(15)* for this task.

FNNs are frequently applied to model (bio)chemical reactors. Thibault et al. *(6)* used an FNN to predict key fermentation variables. Di Massimo et al. *(7)* utilized FNNs to model the production of penicillin. Warnes et al.

(11) tested FNNs and radial-basis function networks to estimate the concentration of biomass and of a recombinant protein produced by *Escherichia coli*—variables that are usually measured off-line. Cruz et al. *(10)* modeled cephalosporin C production using two FNNs, the first one estimating cell concentration and the second inferring antibiotic concentration. They used black box models. The neural network replicates the input-output behavior of the whole system.

With a hybrid model, it is possible to include prior knowledge of the system, even if it is not complete. The ANN would assume the difficult task of emulating aspects that could not be reasonably modeled within the field of a conventional approach. Hybrid models are especially indicated for biochemical processes, in which key variables are not measured on-line and the data training sets are sparse. The ANN would serve as an estimator of unknown parameters or variables of the phenomenologic model of the process. Psichogios and Ungar *(12)*, Thompson and Kramer *(16)*, and van Can et al. *(13,17)* have provided examples of successful applications of the gray box approach to biochemical/enzymatic reactors.

Materials and Methods

Microorganism

C. acremonium ATCC 48272 was used throughout this work. This strain is able to produce more than 1000 mg of cephalosporin C/L, when grown in a synthetic medium *(18)*. It was kept on cryotubes in the presence of glycerol at –50°C.

Culture Media

For inoculum preparation a synthetic medium *(19)* was utilized containing (the following): 30.0 g/L of glucose, 8.8 g/L of ammonium acetate, 5.0 g/L of DL-methionine, 1.5 g/L of oleic acid, 2.3 g/L of KH_2PO_4, 5.8 g/L of K_2HPO_4, 0.16 g/L of $Fe(NH_4)_2(SO_4)_2 \cdot 6H_2O$, and 2.0 g/L of $CaCO_3$. The pH was 7.0 ± 0.1. Micronutrients were provided by a salt solution (50 mL/L) whose composition was 16.2 g/L of Na_2SO_4, 7.68 g/L of $MgSO_4 \cdot 7H_2O$, 1.6 g/L of $CaCl_2 \cdot 2H_2O$, 0.64 g/L of $MnSO_4 \cdot H_2O$, 0.64 g/L of $ZnSO_4 \cdot 7H_2O$, and 0.04 g/L of $CuSO_4 \cdot 5H_2O$. The composition of the production medium was the same as that of the inoculum medium, except for glucose, whose concentration was 27.0 g/L.

Experimental Procedure

The standard procedure for fermentation started with the transfer of 15 mL of cell spores to 135 mL of culture medium. One thousand milliliter flasks were incubated (250 rpm at 26°C) for 48 h (run 1) and 72 h (runs 2 and 3) in a shaker (model G25, New Brunswick Scientific, Edison, NJ). Runs 4 and 5 were repeated batches. The cultivated cells were then used to inoculate the fermentation broth. The inoculum seed was 10% in volume.

The experiments were carried out in a stirred aerated bioreactor with a 5-L working volume (Bioflo II-C; New Brunswick Scientific) at 26°C. Fermentation runs were carried out for about 120 h and samples were taken periodically. The experiments started in batch mode and after depletion of glucose, supplementary medium was added to the fermentor at suitable flow rates. The supplementary medium contained the same components as the production medium, but instead of glucose, hydrolyzed sucrose (glucose + fructose) was used as a carbon source *(20)*. Cell mass and the concentrations of glucose and cephalosporin C were measured off-line.

A paramagnetic oxygen analyzer (model 755; Rosemount Analytical) and an infrared carbon dioxide analyzer (model 880A; Rosemount Analytical) were used to monitor the effluent gas. The analyzers were calibrated with standard gases (5% CO_2/95% N_2, and 99.99% N_2).

A Supervisory Data Acquisition and Control System (Unisoft, São Paulo, Brazil) was used. It includes a Programmable Logic Controller (model 340; GE Fanuc) supervised by a personal computer. The manipulated variables were airflow rate, stirring, cooled water flowrate, and feed rate of supplementary medium. Temperature and dissolved oxygen were the controlled variables. The carbonate buffer was able to maintain the pH.

Analytical Procedures

Cell mass concentration was evaluated in terms of dry weight at 105°C, in grams/liter. Glucose concentration was measured utilizing the enzymatic GOD-PAP method *(21)*. Cephalosporin C concentration was determined by high-performance liquid chromatography (Waters), with a Nova Pak® C-18 column, isocratic operation; eluent; phosphate buffer (1.36% KH_2PO_4, pH 6.0); and 100/3 (v/v) KOH/CH_3CN. Cephalosporin C standards were kindly provided by Laboratório de Tecnologia Enzimática do Instituto de Catalisis y Petróleo Química do CSIC, Madrid, Spain.

Database for Training and Validation of Hybrid Model

Five experimental runs were used for training and validation of the hybrid model. The experiments were conducted in fed-batch and repeated fed-batch modes. On-line measurements included carbon dioxide and oxygen concentrations in the exhausted gas. Table 1 gives the experimental conditions of each data set (runs 1–5). The data-sampling interval was 10 s, but for training purposes a 10-min period was used. The output variables of the ANNs were compared with "experimental" rates during the training procedure. These rates were obtained after smoothing the concentration time course of the state variables. A mechanistic model was used for this purpose *(20)*.

Network Architecture

Two distinct phases are present during fed-batch cephalosporin C production. During the trophophase there is a rapid consumption of substrate to produce biomass. The synthesis of antibiotic occurs during the

Table 1
Experimental Conditions of Bioassays

Run no.	Incubation time in shaker (h)	Inoculum seed (% [v/v])	Flow rate of supplementary medium (mL/h)	Feed starting time (h)	$Y_{x/s}$ (g cell/ g glucose)	pH range
1[b]	48	10	10.00	66.50	0.41	6.3–7.5
2[c]	72	20	10.00	45.00	0.58	6.5–6.9
3[c]	72	20	10.75	44.66	0.58	6.1–6.8
4[b,d]	—	15	13.25	22.75	0.58	6.3–6.6
5[c,d]	—	15	12.00	21.25	0.51	6.3–8.3

[a]T = 26°C; Q_{air} = 3.0 SLPM (standard liter per minute), dissolved oxygen controlled at 40% of saturation, manipulating the stirring speed.
[b]Runs used as validation data set.
[c]Runs used in training data set.
[d]Runs 4 and 5 were sequential repeated fed batch of run 3.

idiophase, when cell growth becomes insignificant. During this phase, glucose should be fed at low flow rates. An increase in carbon dioxide concentration and a decrease in oxygen concentration in the exhausted gas *(22)* accompany the rapid biomass formation in the trophophase. Therefore, these two on-line variables were chosen as inputs for the ANN. Our strategy also included the specific growth rate at five previous sampling times as additional inputs to the net, as suggested by Hernández and Arkun *(23)*. Hence, the topology of the first ANN included 17 inputs, 1 hidden layer with 5 nodes and 1 output, the specific growth rate at the present time (Fig. 1A).

The second network inferred the specific glucose consumption rate at the present time. Inputs to this network were concentrations of carbon dioxide and oxygen in the exhaust gas, specific growth rate (provided by the first ANN), and specific glucose consumption in five previous time intervals (Fig. 1B).

The third network provided the specific production rate at the present time. The inputs to this network were specific growth rate (from first network) and specific production rate in five previous time intervals (Fig. 1C). The three FNNs had five nodes in the hidden layer. The topology of the FNNs was defined by trial and error.

This architecture follows the "modularization" concept of Di Massimo et al. *(7)*. The information provided by one network (the specific growth rate) was used by another two ANNs (which predicted the specific glucose consumption and antibiotic production rates). When we tried a single network to predict the three rates simultaneously, the results were not promising.

The neuron activation function was sigmoidal *(24)*. The objective function during the training phase was the sum of the squares of the deviation between the neural network outputs and experimental outputs, for all data in the set. The training algorithm was the classic backpropagation of

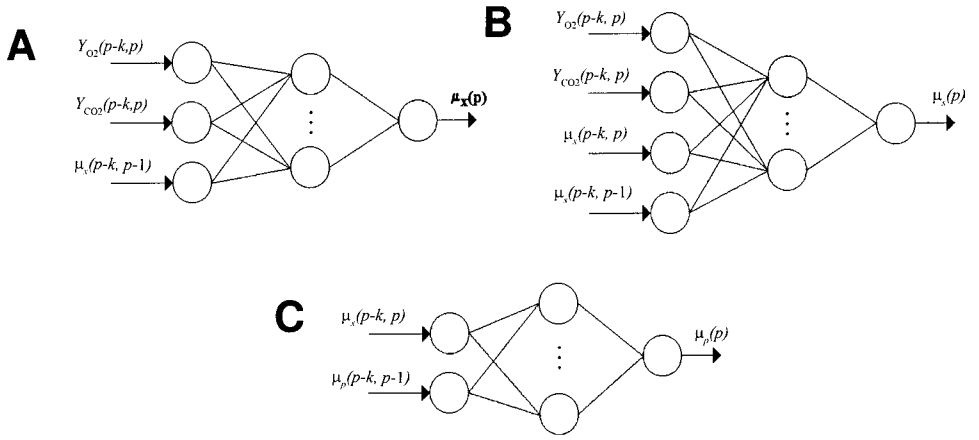

Fig. 1. FNNs used in this study. **(A)** First FNN, to estimate the specific growth rate; **(B)** second FNN, to infer the specific glucose consumption rate; and **(C)** third FNN, for specific production rate. p, present time; k is set to five previous instants; $Y_{O_2}(p-k, p)$, oxygen mole fractions in the exhausted gas; $Y_{CO_2}(p-k, p)$, carbon dioxide mole fractions in the exhausted gas; t, time (h); $\mu_x(p-k, p-1)$, specific growth rates (h^{-1}); $\mu_s(p-k, p-1)$, specific glucose consumption rates (h^{-1}); $\mu_p(p-k, p-1)$, specific product formation rates (h^{-1}).

Rumelhart and McClelland *(14)*. The mean square error (MSE) between the neural network output and the learning data was the objective function for the training algorithm. After 300 presentations of the data set, the MSE did not decrease significantly and the training was interrupted, to avoid overfitting.

Hybrid Model

The mass balance equations of the main components are as follows:

$$\frac{dC_x}{dt} = \mu_x \times C_x \tag{1}$$

$$\frac{dC_s}{dt} = -\mu_s \times C_x \tag{2}$$

$$\frac{dC_p}{dt} = \mu_p \times C_x \tag{3}$$

in which C_p is cephalosporin C concentration (g/L), C_s is glucose concentration (g/L), C_x is cell concentration (g/L), μ_s is specific substrate consumption rate (h^{-1}), μ_x is specific growth rate (h^{-1}), and μ_p is specific production rate (h^{-1}).

The neural networks illustrated in Fig. 2 provided the three specific rates. It is important to stress that μ_s, μ_x, and μ_p are apparent rates, which

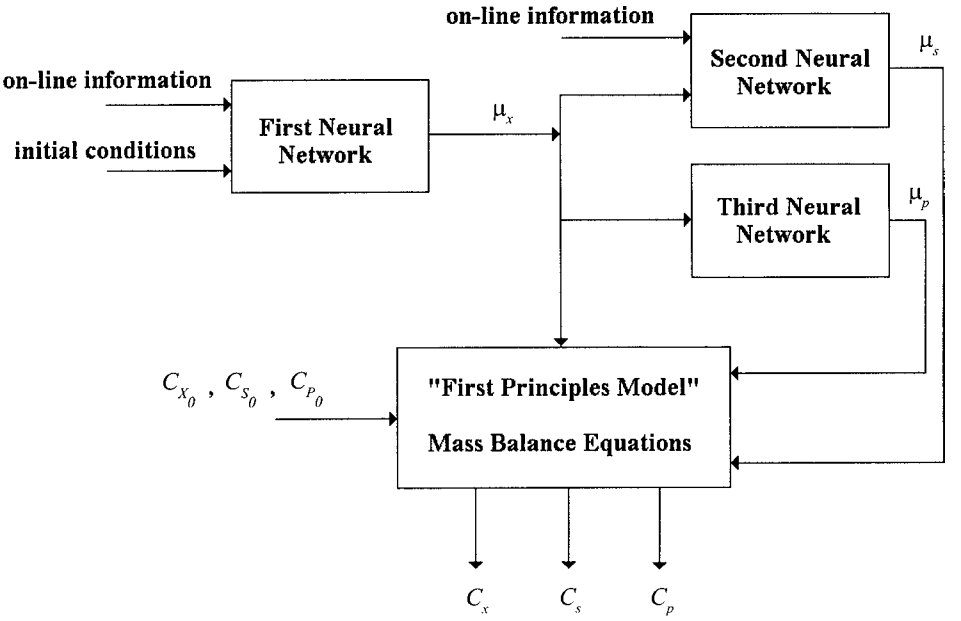

Fig. 2. Schematic hybrid model. Three ANNs estimate three specific rates (cellular growth, glucose consumption, and cephalosporin C production), that are used in the mass balance equations. The rates at the five initial sampling times (initial conditions of the first FNN) were equal to zero.

take into account the process kinetics and the dilution effect caused by the fed-batch operation. This is a characteristic feature of the hybrid approach; the ANNs were trained to lump these two phenomena in one single rate.

Results and Discussions

The network results for μ_x and μ_s for the runs used for training are illustrated in Fig. 3. The two networks provided fairly good estimates at the three different conditions. The experimental rates were obtained by smoothing the raw data. These results showed that the ANNs captured the essential characteristics of the nonlinear process response.

Figure 4 exhibits the performance of the third neural network, inferring the specific production rate for the three training data sets. Figure 5 shows the inferred variables for the training set. The model follows accurately the concentration time course for all the calculated variables.

Run 1 has some distinct features (refer to Table 1) that make it very appropriate to validate our inference algorithm. During this run the lag phase was considerably longer (approx 24 h), owing to the different inoculation procedure. Initial glucose concentration was also higher. And, most important, the overall yield for run 1 was $Y_{X/S} = 0.41$ (g of cell/g of glucose), whereas for all other runs it was in the range $Y_{X/S} = 0.51$–0.58. Therefore, using this specific run to validate the hybrid model is a very strong test of its extrapolation capabilities.

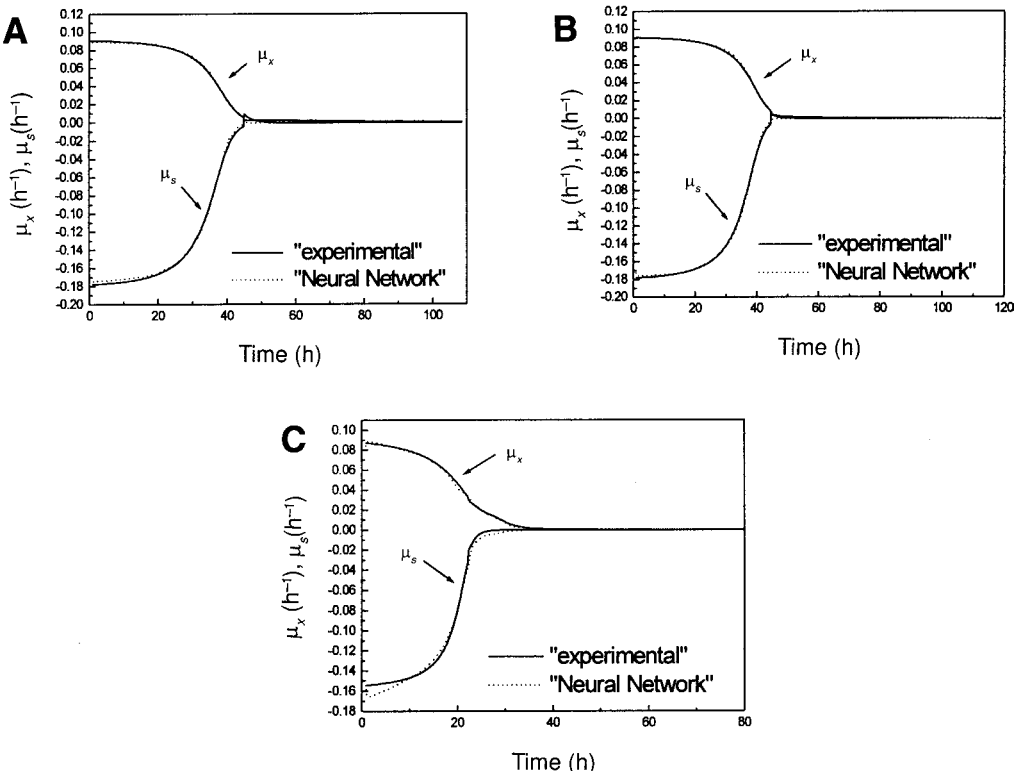

Fig. 3. Specific growth and consumption rates obtained by neural networks during the training phase. **(A)** Run 2, fed batch with standard inoculum preparation; **(B)** run 3, fed batch with standard inoculum preparation; **(C)** run 5, repeated fed batch.

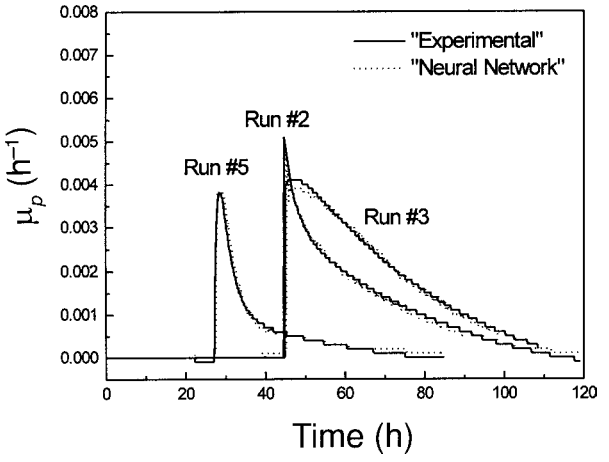

Fig. 4. Specific production rate time course. Run 2, fed batch with standard inoculum preparation; run 3, fed batch with standard inoculum preparation; run 5, repeated fed batch.

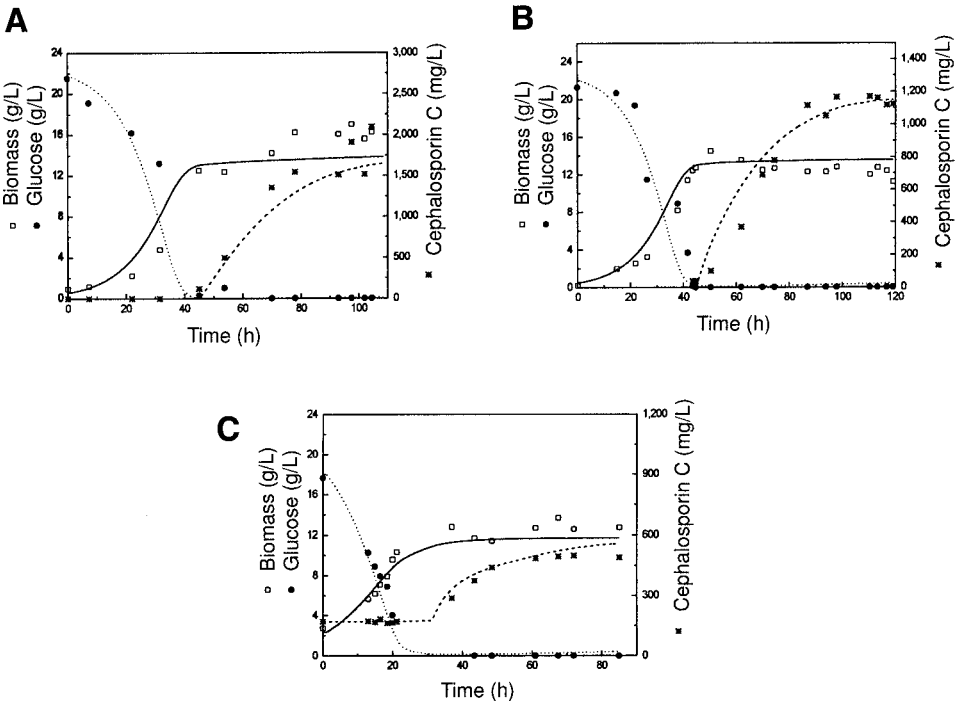

Fig. 5. Hybrid model fitting to the training data sets. **(A)** Run 2, fed batch with standard inoculum preparation; **(B)** run 3, fed batch with standard inoculum preparation; **(C)** run 5, repeated fed batch.

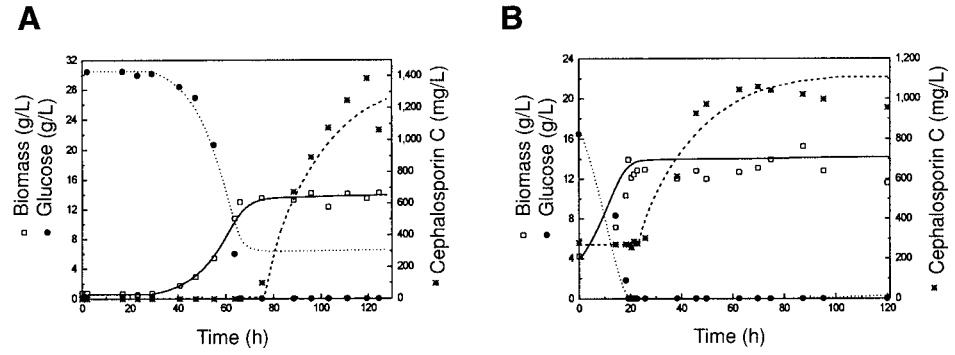

Fig. 6. Hybrid model validation test. **(A)** Run 1, fed batch with cell growth delay; **(B)** run 4, repeated fed batch.

Figure 6 displays the results of the validation tests. Run 4 was not used for the FNN training, and therefore it may be used as a validation set. The inference of cell mass and product concentration was very accurate in both cases. As for glucose concentration, an offset was observed for run 1 after 65 h. Actually, this is an expected result, since the FNN that describes μ_s had no previous information about the possibility of such a different

yield coefficient occurring. Nevertheless, the overall performance of the hybrid model was very good with respect to cell and product concentrations. These are the key state variables to support the decision of interrupting the fermentation and are not measured on-line. Run 4, displayed a very good fitting for all variables.

Despite the offset in final glucose concentration observed in run 1, the hybrid model is superior to the white box approach. When using a first-principles model, such as the one described by Cruz et al. *(20)*, it would be necessary to reset the parameter $Y_{X/S}$ and to establish a time condition to trigger the beginning of the cellular growth. In other words, a conventional, white box model whose parameters were estimated from runs 2, 3, and 5 would not be able to simulate the trend of the bioreactor during run 1. It is important to note that the predictive capabilities of the hybrid model might be further enhanced if run 1 were added to the training set.

Conclusion

ANNs were used to estimate reaction rates from on-line data, which were then used in the conservation equations. Our results indicate that, despite the different lag phase duration that occurred for run 1, a hybrid neural network model was able to capture the complex dynamics of the cephalosporin C production bioprocess.

The hybrid model could predict with accuracy the depletion of glucose at the end of the trophophase and, therefore, may be used to trigger the beginning of the supplementary feed.

Our architecture was based on the modularization concept, in which one ANN provides information to the next one. A single ANN was not able to simultaneously infer cell growth, glucose consumption, and product formation rates.

Acknowledgments

We gratefully acknowledge the financial support of CAPES and CNPq and a scholarship from FAPESP (São Paulo State Foundation, Brazil).

References

1. Savidge, T. A. (1984), in *Biotechnology of Industrial Antibiotics*, vol. 22, Vandamme, E. J., Marcel Dekker, NY, pp. 171–224.
2. Matsumura, M., Imanaka, T., Yoshida, T., and Taguchi, H. (1981), *J. Ferment. Technol.* **59(2),** 115–123.
3. Chu, W. B. Z. and Constantinides, A. (1988), *Biotechnol. Bioeng.* **32,** 277–288.
4. Basak, S., Velayudhan, A., and Ladisch, M. R. (1995), *Biotechnol. Prog.* **11,** 626–631.
5. Araujo, M. L. G. C., Oliveira, R. P., Giordano, R. C., and Hokka, C. O. (1996), *Chem. Eng. Sci.* **51(11),** 2835–2840.
6. Thibault, J., Breusegem, V. V., and Chéruy, A. (1990), *Biotechnol. Bioeng.* **36,** 1041–1048.
7. Di Massimo, C., Montague, G. A., Willis, M. J., Tham, M. T., and Morris, A. J. (1992), *Comput. Chem. Eng.* **16(4),** 283–291.

8. Karim, M. N. and Rivera, S. L. (1992), *Comput. Chem. Eng. Suppl.* S369–S377.
9. Syu, M.-J. and Tsao, G. T. (1993), *Biotechnol. Bioeng.* **42,** 376–380.
10. Cruz, A. J. G., Araujo, M. L. G. C., Giordano, R. C., and Hokka, C. O. (1998), *Appl. Biochem. Biotecnol.* **70–72,** 579–592.
11. Warnes, M. R., Glassey, J., Montague, G. A., and Kara, B. (1998), *Neurocomputing* **20,** 67–82.
12. Psichogios, D. C. and Ungar, L. H. (1992), *AIChE J.* **38(10),** 1499–1506.
13. van Can, H. J. L., te Breake, H. A. B., Hellinga, C., Luyben, K. C. A. M., and Heijnen, J. J. (1997), *Biotechnol. Bioeng.* **54(6),** 549–566.
14. Rumelhart, D. E. and McClelland, J. L. (1986), in *Parallel Distributed Processing*, vol. 1, Massachusetts Institute of Technology, Cambridge, pp. 318–362.
15. Ruck, D. W., Rogers, S. K., Kabrisky, M., Maybeck, P. S., and Oxley, M. E. (1992), *IEEE Trans. Pattern Anal. Machine Intell.* **14(6),** 686–691.
16. Thompson, M. L. and Kramer, M. A. (1994), *AIChE J.* **40(8),** 1328–1340.
17. van Can, H. J. L., te Breake, H. A. B., Bijman, A., Hellinga, C., Luyben, K. C. A. M., and Heijnen, J. J. (1999), *Biotechnol. Bioeng.* **62(6),** 666–680.
18. Shen, Y.-Q., Wolfe, S., and Demain, A. L. (1986), *Bio/Technology* **4,** 61–63.
19. Demain, A. L., Newkirk, J. F., and Hendlin, D. (1963), *J. Bacteriol.* **85,** 339–344.
20. Cruz, A. J. G., Silva, A. S., Araujo, M. L. G. C., Giordano, R. C., and Hokka, C. O. (1999), *Chem. Eng. Sci.* **54,** 3137–3142.
21. Trinder, P. (1969), *Ann. Clin. Biochem.* **6,** 24.
22. Silva, A. S., Cruz, A. J. G., Araujo, M. L. G. C., and Hokka, C. O. (1998), *Braz. J. Chem. Eng.* **15(4),** 320–325.
23. Hernández, E. and Arkun, Y. (1992), *Comput. Chem. Eng.* **16(4),** 227–240.
24. Bhat, N. and McAvoy, T. J. (1990), *Comput. Chem. Eng.* **14(4/5),** 573–583.