Abstract—[Purpose] In order to enhance the extraction rate of polysaccharides, a series of statistical approach was used to optimize the ultrasonic-assisted extraction conditions from the mycelium of *Paecilomyces tenuipes* Pt196. [Method] The optimization of conditions was carried out in two stages. Firstly, the effects of various experimental parameters considered for the investigation (ultrasonic power, ultrasonic time and liquid-solid ratio) were studied using the method of single factor test design experiments. Secondly, a 15-run Box-Behnken design was performed to optimize the extraction conditions of polysaccharides. The experimental results of Box-Behnken design were analyzed by response surface methodology and artificial neural network together with genetic algorithm. [Results] The optimum conditions for polysaccharides extraction obtained by the application of artificial neural network-genetic algorithm were ultrasonic time 345s, ultrasonic power 320 W and liquid-solid ratio 95 mL·g⁻¹. [Conclusions] Artificial neural network-genetic algorithm can effectively select the best extraction conditions of polysaccharides in this study.

Keywords— polysaccharides; *Paecilomyces tenuipes*; response surface methodology; artificial neural network; genetic algorithm

I. INTRODUCTION

*Paecilomyces tenuipes* (also called *Isaria japonica*), a caterpillar fungi, belong to the subphylum Ascomycotina, class Pyrenomycetes, order Clavicipitales. This fungus is traditionally used as health foods for improving blood loss, fatigue and anorexia in Korea, Japan and China. Recently, some valuable constituents and beneficial biological activities of *Paecilomyces tenuipes* were found [1-4]. Polysaccharides are polymeric carbohydrate structure and obtained from various organisms, such as microorganisms, plants, and animals. Many studies show that polysaccharides possess efficient bioactivities including immune regulation, antioxidant, anti-tumor and hypoglycemic activities [5-7].

Extraction of bioactive compounds is an important step in application of *Paecilomyces tenuipes*. In this study, ultrasonic-assisted extraction technology, one of the most effective extraction techniques, was used [8]. Many factors will significantly affect the extraction efficacy of many natural products [9]. Hence, it is necessary to optimize the extraction conditions to obtain highest extraction rate of polysaccharides.

Response surface methodology (RSM) is introduced as one of the most commonly used tool to study the effects of individual measured parameters and their interactions on response variable [10]. This methodology can greatly reduce the number of experiments and enhance the work efficiency [11]. Another approach used in this study is artificial neural networks (ANN) coupled with genetic algorithm (GA). Unlike the regression method, ANN is a powerful and reliable tool for solving nonlinear problems [12]. The genetic algorithm is an efficient optimization procedure based on natural evolution. It can be used to optimize the complex problems. Many experimental studies show that ANN - GA is powerful tool in the design of optimal process in fermentation research and superior to RSM [13].

The present work attempts to relate the effects of the extraction conditions on the extraction rate of polysaccharides from the mycelium of *Paecilomyces tenuipes* Pt196 using a sequential statistical strategy including Box-Behnken design (BBD), multivariate quadratic regression (MQR) model and ANN model.

II. MATERIALS AND METHODS

A. Microorganism and medium

*Paecilomyces tenuipes* Pt196 was obtained from Anhui Agricultural University. This fungus was cultivated on a medium containing (g·L⁻¹) : peptone 10, glucose 40, yeast extract 10. For the production of mycelium, liquid fermentation was employed. The fermentation conditions were: temperature 26, inoculation amount 7%, rotate speed 150rpm, time 4d, liquid volume in 250mL flask 100mL, seed age 3d, natural pH. After incubation, the biomass was harvested by centrifugation for 8 min at 3500×g to separate it from the liquid medium. After repeated washing with distilled water, the mycelium pellets were freeze-dried for this experiment.

B. Sample preparation

The polysaccharide was extracted by water assisted with ultrasound at various ultrasonic time, ultrasonic power and liquid-solid ratio from the powder of lyophilized mycelium. The extracted polysaccharide was centrifuged at 4297×g for 5 min to remove the particulate matters and used in this study.
C. Analysis method

The polysaccharides content were measured using a literature method [14].

D. Experimental design

1) Single-factor test design experiments

Traditional single-factor test, in which one factor varies at a time while all other factors keeps invariant, was used in this study in order to select a suitable ultrasonic power, ultrasonic time and liquid-solid ratio. Firstly, samples (0.1g) were soaked and extracted using 3mL water at the power of 200, 240, 280, 320, 360, 400, 440, 480, 520W, and kept for 480s. The extraction rate of polysaccharides was used as a responsive index. Secondly, the influence of the ultrasonic time on the extraction efficiency was investigated by considering ten ultrasonic times (60, 120, 180, 240, 300, 360, 420s, 480, 540, 600s). Finally, based on the results above, the best liquid-solid ratio (20, 30, 40, 50, 60, 70, 80, 90, 100, 110 mL·g⁻¹) with the scopes of investigation was determined. Considering ten ultrasonic times (60, 120, 180, 240, 300, 360, 420, 480, 540, 600s), the best liquid-solid ratio (20, 30, 40, 50, 60, 70, 80, 90, 100, 110 mL·g⁻¹) with the scopes of investigation was selected.

2) Box-Behnken design experiments

To identify the best conditions, Box-Behnken design was followed. Three different parameters, each at three levels, were studied in 15 combinations. This design is a second-order multivariate technique based on three-level incomplete factorial designs. It has been received a wide application for assessment of critical experimental conditions, that is, minimum or maximum of response function [15]. The levels of independent variables were summarized in Table 1. The design was shown in Table 2. The experiments were carried out in triplicate.

The experimental data were analyzed by two different approaches, including RSM and ANN-GA. The RSM was used to describe the individual and interactive effects of the process variable on the response value. A quadratic equation may be expressed as:

\[ Y = \beta_0 + \sum \beta_i X_i + \sum \sum \beta_{ij} X_i X_j + \sum \beta_{ik} X_i, 
\]

Where \( Y \) is the response variable; \( \beta_0 \) is the constant; \( \beta_i \), \( \beta_{ij} \) and \( \beta_{ik} \) are the regression coefficients for intercept, interaction and quadratic term, respectively; \( X_i \) and \( X_j \) are the coded values of the factors computed by Eq. (2):

\[ X_i = (x_i - X_0)/\Delta X_i \]

Where \( X_i \) is the coded value of the variable, \( x_i \) is the independent variable real value, \( X_0 \) is the independent variable real value at the center point, and \( \Delta X_i \) is the step change value [16].

In addition, a more sophisticated method, ANN techniques, was also developed for maximizing the extraction rate of polysaccharides in present work. ANN techniques are especially suitable for modeling nonlinear and complicated systems. The basic ANN architecture consists of three layers: input layer, hidden layer and output layer. The input and output were the design of BBD and the results of BBD, respectively. Considering the capabilities of the ANN model, the degree of approximation \( (D_a) \) [17] was employed to search the suitable number of hidden nodes. This function can be represented as follow:

\[ D_a = \frac{n_c}{n} \times MSE_{c} + \frac{n_t}{n} \times MSE_{t} + |MSE_{c} - MSE_{t}| \]

Where \( n \): the number of calibration set and test set; \( n_c, n_t \): the number of calibration set and test set; \( MSE_{c}, MSE_{t} \): the mean-square-errors (MSE) of test set and calibration set, respectively; \( c \): the constant number (c was 0.3 in present study). The network was trained using the implemented Levenberg-Marquardt algorithm. The suitable model was subsequently used to search for the optimum parameters of extraction in the test regions based on BBD data using GA. The parameters of GA were: population size as 20, population type as double vector, the initial population as given randomly, crossover fraction as 0.8, selection function as stochastic uniform, crossover function as scattered, elite count as 2, migration fraction as 0.2, penalty factor as 100 and number of generation over which GA evolved as 100.

III. RESULTS AND DISCUSSION

A. Effects of extraction parameters on extraction rate of polysaccharides

On the basis of single variable at a time experiments, suitable ultrasonic power, ultrasonic time and liquid-solid ratio were identified. The results were shown in Fig. 1. It can be seen, the most suitable ultrasonic power, ultrasonic time and liquid-solid ratio were 360W, 360s, and 90 mL·g⁻¹ respectively.

B. Box-Behnken design

To obtain the optimum extraction conditions, the BBD was selected with three process variables at three levels. The results were shown in Table 2.

1) Response surface analysis

The results of Box-Behnken design were analyzed by RSM. A regression analysis was carried out by the software of Minitab 15 and the results were shown in Table 3. A multivariate quadratic regression model was developed based on BBD data for determining the individual effects.
and mutual interaction effects of candidate variables. The final second-order polynomial model obtained in terms of coded factors was presented below:

$$Y = 2.5446 - 0.0652X_1 - 0.1928X_2 + 0.4056X_3 - 0.1946X_1X_2 + 0.3298X_1X_3 - 0.1446X_2X_3 + 1.3751X_1^2 + 0.0198X_2^2 - 0.3458X_3^2$$  \( (4) \)

The determination of coefficient \( R^2 \) of the model was 0.9826 and indicated that only 1.38% of the total variations were not explained by the MQR model. These results demonstrated that this model fitted experiment data fairly well. \( F \)-test method was employed to exam the significances of the coefficients in this model. The results were displayed in Table 3 and the effects of individual variables and the mutual effects between the variables were learnt from these results. As can be seen, the effects of \( X_2, X_3, X_1X_2, X_2^2, X_3^2 \) were significant. The other terms were insignifican. Minitab 15 was employed to search for the maximum value in model. The optimum condition was found as follows: ultrasonic time 352 s, ultrasonic power 339 W and liquid-solid ratio 97 mL·g\(^{-1}\). Then the highest optimal value in model. The optimum condition was determined by the ANN model. The determination of coefficient \( R^2 \) of the optimized ANN model was 0.9618, indicating that 96.18% of the data variances can be explained by ANN model. The root mean square error of calibration set \( (RMSE_c) \), test set \( (RMSE_t) \) and prediction set \( (RMSE_p) \) were found to be 0.074, 0.055 and 0.084, respectively. Therefore, the ANN model had good ability of forecasting and generalization in this work. After the reliability model was developed, GA was employed to search for the optimum extraction parameters on polysaccharides extraction rate. The experimental extraction rates under the conditions recommended by RSM model. Therefore, the ANN-GA model performed better than RSM model in the optimization studies.

2) Artificial neural network and genetic algorithm

In order to achieve the purpose of this study, the data of BBD were analyzed using ANN model in this section. A three-layer feed back ANN model was structured on Matlab version 7.6.0.324 (R2008a) software. One of the data were randomly chosen and used as prediction set. Another one was randomly chosen and used as test set. The other experimental data were used as calibration set. The suitable number of hidden nodes was set to 9 and the results were presented in Fig. 4. The determination of coefficient \( R^2 \) of the optimized ANN model was 0.9618, indicating that 96.18% of the data variances can be explained by ANN model. The root mean square error of calibration set \( (RMSE_c) \), test set \( (RMSE_t) \) and prediction set \( (RMSE_p) \) were found to be 0.074, 0.055 and 0.084, respectively. Therefore, the ANN model had good ability of forecasting and generalization in this work. After the reliability model was developed, GA was employed to search for the optimum extraction parameters on polysaccharides extraction rate. The experimental extraction rates under the conditions recommended by RSM model. Therefore, the ANN-GA model performed better than RSM model in the optimization studies.

![Figure 3. Contour plots of the estimated response surface for extraction rate of polysaccharides](image-url)
ACKNOWLEDGEMENTS

This work has been supported by the Central Lab of General Biology in JiLin University. Authors deeply appreciate their financial assistance and directions.

REFERENCES


