Saudi Journal of Biological Sciences 26 (2019) 1154–1160

Contents lists available at ScienceDirect

Saudi Journal of Biological Sciences

journal homepage: www.sciencedirect.com

Original article

Prediction model of PSO-BP neural network on coliform amount in special food

Yun Deng [a,c], Hanjie Xiao [b], Jianxin Xu [a,d,*], Hua Wang [a,d]

[a] Quality Development Institute, Kunming University of Science and Technology, Kunming 650093, China
[b] Business School, Huzhou University, Huzhou 313000, China
[c] Comprehensive Testing Center for Quality and Technology Supervision of Dehong Prefecture, Mangshi 678400, China
[d] State Key Laboratory of Complex Nonferrous Metal Resources Clean Utilization, Kunming University of Science and Technology, Kunming 650093, China

ARTICLE INFO

Article history:
Received 5 January 2019
Revised 21 June 2019
Accepted 30 June 2019
Available online 2 July 2019

Keywords:
Principal component analysis
Particle swarm algorithm
BP neural network
Coliform bacteria
Sa pie

ABSTRACT

Special food safety supervision by means of intelligent models and methods is of great significance for the health of local people and tourists. Models like BP neural network have the problems of low accuracy and poor robustness in food safety prediction. So, firstly, the principal component analysis was used to extract the key factors that influenced the amount of coliform communities, which was applied to reduce the dimension of this model as the input variable of BP neural network. Secondly, both the particle swarm optimization (PSO) and BP neural network were implemented to optimize initial weights and threshold to obtain the optimal parameter, and a model was constructed to predict the amount of coliform bacteria in Dai Special Snacks, Sa pie, based on PSO-BP neural network model. Finally, the predicted value of the model is verified. The results show that MSE is 0.0097, MAPE is 0.3198 and MAE is 0.0079, respectively. It was clear that PSO-BP model was better accuracy and robustness. That means, this model can effectively predict the amount of coliform. The research has important guiding significance for the quality and the production of Sa pie.

0. Introduction

Special food of the minority nationalities in Yunnan is abundant with unique taste, which is favored by local people and tourists. But its safety supervision is subject to technology, resources and capability building (National Health Commission of the People's Republic of China, 2004). Sa pie, a kind of special food of Dai nationality in Dehong, Yunnan, is made of soup, meat slices and rice noodle. In this paper, we take Sa pie as an example. It is full of special ingredients and spices. Main indicators of current inspection include: 17 amino acid such as histidine, serine and aspartic acid, total acidity (mainly lactic acid), protein, fat, selenium and coliform. Coliform is not named after bacteriology. Instead, it is from sanitary bacteriology. In recent years, food safety incidents occur frequently, which has aroused the concern of most people. As an important indicator of food safety, coliform is related to detect the possibility of enteropathogenic microorganism and determine if there is any potential threat of food poisoning and epidemic diseases, which has been widely used in food sanitation work (Jin and Lu, 2011; Rui-ying et al., 2012). Therefore, from the perspective of food safety, the amount of coliform tested in food must be controlled in a certain range (Qian et al., 2015). How to predict the amount of coliform in Sa pie and determine whether it is safe for people's health has been the hot issue of the watchdog, and prediction of its total amount is significantly important.

Scholars at home and abroad conduct the study of food safety assessment indicator system and evaluation models from different angles like the supply chain, of food safety assessment indicator system and assessment models from different angles like the supply chain, public opinions, and testing data of food safety (Wei-dong et al., 2007; Geng et al., 2017). In terms of evaluation models, according to the category and characteristics of indicator system data, scholars mainly adopt the following methods: (1) model analysis, including artificial neural network, fuzzy clustering method, decision-making tree, extreme learning machine (Wang and Yue, 2016), etc.; (2) data mining methods (Xun-ping et al., 2011); (3) exponential analysis (Ying-hua et al., 2017); (4) hazard analysis (Tie-tao et al., 2012). The above methods mostly are based on

* Corresponding author.
E-mail address: xujianxina@163.com (J. Xu).

Peer review under responsibility of King Saud University.

Production and hosting by Elsevier

https://doi.org/10.1016/j.sjbs.2019.06.016
1319-562X/

modeling from external factors like environment quality, production and processing, circulation process, safety management and monitoring (De-bin et al., 2010; Qiang et al., 2014). Few of them focus on the ingredients of food, physicochemical indicators and bacterial components to evaluate food safety risks (Kai, 2015). Existing models are established for food safety prevention, but limited by the casual inspection of the watchdog and the management value (Tong-qiang and Ye, 2017), which means it is loosely related to routine work of the watchdog. The amount of coliform bacteria is a key inspection indicator. But current experimentation is the only way for inspection, which costs high and takes a long time. To dig out the underlying information and rules from the existing inspection data is a guiding direction. BP neural network model is frequently methods in food safety assessment (Xue et al., 2015; Man et al., 2015). For instance, Cai Qiang, etc. propose a neural network model for evaluating food safety by means of Hazard Analysis and Critical Control Point (HACCP). Chen Kai (Kai, 2015) puts forward a data-driven food safety precautionary analysis method based on improved association rules, AHP method with entropy weight, and BP neural network. But there are some problems of traditional BP neural network. In network training and prediction, weight and the threshold value is produced randomly. With the increase of system order or unknown order, the fast-increasing network structure makes the rate of convergence slow down and may be caught in local minima convergence (Xiao-chuan et al., 2013), which thereby leads to an unsatisfactory result of prediction. Particle swarm algorithm (PSO), emerging in 1990 (Kennedy and Eberhart, 2011), is famous for its easy operation, high accuracy and quick convergence, and is mainly used to cope with the optimization in industrial design.

In this study, a novel PSO-BP approach that hybridizes PSO and BP neural network is proposed to forecast the amount of coliform bacteria in ethnic food. Firstly, the principal component analysis (PCA) was used to extract the key factors that influenced the amount of coliform communities, which was applied to reduce the dimension of this model as the input variable of BP neural network. Secondly, both the particle swarm optimization (PSO) and BP neural network were implemented to optimize initial weights and threshold to obtain the optimal parameter, and a model was constructed to predict the amount of coliform bacteria in Dai Special Snacks, Sa pie, based on PSO-BP neural network model. Finally, the predicted value of the model is verified. PSO combined with BP neural network model may make up for the shortage of BP neural network of slow convergence rate, and improve the accurate prediction of model, which can be a reference for the watchdog to carry out food safety supervision. In order to reveal the prediction capability of this model, we present a comparative analysis with BP model, and GA-BP model. The simulation results show that the proposed PSO-BP model consistently has the minimum statistical error of the mean absolute error (MAE), mean square error (MSE) and mean absolute percentage error (MAPE). It is concluded that the proposed approach is an effective way to improve prediction accuracy.

The remainder of the paper is organized as follows: Section 1 presents in detail materials and methods. Section 2 presents the hybrid PSO-BP approach for the amount prediction of coliform bacteria in ethnic food. The different steps of mathematical method and the experimental steps are also given in this section. Section 3 presents the experimental analysis results. Finally, the conclusions are provided in Section 4.

1. Materials and methods

1.1. Experimental data of Sa pie

This paper randomly selects 30 Sa pie samples. Amino acid analyzer, nitrogen distillation devices are used to measure the main components and determine the PH value of each sample, with the results in Table 1. Main components include: serine, amino acid, arginine, phenylalanine, glycine, methionine, aspartic acid, isoleucine, threonine, lysine, alanine, tyrosine, valine, cystine,

Table 1
Original data obtained from tests of Sa pie (part).

Testing items	Unit	Testing samples								
		N0: SP201705465	N0: SP201705466	N0: SP201705467	N0: SP201705468	. . ...	N0: SP201705492	N0: SP201705493	N0: SP201705494	N0: SP201705495
Histidine	g/100 g	0.265	0.255	0.303	0.362	. . ...	0.185	0.241	0.252	0.274
Serine	g/100 g	0.296	0.320	0.242	0.348	. . ...	0.188	0.294	0.220	0.258
Arginine	g/100 g	0.327	0.375	0.401	0.453	. . ...	0.280	0.289	0.218	0.260
Glycine	g/100 g	0.257	0.294	0.315	0.404	. . ...	0.304	0.265	0.196	0.222
Aspartic acid	g/100 g	0.683	0.637	0.790	0.780	. . ...	0.640	0.500	0.444	0.503
Glutamic acid	g/100 g	2.250	1.740	2.410	1.260	. . ...	0.951	1.910	0.864	1.120
Threonine	g/100 g	0.360	0.375	0.384	0.432	. . ...	0.302	0.355	0.249	0.290
Alanine	g/100 g	0.360	0.403	0.444	0.537	. . ...	0.316	0.360	0.269	0.328
Proline	g/100 g	0.048	0.122	0.053	0.221	. . ...	0.092	0.107	0.078	0.077
Cystine	g/100 g	0.219	0.201	0.094	0.113	. . ...	0.090	0.240	0.123	0.172
Lysine	g/100 g	0.565	0.579	0.642	0.725	. . ...	0.486	0.527	0.400	0.467
Tyrosine	g/100 g	0.355	0.290	0.316	0.324	. . ...	0.274	0.282	0.232	0.273
Methionine	g/100 g	0.253	0.326	0.264	0.301	. . ...	0.237	0.323	0.259	0.313
Valine	g/100 g	0.447	0.380	0.326	0.385	. . ...	0.317	0.439	0.315	0.374
Isoleucine	g/100 g	0.327	0.312	0.336	0.400	. . ...	0.252	0.267	0.205	0.247
Leucine	g/100 g	0.783	0.729	0.775	0.911	. . ...	0.631	0.715	0.540	0.643
Phenylalanine	g/100 g	0.476	0.376	0.347	0.412	. . ...	0.336	0.453	0.336	0.375
Total acid	g/kg	4.260	4.110	2.900	7.950	. . ...	7.140	5.140	3.400	2.650
Protein	g/100 g	9.360	8.300	9.730	10.200	. . ...	6.300	10.300	10.100	9.870
Fat	g/100 g	1.420	1.090	1.550	1.270	. . ...	1.510	1.700	1.540	1.650
PH value	/	5.540	5.430	5.930	4.980	. . ...	4.870	5.050	5.340	5.730
Selenium	μg/100 g	6.360	5.810	4.920	6.090	. . ...	3.890	6.750	9.060	4.600
Lead	mg/kg	<0.04	0.13	0.062	0.19	. . ...	0.2	0.045	<0.04	<0.04
Total arsenic	mg/kg	<0.03	<0.03	<0.03	<0.03	. . ...	<0.03	<0.03	<0.03	<0.03
Total mercury	mg/kg	<0.01	<0.01	<0.01	<0.01	. . ...	<0.01	<0.01	<0.01	<0.01
Cadmium	mg/kg	0.0061	0.0051	0.0063	<0.003	. . ...	0.0059	0.0097	0.01	0.0085
Coliform	CFU/g	680,000	150,000	0	810,000	. . ...	90,000	190,000	21,000	170,000

proline, leucine, glutamic acid, total acid, protein, fat, selenium and coliform groups.

1.1.1. Sa pie data acquisition methods

Amino acid analyzer, nitrogen distillation device, Soxhlet extraction method, atomic fluorescence spectrophotometer (shown in Fig. 1) are used to make tests. Testing methods and standards are in line with GB5009.124 – 2016 *National Food Safety Standard: Determination of Amino Acid in Food*; PH meter is used to test the total acidity of Sa pie, according to GB5009.239-2016: *National Food Safety Standard: Determination of Acidity in Food*; nitrogen distillation device is to testify the content of protein based on GB5009.5-2016: *National Food Safety Standard: Determination of Protein in Food*; soxhlet extractor is used to testify the content of fat based on GB5009.5-2016: *National Food Safety Standard: Determination of Protein in Food*; atomic fluorescence spectrophotometer is used to testify the content of selenium in Sa pie; MPN (Most Probably Amount) quantitative test method is adopted to measure the amount of coliform groups, referring to the GB 4789.3-2016 *National Food Safety Standard: Statistics of Coliform Bacteria.*

1.2. Mathematical theory

Due to the sample limitation, there are 22 testing items, including physiochemical indicator, heavy metal, and pesticide residue. These data are of high dimension and high complexity, which is typically nonlinear. If we directly use them as the input variables, the prediction accuracy cannot be guaranteed. Thus, the main thinking way of modeling is first to lower the dimension, then adopt neural network to make training, observe the training accuracy, explore the optimization and finally build the optimized neural network model.

1.2.1. Principal component analysis

The fewer variables can be used to reconstruct and convert into independent or irrelative variables by Principal component analysis (PCA). Which can lower the data dimension, that is, to select out the principal component which has small amount of original variables and maintain the most original variable information. Main calculation procedures of PCA are as follows.

(1) Standardized processing is carried out on the testing data of Sa pie.
(2) Correlation coefficient matrix R is calculated.
(3) Calculate the characteristic value and feature vector of R, and make up m new indicator variables.
(4) Choose $m$principal components and calculate its information contribution rate b_j and accumulated contribution rate α_j, shown in Table 2.
(5) Calculate the load and scoring function of each principal component on each variable, shown in Table 3.

When α_j is close to 1, we can choose j indicator variables $y_1, y_2, \ldots, y_j$ as the j principal components to replace m variables, and thereby obtain the key factors which we desire. Here, we only choose $\alpha_j \geqslant 0.85$ to make calculation of its effect on the amount of coliform. We can learn from Table 2 that the largest four characteristic values in the matrix are 21.080, 11.504, 2.983 and 0.651, all of which are over 0, and the comprehensive information contribution rate is 87.834%. We can infer that the 1, 2, 3 and 4 principal components have basically maintained the original information of all indicators, which can fully reflect the changing trend. So, it is feasible and valid that we choose the four principal components to replace the original 22 indicators.

Principle component load presents the correlation coefficient between variables and the principal component, i.e., the feature vector mentioned above. As shown in Table 3, for the 1 principal component, protein, selenium, total acid and glutamic acid have the largest contribution successively; for the 2 principal component, selenium, protein and PH value successively; for the 3 principal component, total acid and protein successively; for the 4 principal component, PH value has the largest contribution. Therefore, protein, selenium,

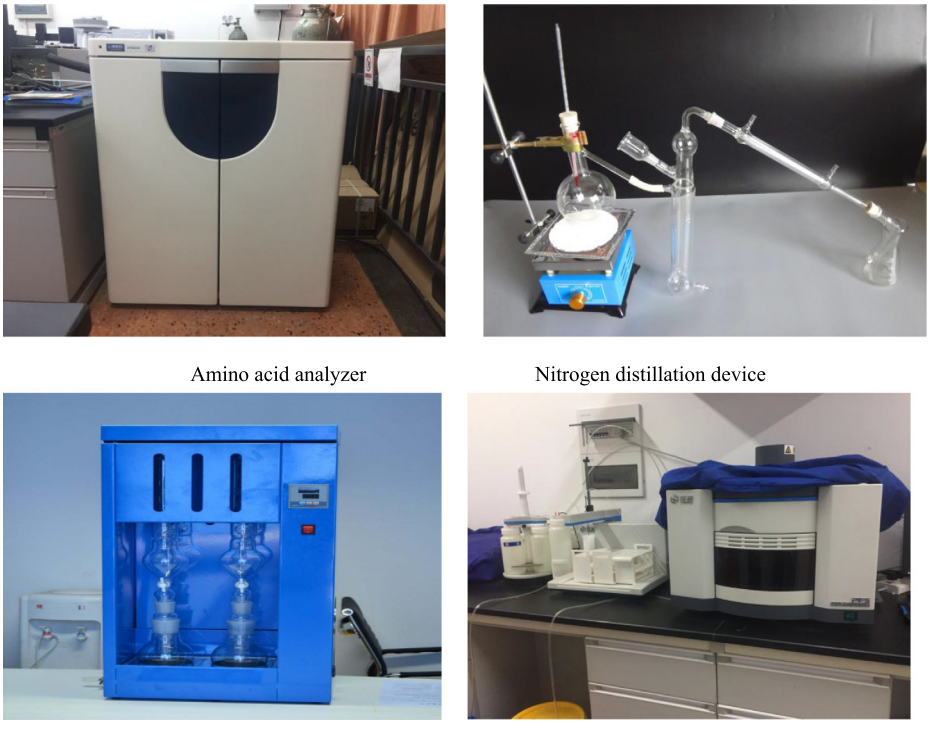

Fig. 1. Experimental devices of data acquisition.

Table 2
Characteristic values, contribution rate and accumulated contribution rate of PCA.

Variable	Characteristic value	Difference value	Contribution rate	Accumulated contribution rate
x1	21.080	9.576	64.128	64.128
x2	11.504	8.522	12.208	76.337
x3	2.983	2.331	8.410	84.747
x4	0.651	0.102	3.087	87.834
x5	0.549	0.419	2.585	90.419
x6	0.130	0.056	1.401	91.819
⋮	⋮	⋮	⋮	⋮
x16	0.016	0.002	0.441	98.270
x17	0.014	0.001	0.416	98.686
x18	0.013	0.000	0.401	99.087
x19	0.013	0.005	0.393	99.479
x20	0.008	0.005	0.309	99.788
x21	0.003	0.003	0.202	99.990
x22	0.000	0.000	0.010	100.000

Table 3
Principal component load matrix (part).

Standardized variables	Prin1′	Prin2′	Prin3′	Prin4′
x1: histidine	−0.001	−0.001	−0.001	0.001
x2: serine	0.001	0.003	0.000	0.000
x3: arginine	−0.001	0.006	0.003	0.007
x4: glycine	0.003	0.003	−0.003	0.004
x5: aspartic acid	0.005	0.006	0.007	0.008
x6: glutamic acid	0.028	0.010	−0.002	0.034
⋮	⋮	⋮	⋮	⋮
x18: total acid	0.209	−0.115	0.969	−0.065
x19: protein	0.926	−0.293	−0.233	0.024
x20: fat	0.023	−0.011	0.016	0.048
x21: PH value	0.001	0.037	0.070	0.994
x22: ‘selenium	0.311	0.948	0.044	−0.039

total acid, glutamic acid and PH value have relatively large effect on the amount of coliform bacteria of Sa pie.

1.2.2. Particle swarm algorithm

Particle Swarm Optimization (PSO) is a heuristic and evolutionary algorithm, presented by Eberhart and Kennedy in 1995 (Kennedy and Eberhart, 2011). Its basic principal is to imitate the predatory behavior of birds that birds can adjust the searching route with experience and communication with the flock. For the optimization, each solution is the way that in the search of a position in a space, particles change the flying distance and directions via changing the speed (Johnson and Wichern, 2001). Each particle remembers its optimal position p_{iD} in the searching history in the iteration process. All the optimal positions of all particles is the global optimal position p_{gD}. The equation and parameter of particle movement are as follows (Shi and Eberhart, 1998; Jain et al., 2018).

$$V_{iD}^{j+1} = \omega V_{iD}^{j} + c_1 r_1 (p_{iD}^{j} - x_{iD}^{j}) + c_2 r_2 (p_{gD}^{j} - x_{iD}^{j}) \quad (1)$$

$$x_{iD}^{j+1} = x_{iD}^{j} + v_{iD}^{j+1} \quad (2)$$

where i stands for the particle; j the current iteration amount; d the particle dimension; x_{iD}^{j} and v_{iD}^{j} are the velocity and position in the j iteration; non-negative constant c_1 and c_2 are the learning factor, which determines the effects of p_{iD} and p_{gD}on the new velocity; r_1 and r_2 are the pseudo random amount evenly distributed in the interval [0, 1]; ω is the inertia weight, adjusting the searching ability in the solution domain.

1.2.3. BP neural network model

BP neural network model is a typical local recurrent network, consisting of the input unit, output unit as well as hidden layer. In the hidden layer it can be divided into single hidden unit and multi hidden unit according to the amount of layers (Huang et al., 2019). The multi hidden unit is made of many single hidden units, which has a stronger generalization capability and a more accurate prediction but relatively longer training time compared with the single hidden unit (Xiao-chuan et al., 2013). When making a decision of the hidden layer, we should take network precision and training time into consideration. For a simple mapping relation, when the network precision is satisfied, we can decide to choose a single hidden layer to quicken the speed; for a complex mapping relation, multi hidden layer is preferred to improve the precision of prediction (Saravanan and Jerald, 2019). In this paper, the single hidden layer is used, including the input/output unit and hidden layer.

2. PSO-BP neural network prediction model

2.1. Setup of key parameters

We use Windows7 as the system platform, MATLAB2014a as the processor. In order to eliminate the effects of input variables, Mapminmax function is used to normalize the sample input data into [−1, 1]. In the network training and prediction, as the weight and threshold value are randomly produced, we make use of PSO to optimize the initial weight and threshold to acquire the optimal parameter. In the learning process, logsig and purelin functions are used to train transition function of output unit and hidden layer of the neural network, where the largest iteration amount is 50, the learning rate for the network is 0.005, and the margin of error is 0.00001.

In BP neural network model, the parameter setup is vital, where the amount of nodes in the hidden layer and neurons in the input unit are the most important variables. The amount of neurons in the input unit is directly linked to the prediction outcome. This paper, firstly, adopts PCA method to make extraction of the key factors from original data of coliform amount in Sa pie, and then, the four principal component is used as the input variable in the neural network, which can make a simplification and improve the convergence of the network.

The amount of nodes in the hidden layer is equally important. Theoretically, in the neural network, if there is a hidden layer, nodes in a necessary amount can be close to any random continuous function. However, there is no detailed calculation criterion so far. In the hidden layer, most are based on experience to set the neurons. According to the empirical formula, the range of the amount of neurons is determined as 3–13. Repeated experiments are carried out. Optimal parameters acquired in the process are used to make valuation and prediction of PSO-BP model. Network errors are compared to obtain the best node amount in the hidden layer and finally the amount is determined as 12.

In a sum, the PSO-BP neural network structure is 4-12-1. PSO parameters are set as follows: population size 40, evolutionary generation 70, acceleration factor c_1=c_2 = 1.49445, intervals of particle position and velocity [−5, 5] and [−1, l] respectively. The parameters of PSO-BP neural network are set as: training time 50, training objectives l.0e−005 and learning rate 0.005.

2.2. PSO-BP flow chart

For improving the precision of neural network model and on the basis of the above two algorithms, we build the PSO-BP prediction model for the amount of coliform bacteria in Sa pie (shown in Fig. 2). Prediction procedures of this model are as follows.

(1) Amino acid analyzer, Soxhlet extraction method, atomic fluorescence spectrophotometer and other devices are used to make tests on the Sa pie samples. 30 original data is acquired to investigate the relation of Sa pie ingredients and coliform amount.
(2) As it costs high and samples are short in amount, another 160 samples are produced randomly based on the fundamental features of the testing results, that is, there are 190 Sa pie testing results. In order to avoid the circumstance that comparison cannot be made among different units, we pre-retreat the original data and normalize them.
(3) PCA is used to determine the key factor of coliform in Sa pie to eliminate the redundancy and correlation among variables, and lower the dimensions of input variables. 170 Sa pie samples serve as a training set and the rest 20 is the testing set.
(4) Initial weight and threshold value of PSO optimized neural network are used to search for the optimal parameters.
(5) According to the reality, make the judgement whether it is over here; otherwise, return to step 4.
(6) Make a valuation on PSO-BP neural network by using the acquired optimal parameters and determine the best node in the hidden layer.
(7) Input the training sample data set and build the PSO-BP prediction model of the amount coliform bacteria in Sa pie.
(8) The testing set is used to testify the capability of the prediction model; we make a comparison between the model and other traditional prediction models to realize an accurate prediction of coliform amount in Sa pie.

2.3. Model evaluation indicators

In order to better observe the processing of data, we decide to use three error indicators: the mean square error (MSE), mean

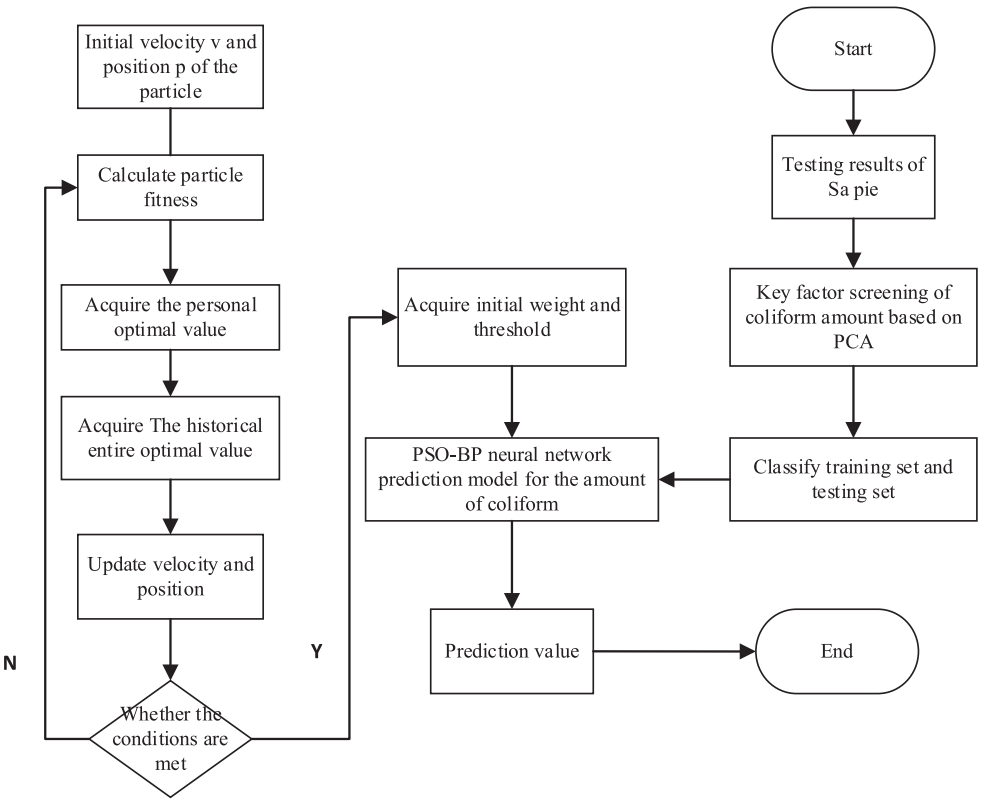

Fig. 2. Flow chart of PSO-BP prediction model for the amount of coliform bacteria in Sa pie.

absolute error (MAE), and mean absolute percentage error (MAPE) (Hannan et al., 2018; Al-Musaylh et al., 2018). MAE reflects the dissociation of prediction error, that is, a smaller value indicates a better prediction. MSE shows the reality of prediction and can evaluate the change of data, where a smaller value is favorable. MAPE explains the difference between the prediction value and original value and a smaller value is preferred. Calculation equations of error indicators are as follows.

(1) MAE

$$MAE = \frac{1}{n}\sum_{t=1}^{n}\left|Y_t - \hat{Y}_t\right| \tag{3}$$

(2) MSE

$$MSE = \frac{1}{n}\sum_{t=1}^{n}\left(Y_t - \hat{Y}_t\right)^2 \tag{4}$$

(3) MAPE

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{Y_t - \hat{Y}_t}{Y_t}\right| \tag{5}$$

where $Y_t$represents the actual value of coliform amount in Sa pie, and $\hat{Y}_t$ the corresponding prediction value.

3. Results and analysis

3.1. Prediction results of the model

Amino acid analyzer, Soxhlet extraction method, atomic fluorescence spectrophotometer and other devices are used to make tests on 30 Sa pie samples. 160 data generated randomly form the testing results of Sa pie is acquired to investigate the relation of Sa pie ingredients and coliform amount, and make the prediction of coliform bacteria via its principal components content. 170 data ranging from N0: SP201705465 to N0: SP201705635 are sample data, which are used to build the PSO-BP prediction model of coliform amount in Sa pie. Data from N0: SP201705636 to N0: SP201705655 are used as testing data for the verification of the model. In order to show a more direct observation of the prediction, a comparison analysis of the prediction value and the real value are implemented in Fig. 3 by using the BP and GA-BP neural network model. The correlation curves of prediction value and real value of coliform amount in Sa pie of each model are shown in Fig. 3.

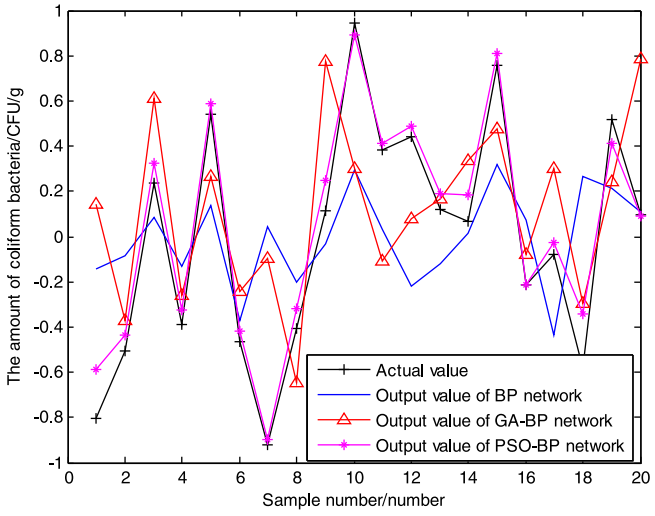

Fig. 3. Comparison curves between the prediction value and actual value of each model.

3.2. Results analysis

As is seen in Fig. 3, compared with BP neural network model, both the PSO-BP and GA-BP models demonstrate a better prediction trend and accuracy; compared with GA-BP model, PSO-BP model is better with good volatility and following quality. In other words, PSO-BP model is more suitable for prediction of the amount of coliform in Sa pie.

In order to make a further comparison with the three models, we adopt three error indicators, i.e., the MAE, MSE and MAPE. Comparison outcome is shown in Table 4. Fig. 4 presents a more visual view of prediction errors of each model.

So, we can arrive at the following conclusions.

(1) Comparing to the BP neural network model: firstly, the MSE of the prediction value of PSO-BP model is 0.0097, significantly superior to BP network 95.29%, which further indicates that there is a relatively flat change of the prediction values of PSO-BP model, that is, the following quality of PSO-BP model is better; next, mean absolute error (MAE) of the prediction value of PSO-BP model is 0.0079, excelling BP model's 0.3751 with 79.23% improvement, which shows a smaller dispersion degree; then, the MAPE of the prediction value of PSO-BP model is 0.3198, far more smaller than BP's 1.0536, which indicates a smaller difference between the prediction value and actual value; finally, the residual curves in Fig. 4 reflect that PSO-BP model's prediction error fluctuates around 0 with a small range. In general, based on the error indicators, PSO-BP model presents better performance than BP network model.

Table 4
Comparison of error indicators of each prediction model.

Prediction model	Error indicators		
	MSE	MAE	MAPE
BP model	0.2059	0.3751	1.0536
GA-BP model	0.1815	0.3239	0.9101
PSO-BP model	0.0097	0.0779	0.3198

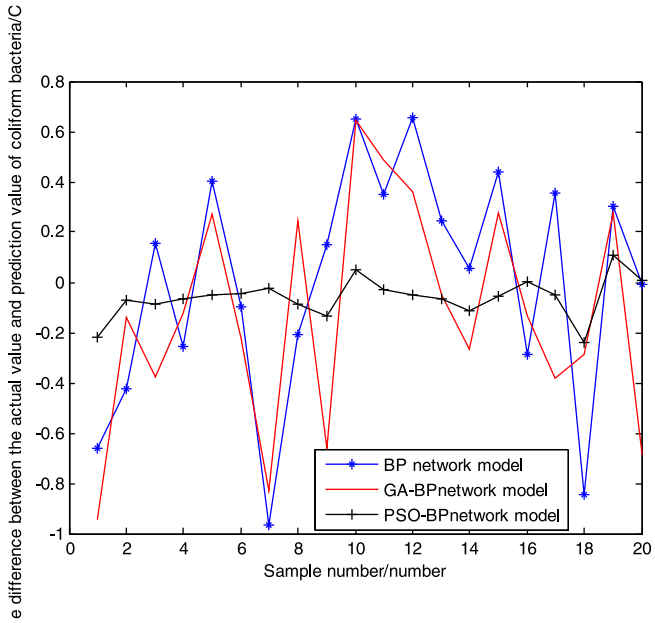

Fig. 4. Comparison of predicted residual curves of each model.

(2) Comparing to GA-BP model: firstly, the MSE of the prediction value of PSO-BP model is 94.66% superior to BP network, which further indicates that there is a relatively small change of the prediction values of PSO-BP model, that is, the following quality of PSO-BP model is better; next, mean absolute error (MAE) of the prediction value of GA-BP model is 0.3239, while PSO-BP model is of 75.95% improvement, which shows a smaller dispersion degree; then, the MAPE of the prediction value of PSO-BP model improves 64.68%, which indicates a smaller difference between the prediction value and actual value; finally, the residual curves in Fig. 4 reflects a smaller fluctuation of prediction error. In a sum, based on the above error indicators, PSO-BP model is more accurate than GA-BP model.

4. Conclusions

Accurate coliform amount prediction offers an effective evidence for the prediction of coliform amount and thereby has some guidance for the actual production of Sa pie. For this purpose, a novel PSO-BP approach, which combines particle swarm optimization (PSO) and BP neural network, is proposed to forecast the coliform amount in Sa pie. The following conclusions can be drawn from the experimental results:

(1) Amino acid analyzer, nitrogen distillation device, Soxhlet extraction method, atomic fluorescence spectrophotometer and other devices are used to make tests on the 30 Sa pie samples. 160 data are generated from the basic features of the testing results, and are used to investigate the relation of the amount of coliform bacteria and the ingredients of Sa pie. PCA analysis indicates that protein, selenium, total acid (lactic acid), glutamic acid and PH value have the large effects on the coliform amount.
(2) A PSO-BP neural network prediction model is built by using contents of main ingredients of Sa pie to make a prediction of the coliform amount, which is feasible.
(3) Comparisons are made among PSO-BP, BP and GA-BP models. Three error indicators are used to explore the relation of the prediction value and actual value of PSO-BP model for predicting the amount of coliform bacteria in Sa pie, whose MAE, MSE and MAPE are 0.0097, 0.0779 and 0.3198 respectively. The model enjoys a better prediction outcome. In other words, it reflects that the PSO-BP model has made favorable improvement, which offers an effective evidence for the prediction of coliform amount and thereby has some guidance for the actual production of Sa pie.

Acknowledgement

Zhejiang Philosophy & Social Sciences Planning Projects (Grant No. 19NDQN312YB).

References

Al-Musaylh, M.S., Deo, R.C., Yan, L., et al., 2018. Two-phase particle swarm optimized-support vector regression hybrid model integrated with improved empirical mode decomposition with adaptive noise for multiple-horizon electricity demand forecasting. Appl. Energy 217, 422–439.
De-bin, Zhang, Jia-peng, Xu, Jian-jun, Xu, et al., 2010. Model for food safety warning based on inspection data and BP neural network. Trans. CSAE 26 (1), 221–226.
Geng, Z.Q., Zhao, S.S., Tao, G.C., et al., 2017. Early warning modeling and analysis based on analytic hierarchy process integrated extreme learning machine (AHP-ELM): application to food safety. Food Control 78, 33–42. https://doi.org/10.1016/j.foodcont.2017.02.045.
Hannan, M.A., Lipu, M.S.H., Hussain, A., et al., 2018. Neural network approach for estimating state of charge of lithium-ion battery using backtracking search algorithm. IEEE Access. PP(99):1-1.
Huang, H., Jiang, M., Ding, Z., et al., 2019. Forecasting emergency calls with a poisson neural network-based assemble model. IEEE Access 7 (99), 18061–18069.
Jain, N.K., Nangia, U., Jain, J., 2018. A review of particle swarm optimization. J. Inst. Eng. 3, 1–5.
Jin, Xu, Lu, Pang, 2011. Comparison and analysis of the aerobic plate count, coliforms, escherichia coli, enterobacteriaceae in national and international food safety standards. Chin. J. Food Hyg. 23 (5), 472–477.
Jing, Qian, Heng-qiu, Xu, Jun, He, et al., 2015. Surveillance on foodborne pathogenic bacteria in foods in Hefei City between 2010 and 2012. Chin. J. Disease Control 19 (5), 533–535.
Johnson, R.A., Wichern, D.W., 2001. Practical Multivariate Statistical Analysis. Tsinghua University Press, Beijing.
Kai, Chen, 2015. Food Safety Early Warning Analysis Method and Application Based on Data – Driven. Beijing University of Chemical Technology, Beijing.
Kennedy, J., Eberhart, R., 2011. Particle swarm optimization. Proc. of 1995 IEEE Int. Conf. Neural Networks, (Perth, Australia), Nov. 27–Dec. 2011, vol. 4, no. 8, pp. 1942–1948.
Man, Zhang, Ting, Li, Yu-han, Ji, et al., 2015. Optimization of CO_2 enrichment strategy based on BPNN for tomato plants in greenhouse. J. Agric. Machinery 46 (8), 239–245.
National Health Commission of the People's Republic of China, 2004. Microbiological Testing of Food Hygiene: General Principles: Microbiological Examination of Food Hygiene. China Standard Press, Beijing.
Qiang, Cai, Jun-jun, Wang, Hai-sheng, Li, et al., 2014. Research on establishment of food safety evaluation model based on neural network. J. Food Sci. Technol. 32 (1), 69–76.
Rui-ying, Chen, Zhen, Cai, Juan, Yang, 2012. Analysis on the monitoring results of food safety in catering industries in Taizhou in 2010. Chin. J. Food Hyg. 24 (1), 63–66.
Saravanan, A., Jerald, J., 2019. Ontological model-based optimal determination of geometric tolerances in an assembly using the hybridised neural network and Genetic algorithm. J. Eng. Des. 15, 1–19.
Shi, Y., Eberhart, R., 1998. A modified particle swarm optimizer. In: Proceeding of the IEEE International Conference of Evolutionary Computation. Anchorage, AK, pp. 69–73.
Tie-tao, Wang, Guo-ying, Wang, Yue, Chen, 2012. Network public opinion threat assessment based on Elman neural network. J. Inform. Eng. Univ. 13 (4), 482–488.
Tong-qiang, Jiang, Ye, Ren, 2017. GA-BP neural network and its application in safety evaluation of liquid milk. Food Ind. Technol. 38 (5), 289–292.
Wang, J., Yue, H., 2016. Food safety pre-warning system based on data mining for a sustainable food supply chain. Food Control 73, 223–229. https://doi.org/10.1016/j.foodcont.2016.09.048.
Wei-dong, Shao, Yu-yue, Shi, Wei, Tian, 2007. Review of domestic and foreign food safety evaluation. Food Eng. 14 (2), 3–5.
Xiao-chuan, Wang, Feng, Shi, Lei, Yu., et al., 2013. MATLAB neural network, 43 case studies. Beijing University of Aeronautics and Astronautics Press, Beijing.
Xue, Liu, Ya-mei, Li, Jiao, Liu, et al., 2015. BP neural network based prediction model for fresh egg's shelf life. J. Agric. Machinery 46 (10), 328–334.
Xun-ping, Lei, Qiu, R., Yang, Wu, 2011. Food safety pre-warning model based on supply chain and expansion decision and its application. Chin. J. Saf. Sci. 21 (11), 136.
Ying-hua, Song, Yu, Zhang, Dan-hui, Fang, et al., 2017. Study on an intelligence flow control strategy for food safety supervision. J. Intell. 36 (10), 1023–1030.