Comparison of artificial neural network (ANN) and response surface methodology (RSM) in fermentation media...

Article in Biochemical Engineering Journal · October 2008

DOI: 10.1016/j.bej.2008.05.009

Biochemical Engineering Journal 41 (2008) 266–273

Contents lists available at ScienceDirect

Biochemical Engineering Journal

journal homepage: www.elsevier.com/locate/bej

Comparison of artificial neural network (ANN) and response surface methodology (RSM) in fermentation media optimization: Case study of fermentative production of scleroglucan

Kiran M. Desai, Shrikant A. Survase, Parag S. Saudagar, S.S. Lele, Rekha S. Singhal*

Food Engineering and Technology Department, Institute of Chemical Technology, University of Mumbai, Nathalal Parikh Marg, Matunga, Mumbai 400 019, Maharashtra, India

A R T I C L E I N F O

Article history:
Received 5 February 2007
Received in revised form 5 May 2008
Accepted 21 May 2008

Keywords:
Scleroglucan
Sclerotium rolfsii
Response surface methodology
Artificial neural network
Genetic algorithms
Sensitivity analysis

A B S T R A C T

Response surface methodology (RSM) is the most preferred method for fermentation media optimization so far. In last two decades, artificial neural network-genetic algorithm (ANN-GA) has come up as one of the most efficient method for empirical modeling and optimization, especially for non-linear systems. This paper presents the comparative studies between ANN-GA and RSM in fermentation media optimization. Fermentative production of biopolymer scleroglucan has been chosen as case study. The yield of scleroglucan was modeled and optimized as a function of four independent variables (media components) using ANN-GA and RSM. The optimized media produced 16.22 ± 0.44 g/l scleroglucan as compared to 7.8 ± 0.54 g/l with unoptimized medium.

Two methodologies were compared for their modeling, sensitivity analysis and optimization abilities. The predictive and generalization ability of both ANN and RSM were compared using separate dataset of 17 experiments from earlier published work. The average % error for ANN and RSM models were 6.5 and 20 and the CC was 0.89 and 0.99, respectively, indicating the superiority of ANN in capturing the non-linear behavior of the system. The sensitivity analysis performed by both methods has given comparative results. The prediction error in optimum yield by hybrid ANN-GA and RSM were 2% and 8%, respectively.

1. Introduction

The development of proper fermentation media is a necessary and important step in efficient utilization fermentation technology [1]. The conventional "one-factor-at-a-time" approach is laborious and time consuming. Moreover, it seldom guarantees the determination of optimal conditions [2]. These limitations of a single factor optimization process can be overcome by using empirical methods. In empirical methods two approaches are possible, *viz.* statistical-based approach and artificial intelligence-based black box approach.

In statistical-based approaches, response surface methodology (RSM) has been extensively used in fermentation media optimization [3–7]. RSM is a collection of statistical techniques for designing experiments, building models, evaluating the effects of factors and searching for the optimum conditions [8]. It is a statistically designed experimental protocol in which several factors are simultaneously varied [3]. In RSM, the experimental responses to design of experiments (DOEs) are fitted to quadratic function. The number of successful applications of RSM suggests that second-order relation can reasonably approximate many of the fermentation systems.

In last two decades, ANN has emerged as an attractive tool for non-linear multivariate modeling [9]. The power of ANN is that it is generic in structure and possesses the ability to learn from historical data. The main advantage of ANN compared to RSM are: (i) ANN does not require a prior specification of suitable fitting function and (ii) ANN has universal approximation capability, i.e. it can approximate almost all kinds of non-linear functions including quadratic functions, whereas RSM is useful only for quadratic approximations.

It is believed that ANN would require much more number of experiments (number of patterns) than RSM to build an efficient model. But in fact, ANN can also work well even with relatively less data, if the data is statistically well distributed in the input domain, which is the case with DOE. Thus experimental data of RSM should be sufficient to build effective ANN model. There are few case studies available in literature where models were developed by RSM and ANN using same DOE; and ANN models have consistently worked better than RSM [10–13]. Bas and Boyaci [13] has

* Corresponding author. Tel.: +91 22 24145616; fax: +91 22 24145614.
E-mail addresses: rekha@udct.org, rsinghal7@rediffmail.com (R.S. Singhal).

1369-703X/$ – see front matter
doi:10.1016/j.bej.2008.05.009

recently reported the comparison of ANN and RSM in enzyme kinetics, which also suggest the superiority of ANN. These reports have compared the two methods mainly from modeling prospective; whereas this paper has taken the comparison further in sensitivity analysis and also in optimization. The other perceived disadvantage of ANN compared to RSM is that RSM, because of its structured nature, is more useful in getting insight information such as sensitivity analysis and interactive effect of two components on the system. There are few reports available on the methods of carrying out sensitivity analysis using inherent structure of ANN [14–16]. Jaiswal et al. [16] has also described the method computing two-way interactions of independent variable on the system for ANN model. Though, aim of this paper is not to get into the intricacies of the sensitivity analysis by ANN model, the results of sensitivity analysis using one of the methods, namely ‘perturb method’, is described.

The input space of quadratic model of RSM can be easily optimized using conventional gradient-based methods. The ANN models being exclusively data-based, cannot be guaranteed to be smooth. Hence, the conventionally used gradient-based optimization methods, which require the objective function to be continuous, differentiable and more importantly smooth, cannot be used efficiently for optimizing the input space of an ANN model.

Genetic Algorithms (GAs) [17,18], an artificial intelligence-based stochastic non-linear optimization formalism is used to optimize the input space of ANN model. This hybrid methodology will be referred as ANN-GA hereafter. The GA mimics the principles of biological evolution namely, “survival-of-the-fittest” and “random exchange of data during propagation” followed by biologically evolving species. GA has been proved to be an ideal technique to solve diverse optimization problems in biochemical engineering [19,20].

The present work has twofold objectives, *viz.* (i) maximizing the fermentative yield of scleroglucan using empirical techniques and (ii) comparing the performances of the statistical- and artificial intelligence-based optimization techniques. Fermentative production of scleroglucan by *Sclerotium rolfsii* MTCC 2156 in submerged culture was chosen as case study. Scleroglucan is a non-ionic, water-soluble homopolysaccharide consisting of a linear chain of β-D-(1-3)-glucopyranosyl groups and β-D-(1-6)-glucopyranosyl groups [21]. In this work, ANN-GA and RSM have been used to optimize and to study the effects of the concentrations of media components namely, sucrose, yeast extract, magnesium sulphate and dipotassium hydrogen phosphate on scleroglucan production with initial pH of 4.5 ± 0.2. The optimum condition given by both approaches have been experimentally verified. The predictive models given by RSM and ANN have also been compared for their efficiencies. To the best of our knowledge, this is the first report on comparison of ANN-GA and RSM in fermentation media optimization as well as scleroglucan fermentation media optimization from *S. rolfsii* MTCC 2156 using either ANN-GA or RSM.

2. Materials and methods

2.1. Medium components

Glucose, sucrose, maltose, lactose, soluble starch, fructose, magnesium sulphate, ferrous sulphate, ammonium chloride, ammonium sulphate, yeast extract, peptone, casein digest, soybean meal, beef extract, urea were purchased from M/S Hi-Media Limited, Mumbai, India. Di-potassium hydrogen phosphate, sodium nitrate, potassium nitrate were purchased from M/S S. D. Fine chemicals Limited, Mumbai, India. Soybean oil, olive oil, sunflower oil and rice bran oil were purchased from Nature Fresh Limited, Mumbai, India.

2.2. Maintenance of culture and seed culture preparation

S. rolfsii MTCC 2156 was procured from MTCC, Chandigarh, India. The culture was grown on potato dextrose agar at 28 °C for 5 days. A 3 ml cell suspension prepared from such plates was used to inoculate 50 ml of sterile seed culture medium in 250 ml conical flasks which was incubated at 28 °C, 180 rpm for 2 days on rotary shaker.

2.3. Optimization of fermentation medium using one-factor-at-a-time method

Production medium contained (%) sucrose 2.0, sodium nitrate 0.3, yeast extract 0.1, magnesium sulphate 0.05, di-potassium hydrogen phosphate 0.13, citric acid 0.07, potassium chloride 0.05 and ferrous sulphate 0.005. pH was adjusted to 4.5 ± 0.2. Fermentation was carried out by inoculating 5 ml seed culture in 50 ml of sterile production medium and incubating at 28 ± 2 °C and 180 rpm for 72 h. The results of the one-factor-at-a-time optimization studies reported in our previous study [2] were used for deciding the input space for further media optimization.

2.4. Estimation of biomass

Fermented broths were neutralized with NaOH or HCl as required, diluted three to four fold with distilled water, heated at 80 °C for 30 min, homogenized and then centrifuged (10,000 × *g*, 30 min). The pellet so obtained was washed with distilled water and dried at 105 °C. The supernatant was used for estimation of scleroglucan production.

2.5. Estimation of scleroglucan production

Two volumes of 96% (v/v) ethanol or isopropanol were added to precipitate the scleroglucan from clear supernatant. The mixture was allowed to stand for 8 h at 4 °C for complete precipitation. Scleroglucan was recovered by filtration under vacuum and dried at 105 °C.

2.6. Sugar utilization during fermentation by S. rolfsii MTCC 2156

For this, 1 ml of broth was taken after every 12 h during the course of 72 h fermentation, centrifuged at 10,000 × *g* for 15 min at 4 °C. After removing the scleroglucan by ethanol precipitation from the cell-free broth, the supernatant was used for the estimation of residual sucrose. A suitably diluted 0.1 ml aliquot was analyzed by phenol sulphuric acid method as follows: to 0.1 ml of diluted sample, 1 ml of 5% (w/v) phenol solution and 5 ml of 95% sulphuric acid were added. The tubes were mixed by shaking after 10 min, and cooled at 25 °C for 30 min. The extinction was read at 490 nm. The standard curve was plotted using glucose in the concentration range of 10–100 μg/ml [22].

3. Predictive modeling and optimization methods

3.1. Artificial neural network

The commonly used feed forward architecture of ANN, also known as multiplayer perceptron (MLP) was used to build predictive model with concentrations of four media components as an input, and yield of scleroglucan as an output to the model. In this architecture, data always flow in a forward direction, i.e. from input layer to output layer. A real number quantity, known as weights, is associated with the connection of two neurons, which an adjustable parameter of the network. The neurons in the input layer simply introduce the scaled input data to the hidden layer *via* weights. The neurons in the hidden layer perform two tasks. First, they sum up the weighted inputs to neurons, including bias as shown by the following equation:

$$\text{sum} = \sum_{i=1}^{n} x_i w_i + \theta \tag{1}$$

where w_i ($i = 1, n$) are the connection weights, u is called bias and x_i is the input parameter.

The weighted output is then passed through an activation function. The activation function shifts the space in non-linearity of input data. The logistic output function is used in this work, shown by the following equation:

$$f(\text{sum}) = \frac{1}{1 + \exp(-\text{sum})} \tag{2}$$

The output thus produced by hidden layer becomes an input to output layer. The neurons in the output layer produce the output using same procedure as that of neurons in the hidden layer. An error function based upon this calculated output and actual experimental output is formulated. Training an ANN is an iterative process where this pre-specified error function is minimized by adjusting the weights appropriately. The commonly employed error function the *root-mean-squared-error* (RMSE) used in this work is defined as

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{N} \sum_{n=1}^{M} (y_n^i - \hat{y}_n^i)^2}{NM}} \tag{3}$$

where N refers to the number of patterns used in the training; M denotes the number output nodes; i denotes the index of the input pattern (vector) and y_n^i and $\hat{y}_n^i$ are the desired (target) and predicted outputs of the nth output node, respectively. The RMSE is minimized using the *error-back-propagation* (EBP) algorithm [23], which uses the gradient-descent technique based on the *generalized delta rule* (GDR). The EBP training algorithm makes use of two adjustable parameters namely, the learning rate (ε) ($0 < \varepsilon \leq 1$), and momentum coefficient (γ) ($0 < \gamma \leq 1$). The magnitudes of both these parameters are optimized heuristically along with the number of hidden layer neurons. The details of training an optimal MLP model possessing good prediction and generalization abilities are described for instance in [9,19,24,25].

3.2. Genetic algorithm

Once a generalized ANN model has been developed, its input space is optimized using genetic algorithm. The input vector comprising of input variables of model becomes the decision variable for the GA. GA treats an optimization through a simple cycle of four stages which consists of initialization of solution populations known as chromosomes, fitness computation based on objective function, selection of best chromosomes, genetic propagation of selected parent chromosomes using genetic operators like crossover and mutation to create the new population of chromosomes. The whole process continues until a suitable result is achieved. The best string that evolves after repeating the above-described loop till convergence forms the solution to the optimization problem.

3.3. Response surface methodology

RSM is an empirical statistical modeling technique employed for multiple regression analysis using quantitative data obtained from properly designed experiments to solve multivariate equations simultaneously [3]. RSM is used to determine the optimum nutrient concentrations, for the production of scleroglucan. A central composite rotatable experimental design (CCRD) for four independent variables was used. The medium components (independent variables) selected for the optimization were sucrose, yeast extract, di-potassium hydrogen phosphate and magnesium sulphate. Regression analysis was performed on the data obtained from the experiments.

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

$$x_i = \frac{X_i - X_{cp}}{\Delta X_i} \quad i = 1, 2, 3, \ldots, k \tag{4}$$

where x_i, dimensionless value of an independent variable; X_i, real value of an independent variable; X_{cp}, real value of an independent variable at the center point; and ΔX_i, step change of real value of the variable i corresponding to a variation of a unit for the dimensionless value of the variable i.

The experiments were carried out at in duplicate, which was necessary to estimate the variability of measurements. Replicates at the center of the domain in three blocks permit the checking of the absence of bias between several sets of experiments. The relationship of the independent variables and the response was calculated by the second-order polynomial Eq. (5).

$$Y = \beta_0 + \sum_{i=1}^{k} \beta_i X_i + \sum_{i=1}^{k} \beta_{ii} X_i X_i + \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} \beta_{ij} X_i X_j \tag{5}$$

Y is the predicted response; β_0 a constant; β_i the linear coefficient; β_{ii} the squared coefficient; and β_{ij} the cross-product coefficient, k is the number of factors.

The second-order polynomial coefficients were calculated using the software package Design Expert Version 6.0.10 to estimate the responses of the dependent variable. Response surface plots were also obtained using Design Expert Version 6.0.10.

4. Results and discussion

4.1. Hybrid ANN-GA optimization

4.1.1. Predictive modeling with ANN

The design of experiments, which is used for training the network and respective experimental yields are given in Table 1. The coded values of independent variables are given in Table 2. ANN-based process model was developed using the most popular feed-forward ANN architecture namely, multi-layer perceptron (MLP) with logistic sigmoidal function. The MLP network has four input nodes representing components concentrations and one output node representing the scleroglucan yield (g/l) at the end of a batch. The data partitioning as training set and test had been done to avoid over-training and overparameterization. The training cycle were performed for varying numbers of neurons in the hidden layer and also for various combinations of ANN-specific parameter like learning rate, random initialization. The generalization capacity of the model was ensured by selecting the weights resulting in the least test set RMSE. The MLP with three nodes in hidden layer resulted in the least value for the test set RMSE, i.e. ($E_{tst} = 0.327$). The corresponding RMSE was ($E_{trn} = 0.102$).

The average error (%) between the experimental and model predicted scleroglucan concentrations for the training and test set data were 1.124 and 2.365, respectively; the values of correlation coefficient (CC) between the model predicted and experimental scleroglucan yield pertaining to the training set (CC_{trn}) and the test

Table 1
Central composite rotatable design (CCRD) matrix of independent variables and their corresponding experimental and predicted yields of scleroglucan

No.	Media concentration (g/l)				Scleroglucan (g/l)		
	Sucrose	Yeast extract	K_2HPO_4	$MgSO_4$	Experimental[a]	RSM predicted	ANN predicted
1	35	1.0	1.25	0.5	7.97	8.39	7.97
2	65	1.0	1.25	0.5	15.22	14.29	15.21
3	35	2.0	1.25	0.5	9.53	8.91	9.62
4	65	2.0	1.25	0.5	12.60	12.38	12.53
5	35	1.0	1.75	0.5	9.44	8.86	9.43
6	65	1.0	1.75	0.5	13.16	13.57	13.25
7	35	2.0	1.75	0.5	9.81	9.42	9.78
8	65	2.0	1.75	0.5	11.65	11.70	11.62
9	35	1.0	1.25	1.0	8.47	8.56	8.47
10	65	1.0	1.25	1.0	14.05	14.38	13.74
11	35	2.0	1.25	1.0	9.50	9.04	9.50
12	65	2.0	1.25	1.0	11.70	12.43	12.51
13	35	1.0	1.75	1.0	7.09	7.25	7.89
14	65	1.0	1.75	1.0	11.10	11.87	11.17
15	35	2.0	1.75	1.0	6.69	7.77	6.67
16	65	2.0	1.75	1.0	10.44	9.97	11.24
17	20	1.5	1.5	0.75	6.29	6.49	6.30
18	80	1.5	1.5	0.75	14.88	14.59	14.88
19	50	0.5	1.5	0.75	12.25	11.96	12.23
20	50	2.5	1.5	0.75	10.38	10.58	10.38
21	50	1.5	1.0	0.75	10.65	11.03	10.62
22	50	1.5	2.0	0.75	9.50	9.03	9.49
23	50	1.5	1.5	0.25	10.16	11.14	11.03
24	50	1.5	1.5	1.25	10.65	9.58	10.62
25	50	1.5	1.5	0.75	11.13	11.12	11.17
26	50	1.5	1.5	0.75	11.50	11.12	11.17
27	50	1.5	1.5	0.75	11.10	11.12	11.17
28	50	1.5	1.5	0.75	10.99	11.12	11.17
29	50	1.5	1.5	0.75	10.98	11.12	11.17
30	50	1.5	1.5	0.75	11.03	11.12	11.17

[a] Average of two readings.

set (CC_{tst}) were 0.992 and 0.98, respectively. The small and comparable magnitudes of the RMSE and average prediction error (%), and the high and comparable values of CC, for both the training and test set outputs suggest that the MLP-based model possesses good approximation and generalization characteristics.

4.1.2. GA-based optimization

The GA-based technique was to optimize the input space of ANN model with objective of maximization of scleroglucan yield. The values of GA-specific parameters used in the optimization simulations were chromosome length = 40, population size = 50, crossover probability = 0.9, mutation probability = 0.05, and number of generations over which GA evolved = 750.

The objective function can be defined as follows:

$$\text{Maximize} \quad y = f(\mathbf{x}, \mathbf{W}); x_i^L \le x_i \le x_i^U, \quad l = 1, 2, \ldots, P \tag{6}$$

where f represents the objective function (ANN model); $\mathbf{x}$ denotes the input vector; $\mathbf{W}$ denotes corresponding weight vectors; y refers to the scleroglucan experimental yield; the input vector, x, denotes the fermentation operating conditions; P denotes number of input variables; and x_l^L and x_l^U represent the lower and upper bounds on x_i, i.e. alpha values for each variable, respectively. The fitness of each chromosome (candidate solution) was evaluated based on following fitness function:

$$\text{error}_j = 1 - \frac{1}{\hat{y}^j_{\text{pred}}}; \quad j = 1, 2, \ldots, N \tag{7}$$

where error_j denotes the fitness value of the candidate solution and $\hat{y}^j_{\text{pred}}$ denotes the MLP model predicted scleroglucan yield for given candidate solution. During GA-implementation, the search for the optimal solutions was restricted between the bounds specified in RSM design (see Table 2).

The GA-based optimization procedure was repeated several times for different randomly initialized population of the candidate solutions (chromosomes) and also for different GA specific parameters parameters. These repetitions at varying initial conditions ensured that entire search space was search rigorously to find global optimum. It was also observed that for most of varied initial conditions GA converged to similar solution, suggesting it to be the global solution. The optimum solution was found heuristically. The ANN predicted yield of scleroglucan at GA optimized condition was 16.19 g/l. This result was verified by carrying out the fermentation at GA-specified optimum conditions. The scleroglucan yield obtained in the verification experiment was 16.42 ± 0.68 (g/l), which is in close agreement with the hybrid ANN-GA solution.

4.2. Response surface methodology

To examine the combined effect of four different medium components (independent variables), on scleroglucan production, a central composite factorial design of $2^4 = 16$ plus 6 *centre points* and $(2 \times 4 = 8)$ *star points* leading to a total of 30 experiments were performed. Second-order polynomial equation was used to correlate the independent process variables, X_i, with scleroglucan production. The second order polynomial coefficient for each

Table 2
Coded values of independent variables

Independent variables	Coded values				
	−2	−1	0	+1	+2
Sucrose	20	35	50	65	80
Yeast extract	0.5	1.0	1.5	2.0	2.5
K_2HPO_4	1.0	1.25	1.5	1.75	2.0
$MgSO_4$	0.25	0.50	0.75	1.0	1.25

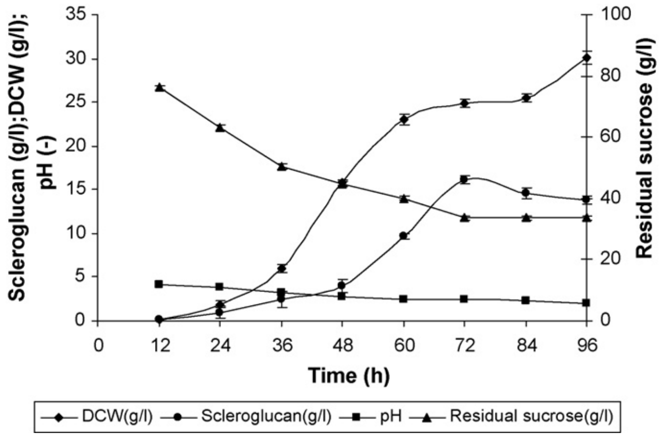

Fig. 1. Production profile of scleroglucan by *Sclerotium rolfsii* MTCC 2156 on the media optimized by RSM.

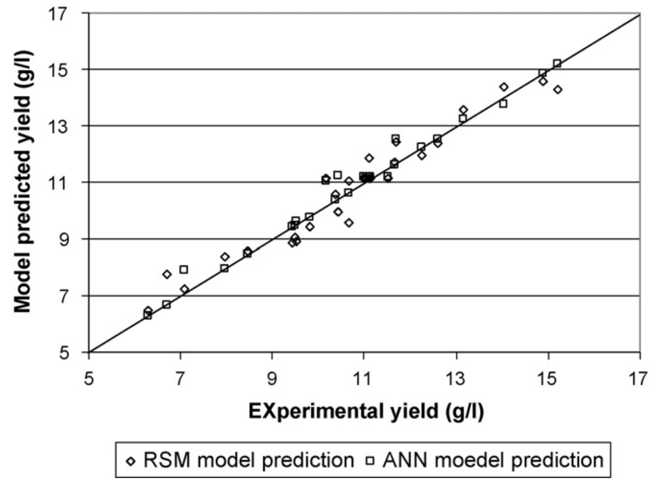

Fig. 2. RSM and ANN predicted vs. experimental yields for scleroglucan production by *S. rolfsii* MTCC 2156.

term of the equation determined through multiple regression analysis using the Design Expert. The same DoE, which used in ANN-based model development was also used to build RSM model.

The results were analyzed by using ANOVA, i.e. analysis of variance suitable for the experimental design. The results are shown in Table 3. The Model *F*-value of 16.88 implies that the model is significant. Model *F*-value is calculated as ratio of mean square regression and mean square residual. Model *P*-value (Prob > *F*) is very low (0.0001). This resignifies the significance of the model.

The *P*-values were used as a tool to check the significance of each of the coefficients, which, in turn, are necessary to understand the pattern of the mutual interactions between the test variables. The *t* ratio and the corresponding *P* values, along with the coefficient estimate, are given in Table 3. The smaller the magnitude of the *P*, the more significant is the corresponding coefficient. Values of *P* less than 0.05 indicate model terms are significant. The coefficient estimates and the corresponding *P* values suggests that, among the test variables used in the study, X_1 (sucrose), X_2 (yeast extract), X_3 (K_2HPO_4), X_4 ($MgSO_4$), $X_1 \times X_2$ (sucrose × yeast extract) and $X_3 \times X_4$ ($K_2HPO_4 \times MgSO_4$) are significant model terms. Sucrose ($P < 0.0001$) has the largest effect on scleroglucan production, followed by K_2HPO_4 ($P < 0.0043$), $MgSO_4$ ($P < 0.0189$) and yeast extract ($P < 0.0336$). The mutual interaction between sucrose and yeast extract ($P < 0.0045$) and K_2HPO_4 and $MgSO_4$ ($P < 0.026$) were also found to be important. Other interactions were found to be insignificant.

The corresponding second-order response model (see Eq. (5)) that was found after analysis for the regression was

$$\begin{aligned} \mathrm{Yield}_{(g/l)} = {} & 11.12 + 2.03X_1 - 0.35X_2 - 0.50X_3 - 0.39X_4 - 0.15X_1^2 \\ & + 0.037X_2^2 - 0.27X_3^2 - 0.19X_4^2 - 0.61(X_1 \times X_2) \\ & - 0.30(X_1 \times X_3) - 0.021(X_1 \times X_4) + 0.011(X_2 \times X_3) \\ & - 0.011(X_2 \times X_4) - 0.45(X_3 \times X_4). \end{aligned} \tag{8}$$

The fit of the model was also expressed by the coefficient of determination R^2, which was found to be 0.88, indicating that 88.0% of the variability in the response could be explained by the model.

The optimal concentrations for the four components as obtained from the maximum point of the model were calculated to be as 80, 1.01, 1.06 and 1.15 g/l for sucrose, yeast extract, K_2HPO_4 and magnesium sulphate, respectively. By substituting levels of the factors into the regression equation, the maximum predictable response for scleroglucan production was calculated and was experimentally verified. The maximum production of scleroglucan obtained experimentally using the optimized medium was 16.22 g/l, which is in correlation with the predicted value of 17.32 g/l by the RSM

Table 3
Analysis of variance (ANOVA) for the experimental results of the central-composite design (quadratic model)

Factor[a]	Coefficients	Sum of squares	Standard error	d.f.[b]	*F*-value	*t* ratio	*P*[c]
Intercept or model	11.12	124.64	0.3	14	16.88	37.06	<0.0001*
X_1	2.03	98.42	0.15	1	186.61	13.53	<0.0001*
X_2	−0.35	2.88	0.15	1	5.47	−2.33	0.0336*
X_3	−0.50	5.96	0.15	1	11.3	−3.33	0.0043*
X_4	−0.39	3.65	0.15	1	6.92	−2.60	0.0189*
X_1^2	−0.15	0.58	0.14	1	1.1	−1.07	0.3103†
X_2^2	0.037	0.037	0.14	1	0.071	0.26	0.7939†
X_3^2	−0.27	2.05	0.14	1	3.88	−1.92	0.0676†
X_4^2	−0.19	1.0	0.14	1	1.89	−1.35	0.1894†
$X_1 \times X_2$	−0.61	5.88	0.18	1	11.15	−3.38	0.0045*
$X_1 \times X_3$	−0.30	1.43	0.18	1	2.71	−1.66	0.1206†
$X_1 \times X_4$	−0.021	0.0072	0.18	1	0.014	−0.116	0.9084†
$X_2 \times X_3$	−0.011	0.0020	0.18	1	0.00380	−0.061	0.9514†
$X_2 \times X_4$	0.011	0.0020	0.18	1	0.00384	0.061	0.9514†
$X_3 \times X_4$	−0.45	3.19	0.18	1	6.04	−2.80	0.0266*

[a] X_1 = sucrose; X_2 = yeast extract; $X_3 = K_2HPO_4$; $X_4 = MgSO_4$.
[b] Degree of freedom.
[c] †: not significant; *$P < 0.05$, $R^2 = 0.88$.

Table 4
RSM and ANN predictions for totally unseen data

No.	Sucrose	YE	K_2HPO_4	$MgSO_4$	Experimental[a]	RSM	ANN
1	20	0.5	1	0.025	3.29	2.91	3.05
2	20	1	1.3	0.05	6.01	4.88	5.47
3	20	1.5	1.6	0.075	5.23	6.01	5.96
4	20	2	1.9	0.1	6.31	6.29	5.89
5	40	0.5	1.3	0.075	9.01	9.98	10.16
6	40	1	1	0.1	7.12	11.74	7.33
7	40	1.5	1.9	0.025	10.25	11.53	9.17
8	40	2	1.6	0.05	7.50	9.95	10.18
10	60	0.5	1.6	0.1	12.49	13.18	12.49
11	60	1	1.9	0.075	12.59	12.92	12.03
12	60	1.5	1	0.05	9.00	14.17	10.85
13	60	2	1.3	0.025	13.30	11.75	12.60
14	80	0.5	1.9	0.05	13.20	18.41	14.04
15	80	1	1.6	0.025	16.58	16.87	15.62
16	80	1.5	1.3	0.1	13.80	16.01	13.01
17	80	2	1	0.075	14.30	16.89	14.13

[a] The fermentation results used in this table are reported in Survase [2].

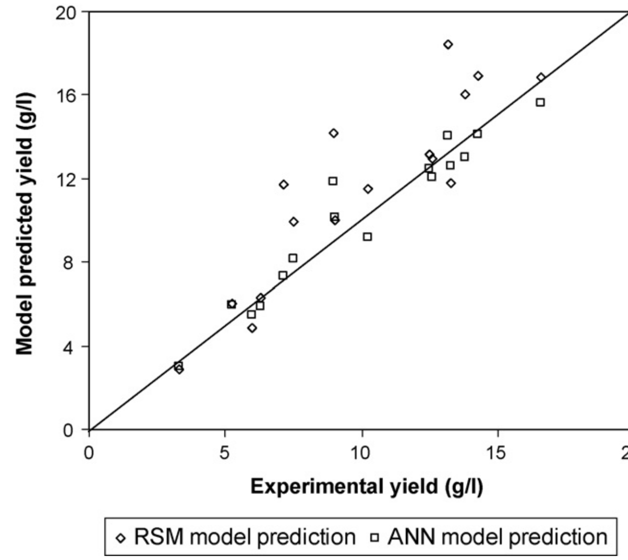

Fig. 3. Comparison of generalization ability RSM and ANN model for unseen dataset. Dataset is taken from Survase [2].

regression study. The result of the batch carried out at optimum concentration is shown in Fig. 1.

4.3. Comparison of RSM and hybrid ANN-GA

4.3.1. Predictive capabilities

The ANN and RSM model were compared for DoE, using which the both models were trained. The comparison was made on the basis of various parameters such as average % error, RMSE and CC. The predicted values by ANN as well RSM model are tabulated in Table 1. Fig. 2 shows the comparative parity plot for ANN and RSM predictions for DoE. The MLP-based model had fitted the experimental data with an excellent accuracy. The RSM-based prediction shows greater deviation than ANN. The comparative values average % error, RMSE and CC were given in Table 5.

The generalization ability can be best judged only with totally unseen dataset. Thus, it was decided to test both the models using completely unseen data of the same fermentation system reported in Survase [2]. The experimental and predicted yields are summarized in Table 4. The comparative values average % error, RMSE and CC were given in Table 5. The correlation coefficient for unseen data by RSM and ANN are 0.89 and 0.98; and average percentage error is 20 and 6.5. Fig. 3 shows the comparative parity plot for ANN and RSM predictions for DoE. Thus, ANN has shown significant higher generalization capacity than RSM. This higher predictive accuracy of ANN can be attributed to its universal ability to approximate non-linearity of the system whereas RSM only restricted to second-order polynomial.

4.3.2. Sensitivity analysis

The effect of the individual components and interactions of the components on the system can be studied in more obvious way in RSM than ANN. Since the variables in the quadratic equation of RSM are in the normalized form, the coefficients of the equation gives direct measure of the contribution of the various components in the system. As shown in Table 3, sucrose (X_1) has the largest coefficient (2.03), which indicates that sucrose is by far the most dominating factor. In the interactions terms, the coefficient of sucrose–yeast extract interaction ($X_1 \times X_2$) and K_2HPO_4–$MgSO_4$ interaction ($X_3 \times X_4$) have higher coefficients (i.e. 0.61 and 0.45, respectively), indicating these two interactions have significant effect on the system compared to other interactions. The same observations are further quantified using ANOVA earlier. ANN being a *black box model*, it does not give such insights of the system directly. But there are numerous methods available which gives the sensitivity analysis of the system using inherent nature of ANN. Fig. 4 shows the ANN sensitivity analysis of the system using 'perturb method'. Each series in the graph represents the rate of change of response with change in the given input variable. Higher the slope and range of change in the response, greater the influence of the variable. It could be observed that glucose has the highest influence on the system whereas all other variables have much lesser and almost equal influence. The slopes of the linear fitting of each variable are 2.06 (glucose), 0.51 (yeast extract), −0.32 (K_2HPO_4) and −0.18 (M_gSO_4), which are interestingly quite comparable to the coefficient of first-order terms in the quadratic RSM

Table 5
Comparison of predictive capacity of RSM and ANN

Parameters	Design data[a]		Validation data[b]	
	RSM	ANN	RSM	ANN
Correlation coefficient	0.93	0.99	0.89	0.98
Average % error	4.61	1.67	20	6.5
RMSE	0.31	0.11	2.50	0.73

[a] Central composite rotatable design (CCRD) which is used for training both RSM and ANN model.
[b] Unseen dataset (Survase [2]), i.e. not used for modeling.

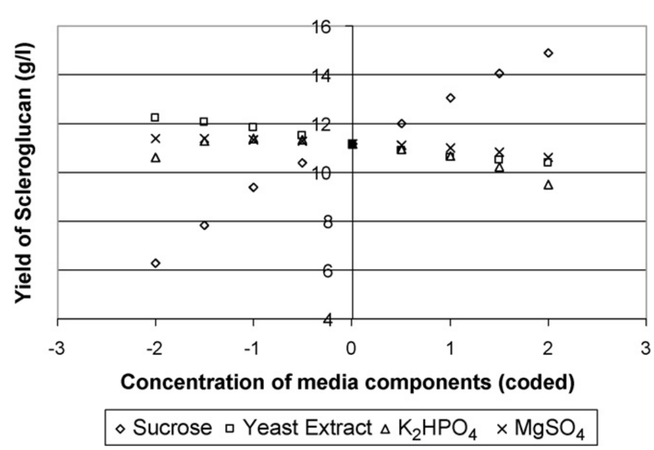

Fig. 4. Sensitivity analysis of fermentation system using ANN model.

Table 6
Optimized medium composition for scleroglucan production by *Sclerotium rolfsii* MTCC 2156 using different methodologies

	Component concentration (g/l)				Scleroglucan (g/l)	
	Sucrose	Yeast extract	K_2HPO_4	$MgSO_4$	Predicted	Experimental
Before optimization	20.0	1.00	1.30	0.50	–	7.8 ± 0.54
At center point of DOE	50.0	1.50	1.50	0.75	–	11.23 ± 1.21
RSM	80.0	1.01	1.06	1.15	17.32 ± 0.00	16.22 ± 0.44
Hybrid ANN-GA	80.0	0.81	1.20	0.50	16.19 ± 0.00	16.42 ± 0.68

equation. Thus, ANN is also equally efficient in sensitivity analysis.

4.3.3. Optimization

The comparison of yields of scleroglucan for optimized media using different technique is given in Table 6. The unoptimized yield was 7.8 ± 0.54 g/l. The optimized media concentrations obtained by RSM and hybrid ANN-GA are almost similar except for the concentrations of $MgSO_4$. RSM has predicted scleroglucan yield of 17.32 g/l at optimized condition. The experimental verification has given yield of 16.22 ± 0.44 g/l. Similarly, predicted and experimental yield for hybrid ANN-GA were 16.19 and 16.42 ± 0.68 g/l, respectively. The prediction error in optimum yield by hybrid ANN-GA and RSM were 2% and 8%, respectively. The maximum experimental scleroglucan yield obtained in this case is almost same for ANN-GA as well as RSM optimized inputs. But the point that authors want to emphasize here is that RSM has over predicted the yield. This difference between predicted and experimental yield can be contributed to the extent deviation in predictive capacity of model. Since ANN is more accurate and more generalized model than quadratic RSM, it is better equipped to reach the global optimum. Thus even though, in this particular case, maximum experimental yield obtained by ANN-GA and RSM are not statistically different, ANN has predicted the optimum condition and yield more accurately than RSM.

RSM is most widely used method in fermentation media optimization. It is one of the efficient methods for non-linear optimization. But its main limitation of RSM is that is assumes only quadratic non-linear correlation. So if we want to use RSM effectively, we need to narrow down search window appropriately (if we shrink the search window narrow enough, linear correlation may also suffice). This makes the search process highly dependent upon search space. It will require either extra experiments or good priory knowledge of the system to fix search window. Since ANN can inherently capture almost any form of non-linearity, it can easily overcome above discussed limitation of RSM. Thus, in case of ANN, more liberal search space can be chosen; even if, the correlation in that search space is more complex than quadratic.

5. Conclusion

In the present work, ANN and RSM methodologies are compared for their predictive and generalization capabilities, sensitivity analysis and optimization efficiency in fermentation media optimization. ANN showed better accuracy and generalization capability than RSM even with limited number of experiments. The prediction accuracy of ANN was almost three times better than RSM. Because of its structured nature, RSM is useful in getting insight information (e.g. interactions between different components) of the system directly. But ANN was also found to be equally useful in the sensitivity analysis (further investigation in regards are required). ANN has also shown higher accuracy in finding optimum condition and predicting optimum yield. Thus, ANN has consistently performed better than RSM in all the aspects. Using artificial intelligence-based methods; the yield of scleroglucan was significantly increased (from 7.8 ± 0.54 g/l in unoptimized media to 16.42 ± 0.68 g/l) with minimum number of experiments. Thus, it can be concluded that even though RSM is most widely used method for fermentation media optimization, ANN-GA methodology may present a better alternative.

References

[1] A. Margaritis, G.W. Pace, Microbial polysaccharides, in: Moo-Young, C.W. Robinson (Eds.), Advances in Biotechnology, vol. 2, 1985, pp. 1005–1044.
[2] S.A. Survase, P.S. Saudagar, R.S. Singhal, Production of scleroglucan from *Sclerotium rolfsii* MTCC 2156, Bioresour. Technol. 97 (2006) 989–993.
[3] J.K. Rao, K. Chul-Ho, S.K. Rhee, Statistical optimization of medium for the production of recombinant hirudin from *Saccharomyces cerevisiae* using response surface methodology, Process Biochem. 35 (2000) 639–647.
[4] P. Rama Mohan Reddy, B. Ramesh, S. Mrudula, G. Reddy, G. Seenayya, Production of thermostable b-amylase by *Clostridium thermosulfurogenes* SV2 in solid-state fermentation: optimization of nutrient levels using response surface methodology, Process Biochem. 39 (2003) 267–277.
[5] W. Wei, Z. Zheng, Y. Liu, Z. Zhu, Optimizing the culture conditions for higher inulinase production *Kluyveromyces* sp. Y-85 and scaling-up fermentation, J. Fermentation Bioeng. 26 (4) (1998) 395–399.
[6] J.R. Dutta, P.K. Dutta, R. Banerjee, Optimization of culture parameters for extracellular protease production from a newly isolated *Pseudomonas* sp. using response surface and artificial neural network models, Process Biochem. 39 (2004) 2193–2198.
[7] Y.-H. Xiong, J.-Z. Liu, H.-Y. Song, L.-N. Ji, Enhanced production of extracellular ribonuclease from *Aspergillus niger* by optimization of culture conditions using response surface methodology, Biochem. Eng. J. 21 (2004) 27–32.
[8] S.J. Kalil, F. Maugeri, M.I. Rodrigues, Response surface analysis and simulation as a tool for bioprocess design and optimization, Process Biochem. 35 (2000) 539–550.
[9] K.M. Desai, B.K. Vaidya, R.S. Singhal, S.S. Bhagwat, Use of an artificial neural network in modeling yeast biomass and yield of b-glucan, Process Biochem. 39 (2004) 2193–2198.
[10] W. Lou, S. Nakai, Application of artificial neural networks for predicting the thermal inactivation of bacteria: a combined effect of temperature, pH and water activity, Food Res. Int. 34 (2001) 573–591.
[11] J. Bourquin, H. Schmidli, P.V. Hoogevest, H. Leuenberger, Advantages of artificial neural networks (ANNs) as alternative modeling technique for data sets showing non-linear relationships using data from a galenical study on a solid dosage form, Eur. J. Pharm. Sci. 7 (1998) 5–16.
[12] S. Agatonovic-Kustrin, M. Zecevic, L.J. Zivanovic, I.G. Tucker, Application of artificial neural networks in HPLC method development, J. Pharm. Biomed. Anal. 17 (1998) 69–76.
[13] D. Bas, I. Boyac, Modeling and optimisation. II. Comparison of estimation capabilities of response surface methodology with artificial neural networks in a biochemical reaction, J. Food Eng. 78 (2007) 846–854.
[14] M. Gevrey, I. Dimopoulos, S. Lek, Review and comparison of methods to study the contribution of variables in artificial neural network models, Ecol. Model. 160 (2003) 249–264.
[15] S. Jaiswal, E.R. Benson, J.C. Bernard, G.L. Van Wicklen, Neural network modelling and sensitivity analysis of a mechanical poultry catching system, Biosyst. Eng. 92 (1) (2005) 59–68.
[16] S. Jaiswal, E.R. Benson, J.C. Bernard, G.L. Van Wicklen, Two-way interaction of input variables in the sensitivity analysis of neural network, Ecol. Model. 195 (2006) 43–50.
[17] L. Davis, Handbook of Genetic Algorithms, Van Nostrand Reinhold, New York, 1991.
[18] D.E. Goldberg, Genetic Algorithms in Search, Optimization, and Machine Learning, Addison-Wesley, New York, 1989.
[19] D. Sarkar, J.M. Modak, Optimization of fed-batch bioreactor using genetic algorithm, Chem. Eng. Sci. 58 (2003) 283–2296.
[20] S. Nandi, P. Mukharjee, S.S. Tambe, R. Kumar, B.D. Kulkarni, Reaction modeling and optimization using neural networks and genetic algorithms: case study involving TS-1 catalyzed hydroxylation of benzene, Ind. Eng. Chem. Res. 41 (2002) 2159–2169.
[21] J.I. Farina, F. Sineriz, O.E. Molina, N.I. Perotti, High scleroglucan production by *Sclerotium rolfsii*: influence of media composition, Biotechnol. Lett. 20 (1998) 825–831.

[22] M. Dubois, K.A. Gilles, J.K. Hamilton, P.A. Robers, F. Smith, Colorimetric methods for determination of sugars and related substances, Anal. Chem. 28 (1956) 350–356.
[23] D. Rumelhart, G. Hinton, R. Williams, Learning representations by backpropagating errors, Nature 323 (1986) 533–534.
[24] S.S. Tambe, B.D. Kulkarni, P.B. Deshpande, Elements of Artificial Neural Network with Selective Applications in Chemical and Biological Sciences, Simulation and Advance Control, Inc., 1996.
[25] J.A. Freeman, D.M. Skapura, Neural Networks: Algorithms, Applications, and Programming Techniques, Addison-Wesley, Reading, MA, 1991.