ATINER CONFERENCE PAPER SERIES No: IND2013-0821

Athens Institute for Education and Research
ATINER

ATINER's Conference Paper Series
IND2013-0821

Modeling of Biotransformation of Crude Glycerol from Biodiesel Production by Using Dynamic Artificial Neural Network

Elisaveta Kirilova
Assistant Professor
Bulgarian Academy of Sciences
Bulgaria
Sophia Yankova
Assistant Professor
Bulgarian Academy of Sciences
Bulgaria
Bilyana Ilieva
PhD Student
Bulgarian Academy of Sciences
Bulgaria
Natasha Vaklieva-Bancheva
Professor
Bulgarian Academy of Sciences
Bulgaria

ATINER CONFERENCE PAPER SERIES No: IND2013-0821

Athens Institute for Education and Research
8 Valaoritou Street, Kolonaki, 10671 Athens, Greece
Tel: + 30 210 3634210 Fax: + 30 210 3634209
Email: info@atiner.gr URL: www.atiner.gr
URL Conference Papers Series: www.atiner.gr/papers.htm

Printed in Athens, Greece by the Athens Institute for Education and Research.
All rights reserved. Reproduction is allowed for non-commercial purposes if the source is fully acknowledged.

ISSN **2241-2891**
23/1/2014

An Introduction to ATINER's Conference Paper Series

ATINER started to publish this conference papers series in 2012. It includes only the papers submitted for publication after they were presented at one of the conferences organized by our Institute every year. The papers published in the series have not been refereed and are published as they were submitted by the author. The series serves two purposes. First, we want to disseminate the information as fast as possible. Second, by doing so, the authors can receive comments useful to revise their papers before they are considered for publication in one of ATINER's books, following our standard procedures of a blind review.

Dr. Gregory T. Papanikos
President
Athens Institute for Education and Research

This paper should be cited as follows:

Kirilova, E., Yankova, S., Ilieva, B. and Vaklieva-Bancheva, N. (2013) **"Modeling of Biotransformation of Crude Glycerol from Biodiesel Production by Using Dynamic Artificial Neural Network"** Athens: ATINER'S Conference Paper Series, No: **IND2013-0821**.

Modeling of Biotransformation of Crude Glycerol from Biodiesel Production by Using Dynamic Artificial Neural Network

Elisaveta Kirilova
Assistant Professor
Bulgarian Academy of Sciences
Bulgaria

Sophia Yankova
Assistant Professor
Bulgarian Academy of Sciences
Bulgaria

Bilyana Ilieva
PhD Student
Bulgarian Academy of Sciences
Bulgaria

Natasha Vaklieva-Bancheva
Professor
Bulgarian Academy of Sciences
Bulgaria

Abstract

Glycerol is a by-product from biodiesel industry. The fermentation of glycerol by different bacterial strains is a complex biotechnological process producing valuable products as organic acids, diols, and ethanol, which are used as reagents in chemical industry. In this study results obtained from biotransformation of glycerol by the help of bacterium from the genera of *Pseudomonas*, in particular *Pseudomonas denitrificans 1625* to formation of 1,3-propanediol (1,3-PD) are represented for the first time. Mathematical modeling is the approach that can give an answer for the process performance in different scenarios. Having in mind the lack of information about the kinetics of this process we have proposed a dynamic Nonlinear AutoRegressive with eXogenous inputs – NARX Artificial Neural Network (ANN) for the time series modeling of the process. For that purpose real sets of time series of parallel batch experiments of the process at different initial concentration of the glycerol have been used. We have developed dynamic ANN model fed with the concentration values of the biomass as independent (exogenous) inputs and 5 and 7 time series interval of past values of the concentrations of the glycerol

and 1,3-propanediol fed with tapped delay line of 1 for prediction of their next values (glycerol and 1,3-propanediol). The developed model have been trained and validated with selected for the purpose data sets. The obtained after validation results have shown acceptable coincidence between measured and calculated data.

Key words: Modeling, NARX ANN, Glycerol, 1,3-Propanediol, *Pseudomonas denitrificans 1625*

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

Corresponding Author:

Introduction

1,3-propanediol is valuable chemical reagent on a great commercial scale in production of composites, adhesives, laminates, polyesters, solvents, antifreeze as well as supplements in fuel oil and etc. Since its production is quite limited and very expensive, alternative for its obtaining from crude glycerol by the help of strains of microorganisms is quite promising. On the other hand the crude glycerol appears in big amounts as by-product from the biodiesel production. Due to its poor quality it couldn't be used directly in the industry and needs further treatment. For that reason, its utilization is a valuable for that production and leads as to reduction of production cost as well as environmental impact minimization.

Many experimental investigations of the process of biotransformation of crude glycerol to formation of 1,3-propanediol by different bacteria strains so as *Citrobacters, Enterobacter* and *Clostridia* and so on (Biebl et al., 1999) have been carried out in order to evaluate its performance. However, until now there have been no records for realization of this process by the help of glycerol-degradation strain from the genera *Pseudomonas*. In this study for the first time results, obtained during experimental investigation of process of biotransformation of crude glycerol by the help of *Pseudomonas 1625* to formation of 1,3-propanediol are presented.

Taking into account that the kinetics of bioconversion of glycerol is highly complicated since microbial growth is subjected to multiple inhibitions of substrate and products, mathematical modeling is the approach which can overcome many of defenses of experimental studies, investigating concurrently substantial aspects of the process. Most of developed mathematical models are based namely on kinetic models for substrate consumption and product formation in a large range of feed glycerol concentrations into medium (Zeng et al., 1994; Zeng&Deckwer, 1995; Zeng, 1995; Xiu et al., 2000; Ya Qin Sun, 2008). However they don't take into account the formation of intermediate and intracellular substances along the process as well as the enzymes for the glycerol metabolism.

Having in mind that the kinetics of biotransformation of crude glycerol by the help of *Pseudomonas 1625* to formation of 1,3-propanediol is not clear enough, as the most suitable for its modeling appear Artificial Neural Networks which treat the process as a "black box".

ANN are inspired from the way how biological neurons transmit and process information, comprising a large number of highly interconnected artificial neurons which receive input data and process them so as to obtain output. They are applicable for modeling wide range phenomena in physics, computers science, biochemistry, mathematics, sociology, economics, telecommunications, bio-medical instrumentation and many other areas, through consideration of only the available values of the process variables, developing conditional nonlinear functions based on the extraction of the data.

The main objective of this study is to develop dynamic recurrent ANN model for prediction of the obtained amounts of 1,3-propanediol and degradation of the glycerol at different initial concentrations of the substrate.

The rest of the study is structured as follows: Section 2 provides description of the experiments carried out, involving specification of producer strain, composition of used bacterial culture medium, conditions under which the fermentation is conducted and used equipment for determination of concentrations of product obtained and the biomass. Section 3 is devoted to developed dynamic ANN model of the process. The results of models' validation are also shown. Finally, short conclusions are done.

Materials and Methods

Biotransformation of glycerol is affected by a number of factors so as: the specification of producer strain, the composition of used bacterial culture medium and the conditions under which the fermentation is conducted so as temperature, initial concentration of substrate, aeration, pH.

Bacterial Strain

For the purpose of this study is used glycerol-degradation strain from the genera of *Pseudomonas*, in particular *Pseudomonas denitrificans 1625***.** It is anaerobic bacterium for denitrification. Denitrification bacteria are heterotrophic, i.e. they received carbon from organic resources so as sucrose, glucose, alcohol, organic acids and others. In general, *Pseudomonas denitrificans* has a wide industrial application mainly in the production of vitamin B_{12}. Effective method for its production is by using of beet molasses that is obtained as a by-product in sugar industry, (Li et al., 2013).

Bacterial Culture Medium

During the batch process the bacterial strain of *Pseudomonas denitrificans 1625* is cultivated with two culture media for seed and for operation. Their compositions are listed in Tables 1 and 2.

Table 1. *Bacterial Culture medium №1 (seed)*

Component	Amount
Yeast extract	1.0g/l
Pepton	10g/l
NaCl	10g/l
Agar	15g/l
Distilled water	1000ml
Medium pH	7.2 – 7.4

Table 2. *Bacterial Culture medium №2 (operating)*

Component	Amount
$MgSO_4.7H_2O$	0.2g/l
$CaCl_2.2H_2O$	0.2g/l
NaCl	5.0g/l
KH_2PO_4	379ml
Na_2HPO_4	621ml
Saline solution	5g/l
Glycerol	2-6g/l

Equipment Cultivation

The culture media have been sterilized in 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 at temperature of 30ºC with stirring. Throughout the processes the initial concentration of glycerol has varied in between 10 – 30 g/l.

Fermentation Conditions

The main factors of glycerol fermentation are: temperature, substrate concentration, aeration and pH.

Temperature. The optimal temperature for growth of *Pseudomonas aeruginosa* is 37^0C. Despite the fact that 37^0C is approved as the optimal temperature due to the fact that the bacteria *Pseudomonas denitrificans* and *Pseudomonas aeruginosa* are from the same genera, the fermentation of glycerol has been conducted at temperature of 30^0C.

Substrate concentration. The process is carried out at initial concentration of glycerol of 10-30 g/l.

pH. It is experimentally proven that the optimal pH for *Pseudomonas sp.* is in the range 7.0 - 7.5.

Analytical Methods

Determination of Metabolites Concentrations

Quality and quantity analysis of dissolved product from the fermentation and the substrate glycerol by using HPLC have been conducted. All samples have been previously filtered through a filter with a pore size of 0.45 μm (Boeco, Germany).

Determination of Concentration of Biomass

The concentrations of biomass during the experiments have been determined by measuring the extinction of the sample at a wavelength of λ = 660 nm. For this purpose spectrophotometer Specol (CARL Zeiss, Jena, Germany) has been used. The concentration of biomass is calculated from a calibration curve.

3. Dynamic Artificial Neural Network modeling for time series forecasting the formation of 1,3-propanediol and glycerol degradation at different initial concentrations of the substrate.

Formulation of the Proposed Model

Artificial neural networks are mathematical models inspired from the way how biological neurons transmit and process information. They are flexible computing frameworks that can be applied to a wide range of time series forecasting problems with a high degree of accuracy. ANNs have different advantages so as: they don't need any information about the process under consideration, only the available data for the input parameters and output parameters are needed; ANNs can generalize, i.e. after ANN are trained with selected for that data patterns, ANNs can often correctly infer the unseen part of the time series data even if the data patterns contain noisy information. The power of ANNs comes from the parallel processing of the information from the data.

In general, 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 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 *bias*.

All data propagate along the connections in the direction from the network inputs to the network outputs, hence the term *feed-forward*.

The most widely used feed-forward dynamic ANN model for the time series modeling and forecasting is Nonlinear AutoRegressive with eXogenous inputs – NARX model, (Diaconescu, 2008; Benne et al., 2000). The NARX is a recurrent dynamic network, with feedback connections enclosing several layers of the network. In the figure 1 is represented a two-layer NARX feed-forward. It is fed with a sequence of adjacent values of time series of input parameters formed tapped delay line. This input is usually referred to as a time window since it provides a limited viewed 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.

The general prediction equations for computing the next value of time series $y(t+1)$ - output by using the past observation of independent (exogenous) inputs $x(t), x(t-1),..., x(t-n_x)$ and past observation $y(t), y(t-1),..., y(t-n_y)$ also considered as inputs are written as follows:

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

Where

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

is hyperbolic tangent function used as activation function applied for aggregated weighted sums:

$$net_i = \sum_j w_{i,j} \cdot x_j + b \qquad (3)$$

i - index for a given neuron;

j - index for a given input;

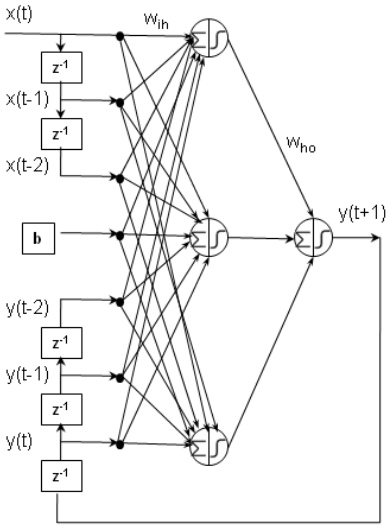
b - bias input;

n_y and n_x - the maximum time delays for the inputs;

w_{ih} and w_{ho} - the weighted coefficients of input-to-hidden connections and hidden-to-output connections;

$v(t)$ - the noise of the data.

Figure 1. *Two-layer NARX feedforward ANN*



As can be seen from the Figure 1, the neurons of hidden layers aggregate weighted values of all inputs to single values, eq. (3). Then, an *activation function* is applied to the aggregated weighted value to produce an individual *output* for the specific neuron, eq. (2).

Using *a least-square function* (LSF) as optimization criteria 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 criteria. For this purpose a powerful *optimization algorithm* should be used. This process is called *training* of ANN or *supervising.*

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.

Choice of ANN architecture is preceded by the analysis of collected data, so as to reveal clearly the relations between inputs and outputs and to choose patterns for ANN training and validation of designed models.

Determination Inputs and Outputs of the Model

For the purpose of ANN modeling, data obtained from conducting of parallel experiments at three initial concentration groups of crude glycerol as follows 10g/l, 20g/l and 30g/l have been used. Product - 1,3-propanediol and the biomass are measured in the interval of 20-24 h during the parallel experiments last.

The purpose of ANN modeling of the process of biotransformation of crude glycerol is 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. The values of concentrations of the biomass are very small and almost unchangeable due to the fact that the process is led in the initial stage of biomass growth. For that reason they are used as Exogenous (independent) inputs.

Thus, we have selected required inputs and output for the process modeling as is listed in Table 3.

Table 3. *Inputs and outputs for ANN modelling*

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

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

ANN Architecture

Based on the available data, number of inputs and outputs, we have selected as the most promising architecture consisting 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 5 time series and 7 time series of the inputs forming tapped delay line of 1. The values of input data in the 5 time series are taken at 0th, 16th, 32th, 48th and 64th hours and in the 7 time series are taken at 0th, 12th, 24th, 36th, 48th, 60th and 72th hours, respectively.

Training and Validation of ANN

By using the patterns for training we have implemented BASIC genetic algorithm (Shopova&Vaklieva, 2006) for the ANN architecture selected in order to obtain the values of the weighting coefficients (models' parameters) at which the criterion LSF has a minimal value.

As a result of multiple calculations, two equally optimal solutions with different values of weighting coefficients - ANN model parameters have been selected for each one of the modeled 5 and 7 time series.

These solutions have been validated with selected for the purpose data patterns.

In order to estimate the efficiency of obtained solutions as well as their ability to make precise predictions we have used a Root Mean Square Error (RMSE) as a performance accuracy measure.

The RMSE represents square root of the average of the summing square predicting errors and defines as follows:

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

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 different solutions are summarized in the Table 4.

Table 4. *Values of RMSE for obtained Solutions for 5 and 7 Time Series*

Time series	RMSE	
	Glycerol, g/l	1,3-propanediol, g/l
5 time series		
Solution 1	2.9625	0.6148
Solution 2	3.0234	0.6190
7 time series		
Solution 1	2.3953	0.7788
Solution 2	2.4058	0.7784

Figures 2 and 3 demonstrate comparison between measured and calculated data for three groups of experiments carried out at three different concentrations of the substrate - 10 g/l, 20 g/l and 30 g/l. The respective curves of biodegradation of the glycerol are mentioned on Figures 2 and 3 as 1a, 1b and 1c. Correspondingly, the concentrations of products formation are mentioned on both figures as 2a, 2b and 2c.

Figure 2. *Five time series. Comparison between measured and calculated data of both glycerol and 1,3-propanediol at regarded three different initial concentration for the glycerol*

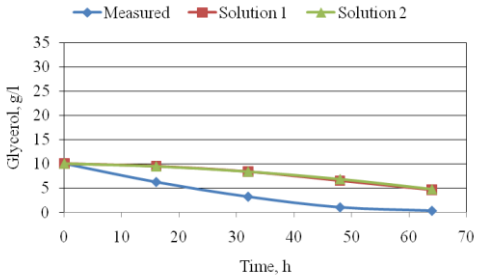


Fig. 2-1a

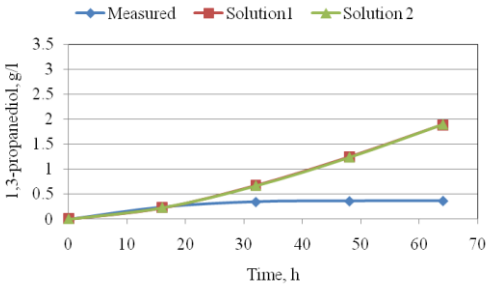


Fig. 2-2a

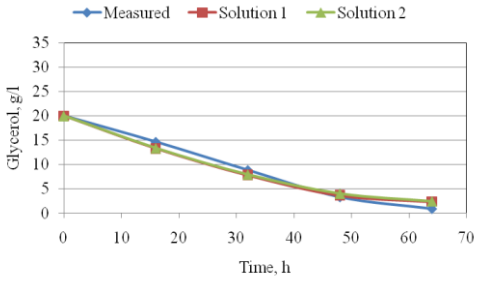


Fig. 2-1b

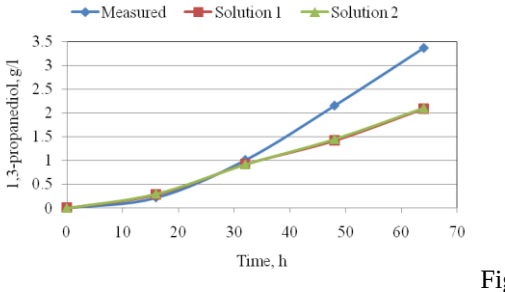


Fig. 2-2b

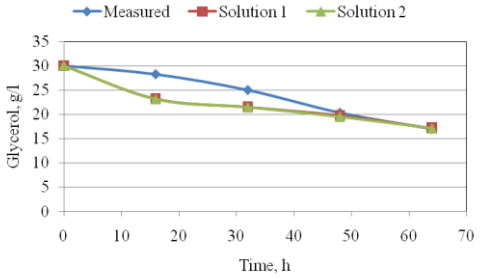


Fig. 2-1c

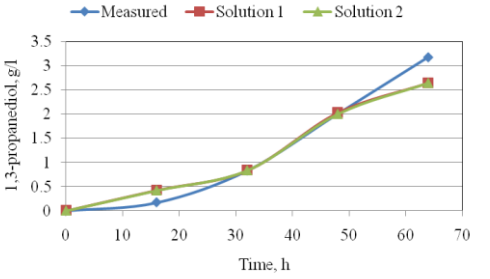


Fig. 2-2c

Figure 3. *Seven time series. Comparison between measured and calculated data of both glycerol and 1,3-propanediol at regarded three different initial concentration for the glycerol*

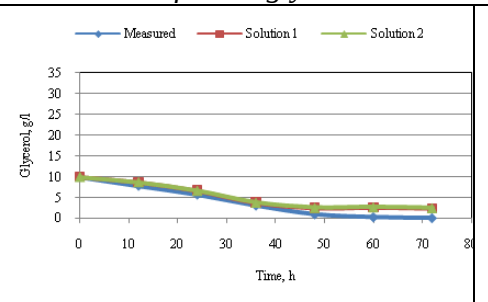


Fig. 3-1a

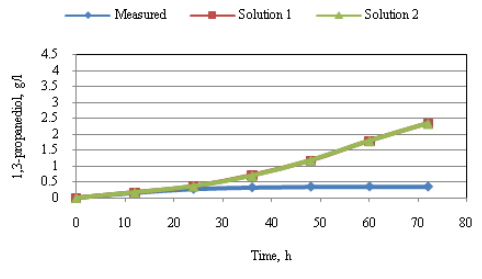


Fig. 3-2a

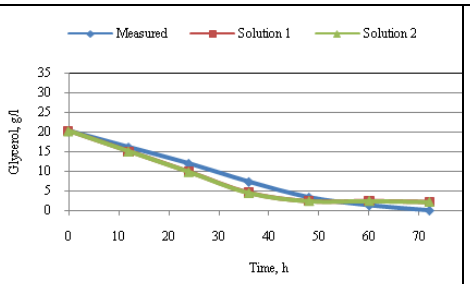


Fig. 3-1b

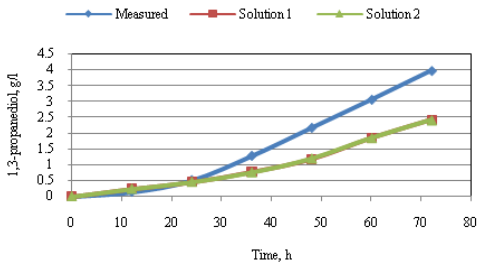


Fig. 3-2b

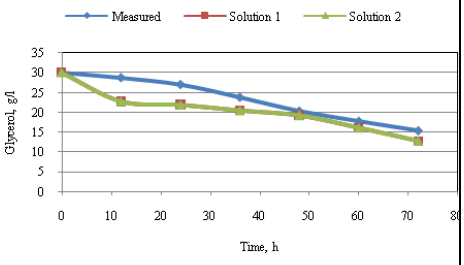


Fig. 3-1c

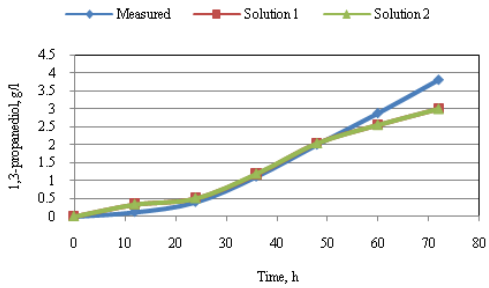


Fig. 3-2c

RMSEs summarized in Table 4, for the solutions obtained for both time series, show that the biotransformation of glycerol is modeled better for the case of the 7 time series than in 5 time series, while the product 1,3-propanediol formation is modeled better for the case of the 5 time series.

Likewise, it can be seen on the Figures 2 and 3 that there exists a best coincidence between measured and calculated data for the biotransformation of glycerol in the case of the 5 time series at initial concentration of glycerol of 20 g/l. For the case of 7-time series, the degradation of the glycerol is modeled very well at initial concentration of the glycerol 10 and 20 g/l as at 30 g/l a good coincidence can also be seen.

Concerning to the 1,3-propanediol formation, for both time series a very good coincidence between measured and calculated data can be seen at initial concentration of glycerol of 30 g/l, while at initial concentrations of glycerol of 10 and 20 g/l, a good coincidence between measured and calculated data for the product formation exists only until 32th hour for the case of 5 time series and until 24th hour for the case of 7 time series. Since then for both time series a discrepancy between measured and calculated data can be observed. We accept that the latter is due to the lack of information about the process of biotransformation of crude glycerol by using of bacterium *Pseudomonas* to formation of 1,3-propanediol and also to the fact that the ANN dynamic model doesn't account some important for the process parameters so as to be taken as exogenous inputs.

Conclusions

In this study for the first time the process of biotransformation of crude glycerol by the help of the bacteria *Pseudomonas denitrificans 1625* to formation of 1,3-propanediol is commented. The study provides a short description of process operational conditions. The results obtained from three groups of parallel experiments conducted at three different initial concentrations of the glycerol are used for its mathematical modeling. Taking into account that the process is not clear enough as the most suitable for its modeling appears Artificial Neural Network. Dynamics of the glycerol biotransformation and product formation are modeled by using Nonlinear AutoRegressive with eXogenous inputs – NARX ANN. There ANN architecture comprising three layers - one input, one hidden layer and one output layer has been investigated. The dynamic recurrent ANN is constructed so as to be fed with 5 time series and 7 time series of the inputs forming tapped delay line of 1. The ANN model is validated with selected for this purpose data sets and for estimation the efficiency of obtained solutions the Root Mean Square Error has been evaluated. Obtained values for RMSE for both cases of 5 and 7 time series have shown that designed ANN models predict with enough degree of accuracy the processes of biotransformation of the glycerol and product formation. Observed differences between measured values for the concentrations of the glycerol and the product 1,3-propanediol with these ones predicted by the ANN model can be explained by the fact that the process is a very new and will be studied in detail in the future.

References

Benne, M., Grondin-Perez, B., J.-P. Chabriat, Herve, P. (2000). 'Artifcial neural networks for modelling and predictive control of an industrial evaporation process.' *Journal of Food Engineering* 46: 227-234.

Biebl, H., Menzel, K., Zeng, A.P., Deckwer, W.D. (1999). 'Microbial production of 1,3-propanediol.' *Appl. Microbiol. Biotechnol.* 52: 289–297.

Diaconescu, E. (2008). 'The use of NARX neural networks to predict chaotic time series'. *WSEAS Transactions on Computer Research* 3(3): 182-191.

Li, K., Peng W., Zhou J., Wei, S., Cheng, X. (2013). 'Establishment of beet molasses as the fermentation substrate for industrial vitamin B_{12} production by *Pseudomonas denitrificans.*' *Journal of Chemical Technology and Biotechnology*, Article in Press, Published online: 19 February 2013.

Shopova E.G., Vaklieva-Bancheva, N.G. (2006). 'Basic – a genetic algorithm for engineering problem solution.' *Computers and Chemical Engineering* 30(8): 1293-1309.

Xiu, Z.L., Zeng, A.P., An L.J. (2000). 'Mathematical modeling of kinetics and research on multiplicity of glycerol bioconversion to 1,3-propanediol.' *J. Dalian Univ. Technol.* 40: 428–433.

Ya Qin Sun, Wen Tao Qi, Hu Teng, Zhi Long Xiu, An Ping Zeng, (2008). 'Mathematical modeling of glycerol fermentation by Klebsiella pneumoniae:

concerning enzyme-catalytic reductive pathway and transport of glycerol and 1,3-propanediol across cell membrane.' *Biochemical Engineering Journal* 38: 22–32.

Zeng, A.P. (1995). 'A kinetic model for product formation of microbial and mammalian cells'. *Biotechnol. Bioeng.* 46: 314–324.

Zeng, A.P., Deckwer, W.D. (1995). 'A kinetic model for substrate and energy consumption of microbial growth under substrate-sufficient conditions.' *Biotechnol. Prog.* 11: 71–79.

Zeng, A.P., Rose, A., Biebl, H., Tag, C., Guenzel, B., Deckwer, W.D. (1994). 'Multiple product inhibition and growth modeling of Clostridium butyricum and Klebsiella pneumoniae in glycerol fermentation.' *Biotechnol. Bioeng.* 44: 902–911.