CHEMICAL ENGINEERING TRANSACTIONS Volume 18, 2009
Editor Jiří Klemeš

DOI: 10.3303/CET0918073

Development and implementation of an automated monitoring system for improved bioethanol production

Elmer Ccopa Rivera, Félix de Farias Junior, Daniel Ibraim Pires Atala, Rafael Ramos de Andrade, Aline Carvalho da Costa, Rubens Maciel Filho
School of Chemical Engineering, State University of Campinas, P.O. Box 6066, 13081-970, Campinas, SP, Brazil

This study presents results from the implementation and testing of an automated monitoring system for a bioethanol production process using MLP-based soft-sensors (MLP-SS). The system is based on an array of primary sensor, a communication module and a monitoring and data acquisition subsystem. This integrated framework provides a real-time monitoring solution, which is one of the most important aspects of the decision making in the strategies of optimization and control of bioprocesses.

1. Introduction

Although sensors for measuring state variables in real-time are available today, a review of the latest research has shown that analytical instruments are not robust in the industrial environment (Kano and Nakawa, 2008). In alcoholic fermentation processes these instruments are difficult to calibrate, mainly due to the characteristics of industrial culture media, such as turbidity of the culture, presence of dissolved CO_2, among others (Macedo, 2003).

On the other hand, Artificial Neural Networks (ANN) have been used successfully for solving biotechnological complex problems related to the field of measurements and instrumentation. Probably, the most popular ANN used in engineering applications is the Multilayer Perceptron Neural Network (MLP) due to its easily understandable architecture and a simple mathematical form, which results in an easy tool for modeling and implementation. In addition, it is known from previous works that MLPs can be used to offer adaptive solutions, since the reestimation of their parameters is a straightforward procedure (Rivera et al., 2007). By exploiting the relationship among the process variables of bioprocesses, MLPs could be used to implement advanced control techniques, such as software sensors, algorithms for on-line estimation of state variables and model parameters that are not measurable in real-time (Gonzaga et al., 2008, Lee et al., 2008). This is a promising research area with significant impact on biotechnological industry, which requires an efficient monitoring with reliable sensors to control the setting of the process. Thus, for a reliable performance prediction through modeling, the MLP-SS should be implemented on a platform able to provide a powerful toolset for process identification and control with a direct interface to instruments, sensors and actuators. A programme to fully automate the implementation of software

Please cite this article as: Rivera E.A.C., Farias F.J., Atala D.I.P., Ramos R.D.A., Carvalho A.D.A. and Maciel Filho R., (2009), Development and implementation of an automated monitoring system for improved bioethanol production, Chemical Engineering Transactions, 18, 451-456 DOI: 10.3303/CET0918073

sensors can be developed using a graphical programming environment, such as LabVIEW (Laboratory Virtual Instrument Engineering Workbench). LabVIEW has an extensive library of functions and subroutines for most programming tasks. It also contains an application specific library for data acquisition, serial instrument control, data processing, analysis presentation and storage. Applications created with LabVIEW are referred to as virtual instruments (VIs) (Alford, 2006).

In this study, built-in LabVIEW functions and library VIs were used to develop a software sensor based on MLPs for on-line monitoring of a bioethanol production process. This program can be adapted to a wide range of instrumentation, control and optimization applications. The proposed monitoring system is based on LabVIEW as a software driver, the data acquisition system and the sensors for the secondary measurements (input variables of the software sensor): pH; turbidity, Tb(%); CO_2 flow rate; F(m^3/h) and temperature, T(^oC). The system acquires process variables data through sensors. The measured values are delivered to the computer program through the data acquisition system for data processing and prediction of state variables (output variables of the software sensor): concentrations of biomass, X (g/L); substrate, S (g/L) and product, P(g/L).

2. Experiments

The microorganism used was *Saccharomyces cerevisiae* cultivated in the Bioprocess Engineering Laboratory in the Faculty of Food Engineering/State University of Campinas and obtained from an industrial fermentation plant. Eight experiments performed in a New Brunswick Scientific Bioflow III bioreactor (5 L working volume) were used to develop the MLP-SS. These experiments were carried out in a temperature range of 30-38^oC and an initial substrate concentration range of 71.7-168.0g/L. Another two batch experiments were used for the validation (at 36.8^oC and 133.0g/L) and prediction (at 31.2^oC and 156.0g/L) tests. Material and analytical methods for the determination of the state variables (X, S and P) are described elsewhere (Andrade et al., 2007).

On-line measurements (pH, Tb, F and T) were stored in the computer through a data acquisition board associated to a management software application. The sample time was 3 min for all on-line data. Carbon dioxide flow rate was measured by a digital gas volumetric flow sensor, pH by glass pH-electrode, (both from Cole-Parmer Instrument, London, England). Production medium turbidity was measured by a turbidity transmitter (FSC 402 Mettler Toledo Ingold Inc., USA) and temperature by thermocouple (N. Brunswick Scientific Co.).

3. Automated Monitoring System

In this study, an intelligent system is developed with primary on-line sensors, which capture large volumes of real-time bioprocesses data, and a model building software package that interprets the knowledge content in the stored data. As shown in Fig. 1, the proposed on-line monitoring system comprises three essential elements: (i) A bioreactor with on-line sensors (pH, CO_2 flow rate, temperature and turbidity). Such features as relatively simple instrumentation, short measuring time and low prices make the primary on-line sensors a right choice for the approach proposed in this study.

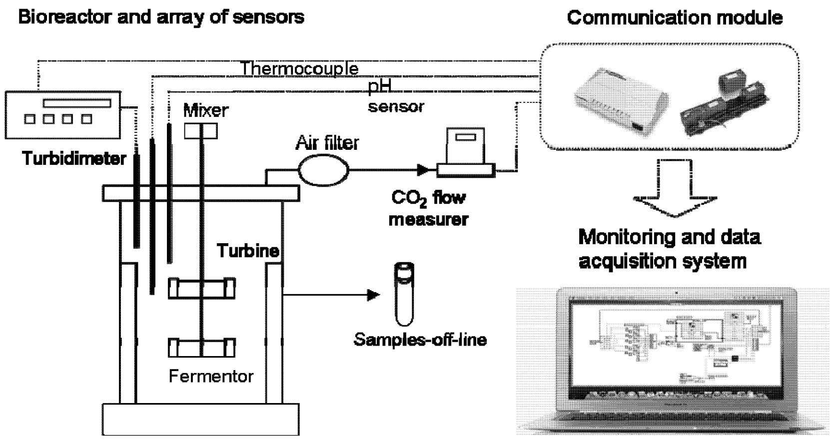

Figure 1. Framework of the automated monitoring system

(ii) A communication module that transfers the measured data from the bioreactor to a monitoring and data acquisition system. It consists of devices to convert the protocol of the output of the primary sensors (serial digital signals from gas volumetric flow sensor, pH-electrode and thermocouple) from RS232C to TCP/IP via a National Instruments ENET-232/4. The received analog signal from the turbidimeter (analog input 0-20 mA) is also transformed to standard TCP/IP using a Field Point network module 1600. TCP/IP protocol (Transmission Control Protocol/Internet Protocol) facilitates the communication across the internet, even remote access.
(iii) The monitoring and data acquisition system, which monitor the bioprocess based on the built-in LabVIEW functions, library VIs and MLP -based soft-sensor. Sampling and data acquisition from primary sensors (pH, CO_2 flow rate, turbidity and temperature) and postprocessing of these data are performed in the LabVIEW environment, which is a general purpose graphical programming environment.

3.1 MLP-based soft-sensor development

It is shown in Rivera et al. (2007) (Cybenko's Theorem) that the continuous functions which describe the dynamic behavior of an alcoholic fermentation can be approximated to any desired accuracy with a MLP neural network of one hidden layer of sigmoidal hidden neurons and a layer of linear output neurons. Such structure was selected as the neural paradigm. This neural network consists of three types of layers: an input layer, an output layer and one or more hidden layers, whose numbers of neurons are N, M and K, respectively. Each layer may have a different number of neurons which are interconnected by adjustable parameters (weights and biases) associated with them. The relationship is given mathematically as:

$$y_j = f\left(\sum_{i=1}^{N} w_{ji} x_i + \theta_j\right) = \frac{1}{1+\exp^{A}} \tag{1}$$

$$A = -((w_{M1}pH) + (w_{M2}Tb) + (w_{M3}F) + (w_{M4}T) + (w_{M5}t) + \theta_M) \tag{2}$$

$$\text{Concentration} = G\left(\sum_{j=1}^{M} W_{kj} y_j + \beta_k\right), \text{(j = 1,...,M), (k = 1)} \quad (3)$$

where w_{ji} is the weight connecting the ith neuron in the input layer and the jth neuron in hidden layer. w_j is the bias of the jth neuron in the hidden layer. W_{kj} is the weight connecting the jth neuron in the hidden layer and the kth neuron in the output layer. W_k is the bias in the kth neuron in the output layer. $f(\cdot)$ and $G(\cdot)$ are the sigmoidal activation functions of the jth neuron in the hidden layer and of the kth neuron in the output layer, respectively. In the present study, three MLPs, one for each output (concentration of biomass, X, substrate, S, and ethanol, P), were adopted as an alternative to the use of one MLP-SS with three outputs. Tests in the current study and previous investigations (Rivera et al., 2007) have shown that this approach lead to better results avoiding unnecessarily complex structures.

Small random values are used to initialization of weights and biases. Subsequently, the standard backpropagation learning algorithm, based on a gradient descendent method implemented in FORTRAN is employed to train each network describing the concentrations (Eq. 3). In this study, both input and output data were normalized to the range [0.1, 0.9]. The number neurons in the hidden layer was varied from 10 to 70, and the optimal number chosen by the cross-validation criterion with the number of epochs fixed at 2000 for all the studied architectures. The neural network with sixty hidden nodes for describing X, forty hidden nodes for describing S, and twenty hidden nodes for describing P were found to present the lowest mean square error for the validation sample. The learning rate η, and the momentum coefficient α, used in this work were optimized both to be 0.95 in the backpropagation learning.

After training, the appropriate network architecture and the parameters set (weights and biases) are saved in a file. This information is used as an input to a program that is written in Phyton to an automatic conversion of the optimal MLP model into a program based LabVIEW formula node that is used to infer the concentrations of biomass, substrate and ethanol.

4. Results and Discussion

Figure 2 shows the corresponding profiles of on-line measurements (input variables) in the batch fermentation at 31.2^{o}C used for the prediction test.

The prediction test results are illustrated in Figure 3 and quantified through the R.S.D. (Residual Standard Deviation) (Andrade et al., 2007). It can be observed that the prediction of the software sensor has a good agreement with the experiment data and it has not been affected by noise on any input. The concentrations of biomass, substrate, and ethanol calculated using the resulting mathematical model presented deviations of 10.1, 25.4 and 17.8% from the experimental data, respectively. The results have shown that it is possible to accurately infer these concentrations using pH, turbidity, CO_2 flow rate and temperature on line measurements and a MLP-SS.

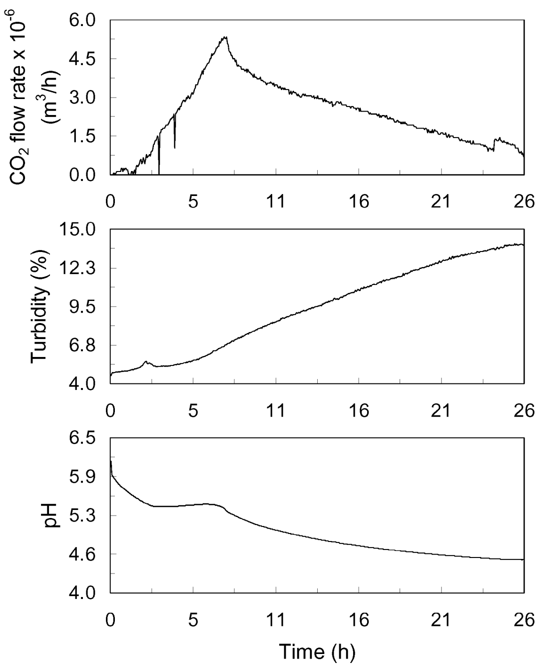

Figure 2. On-line primary sensor measurements at 31.2oC

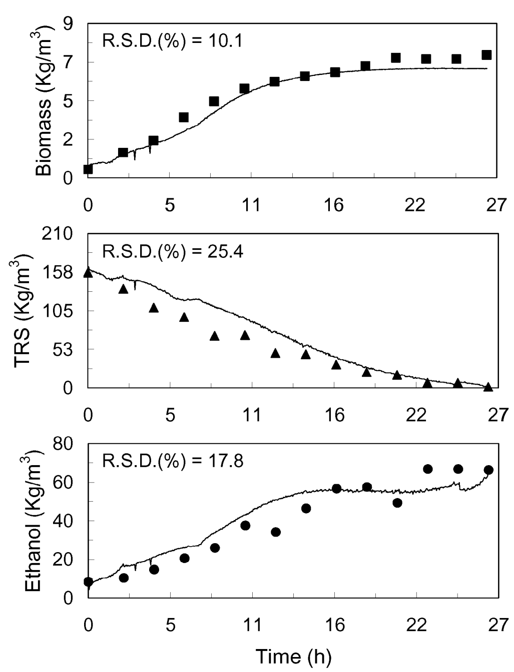

Figure 3. Experimental (cell mass, X (■); substrate (total reducing sugars), S (▲) and ethanol, P (●)) and software sensors (solid lines) results

5. Concluding Remarks

The application presented in this study illustrates the usefulness of an automated monitoring system carried out in LabVIEW environment. A data acquisition module is implemented to read all influencing variables, which are first used to train the MLPs. The optimal MLPs architecture is placed in a LabVIEW based program formula node that monitors the concentration of biomass, ethanol and substrate. The MLP-based monitoring system represents thus a robust model-based approach which is expected to contribute for improving the implementation of suitable operating strategies of optimization as well as advanced control to achieve high operational performance.

Acknowledgements

The authors acknowledge FAPESP (process number 06/51646-4), and CNPq for financial support.

References

Alford J.S., 2006, Bioprocess control: Advances and challenges, Comput. Chem. Eng., 30, 1464-1475.

Andrade R.R., Rivera E.C., Costa A.C., Atala D.I.P., Maugeri Filho F. and Maciel Filho R., 2007, Estimation of temperature dependent parameters of a batch alcoholic fermentation process, Appl. Biochem. Biotech., 136-140, 753-764.

Cybenko G., 1989, Approximation by super position of a sigmoidal function. Math Control Signal, 2, 303-314.

Gonzaga J.C.B., Meleiro L.A.C., Kiang C. and Maciel Filho R., 2008, ANN-based soft-sensor for real-time process monitoring and control of an industrial polymerization process, Comput. Chem. Eng., 33, 43-49.

Kano M. and Nakagawa Y., 2008, Data-based process monitoring, process control, and quality improvement: Recent developments and applications in steel industry, Comput. Chem. Eng., 32, 12-24.

Lee M.W., Hong S.H., Choi H., Kim J.H., Lee D.S. and Park J.M., 2008, Real-time remote monitoring of small-scaled biological wastewater treatment plants by a multivariate statistical process control and neural network-based software sensors, Process Biochem., 43, 1107-1113.

Macedo I.C., 2003, Estado da arte e tendências das tecnologias para energia. Brasília: Ministério da Ciência e da Tecnologia. CTEnerg Secretaria Técnica do Fundo Setorial de Energia. CGEE Centro de gestão e Estudos estratégicos (in Portuguese).

Rivera E.C., Costa A.C., Andrade R.R., Atala D.I.P., Maugeri Filho F. and Maciel Filho R., 2007, Developed of adaptive modeling techniques to describe the temperature dependent kinetics of biotechnological processes, Biochem. Eng. J., 36, 157-166.