Journal of Chemical Technology and Metallurgy, 49, 5, 2014, 473-478

A NEW APPROACH FOR MODELING THE BIOTRANSFORMATION OF CRUDE GLYCEROL BY USING NARX ANN

Elisaveta Kirilova, Sophia Yankova, Bilyana Ilieva, Natasha Vaklieva-Bancheva

Institute of Chemical Engineering,
Bulgarian Academy of Sciences,
"Acad. G. Bonchev" str. , bl. 103
Sofia 1113, Bulgaria
E-mail: eshopova@gmail.com

Received 10 November 2013
Accepted 04 July 2014

ABSTRACT

In this study a new approach for the modeling of a biotransformation process of crude glycerol to formation of 1,3-propanediol by the help of bacteria Pseudomonas denitrificans 1625 is proposed. Results obtained from the experimental studies of the process are presented for the first time. For the purpose of modeling dynamic Nonlinear AutoRegressive with eXogenous inputs – NARX Artificial Neural Networks are developed. They are used for prediction of the change in the time of the concentration of the substrate and the product as a function of their previous values. The models are trained and validated with real data samplings, obtained from parallel batch experiments carried out at three different initial concentrations of the substrate selected for this purpose. The obtained results show an acceptable coincidence between the measured and the predicted from the model values.

Keywords: modeling, NARX ANN, glycerol, 1,3-propanediol, Pseudomonas denitrificans 1625.

INTRODUCTION

Crude glycerol is obtained as a by-product from biodiesel production. Its utilization with strains of microorganisms leads to valuable chemical reagents as 1,3-propanediol.

Numerous experimental investigations of the process of biotransformation of crude glycerol to 1,3-propanediol by different bacteria strains have been carried out in order to evaluate their performance [1]. However, up to now there are no known results in the literature for the crude glycerol biotransformation by using a strain from the genera Pseudomonas. In this study, for the first time results obtained in the experimental investigation of the process of biotransformation of crude glycerol to 1,3-propanediol with Pseudomonas 1625 are presented.

Mathematical models have been developed to describe the process kinetics in a large range of values for feed glycerol concentrations into the medium [2 - 6].

Having in mind that the kinetics of biotransformation of crude glycerol to 1,3-propanediol, with the help of Pseudomonas 1625 is not clear enough, for the purpose of modeling, the application of models based on Artificial Neural Networks appears more suitable.

The aim of this study is to develop a dynamic recurrent ANN (Artificial Neural Networks) model for prediction of the obtained amounts of 1,3-propanediol and glycerol degradation, at different initial concentrations of the glycerol substrate.

The study is structured as follows: Section 2 provides description of the experimental work carried out. Section 3 is devoted to the developed ANN model of the process and its validation. Finally, short conclusions are formulated.

MATERIALS AND METHODS

Biotransformation of glycerol is affected by different factors such as: the specification of the producer strain; the composition of the used bacterial culture medium; and conditions under which the fermentation is carried

out like temperature, initial concentration of substrate, aeration, pH, etc.

Bacterial Strain. Pseudomonas denitrificans 1625 was used. It has a wide industrial application mainly in the production of vitamin B_{12}. An effective method for its production is by using of beet molasses, obtained as a by-product in the sugar industry [7].

Bacterial Culture medium. The bacterial strain of Pseudomonas denitrificans 1625 was cultivated batch-wise with two culture media for seeding and for operation. Their compositions are listed in Table 1.

Cultivation conditions. The culture media were sterilized in an autoclave at a pressure of 1 atm and temperature of 121^0C for 20 minutes. Inoculums have been grown for 24 hours on a rotary shaker at 200 rpm and temperature of 30ºC with stirring. During the process the initial concentration of glycerol has been varied in between 10 - 30 g/l.

Fermentation conditions. The main factors of the glycerol fermentation are: temperature, substrate concentration, aeration and pH. The fermentation of the glycerol was conducted at temperature of 30^0C. The process was carried out at initial concentrations of the glycerol of 10 to 30 g/l. The optimal pH for Pseudomonas sp. is in the range of 7.0 - 7.5.

Analytical methods. Qualitative and quantitative analysis of the dissolved product from the fermentation and the substrate glycerol by HPLC were carried out to determine metabolites concentrations.

The concentration of biomass was determined during the experiments by measuring the extinction of the sample at a wavelength of $\lambda = 660$ nm. For this purpose, a spectrophotometer „Specol" was used. The concentration of biomass was calculated from a calibration curve.

DYNAMIC ANN FOR FORECASTING OF 1,3-PROPANEDIOL FORMATION AT DIFFERENT INITIAL CONCENTRATIONS OF THE SUBSTRATE GLYCEROL

Formulation of the proposed model

Artificial Neural Networks are mathematical models inspired from the way the biological neurons transmit and process information. They are flexible computing techniques that can be applied to a wide range of time series, forecasting problems solution with a high degree of accuracy.

Generally, a feed-forward ANN consists of inputs (like synapses through which natural neurons receive the signals from other ones), outputs and one or more hidden layers with multiple neurons in them. Connections between them are modified by weights (strength of the respective signals), which represent the model

Table 1. Compositions of bacterial culture media.

Seed culture medium		Operation culture medium	
Component	Amount	Component	Amount
Yeast extract	1.0 g/l	$MgSO_4.7H_2O$	0.2 g/l
Pepton	10 g/l	$CaCl_2.2H_2O$	0.2 g/l
NaCl	10 g/l	NaCl	5.0 g/l
Agar	15 g/l	KH_2PO_4	379 ml
Distilled water	1000 ml	Na_2HPO_4	621 ml
Medium pH	7.2 - 7.4	Saline solution	5.0 g/l
		Glycerol	2-6 g/l

parameters. In addition, each neuron has an extra input that is assumed to have a constant value of one. The weight that modifies this extra input is called the bias. All data propagate along the connections in the direction from the network inputs to the network outputs, hence the term feed-forward ANN.

Then, the neurons of hidden layers aggregate these weighted values to single values, as follows:

$$net_i = \sum_j w_{i,j} \cdot x_j + b \qquad (1)$$

$$\forall i, i \in k, \quad \forall j, j \in l,$$

where:
i are indices for the neurons;
j are indices for the inputs;
$w_{i,j}$ are the weighted coefficients of input-to-hidden connections and hidden-to-output connections;
x_j are the inputs of the neural network;
b are the bias inputs for each one of the neurons from the hidden layers.

Then, an activation function is applied to the aggregated weighted value to produce an individual output for the specific neuron (like activated natural neuron which emits a signal through the axon which might be sent to another synapse, and might activate other neurons). For the purpose of modeling the following sigmoid function is used:

$$F(x)_i = \frac{1}{1 + e^{-a \cdot net_i}} \qquad (2)$$

where
a is a coefficient which determines the slope of sigmoid function. In our case we have chosen $a = 2$ as most frequently used value.

The most widely used feed-forward dynamic ANN models for the time series modeling and forecasting are Nonlinear AutoRegressive with eXogenous inputs – NARX models [8 - 9]. The NARX is a recurrent dynamic network, with feedback connections, enclosing several layers of the network. A two-layer NARX feed-forward is represented in Fig. 1. It is fed with a sequence of adjacent values of time series of input parameters forming a tapped delay line. This input is usually referred to as a time window since it provides a limited view on part of the series. The NARX also involves a delayed connection from the output of the second layer to input - a second tapped delay line.

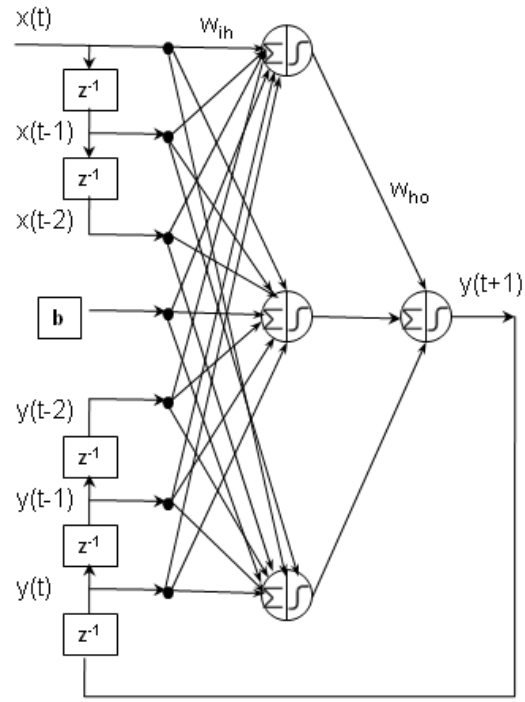

Fig. 1. The two-layer NARX feedforward ANN.

Thus, the general prediction equations for computing the next value of time series $y(t+1)$ as the NARX outputs, are functions of the past observations of the independent (exogenous) inputs $x(t)$ $x(t-1),...,x(t-n_x)$ and past values of $y(t)$ $y(t-1),...,y(t-n_y)$ which are already considered as inputs:

$$y(t+1) = F(y(t)\ y(t-1),...,y(t-n_y)\ x(t-1),...,x(t-n_x)\ + v(t) \qquad (3)$$

where:
n_y and n_x - the maximum time delays for the inputs;
$v(t)$ - the noise of the data.

Using a least-square function (LSF) as an optimization criterion, the weights of the artificial neurons are adjusted in a way so as the required outputs for specific inputs are thus obtained to minimize the optimization criterion. For this purpose a powerful optimization algorithm should be used. This process is called supervised learning of ANN.

The performance of ANN is influenced substantially

by the number of inputs and time delays for each one of the inputs, as well as its architecture, i.e. the number of hidden layers and neurons in each hidden layer.

Determination of the inputs and outputs of the model

For the purpose of the ANN modeling, data obtained from the parallel experiments conducted at three initial concentrations of crude glycerol, respectively 10 g/l, 20 g/l and 30 g/l, have been used. Concentrations of the product - 1,3-propanediol and the degraded crude glycerol were measured in the interval of 20 - 24 h during the experiments.

The ANN modeling of the process of biotransformation of crude glycerol aims to predict the next values of the concentrations of both, the glycerol and the product 1,3-propanediol, as a function of their past concentrations values. Due to the fact that the process is led in the initial stage of biomass growth, the concentration values of the biomass are very small and almost unchangeable. For that reason they are used as exogenous (independent) inputs, which permit us to select the required inputs and outputs for the modeling. They are listed in Table 2.

Using available data, the patterns for training and validation of ANN models have been selected.

ANN Architecture

Based on the available data and number of inputs and outputs, we have selected the most promising architecture consisting of one input layer with 3 inputs, one hidden layer with 5 neurons and one output layer with 2 outputs. The dynamic recurrent ANN is constructed so as to be fed with 7 time series of the inputs forming tapped delay line of 1.

Training and validation of ANN

We have implemented the BASIC genetic algorithm [10] to obtain the values of the weighting coefficients at which the criterion LSF has a minimal value.

As a result of multiple calculations, two equally optimal solutions with different values of ANN model parameters have been selected. These solutions were validated with selected data samplings. The Root Mean Square Error (RMSE) was used to estimate their efficiency and ability to make precise predictions:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(P_i - Q_i)^2} \quad (4)$$

where P_i and Q_i are calculated and measured values for the concentrations of the glycerol and 1,3-propanediol, respectively, and n is the number of data for the time series.

The RMSE values obtained for the two best solutions are summarized in the Table 3.

RMSEs summarized in Table 3, show that the biotransformation of the glycerol is modeled better than the 1,3-propanediol formation.

Fig. 2 demonstrates the comparison between measured and calculated data for the three groups of experiments at the concentrations of the substrate - 10 g/l, 20 g/l and 30 g/l.

It can be seen from the Fig. 2 that the degradation of the glycerol is modeled better at initial concentration of the glycerol of 10 and 20 g/l, than at 30 g/l.

Table 2. Inputs and outputs for ANN modelling.

Inputs	Outputs
Concentrations of the biomass, g/l - Exogenous (independent) inputs, generated by using first delay line	-
Past in the time series values of the concentrations of the glycerol, g/l, generated by using tapped delay line	Next in the time series values of the concentrations of the glycerol concentration, g/l
Past in the time series values of the concentrations of the 1,3-propanediol, g/l, generated by using tapped delay line	Next in the time series values of the concentrations of the 1,3-propanediol, g/l

Table 3. Values of RMSE for obtained optimal solutions.

7 time series	RMSE	
	Glycerol, g/l	1,3-propanediol, g/l
Solution 1	2.3953	0.7788
Solution 2	2.4058	0.7784

Concerning the 1,3-propanediol formation, a very good coincidence between measured and calculated data is seen at 30 g/l initial concentration of the glycerol, while at initial concentrations of 10 and 20 g/l, a good coincidence between measured and calculated data for the product formation exists only in the beginning of the process until the first 24 hours. After that a discrepancy between measured and calculated data is observed. We accept that the latter is due to the lack of information about the process of biotransformation, because the process is very new and not well studied.

CONCLUSIONS

In this study, the process of biotransformation of crude glycerol to 1,3-propanediol by help of the bacteria Pseudomonas denitrificans 1625 is discussed for the first time. Short description of the operational conditions is provided. The results obtained from the parallel experiments conducted at three different initial concentrations of the glycerol were used for its mathematical modeling. Taking into account that the process is not sufficiently studied an ANN model has been used. Dynamics of the

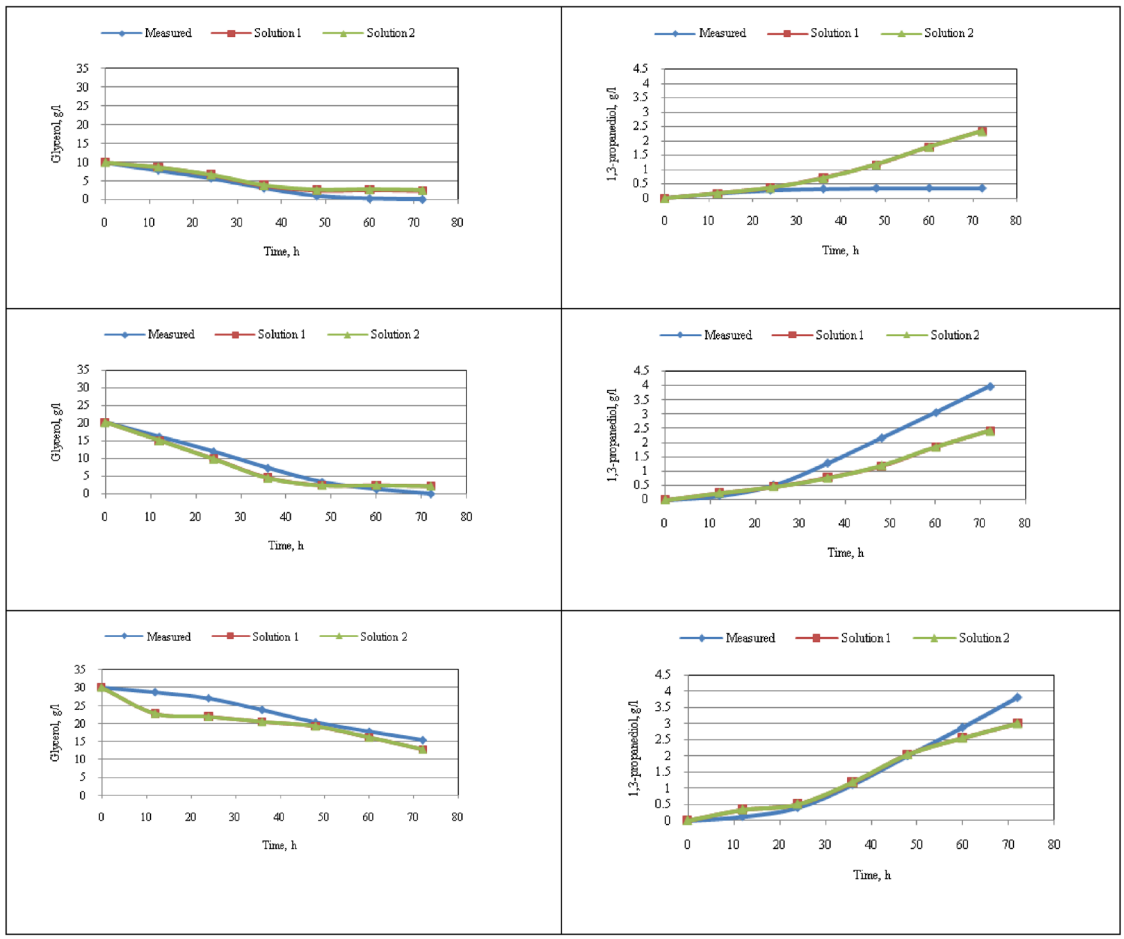

Fig. 2. Comparison between measured and calculated data for glycerol and 1,3-propandiol.

glycerol biotransformation and product formation have been modeled by a Nonlinear AutoRegressive with eXogenous inputs – NARX ANN. An ANN architecture comprising three layers - one input layer, one hidden layer and one output layer has been chosen. The NARX ANN is constructed to be fed with 7 time series of the inputs forming a tapped delay line of 1. The model is validated and the Root Mean Square Error is used to estimate the efficiency of the obtained solutions. The obtained results show that designed ANN model predicts with suficient degree of accuracy the processes of biotransformation of the glycerol to 1,3-propanediol. Observed differences between measured values for the concentrations of the glycerol and the product 1,3-propanediol and those predicted from the model, can be explained by the fact that the process is very new and not well studied, which requires further investigations.

Acknowledgements

This study has been carried out with the financial support of the National Science Fund of Ministry of Education and Science, under the Contract № DMU 03/104/2011.

REFERENCES

1. H. Biebl, K. Menzel, A.P. Zeng, W.D. Deckwer, Microbial production of 1,3-propanediol, Appl. Microbiol. Biotechnol., 52, 1999, 289-297.
2. A.P. Zeng, A. Rose, H. Biebl, C. Tag, B. Guenzel, W.D. Deckwer, Multiple product inhibition and growth modeling of Clostridium butyricum and Klebsiella pneumoniae in glycerol fermentation, Biotechnol. Bioeng., 44, 1994, 902-911.
3. A.P. Zeng, W.D. Deckwer, A kinetic model for substrate and energy consumption of microbial growth under substrate-sufficient conditions, Biotechnol. Prog., 11, 1995, 71-79.
4. A.P. Zeng, A kinetic model for product formation of microbial and mammalian cells, Biotechnol. Bioeng., 46, 1995, 314-324.
5. Z.L. Xiu, A.P. Zeng, L.J. An, Mathematical modeling of kinetics and research on multiplicity of glycerol bioconversion to 1,3-propanediol, J. Dalian Univ. Technol., 40, 2000, 428-433.
6. Ya Qin Sun, Wen Tao Qi, Hu Teng, Zhi Long Xiu, An Ping Zeng, Mathematical modeling of glycerol fermentation by Klebsiella pneumoniae: concerning enzyme-catalytic reductive pathway and transport of glycerol and 1,3-propanediol across cell membrane, Biochem. Eng. J., 38, 2008, 22-32.
7. K. Li, W. Peng, J. Zhou, S. Wei, X. Cheng, Establishment of beet molasses as the fermentation substrate for industrial vitamin B_{12} production by Pseudomonas denitrificans, J. of Chem. Tech. and Biotechnol., 88, 9, 2013, 1730-1735.
8. E. Diaconescu, The use of NARX neural networks to predict chaotic time series, WSEAS Transactions on Computer Research, 3, 3, 2008, 182-191.
9. M. Benne, B. Grondin-Perez, J.-P. Chabriat, P. Herve, Artifcial neural networks for modelling and predictive control of an industrial evaporation process, J. of Food Eng., 46, 2000, 227-234.
10. E.G. Shopova, N.G. Vaklieva-Bancheva, Basic – a genetic algorithm for engineering problem solution, Comput. and Chem. Eng., 30, 8, 2006, 1293-1309.