Biotechnology and Bioprocess Engineering **2006**, 11: 351-356

An Artificial Neural Network for Biomass Estimation from Automatic pH Control Signal

Won Hur* and Yoon-Keun Chung

Department of Bioengineering and Technology, College of Engineering, Kangwon National University, Chunchon 200-701, Korea

Abstract This study developed an artificial neural network (ANN) to estimate the growth of microorganisms during a fermentation process. The ANN relies solely on the cumulative consumption of alkali and the buffer capacity, which were measured on-line from the on/off control signal and pH values through automatic pH control. The two input variables were monitored on-line from a series of different batch cultivations and used to train the ANN to estimate biomass. The ANN was refined by optimizing the network structure and by adopting various algorithms for its training. The software estimator successfully generated growth profiles that showed good agreement with the measured biomass of separate batch cultures carried out between at 25 and 35°C.

Keywords: on-line monitoring, software sensor, alkali consumption, buffer capacity, biomass

INTRODUCTION

Biomass is one of the most important parameters in industrial fermentation processes [1]. The continuous monitoring of the cell mass could provide relevant information for determining the level of contamination, seed development, and the optimum time for the induction or harvesting time. Therefore, many sensors have been introduced for the in-line or *in situ* monitoring of this variable. However, there is no generally applicable method for the on-line monitoring of the biomass concentration in a process environment [2]. Many papers have reported the continuous on-line measurement of the cell density or the monitoring of biomass, based mainly on the optical methods such as turbidometry [3] or fluorescence measurements [4,5]. The techniques for estimating the biomass such as agitation speed of DO-stat culture [6], the capacitance of the culture broth [7], and impedance measurements [8] have also been reported. However, these methods often require sophisticated instrumentation, which require calibrations and laborious maintenance.

It has been suggested that software sensors can be used to estimate the state of the bioprocess. This estimation is based on mathematical algorithms that can produce estimates of the unmeasured variables from those measured using hardware sensors [9]. In estimating the biomass, other variables such as the dissolved oxygen level, pH and/or exit gas can be measured and mathematically processed using the chemical stoichiometry [10], mathematical models [11,12], Kalman filters [13,14], or artificial neural networks (ANN) [15,16]. An adaptive estimator and the use of a Kalman filter rather than a growth-model-based estimation have been suggested due to the inevitable change in culture conditions between different batches. Artificial neural networks have also been suggested to be an attractive alternative for modeling nonlinear fermentation systems on account of their ability to learn process non-linearity directly from plant data [17].

The successful estimation of biomass using an ANN requires a good set of input variables to produce the proper input-output mapping of the ANN [18]. It would be better to choose an input variable that is more directly related to the process as long as it is measurable. It is expected that the level of alkali consumption, which represents the metabolic activity of the microbial culture would be directly related to the level of biomass. We recently reported a method for the on-line measurement of the buffer capacity and alkali consumption rate through the simple modification of an automatic pH controller in a jar fermentor and by data acquisition of the pH control signal [19]. A previous study suggested that the alkali consumption rate is proportional to the specific growth rate during batch cultivation, and the buffer capacity of a culture medium is determined by the ionic components that dissociate at the culture pH. Therefore, this paper proposes an ANN to estimate the biomass using the variables derived from an automatic pH control signal.

***Corresponding author**
Tel: +82-33-250-6276 Fax: +82-33-243-6350
e-mail: wonhur@kangwon.ac.kr

MATERIALS AND METHODS

On-Line Measurement of Buffer Capacity and Alkali Consumption Rate

Automatic pH control during microbial cultivation generates a typical pH pattern, as shown in Fig. 1A. This pH pattern was generated using a pulse-type feeding of a concentrated solution of alkali or acid or gaseous ammonia, in order to minimize the dilution of the fermentation broth. Hence, the pH of a medium fluctuates between a lower and upper bound. Fig. 1B shows a schematic diagram of an alkali pump feeding and the dynamics of the resulting pH variations. Therefore, the alkali consumption rate and buffer capacity can be derived from the pH response patterns. Using the knowledge of the pump feeding rate (F_{pump}), alkali concentration (N_{alk}), culture volume (V), and pulse length for a pump operation (T_{pump}), the normality of the alkali fed ($[OH]_{fed}$) for a pH adjustment can be calculated using the following formula:

$$[OH]_{fed} = \frac{N_{alk} \cdot F_{pump} \cdot T_{pump}}{V} \quad (1)$$

The net increase in pH as a result of the alkali feed can be estimated from an extrapolation of the decreasing part of the slope of the pH response, as shown in Fig. 1B. Therefore, the buffer capacity of a culture medium can be calculated by dividing the concentration of the alkali fed by the net increase in pH.

$$\beta = \frac{[OH]_{fed}}{\Delta pH} \quad (2)$$

Culture Conditions

The *E. coli* cells were maintained in a 20% (v/v) glycerol stock at −80°C, after growing in a Luria-Bertani (LB) medium (10 g/L yeast extract, 5 g/L tryptone, and 10 g/L NaCl). The culture medium for the batch cultivation contained the following compositions per liter: 0.5 g of KH_2PO_4, 5 g of NH_4Cl, 5 g of NaCl, 0.4 g $MgSO_4$, 3 g of yeast extract, 10 g of casamino acid, and 20 g of glucose. The cultivations were performed in a 5 L jar fermentor (BK5L, Boksung Eng. Korea) that was equipped with a flat-blade impeller, a pH electrode, and pH measurement and control unit. The control unit was modified to monitor the rate of alkali consumption and the buffer capacity of the culture broth.

Analytical Methods

The level of cell growth was monitored by measuring the absorbance at 680 nm for conversion to the dry cell weight. A conversion table was prepared from several samples by measuring the absorbance, as well as by weighing the saline-washed cells after drying them in an oven at 105°C for 24 h. The residual glucose level in the culture medium was analyzed using the DNS method [20]. The acetic acid concentration was determined using a gas chromatograph (3400cx, Varian, Palo Alto, CA, USA) that was equipped with a Gaschrom 220 column and an FID detector.

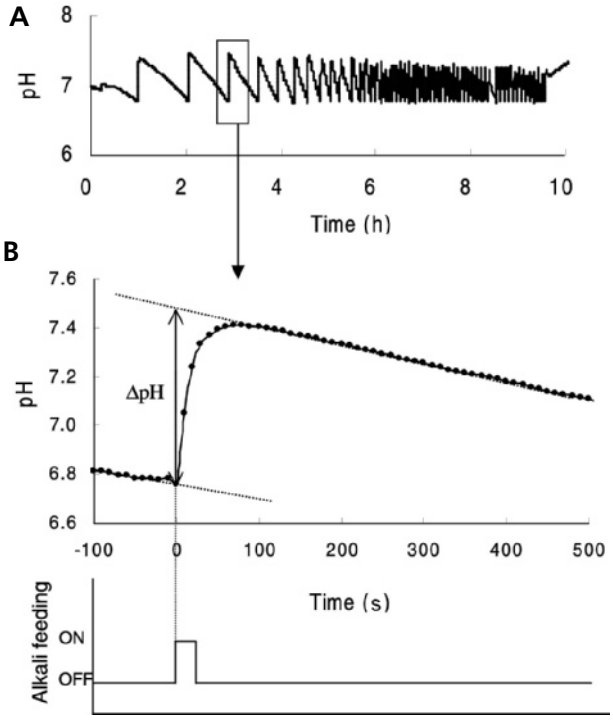

Fig. 1. Schematic for obtaining the ΔpH profile generated by an automatic pH controller in response to the pulse-type feeding of alkali.

Artificial Neural Network

An ANN software package (Quicknet V2.1, ICSI, CA, USA) was used to construct the network topology by changing the node parameters. The ANN was trained by loading the data file for the buffer capacity, cumulative alkali consumption, alkali consumption rate, and cultivation time. The input and output parameters were normalized and the weight parameters of the ANN were optimized using various methods. The trained ANN software generates the estimated output and the root mean square error. The result of the matching between the predicted values by ANN and the measured values was evaluated statistically based on the coefficient of determination.

RESULTS AND DISCUSSION

Batch Cultivations and On-Line Measurements of the Derived Variables

A series of batch cultivations were carried out using five different cultivation conditions in order to train the ANN to estimate the levels of cell growth, as shown in

Table 1. The environmental conditions used in the batch cultivations of *E. coli* for ANN training and evaluation

	No.	Media	Glucose supplement	PH set point	Temperature (°C)	Agitation (rpm)	Aeration (L/min)
Training runs	1	LB + 5 g/L Glucose	60 g/L	6.5	37	300	1.0
	2	LB + 10 g/L Glucose		6.5	37→30	300	1.5
	3	LB + 10 g/L Glucose		6.5	37→30	300	1.5
	4	LB + 20 g/L Glucose	20 g/L	6.5	37	300	1.5
	5	LB + 20 g/L Glucose		6.5	37→25→30→37	300	0.5∼1.5
Test runs	1	LB + 40 g/L Glucose	20 g/L	6.5	37	200∼300	1.5
	2	LB + 40 g/L Glucose		6.5	37→32	200	0.5

Table 1. The cultivation conditions were varied by changing the initial glucose level, culture temperature, agitation, and aeration rate. For example, one batch was grown at 37°C, while another was grown at 30°C, which was then shifted down to 25°C in the mid stage of the exponential phase. The aeration was also varied from 0.5 to 1.5 vvm. The cultivation of *E. coli* was carried out in a 5 L jar fermentor that was modified to allow monitoring of the pH control signals. Since the intention was to develop an ANN that can utilize the parameters derived from pH control signals, the cumulative alkali consumption, its rate and buffer capacity were monitored on-line during the cultivations.

Design of Artificial Neural Network

A good set of input variables is essential for the proper input-output mapping of an ANN. Therefore, combinations of the measured variables and the cultivation time were tested as input variables using an ANN with varying numbers of neurons on a single layer. When the ANN was trained using the cumulative alkali consumption and buffer capacity, the estimated values were within 88% of the measured biomass levels. The other combinations of input variables produced poorer results. For example, the ANN with 3 input variables, the buffer capacity, alkali consumption rate, and cultivation time, produced valued that were only 40∼50% of the measured values. As a result, the buffer capacity and alkali consumption was found out to be the best combination of input variables.

However, the combination of culture time, alkali consumption, and its rate showed the worst result, as shown in Fig. 2, which did not include the buffer capacity as an input variable. This suggests that the buffer capacity is a key input parameter. A previous study demonstrated that the cumulative alkali consumption is usually proportional to the level of cell growth and organic acid production [19]. The organic acid also increases the buffer capacity of the culture medium. Therefore, alkali consumption can be used to estimate the production of biomass provided that it is compensated for by the buffer capacity.

Therefore, two input variables, buffer capacity and alkali consumption, were used to construct an ANN for biomass estimation. The optimum topology of the ANN was examined by varying the number of nodes and the

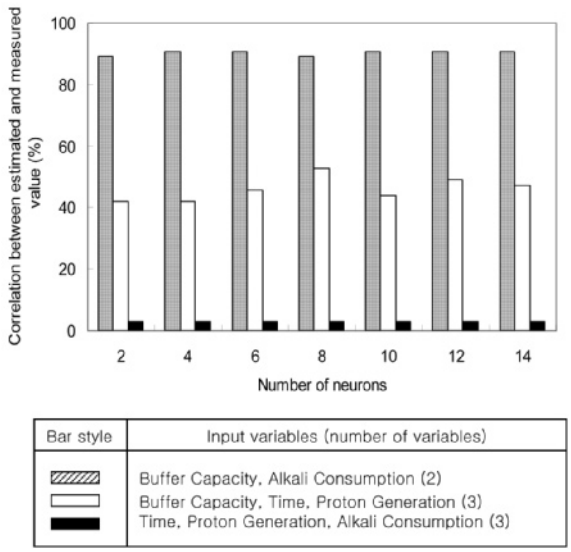

Fig. 2. The selection of input variables by a single layer ANN with different numbers of neurons.

transfer functions *i.e.* sigmoid, hyperbolic tangent, and Gaussian. Fig. 3A gives a comparison of estimated and measured values. The best fit of 92.2% was obtained through the application of a sigmoid function with a 2-3-3-1 topology of the ANN and by the application of a hyperbolic tangent function with a 2-3-4-1 and 2-4-4-1 topology. However, the additional hidden layers did not improve the ANN. Therefore, an ANN with the simplest topology of 2-3-3-1 of Fig. 3B with a sigmoid function was used in the subsequent experiments.

Training of the ANN and Estimation of Biomass

The constructed ANN was trained using the variables measured on-line and the biomass as the output data measured off-line. Various training algorithms, On-line backprop, Batch backprop, Rprop, Backprop-rand, Delta-bar-delta, and Quickprop, were tested. Fig. 4 shows the Root Mean Square (RMS) error for the training algorithms ranging from 2,000 to 10,000 iterations. The training of the ANN was carried out using a back-propagation algorithm and the RMS error was reduced to

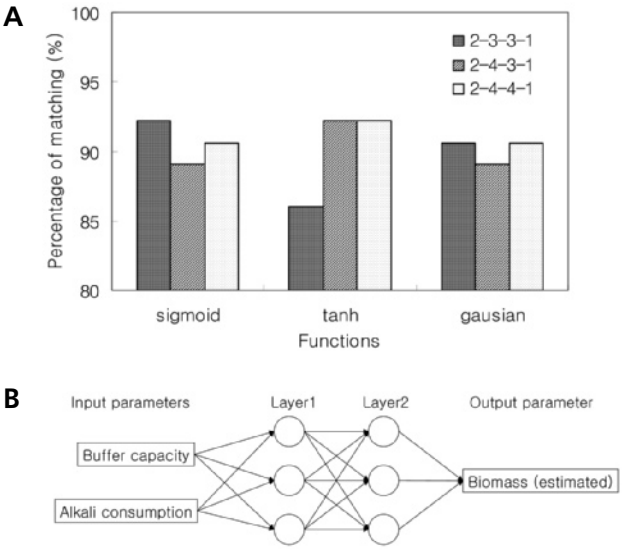

Fig. 3. Fine refinement of the ANN topology and transfer function and the optimized topology.

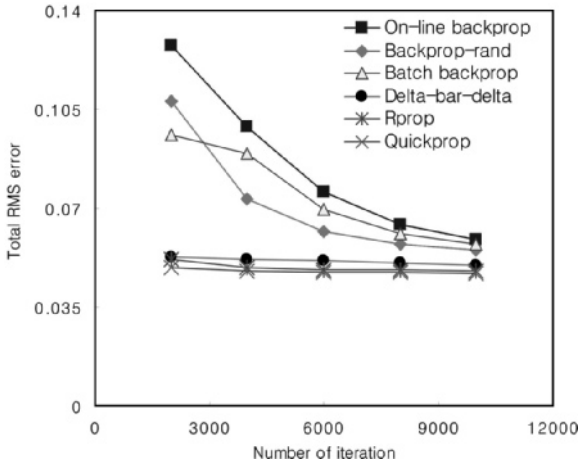

Fig. 4. Training of the ANN using various algorithms.

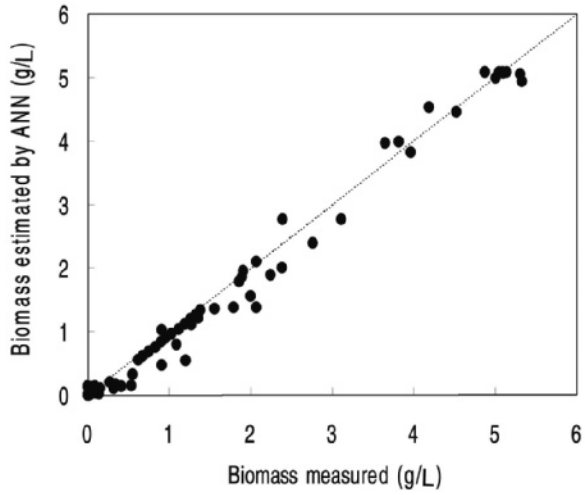

Fig. 5. Comparison of the biomass estimated by the ANN with those measured off-line.

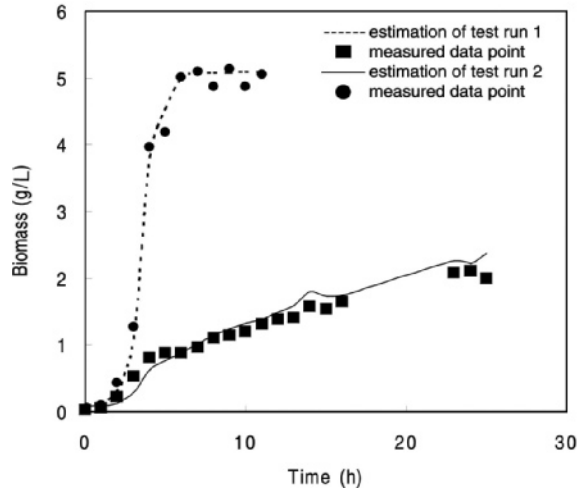

Fig. 6. ANN estimation of the biomass (run 1, ·····; run 2, ——) compared with the off-line measured values (run 1, ■; run 2, ●) cultivated under two different conditions.

0.047 after 10,000 iterations. The learning rate was significantly enhanced by applying the Quickprop algorithm, which gave the lowest RMS error of 0.049, even at less than 2,000 iterations. The levels of biomass estimated by the ANN were compared with the values measured off-line. The estimated values generated by the ANN showed good agreement with the biomass values measured off-line for five cultivations (Fig. 5).

Two different batch cultivations were carried and the ANN estimated the biomass. The batch cultivations were performed under significantly different environmental conditions. The culture temperature in one batch was 37°C, and the temperature of the other batch was shifted to 32°C after 5 h of cultivation at 37°C. The growth profiles were significantly different but the estimated profile agreed well with the levels of biomass measured off-line (Fig. 6).

The results of this study demonstrated that an ANN utilizing the pH control signal successfully estimated the biomass levels during the batch cultivation under significantly different environmental conditions. The correct choice of input variables is essential for the successful estimation of the variables of interest. The input variables must carry at least the indirect information of the output variable for a successful estimation even in an ANN. The two input variables, buffer capacity and cumulative alkali consumption, are expected to increase with the growth of cells. However, the buffer capacity is affected by the production of organic acid during cultivation. For example, acetic acid is one of the major organic acids produced by *E. coli*. However, acetate production depends largely on the cultivation conditions, as well as on the specific growth rate [21]. *E. coli* produces a significant amount of acetate under anaerobic conditions, and even under aerobic conditions when high glucose levels are present in

the medium [22]. Therefore, the cumulative alkali consumption and buffer capacity can provide a mixed signal consisting of metabolic activity and acetate production.

There are software sensors that have been reported to successfully estimate the biomass as well as the other variables of bioprocess. These sensors usually require the carbon evolution rate (CER) or/and oxygen utilization rate (OUR) as input variables in combination with the mass balance [10], mathematical models [11,12], a model with a Kalman filter [13,14], and software technologies. Artificial neural networks [15,16] and more advanced neuro-fuzzy algorithm [23] have been adopted using the CER and OUR as input variables. Hence, the exit-gas composition is essential for software sensors. However, one study used the alkali consumption rate as the sole input variable for estimating the substrate and biomass [24]. This report demonstrated that the buffer capacity and the level of alkali consumption are an appropriate set of input variables to give an acceptable estimation of the levels of biomass grown under significantly different conditions.

It should be noted that the on-line estimation of the biomass does not require any sophisticated analytical instrumentation except for a minor modification of the existing pH controller. The use of the pH response signal simply means the on-line measurement of the buffer capacity and cumulative alkali consumption during cultivation. The online measurement was used as a part of the control software in combination with the ANN. It is possible that the software sensor for biomass estimation does not require frequent calibration and manual labour. This is because the pH probe, which is one of the few reliable and steam-sterilizable sensors for fermentation, was already inserted into the broth [1]. The pH value is a reliable process parameter that can be used to monitor industrial fermentation processes.

The use of the pH control signal is an interesting target because fermentation usually requires automatic pH control to maintain the pH of the culture medium within a specified range. The pH control is generally applied in order to automatically supply the substrate for a fed-batch culture in a pH-stat model [25]. The use of pH measurements to estimate the product level and the level of biomass in aerobic yeast fermentation has been reported [26]. The pH monitor was also used to probe the metabolic shift that occurs at the IPTG induction of recombinant protein production [27]. In this study, it was demonstrated that pH control enables process control and gives an estimation of the key parameters during microbial growth. For further development in utilizing pH monitoring, metabolic flux analysis or reverse engineering [28] of pH change could be applied for developing more advanced software sensor for probing cell growth and its metabolism.

CONCLUSION

This study demonstrated the utility of the pH response signal for the on-line measurement of the buffer capacity and alkali consumption, was further used to estimate the biomass during cultivation. The pH response signal was implemented as a software sensor and a minor modification of the existing pH controller was made. It should be noted that the on-line estimation of biomass does not require any sophisticated analytical instrumentation. Because the pH probe is one of the few reliable and steam-sterilizable sensors for fermentation and is already inserted into the broth, the estimation also can be a reliable process parameter and be applied for monitoring various industrial fermentation processes. Combinations of input variables such as the alkali consumption rate, alkali consumption, buffer capacity, and culture time were used but the use of the alkali consumption and buffer capacity as input parameters produced the best results.

REFERENCES

[1] Schügerl, K. (1991) Common instruments for process analysis and control. pp. 6-25. In: R. J. Rehm, G. Reed, A. Puhler, and P. Stadler (eds.). *Biotechnology 4*. VCH Publishers Inc., New York, NY, USA.

[2] Olsson, L. and J. Nielsen (1997) On-line and *in situ* monitoring of biomass in submerged cultivations. *Trends Biotechnol*. 15: 517-522.

[3] Salgado, A. M., R. O. M. Folly, and B. Valdman (2001) Biomass monitoring by use of a continuous on-line optical sensor. *Sens. Actuators B: Chem.* 75: 24-28.

[4] MacMichael, G., W. B. Armiger, J. F. Lee, and R. Mutharasan (1987) On-line measurement of hybridoma growth by culture fluorescence. *Biotechnol. Tech*. 1: 213-218.

[5] Li, J. K. and A. E. Humphrey (1991) Use of fluorometry for monitoring and control of a bioreactor. *Biotechnol. Bioeng*. 37: 1043-1049.

[6] Na, J.-G., H. H. Kim, and Y. K. Chang (2005) On-line estimation of cell growth from agitation speed in DO-stat culture of a filamentous microorganism, *Agaricus blazei*. *Biotechnol. Bioprocess Eng.* 10: 571-575.

[7] Ferreira, A. P., L. M. Vieira, J. P. Cardoso, and J. C. Menezes (2005) Evaluation of a new annular capacitance probe for biomass monitoring in industrial pilot-scale fermentations. *J. Biotechnol*. 116: 403-409.

[8] Harris, C. M. and D. B. Kell (1983) The radio-frequency dielectric properties of yeast cells measured with a rapid, automated, frequency-domain dielectric spectrometer. *Bioelectrochem. Bioenerg*. 11: 15-28.

[9] Bogaerts, P. h. and R. Hanus (2001) On-line state estimation of bioprocesses with full horizon observers. *Math. comput. Simul.* 56: 425-441.

[10] Chattaway, T., A. L. Demain, and G. Stephanopoulos (1992) Use of various measurements for biomass estimation. *Biotechnol. Prog.* 8: 81-84.

[11] Di Massimo, C., P. A. Lant, A. Saunders, G. A. Montague, M. T. Tham, and A. Morris (1992) Bioprocess application of model-based estimation technique. *J. Chem. Technol. Biotechnol*. 53: 265-277.

[12] Maher, M., G. Roux, and B. Dahhou (1995) A method for estimating the state variables and parameters of fermenta-

tion systems. *J. Chem. Technol. Biotechnol*. 63: 153-159.
[13] Thibault, J., V. V. Breusegem, and A. Cheruy (1990) On-line prediction of fermentation variables using neural networks. *Biotechnol. Bioeng.* 36: 1041-1048.
[14] Pons, M. N., A. Rajab, J. M. Flaus, J. M. Engasser, and A. Cheruy (1988) Comparison of estimation methods for biotechnological processes. *Chem. Eng. Sci.* 43: 1909-1914.
[15] Shi, Y. and W.-K. Yuan (1988) Application of adaptive estimation in microbial fermentation processes. *Chem. Eng. Sci.* 43: 1915-1920.
[16] Di Massimo, C., G. A. Montague, M. J. Willis, M. T. Tham, and A. J. Morris (1992) Towards improved penicillin fermentation via artificial neural networks. *Comp. Chem. Eng.* 16: 283-291.
[17] James, S., R. Legge, and H. Budman (2002) Comparative study of black-box and hybrid estimation methods in fed-batch fermentation. *J. Process Control* 12: 113-121.
[18] Baughman, D. R. and Y. A. Liu (1995) *Neural Networks in Bioprocessing and Chemical Engineering.* pp. 1-29. Academic Press, San Diego, CA, USA.
[19] Jung Y.-K. and W. Hur (1997) Analysis of pH change and an automatic pH control with a new function: on-line estimation of acetic acid. *Biotechnol. Bioprocess Eng*. 90: 580-582.
[20] Miller, G. L. (1959) Use of dinitrosalicylic acid reagent for determination of reducing sugar. *Anal. Chem.* 31: 426-428.
[21] Luli, G. W. and W. R. Strohl (1990) Comparison of growth, acetate production, and acetate inhibition of *Escherichia coli* strains in batch and fed-batch fermentations. *Appl. Environ. Microbiol*. 56: 1004-1011.
[22] Han, K. (1992) *A Study of Acetic Acid Formation in* Escherichia coli *Fermentation*. Ph.D. Thesis. University of California, Irvine, CA, USA.
[23] Beluhan, D. and S. Beluhan (2000) Hybrid modeling approach to on-line estimation of yeast biomass concentration in industrial bioreactor. *Biotechnol. Lett.* 22: 631-635.
[24] Feitkenhauer, H. and U. Meyer (2004) Software sensors based on titrimetric techniques for the monitoring and control of aerobic and anaerobic bioreactors. *Biochem. Eng. J*. 17: 147-151.
[25] Suzuki, T., T. Yamane, and S. Shimizu (1990) Phenomenological background and some preliminary trials of automated substrate supply in pH-stat model fed-batch culture using a set point of high limit. *J. Ferment. Bioeng.* 69: 292-297.
[26] San, K. and G. Stephanopoulos (1984) Studies on on-line bioreactor identification. Ⅳ. Utilization of pH measurements for product estimation. *Biotechnol. Bioeng*. 26: 1209-1218.
[27] Hur, W. and Y.-K. Chung (2005) On-line monitoring of IPTG induction for recombinant protein production using an automatic pH control signal. *Biotechnol. Bioprocess Eng*. 10: 304-308.
[28] Lee, D. (2005) Component-based software architecture for biosystem reverse engineering. *Biotechnol. Bioprocess Eng.* 10: 400-407.

[Received April 12, 2006; accepted June 26, 2006]