ANN and RSM based modelling for optimization of cell dry mass of Bacillus sp. strain B67 and its antifungal activity against Botrytis cinerea

Jamil Shafi a, Zhonghua Sun a, Mingshan Ji a, Zumin Gu a and Waqas Ahmad b

aDepartment of Pesticide Science, Plant Protection College, Shenyang Agricultural University, Shenyang, China; bDepartment of Horticulture, College of Agriculture, BZU Bahadur Sub-Campus Layyah, Layyah, Pakistan

ABSTRACT
The present study was conducted to present the comparative modelling, predictive and generalization abilities of response surface methodology (RSM) and artificial neural network (ANN) for optimization of fermenting medium. Cell dry mass and inhibition zone of strain B67 against Botrytis cinerea were used as response variables. The response variables were optimized and modelled as a function of five independent variables (pH, gelatine percentage, incubation period, agitation speed, and temperature) using response surface methodology and artificial neural network. The results of both approaches were compared for their modelling abilities in terms of root-mean-squared error (RMSE), mean absolute error (MAD), chi-square, and correlation coefficient, computed from experimental and predicted data. ANN models were proved to be superior to RSM with lower RMSE, MAD, and chi-square and higher values for correlation coefficient, coefficient of determination, and predictive coefficient of determination. The optimum fermenting conditions predicted were pH 6.65, gelatine 3.30%, incubation period 35 h, agitation speed 163 rpm, and incubation temperature 33.64 °C, with 15.00 g/L and 31.64 mm cell dry mass and inhibition zone, respectively. The predictive models were validated experimentally and were found in agreement with experimentally obtained values.

ARTICLE HISTORY
Received 25 January 2017
Accepted 11 September 2017

KEYWORDS
Artificial neural network; Bacillus sp.; Botrytis cinerea; fermentation; response surface methodology

Introduction
The excessive use of synthetic pesticides for the management of plant pathogens has given rise to problems due to the development of resistance in pathogens against traditional fungicides [1]. Therefore the biological management of pathogens has gained interest as an alternative approach that is sustainable and safe both economically and environmentally [2]. Biological control agents have an important position in the integrated management of plant diseases and especially the bacteria from Bacillus genus have potential to serve as biological control agents [3–6].

Although a large number of biological control agents have been discovered, the number of commercially available ones is very small [7]. The contribution of biological control agents is less than 5% of the total use of pesticides in agriculture [3], which might be due to complications in the formulation and fermentation processes of these biological control agents [7]. Low yield and high production cost are major limiting factors for the development, commercialization, and implementation of biological control agents [8]. Thus, the optimization of fermenting medium is crucial to overcoming the above limitations [7]. Efficient fermentation is an important step for industrial use of bacteria as biological control agents for enhancing their cell density and yield of products of interest.

Single-factor optimization is time consuming and laborious. Furthermore, it hardly ever promises the determination of optimal conditions for effective and efficient production of the target product [9]. These limitations of the one-factor-at-a-time approach can be overcome by using alternative approaches like response surface methodology (RSM) and artificial neural network (ANN). RSM has been largely used in optimization of fermentation medium [10–15]. RSM encompasses a group of statistics-based approaches for building models, designing experiments, assessing effects of factors and searching for optimal conditions [16]. It is a designed experimental procedure which allows the simultaneous optimization of several factors [10]. The experimental responses in RSM are fitted to a quadratic function and the successful application of RSM in a number of studies suggested to be the best choice for many of the fermentation systems [17].

In recent years, ANN has arisen as an efficient and attractive approach for nonlinear multifactor modelling due to its generic structure and ability to learn from historical data [18,19]. ANN is more efficient than RSM because it does not require a prior description of proper
fitting function and it has the ability of universal approximation, i.e. approximation of almost all kinds of nonlinear functions, while RSM is only useful for quadratic approximation [17]. Moreover ANN is structured in nature and useful for getting more insight information, i.e. it has the ability to provide sensitivity analysis and to reveal the interactive effect of two factors on the system [20–23].

The present study had two objectives: (i) to maximize the cell dry mass of Bacillus sp. strain B67 and its antifungal potential against Botrytis cinerea and (ii) to compare the modelling ability of RSM and ANN for optimization of fermenting medium components, namely pH, gelatine percentage, incubation period, rotation speed, and incubation temperature. The optimized conditions predicted by both RSM and ANN were compared and experimentally verified.

Materials and methods

Biocontrol agent and fungal pathogen

The Bacillus sp. strain B67 was obtained from the Department of Pesticides Science, Plant Protection College, Shenyang Agricultural University, Liaoning, China. It was isolated from the rhizosphere of a tomato plant and was maintained in Tryptic Soy Broth (TSB) with 20% glycerol at −80 °C after initial testing of its antifungal activity against B. cinerea [24], it was activated in TSB for 48 h before use [2]. The fungal pathogen (B. cinerea) was also obtained from the Department of Pesticides Science, Plant Protection College, Shenyang Agricultural University, Liaoning, China, where it was isolated from tomato plant leaves. The biological control agent (Bacillus sp. strain B67) and the pathogen (B. cinerea) were deposited in China General Microbial Cultural Collection Centre under accessions numbers CGMCC1.15933 and CGMCC 3.18224, respectively.

Preliminary antifungal activity of biocontrol agent

Pure cultures of fungus were initially grown in Petri dishes containing standard potato dextrose agar (PDA) (20% potato extract, 2% dextrose, and 1.5% agar) medium and incubated at 28 °C for 5 days. After this period, 8 mm disks were cut from the edge of actively growing colonies of fungus with the aid of a cork borer. Two plugs were placed at opposite edges of each dish. The strain B67 was streaked on the centre of the PDA plate at the time of fungi transplanting. After incubation for 5 days at room temperature, the radii of the zones of inhibition of the fungus were measured in two perpendicular directions.

Growth medium and culture conditions

The inoculum of strain B67 was prepared by transferring a single colony to a 250 mL conical flask containing 100 mL of basal medium (standard Luria-Bertani medium) composed of tryptone 10 g/L, yeast extract 5 g/L, and NaCl 10 g/L and grown for 12 h at 28 °C. Then, 10% (v/v) of the inoculum culture was added to 250 mL conical flasks containing 100 mL of the experimental medium consisting of the basal medium adjusted according to experimental design (Table 1). The cultures were incubated at different temperatures, using different rotator speeds for the different incubation periods as determined by the experimental design (Table 1).

Estimation of cell dry mass and antifungal activity

Cells were harvested by centrifugation at 10,000 g for 20 min. The cell pellet was washed with phosphate buffer and recentrifuged. After that, the cell pellet was dried to constant mass in an oven at 80 °C for 24 h.

The cultural supernatants were used to evaluate the antifungal activity. The antifungal activity was determined by the Oxford cup plate assay system. PDA medium (70 mL) was heated until completely melted and then slowly cooled to 50 °C, and rapidly mixed with 8 mL of a spore suspension of B. cinerea (107 spores/mL) before being poured onto the 180 mm plates and allowed to solidify at room temperature. The supernatant (150 μL) was added into each Oxford cup on the plate, incubated at 28 °C for 5 days in the dark. The antifungal activity was determined by measuring the diameter of inhibition zones in two perpendicular directions.

RSM based modelling

In the present study, RSM was employed to determine the optimum medium components for cell dry mass and antifungal activity against B. cinerea. A central composite design (CCD) with five factors and five levels for each was used to design the experiment (Table 1). The medium components selected for optimization included pH (X1), gelatine percentage (X2), incubation period (X3), agitation speed (X4) and temperature (X5). Central composite designed (CCD) consisted of 16 cube, 10 axial, and 6 central pints as 32 triplicate experiments were performed in a randomized order. The central values selected for experimental design were: medium pH 6.5, gelatine 3%, incubation period 35 h, rotator speed 150 rpm, and incubation temperature 34 °C. Each variable was coded at 5 levels between −2 and 2 and the coding of the variables was done using Equation (1).

$$x_i = \frac{X_i - X_{cp}}{\Delta x_i} \quad i = 1, 2, 3, \ldots, k$$  \hspace{1cm} (1)$$

where $x_i$ is a dimensionless value of the independent variable, $X_i$ is the real value of the independent variable,
and input parameters was determined by Equation (2).

\[ Y = b_0 + \sum_{j=1}^{k} b_j x_j + \sum_{i=1}^{k} \beta_{ij} x_i^2 + e \]  

(3)

where \( Y \) is the predicted response; \( b_0 \) is the intercept constant; \( \beta_j \), \( \beta_{ij} \) and \( \beta_{i j} \) are the interaction coefficients of the linear, quadratic and the second-order terms, respectively; \( k \) is the number of factors; \( x_i \) and \( x_j \) are variables (\( i \) and \( j \) range from 1 to \( k \)); \( e \) is the error [25-26].

Multiple regression was employed to analyse the experimental data, and the significance of regression coefficients was evaluated by \( F \)-test. Modelling was started with a quadratic model including linear, squared, and interaction terms and the model competences were tested in terms of \( R^2 \), adjusted \( R^2 \), and predicted \( R^2 \) values. The significant terms in the model were determined by analysis of variance (ANOVA) for each response, and ANOVA tables were constructed. The regression coefficients were used to make statistical calculations to generate contour plots from the regression models.

**ANN based modelling**

The feed forward architecture of ANN, also known as multilayer perceptron (MLP) was further used to provide non-linear mapping between the input and the output variables. ANN was employed to construct the predictive model with five input variables (pH, gelatine percentage, incubation period, agitation speed, and temperature) and two output variables (cell dry mass and inhibition zone) for the same set of experimental data used for RSM.

The input layer consisted of five neurons, i.e. pH, gelatine percentage, incubation period, agitation speed, and temperature. The hidden layers were two neurons each. The output layer consisted of two neurons for each response. The back propagation algorithm was used to train the model for 1000 epochs, and the mean square error was used as the performance criterion.
temperature, while the output layer comprised of two neurons, i.e. cell dry mass and inhibition zone. The data flow in a forward direction, i.e. from an input layer to an output layer. A real number quantity, known as weights, is linked with the assembly of neurons, which is a variable parameter of the network. The neurons in the input layer just introduce the scaled input data to the hidden layer via weights. The neurons in the hidden layer perform two tasks, i.e. they sum up the weighted inputs to neurons, including bias as shown by Equation (4).

$$\text{sum} = \sum X_i W_{ij} + \theta_j$$

where $W_{ij}$ are the connection weights; $X_i$ is the input variable; $\theta_j$ is bias.

The weighted output is then passed through an activation function $f(\text{sum})$. The activation shifts the space in non-linearity of input data. The sigmoid transfer function was used in this study, as shown in Equation (5).

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

The input for the output layer comes from the output produced by the hidden layer. The same function is used by the neuron in the output layer and produces the same output using the same technique as that of the neuron in the hidden layer. ANN training is an iterative process where a pre-specified error function is minimized by adjusting the weights appropriately. An error function (the root-mean-squared error (RMSE) based upon the calculated output and actual experimental output was formulated using Equation (6).

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

where $i$ is the index of pattern; $N$ is the number of patterns used in data training; $M$ is the number of output nodes; $y_n^i$ and $\hat{y}_n^i$ are the target and predicted outputs of the $n$th node, respectively.

**Comparison of RSM and ANN**

The goodness of fit and prediction abilities of computed models were evaluated by calculating RMSE [27], mean absolute error (MAE) [28], chi square ($\chi^2$) [29] and correlation coefficients ($R^2$) [30]. The formulas used for these analyses are given in Equations (6)–(9), respectively. Furthermore, the values predicted by RSM and ANN were plotted against the corresponding experimental values.

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^{n} | Y_{ie} - Y_{ip} |$$

$$\chi^2 = \sum_{i=1}^{n} \frac{(Y_{ip} - Y_{ie})^2}{Y_{ip}}$$

$$R^2 = \frac{\sum_{i=1}^{n}(Y_{ip} - Y_{ie})^2}{\sum_{i=1}^{n}(Y_{ip} - Y_e)^2}$$

where $n$ is the number of experiments; $Y_{ie}$ is the experimental value for $i$th experiment; $Y_{ip}$ is the predicted value for the $i$th experiment; $Y_e$ is the average of the experimental values.

**Results and discussion**

**RSM modelling**

During the preliminary confirmation, strain B67 showed strong antifungal activity against $B.\,\text{cinerea}$ (Figure 1). Then the study was mainly focused to inspect the effect of five independent variables, i.e. pH, gelatine, incubation period, temperature, and agitation speed on cell dry mass and inhibition zone. A central composite design

![Figure 1. Antifungal activity of strain B67 against Botrytis cinerea. Strain B67 dual cultured with Botrytis cinerea (A) vs. sterile distilled water as control (B).](image-url)
(CCD) of 16 cube, 10 axial, and 6 central points, leading to a total 32 experiments, was performed. Linear interactive and quadratic coefficients for experimental variables were calculated to develop the regression models. Consecutive sum of squares and descriptive statistics were performed to determine the models’ capabilities. For developing predictive models for cell dry mass and inhibition zone, a second-order polynomial regression analysis was performed. Based on CCD, the second-order polynomial equations were derived. The final equations are given below (Equations (10) and (11)).

\[
Y_1 = -418.0 + 81.63X_1 + 8.314X_2 + 4.668X_3 \\
+ 0.2057X_4 + 3.017X_5 - 6.654X_1X_1 \\
- 1.314X_2X_2 - 0.05504X_3X_3 - 0.000466X_4X_4 \\
- 0.05354X_5X_5 + 0.1708X_1X_5 \\
- 0.001210X_2X_4 - 0.01533X_3X_5 \tag{10}
\]

\[
Y_2 = -1028.4 + 192.4X_1 + 16.63X_2 + 10.959X_3 \\
+ 0.4721X_4 + 9.713X_5 - 15.141X_1X_1 \\
- 2.610X_2X_2 - 0.11891X_3X_3 - 0.001064X_4X_4 \\
- 0.13466X_5X_5 + 0.2104X_1X_5 \\
- 0.00287X_2X_4 - 0.05812X_3X_5 \tag{11}
\]

where \(Y_1\) and \(Y_2\) are the predicted values for cell dry mass and inhibition zone respectively while \(X_1, X_2, X_3, X_4, X_5\) and \(X_6\) are medium pH, gelatine percentage, incubation period, agitation speed, and incubation temperature, respectively.

The fitness and adequacy of the models were analysed by ANOVA. The \(F\) values (81.82, 131.08) indicate that both of the models have satisfactory goodness of fit (Table 2). The models \(p\) values were very low, i.e. 0.000 also confirm the models results (Table 3). \(P\) values were used to evaluate the significance of different coefficients, which provides the information required to understand the interaction pattern among the experimental variables. Smaller \(P\) values refer to larger significance of the respective coefficient. The coefficients estimates and their respective \(P\) values showed that, among all tested variables, \(X_1\) (pH), \(X_2\) (gelatine), \(X_3\) (agitation speed), \(X_4\) (incubation period), and \(X_5\) (temperature) had a significant effect in both models, i.e. on cell dry mass and inhibition zone (Table 3).

The models’ fitness was expressed by the coefficient of determination (\(R^2\)) and adjusted coefficient of determination \(R^2\). The values for \(R^2\) were 0.9834 and 0.9881 for cell dry mass and inhibition zone, respectively, which showed that both of the models were >98% accurate to express the variability of the data (Table 2). The adjusted coefficient of determination corrects the \(R^2\) value for sample size and number of terms in the model, and its values (0.9713 for cell dry mass and 0.9796 for inhibition zone) were also high enough to express the high significance of models (Table 2).

The effects of the independent variables and the interactive effects of each independent variable on the response variables were illustrated by contour plots.

| Table 2. ANOVA for the cell dry mass and inhibition zone based on central composite design. |
| --- | --- |
| Source | SS | MS | F-Value | P-Value | SS | MS | F-Value | P-Value |
| \(X_1\) | 5.208 | 5.208 | 17.45 | 0.001 | 44.55 | 44.554 | 39.17 | 0.000 |
| \(X_2\) | 4.489 | 4.489 | 15.04 | 0.001 | 22.62 | 22.620 | 19.89 | 0.000 |
| \(X_3\) | 7.594 | 7.594 | 25.45 | 0.000 | 31.05 | 31.054 | 27.30 | 0.000 |
| \(X_4\) | 33.229 | 33.229 | 111.35 | 0.000 | 163.80 | 163.804 | 144.00 | 0.000 |
| \(X_5\) | 2.089 | 2.089 | 7.00 | 0.016 | 10.53 | 10.534 | 9.26 | 0.007 |
| \(X_1^2\) | 81.174 | 81.174 | 272.01 | 0.000 | 420.29 | 420.286 | 369.47 | 0.000 |
| \(X_2^2\) | 50.610 | 50.610 | 169.59 | 0.000 | 199.86 | 199.856 | 175.69 | 0.000 |
| \(X_3^2\) | 55.541 | 55.541 | 186.12 | 0.000 | 259.22 | 259.222 | 227.88 | 0.000 |
| \(X_4^2\) | 39.796 | 39.796 | 133.36 | 0.000 | 207.59 | 207.586 | 182.49 | 0.000 |
| \(X_5^2\) | 108.955 | 108.96 | 365.11 | 0.000 | 689.35 | 689.347 | 605.99 | 0.000 |
| \(X_1X_2\) | 4.202 | 4.202 | 14.08 | 0.001 | 6.38 | 6.376 | 5.60 | 0.029 |
| \(X_1X_3\) | 1.464 | 1.464 | 4.91 | 0.040 | 8.27 | 8.266 | 7.27 | 0.015 |
| \(X_1X_4\) | 3.386 | 3.386 | 11.35 | 0.003 | 48.65 | 48.651 | 42.77 | 0.000 |
| \(X_1X_5\) | 317.431 | 24.418 | 81.82 | 0.000 | 1704.0 | 131.079 | 115.23 | 0.000 |
| \(X_2X_3\) | 5.157 | 0.397 | 9.25 | 0.011 | 13.94 | 1.073 | 0.82 | 0.645 |
| \(X_2X_4\) | 0.214 | 0.043 | 5 | 0.653 | 1.307 |
| \(X_2X_5\) | 322.802 | 31 | 1724.5 |
| \(R^2\) | 0.9834 | 0.9881 |
| \(Adj R^2\) | 0.9713 | 0.9796 |
| \(Pred R^2\) | 0.9425 | 0.9659 |

Note: \(X_1\), pH; \(X_2\), gelatin (%); \(X_3\), incubation period (h); \(X_4\), agitation speed (rpm); \(X_5\), temperature (°C).
The oval shape of the contour plots indicates a significant interaction between the independent variables. The smallest ellipses in the contour plots represent the maximum predicted values. The cell dry mass and inhibition zone for each pair of independent variables are presented in Figures 2(A–J) and 3(A–J). It is apparent that the cell dry mass and inhibition zone increase with the increase in pH from 5.5 to 6.65 but a further increase in pH caused a reduction in both cell dry mass and inhibition zone. A similar relation was observed for gelatine, incubation period, agitation speed, and incubation temperature. An initial increase in these variables resulted in an increase in the cell dry mass and inhibition zone but, after a certain level, further increase resulted in reduced cell dry mass and inhibition zone (Figures 2 and 3). Therefore, the maximum cell dry mass and inhibition zone were predicted at pH 6.5, gelatine concentration of 3.30%, incubation period of 35 h, agitation speed of 163 rpm and incubation temperature of 33.63 °C (Figures 2 and 3). The cell dry mass (14.80 g/L) and antifungal activity (32.25 mm inhibition zone) obtained in the present work were much higher than the values reported previously [2,31,32].

The growth of Bacillus species is generally considered an aerobic process, so the amount of dissolved oxygen is an important parameter in the fermentation process [33]. The oxygen supply is directly related to the agitation speed, and therefore, the agitation speed was found to be an important factor affecting the cell dry mass and antifungal activity. Lower or higher pH of the fermenting medium affects the behaviour of the secondary metabolites production [34,35]. Similarly, lower or higher temperatures affect the enzymatic activity of organisms [36]. The present findings are in agreement with these observations, as both of the factors, i.e. pH and temperature, were found to significantly affect the cell dry mass and the inhibition zone. The results revealed that lower or higher levels of pH, gelatine, incubation period, agitation speed, and incubation temperature were not favourable for cell dry mass and antifungal activity [37].

### ANN modelling

Multilayer perceptron (MLP) with logistic sigmoid function was used to develop ANN-based process models. An MLP with feed-forward architecture comprised of five input nodes (pH, gelatine, incubation period, agitation speed, and temperature) and two output nodes (cell dry mass and inhibition zone) was used (Figure 4). First, the ANN model was optimized to obtain minimum dimension and error during the training and testing of the data. The design of the experiment and the respective experimental yield was used for training the network. The data were divided into training, validation and test sets to avoid over-training and over-parameterization. Data training was performed by changing the number of neurons in the hidden layer and for different combinations of ANN-specific parameters, i.e. initialization and learning rate. The generalization capacity of the model was confirmed by selecting the weights of parameter resulting in the least test set of RMSE. The MLP with 10 nodes in the hidden layer resulted in the least value for the test set RMSE (Figure 4). The correlation values between the predicted and the experimental sets were 0.999 for cell dry mass and 0.997 for inhibition zone (Table 2).

### Validation of model predictions

The predictive models were validated by conducting experiments in triplicate, using the optimized conditions. The maximum predicted values for cell dry mass and inhibition zone were 15.00 g/L and 31.12 mm.

### Table 3. Regression analysis of central composite design showing models coefficients and their significance for cell dry mass and inhibition zone.

| Model term | Coef  | SE Coef | T-Value | P-Value |
|------------|-------|---------|---------|---------|
| X1         | 0.466 | 0.112   | 4.18    | 0.001   |
| X2         | 0.433 | 0.112   | 3.88    | 0.001   |
| X3         | 0.563 | 0.112   | 5.04    | 0.000   |
| X4         | 1.177 | 0.112   | 10.55   | 0.000   |
| X5         | -0.295| 0.112   | -2.65   | 0.016   |
| X1X2       | -1.664| 0.101   | -16.49  | 0.000   |
| X1X3       | -1.314| 0.101   | -13.02  | 0.000   |
| X1X4       | -1.376| 0.101   | -13.64  | 0.000   |
| X1X5       | -1.165| 0.101   | -11.55  | 0.000   |
| X2X3       | -1.927| 0.101   | -19.11  | 0.000   |
| X2X4       | 0.513 | 0.137   | 3.75    | 0.001   |
| X2X5       | -0.302| 0.137   | -2.21   | 0.040   |
| X3X4       | -0.460| 0.137   | -3.37   | 0.003   |
| X3X5       |        |         |         |         |

Note: X1, pH; X2, gelatin (%); X3, incubation period (h); X4, agitation speed (rpm); X5, temperature (°C).
respectively. The optimized predicted conditions were pH 6.65, gelatine 3.30%, incubation period of 35.0 h, agitation speed of 163.13 rpm, and incubation temperature of 33.64 °C, with the desirability of 99% and 96% for cell dry mass and inhibition zone, respectively. The average values obtained in the experiments were 4.80 g/L and 32.25 mm for cell dry mass and inhibition zone, respectively. This was in close agreement with the values predicted by the models. The good correlation between the predicted and the experimental values was observed after optimization which further justified the validity of models.
Comparison of RSM and ANN models

Nutritional constituents and culture conditions play a crucial role in the development of an effective fermentation process [38–40]. The statistical techniques allowed determination of the best fermentation conditions that give the maximum cell dry mass and antifungal activity for Bacillus sp. strain B67. The RSM and ANN models were compared for their predictive capabilities. The values predicted by RSM and ANN are presented in Table 1. Descriptive statics in terms of RME, MAE, chi-square ($\chi^2$) and correlation coefficient ($R^2$) between experimental and predicted values were calculated to compare the results (Table 4). These results showed that the ANN models have higher modelling capability compared to
the RSM models for both cell dry mass and inhibition zone. The linear regression analysis between the values predicted by RSM and ANN showed that the values predicted by ANN are much closer to the line of perfect prediction than those predicted by RSM (Figure 5). The better modelling ability of ANN can be attributed to its universal approximation ability for nonlinearity, whereas RSM is only limited to a second-order polynomial regression [26,41]. Modelling using RSM is easier compared to ANN, as ANN needs a higher number of inputs than RSM for better predictions. ANN has excellent prediction and optimization abilities, while sensitivity analysis is more precise in RSM. RSM is recommended for modelling of a new process, while ANN is best suited for nonlinear systems that include interactions higher than quadratic. Moreover ANN does not require any prior specification for suitable fitting function [42].

Conclusions

In this work, fermentation conditions for strain B67 were optimized for maximizing the cell dry mass and antifungal activity by using RSM and ANN approaches. Furthermore, the predictive and generalization abilities of both techniques were compared. ANN showed more precise optimum conditions for maximizing the cell dry mass and inhibition zone than RSM. The ANN predictive ability

Table 4. Error based comparison of RSM and ANN models.

| Parameters                  | Cell dry mass | Inhibition zone |
|-----------------------------|---------------|-----------------|
| Root mean square error (RMSR) | 0.410, 0.116  | 0.800, 0.384    |
| Mean absolute error (MAD)   | 0.334, 0.097  | 0.659, 0.220    |
| Chi-square ($x^2$)          | 0.808, 0.053  | 1.323, 0.180    |
| Correlation coefficient ($R^2$) | 0.992, 0.999  | 0.994, 0.999    |

Figure 4. A general graphical representation of a feed forward multilayer perceptron (MLP) neural network consisting of 5 input nodes in the input layer, 10 nodes in the hidden layer and 2 nodes in the output layer.

Figure 5. Comparison of experimentally obtained values with the values predicted by RSM and ANN for cell dry mass (A) and inhibition zone (B).
was proved to be better than that of RSM and it can be concluded that ANN gives a more accurate replacement of RSM.

**Disclosure statement**

No potential conflict of interest was reported by the authors.

**Funding**

This study was supported by the Ministry of Agriculture of the People’s Republic of China to the Department of Pesticide Science, College of Plant Protection, Shenyang Agricultural University, as special funds for agro-scientific research in the public interest [grant number 20133225].

**ORCID**

Jamil Shaﬁ http://orcid.org/0000-0002-8670-6631

**References**

[1] Marin DH, Romero RA, Guzman M, et al. Black Sigatoka: an increasing threat to banana cultivation. Plant Dis. 2003;87:208–222.

[2] Mosquera S, González-Jaramillo LM, Orduz S, et al. Multiple response optimization of *Bacillus subtilis* EA-CB0015 culture and identification of antifungal metabolites. Biotechnol Agric Biotechnol. 2014;3:378–385.

[3] Bailey K, Boyetchko S, Langle T. Social and economic drivers shaping the future of biological control: a Canadian perspective on the factors affecting the development and use of microbial biopesticides. Biol Control. 2010;52:221–229.

[4] Raaijmakers JM, De Bruijn I, Nybroe O, et al. Natural functions of lipopeptides from *Bacillus* and *Pseudomonas*: more than surfactants and antibiotics. FEMS Microbiol Rev. 2010;34:1037–1060.

[5] Heins SD, Manker DC, Jimenez DR, et al., inventors; Agraculture and Food Technology Corp, assignee. Method and system for controlling plant pests. United States patent US 6,103,228. 2000 Aug 15.

[6] Shaﬁ J, Tian H, Ji M. *Bacillus* species as versatile weapons for plant pathogens: a review. Biotechnol Biotechnol Equip. 2017;31:446–459.

[7] Hynes RK, Boyetchko SM. Research initiatives in the art and science of biopesticide formulations. Soil Biol Biochem. 2006;38:845–849.

[8] Schisler D, Slininger P, Behre R, et al. Formulation of *Bacillus* spp. for biological control of plant diseases. Phytopathology. 2004;94:1267–1271.

[9] Niemann H, Rath D, Wrenzycki C. Advances in biotechnology: new tools in future pig production for agriculture and biomedicine. Reprod Domest Anim. 2003;38:82–89.

[10] Rao KJ, Kim C-H, Rhee S-K. Statistical optimization of medium for the production of recombinant hirudin from *Saccharomyces cerevisiae* using response surface methodology. Process Biochem. 2000;35:639–647.

[11] Reddy PRM, Ramesh B, Mrudula S, et al. Production of thermostable β-amylase by *Clostridium thermosulfurogenes* SV2 in solid-state fermentation: optimization of nutrient levels using response surface methodology. Process Biochem. 2003;39:267–277.

[12] Wei W, Zheng Z, Liu Y, et al. Optimizing the culture conditions for higherulinase production by *Kluveromyces* sp. Y-85 and scaling-up fermentation. J Ferment Bioeng. 1998;86:395–399.

[13] Dutta JR, Dutta PK, Banerjee R. Optimization of culture parameters for extracellular protease production from a newly isolated *Pseudomonas* sp. using response surface and artificial neural network models. Process Biochem. 2004;39:2193–2198.

[14] Xiong Y-H, Liu J-Z, Song H-Y, et al. Enhanced production of extracellular ribonuclease from *Aspergillus* niger by optimization of culture conditions using response surface methodology. Biochem Eng J. 2004;21:27–32.

[15] Shaﬁ J, Mingshan J, Zhiqiu Q, et al. Optimization of *Bacillus subtilis* strain JS-786 cell dry mass and its antifungal activity against *Botrytis cinerea* using response surface methodology, Arch Biol Sci. 2017;69:469–480.

[16] Kalil S, Maugeri F, Rodrigues M. Response surface analysis and simulation as a tool for bioprocess design and optimization. Process Biochem. 2000;35:539–550.

[17] Desai KM, Survase SA, Saudagar PS, et al. 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;41:266–273.

[18] Desai KM, Vaidya BK, Singhal RS, et al. Use of an artificial neural network in modeling yeast biomass and yield of β-glucan. Process Biochem. 2005;40:1617–1626.

[19] Yang Y, Gao M, Yu X, et al. Optimization of medium composition for two-step fermentation of vitamin C based on artificial neural network–genetic algorithm techniques. Biotechnol Biotechnol Equip. 2015;29:1128–1134.

[20] Gevrey M, Dimopoulou I, Lek S. Review and comparison of methods to study the contribution of variables in artificial neural network models. Ecol Model. 2003;160:249–264.

[21] Jaiswal S, Benson E, Bernard J, et al. Neural network modelling and sensitivity analysis of a mechanical poultry catching system. Biosys Eng. 2005;92:59–68.

[22] Gevrey M, Dimopoulou I, Lek S. Two-way interaction of input variables in the sensitivity analysis of neural network models. Ecol Model. 2006;195:43–50.

[23] Sewsenyker Y, Kana EBG, Lateef A. Modelling of biohydrogen generation in microbial electrolysis cells (MECs) using a committee of artificial neural networks (ANNs). Biotechnol Biotechnol Equip. 2015;29:1208–1215.

[24] Ji M, Wang Y, Cheng G, et al. Selection of antagonistic strains against watermelon wilt and their biocontrol efficiency. Chin J Biol Control. 2002;18:74–78.

[25] Montgomery DC. Design and analysis of experiments. New York (NY): Wiley; 2017.

[26] Maran JP, Mekala V, Manikanadan S. Modeling and optimization of ultrasound-assisted extraction of polysaccharide from *Cucurbita moschata*. Carbohydr Polym. 2013;92:2018–2026.

[27] García-Gimeno R, Hervas-Martinez C, Rodríguez-Perez R, et al. Modelling the growth of *Leuconostoc mesenteroides*
by artificial neural networks. Int J Food Microbiol. 2005;105:317–332.

[28] Youssefi S, Emam-Djomeh Z, Mousavi S. Comparison of artificial neural network (ANN) and response surface methodology (RSM) in the prediction of quality parameters of spray-dried pomegranate juice. Dry Technol. 2009;27:910–917.

[29] Çelekli A, Geyik F. Artificial neural networks (ANN) approach for modeling of removal of Lanaset Red G on Chara contraria. Bioresour Technol. 2011;102:5634–5638.

[30] Rajendra M, Jena PC, Raheman H. Prediction of optimized pretreatment process parameters for biodiesel production using ANN and GA. Fuel. 2009;88:868–875.

[31] Huang C-J, Wang T-K, Chung S-C, et al. Identification of an antifungal chitinase from a potential biocontrol agent, Bacillus cereus 28-9. BMB Reports. 2005;38:82–88.

[32] Wei X, Yan P-s, Wu H-q, et al. Antagonizing Aspergillus parasiticus and promoting peanut growth of Bacillus isolated from peanut geocarposphere soil. J Integr Agric. 2014;13:2445–2451.

[33] Gao X, He Q, Jiang Y, et al. Optimization of nutrient and fermentation parameters for antifungal activity by Streptomyces lavendulae Xjy and its biocontrol efficacies against Fulvia fulva and Botryosphaeria dothidea. J Phytopathol. 2016;164:155–165.

[34] Yang X, Yang H, Jian H, et al. Effect of fermentation conditions on antibiotic production of Xenorhabdus nematophilus. Wei sheng wu xue tong bao. 2000;28:12–16.

[35] Yang X, Qiu D, Jiao N, et al. Cultural medium and fermentation conditions of (sl Xenorhabdus) sp. strain D43. Chin J Biol Control. 2005;22:58–62.

[36] Lai L-ST, Tsai T-H, Cheng T-Y. The influence of culturing environments on lovastatin production by Aspergillus terreus in submerged cultures. Enzyme Microb Technol. 2005;36:737–748.

[37] Wang Y-H, Zhang X. Influence of agitation and aeration on growth and antibiotic production by Xenorhabdus nematphilus. World J Microbiol Biotechnol. 2007;23:221–227.

[38] Ripa F, Nikkon F, Zaman S, et al. Optimal conditions for antimicrobial metabolites production from a new Streptomyces sp. RUPA-08PR isolated from Bangladeshi soil. Mycobiology. 2009;37:211–214.

[39] Usha Kiranmayi M, Sudhakar P, Sreenivasulu K, et al. Optimization of culturing conditions for improved production of bioactive metabolites by Pseudonocardia sp. VUK-10. Mycobiology, 2011;39:174–1781.

[40] Song Q, Huang Y, Yang H. Optimization of fermentation conditions for antibiotic production by actinomycetes YJ1 strain against Sclerotinia sclerotiorum. J Agric Sci. 2012;4:95–102.

[41] Maran JP, Sivakumar V, Thirugnanasambandham K, et al. Artificial neural network and response surface methodology modeling in mass transfer parameters predictions during osmotic dehydration of Caricapapaya L. Alexandria Eng J. 2013;52:507–516.

[42] Mansour Ghaffari M, Mostafa K. Comparison of response surface methodology and artificial neural network in predicting the microwave-assisted extraction procedure to determine zinc in fish muscles. Food Nutr Sci. 2011;2:803–808.