International Journal of Research in Advent Technology, Vol.6, No.9, September 2018
E-ISSN: 2321-9637
Available online at www.ijrat.org

Usage of Response Surface Methadology and Artificial Neural Networks (ANN) for Optimizing Process Variables to Produce Lipase

Shiva Naresh.M[1*,] G.H.Rao[1]
[1]*Department of Chemical Engineering , ANITS, Visakhapatnam, India*
Email: shivanaresh.che@anits.edu.in, shivanaresh@gmail.com

Abstract: Lipases (triacylglycerol acylhydrolases E.C.3.1.1.3) are one of the most important classes of industrial enzymes. The treatment of wastes by lipases looks to become very important, this includes the breakdown of fat solids, the prevention or cleaning of fat films and the clean of fat containing waste effluents from industries. The present study deals with the production of lipase enzyme in submerged fermentation using the marine yeast *Yarrowia lipolytica* NCIM3589 and fermentation medium optimization for the production of lipase using Box-Behnken design of Response Surface Methodology (RSM) and global optimization routine of Artificial Neural Networks (ANN).

Keywords: Lipase, Response Surface Methodology, Artificial Neural Networks

1. INTRODUCTION

Lipases (triacylglycerol acylhydrolases E.C.3.1.1.3) are one of the most important classes of industrial enzymes. The treatment of wastes by lipases looks to become very important, this includes the breakdown of fat solids, the prevention or cleaning of fat films and the clean of fat containing waste effluents from industries (1-5). The classical fermentation experiments for production of lipase use 'one at a time strategy' of improving fermentation conditions but this approach is time consuming and ignores the combined interactions between physicochemical parameters (6).

Response surface methodology (RSM) overcomes the limitation of 'one time approach' and uses mathematical models to analyze the data and predict the relationship between the yield of the products and the input variables. The response curves drawn from RSM facilitate to find optimum variables with minimum experiments (7-10).

ANN is a modeling technique that predicts the output with limited knowledge of the process. The inputs are given and the weight are given to the network and the output is compared with the target, if the target is not reached the weights are adjusted and the process continues until the target is reached. The entire simulation nowadays is conducted using computer software's like Trazan, MATLAB,STATISTICA etc. (11-19).Many fermentation studies are carried out using RSM and ANN which is a useful tool for bioprocess monitoring and control. (20-22).In this paper we report the optimization of lipase production variables using RSM and ANN.

2. METHODOLOGY

2.1 Microorganism

For the present investigation the yeast *Yarrowia lipolytica*NCIM 3589 was procured from National collection for Industrial Microorganisms (NCIM), Pune, India. This culture was maintained on MGYP medium and stock culture was stored at 5^0C and subcultured weekly in the laboratory. Fresh working slants were prepared whenever required for carrying out the experiments.

2.2 Chemicals

Chemicals were all of analytical grade. The oils used in the medium were purchased from the local market.

2.3 Composition of medium

Production media contains mustard oil 20g; Urea 2g, KH_2PO_4 2.5g; $MgSO_4.7H_2O$ 1.5g; $CaCl_2$ 1.0g; H_3BO_3 2.5mg; $FeCl_3.4H_2O$ 0.1mg; ZnSO4.7H2O 1.0mg; KI 1.0mg; CuSO4.5H2O 0.44mg; Thiamin 200µg; Biotin 8µg are dissolved in 1000 ml distilled water and pH is adjusted to 4.5.

2.4 Inoculum preparation

The yeast *Yarrowia* strain was cultivated in a medium of seed culture. The cells were cultivated in

this medium at 30^0C on a shaker at 200 rpm for 24 hrs.

2.5 Production of lipase by submerged fermentation

The experiments were performed in 250 ml Erlenmeyer flasks containing required quantities of production medium and distilled water is added, so that total medium volume is 50 ml. The entire medium is sterilized in an autoclave (at 121^0C; 15 lb pressure for 20 minutes). The required volume 1% (v/v) of inoculum is added to the above production medium from seed culture aseptically. The flasks were then incubated in an orbital shaker at 30^0C at 185 rpm. After 3 days the fermented liquid was collected in ependiff tubes and centrifuged for 15 min. The supernatant was used to determine enzyme activity. The effect of different oils, effect of level of carbon source, effect of inoculums age and incubation period, effect of inoculums level, effect of pH, effect of different nitrogen sources, effect of ammonium sulphate, effect of salt solution, effect of temperature, on the production of lipase was studied.

2.6 Lipase assay

Lipase activity was determined by olive oil substrate emulsion method. 1.0 ml of substrate emulsion (70 ml Emulsifying reagent (NaCl 17.9gm, KH_2PO_4 0.41ml, Glycerol 540 ml Acacia 100 gm in 1000 ml) + 30 ml olive oil) is taken with 0.8ml of 0.2M potassium and phosphate buffer (pH-7.0) and 1ml of crude enzyme broth. Enzyme substrate mixture is incubated at 37^0C for 30 min. The reaction was terminated by adding 2ml cold acetone ethanol mixture (1:1 v/v). The amount of fatty acids liberated was determined by titration with 0.01N NaOH. Phenolphthalein indicator is used to detect the pink end point. One unit of lipase activity was defined as the amount of enzyme, which liberated 1µmol of fatty acids per minute.

$$Lipase\ activity = \frac{N_2 \times (A - B) \times 10^3}{30}$$

N_2= normality of NaOH; A-B = difference between the volumes of alkali solution consumed for the test solution and for the blank (1 ml of substrate emulsion, 0.8 ml of 0.2 M phosphate buffer (pH 7.0) and 1 ml of distilled water is titrated against 0.01N NaOH).

2.7 Response Surface Methodology

Response surface methodology (RSM) is a collection of mathematical and statistical techniques that are useful for modeling and analysis of problems in which a response of interest is influenced by several problems and the overall objective is to optimize the operating conditions for the system.

Factors (x) → System/Process → Responses (Y)

Temp, pH — Biological system — Yield

The yield (Y) can be expressed in a mathematical form as Y= f(x1, x2, x3) + ε.

Classical biological experiments were performed by studying only one variable at a time which usually results in more experimentation than necessary. It ignores the interaction effect of other variables on the yield of product. An experiment design is composed of a definite number of experimental runs, each of which requires a distinct setting or level of one or more experimental variables or factors, which are direct control of the experimenter. Selection of a particular design is an extremely important step because the most clever data analyst cannot answer a question which the design did not ask. Therefore considerable thought should be given to the choice of the design. In order to evaluate the quadratic coefficients it is necessary for each variable to have at least three levels. The three levels give us an idea which variable causes the curvature. If the responses are well modeled by a linear function of independent variables, the approximating equation is the first order model.

$Y=\beta_0 +\beta_1 x_1+\beta_2 x_2+------\beta_k x_k+ c$

Where β_0 is constant, β_1,β_2…….. β_k are linear co-efficients. $x_1=(A-x_0) / \Delta x$, x_1= coded value of the input variable A, x_0 = value of A at the centre point. Δx = step change.

If there is a curvature in the system, then a polynomial of higher degree must be used, such as second order model as follows.

$$Y = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{i=1}^{k} \beta_{ii} x_i^2 + \sum \sum_{i<j}^{k} \beta_{ij} x_i x_j + c$$

Where β_{ii}, β_{ij} are quadratic and cross product coeffecients respectively.

2.7.1 Location of stationary point

We find the levels of x_1, x_2, x_3…….x_k that optimize the predicted response. This point if it exists will be the set of x_1, x_2, x_3…..x_k for which partial derivatives is zero.

$$\frac{\partial y}{\partial x_1} = \frac{\partial y}{\partial x_2} = \frac{\partial y}{\partial x_3} = - - - - - = \frac{\partial y}{\partial x_k} = 0$$

The point x1, x2,x3….xk is called stationary point. This stationary point could represent (i) a point of maximum response (ii) appoint of minimum response (iv) a saddle point

2.7.2 Box-Behnken Design

. Box-Behnken design is a spherical, revolving design viewd as a cube (Figure 1), consists of a central point and the middle points of the edges. This design requires an experimental number according to $N=k^2+k+c_p$, where (k) is the factor number and $c_p$is the replicate number of the central points.

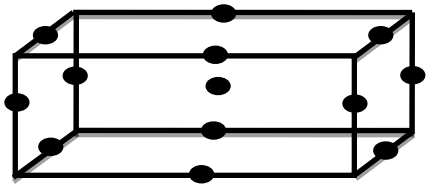

Fig 1: Box-Behnken design: Design as derived from a cube.

The repetition of center points makes one to compute an estimate of the error term that does not depend on the fitted model. For this design except the center point appear at a distance $\sqrt{2}$ from the origin. A three variable Box-Behnken design is used in this investigation. 15 experimental runs were carried out according to this design for a period of three days and results were summarized in next section.A quadratic equation was generated using multiple linear regressionsavailable in STATISTICA software.

2.8 Artificial Neural Network (ANN)

ANN methodology was applied to provide a nonlinear mapping between input variables (mustard oil, ammonium sulphate and salt solution concentration) and the output variable (enzyme production) for the data obtained with Box-Behnken design of RSM. The type of ANN chosen was the back propagation network having a feed forward structure. The choice of the neural network developed in this work (a 3-3-1 structure: three input neurons-three neurons in a hidden layer-one output neuron) was determined by trail and error.

3. RESULTS

3.1 Preliminary experimental runs

The selection of the factor range is extremely important when planning the experimental design; otherwise, after completion of the experimental runs, the optimal conditions obtained either by RSM or ANN may not be found inside the experimental regions.

3.1.1 Effect of different oils

Different oils (2% v/v) jatrophoa oil, niger oil, mustard oil, pongam oil, were studied to select the best oil, which can be used as the carbon source and inducer for the production of lipase. The study was carried out by fixing the process conditions of temperature (30^0C) and inoculum level (1% v/v). The maximum lipase activity 1.44 (U/ml) was observed for mustard oil.

3.1.2 Effect of level of carbon source

The study was carried out by varying the levels of mustard oil (1% - 5% v/v) by fixing the process conditions of temperature (30^0C), inoculum level (1%v/v) and incubation time 3 days. The maximum lipase activity 1.5 U/ml was observed at a concentration of 2%v/v.

3.1.3 Effect of inoculum age and incubation period

By fixing the process conditions of temperature 30^0C, inoculum level (1%v/v) and carbon (2% v/v), the amount of lipase was observed daily during a period of six days. The maximum lipase activity 1.5 U/ml was observed on third day.

3.1.4 Effect of inoculum level

By fixing the process conditions of temperature 30^0C, carbon (2%v/v) and incubation period (3 days) the amount of lipase 1.56 U/ml was maximum for 1%v/v, where the studies are done for levels ranging from 0.5-1.75 %v/v.

3.1.5 Effect of pH

By fixing the process conditions of temperature 30^0C, carbon (2%v/v) and incubation period (3 days),

inoculum level (1%v/v), the amount of lipase produced was maximum (1.56 U/ml) at pH 4.0.

3.1.6 Effect of different nitrogen sources

Different nitrogen sources, urea, ammonium sulphate, yeast extract, ammonium nitrate of 2% w/v were studied to select the best nitrogen source. The maximum lipase activity 1.8 U/ml was observed for ammonium sulphate.

3.1.7 Effect of ammonium sulphate

The study was carried out by fixing the process conditions temperature (30^0C), inoculum level (1% v/v) and carbon source (2% v/v). The maximum lipase activity was observed at 2% w/v concentration of ammonium sulphate.

3.1.8 Effect of salt solution

By fixing the process conditions, the effect of different salt solution level was studied. The maximum lipase yield of 2.1 (U/ml) was obtained at 50% salt solution concentration.

3.1.9 Effect of temperature

As yeast strains are very sensitive towards temperature, a slight change in temperature causes reduction in the yield. By fixing the process conditions of carbon source (2% v/v), inoculum level (1% v/v), salt solution (50% v/v) and incubation time of three days, the experiments were carried out at different temperatures in the range of 25^0C to 40^0C. The maximum lipase activity of 1.96 U/ml at 30^0C.

3.2 Response Surface Methodology

From the results of preliminary experiments, the following factor levels were selected. Mustard oil (05-3.5 % v/v), ammonium sulphate (0.5-3.5% w/v) and salt solution concentration (40 to 60% v/v) as listed in Table 1

Table 1: Process variables and levels

Factors	Lower (-1)	Centre (0)	Upper (1)
Mustard oil	0.5	2.0	3.5
Ammonium sulphate	0.5	2.0	3.5
Salt solution	20	40	60

15 experimental runs were carried out according to Box-Behnken three variable design (Table 2) for a period of three days and the results were summarized in Table 2. The best model for maximizing lipase production by a response surface analysis was the following quadratic polynomial model. The significance of each coefficient was determined by student's t-test and p-values which are listed in Table 3. The larger the magnitude of the t-value and smaller the p-value, the more significant is the corresponding co-efficient. Each experiment run of the Box-Behnken design was conducted three times and the mean value was reported in Table 2. The analysis of lipase assay was done as discussed in section 2.6.

Table 2:Box-Behnken three variable experimental design and comparison of experimental values and predicted values of RSM and ANN

S.No	Coded Variables			Lipase Activity (U/ml)		
	x1	x2	x3	Expt	Predicted RSM	Predicted ANN
1	1	1	0	2.06	2.03	2.06
2	1	-1	0	1.61	1.67	1.60
3	-1	1	0	1.92	1.82	1.61
4	-1	-1	0	1.43	1.42	1.40
5	1	0	1	2.44	2.34	2.99
6	-1	0	-1	1.45	1.42	1.40
7	-1	0	1	2.41	2.40	2.40
8	-1	0	-1	1.42	1.42	1.40
9	0	1	1	1.95	1.97	1.90
10	0	1	-1	1.76	1.72	1.70
11	0	-1	1	2.06	2.03	2.06
12	0	-1	-1	1.00	0.92	1.39
13	0	0	0	2.72	2.81	2.79
14	0	0	0	2.72	2.81	2.79
15	0	0	0	2.72	2.81	2.79

Table 3: Model co-efficient estimated by multiple linear regression

	Co-efficient	**Std. error**	**t-value**	**p-value**
Intercept	1.811	0.034	53.24	0.0001
x_1	0.117	0.0439	2.66	0.044
x_2	0.187	0.0367	5.101	0.0037
x_3	0.342	0.0439	7.788	0.0005
x_1^2	0.177	0.0296	6.003	0.0018
x_2^2	0.3605	0.0296	12.18	0.00006
x_3^2	0.2512	0.0296	7.27	0.00007
x_1x_2	-0.01	0.0519	-0.19	0.855
x_2x_3	-0.215	0.0519	-4.14	0.009
x_1x_3	-0.144	0.0708	-2.3	0.097

The Table 3 imply that first order main effect of nitrogen source, salt solution concentration and their second order main effects of carbon source, nitrogen source and salt solution concentration are highly significant as is evident from their respective p-values (px_1,px_2,px_3,px_1^2,px_2^2,px_3^2) which are less than or equal to 0.05. The best model for maximizing lipase production by a response surface analysis was the following quadratic model (Eq 3.1).

Lipase activity (U/ml) = $1.811+0.117x_1+0.187x_2+0.32x_3+0.177x_1^2+0.3605x_2^2+0.215x_3^2-0.01x_1x_2-0.144x_1x_2-0.215x_2x_3$ (3.1)

The fit of the model was checked by the coefficient of determination R^2 which was calculated to be 0.98849, indicating that 98.84% of the variability in the response could be explained by the model. By optimizing the above equation, the maximum lipase activity is 2.9 U/ml and the optimum concentration of carbon source i.e. mustard oil, x_1=2.16(% v/v), optimum concentration of nitrogen source i.e. ammonium sulphate, x_2=2.2(%w/v) and optimum concentration of salt solution, x_3=48.59 (%v/v). Experiments in triplicate were carried out at the above conditions of mustard oil, ammonium sulphate and salt solution concentration and an average response of 2.87 U/ml lipase yield was observed which was very close to the predicted value and justifies the validity of response model and the existence of an optimum point.

The three dimensional response surfaces are plotted in Figs. 1-3 corresponding to the combined effects of ammonium sulphate-mustard oil, salt solution-mustard oil, salt solution-ammonium sulphate. The response obtained were convex in shape suggesting that there were well defined operating conditions

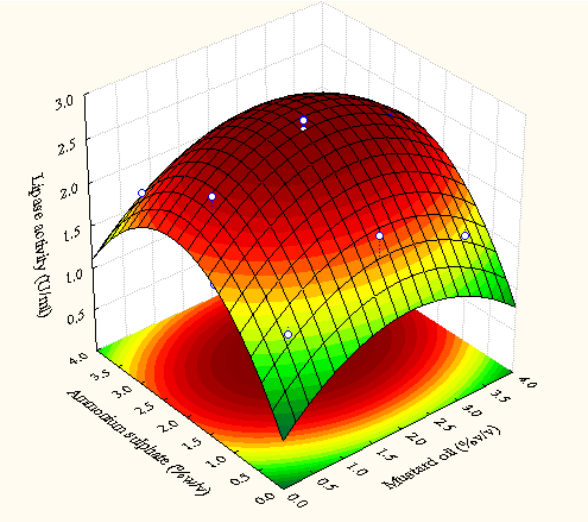

Fig 1: Response surface and contour plot of mustard oil vs. ammonium sulphate on lipase activity (salt solution was kept constant at 40% (v/v))

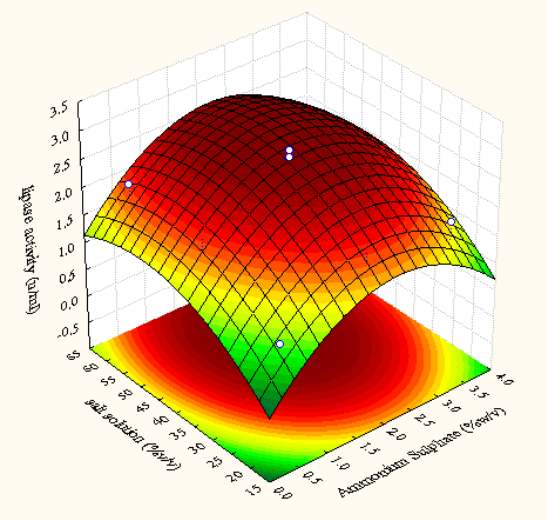

Fig 2: Response surface and contour plot of ammonium sulphate vs. salt solution on lipase activity (mustard oil was kept constant at 2% (v/v))

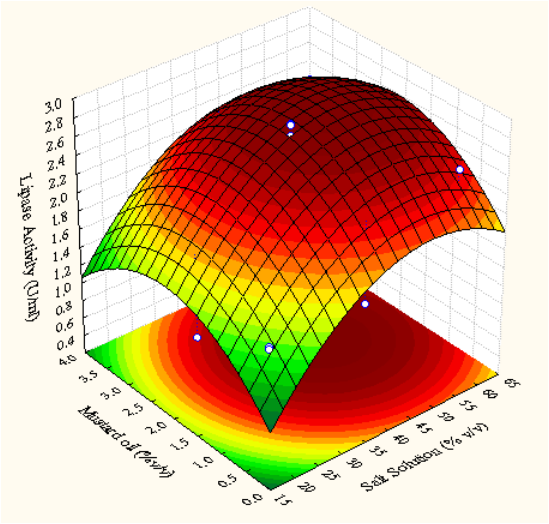

Fig 3: Response surface and contour plot of salt solution vs. mustard oil on lipase activity (ammonium sulphate was kept constant at 2% (w/v))

3.3 ANN

The data set of 15 experimental runs obtained with the Box-Behnken design was split into two categories: a training set comprising 12 experimental runs was used to optimize the weights of the networks and a testing set comprising 3 experimental runs (3,5 and 12) was used to evaluate their predictive capability. The optimum enzymatic activity of 3.12 U/ml was obtained with three neurons in the hidden layer.

The transfer functions used in the neural networks are *transig* and *purelin* at the hidden layer and outer layer respectively. Newff function is used for the training of the neural networks. The training function trainbr is used in this work. The following equation (Eq. 3.2) is the outcome of the neural network training, relating the input variables (x_1,x_2,x_3) to the output variable, y, in terms of weights and biases.

$$Y=\mathbf{w2}*(2./(1.0+\exp(-2*(\mathbf{w1}*\mathbf{pt1}+\mathbf{b1})))-\mathbf{1})+\mathbf{b2} \quad (3.2)$$

Where**w1** and **w2**are the weights, b1 and b2 are the biases, shown in Table 4. Y is the predicted value from the neural network and pt is the row vector of three independent variables, [x1, x2, x3], while pt1 represents the transpose of this vector with a dimension of (3x1).

Table 4: Weights and Biases of the Neural Networks

Training function	**weights and biases of ANN**			
trainbr	w1 matrix: (3×3)	w2 (1×3)	b1 (3×1)	b2 (1×1)
	0.34 1.31 0.32 -0.64 1.16 -0.07 0.76 0.12 0.02	1.49 -1.34 -0.80	0.76 -0.66 -0.60	0.63

The equation (3.2) represents the output, y(enzyme yield), for the given set of independent variables represented in '**pt**' when 'tansig' was used as the transfer function in the hidden layer and 'purelin' was used as the transfer function in the outer layer. The input data of the independent variable were transformed between -1 and +1 using the built in function 'premmnf' prior to neral network training while 'postmnmx' was used to transform back the optimized set of independent variables into the original scale, after the global optimization method was applied. The simulated values of the lipase yield as predicted by equation (3.2) are in close agreement with those of experiment values as evident from Table 4. The global optimization routine gbsolve was used to optimize the equation (3.2) by writing MATLAB code and the following optimum values are obtained.

The maximum lipase yield, y=3.12 U/ml, the optimum concentration of carbon source i.e. mustard oil, x_1=2.43 (% v/v), optimum concentration of nitrogen source i.e. ammonium sulpate, x_2=1.88 (%w/v), optimum concentration of salt solution, x_3=59.97(% v/v).

4. CONCLUSIONS

The present investigation deals with the production of lipase enzyme in submerged fermentation using the marine yeast *Yarrowia lipolytica*NCIM 3589. The present study reveals that the mustard oil can be used as carbon source and also as an inducer. Preliminary experiments were carried out to fix the range of the concentrations of process variables: mustard oil (carbon source), ammonium sulphate (nitrogen source) and salt solution. The respective uses of Box-

Behnken design (statistical experimental design) and neural networks were found to be effective in locating the optimum conditions within the range fixed from the preliminary runs.

15 experimental runs were carried according to the chosen experimental design and a quadratic equation was fitted to the data relatin the yield of enzyme (as the dependent variable) to the concentration of three process variables (as the independent variables). The yields predicted by the equation based on Box-Behnken design were found to be in close agreement with the experiment values. The optimum lipase yield obtained by solving the empirical quadratic equation was found to be 2.90 U/ml as compared to 2.1 U/ml, the enzyme yield obtained with the initial experiments. Thus an increase of 38% enzyme yield was obtained by using the optimization strategy of Box-Behnken design of the response surface methodology. The final experiment repeated at the optimal settings of the process variables produced a lipase yield of 2.87 U/ml which is quite closer to the predicted optimal value of the Box-Behnken design.

The data derived from 15 runs of Box-Behnken design were also fitted by the backpropogation method of the artificial neural networks and the predicted values of enzyme yield were found to be quite closer to the experimental values. When a global optimization routine was employed to optimize the equation resulted from neural networks, the optimum predicted enzyme yield was found to be 3.12 U/ml which is 53% higher than the value obtained from preliminary runs.

This work established the potential use and superiority of the artificial networks in combination with global optimization technique for optimizing the fermentation medium for enzyme production.

REFERENCES

[1] Sharma .R et al., Biotechnology Advances 19 (2001) 627-662.
[2] Yonada .T and Takada .H., Bioresource Technology 19 (1995) 245-247.
[3] Talon .R et al., Enzyme Microbial Technology 19 (1996) 620-622.
[4] Jagger K.E., and Eggert .T., Current Opinion Biotechnology 13(2002) 390-397.
[5] Hasan .F et al., Enzyme Microbial Technology 39(2006)235-251.
[6] Haaland .P.D., Stattistical problem solving. In: Haaland PD, editor. Experiment desin in biotechnology. New York: Marcel Dekker (1989).
[7] Montogomery .C.D., Design and Analysis of Experiments 5th ed., John Wiley and Sons Inc., New York (2000)
[8] Paresh .G et al., Journal of Current Microbiology ., 45 (2002) 203-208.
[9] Wen-Kai Lu et al., Journal of Applied Science and Engineering 2(1) (2004) 49-58.
[10] Newell. J, "Review of STATISTICA 6.0", available online at: www.statsoft.com/uk/welcome.html
[11] Kamimura .E.S et al., Journal of Biotech and Applied Biochemical., 33 (2001) 153-159.
[12] Nagata .Y and Chu .K.H., Biotech Letters., 25(2003) 1837-1842.
[13] Demuth .H and Hagan .M.B.M., "Neural Network Toolbox, For Use with MATLAB", The MathWorks, Inc. 3 Apple Hill Drive Natick, MA, (2006)
[14] Sivanadan .S.N et al., Introduction to Neural Networks Using MATLAB, Tata Mc.Graw Hill, New-Delhi (2006)
[15] Mattias .B and Kenneth .H. Advance Modeling and Optimization, 1(1999) 17-36.
[16] Wagner .R.W and Heitor .S.L., Biotech Letters, 43 (2000) 29-37.
[17] Liu .C.H et al., Process Biochemistry, 34 (1999) 17-23.
[18] Dutta .J.R et al., Process Biochemistry, 39 (2004) 2193-2198.
[19] Susan .L et al., Journal of Biotech., 52 (1997) 257-266.
[20] Jayati .R.D et al., Process Biochemistry., 39 (2004) 2193-2198.
[21] Rubina .N et al., Journal of Indian Mocrobial Biotechnology 39 (2012) 23-254.
[22] Sarat Babu. I at al., Brazilian Journal of Microbiology 3 (2013) 915-921.