CrossMark

User-friendly optimization approach of fed-batch fermentation conditions for the production of iturin A using artificial neural networks and support vector machine

Fudi Chen [a], Hao Li [b], Zhihan Xu [c], Shixia Hou [c], Dazuo Yang [a,d,*]

[a] *Key Laboratory of Marine Bio-Resources Restoration and Habitat Reparation in Liaoning Province, Dalian Ocean University, Dalian, China*
[b] *College of Chemistry, Sichuan University, Chengdu, Sichuan, China*
[c] *College of Light Industry, Textile and Food Engineering, Sichuan University, Chengdu, Sichuan, China*
[d] *College of Life Science and Technology, Dalian University of Technology, Dalian, China*

A R T I C L E   I N F O

A B S T R A C T

Background: In the field of microbial fermentation technology, how to optimize the fermentation conditions is of great crucial for practical applications. Here, we use artificial neural networks (ANNs) and support vector machine (SVM) to offer a series of effective optimization methods for the production of iturin A. The concentration levels of asparagine (Asn), glutamic acid (Glu) and proline (Pro) (mg/L) were set as independent variables, while the iturin A titer (U/mL) was set as dependent variable. General regression neural network (GRNN), multilayer feed-forward neural networks (MLFNs) and the SVM were developed. Comparisons were made among different ANNs and the SVM.
Results: The GRNN has the lowest RMS error (457.88) and the shortest training time (1 s), with a steady fluctuation during repeated experiments, whereas the MLFNs have comparatively higher RMS errors and longer training times, which have a significant fluctuation with the change of nodes. In terms of the SVM, it also has a relatively low RMS error (466.13), with a short training time (1 s).
Conclusion: According to the modeling results, the GRNN is considered as the most suitable ANN model for the design of the fed-batch fermentation conditions for the production of iturin A because of its high robustness and precision, and the SVM is also considered as a very suitable alternative model. Under the tolerance of 30%, the prediction accuracies of the GRNN and SVM are both 100% respectively in repeated experiments.

1. Introduction

Produced by *Bacillus subtilis*, the nonribosomal lipopeptide antifungal antibiotic iturin A is structurally composed of two parts. The first part consists of seven amino acid residues (L-Asn–D-Tyr–D-Asn–L-Gln–L-Pro–D-Asn–L-Ser) which are formed into a peptide circle. The second part is a hydrophobic tail with 11–12 carbons [1,2,3]. In terms of treating both human and animal mycoses, iturin A has been showed to be a potential bio-resource due to its wide-scale-spectrum antifungal activity [4,5]. According to recent research, the iturin A can also be applied as a controlling agent to fight against plant pathogens causing a decrease in crop production, such as southern corn leaf blight [6].

During the past decades, researchers have paid much attention to the practical production of iturin A due to its foreseeable potential in biological fields. In order to increase the yield of iturin A, the optimization method is commonly adopted in creating better fermentation conditions. For decades, the optimization of fermentation has been studied in many ways [7,8]. In a laboratory environment, the majority of the methods to optimize the fermentation process are largely based on data obtained from a large amount of experimental works, which cannot be used in practical applications. Additionally, the statistic-based methods such as the orthogonal experiment method and response surface methodology (RSM) [9] cost more manpower and resources than expected. In order to gain statistics that are suitable for practical production, researchers brought up the uniform design (UD) method. So far, the UD method has been successfully applied in many optimization processes [6,10,11]. Compared with the traditional statistical methods, the UD can enormously save manpower and resources in the lab by reducing the number of essential experiments in different dimensions and allows as many different levels of factors as it can [6].

With the development of artificial intelligence (AI), artificial neural networks (ANNs) have been widely applied in predictive modeling. With a comparatively higher accuracy in modeling and better ability in generalization, ANNs are able to simulate the bio-process and predict the results [12,13,14,15]. Compared with

* Corresponding author.
E-mail address: dzyang1979@outlook.com (D. Yang).

the traditional statistical methods, ANNs can also model all non-linear multivariate functions, while the traditional statistical methods can only model the quadratic functions [16,17,18]. Also, it is reported that the ANNs are more accurate than the RSMs in many cases [19,20]. Normally, UDs have relatively-representative and regularly-distributed patterns. Based on these patterns with high quality, ANNs are also able to establish equally accurate models with a comparatively smaller amount of data than it is supposed to require obtain.

Despite the advantages of ANNs modeling, few studies have reported using ANNs to reduce the number of experiments. An ANN model was established based on UD data was conducted by Peng and colleagues [6]. In their research, the ANN model based on UD data was adopted in the optimization of iturin A yield and a comparison of the ANN-GA methods and the UD methods was conducted for the first time. Adopted widely during variable chemical process [21,22], this method can be effectively used for applications. However, as a technician, one may find it difficult to use this method to practical applications because of its complexity. People may feel confused using related approaches. Here, an alternative series of user-friendly ANNs and a support vector machine (SVM) are proposed to seek a better optimization method in order to increase the yield of iturin A based on the data from Peng's research [6]. We aim at creating more alternative methods to improve the simplification of the fed-batch fermentation conditions for the production of iturin A, so that the maneuverability of the practical applications can be improved using novel modeling methods.

2. Materials and methods

2.1. Fed-batch fermentation of iturin A

According to Peng and colleagues' research [6], the separated *B. subtilis* ZK8 strain was used for the production of iturin A. The seed culture-medium contained 2.86 g/L KH_2PO_4, 3 g/L $MgSO_4$, 25 g/L glucose and 30 g/L peptone. The slant culture-medium contained 1.5 g/L K_2HPO_4, 1.8 g/L agar, 1.8 g/L $MgSO_4 \cdot 7H_2O$, 20 g/L peptone and 10 mL/L glycerol. The fermentation culture-medium was prepared with 0.79 g/L KH_2PO_4, 0.8 g/L yeast extract, 2.4 g/L soybean protein powder hydrolysate, 3.8 g/L $MgSO_4$ and 31 g/L glucose. Strain ZK8 was activated in the slant culture-medium. The activated strain was then inoculated and incubated in the seed culture medium in a shaker at 30°C with 150 rpm for 20 h. Then, the seed culture was inoculated in fermentation culture by 10% amount of inoculum for 48 h at 30°C with 150 rpm. After 24 h of fermentation, the asparagine (Asn), glutamic acid (Glu) and proline (Pro) were added to the broth in different concentrations [6]. The yield of iturin A was determined by titer measurement and the cylinder-plate method was used to measure the titer of iturin A [6,23,24,25]. According to the experimental results [6], statistical results were obtained (Table 1).

2.2. ANNs

ANNs [26,27,28] are powerful machine learning techniques with the functions of estimation and approximation based on the

Table 1
Statistical experimental results of the amino acid concentration (mg/L) and iturin A titer (U/mL) (data extracted from Peng's research [6]).

Statistical item	Factor (mg/L)			Iturin A titer (U/mL)
	Asn	Glu	Pro	
MIN	50	200	50	10,108
MAX	200	400	200	13,064.1
AVERAGE	119.6	293.52	119.6	12,033.2

inputs. Interconnected artificial neural networks usually consist of neurons that can calculate values from inputs and adapt to different situations. Therefore, ANNs are capable of numeric prediction and pattern recognizing. Recent years, ANNs have gained wide popularity in inferring a function from observation especially when the data or the task is too complicated to be dealt with human brains. In our studies, multilayer feed-forward neural networks (MLFNs) and general regression neural network (GRNN) were used for developing alternative models for optimizing the fed-batch fermentation conditions of iturin A.

2.2.1. MLFNs

MLFNs trained with a back-propagation learning algorithm, are the most popular neural networks [29,30,31]. They are applied to a wide variety of chemistry related problems [29].

An MLFN model consists of neurons that are ordered into layers (Fig. 1). The first layer is called the input layer, the last layer is called the output layer, and the layers between are hidden layers. For the formal description of the neurons we can use the so-called mapping function Γ, that assigns for each neuron i a subset $\Gamma(i) \subseteq V$ which consists of all ancestors of the given neuron. A subset $\Gamma(i)^{-1} \subseteq V$ consists of all predecessors of the given neuron i. Each neuron in a particular layer is connected with all neurons in the next layer. The connection between the ith and jth neuron is characterized by the weight coefficient ω_{ij}, and the ith neuron by the threshold coefficient ϑ_i (Fig. 2). The weight coefficient reflects the importance degree of the given connection in the neural network. The output value of the ith neuron x_i is determined by [Equation 1 and Equation 2]. It holds that:

$$x_i = f(\xi_i) \qquad \text{[Equation 1]}$$

$$\xi_i = \vartheta_i + \sum_{j \in r_i^{-1}} \omega_{ij} x_j \qquad \text{[Equation 2]}$$

where ζ_i is the potential of the ith neuron, and function $f(\zeta_i)$ is the so-called transfer function (the summation in [Equation 2] is carried out over all neurons j transferring the signal to the ith neuron). The threshold coefficient can be understood as a weight coefficient of the connection with formally added neuron j, where $x_j = 1$ (so-called bias).

For the transfer function, it holds that

$$f(\zeta) = \frac{1}{1 + \exp(-\zeta)}. \qquad \text{[Equation 3]}$$

The supervised adaptation process varies the threshold coefficient ϑ_i and weight coefficient ω_{ij} to minimize the sum of the squared
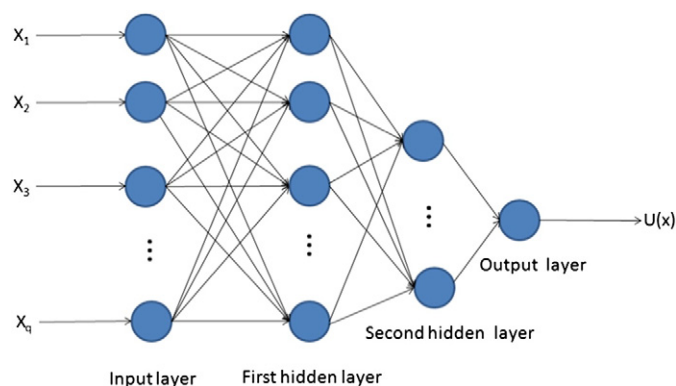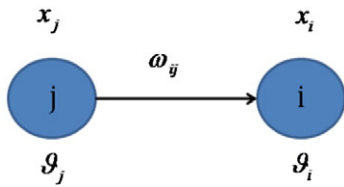

Fig. 1. Structure of the MLFN.

Fig. 2. Connection between the two neurons *i* and *j*.

differences between the computed and required output values. This is accomplished by minimization of the objective function *E*:

$$E = \sum_0 \frac{1}{2}(x_0 - \hat{x}_0)^2 \qquad \text{[Equation 4]}$$

where x_0 and $\hat{x}_0$ are the vectors composed of the computed and required activities of the output neurons and summation runs over all output neurons.

2.2.2. GRNN

GRNN was firstly created by Specht [32]. It has a strong prediction capacity in approximation, prediction, medical diagnosis, chemical engineering, pattern recognition, and 3D modeling [33,34,35,36,37,38]. In terms of approximation, GRNN usually performed better than other neural networks in practical applications [38]. The features of the GRNN are fast learning, consistency, and optimal regression with large number of samples [39]. A GRNN has four layers: input, pattern, summation, and output, which are shown in Fig. 3 [39].

Input layer keeps corresponding input automatically and transfers input vector *x* to pattern layer. Pattern layer consists of neurons for training datums. In this layer, the weighted squared Euclidean distance can be calculated by [Equation 5]. Test inputs applied to the network are first subtracted from values of pattern layer neurons. And either squares or absolute values of subtracts applied to exponential activation function will be summed. Results are transferred to the summation layer. Dot product of pattern layer outputs and weights is added by neurons of summation layer. In Fig. 3, weights are shown by A and B, y values of training data stored at pattern layer determine their values, and $f(x)K$ denotes weighted outputs of pattern layer where *K* is a Parzen window associated constant. $Yf(x)K$ denotes multiplication of pattern layer outputs and training data output *Y*. At output layer, $f(x)K$ divides $Yf(x)K$ to estimate the desired *Y*, given in [Equation 6 and Equation 7] [32,38]:

$$D_j = (x - x_j)^T (x - x_j), \qquad \text{[Equation 5]}$$
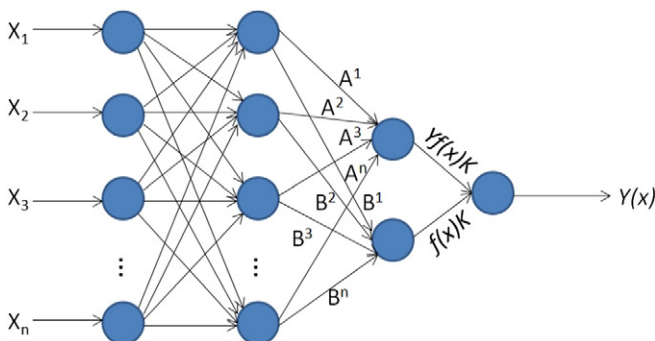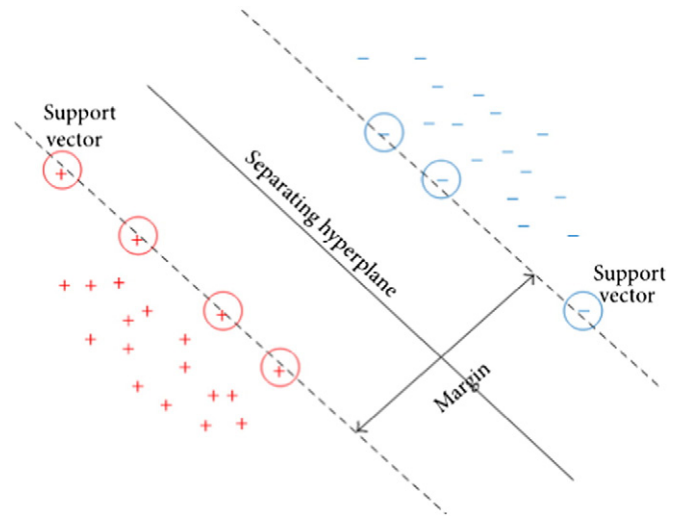

Fig. 4. Support vectors determine the position of the optimal hyperplane.

$$Y(x) = \frac{\int_{-\infty}^{\infty} Yf(x, Y)dY}{\int_{-\infty}^{\infty} Yf(x, Y)dY}, \qquad \text{[Equation 6]}$$

$$Y(x) = \frac{\sum_{j=1}^{p} y_j e[-D_j/2\sigma^2]}{\sum_{j}^{p} e[-D_j/2\sigma^2]}. \qquad \text{[Equation 7]}$$

2.3. SVM model

SVM is a learning algorithm mainly based on statistical learning theory [40]. On the basis of the limited information of samples between the complexity and learning ability of models, this theory has an excellent capability of global optimization to improve generalization. In regard to linear separable binary classification, finding the optimal hyperplane, a plane that separates all samples with the maximum margin, is an essential principle of SVM [41,42]. Not only does the plane help improve the predictive ability of the model, but also it helps reduce the error which occurs occasionally in classifying. Fig. 4 illustrates the optimal hyperplane, with "+" indicating the samples of type 1 and "-" representing the samples of type -1.

Fig. 5 shows the main structure of SVM. The letter "*K*" stands for kernels [43]. As we can see from the figure, it is a small subset
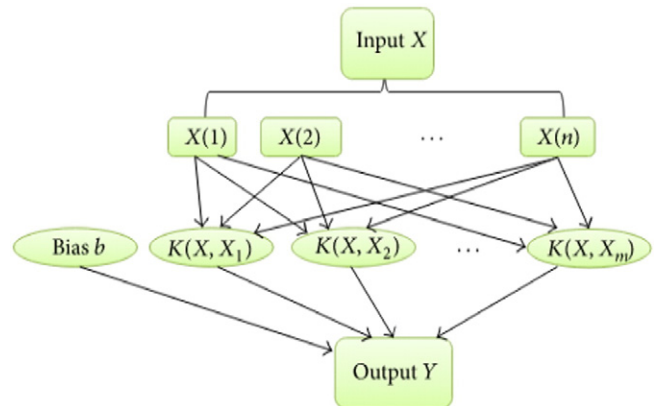

Fig. 3. Structure of the GRNN.

Fig. 5. Main structure of support vector machine.

Table 2
Best model search in different machine learning models.

Model type	Mean RMS error	Training time	Forecast accuracy
GRNN	457.88	0:00:01	100%
SVM	460.13	0:00:01	100%
MLFN (2 nodes)	760.86	0:00:23	88.89%
MLFN (3 nodes)	526.38	0:00:26	88.89%
MLFN (4 nodes)	848.09	0:00:39	77.78%
MLFN (5 nodes)	1410.73	0:00:50	55.56%
MLFN (6 nodes)	583.13	0:01:07	88.89%
MLFN (7 nodes)	878.71	0:01:23	77.78%
MLFN (8 nodes)	866.83	0:01:51	77.78%
MLFN (9 nodes)	1380.12	0:02:13	55.56%
MLFN (10 nodes)	972.60	0:02:42	66.67%
…	…	…	…
MLFN (25 Nodes)	3032.92	0:02:02	0.00%

extracted from the training data by relevant algorithm that consists of the SVM. For classification, choosing suitable kernels and appropriate parameters is of great importance to get prediction accuracy. However, a mature international standard currently for us to choose these parameters is nonexistent. In most circumstances, the comparison of experimental results, the experiences from copious calculating, and the use of cross validation that is available in software package are helping us to solve that problem to some extent [44,45].

3. Results and discussion

3.1. Model development

According to previous research, the production of iturin A yields by adding various concentrations of Asn, Glu and Pro during the fed-batch fermentation process [6]. Here, we aim at using novel ANNs and SVM to fit the concentration levels of the added component of Asn, Glu and Pro, from which we can use for the prediction of the iturin A titer.

The concentration levels of Asn, Glu and Pro (mg/L) were set as independent variables, while the iturin A titer (U/mL) was set as dependent variable. Since numeric predictions of machine learning techniques are completely based on existing data, the data should be divided into two sets before model developments, the training and testing sets. Training set help programs "learn" the regulation of data while testing set is used for validating the trained model after a training process. Here, 65% data group was set as training set, while

35% data group was set as testing set. The ANN prediction models were constructed by the NeuralTools® software (trial version, Palisade Corporation, NY, USA) [46,47,48]. We chose the GRNN and MLFN as the training algorithms. The SVM was developed with Matlab software.

We used root mean square (RMS) error and training time as the indicators to measure the performances of the ANNs and SVM (Table 2). The number of nodes of MLFNs were set from 2 to 25, from which we tried to find out the change regulation of the MLFNs when dealing with the development process.

Table 2 indicates that the GRNN, SVM and MLFNs with 3 and 6 nodes have comparatively low mean RMS errors (477.88, 460.13, 526.38 and 583.13 respectively). It is clear that the GRNN and SVM have the lowest RMS errors and the shortest training times, while the MLFNs have comparatively higher RMS errors and longer training times. To determine the accuracy of predictions, the forecast accuracy was used as an indicator. In current applications, the empirical tolerance of ANNs is 30%, which means that a single prediction result can be considered as "good prediction" when the relative error is lower than 30% of the actual value. Here, the forecast accuracy is the percentage of the tested sample of "good prediction" in the total testing set. Table 2 shows that the forecast accuracy (under the tolerance of 30%) of the GRNN and SVM are both 100%. Here, we discuss the availability of the GRNN, SVM and MLFNs respectively in order to determine the most suitable model for the design of the fed-batch fermentation conditions for the production of iturin A.

3.2. Comparison between the GRNN and MLFNs

As for the GRNN, it has the lowest RMS error and the shortest training time during our research, compared with other 24 MLFNs. And according to the robustness of the principles of the GRNN [32,38], it has a high reproducibility, which has an overwhelming advantage compared to other ANNs during our research. In order to test the robustness of the GRNN, computational experiments for the GRNN were repeated, which are shown in Fig. 6.

Fig. 6 shows the RMS errors of the GRNN in repeated experiments. It is significant that there is a stable fluctuation during the experiments, which shows that the GRNN for the optimization process is robust. More importantly, the mean RMS error is relatively low, which ensures the availability of the GRNN. Under the tolerance of 30%, the prediction accuracy of the GRNN is 100% in all repeated experiments.
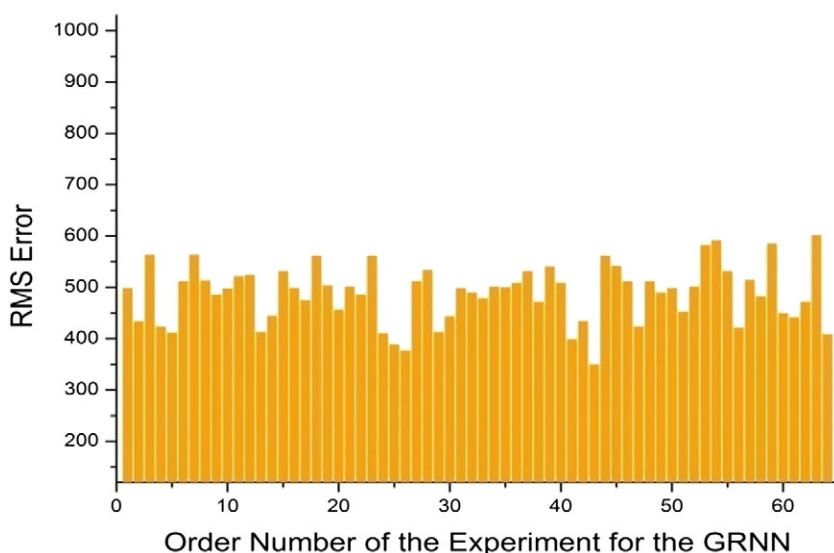
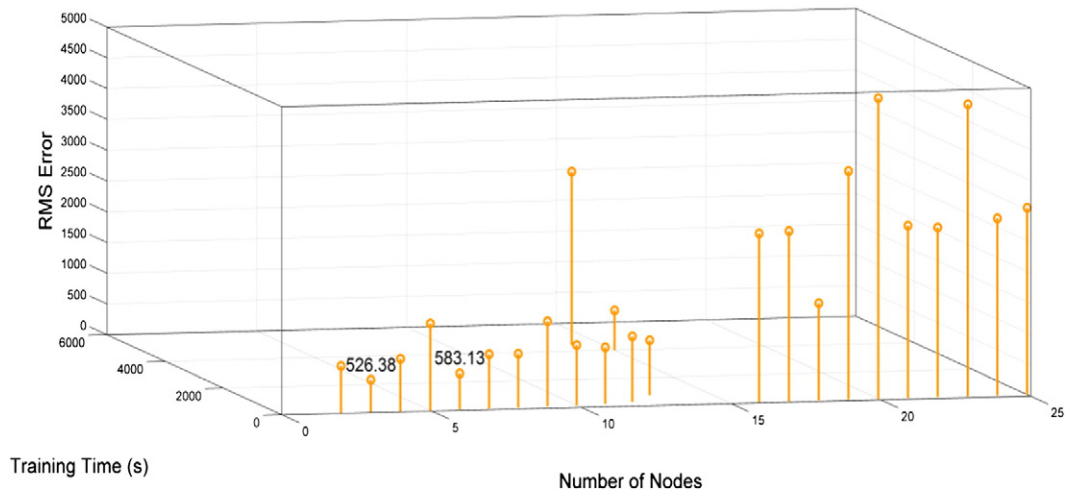**Fig. 6.** Results of repeated computational experiments for the GRNN.

Fig. 7. RMS errors and training times of MLFNs with the change of nodes.

To illustrate the change regulation of different MLFNs, Fig. 7 is used for showing the RMS errors and training times of MLFNs with different nodes.

It can be seen that with the increase of nodes, the RMS errors and training times of MLFNs become unsteadily fluctuant, which highly corresponds to the fluctuation character of the MLFN principle. It should be mentioned that results in different MLFNs presented by Table 2 are not the fixed results, because of the effects of the different random initial values chosen by the computer when training. However, it is still clear that MLFNs may have good results (relatively low RMS errors and short training times) with a relatively low number of nodes.
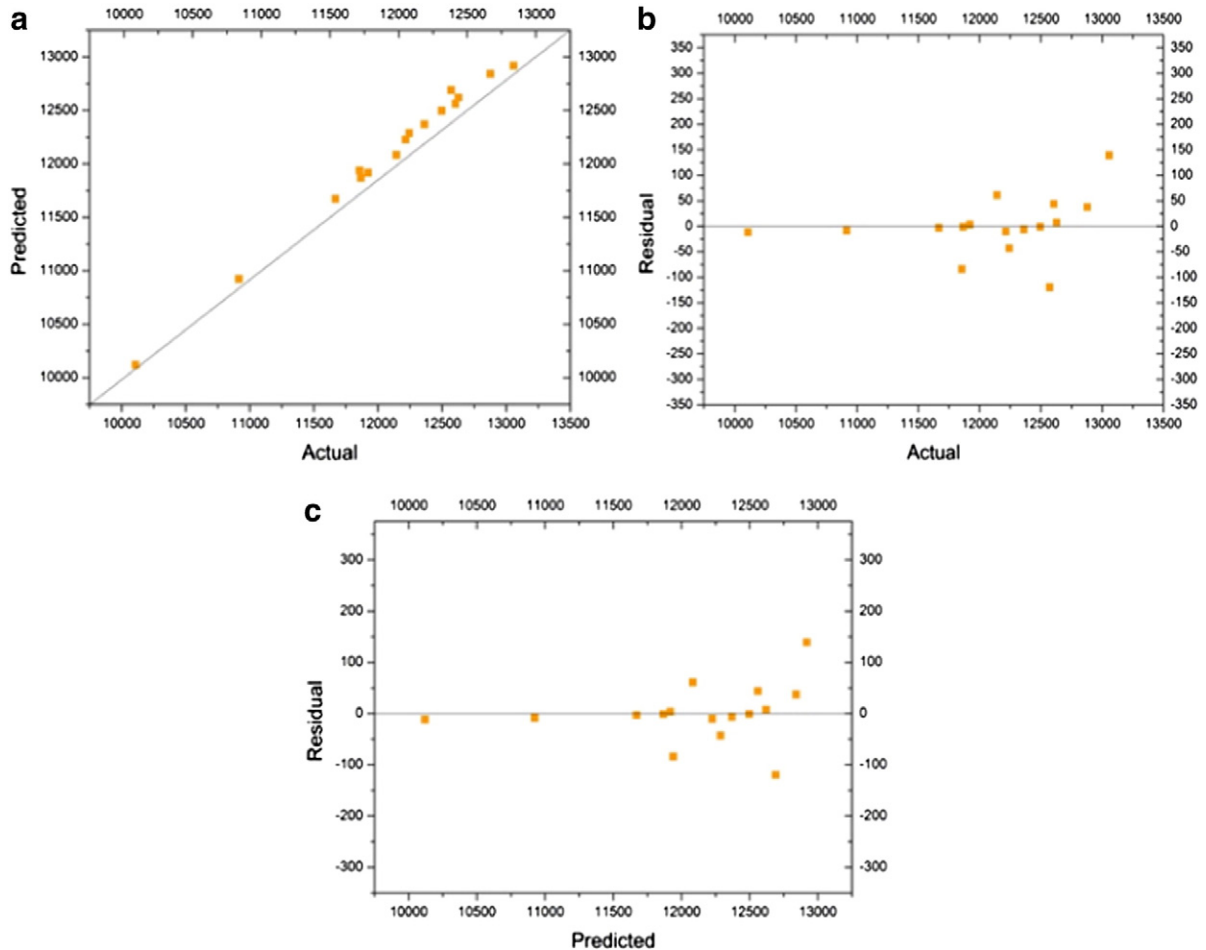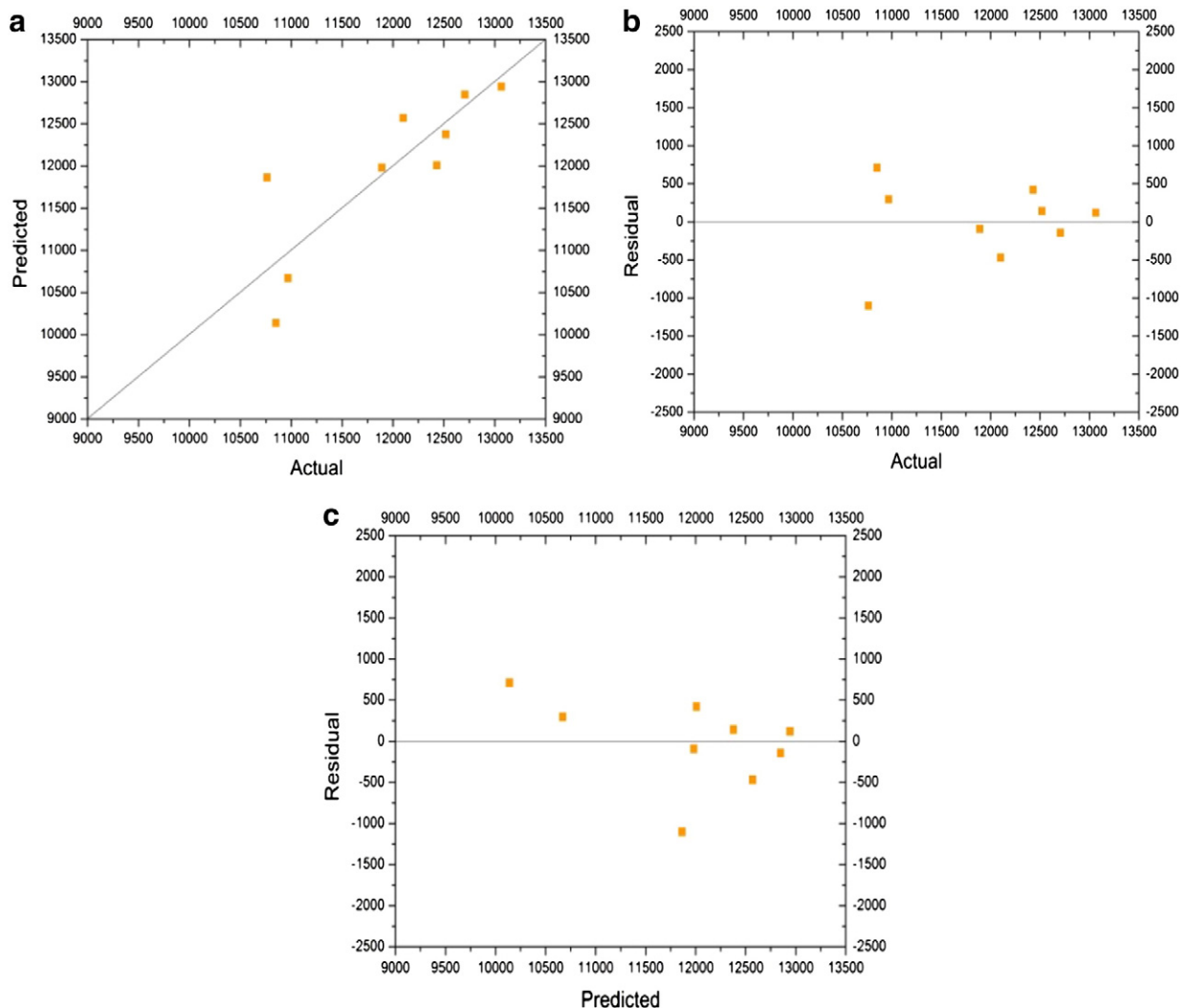

Fig. 8. Training results of the GRNN. a) Predicted values versus actual values; b) residual values versus actual values; c) residual values versus predicted values.

Fig. 9. Testing results of the GRNN. a) Predicted values versus actual values; b) residual values versus actual values; c) residual values versus predicted values.

For practical applications, one should use related software to find out the most suitable model for the optimization of the fed-batch fermentation condition in the range of low number of nodes. Compared to the GRNN, MLFNs cost longer time and the fluctuations are not as stable as what the GRNN presents. Therefore, we still consider that the GRNN is a more suitable model for the optimization of the fed-batch fermentation conditions.

3.3. Training and testing results of the GRNN and SVM

Here, we use one of the typical examples of the training and testing results to present the availability of the GRNN and SVM respectively. Fig. 8 and Fig. 9 are used to illustrate the training and testing results of the GRNN, while Fig. 10 is used to illustrate the testing results of the SVM. The training and testing sets of the GRNN and SVM are the same.

The capacity for recall of the GRNN for the optimization of the design is illustrated in Fig. 8, showing the training results of the GRNN. It shows that the GRNN has a strong capacity for recall. The predicted values are highly close to the actual values (Fig. 8a), which indicates that the non-linear fitting effects of the model is highly decent. The comparisons between the residual values and actual/predicted values (Fig. 8b and Fig. 8c) also show that the

residual values are relatively low, which suggests the robustness of the development of the GRNN.

To show the availability of the GRNN after a training process, we use the data set which has not been used for the training process. Results are shown in Fig. 9.

Fig. 9 shows the precise predicted results during the testing process. Predicted values are close to the actual values (Fig. 9a). Residual values presented by Fig. 9b and Fig. 9c show that the residual values are relatively low. Results present the robustness and availability of the GRNN model when testing.

In terms of the testing results of the SVM, Fig. 10 illustrates the correctness and robustness of the SVM in the prediction section.

Being similar to the results of the GRNN in the aspects of the RMS error and the training time, the testing results of the SVM are also highly similar to those of the GRNN. We can see that the SVM can generate a fairly analogical and precise result, compared to the testing results of the GRNN.

In sum, the GRNN and SVM are both available for the optimization of fed-batch fermentation conditions for the production of iturin A. Both the GRNN and SVM have the lowest RMS errors and the shortest training times. Compared to Peng's research [6], the GRNN and SVM are more convenient because of the user-friendly packed software [46,47,48,49]. Technicians can use the models and approaches provided by this article in practical applications without complex operation works.
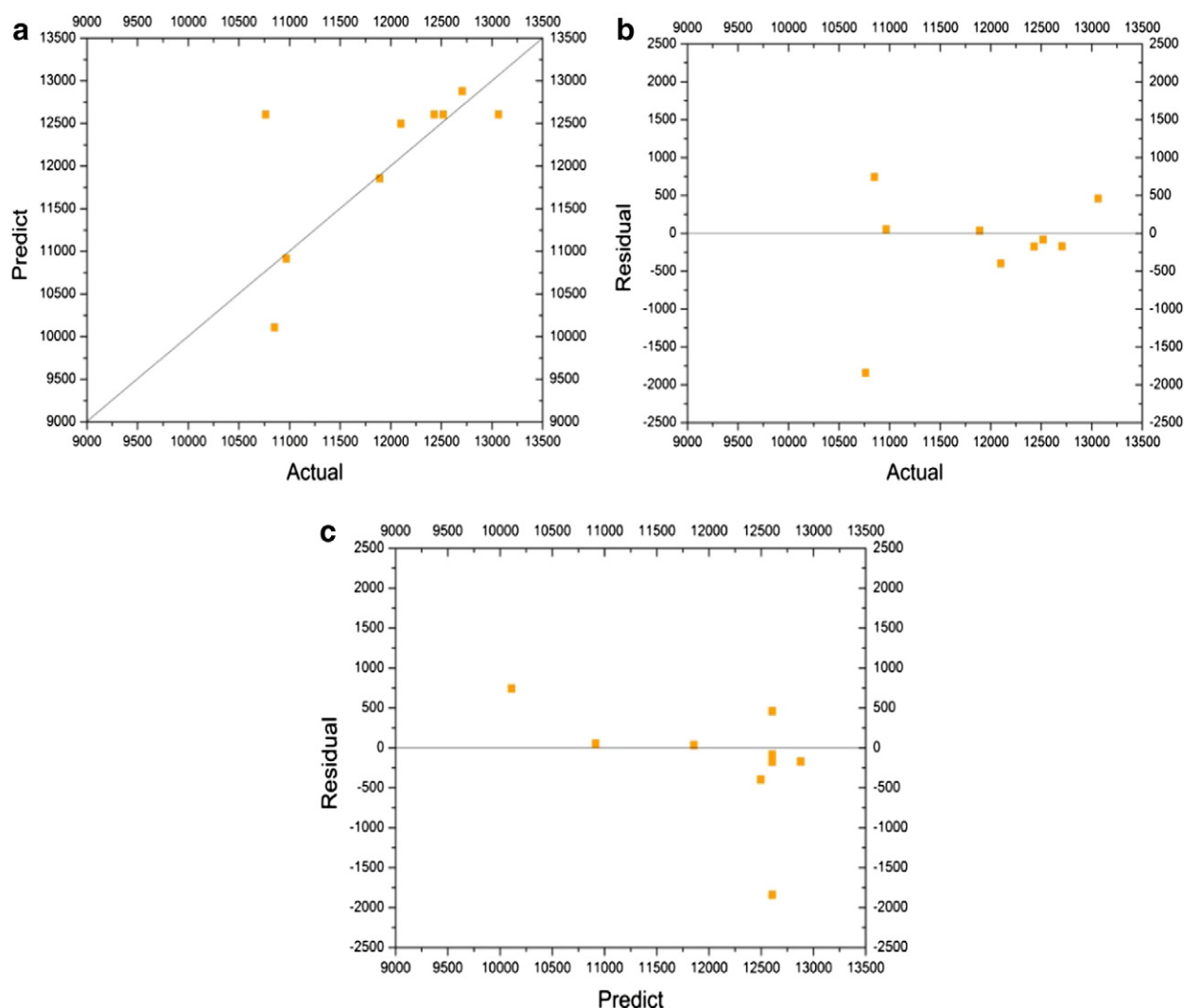
Fig. 10. Testing results of the SVM. a) Predicted values versus actual values; b) residual values versus actual values; c) residual values versus predicted values.

3.4. Comparisons with other optimization methodologies

Related different optimization methodologies for biotechnology were presented in previous reports [6,9,19,20], including regression analysis, orthogonal experiment method and the RSM. Though these models have their own advantages (*e.g.* they do not have high requirements to computers), they also have many disadvantages compared to machine learning techniques like ANNs and SVM. Generally, the overwhelming advantages of ANNs and SVM in optimization process of biotechnological production are precision, robustness and time-saving. ANNs and SVM make predictions strongly based on the "well-trained" training data set and their programs can run automatically without too much human intervention. The non-linear function of ANNs can even develop a powerful non-linear prediction system, which ensures the precision of predictions [48,49]. The principle of the SVM can strongly ensure robust results [40]. With the development of computers and programming tools, ANNs and SVM now can be easily established, which are more time-saving and user-friendly.

4. Conclusion

According to the modeling results, the GRNN is considered as the most suitable ANN model due to its highly robustness and precision.

The SVM is also considered as a suitable alternative model due to its robust and precise testing results. Under the tolerance of 30%, the prediction accuracies of the GRNN and SVM are both 100% in repeated experiments. Results indicate that the GRNN and SVM are strong alternative and operable models for the optimization for the fermentation conditions of iturin A. Being compared to the MLFNs and other models provided by previous studies, the GRNN and SVM have overwhelming advantages including low RMS error, time-saving and user-friendliness. According to the characteristic of machine learning models, over-fitting can be avoided with a large scale of training data because it can get rid of the local over-fitting phenomenon [49,50]. Therefore, with a larger scale of samples, the prediction results may be improved. We can rationally assume that in further practical applications, a larger amount of data obtained from mass production in industry can ensure higher availability and robustness of a model for optimizing the fermentation conditions of iturin A.

References

[1] Isogai A, Takayama S, Murakoshi S, Suzuki A. Structure of β-amino acids in antibiotics iturin A. Tetrahedron Lett 1982;23:3065–8. http://dx.doi.org/10.1016/S0040-4039(00)87534-6.

[2] Besson F, Peypoux F, Michel G, Delcambe L. Characterization of iturin A in antibiotics from various strains of *Bacillus subtilis*. J Antibiot 1976;29:1043–9. http://dx.doi.org/10.7164/antibiotics.29.1043.

[3] Delcambe L, Peypoux F, Besson F, Guinand M, Michel G. Structure of iturin and iturin-like substances. Biochem Soc Trans 1977;5:1122–4 [http://europepmc.org/abstract/med/913800.].

[4] Moyne A-L, Shelby R, Cleveland TE, Tuzun S, Bacillomycin D. An iturin with antifungal activity against *Aspergillus flavus*. J Appl Microbiol 2001;90:622–9. http://dx.doi.org/10.1046/j.1365-2672.2001.01290.x.

[5] Phae CG, Shoda M. Investigation of optimal conditions for foam separation of iturin, an antifungal peptide produced by *Bacillus subtilis*. J Ferment Bioeng 1991;71: 118–21. http://dx.doi.org/10.1016/0922-338X(91)90235-9.

[6] Peng W, Zhong J, Yang J, Ren Y, Xu T, Xiao S, et al. The artificial neural network approach based on uniform design to optimize the fed-batch fermentation condition: Application to the production of iturin A. Microb Cell Fact 2014;13:54. http://dx.doi.org/10.1186/1475-2859-13-54.

[7] Zhang XY, Zhou JY, Fu W, Li ZD, Zhong J, Yang J, et al. Response surface methodology used for statistical optimization of jiean-peptide production by *Bacillus subtilis*. Electron J Biotechnol 2010;13:5. http://dx.doi.org/10.2225/vol13-issue4-fulltext-5.

[8] Iwase N, Rahman MS, Ano T. Production of iturin A homologues under different culture conditions. J Environ Sci 2009;21:28–32. http://dx.doi.org/10.1016/S1001-0742(09)60031-0.

[9] Zhou WW, He YL, Niu TG, Zhong JJ. Optimization of fermentation conditions for production of anti-TMV extracellular ribonuclease by *Bacillus cereus* using response surface methodology. Bioprocess Biosyst Eng 2010;33:657–63. http://dx.doi.org/10.1007/s00449-009-0330-0.

[10] Liang YZ, Fang KT, Xu QS. Uniform design and its applications in chemistry and chemical engineering. Chemom Intell Lab Syst 2001;58:43–57. http://dx.doi.org/10.1016/S0169-7439(01)00139-3.

[11] Liu D, Wang P, Li F, Li J. Application of uniform design in L-isoleucine fermentation. Chin J Biotechnol 1991;7:207–12.

[12] Massimo CD, Willis MJ, Montague GA, Tham MT, Morris AJ. Bioprocess model building using artificial neural networks. Bioproc Eng 1991;7: 77–82. http://dx.doi.org/10.1007/BF00383582.

[13] Simutis R, Lübbert A. Exploratory analysis of bioprocesses using artificial neural network-based methods. Biotechnol Prog 1997;13:479–87. http://dx.doi.org/10.1021/bp9700364.

[14] Vlassides S, Ferrier JG, Block DE. Using historical data for bioprocess optimization: Modeling wine characteristics using artificial neural networks and archived process information. Biotechnol Bioeng 2001;73:55–68. http://dx.doi.org/10.1002/1097-0290(20010405)73:1%3C55::AID-BIT1036%3E3.0.CO;2-5.

[15] Schubert J, Simutis R, Dors M, Havlik I, Lübbert A. Bioprocess optimization and control: Application of hybrid modelling. J Biotechnol 1994;35:51–68. http://dx.doi.org/10.1016/0168-1656(94)90189-9.

[16] Walczak B, Massart DL. The radial basis functions-partial least squares approach as a flexible non-linear regression technique. Anal Chim Acta 1996;331:177–85. http://dx.doi.org/10.1016/0003-2670(96)00202-4.

[17] Gemperline PJ, Long JR, Gregoriou VG. Nonlinear multivariate calibration using principal components regression and artificial neural networks. Anal Chem 1991; 63:2313–23. http://dx.doi.org/10.1021/ac00020a022.

[18] Kramer MA. Nonlinear principal component analysis using autoassociative neural networks. AIChE J 1991;37:233–43. http://dx.doi.org/10.1002/aic.690370209.

[19] Desai KM, Survase SA, Saudagar PS, Lele SS, Singhal RS. Comparison of artificial neural network (ANN) and response surface methodology (RSM) in fermentation media optimization: Case study of fermentative production of scleroglucan. Biochem Eng J 2008;41:266–73. http://dx.doi.org/10.1016/j.bej.2008.05.009.

[20] Bas D, Boyac IH. Modeling and optimization II: Comparison of estimation capabilities of response surface methodology with artificial neural networks in a biochemical reaction. J Food Eng 2007;78:846–54. http://dx.doi.org/10.1016/j.jfoodeng.2005.11.025.

[21] Lakshminarayanan AK, Balasubramanian V. Comparison of RSM with ANN in predicting tensile strength of friction stir welded AA7039 aluminium alloy joints. Trans Nonferrous Metals Soc China 2009;19:9–18. http://dx.doi.org/10.1016/S1003-6326(08)60221-6.

[22] Nannipieri P, Kandeler E, Ruggiero P. Enzyme activities and microbiological and biochemical processes in soil. In: Burns RG, Dick RP, editors. Enzymes in the environment. New York: Marcel Dekker Inc.; 2002. p. 1–33.

[23] Hamedi J, Shahverdi AR, Samadi N, Mohammadi A, Shiran M, Akhondi S. A cylinder-plate method for microbiological assay of clavulanic acid. Pharmeur Sci Notes 2006;2006:53–4.

[24] Souza MJ, Rolim CM, Melo J, Souza Filho PS, Bergold AM. Development of a microbiological assay to determine the potency of ceftiofur sodium powder. J AOAC Int 2007;90:1724–8.

[25] Lopes CCGO, Salgado HRN. Development and validation of a stability indicative agar diffusion assay to determine the potency of linezolid in tablets in the presence of photodegradation products. Talanta 2010;82:918–22. http://dx.doi.org/10.1016/j.talanta.2010.05.056.

[26] Hopfield JJ. Artificial neural networks. IEEE Circuits Devices Mag 1988;4:3–10. http://dx.doi.org/10.1109/101.8118.

[27] Yegnanarayana B. Artificial neural networks. New Delhi: PHI Learning Pvt. Ltd; 2009.

[28] Dayhoff JE, DeLeo JM. Artificial neural networks. Cancer 2001;91:1615–35. http://dx.doi.org/10.1002/1097-0142(20010415)91:8+%3C1615::AID-CNCR1175%3E3.0.CO;2-L.

[29] Svozil D, Kvasnicka V, Pospichal J. Introduction to multi-layer feed-forward neural networks. Chemom Intell Lab Syst 1997;39:43–62. http://dx.doi.org/10.1016/S0169-7439(97)00061-0.

[30] Johansson EM, Dowla FU, Goodman DM. Backpropagation learning for multilayer feed-forward neural networks using the conjugate gradient method. Int J Neural Syst 1991;2:291–301. http://dx.doi.org/10.1142/S0129065791000261.

[31] Smits JRM, Melssen WJ, Buydens LMC, Kateman G. Using artificial neural networks for solving chemical problems: Part I. Multi-layer feed-forward networks. Chemom Intell Lab Syst 1994;22:165–89. http://dx.doi.org/10.1016/0169-7439(93)E0035-3.

[32] Specht DF. A general regression neural network. IEEE Trans Neural Netw 1991;2: 568–76. http://dx.doi.org/10.1109/72.97934.

[33] Goulermas JY, Liatsis P, Zeng XJ, Cook P. Density-driven generalized regression neural networks (DD-GRNN) for function approximation. IEEE Trans Neural Netw 2007;18:1683–96. http://dx.doi.org/10.1109/TNN.2007.902730.

[34] Khan J, Wei JS, Ringnér M, Saal LH, Ladanyi M, Westermann F, et al. Classification and diagnostic prediction of cancers using gene expression profiling and artificial neural networks. Nat Med 2001;7:673–9. http://dx.doi.org/10.1038/89044.

[35] Baxt WG. Use of an artificial neural network for the diagnosis of myocardial infarction. Ann Intern Med 1991;115:843–8. http://dx.doi.org/10.7326/0003-4819-115-11-843.

[36] Hoskins JC, Himmelblau DM. Artificial neural network models of knowledge representation in chemical engineering. Comput Chem Eng 1988;12:881–90. http://dx.doi.org/10.1016/0098-1354(88)87015-7.

[37] Li H, Leng WJ, Zhou YB, Chen FD, Xiu ZL, Yang DZ. Evaluation models for soil nutrient based on support vector machine and artificial neural networks. Sci World J 2014; 2014:7. http://dx.doi.org/10.1155/2014/478569.

[38] Kandirmaz HM, Kaba K, Avci M. Estimation of monthly sunshine duration in turkey using artificial neural networks. Int J Photoenergy 2014;2014:9. http://dx.doi.org/10.1155/2014/680596.

[39] Popescu I, Kanatas IA, Constantinou AP, Nafornita I. Application of general regression neural networks for path loss prediction. Proceedings of international workshop trends and recent achievements in information technology; 2002.

[40] Deng N, Tian Y, Zhang C. Support vector machines: Optimization based theory, algorithms, and extensions. 1st ed. New York: Chapman & Hall/CRC; 2012.

[41] Zhong X, Li J, Dou H, Deng SJ, Wang GF, Jiang Y, et al. Fuzzy nonlinear proximal support vector machine for land extraction based on remote sensing image. PLoS One 2013;8:e69434. http://dx.doi.org/10.1371/journal.pone.0069434.

[42] Shen Y, He Z, Wang Q, Wang Y. Feature generation of hyperspectral images for fuzzy support vector machine classification. IEEE Instrum Meas Technol Conf 2012: 1977–82. http://dx.doi.org/10.1109/I2MTC.2012.6229278.

[43] Kim DW, Lee K, Lee D, Lee KH. A kernel-based subtractive clustering method. Pattern Recogn Lett 2005;26:879–91. http://dx.doi.org/10.1016/j.patrec.2004.10.001.

[44] Fan RE, Chang KW, Hsieh CJ, Wang XR, Lin CJ. LIBLINEAR: A library for large linear classification. J Mach Learn Res 2008;9:1871–4.

[45] Guo Q, Liu Y. ModEco: An integrated software package for ecological niche modeling. Ecography 2010;33:637–42. http://dx.doi.org/10.1111/j.1600-0587.2010.06416.x.

[46] Pollar M, Jaroensutasinee M, Jaroensutasinee K. Morphometric analysis of *Tor tambroides* by stepwise discriminant and neural network analysis. World Acad Sci Eng Technol 2007;33:16–20.

[47] Friesen D, Patterson M, Harmel B. A comparison of multiple regression and neural networks for forecasting real estate values. Reg Bus Rev 2011; 30:114–36.

[48] Vouk D, Malus D, Halkijevic I. Neural networks in economic analyses of wastewater systems. Expert Syst Appl 2011;38:10031–5. http://dx.doi.org/10.1016/j.eswa.2011.02.014.

[49] Tetko IV, Livingstone DJ, Luik AI. Neural network studies. 1. Comparison of overfitting and overtraining. J Chem Inf Comput Sci 1995;35:826–33. http://dx.doi.org/10.1021/ci00027a006.

[50] Tu JV. Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. J Clin Epidemiol 1996;49: 1225–31.