Mixture of Xylose and Glucose Affects Xylitol Production by *Pichia guilliermondii*: Model Prediction Using Artificial Neural Network

Magharei, Azadeh; Vahabzadeh, Farzaneh*+; Sohrabi, Morteza,
Rahimi Kashkouli, Yousef•; Maleki, Mohammad
Faculty of Chemical Engineering, Amir Kabir University of Technology, Tehran, I.R. IRAN

ABSTRACT: *Production of several yeast products occur in presence of mixtures of monosaccharides. To study effect of xylose and glucose mixtures with system aeration and nitrogen source as the other two operative variables on xylitol production by Pichia guilliermondii, the present work was defined. Artificial Neural Network (ANN) strategy was used to mathematically show interplay between these three controllable factors and the xylitol productivity response. In the first stage, model fitting was performed using Response Surface Methodology (RSM) and the appropriate fraction of this design then was applied for the ANN training step (Levenberg Marquardt 'LM' algorithm). The best ANN model configuration with the three test input variables composed of six neurons in the hidden layer and tangent sigmoid (TANSIG) and linear transfer function (PURELIN) were used as the activation functions for the data processing from inputs to the hidden layer and from the constructed neurons to the output nodes. The network performance was evaluated by Mean Squared Error (MSE) and the regression coefficient of determination (R^2). These values respectively, for the RSM model fitting were 2.327×10^{-4} and 0.9817, and for the ANN training data were 2.29×10^{-8} and 0.9999. While MSE and R^2 values for the other two steps of ANN were 4.56×10^{-3} and 0.9741 (validating step) and1.52×10^{-3} and 0.9325 (testing step), respectively. Positive synergism of ANN with RSM was confirmed.*

KEY WORDS: *Artificial neural network, Glucose and xylose mixture, Pichia guilliermondii, Response surface methodology, Xylitol production.*

INTRODUCTION

Xylitol as a naturally occurring pentahydroxy sugar alcohol has high sweetening power and its ability in being used as an alternative to sucrose, has directed attentions toward its different applications specially in foods, pharmaceuticals, and related medical products [1,2]. Points of interest in these areas are anticariogenic properties of this polyol and its insulin-independency metabolism [3,4]. Low concentration of xylitol in fruits and vegetables has made extraction of this sugar alcohol less economical and more difficult to manage efficiently

* *To whom correspondence should be addressed.*
+ *E-mail: far@aut.ac.ir*
• *Other Address: Young Research Club, Islamic Azad University, Gachsaran Branch, Gachsaran, I.R. IRAN*
1021-9986/12/1/119 13/$/3.30

the process. On an industrial scale, production of xylitol is based on chemical separation of xylose fraction of hemicellulosic materials but the process is still considered expensive since the hemicellulosics also contain other types of sugars and biopolymers, which should be separated from the test matrix. Attentions are now directed towards use of biochemical routes and despite of capability of yeasts for xylitol production, there are still some problems in these types of bioconversions like potentiality of simultaneous production of other metabolites (ethanol, glycerol, arabinol,…). The extent of these fermentation issues could be reduced through use of microbial screening program and preparation of different nutritive regimes for the test microorganism and choosing the microorganism based on its potential to act selectively in response to one particular substrate in a multisubstrate system.

Results of many researches show that yeast is capable of growing on glucose in essentially four modes, aerobically or anaerobically with or without the substrate repression. Oxidative and fermentative patterns of the growth are differentiated based on glucose concentration and availability of molecular oxygen as the ultimate electron acceptor [5, 6]. Fig. 1 shows interplay between these biochemical pathways operative in yeasts when they grow on mixtures of xylose and glucose under aerobic conditions. The oxidative branch of the Pentose Phosphate Pathway (PPP) uses $NADP^+$ as the electron acceptor in the two oxidation steps and ribulose 5-phosphate (ribulose 5-p) forms in the last step which is followed by the formation of ribose 5- phosphate (ribose 5-p) through the isomerization reaction in the non-oxidative branch of the pathway. When cells need much more NADPH than ribose 5-p then a reversible link between the PPP and glycolysis is created with use of some enzymes (mainly transketolase and transaldolase). Only the non- oxidative branch of the PPP is active when cellular need for ribose 5-p is much more than NADPH (conversion of fructose 6-phosphate 'fructose 6-p' and glyceraldehyde 3-phosphate 'glyceraldehyde 3-p' as the glycolytic intermediates to ribose 5-p without the formation of NADPH) [5, 6].

A systematic approach for optimization of fermentation for xylitol production hence could be centered not only the mixtures of xylose and glucose but also on the system aeration, and availability of nitrogen, i.e., to establish balanced states between energy production, biomass formation, and biosynthesis of xylitol metabolite. Studies on the operational variables used for the cultural conditions set up for the test yeast performance in xylitol production, show that biochemical conditions of the culture affect delicately redox balance between two coenzyme-dependent enzymes: $NADP^+$/NADPH dependent Xylose Reductase (XR), and NAD^+/NADH dependent Xylitol DeHydrogenase (XDH) [7- 10].

Different nitrogen sources such as ammonium nitrate, sulfate, chloride, and acetate besides urea and yeast extract are investigated for xylitol production by some yeast strains [11]. In their report, preference of AmS to different nitrogen sources for xylitol production by *Hansenula polymorpha* was indicated. However there are some reports in advantage of using urea as well as yeast extract as nitrogen source to have high xylitol production [12].

Candida guilliermondii, the anamorph stage of *Pichia guilliermondii,* is among different *Candida* species capable to produce efficiently xylitol and because of the pathogenic nature of some *candida* species *P. guilliermondii* has been claimed to be the preferred yeast species for use in food and health related industries [13, 14]. The maximum value for xylitol yield for *Candida* species reported in literature is 0.85 g/g xylose although one may consider the theoretical yield as 1 g/g (with the negligible growth of the microorganism) [15]. In fact, typical yields for xylitol vary between 0.4 and 0.7 g/g [16, 17] and operational conditions affect strongly in these values.

Artificial Neural Networks (ANNs) are software tools act mathematically to deal with certain type of problems in a similar way, as the human brain is capable to solve a particular issue [18]. Briefly, many steps which are involved in the ANN design as a complex dynamic process can be summarized as follows: choosing a desired input-output mapping; selecting an appropriate neural architecture; training the network and repeating it several times with changing the size of the network and/ or varying the learning process parameters; cross validation of the data set results the selection of the network, which performs best; and using a new set of data to study the ability of the network in generalization of its developed functions [19, 20]. The model building and prediction properties of ANNs would be better considered as complementary to the formal statistical tools of regression analysis, which are actually based on the underlying knowledge of how the system of

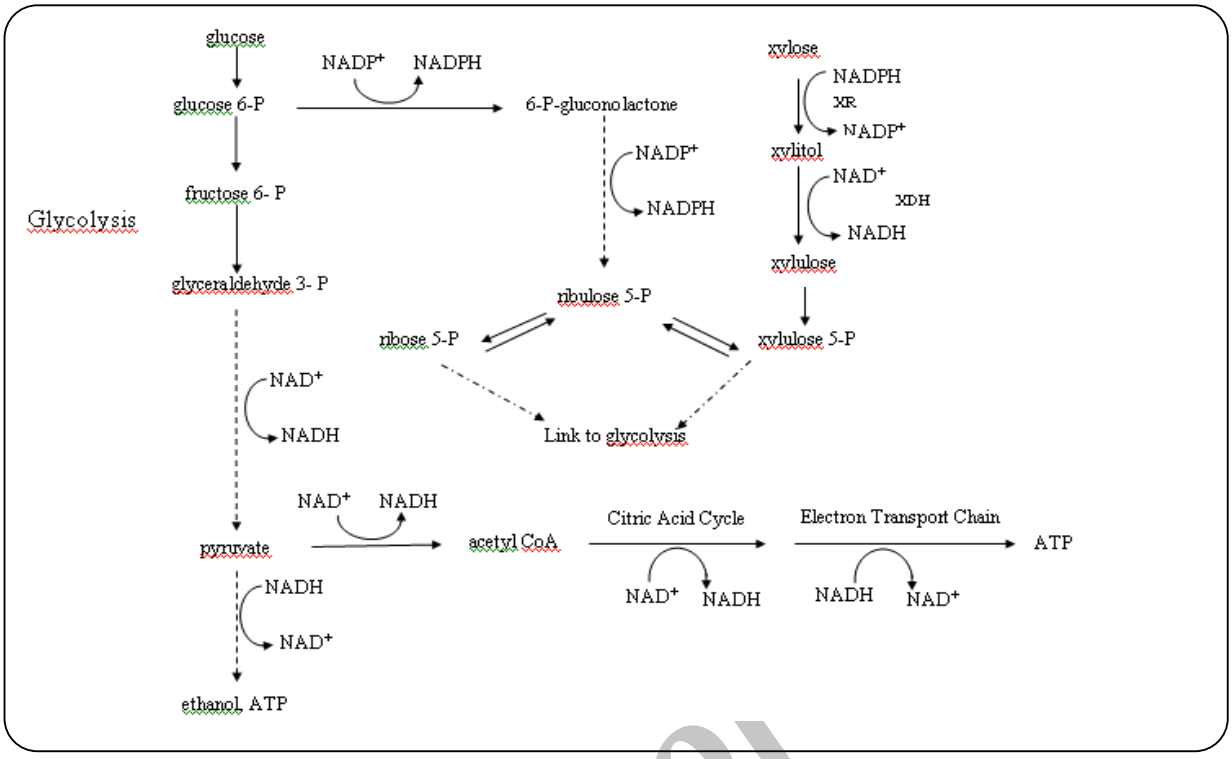

Fig. 1: NAD+ and NADP+- dependency of some enzymes used in xylose and glucose metabolism in yeasts - the interplay between some pathways are shown (glycolytic and pentose phosphate pathways, citric acid cycle and electron transport chain) (see ref. [6] for the details).

the interest works. RSM uses a sequential approach to experimentation and Central Composite Designs (CCD) as a statistical tool can model the key factors on basis of constructing a path of steepest ascent when the goal of the estimation is the maximization of the response [20, 21].

Extent of each fermentation process is greatly under influence of the substrate (s) and in this work mixtures of the two monosaccharides namely, xylose and glucose were used as the substrates to study xylitol production by *P. guilliermondii*. Three fermentative factors then were glucose to xylose (G:Xyl) ratio (wt/wt), system aeration based on the ratio of the volume of the fermentation medium to the volume of the flask (V_m: V_f), and the concentration of nitrogen source (ammonium sulfate 'AmS'). Statistical modeling approach was undertaken by applying RSM in the first stage and the evaluation of the results by ANN in the next step. Although literature survey shows use of RSM for optimization of xylitol production and the importance of defining experimental conditions have been emphasized [1], but to our knowledge there is no report regarding design of experiments based on simultaneous evaluation of ANN and RSM for xylitol production by *p. guilliermondii*.

EXPERIMENTAL SECTION

Fermentation studies

P. guilliermondii DSMZ 70057 was obtained from Deutsche Sammlung von Mikroorganismen und Zelkulturen (DSMZ). The medium with the following composition was used for the maintenance of the yeast cells (YM agar) (g/L): yeast extract 3, malt extract 3, peptone from soybeans 5, glucose 10, and agar 15. Table 1 shows the growth medium composition and conditions which obtained according to the available literature used in the present study with the slight modifications [2, 4]. In this work for providing the pre-determined size of the inoculum (1 g/L), the absorbance curve corresponded to the amount of dried cells was established (Jasco V-550 spectrophotometer, A_{660}). G:Xyl ratios (wt/wt) were set at the constant level of xylose and change of the volume of fermentation medium in the test flasks (250 mL) was used to provide different system aerations: ≤ 65 mL and ≤

Table 1: Composition of growth medium and cultivation conditions.

Component (g/L) or condition	*P. guilliermondii* growth medium
glucose	20
$(NH_4)_2 SO_4$ [a]	5
$MgSO_4. 7H_2O$	0.5
KH_2PO_4	1.0
$CaCl_2. 2H_2O$	0.1
yeast extract	1.0
growth medium (mL) in a 125-mL Erlenmeyer flask	50
incubation time (h)	16
temperature (°C)	30
shaking rate (rpm)	200

[a] *: source of inorganic nitrogen used in the present study (AmS Conc.).*

100 mL for setting aerobic and semi-aerobic conditions, respectively[15].. As shown in Table 2 (a) , the x_1,x_2, and x_3 as three independent variables used in the present study were defined as nitrogen source concentration, G:Xyl ratio, and Vm :Vf ratio, respectively. Details regarding the actual levels of these variables are shown in that table. The experiments were performed using the following equation (Table 2 (b)): $N= 2^k+2k+x_0$, where N is the number of experiments, k is the number of independent variables, and x_0 is the number of center points. Thus for this design used in the present work the total number of experiments were 20 (k=3, x_0 = 6). All the reported productivities were determined based on 90 h fermentation time.

Analytical methods

Cell dry weight was measured by centrifuging the cell suspension (8000 × *g* for 10 min), washing the sediment with sterile distilled water, and drying the washed sediment in an oven (105 °C for about 24 h). Spectrophotometric method was used to obtain the absorbance of the cell suspension at 660 nm and the calibration curve was prepared by plotting the cell dry weight versus A_{660} [11].

Xylose and xylitol measurements were performed with use of HPLC technique: a carbohydrate high performance column, 4.6 × 250 mm (Waters Co., USA) equipped with a refractive index detector (Jasco RI-1530, Intelligent Co.) was used to resolve samples with a solvent system of acetonitrile and water (80:20 (v/v)) at a flow rate of 1 mL/min and the sample size for the injection was 20 µL [11].

Modeling analyses

RSM analysis

RSM with use of CCD considering three independent variables was applied in the first stage of the study. This type of design was used to obtain full second order polynomial model with information on the linear and quadratic effects as well as two factor interactions. The data through these defined experiments were collected according to the arrangement presented in Table 2 (b). Twenty runs were used to optimize xylitol production in terms of the productivity (response). Analysis of variance (ANOVA) and regression calculations were carried out using Design-Expert software version 7.

ANN analysis

All ANN calculations were carried out using MATLAB mathematical software with ANN toolbox (version 7.9). The proposed ANN used for xylitol productivity estimation consisted of three layers, one input layer comprising three nodes or independent variables (AmS concentration, G:Xyl ratio, and system aeration), one hidden layer consisting of several nodes, which were changed to obtain the best configuration, and an output layer containing one output node (xylitol productivity). Fig. 2 shows structural organization of the ANN used for estimation of xylitol productivity.

The ANN network in the present study was trained using Levenberg Marquardt (LM) error back propagation algorithm and with use of this protocol, randomly assigned weights to the initial connectors are propagated

Table 2: Independent variables and their selected levels used to study xylitol fermentation by P. guilliermondii (a). Arrangement of the CCD for the three variables used in the present study and actual values for the response are shown (b). RSM second order model of the xylitol productivity by P. guilliermondii is presented (c). Statistics used to test the adequacy of the constructed model obtained according to the RSM approach (d).

(a)

Independent variables	Level of variables		
	-1	0	1
x_1: $(NH_4)_2SO_4$ Concentration (g/L)	2	3.5	5
x_2: G: Xyl ratio[a]	1:6	1:3	1:2
x_3: V_m:V_f ratio[b]	100:250	75:250	50:250

(b)

Experiments	Predictors (coded units)			Response
	x_1	x_2	x_3	Productivity (g/L/h)
1	-1	-1	-1	0.2393
2	+1	-1	-1	0.1464
3	-1	+1	-1	0.0979
4	+1	+1	-1	0.2015
5	-1	-1	+1	0.1545
6	+1	-1	+1	0.1552
7	-1	+1	+1	0.2127
8	+1	+1	+1	0.4926
9	-1	0	0	0.2039
10	+1	0	0	0.2862
11	0	-1	0	0.1486
12	0	+1	0	0.2173
13	0	0	-1	0.0355
14	0	0	+1	0.0956
15	0	0	0	0.1434
16	0	0	0	0.1407
17	0	0	0	0.1537
18	0	0	0	0.1722
19	0	0	0	0.1694
20	0	0	0	0.1756

(d)

Response	R^2_{Adj}	R^2_{Pred}	Model adequacy ($p<0.05$)[+]	Lack of fit ($p>0.05$)[++]	Adequate precision (ratio>4)	Std. Dev.[†] (s)	Mean[†] ($\bar{x}$)
productivity	0.9652	0.835	0.0001	0.3381	36.894	0.017	0.18

a) xylose at a constant concentration was used (60 g/L),

b) System aeration was carried out according to the ratio of the volume of the fermentation medium to the volume of the flask,

[*] *Regression equation calculated form the significant effects of the ANOVA values of the predicted are specified in their coded units,*

[+] *F- value= (MS for the model/ MS for residual) = 59.59;*

[++] *F =(MS for lack of fit = 3.45×10^{-4}/ MS for pure error = 2.327×10^{-4})= 1.48,*

[†] *CV= $s/\bar{x}$*

forward through the network and calculated for comparison against some known output values. Differences between all the known and propagated output values for every observation are calculated as Mean Squared Error (MSE) and as each of the value is propagated back iteratively through the network, the total error is distributed and reduced. Use of LM as a learning algorithm is advantageous especially with the small sample size, since the data set used for the validation step do not need to be separated from the training one [20, 22, 23]. Number of complete presentation of all of data to the ANN (epoch as a step in the training of the ANN which corresponds to a single pass through the entire data set) is fewer for the network training with LM algorithm. The network training step is automatically stops as the MSE of the validation phase starts to increase and in this way the generalization of the developed ANN is improved. The training cycles in the present study were carried out for varying number of neurons in the hidden layer and when the tested number of the hidden neurons was smaller than four, the ANN was not able to adapt modeling for xylitol production process. Generally time for the training network increases as the number of neurons increases and in fact as the number of parameters involved in a complex nonlinear function increases, the danger of overfitting the network becomes more realistic. This means that although ANN is able to provide a nearly perfect fit to the set of training data, but its prediction ability for new data set is not high [20, 22].

In the present work, the tangent sigmoid (TANSIG) and linear transfer function (PURELIN) were used as the transfer functions in the hidden and output layers of the ANN, respectively. Testing as the final step of the developing ANN is achieved by running the ANN with the trained network data files without known output values (those data that have been previously used). A new set of the output values should be used in the testing step (data not used and seen by the ANN). The effectiveness of the network's developed pattern thus is evaluated.

An analysis of the residuals $(y_i - \hat{y}_i)$ is the key factor in evaluating the appropriateness of the predicted model. The extent of the error reduction by use of the predicted output value is calculated as the coefficient of determination:

$$R^2 = \frac{\text{SS 'total system variability'} - \text{SS 'unexplained variability'}}{\text{SS 'total system variability'}} \quad (2)$$

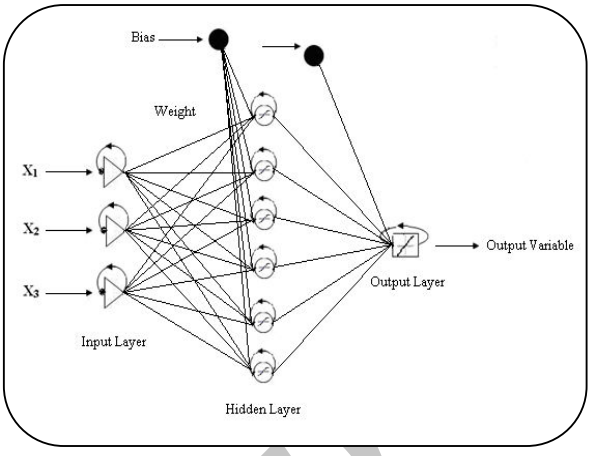

Fig. 2: Schematic representation of the three layers FFEBP used in the present study to estimate xylitol productivity.

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (Y_{i,Target} - \hat{Y}_{i,Pred})^2}{\sum_{i=1}^{n} (Y_{i,T} - \bar{Y})^2} \quad (3)$$

The MSE is defined as follows

$$MSE = \frac{\sum_{i=1}^{n} (Y_{i,T} - \hat{Y}_{i,P})^2}{n} \quad (4)$$

Where $Y_{i,\ Target}$ ($Y_{i,T}$) represents the actual value for the test response (experimental value), $\hat{Y}_{i,Pred}$ ($\hat{Y}_{i,P}$) is the predicted value (either the value predicted by RSM or ANN), $\bar{Y}$ is the mean value, and n is the number of data points.

RESULTS AND DISCUSSION

Parameter selection- design of experiment (DOE)

The high and low values for the each test predictor were chosen based on the results obtained through performing several preliminary experiments (unpublished data). Values for G:Xyl ratio at the center point of the CCD were obtained on the basis of the reported quantities of these sugars in sugar cane bagasse hydrolyzate [15]. The range for system aeration was selected according to the findings reported elsewhere and was based on the fermentation medium volume in a 250-mL erlenmeyer flask [11, 15]. Type of nitrogen source and the range used in the present study both were selected according to the results of our preliminary works, which were in agreement with the findings of the others (preference of AmS to different nitrogen sources for xylitol production

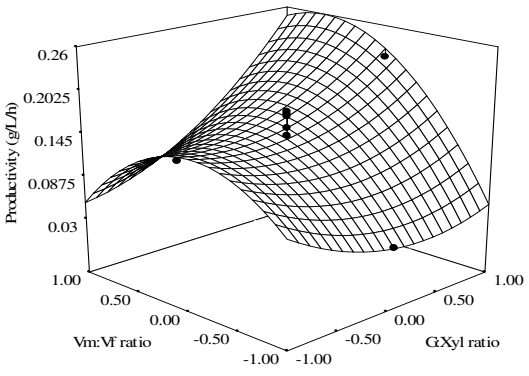

Fig. 3: The effects of system aeration (V_m:V_f ratio) and the G:Xyl ratio on xylitol productivity as shown by a 3-dimensional response surface with the third regressor (AmS concentration) held constant at the center point.

by *Hansenula polymorpha*) [11]. The productivity response values (actual) are given in Table 2 (b).

RSM model fitting

According to RSM approach taken in the first stage of the present study a quadratic regression model was fitted to the observations to estimate surface for the xylitol productivity response (Table 2 (c)). The results of F-test for ANOVA, which are given in Table 2 (d), revealed that this regression is statistically significant (p-value=0.0001) at 95 % confidence level. The determination coefficient of the suggested model (R^2= 0.98) demonstrates a good correlation between the selected parameters and by using these variables, the fitted model could explain 98 % of total variation. All the linear and quadratic terms as well as the interactions of the independent variables present in the predicted model while the most significant effect is due to the concentration of ammonium salt (quadratic) followed by system aeration (quadratic). The interaction of system aeration and the substrate is also significant but with lower absolute value. Table 2 (c) shows the developed model for predicting xylitol production which can be used to navigate the space defined by the CCD. By using the numerical approach- desirability in the Design Expert software the optimal conditions for xylitol productivity was obtained (data are not shown). These data were coincided on the conditions of maximum value of the productivity response (0.49g/L/h) which was obtained through experiment no. 8 in Table 2. Therefore the values of independent variables were 5 g/L AmS, a G: Xyl ratio of 1: 2, and V_m:V_f = 50: 250 mL/mL (Table 2). The response surface was used for finding the combination of G: Xyl ratio and system aeration at the 3.5 g/L ammonium salt concentration (Fig. 3).

Kinetics

Time course of the reaction under optimum conditions is presented in Fig. 4 where at 60 g/L as the initial xylose concentration, utilization of this sugar by *P. guilliermondii* was greatest and peak of xylitol accumulation was at 90 h (the trends of depletion of glucose and xylose substrates are also shown ,which were different). Catabolic repression as a phenomenon discovered long time ago, is due to the diauxic pattern of growth seen in many microorganisms when they grow on the mixture of the substrates [5]. The biphasic growth behavior of several yeasts growing on the mixed monosaccharides has long been recognized and is the focusing point in researches in recent years [7, 8, 24].

Results presented in Fig. 4 show that presence of glucose under this batch fermentation, did not exert a repressing effect on consumption of xylose by *P. guilliermondii*: the biocatalytic reduction of xylose to xylitol efficiently was occurred. While the produced xylitol was consumed gradually after 90 h (90 h as the time for the peak of this polyol accumulation). Similar decreasing trends were observed for the both substrates and by almost complete consumption of xylose, a plateau was reached (Fig. 4). *P. guilliermondii* did not utilize all the glucose and reaching to a plateau while fermentation medium still contained 12.65 g/L glucose (xylitol yield based on initial xylose concentration= 0.74 g/g) (Fig.4). Monitoring fate of appearance of several enzymes and their activities would give conclusive evidences for how to conduct an efficient xylitol fermentation process with use of *P. guilliermondii.*

ANN approach to RSM

Robustness of ANN and its computation style is adaptable and depends on appropriately selecting the model, the learning algorithm, and the cost function (a measure of the distance between the output and its optimal value) [25]. In the present work, Feed Forward Error Back Propagation (FFEBP) algorithm was used for training of the ANN (Fig. 2) and TANSIG and PURELIN were activation functions used for the data transformation

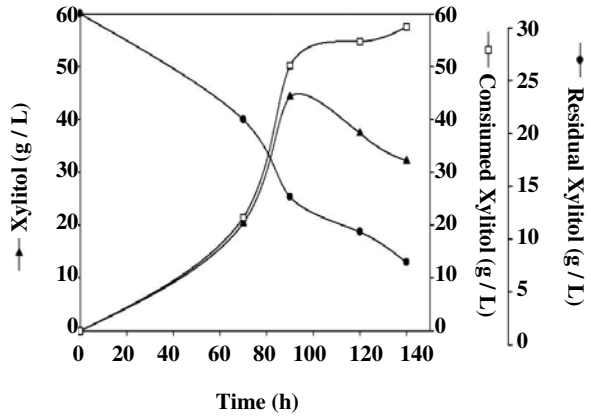

Fig. 4: Time course of xylitol production by P. guilliermondii under the optimal conditions which were as follows: AmS concentration 5 g/L; G:Xyl ratio 1: 2; aeration conditions 50: 250 mL/ mL (experiment no. 8 – see the text for the details).

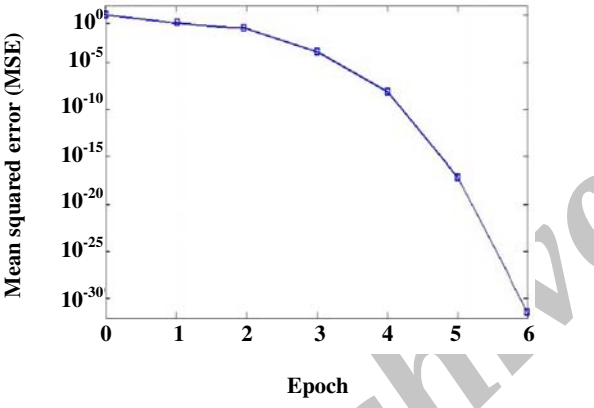

Fig. 5: The performance evaluation during iterative training of the ANN.

from input to the hidden and from the latter layer to the output (xylitol productivity), respectively. Among different methods of computation and storage requirements (memory efficiency) for the network training, LM algorithm was selected. The data obtained experimentally according to the RSM arrangement were used as input data for developing the ANN. According to the data set shown in Table 2 the total data of 15 points were randomly divided into 50 % for the network training process and the remaining 50 % of the total data were used for the validation and the testing steps (each 25% of the total data set). Similar approach has been taken by the others [26].

Fig. 5 shows the network performance during the training process (MSE versus epoch) and the training process was stopped as the MSE reached the lowest value (2.29×10^{-8}). In fact, the hidden neurons number is related to the converging performance of the output error function during the network training. LM algorithm was used for testing several ANN models having the different number of hidden neurons. Finding an optimal neural network architecture and topology is of critical importance for the efficient ANN development. Start of the learning process was with the simplest topology (3-4-1), an architecture that was not able to give the predictive modeling. Low number of hidden neurons corresponds to limited capability of the ANN learning during the weight adjustment while using increased number of the constructed variables in the hidden layer may allow the weight adjustment occurs freely and the network learns the noise in the database during the training process [26]. The selection study was continued with increased number of the hidden neurons and despite of the complexity of the neural network, the optimal number of neurons in the hidden layer was obtained by the trial and error method. Moreover, the topology selection process was ended when the selected architecture was able to give the optimal values employed as the indexes of the system performance. In Fig. 6 MSE and R^2 both are shown as the function of the hidden neurons number. Therefore, six hidden neurons were selected as the optimal number of the constructed variables in the hidden layer. This selection was further confirmed with use of the appropriate sets of data for the validation and testing steps. Fig. 7 (a) presents the network performance with respect to the set of data used in the training step. The accuracy of the developed ANN model is always evaluated with a set of data, which has not been used in the training step (a set of data newly seen by the network (Fig. 7 (b)). The quality of the results predicted by the RSM modeling is shown graphically in Fig. 7 (c) as can be seen, the constructed quadratic model fits very well the data ($R^2 > 0.98$).

The predicted values for the xylitol productivity for either ANN or RSM were presented in Table 3. The comparative values for MSE and R^2 were also included. These values show capability of the model developed here in describing the fermentation process.

The weights for input layer nodes to the hidden layer neurons (w_{ij}) and latter layer to the output layer neuron (v_{jk})

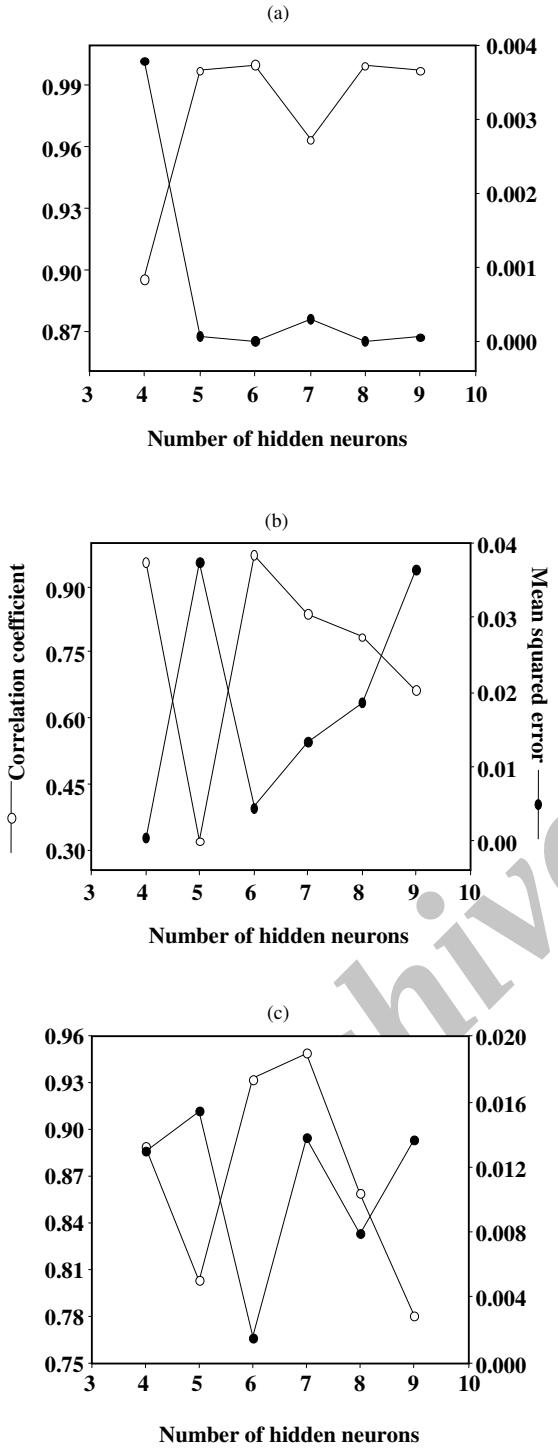

Fig. 6: Selection of the hidden neuron number with respect to training data (a), validating data (b), and testing data (c) for xylitol productiviy, mean squared error (●) and correlation coefficient (○).

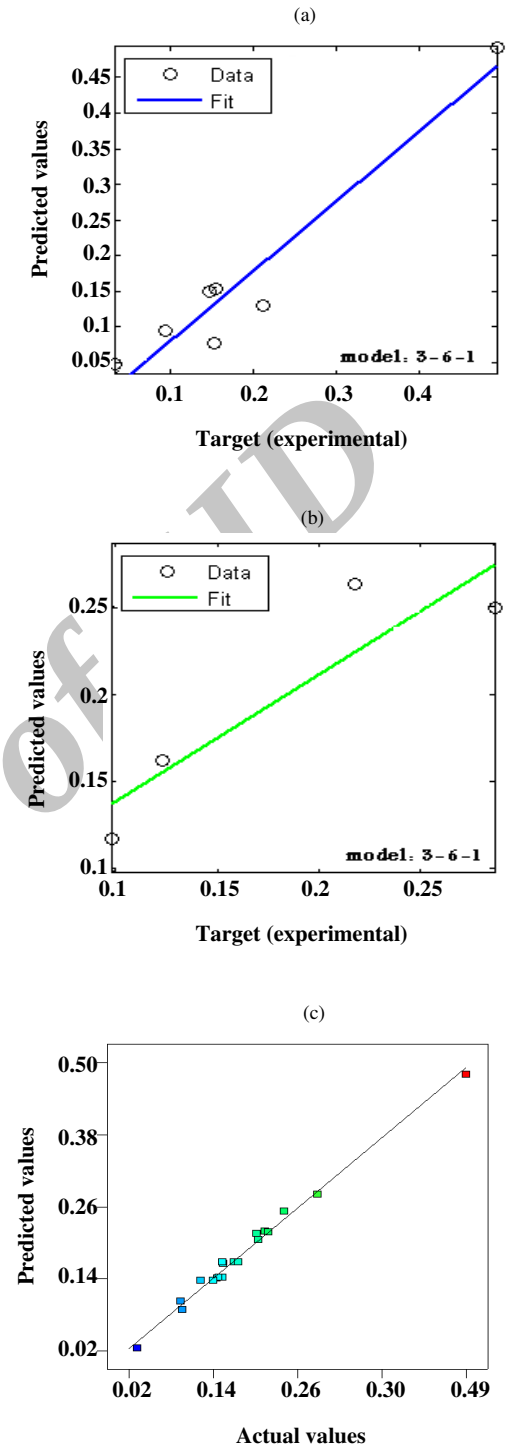

Fig. 7: Plot of experimental versus predicted values for xylitol productivity according to obtained training data (a), and testing data (b) used in the ANN method; plot of the RSM is also presented (c).

Table 3: Predicted values for xylitol productivity obtained by RSM and ANN modeling (a). Comparitive values of R^2 and MSE are also shown (b)

(a)

Experiments	Predicted values	
	productivity (g/L/h)	
	RSM	ANN
1	0.25	0.1831
2	0.14	0.1455
3	0.09	0.1016
4	0.21	0.2054
5	0.14	0.1167
6	0.16	0.1745
7	0.22	0.2171
8	0.48	0.5294
9	0.22	0.2164
10	0.29	0.3048
11	0.15	0.0681
12	0.23	0.2154
13	0.04	0.0980
14	0.11	0.1715
15	0.15	0.1699

(b)

Response	Parameter*	RSM	ANN		
			training	validating	testing
productivity	MS	2.327×10^{-4}	2.29×10^{-8}	4.56×10^{-3}	1.52×10^{-3}
	R^2	0.9817	0.9999	0.9741	0.9325

** Relationship between coefficient of determination and correlation coefficient: coefficient of determination = (correlation coefficient)^2*

in the developed ANN for the Q_p are presented in Table 4. The same Table includes the biases (b_j, b'_k) associated with the hidden layer nodes and with the output layer neuron (Table 4(b)).

CONCLUSIONS

Optimization of bioconversion processes through developing statistical models is a practical approach in highlighting yeast performance in selectively use one substrate over the other i.e., an inducing substrate that is metabolized according to catalytic activities of some inducible enzyme(s) while consumption of the other substrate is under influence of some constitutive enzymes. The capability of the ANN as nonlinear statistical modeling tool was successfully applied in this study to define the relationship between the three test predictor variables and the output in xylitol production using one

Table 4: Weights in the developed ANN for xylitol productivity (Q_p) (a). Values for the bias associated with the neurons in the developed ANN are also given (b).

(a)	
Q_p	
w_{ij} (i =1 to 3; j = 1 to 6)	v_{jk} (j = 1 to 6; k = 1)
w_{11} = - 0.1167	v_{11} = 0.5291
w_{12} = - 1.0001	v_{21} = 0.3079
w_{13} = - 2.8232	v_{31} = -0.4194
w_{14} = 1.7794	v_{41} = -0.3654
w_{15} = 0.0849	v_{51} = -0.5865
w_{16} = - 1.5841	v_{61} = 0.3035
w_{21} = - 2.1109	
w_{22} = - 1.4086	
w_{23} = - 0.5535	
w_{24} = 0.8768	
w_{25} = 0.0459	
w_{26} = - 2.5677	
w_{31} = 0.1979	
w_{32} = -1.3813	
w_{33} = 1.6594	
w_{34} = -2.1581	
w_{35} = -1.5816	
w_{36} = -0.457	

(a)	
bj = (j = 1 to 6)	b'_k (k = 1)
b_1 = 2.4058	b'_k = 0.3212
b_2 = -1.9534	
b_3 = 1.3583	
b_4 = -0.2375	
b_5 = 2.5769	
b_6 = -2.4946	

hidden layer consisted of a set six constructed variables for the productivity. Quality of contribution of the input parameters to the response has been found to be comparable for both RSM and ANN and when one emphasizes on nonlinearity function then, ANN appeared to have higher rank in treating the bioconversion data.

Nomenclatures

AmS	Ammonium sulfate
ANOVA	Analysis of variance
ANN	Artificial neural network
CCD	Central composite design
R^2	Coefficient of determination
CV	Coefficient of variation
df	Degrees of freedom
DOE	Design of experiment
ETC	Electron transport chain
FFEBP	Feed forward error back propagation
Fructose 6- p	Fructose 6- phosphate
G:Xyl	Glucose to xylose
glyceraldehyde 3- p	Glyceraldehyde 3- phosphate
x_i,x_j's	Independent variables
LM	Levenberg Marquardt
PURELIN	Linear transfer function
MSE	Mean squared error
PPP	Pentose phosphate pathway
PRESS	Predicted sum of squares
β_i, β_{ij}'s	Regression coefficients of the fitted quadratic models
RSM	Response surface methodology
ribose 5- p	Ribose 5- phosphate
ribulose 5- p	Ribulose 5- phosphate
TANSIG	Tangent sigmoid
XDH	Xylitol dehydrogenase
Q_p	Xylitol productivity [g/L/h]
XR	Xylose reductase

Received : Apr. 7, 2011 ; Accepted : July 3, 2011

REFERENCES

[1] Mayerhoff Z.D.V.L., Roberto I.C., Franco T.T., Response Surface Methodology as an Approach to Determine the Optimal Activities of Xylose Reductase and Xylitol Dehydrogenase Enzymes from *Candida mogii*, *Appl. Microbiol. Biotechnol.*, **70**, p. 761 (2006).

[2] Mussatto S.I., Silva C.J.S.M., Roberto I.C., Fermentation Performance of *Candida guilliermondii* for Xylitol Production on Single and Mixed Substrate Media, *Appl. Microbiol. Biotechnol.*, **72**, p. 681 (2006).

[3] Jeppsson H., Alexander N.J., Hahn-Ha¨Gerdal B., Existence of Cyanide-Insensitive Respiration in the Yeast *Pichia stipitis* and its Possible Influence on Product Formation During Xylose Utilization, *Appl. Environ. Microbiol.*, **61**, p. 2596 (1995).

[4] Oh D.K., Kim S.Y., Kim J.H., Increase of Xylitol Production Rate by Controlling Redox Potential in *Candida parapsilosis*, *Biotechnol. and Bioeng.*, **58**, p. 440 (1998).

[5] Stanier R.Y., Ingraham J.L., Wheelis M.L., Painter P.R., "General Microbiology", Fifth ed. Macmillan Education LTD., ISBN: 0-333-41768-2 (1986).

[6] Stryer L., "Biochemistry" (4th Ed.) WH Freeman Co., NewYork: ISBN: 0-7167-2009-4 (1995).

[7] Hsiao H., Chiang L., Ueng P.P., Tsao G.T., Sequential Utilization of Mixed Monosaccharides by Yeasts, *Appl. & Environ. Microbiol.*, **43**(4), p. 840 (1982).

[8] Kumar S., Gummadi S.N., Metabolism of Glucose and Xylose as Single and Mixed Feed in *Debaryomyces nepalensis* NCYC 3413: Production of Industrially Important Metabolites, *Appl. Microbiol. Biotechnol.*, **89**, p. 1405 (2011).

[9] Nolleau V., Preziosi-Belloy L., Delgenes J.P., Navarro, J.M., Xylitol Production from Xylose by Two Yeast Strains: Sugar Tolerance, *Curr. Microbiol.*, **27**, p. 191 (1993).

[10] Parajo J.C., Dominguez H., Dominguez J.M., Biotechnological Production of Xylitol Part1: Interest of Xylitol and Fundamentals of its Biosynthesis, *Bioresource Technol.*, **65**, p. 191 (1998).

[11] Suryadi H., Katsuragi T., Yoshida N., Suzuki S., and Tani Y., Polyol Production by Culture of Methanol-Utilizing Yeast, *J. Biosci. & Bioeng.*, **89**, p. 236 (2000).

[12] Barbosa M.F.S., Medeiros T.M.M.B., Mancilha I.M. Lee H.S., Screening of Yeasts for Production of Xylitol from D-Xylose and Some Factors Which Affect Xylitol Yield, *Candida guilliermondii*, *J. Ind. Microbiol.*, **3**, p. 241 (1988).

[13] Leathers T.D., Gupta S.C., Xylitol and Riboflavin Accumulation in Xylose-Grown Cultures of *Pichia guilliermondii*, *Appl. Microbiol. Biotechnol.*, **47**, p. 158 (1997).

[14] Santos D.T., Sarrouh B.F., Rivaldi J.D., Converti A., Silva S.S., Use of Sugarcane Bagasse as Biomaterial for Cell Immobilization for Xylitol Production, *J. Food Eng.*, **86**, p. 542 (2008).

[15] Walther T., Hensirisak P., Agblevor F.A., The Influence of Aeration and Hemicellulosic Sugars on Xylitol Production by *Candida tropicalis, Bioresource Technol.*, **76**, p. 213 (2001).

[16] Aguiar W.B., Faria L.F.F., Cuoto M.A.P.G., Araujo O.Q.F., Pereira N., Growth Model and Prediction of Oxygen Transfer Rate for Xylitol Production from D- Xylose by *Candida guilliermondii*, *Biochem. Eng. J.*, **12**, p. 49 (2002).

[17] Arranda-Barradas J.S., Garibay-Orijel C., Badillo-Corona J.A., A Stoichiometric Analysis of Biological Xylitol Production, *Biochem. Eng. J.*, **50**, p. 1 (2010).

[18] Agatonovic-Kustrin S., Beresford R., Basic Concepts of Artificial Neural Network (ANN) Modeling and Its Application in Pharmaceutical Research, *J. Pharmaceutical and Biomedical Analysis*, **22**, p. 717 (2000).

[19] Bas D., Boyaci I.H., Modeling and Optimization II: Comparison of Estimation Capabilities of Response Surface Methodology with Artificial Neural Networks in a Biochemical Reaction, *J. Food Eng.*, **78**, p. 846 (2007).

[20] Myers R., Montgomery D., "Response Surface Methodology- Process and Product Optimization Using Designed Experimental", New York: John Wiley & Sons, Inc. ISBN 0-471-41255-4 (2002).

[21] Vining, G.G., "Statistical Methods for Engineers", Brooks/Cole publishing Co. A division of International Thomson Publishing Inc. ISBN: 0-534-23706-1(1998).

[22] Torrecilla J.S., Mena M.L., Yanez-Sedeno P., Garcia J., Field Determination of Phenolic Compounds in Olive Oil Mill Wastewater by Artificial Neural Network, *Biochem. Eng. J.*, **38**, p. 171 (2008).

[23] Ebrahimpour A., Abd Rahman R.N.Z.R., Chng D.H.E., Basri M., Salleh A.B., A Modeling Study by Response Surface Methodology and Artificial Neural Network on Culture Parameters Optimization for Thermostable Lipase Production from a Newly Isolated Thermophilic *Geobacillus* sp. Strain ARM, *BMC Biotechnology,* 8, p. 96 (2008).

[24] Rodrigues R.C., Lu C., Lin B., Jeffries T.W., Fermentation Kinetics for Xylitol Production by a *Pichia stipitis* D: Xylulokinase Mutant Previously Grown in Spent Sulfite Liquor, *Appl. Biochem. Biotechnol.*, **148** (1-3), p. 199 (2008).

[25] Ahadian S., Moradian S., Sharif F., Prediction of Time of Time of Capillary Rise in Porous Media Using Artificial Neural Network (ANN), *Iranian Journal of Chemistry and Chemical Engineering*, **26**(1) (2007).

[26] Abdul-Rahman M.B., Chaibakhsh N., Basri M., Salleh A., Zaliha R.N., Abdul-Rahman R., Application of Neural Network for Yield Prediction of Lipase- Catalyzed Synthesis of Dioctyl Adipate, *Appl. Biochem. Biotechnol.*, **158** (3), p. 722 (2009).