MEHRI ESFAHANIAN
MARYAM NIKZAD
GHASEM NAJAFPOUR
ALI ASGHAR GHOREYSHI

Faculty of Chemical Engineering, Noushirvani University of Technology, Babol, Iran

MODELING AND OPTIMIZATION OF ETHANOL FERMENTATION USING *Saccharomyces cerevisiae*: RESPONSE SURFACE METHODOLOGY AND ARTIFICIAL NEURAL NETWORK

In this study, the capabilities of response surface methodology (RSM) and artificial neural networks (ANN) for modeling and optimization of ethanol production from glucose using Saccharomyces cerevisiae in batch fermentation process were investigated. The effects of three independent variables in a defined range of pH (4.2-5.8), temperature (20-40 °C) and glucose concentration (20-60 g/l) on the cell growth and ethanol production were evaluated. The results showed that the prediction accuracy of ANN was apparently similar to RSM. At optimum conditions of temperature (32 °C), pH (5.2) and glucose concentration (50 g/l), suggested by the statistical methods, the maximum cell dry weight and ethanol concentration obtained from RSM were 12.06 and 16.2 g/l, whereas experimental values were 12.09 and 16.53 g/l, respectively. The present study showed that using ANN as a fitness function, the maximum cell dry weight and ethanol concentration were 12.05 and 16.16 g/l, respectively. Also, the coefficients of determination for biomass and ethanol concentration obtained from RSM were 0.9965 and 0.9853 and from ANN were 0.9975 and 0.9936, respectively. The process parameters optimization was successfully conducted using RSM and ANN; however, prediction by ANN was slightly more precise than RSM. Based on experimental data, the maximum yield of ethanol production of 0.5 g ethanol/g substrate (97% of theoretical yield) was obtained.

Keywords: artificial neural network, ethanol fermentation, response surface methodology, Saccharomyces cerevisiae, ethanol yield.

Fermentation is a complicated phenomenon intensely investigated in many bioprocesses [1]. The cost of alcohol production needs to be lowered *via* fermentation processes. Today, bioethanol production by fermentation is one of the popular subjects in the world with regards to the biological environment and economic challenges. In fact, bioethanol feedstock is produced from biomass which is depends on solar energy for converting simple raw lignocellulosic materials to chemical fuel via biological routes [2]. In this view, knowing the optimum condition and estimation of bioethanol production from glucose can be very useful in industrial applications as the main goals in the present work, because any kind of raw materials as carbon sources first must be converted to glucose before the ethanol fermentation is performed. The fermentation process has both nonlinear and dynamic properties. Modeling such a process is difficult and challenging [1]. Considerable attempts have been made by several researchers to propose a methodology based on mathematical models [3,4]. A major disadvantage of these models is that they need a large number of experiments and it is often very complicated to describe the experimental observation [5]. A way of dealing with such problems is to use simple and efficient models, such as response surface methodology (RSM) and artificial neural network (ANN).

Modeling and optimization to enhance the efficiency of a process is one of the most significant stages in a biochemical process [6]. The conventional one-factor-at-a-time approach of optimization is not only tiresome, but also ignores the merged interactions of each factor [7]. One of the most common optimization methods used in the last two decades is the Response Surface Methodology (RSM). RSM is a statistical technique based on the essential principles of statistics, randomization and duplication, which makes the optimization easier by studying the reciprocal interactions among the variables over a wide range of values in a statistically logical manner [8]. Thus, RSM is an effective approach to deal with a large number of variables and there are various reports on the application of RSM for fermentation [9,10]. Although RSM has so many advantages, it is hard to say that it is applicable to all optimization and modeling studies [6].

Artificial Neural Network (ANN) has also been successfully applied in modeling of biological systems [10,11]. The advantages of ANN are that they need less time for development of the extensive experimentation. That is performed with limited numbers of experiments to predict the degree of non-linearity and their capability to learn complex relationships without requiring the knowledge of the model structure [12]. Moreover, it is a superior and more accurate modeling technique than RSM, as ANN depicts the non-linearity of the model [7]. ANN cannot produce a model equation similar to RSM, but it estimates the response based on the trained data [6].

In the present work, ANN and RSM were used for modeling and optimization of ethanol production as a function of temperature, pH and glucose concentration in a batch fermentation using limited experimental runs. The accuracy of the estimated data was defined and then the overall prediction ability of these two techniques was assessed.

MATERIALS AND METHODS

Microorganism

The pure stock culture of *Saccharomyces cerevisiae* was used for ethanol production. The strain was originated from Persian Type Culture Collection (PTCC 24860), supplied by the Iranian Research Organization for Science and Technology (IROST). The medium used for cultivation contained glucose, NH_4Cl and yeast extract: 10, 0.45 and 1 g/l, respectively. The medium was autoclaved at 121 °C and 15 psig for 20 min. The sterilized medium was inoculated with 5% of pure seed culture of the microorganism and then the culture was cultivated in an incubator and shaker at 30 °C for 24 h.

Growth media and culture conditions

The media were used for ethanol production contained various concentrations of glucose (20, 30, 40, 50 and 60 g/l), yeast extract and NH_4Cl (3 and 5 g/l). A suitable proportion of potassium hydrogen phthalate 0.1 M and sodium hydroxide 0.1 M, as buffer solution for pH values of 4.2, 4.6, 5, 5.4 and 5.8 were prepared. The media were autoclaved at 121 °C and 15 psig for 20 min. The sterilized media were inoculated and cultivated in an incubator at 30 °C for 24 h. Batch fermentation experiments were carried out in 250 ml shaking flasks with working volume of 200 ml medium in incubator shaker (IKA, Japan) at various incubation temperatures (20, 25, 30, 35 and 40 °C) and 200 rpm for 30 h. Samples were periodically taken from the culture every 2 hours for monitoring the cell optical density, determination of the cell dry weight and concentration of produced ethanol. The optical density was measured at 620 nm by a spectrophotometer (Unico, USA). The cell dry weight was calculated based on a pre-developed calibration curve. The calibration curve was obtained based on optical density of broth solution with certain concentration of microorganism at 620 nm by a spectrophotometer. The experiments were conducted in triplicates and analysis of each sample was performed in duplicate. The volume changes caused by samples at the end of each run were less than 7%.

For determination of the glucose concentration, 2 ml sample was collected and cells were separated by centrifugation at 7000 g for 7 min by a micro centrifuge Hermle, model: Z 233 M-2 (Germany). The concentration of glucose in the media was determined by color-metric method using DNS reagent [13].

Ethanol analysis

The produced ethanol concentration was measured using a gas chromatograph (Agilent, 7890A) equipped with a flame ionization detector (FID). The stainless steel packed column 1.83 m length and 2.1 mm I.D, 80/100 mesh Porapak Q (Supelco, USA) was used. The initial oven temperature was set at 120 ºC and hold for 1 min. Then, the oven temperature was increased at a rate of 40 °C/min until it reached to 185 °C and remained at this temperature for 8.5 min. The detector temperature was 225 °C. The carrier gas was nitrogen with a flow rate of 30 ml/min. The propionic acid (Merck, Germany) or 2-propanol (Merck) prepared in 3 vol.% was used as internal standard. An exact volume of 50 µl the internal standard was added into 0.5 ml of the sample.

Response surface methodology

One of the best models to describe the biological processes in analytical methods is quadratic model in comparison with linear and cubic equation. The relationship among the three factors is expressed in a second-order equation (Quadratic model):

$$Y = a_0 + \sum_{i=1}^{3} a_i X_1 + \sum_{i=1}^{3} \sum_{j=1}^{3} a_{ij} X_i X_j \qquad (1)$$

where Y is the predicted value, a_0 is a constant and a_i and a_{ij} are the regression coefficients of the RSM model. The X_i and X_j are the factor variables [7]. Statistical analysis of the experimental data was performed to evaluate the analysis of variance (ANOVA) using Design Expert software 7.01 (Stat-Ease, Inc, Minneapolis, MN, USA). In this study, maximum cell dry weight (X_{max}) and ethanol concentration were investigated as a function of temperature, pH and glucose concentration. *S. cerevisiae* was used in the same condition. Table 1 shows the coded values of different media factors and level of independent variables.

Coding of the variables was done according to the following equation:

$$n_{Coded} = \frac{N_{Actual} - (N_{Hi} - N_{Low})/2}{(N_{Hi} - N_{Low})/2} \qquad (2)$$

where n_{Coded} is the coded variables and N_i is the actual variable.

All experiments were carried out in 250 ml flasks with working volume of 200 ml medium in an incubator-shaker at an agitation rate of 200 rpm for duration of 30 h. The predicted values of the cell dry weight and ethanol concentration were based on the equation incorporated in the software. In fact, the predicted values were the responses of software for modeling in comparison to the actual values. Thus, the mathematical regression model for the maximum cell dry weight and ethanol concentration responses fitted in terms of coded factors was obtained as stated in Eqs. (8) and (9).

Artificial neural network

Artificial neural network (ANN) is a mathematical tool that tries to represent the low-level intelligence in natural organisms and it is a flexible structure, capable of making a non-linear mapping between input and output spaces [14]. Especially in the past two decades, it has been greatly used in many fields of science and engineering. ANN consists of many simple computational elements, called nodes or neurons, organized in layers, operating in parallel [15]. Neurons are connected by weights that are modified during the learning phase [14]. All neural networks have three main layers which are called input, hidden and output layers [16]. Many classes of neural networks exist in the literature such as, feed forward back propagation, recurrent neural networks, cascade correlation neural networks and radial basis function neural networks. All types of these network architectures have the same elements: neurons, layers and weights. The most common type of ANN in chemical engineering application is the multi-layer perceptron (MLP), which is a back propagation feed forward network [17]. In this work, a multi-layer feed forward neural network has been used. MLP network consists of input and output layers, with several, but usually only one hidden layer [18]. In this particular ANN, information moves in only one direction, forward from the input layer, through the hidden layer and to the output layer [16]. The number of neurons in the input layer depends on the number of independent variables and the number of neurons in the output layer which corresponds to the dependent variables. In this work, the independent variables were temperature, pH and glucose concentration and the output variable were ethanol concentration and maximum cell dry weight. The schematic representation of the neural network used for this work is illustrated in Figure 1.

It is necessary to train an artificial neural network before using it for a particular application [16]. Feed forward network training starts by applying the input vector to the input layer having network processing element [17]. During the training, the network learns to create new outputs through a repetitive method [16]. Generated outputs by network are compared to the target. The network is adjusted, based on a comparison of the output and the target, until the network output matches the target. Basically, the purpose of training patterns is to reduce the global error. The aim of every training algorithm is to decrease this global error by adjusting the weights and biases [14].

Table 1. Experimental range of different media factors and level of independent variables

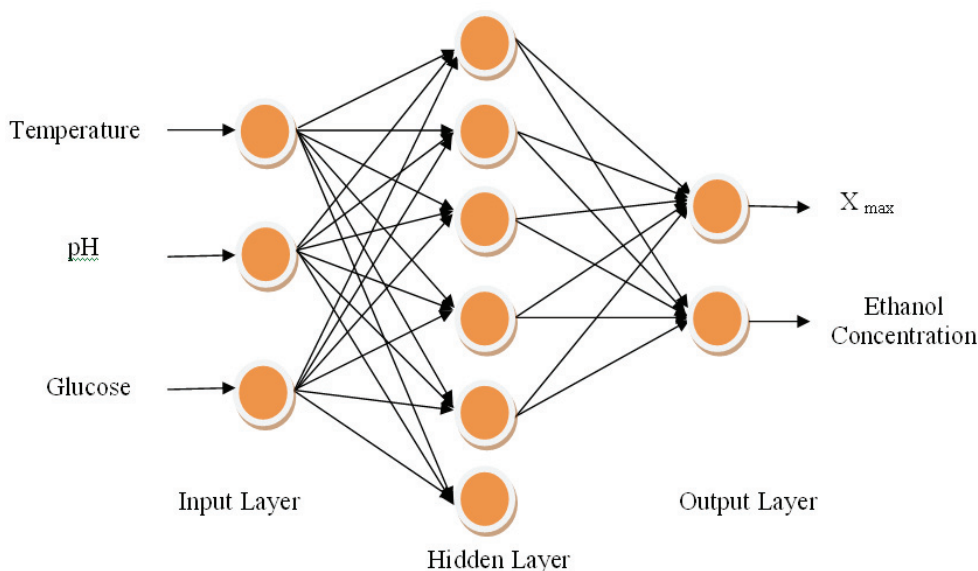| Factor	Symbol	Range and levels				
		-2	-1	0	+1	+2
Temperature, °C	A	20	25	30	35	40
pH	B	4.2	4.6	5	5.4	5.8
Glucose concentration, g/l	C	20	30	40	50	60

Figure 1. Schematic representation of the three layer feed forward neural network.

After training, the trained network was tested by testing data to further check if the network achieved good generalization. In this work, input and output data were scaled between 0 and 1, before training network as follows:

$$
(\text{Scaled})_{\text{value}} = \\
= \frac{(\text{Actual})_{\text{value}} - \text{Minimum}_{(\text{Actual value})}}{\text{Maximum}_{(\text{Actual value})} - \text{Minimum}_{(\text{Actual value})}} \quad (3)
$$

Input data were randomized into three sets: learning, validation and testing. Usually, 30% of data are used for testing and the remaining 70% for training and validation [16]. Activation function of the hidden layer was "logsig" and the one in the output layer was "purelin".

$$
\text{logsig}(x) = \frac{1}{1 + e^{-x}} \quad (4)
$$

$$
\text{purelin}(x) = x \quad (5)
$$

Training of the network was performed with the function of "trainlm", which updates weight and bias values according to the Levenberg-Marquardt optimization. In learning of the network, "learngd" was used as the adaption learning function. The maximum training epochs were 1000, and the mean square error was 0.0001. The other parameters of neural network were taken as defaults of neural network toolbox, MATLAB. The performance of the ANN was statistically measured by the mean squared error (MSE) and the coefficient of determination (R^2) as follows:

$$
MSE = \frac{\sum\limits_{i=1}^{n} \left(X_{\text{pre},i} - X_{\text{exp},i} \right)^2}{n} \quad (6)
$$

$$
R^2 = 1 - \frac{\sum\limits_{i=1}^{n} \left(X_{\text{pre},i} - X_{\text{exp},i} \right)^2}{\sum\limits_{i=1}^{n} \left(X_{\text{exp},i} - \bar{X} \right)^2} \quad (7)
$$

where $X_{\text{pre},i}$ is the predicted output from observation i, $X_{\text{exp},i}$ is the experimental (target) output from observation i, $\bar{X}$ is the average value of experimental output, and n is the total number of data.

RESULTS AND DISCUSSION

Response surface methodology

RSM is a frequently useful technique for modeling and determining the optimal process conditions. Based on the experimental results of the central composite design (CCD) in Tables 2 and 3, a second-order polynomial equation (quadratic model) was established to identify the relationship between maximum cell dry weight, produced ethanol concentration and three batch fermentation conditions.

Statistical testing of the model for each response was done by Fisher's statistical test for analysis of variance (ANOVA). The fitted polynomial equation was expressed as three-dimensional surface plots to imagine the relationship between responses and experimental levels of each factor used in the design [19]. The goodness of fit of the model can be investigated by determination coefficient (R^2) and using adjusted R^2 and predicted R^2. The value of R^2 is

Table 2. CCD matrix of three factors and experimental data, RSM and ANN for determined values of maximum cell dry weight

A	B	C	X_{max} / g l^{-1}				
				RSM		ANN	
Temperature, °C	pH	Concentration, g/l	Experimental value	Model value	Relative error, %	Model value	Relative error, %
25	4.6	30	9.30	9.250	0.54	9.321	0.22
35	4.6	30	10.9	10.945	0.41	10.923	0.21
25	5.4	30	10.83	10.835	0.04	10.812	0.16
35	5.4	30	10.99	11.041	0.46	10.975	0.13
25	4.6	50	10.12	10.001	1.17	10.094	0.25
35	4.6	50	11.89	11.835	0.46	11.863	0.22
25	5.4	50	11.80	11.705	0.80	11.692	0.91
35	5.4	50	12.09	12.059	0.25	12.051	0.32
20	5	40	9.40	9.500	1.06	9.441	0.48
40	5	40	11.58	11.540	0.34	11.563	0.14
30	4.2	40	9.93	9.980	0.50	9.900	0.30
30	5.8	40	11.76	11.780	0.17	11.743	0.14
30	5	20	9.60	9.560	0.41	9.541	0.61
30	5	60	11.20	11.320	1.07	11.163	0.33
30	5	40	12.000	12.000	0.00	11.921	0.65
30	5	40	11.990	12.000	0.08	11.970	0.16
30	5	40	11.980	12.000	0.16	12.013	0.27
30	5	40	11.990	12.000	0.08	12.018	0.23
30	5	40	12.000	12.000	0.00	11.897	0.85
30	5	40	12.000	12.000	0.00	11.965	0.29

the percentage variability in the cell dry weight and ethanol concentration values that are defined by this model. It falls between 0 and 1. The adjusted R^2 is the value adjusted down for a higher number of variables in the model which makes it much more useful than the usual value of R^2 and it is suggested to use R^2 adjusted rather than the regular R^2 to estimate the fit of a multiple regression model. Predicted R^2 is used in regression analysis to show how well the model predicts responses for new observations; predicted R^2 can impede over fitting. Larger values of predicted R^2 suggest models of greater predictive ability [20]. In this study, it was aimed to model the microorganism growth, ethanol biosynthesis and also to model culture cell dry weight for enhanced maximal ethanol production. The results from the analysis of variance of the designed experiments showed that the adjusted and predicted R^2 for X_{max} were 0.9933 and 0.9712, and for ethanol concentration 0.9721 and 0.8770, respectively. The very high values of R^2 imply that the model fitted very well and can predict the responses satisfactorily.

After square root transformation of the X_{max} and ethanol concentration, the responses recommended by the Box-Cox plot, regression analysis of the experiments exhibited that the quadratic model terms for

X_{max} (A^2, B^2 and C^2) and interactive model terms (AB) were significant ($P < 0.05$, without insignificant terms). Values of "Prob > F" less than 0.05 indicate that the model terms were significant. Also, the linear model terms (A, B and C) were found to be significant ($P < 0.05$) in modeling of the cell dry weight. However, the interactive model terms (AC and BC) were insignificant ($P > 0.05$). But for ethanol concentration, the linear model terms (A, B and C) and the quadratic model terms (A^2, B^2 and C^2) were significant model terms, while the interactive model terms (AB, AC and BC) were found to be insignificant. It was noted that A, B and C are temperature, pH and glucose concentration, respectively.

By applying the multiple regression analysis, the results were fitted to a second-order polynomial equation. The ANOVA test output can be investigated in Tables 4 and 5. Thus, the mathematical regression model for maximum cell dry weight and ethanol concentration responses fitted in terms of coded factors was obtained as follows:

$$X_{max} = 12 + 0.51A + 0.45B + 0.44C - 0.37AB + \\ + 0.037AC + 0.032BC - 0.37A^2 - \\ - 0.28B^2 - 0.39C^2 \qquad (8)$$

Table 3. CCD matrix of three factors and experimental data, RSM and ANN for determined values of ethanol concentration

A	B	C	Ethanol concentration, g/l				
			RSM			ANN	
Temperature, °C	pH	Concentration, g/l	Experimental value	Model value	Relative error, %	Model value	Relative error, %
25	4.6	30	10.03	10.261	2.30	9.91	1.19
35	4.6	30	10.56	10.859	2.83	10.719	1.50
25	5.4	30	11.94	12.121	1.51	11.832	0.90
35	5.4	30	12.90	13.239	2.62	13.094	1.50
25	4.6	50	13.89	13.459	3.10	13.648	1.74
35	4.6	50	14.62	14.341	1.90	14.900	1.91
25	5.4	50	16.53	16.119	2.48	16.163	2.22
35	5.4	50	17.86	17.521	1.89	17.663	1.10
20	5	40	11.72	11.860	1.19	12.010	2.47
40	5	40	13.93	13.860	0.50	13.951	0.15
30	4.2	40	9.78	9.820	0.40	9.810	0.30
30	5.8	40	14.78	14.860	0.54	14.482	0.41
30	5	20	10.08	9.520	5.55	9.715	3.62
30	5	60	16.32	17.000	4.16	16.461	0.86
30	5	40	15.49	15.500	0.06	15.361	0.83
30	5	40	15.48	15.500	0.12	15.297	1.18
30	5	40	15.47	15.500	0.19	15.402	0.43
30	5	40	15.50	15.500	0.00	15.511	0.07
30	5	40	15.49	15.500	0.06	15.509	0.12
30	5	40	15.48	15.500	0.12	15.541	0.39

$$[\text{Ethanol concentration}] = 15.50 + 0.5A + 1.26B +$$
$$+1.87C + 0.13AB + 0.071AC + 0.20BC - \qquad (9)$$
$$-0.66A^2 - 0.79B^2 - 0.56C^2$$

The analysis of the 3D surface and contour plot of the maximum cell dry weight and produced ethanol concentration using response surface quadratic model, and also based on Eqs. (8) and (9), are shown in Figures 2 and 3. These figures explain the effect of three independent variables: temperature, glucose concentration and pH on the two responses, including X_{max} and ethanol concentration. In order to determine the optimal levels of each variable for X_{max} and maxi-

mum ethanol production, isoresponse contour plots were constructed by plotting the responses (X_{max} and ethanol concentration) on the Z-axis *versus* the two independent variables, while other variables were at their optimal levels, which was useful for understanding both the main and the interaction effects of these two factors. The response surface can be used to predict the optimum range for different values of the test variables, and the main interactions between the experiment variables can be identified from the circular or elliptical nature of the contours. The circular nature of the contours imply that the interactive effects between the test variables are not significant

Table 4. Analysis of variance (ANOVA) for experimental results of the maximum cell dry weight design (Quadratic model); values of "Prob > F" less than 0.05 indicate model terms are significant

Factor	Coefficient estimate	Sum of squares	Degree of freedom	Mean square	Standard error	P-value, Prob > F
A	0.51	4.18	1	4.18	0.020	<0.0001
B	0.45	3.20	1	3.20	0.020	<0.0001
C	0.44	3.13	1	3.13	0.020	<0.0001
AB	-0.37	1.07	1	1.07	0.028	<0.0001
AC	0.037	0.011	1	0.011	0.028	0.2146
BC	0.032	8.450E-3	1	8.450E-3	0.028	0.2775
A^2	-0.37	3.44	1	3.44	0.016	<0.0001
B^2	-0.28	1.99	1	1.99	0.016	<0.0001
C^2	-0.39	3.88	1	3.88	0.016	<0.0001

Table 5. Analysis of variance (ANOVA) for experimental results of produced ethanol concentration design (quadratic model); values of "Prob > F" less than 0.05 indicate model terms are significant

Factor	Coefficient estimate	Sum of squares	Degree of freedom	Mean square	Standard error	P-value, Prob > F
A	0.50	3.97	1	3.97	0.10	0.0006
B	1.26	25.33	1	25.33	0.10	<0.0001
C	1.87	56.06	1	56.06	0.10	<0.0001
AB	0.13	0.13	1	0.13	0.14	0.3907
AC	0.071	0.041	1	0.041	0.14	0.6302
BC	0.20	0.33	1	0.33	0.14	0.1860
A^2	-0.66	10.84	1	10.84	0.081	<0.0001
B^2	-0.79	15.80	1	15.80	0.081	<0.0001
C^2	-0.56	7.96	1	7.96	0.081	<0.0001

and optimum values of the test variables can be easily obtained with regarding to the center point of contour plot. According to Figure 2 (a and c), the contours around the stationary point were circular and it means that the interaction between temperature and glucose concentration and also pH and glucose concentration on maximum biomass concentration were insignificant. On the other hand, according to Figure 2b, the contour around the stationary point was elliptical and it became elongated more and more along the both pH and temperature axis. This means that the interaction between temperature and pH on biomass concentration was significant and a small change of these two variables causes a big change of cell dry weight. It was evident that X_{max} increased in range of 4 to 12 g/l with increasing of the pH and temperature to 5.2 and 32 °C, respectively, at the optimum glucose concentration of 50 g/l and then decreased. The significant interaction between temperature and initial pH were apparent not only from the elliptical nature of the contour plot, but also from the low probability value ($P < 0.05$). In addition, the maximum ethanol concentration was 17.86 g/l at the optimum condition, as illustrated in Figure 3. According to the contour plot of ethanol concentration, the interaction between each pair of the mentioned independent variables shows a less interactive effect.

Artificial neural network

Experimental values for ethanol concentration and maximum cell dry weight in Tables 2 and 3 were used with an ANN to model the fermentation process.

Figure 4 shows the average relative deviation (ARD) of testing data versus number of neurons in the hidden layer. This graph clearly shows that the 3-6-2 topology is the best topology, with a minimum value of error for ethanol concentration and X_{max}. ARD is selected as the criterion for the accuracy of each topology. ARD can be calculated as follows:

$$ARD = \frac{100}{n} \times \sum_{i=1}^{n} \left| \frac{X_{exp,i} - X_{pre,i}}{X_{exp,i}} \right| \qquad (10)$$

Therefore, a 3-6-2 topology was chosen as the best topology for modeling the fermentation and was built with Neural Network toolbox, MATLAB R2009b. The best network model was the network having the minimum *MSE* and maximum R^2 using predicted and experimental data. After training, the trained network was tested using the experimental data to further check if the network achieved good generalization. The results after training and testing are presented in Tables 2 and 3.

From the training and testing process, *MSE* and R^2 between the predicted values and experimental values were obtained. The R^2 for the maximum biomass and ethanol concentrations were 0.9976 and 0.9936, respectively. In addition, the values of *MSE* for the two responses were 0.0022 and 0.0358, respectively.

Comparison of RSM and ANN

After the model was built, the performance of RSM and ANN models was also compared. Results of the two methods are summarized in Tables 2 and 3. Figures 5 and 6 show the plots of predicted X_{max} and ethanol concentration by ANN and RSM against the experimental values, respectively. The results showed that ANN predictions were closer to 45° diagonal; which means perfect prediction than those of RSM. It means that ANN has higher performance than RSM for the prediction of X_{max} and ethanol concentrations.

Also, for similar conditions, the relative errors of ANN related to experimental data were less than RSM. At optimum conditions, the maximum predicted biomass and ethanol concentration with ANN were 12.32 and 17.58 g/l, respectively. These values were
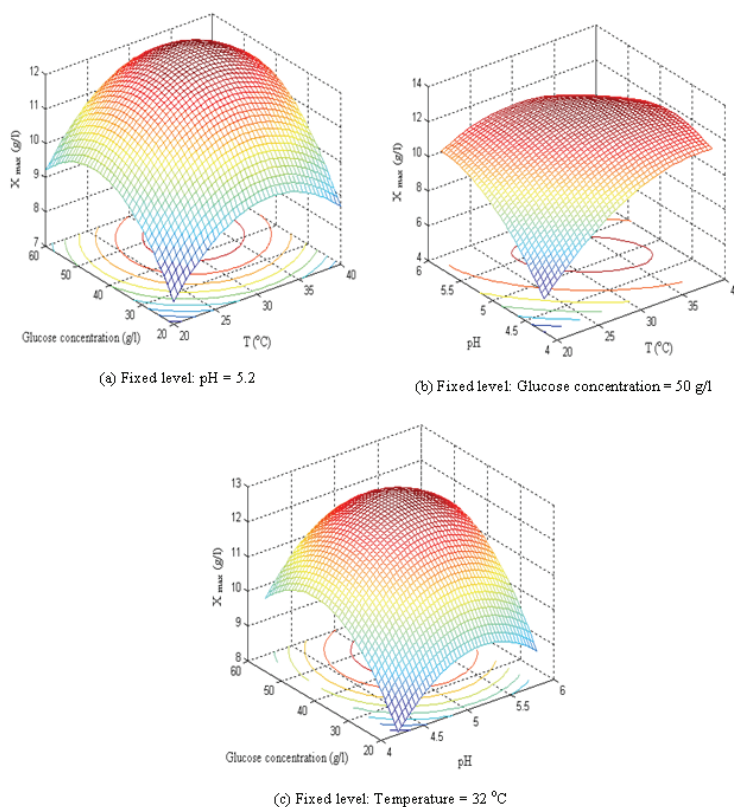
(a) Fixed level: pH = 5.2

(b) Fixed level: Glucose concentration = 50 g/l

(c) Fixed level: Temperature = 32 °C

Figure 2. Three-dimensional response surfaces of maximum cell dry weight showing the interactions between pH, temperature and glucose concentration.
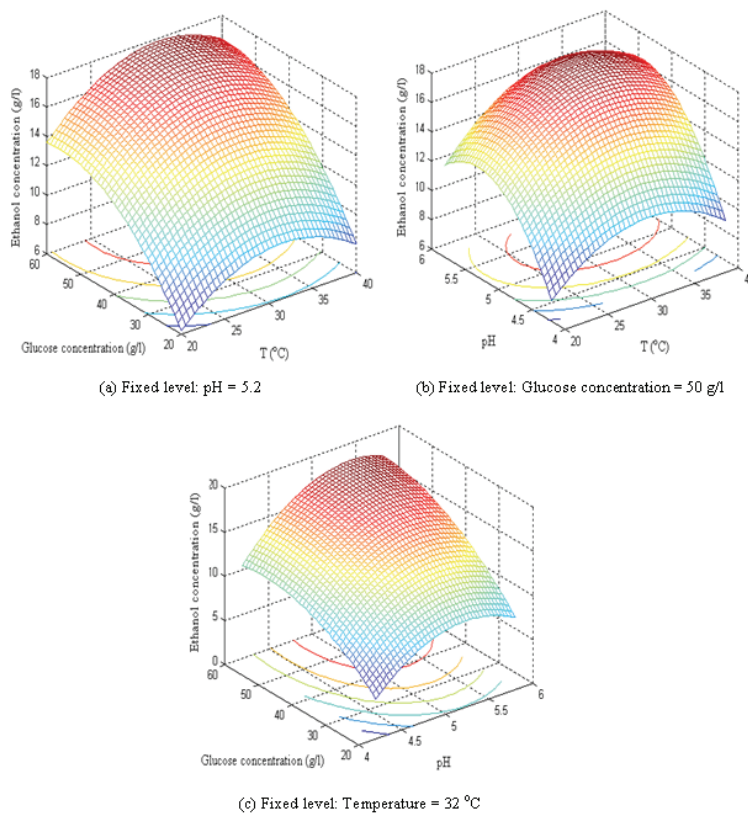

(a) Fixed level: pH = 5.2

(b) Fixed level: Glucose concentration = 50 g/l

(c) Fixed level: Temperature = 32 °C

Figure 3. Three dimensional response surfaces of ethanol concentration showing the interactions between pH, temperature and glucose concentration.
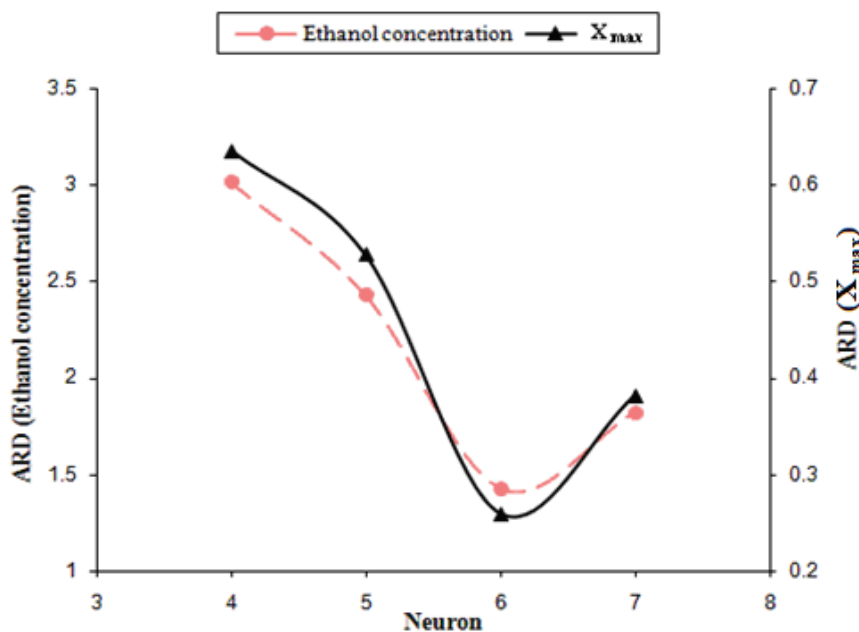
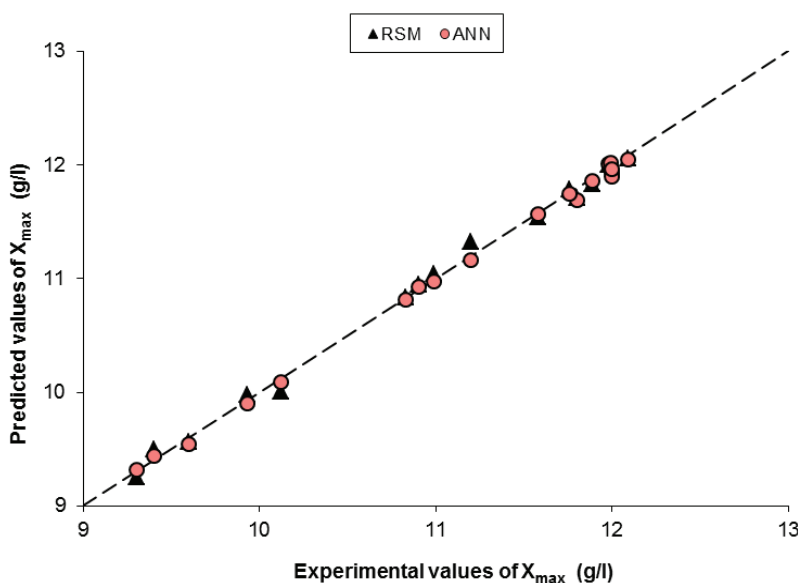Figure 4. ARD of testing data versus neuron number in hidden layer.

Figure 5. The values of X_{max} predicted by ANN and RSM versus the experimental ones.

in good agreement with experimental data, as maximum cell and ethanol concentration at optimum condition were 12.35 and 17.92 g/l, respectively. It is worth to mention that those values by RSM were 12.31 and 17.49 g/l, respectively. Therefore, at optimum conditions, the relative error in RSM was more than ANN. So, ANN is more capable than RSM for prediction of X_{max} and ethanol concentration compared to the obtained experimental data.

Table 6 shows higher R^2, and lower MSE and ARD values obtained by the neural network compared to those of the RSM model. Thus, the accuracy

of neural network model was more desired and was better fitted to the data compared to the RSM method.

Literatures about optimization and modeling of ethanol fermentation by RSM and ANN from various biomass sources and microorganisms as function of different parameters such as temperature, pH, fermentation time and inoculums size were discussed. Ezhumalai *et al.* [10] had investigated optimization of incubation temperature (25-45 °C), pH (5-7) and fermentation time (24-120 h) using RSM and ANN in bioconversion of steam pretreated sugarcane bagasse into ethanol by cellulase and thermotolerant yeast
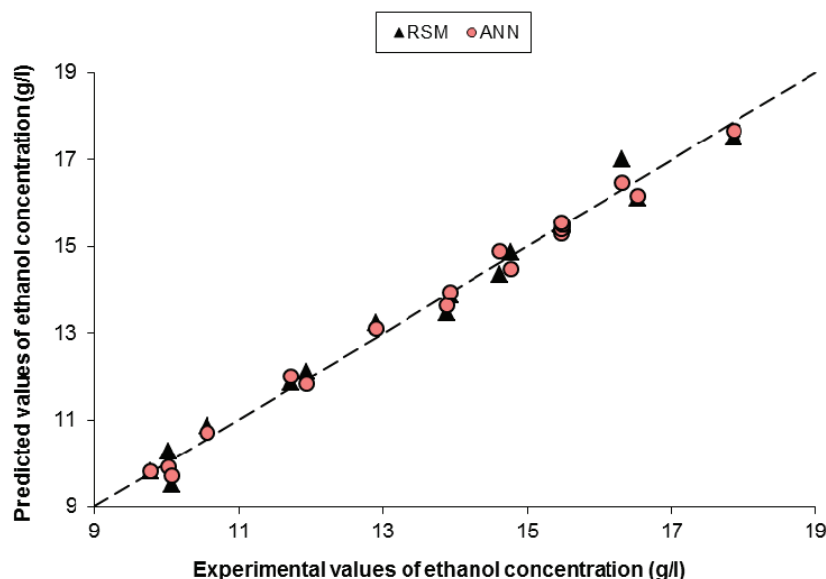
Figure 6. The values of ethanol concentration predicted by ANN and RSM versus the experimental ones.

Table 6. Statistical results of the neural network and RSM models

Model name	Statistical criterions	Ethanol concentration, g/l	X_{max}
ANN	*MSE*	0.0358	0.0022
	R^2	0.9936	0.9975
	ARD	1.1140	0.3435
RSM	*MSE*	0.0824	0.0033
	R^2	0.9853	0.9965
	ARD	1.576	0.395

Kluyveromyces marxianus var. marxianus MTCC 3013. They have reported that the optimum values of temperature, pH and fermentation time were 39 °C, 5.7 and 110 h, respectively. At optimum conditions, maximum ethanol concentration of 5.89 g/l was achieved from 50 g/l pretreated sugarcane bagasse in aerobic batch fermentation. Also, in another investigation [21], studied on bioconversion of lignocellulosic material such as pretreated sugarcane bagasse into ethanol by cellulase and *candida wickerhamii* MTCC 3013 based on CCD experiments. Optimum conditions were obtained at a temperature of 33 °C, pH of 5.7 and fermentation time of 104 h. Maximum ethanol concentration at optimum conditions was 4.28 from 50 g/l pretreated sugarcane bagasse in aerobic batch fermentation. Beside, Yan *et al.* [22] have assessed the optimization of the alcoholic fermentation of blueberry juice by AS 2.316 *S. cerevisiae* wine yeast. Through statistically designed optimization, the optimal conditions of alcoholic fermentation were found to be temperature of 22.65 °C, pH value of 3.53 and inoculum size of 7.37%. At the optimal condition, the production of ethanol and volatile acid of blueberry wine had achieved up to 7.63% and 0.34 g /l, respec-

tively. High ethanol concentration (47g/l) for an immobilized cell reactor, using high substrate concentration (150g/l) has been reported in the literature as the enhanced ICR system was continuous [23]. The yield of ethanol produced in ICR was 0.32 g ethanol/g sugar; however, in this work, a batch process using low concentration of sugar resulted in 0.36 g ethanol/g sugar.

In this work, optimization of ethanol fermentation from glucose using *Saccharomyces cerevisiae* PTCC 24860 was performed by RSM and ANN. The investigated parameters were temperature (20-40 °C), pH (4.2-5.8) and glucose concentration (20-60 g/l). At optimum conditions, including temperature of 32 °C, pH of 5.2 and glucose concentration of 50 g/l obtained by RSM, the maximum ethanol and biomass concentration were 17.92 and 12.35 g/l, which was in good agreement with the predicted values of ANN. In addition, the relative error in RSM was more than ANN only at optimum conditions. Therefore, ANN is more capable than RSM for prediction of X_{max} and ethanol concentration in comparison with experimental data.

CONCLUSION

This study investigated ethanol production using *Saccharomyces cerevisiae* from glucose. Interaction of three independent variables, pH of media (4.2-5.8), temperature (20-40 °C) and glucose concentration (20-60 g/l), in the batch fermentation was investigated. Also, this work described the use of RSM for optimization of ethanol production. Two models were developed for ethanol concentration. In this process, using RSM and ANN methods, maximum cell dry weight prediction was obtained. The best models for RSM and ANN methodologies by quadratic model and multilayer feed forward network were obtained. The optimum process conditions, glucose concentration of 50 g/l, pH value of 5.2 and media temperature of 32 °C were obtained. Finally, these two methods were compared for their predictive abilities. The highest R^2 and lowest MSE and ARD for ANN model with high accuracy and capability of prediction of the given model were identified.

REFERENCES

[1]   J.R. Kwiatkowski, A.J. McAloon, F. Taylor, D.B. Johnston, Ind. Crops Prod. **23** (2006) 288-296

[2]   H. Shafaghat, G.D. Najafpour, P.S. Rezaei, M. Sharifzadeh, World Appl. Sci. J. **7** (2009) 140-144

[3]   J. Albiol, J. Robuste, C. Casas, M. Poch Biotechnol. Prog. **9** (1993) 174-178

[4]   P.M. Schlosser, J.E. Bailey, Math. Biosci. **100** (1990) 87--114

[5]   J. Albiol, C. Campmajo, C. Casas, M. Poch Biotechnol. Progr. **11** (1995) 88-92

[6]   [6]D. Bas, I.H. Boyaci, J. Food Eng. 78 (2007) 846-854.

[7]   Y. Zhang, J. Xu, Z. Yuan, H. Xu, Q. Yu Bioresour, Technol. **101** (2010) 3153-3158

[8]   G. Baskar, C. Muthukumaran, S. Renganathan, Int. J. Chem. Biomol. Eng. **1** (2008) 156-160

[9]   R. Balusu, R.R. Paduru, S. Kuravi, G. Seenayya, G. Reddy, Process Biochem. **40** (2005) 3025-3030

[10]   S. Ezhumalai, V. Thangavelu, N. Thanikachalam, Int. J. Recent Sci. Res. **5** (2010) 125-134

[11]   D. Bas, F.C. Dudak, I.H. Boyaci, J. Food Eng. **79** (2007) 622-628

[12]   H. Mingzhi, Y. Ma, W. Jinquan, W. Yan, Expert Syst. Appl. **36** (2009) 5064-5070

[13]   L.C. Thomas, G.J. Chamberlin, G. Shute, S. Tintometer limited, England, Colorimetric chemical analytical methods, Tintometer, 1980

[14]   M. Kashaninejad, A. Dehghani, M. Kashiri, J. Food Eng. **91** (2009) 602-607

[15]   J.A. Silva, E.H. Costa Neto, W.S. Adriano, A.L.O. Ferreira, L.R.B. Goncalves, World J. Microbiol. Biotechnol. **24** (2008) 1761-1767

[16]   B. Mehdizadeh, K. Movagharnejad, Fluid Phase Equilib. **303** (2011) 40-44

[17]   A. Singh, D. Tatewar, P. Shastri, S. Pandharipande, Indian J. Chem. Technol. **15** (2008) 53-58

[18]   D.E. Rumelhart, G.E. Hintont, R.J. Williams, Nature **323** (1986) 533-536

[19]   L. Sharma, A. Kumar Singh, B. Panda, N. Mallick, Bioresour. Technol. **98** (2007) 987-993

[20]   A. Daneshi, H. Younesi, S.M. Ghasempouri, M. Sharifzadeh, J. Chem. Technol. Biotechnol. **85** (2010) 1528-1539

[21]   S. Ezhumalai, V. Thangavelu, Bioresources **5** (2010) 1879-1894

[22]   H.G. Yan, W.H. Zhang, J.H. Chen, Z.E. Ding, Afr. J. Biotechnol **11** (2012) 3623-3630

[23]   G.D. Najafpour, H. Younesi, K. S. Ku, Ismail, J. Bioresour. Technol. **92** (2004) 251-260.

MEHRI ESFAHANIAN
MARYAM NIKZAD
GHASEM NAJAFPOUR
ALI ASGHAR GHOREYSHI

Faculty of Chemical Engineering,
Noushirvani University of Technology,
Babol, Iran

NAUČNI RAD

MODELOVANJE I OPTIMIZACIJA ALKOHOLNE FEREMENTACIJE SA *Saccharomyces cerevisiae*: METODOLOGIJA POVRŠINE ODZIVA I VEŠTAČKA NEURONSKA MREŽA

U radu su ispitane mogućnosti metodologije površine odziva (MPO) i veštačke neuronske mreže (VNM) u modelovanju i optimizacije produkcije etanola iz glukoze primenom Saccharomyces cerevisiae u šaržnom procesu. Ocenjivan je uticaj tri nezavisne promenljive u definisanim opsezima: pH (4,2-5,8), temperatura (20-40 °C) i koncentracija glukoze (20-60 g/l) na mikrobni rast i produkciju etanola. Dobijeni rezultati su pokazali da je tačnost predviđanja VNM i MPO slična. Pri optimalnim uslovima temperature (32 °C), pH (5,2) i koncentracije glukoze (50 g/l), koji su određeni statističkim metodama, maksimalne koncentracije suve biomase i etanola određene MPO bile su 12,06 and 16,2 g/dm^3, dok su eksperimentalne vrednosti bile 12,09 i 16,53 g/dm^3, redom. Primenom VNM određene su maksimalne koncentracije suve biomase i etanola od 12,05 i 16,16 g/l, redom. Vrednosti koeficijenta determinacije za maksimalne koncentracije suve biomase i etanola bile su 0,9965 i 0,9853 za MPO, odnosno 0,9975 i 0,9936 za VNM. Optimizacija procesnih parametara je izvršena uspešno obema metodama, ali su predviđanja malo preciznija VNM nego MPO. Maksimalni prinos etanola je bio 0,5 g/g supstrata (97% od teorijskog prinosa).

Ključne reči: veštačka neuronska mreža, alkoholna ferementacija, metodologija površine odziva, Saccharomyces cerevisiae, prinos etanola.