ORIGINAL PAPER

Lei Zhi Chen · Sing Kiong Nguang · Xue Mei Li
Xiao Dong Chen

Soft sensors for on-line biomass measurements

Abstract One of the difficulties encountered in control and optimisation of bioprocesses is the lack of reliable on-line sensors for their key state variables. This paper investigates the suitability of using on-line recurrent neural networks to predict biomass concentrations. Input variables of the proposed recurrent neural network are feed rate, liquid volume and dissolved oxygen. Experimental results revealed that the proposed neural network is able to predict biomass concentrations with an accuracy of $\pm 11\%$.

Introduction

Bioprocesses are the most complex in all fields of process engineering due to high dimension, nonlinearity and dynamical characteristics. Ability to control bioprocesses accurately and automatically at optimal states is of considerable interests to many fermentation industries because it can reduce production costs and increase yield while at the same time maintaining the quality of the metabolic product [1, 2, 3, 4]. However, designing a control system for bioprocesses is not a straightforward task due to: (i) significant uncertainty in the model, (ii) lack of reliable on-line sensors which can accurately detect the key state variables, (iii) nonlinear and time varying nature of the process, and (iv) slow response of the process, in particular for cell and metabolic concentrations. This work is mainly focused on the second

above mentioned problem. Throughout the years, the key state variables, such as biomass concentrations in fermentation processes are usually measured off-line with long measurement delay. This limits the range of control algorithms that can be applied to the process. Recently, artificial neural networks (ANNs) have drawn considerable attention to the development of on-line soft sensors [5, 6, 7, 8].

ANNs are computational systems whose architecture and operation are inspired from our knowledge about biological neural cells (neurons) in the brain. These are not simulation of real neurons in the sense that they do not model the biology, chemistry, or physics of a real neuron. They do, however, model several aspects of the information combining and pattern recognition behaviours of real neurons in a simple yet meaningful way. Neural networks have an incredible capability for emulation, analysis, prediction, and association. They are able to solve difficult problems in a way that resembles human intelligence. What is unique about neural networks is their ability to learn by example. However, ANNs can and should be retrained on or off-line whenever new information becomes available. Recurrent neural networks (RNNs), a member of the ANN family, have proven to be a valuable tool and are extensively used in modelling and control of nonlinear dynamic systems [9, 10, 11, 12].

Apart from the selection of neural networks, another important issue is the selection of appropriate state variables to be measured online. Soft sensors work in a manner of cause and effect; the inherent biologic relation between measured and unmeasured states could significantly affect the prediction accuracy. Carbon dioxide, pH, ethanol and dissolved oxygen (DO) can be easily measured online in a research laboratory using standard sensors. Among them, dissolved oxygen, which reflects the fundamental level of energy transduction in bioreaction, is intricately linked to cellular metabolism. Nor et al. [13] studied the online application of dissolved oxygen and mass balances used to estimate the specific growth rate of a fed-batch culture of *Kluyveromyces*

L. Z. Chen · S. K. Nguang (✉)
Department of Electrical and Electronic Engineering,
The University of Auckland, Private Bag,
92019 Auckland, New Zealand
E-mail: sk.nguang@auckland.ac.nz

X. M. Li · X. D. Chen
Department of Chemical and Materials Engineering,
The University of Auckland, Private Bag,
92019 Auckland, New Zealand

fragilis. Dissolved oxygen was also employed in [14] to detect the Acetate formation in *Escherichia coli*. In [15], neural networks have been used to relate the increase of biomass concentration to the increase of lactic acid concentration. This approach requires an additional sensor for measuring the lactic acid concentration.

This paper investigates the suitability of using RNNs to predict on-line biomass concentrations in a fermentation process. RNN input variables are feed rate, liquid volume and dissolved oxygen. All inputs are chosen because they can easily be measured on-line. Output of the RNN is the biomass concentration. Selection of a suitable RNNs topology is done by data generated from a mathematical model; the topology is then re-trained by experimental data. The layout of the remainder of the paper includes determination of a soft sensor structure in Sect. 2, experimental results in Sect. 3, and conclusions in Sect. 4.

Soft sensor structure determination and implementation

RNNs are chosen to estimate the biomass because of their strong capability to capture dynamic information underlying the input-output data pairs. The structure of the proposed neural soft sensor is given in Fig. 1. This neural network consists of one hidden layer, one output neuron, feedforward paths, feedback paths and tapped delay lines (TDLs). In order to enhance dynamical behavior of the sensor, outputs from the output layer and the hidden layer are fed back through TDLs. The output of the ith neuron in the hidden layer is of the form:

$$
h_i(t) = f_h\left(\sum_{j=0}^{n_a} W_{ij}^I p(t-j) + \sum_{k=1}^{n_b} W_{ik}^R \hat{y}(t-k) \right.
$$

$$
\left. + \cdots + \sum_{l=1}^{n_c} W_{il}^H h_l(t-1) + b_i^H \right) \tag{1}
$$

where p is the neural network input, $\hat{y}$ is the neural network output and h is the hidden neuron's output.
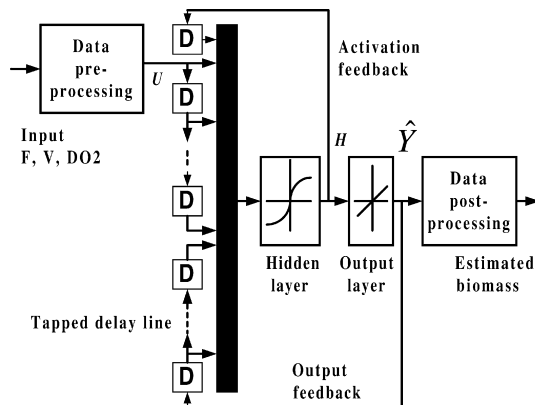
b_i^H is the bias of the ith hidden neuron. n_a, n_b, n_c are the number of input delays, the number of output feedback delays and the number of hidden neurons, respectively. f_h is a sigmoidal function, W_{ij}^I is the weight connecting the jth delayed input to ith hidden neuron, W_{ik}^R is the weight connecting the kth delayed output feedback to ith hidden neuron, and W_{il}^H is the weight connecting the lth hidden neuron output feedback to the ith hidden neuron.

Only one neuron is placed in the output layer, so the output is:

$$
\hat{y}(t) = f_Y\left(\sum_{m=1}^{n_c} W_m^Y h_m(t) + b_Y \right) \tag{2}
$$

where, f_Y is a pure linear function, W_m^Y is the weight connecting the mth hidden neuron's output to the output neuron, and b_Y is the output neuron bias.

A mathematical fermentation model given in [16] consists of six differential equations and is used to generate simulation data. Four different feed rate profiles are chosen to excite the mathematical fermentation model: the random step, the square wave, the saw wave and an industrial feeding policy-like sequence. They are shown in Fig. 2. Each feed rate profile yields 150 input-output (target) pairs corresponding to 6 min sampling time during a 15 hr fermentation.

A well-known fact for choosing the training data set is that it must cover the entire state space of the system as many times as possible. In this study, the random step, which excites the process the most, is used to generate the training data set. Before training an RNN, the inputs and target data have to be pre-processed (scaled) so that they are within a specified range [-1,1]. This specified range is the most sensitive area of the hidden layer activation function. In this case, the output of the trained network will also be in the range [-1,1].

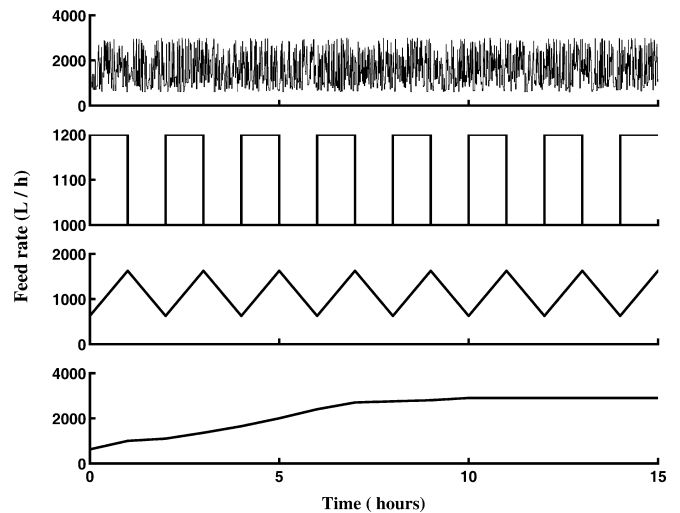Fig. 1 Structure of the proposed neural soft sensor

Fig. 2 Four different feed rate profiles

The performance function used for training and testing the neural networks is a percentage mean square error index [15], and is defined as:

$$E = \sqrt{\frac{\sum_{t=1}^{N}\left(X_t^m - \hat{X}_t\right)^2}{\sum_{t=1}^{N}\left(X_t^m\right)^2}} \times 100\% \qquad (3)$$

where N is the number of sampling data pairs, X_t^m is the measured (actual) value of biomass, and $\hat{X}_t$ is the corresponding estimated value predicted by the neural soft sensors.

The Levenberg-Marquardt backpropagation training algorithm is adopted to train the neural networks due to its faster convergence and memory efficiency [17, 18]. The algorithm can be summarised as follows:

1. Present input sequences to the network. Compute the corresponding network outputs with respect to the parameters $\mathbf{X}_k$ (i.e. weights and bias), the error $\mathbf{e}$ and the overall MSE.
2. Calculate the Jacobian matrix $\mathbf{J}$ through the backpropagating of Marquardt sensitivities from the final layer of the network to the first layer.
3. Update the network parameters using:

$$\Delta \mathbf{X}_k = -\left[\mathbf{J}^T(\mathbf{x}_k)\mathbf{J}(\mathbf{x}_k) + \mu_k\mathbf{I}\right]^{-1}\mathbf{J}^T(\mathbf{x}_k)\mathbf{e} \qquad (4)$$

where, μ_k is initially chosen as a small positive value (e.g., $\mu_k = 0.01$).
4. Recompute the MSE using $\mathbf{X}_k + \Delta \mathbf{X}_k$. If this new MSE is smaller than that computed in step 1, then decrease μ_k, let $\mathbf{X}_{k+1} = \mathbf{X}_k + \Delta \mathbf{X}_k$ and go back to step 1. If the new MSE is not reduced, then increase μ_k and go back to step 3.

The algorithm terminates when: (i) the norm of gradient is less than some predetermined value, or (ii) the MSE has been reduced to some error goal, or (iii) the μ_k is too large to be increased practically, or (iv) the predefined maximum number of iterations has been reached. During training, an early stopping method is useful to prevent a neural network from being overtrained. When random step data are used to train the network, another set of data with a different feed rate profile will be used as a validation data set. The error on the validation set is monitored during the training process. The validation error will normally decrease during the initial phase of training. However, when the network begins to over fit the data, the error on the validation set will typically begin to rise. When the validation error increases for a specified number of iterations, the training is stopped, and the weights and biases at the minimum of the validation error are obtained.

The rest of the data sets, which are not seen by the neural network during the training period, are used in examining the trained network. For each network structure, 50 networks are trained; the one that produces the smallest mean error of the test data sets is retained. The error between the network output and the target output is used to evaluate the "goodness" of the network. Errors for various combinations of input and output delays (hidden layer output feedback delay is fixed at 1) are shown in Fig. 3 from which can be seen that 12 hidden neurons networks frequently out-perform the 6 hidden neurons networks. The errors produced by the 0/4/1 structure (four output feedback delays, one hidden layer output feedback delay and no input delays) are smaller than others, and are very close for two different numbers of hidden neurons. The six hidden neurons network with the above structure is chosen for the online biomass estimation because it has less hidden neurons.

One of the simulation results is plotted in Fig. 4. The soft sensor provides a good prediction of the growth of biomass. The percentage mean square error of predic-
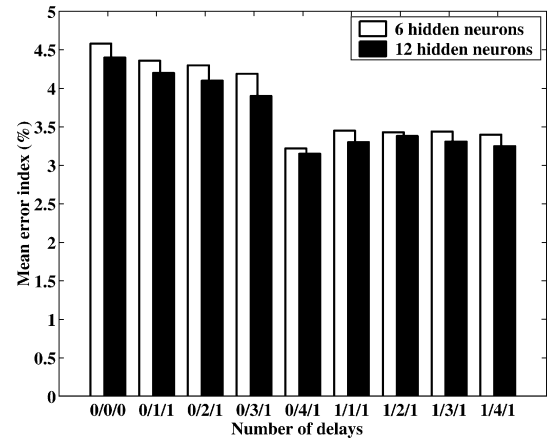
Fig. 3 Estimation mean error on testing data sets for neural networks with different combinations of delays. '0/4/1' indicates that the number of input delays = 0, the number of output feedback delays = 4, the number of hidden layer output feedback delays = 1
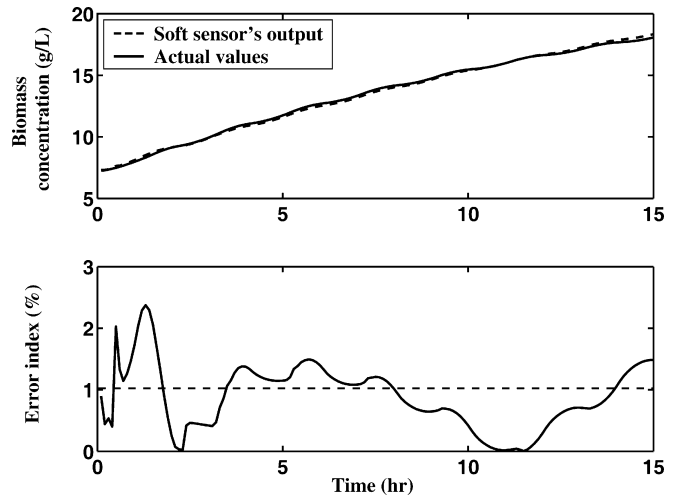
Fig. 4 Simulation result of biomass prediction using 6 hidden neuron network for a fed-batch fermentation process

tion is less than 3%. It should be mentioned that this prediction is based on the assumption that data sets for training, validation and testing are free from noise. However, this assumption is not true in real environments. Practically, errors in measurement, noisy input and output data may affect the accuracy of biomass estimation.

Experimental results

Yeast strain, *Saccharomyces cerevisiae*, produced by Goodman Fielder Milling and Baking N.Z. Ltd. was grown in a YEPD medium [19] with the following composition: Dextrose 20 g/L, Yeast extract 10 g/L, Peptone 20 g/L, and commercial anti-foam 10 drops/L. Starter culture was performed in the shaker at 30°C and 200 RMP for 60–90 min.

Three laboratory experiments with different feed rates (Fig. 2) have been carried out using 3-liter fermentors (New Brunswick Scientific Co., Inc., USA); see Fig. 5. These experiments are used to examine the suitability of the proposed soft sensor for real fermentation processes. Three sets of data are collected, 27 samples are taken during an 8 hr fermentation for each run. One set of data is used for re-training the neural network, one is used for validation and one is
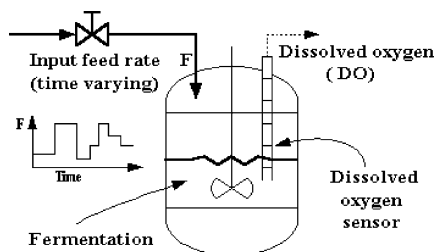
used for testing. Linear interpolation is used to make the sampling time equal.

One of the prediction trajectories is presented in Fig. 6. The network starts from an arbitrary initial point. As can be seen from Fig. 6, the soft sensor is able to converge within a very short time and can predict the trend of the growth of biomass, but it gives a high fluctuation in biomass estimation. In order to overcome this problem, two additional delayed inputs are added to the proposed network (RNN obtained in the previous section). As shown in Fig. 7, a smooth prediction has been achieved. However, the error is slightly higher than that in Fig. 6. By incorporating the hidden layer feedback (with unit delay, see Fig. 1) to the input of the network, the prediction mean error is reduced to 10.3%. The biomass estimation is given in Fig. 8. The experimental results show that the smallest percentage mean error is obtained from the neural soft sensor with four output feedback delays, one hidden layer output feedback delay and two input delays.
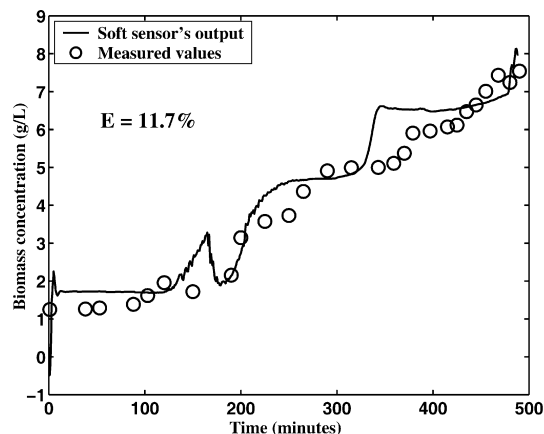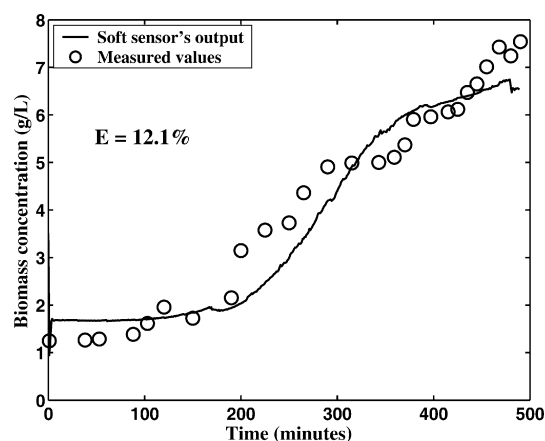

Fig. 7 On-line biomass prediction in a fed-batch baker's yeast fermentation process with two network input delays, four output feedback delays, and no hidden layer output feedback delays

Fig. 5 Experimental setup

Fig. 6 On-line biomass prediction in a fed-batch baker's yeast fermentation process with zero network input delays, four output feedback delays, one hidden layer output feedback delay
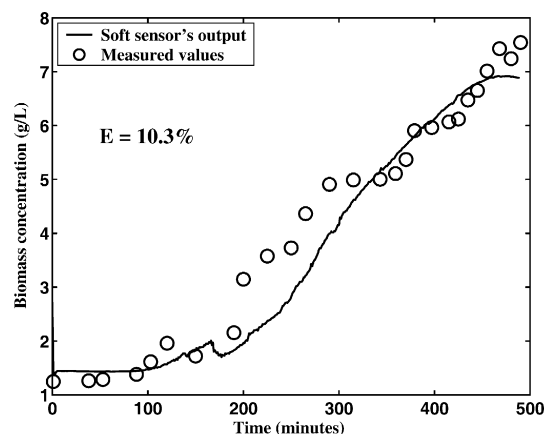

Fig. 8 On-line biomass prediction in a fed-batch baker's yeast fermentation process with two network input delays, four output feedback delays, and one hidden layer output feedback delay

Conclusion

This paper has investigated the suitability of using recurrent neural networks to predict biomass concentrations online. Inputs to the proposed recurrent neural network are the feed rate, liquid volume and dissolved oxygen. A suitable topology of the neural network is obtained via data generated by a mathematical model. The suitable topology of the RNN is then evaluated via experimental data. From the results obtained in both simulation and real processes, it can be concluded that RNNs are indeed a powerful tool for implementing an on-line biomass soft sensor for fermentation processes. The proposed neural network is able to predict the biomass concentration within 11% of its true value.

References

1. Schügerl K (2001) Progress in monitoring, modeling and control of bioprocesses during the last 20 years. J Biotechnol 85:149–173
2. Lee J, Lee S, Park S, Middeelberg A (1999) Control of fed-batch fermentations. Biotechnol Adv 17:29–48
3. Rani K, Rao V (1999) Control of fermentations—a review. Bioprocess Eng 21:77–88
4. Liu Y, Wang F, Lee W (2001) On-line monitoring and controlling system for fermentation process. Biochemical Eng J 7:17–25
5. Jose' de Assis A, Filho R (2000) Soft sensors development for on-line bioreactor state estimation. Comput Eng 24:1355–1360
6. Dacosta P, Kordich C, Williams D, Gomm J (1997) Estimation of inaccessible fermentation states with variable inoculum sizes. Artif Intell Eng 11:383–392
7. Teissier P, Perret B, Latrille E, Barillere J, Corrieu G (1997) A hybrid recurrent neural network model for yeast production monitoring and control in a wine base medium. J Biotechnol 55:157–169
8. Alippi C, Piuri V (1997) Neural modeling of dynamic systems with non-measurable state variables. IEEE instrumentation and measurement technology conference, Ottawa, Canada, May 1997
9. Chow T, Li X (2000) Modeling of continuous dynamical systems with input by recurrent neural networks. IEEE T Circuits-I 47:575–578
10. Shaw A, Doyle III F, Schwaber J A (1997) Dynamic neural network approach to nonlinear process modeling. Comput Chem Eng 21:371–385
11. Alippi C, Piuri V (1996) Experimental neural networks for prediction and identification. IEEE T Instrum Meas 45:670–676
12. Iatrou M, Berger T, Marmarelis V (1999) Modelling of non-linear nonstationary dynamic systems with a novel class of artificial neural networks. IEEE T Neural Netw 10:327–339
13. Nor Z, Tamer M, Scharer J, Moo-Young M, Jervis E (2001) Automated fed-batch culture of *Kluyveromyces fragilis* based on a novel method for on-line estimation of cell specific growth rate. Biochemical Eng J 9:221–231
14. Åkesson M, Karlsson E, Hagander P, Axelsson J, Tocaj A (1999) On-line detection of acetate fromation in *Escherichia coli* cultures using dissolved oxygen responses to feed transients. Biotechnol Bioeng 64:590–598
15. Acuña G, Latrille E, Béal C, Corrieu G (1998) Static and dynamic neural network models for estimating biomass concentration during thermophilic lactic acid bacteria batch cultures. J Ferment Bioeng 85:615–622
16. Berber R, Pertev C, Tarker M (1998) Optimization of feeding profile for baker's yeast production by dynamic programming. Bioprocess Eng 20:263–269
17. Hagan M, Demuth H, Beale M (1996) Neural network design. PWS, Boston
18. Scales L (1985) Introduction to non-linear optimization. Springer, Berlin Heidelberg New York
19. Atkinson B, Mavituna F (1991) Biochemical Engineering and Biotechnology Handbook, 2nd edn. New York