



HAL
open science

Research on GA-RBF Optimization Algorithm in the Prediction of Yield Loss of Maize Diseases and Pests

Guifen Chen, Dongxue Wang, Shan Zhao, Siwei Fu

► **To cite this version:**

Guifen Chen, Dongxue Wang, Shan Zhao, Siwei Fu. Research on GA-RBF Optimization Algorithm in the Prediction of Yield Loss of Maize Diseases and Pests. 10th International Conference on Computer and Computing Technologies in Agriculture (CCTA), Oct 2016, Dongying, China. pp.257-267, 10.1007/978-3-030-06155-5_25 . hal-02179963

HAL Id: hal-02179963

<https://inria.hal.science/hal-02179963>

Submitted on 11 Jul 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Research on GA-RBF optimization algorithm in the prediction of yield loss of maize diseases and pests

Guifen Chen^{1(✉)}, Dongxue Wang¹, Shan Zhao¹, Siwei Fu¹

¹ School of Information Technology, Jilin Agricultural University, Changchun 130118, China
guifchen@163.com

Abstract. In view of the high complexity and nonlinearity of crop pests and diseases, using the traditional BP network and RBF network model method to predict is pretty difficult. And the prediction accuracy is low. Also the effect is not ideal when the sample size is small and the noise is more, therefore, this article presents a fusion optimization algorithm based on genetic-algorithm (GA) and radial basis function neural network (RBF). By unified coding the data center of RBF neural network and its corresponding expansion constant and weight, strengthened the cooperation between the hidden and output layer, furthermore using the functional characteristics of global search using genetic algorithm to obtain the optimal model of yield loss, finally predict on yield loss of maize diseases and pests. By making simulation test data of the National 863 project demonstration area - 13 village, Gong' peng town in Jilin province Yu'shu County, the experimental results show that: After using the GA to optimize the RBF in the network's structure and approximation has obvious improvement and enhancement, can effectively reflect the fluctuation characteristics of maize diseases and pests, has been widely application prospect in the agricultural field.

Keywords: Genetic algorithm; RBF neural network; fusion algorithm; maize diseases and pests; yield loss predict

1 Introduction

One of the main grain crops in our country is maize. It is also the main growing crop in the northeast region. Maize production plays an important role in food production as a whole. Crop diseases and pests is one of the important factors affecting crop yields. Therefore, plant diseases and insect pests of agricultural yield loss forecasting has become an important issue in the study of the economic development and its research and analytical prediction has a strong practical significance.

In view of the high complexity and nonlinearity of crop pests and diseases, using the traditional BP network and RBF network model method to predict is difficult. And the prediction accuracy is low. Also in solving small sample size and more noise effect is not ideal.

To solve the above problem, this article proposes an optimization algorithm based on genetic algorithm (GA) and radial basis function neural network (RBF) fusion to predict yield loss of maize diseases and pests. Build GA-RBF model to predict effects of the occurrence dynamics of plant diseases and pests on maize yield, growers can make timely disaster prevention and cost-effectively reduce the extent of the disaster, increase maize yield.

2 GA-RBF fusion algorithm

2.1 Genetic Algorithm (GA)

Genetic Algorithm referred to as GA (Genetic Algorithm), in essence, is a direct problem-free search method, composed of 4 parts: coding, fitness function, selection of Genetic Operators and parameters of the operation.

Genetic Algorithm optimization basic steps:

1. Initialize, set up an evolutionary algebra device $t \leftarrow 0$, T is the max evolutionary algebra; M individuals were randomly generated as $P(0)$.
2. Fitness function, calculate each individual fitness of $P(t)$ (expressed in $\text{Eval}(V)$).
3. Select operation, use selection operator acts on the group.
4. Cross - operation, take the cross operator P_c operates on the group.
5. Mutation operation, take the mutation operator P_m operates on the group. The group $P(t)$ gets the next generation group $P(t+1)$ after many indeed operations.
6. Termination condition judgment. If $t \leq T$, go to step two; If $t > T$, the optimal solution output is obtained by evolutionary process, during the process can find the maximum fitness individual as the optimal output. Finally, terminate the calculation.

Genetic algorithm flow chart is shown in Figure 1.

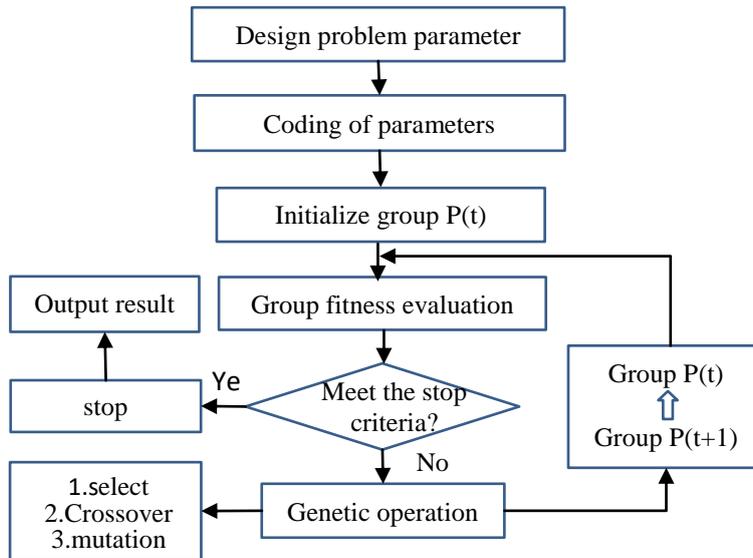


Fig. 1. Genetic algorithm flowchart

GA is a kind of optimization to the problem space encoding, with guided random search optimization method and its implied parallelism and robustness in nature is the symbol of genetic algorithm is different from the traditional optimization search method.

2.2 Radial basis function (RBF)

RBF neural network structure and multilayer forward networks are similar, it is a three layer to the network, input, hidden and the output layers. From input to hidden, the transformation is nonlinear, and from hidden to output transformation is linear.

A typical three layer topology is shown in Figure 2.

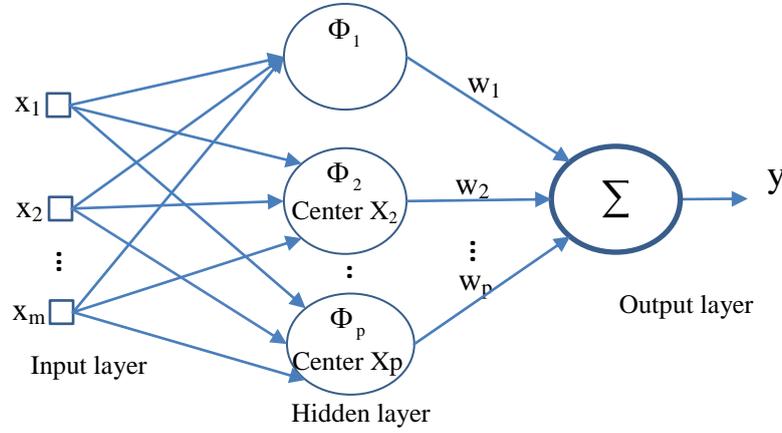


Fig.2. Genetic algorithm flowchart

RBF network hidden layer basis function $g(x)$ can have a lot of choices, without loss of generality. In this article, choice of hidden layer radial basis function is base of gaussian function is commonly used in neural network, so the RBF neural network activation function expressed as:

$$\varphi_i(x) = \exp\left(-\frac{\|x^T - c_i\|^2}{2\sigma_i^2}\right), i = 1, 2, \dots, m \quad (1)$$

$x = (x_1, x_2, \dots, x_n)$ is network input vector; $c_i = (c_{i1}, c_{i2}, \dots, c_{in})$ is the center vector of NO.i neuron in the hidden layer, is a vector with same dimension to the X ;

σ_i is the width of NO.i basis function; Next, M is the number of neurons in hidden layer;

$\|x - c_i\|$ represents between x and c_i vector Euclidean norm; The output layer is a linear layer, is the linear weighted sum of the hidden layer output. Complete following nonlinear mapping.

$$f_n(x) = w_0 + \sum_{i=1}^n w_i \varphi(\|X - c_i\|) \quad (2)$$

$X \in \mathbb{R}^n$ is input vector, is $\mathbb{R}^+ \rightarrow \mathbb{R}$ nonlinear function.

RBF can solve the nonlinear problems, it has good generalization ability, for the problems with complex function relation of functional approximation, its precision is higher. But this method just accelerate the convergence speed in a certain extent. Therefore, this paper proposes an optimization algorithm based on genetic algorithm (GA) and radial basis function neural network (RBF), which can solve this problem better.

2.3 GA-RBF

As mentioned before, neural network technology has penetrated into many fields, but there are still some problems that can not be solved. Because of the GA can converge to the global optimal solution and robustness is strong. Combine GA and RBF can not only play the ability of neural network-generalization, also can make RBF has fast-convergence and better learning-ability.

The three layers of RBF have different function. While designing the network, the most critical step is hidden layer data center C_i , the propagation constant and the value of between hidden and output the size of the weights w_i choice, in order to make RBF neural network has the best approximation effect, it needs select the most accurate parameter values.

The basic problems to be solved when using GA to optimize the parameters of RBF are:

(1) Genetic encoding corresponding to RBF

In this algorithm, the width of the RBF network is the same as the center, which is obtained by the genetic algorithm. The code includes the binary method and the real number coding method

In this article, we use a real coding method, each individual is expressed by a letter of the alphabet, and include its range, not only can improve the accuracy, but also to find the best value in a larger range

(2) Create initial group

In this paper, the population size is set to 20. According to the range of the input vector, the initial group is randomly generated in the range, the center of the initial basis function is generated, which is the initial cluster center.

(3) Construct the fitness function

The size of individual fitness in genetic algorithm to determine the probability of the individual to be inherited to the next generation of individuals. Fitness is the only information that guides the search of genetic algorithm, its good or bad is a measure of the merits of the key algorithm. The fitness function formula is as follows:

$$F(Y) = \frac{1}{E(Y)} \quad (3)$$

Among them, $E(Y)$ is the sum of the error of the network output value and the expected value of all training samples for neural network training. This selection of fitness function can directly translate the differences in performance of each individual.

(4) Genetic operator

a. Select operation: In this paper, the method of proportion selection were adopted.

b. Crossover operation: The crossover rate is the core of the genetic algorithm. Through the cross, the search ability of the algorithm can be improved.

c. Mutation operation: Mutation operation through determining locus for all individuals at random to change the gene mutation probability values, the purpose is to strengthen the regional search ability of genetic algorithm, and increase the various kind of the group.

3 Simulation experiment and result analysis

3.1 Experimental materials

The research area is the farm field of the village of the town of 13 village, Gong' peng town in Jilin province Yu'shu County, is a plain area of Changbai Mountain foreground, black calcium soil, major crops are corn and soybeans, etc. It is the most important commodity grain bases in Jilin province. During the period of rain and heat, sunshine is longer, which is beneficial to the growth of maize crops. Part of data in the experimental area is shown in Table 1.

Table 1. FontPartial Sample data in experimental area

Block area	Diseased plant rate (%)	Pest rate (%)	Yield (kg / HA)
A1	1.82	0.00	4756.98
A2	3.51	0.00	5272.62
A3	1.85	3.70	4800.49
A4	1.85	3.70	4271.18
E1	3.57	5.36	6232.88
E2	2.00	4.00	5665.62
E3	5.66	1.89	5210.89
F1	3.51	5.26	5928.16
F2	1.96	7.84	4682.38
F3	1.85	1.85	5138.17
H1	1.72	3.45	4664.02
H2	0.00	5.36	4808.12
L8	1.69	5.08	4673.46
L9	6.56	3.28	3969.84

3.2 Partial Matlab code

This article uses Matlab neural network and genetic algorithm toolbox to train the neural network, it can directly train and simulate the network by using the function provided by it.

Part of the Matlab code is as follows:

```
net=newrb(train_x,train_y,0.01,1,5,1);
```

```

6
% Constructing neural networks

Test_y=sim(net,test_x);

FieldDD=rep(RANGE,[1,NVAR]);

Chrom=crtrp(NIND,FieldDD); % Create
initial population

gen=0;

ObjV=objfuns(net,Chrom,train_x,train_y);

tic

while gen<MAXGEN

    FitnV=ranking(ObjV); % Assign fitness values

    SelCh=select('sus',Chrom,FitnV,GGAP); % Choose

    SelCh=recombin('xovsp',SelCh,0.7); %
Recombination

    SelCh=mutbga(SelCh,FieldDD,[0.05,1]); % Variation

ObjVSel=objfuns(net,SelCh,train_x,train_y);

%Calculate the value of the child generation objective
function

[Chrom ObjV]=reins(Chrom,SelCh,1,1,ObjV,ObjVSel);

gen=gen+1;

trace(gen,1)=min(ObjV);

trace(gen,2)=sum(ObjV)/length(ObjV);

```

3.3 Simulation results analysis

3.3.1 Simulation parameter setting

By using the GA and neural network toolbox provided by Matlab, the initial population of genetic algorithm is 20, and the maximum generation is 1000. Assignment of genetic operators:

(1) Selection rate (generation gap): After repeated calculation or test view, when the ratio is 0.8, the selection method is more appropriate. Principle of the method is to calculate the fitness value of each parent individually and sum all the parent individual fitness value first, then calculate the ratio of each fitness. The calculation formula is as follows:

(2) Crossover rate: The crossover rate is generally between 0.5~1.0, because the crossover rate is too small to make the search process stalled. In this article, the crossover rate is 0.7.

$$p_i = \frac{F_i}{\sum F}, i = 1, 2, 3 \dots m \quad (4)$$

(3) Mutation rate: If there is no mutation operation in genetic algorithm, it will not generate new genes, but if the mutation rate is too large, it will make the genetic algorithm into a random search, the general mutation rate is between 0.01~0.1. In this paper, the variation rate of 0.02 and 0.05 were compared.

According to the sample data, setting RBF neural network input layer neurons for two, namely Diseased plant rate and Pest rate; Output layer neuron for one, namely maize yield; Using iterative method for design of RBF network, for each iteration a neuron is added, the iteration will stop until the error sum of squares down to the target error or number of neurons reached the maximum value. Finally, a network model with 5 hidden layer neurons is obtained.

3.3.2 Experimental process and results analysis

(1) Forecast error comparison

After several simulation experiments, when the number of training samples increases, compared with the traditional RBF network, the GA-RBF fusion optimization algorithm is improved in the convergence speed and the accuracy of approximation, and the global search ability of the network is better. In order to verify and let the error between predicted and actual reflect more clearly. So, in this article, the error curve were outputted by the matlab GA and RBF neural network toolbox.

The error curve is shown in Figure 3.

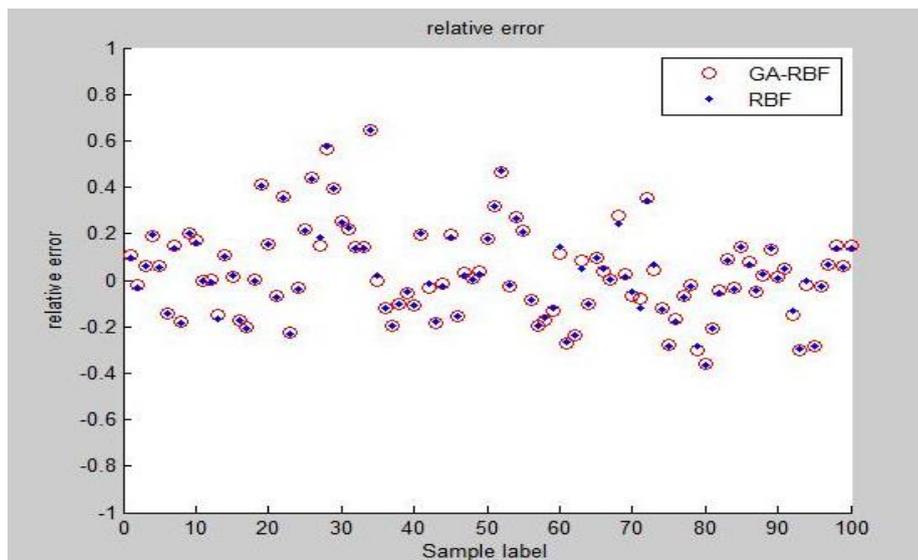


Fig.3. Forecast error results as compared

Through the analysis of the error curve, it is further explained that the GA-RBF fusion algorithm is more accurate than the traditional single RBF neural network algorithm.

(2) Determine prediction accuracy

If there is no mutation operation in genetic algorithm, it will not generate new genes, but if the mutation rate is too large, it will make the genetic algorithm into a random search, select the appropriate mutation rate is one of the key points in the

optimization operation of GA. Therefore, in this article, uses the crossover rate as 0.7 and the mutation rate as 0.02 and 0.05 to do the simulation predictions. 0.02 mutation rate forecast results were shown in Figure 4, 0.05 were shown in Figure 5.

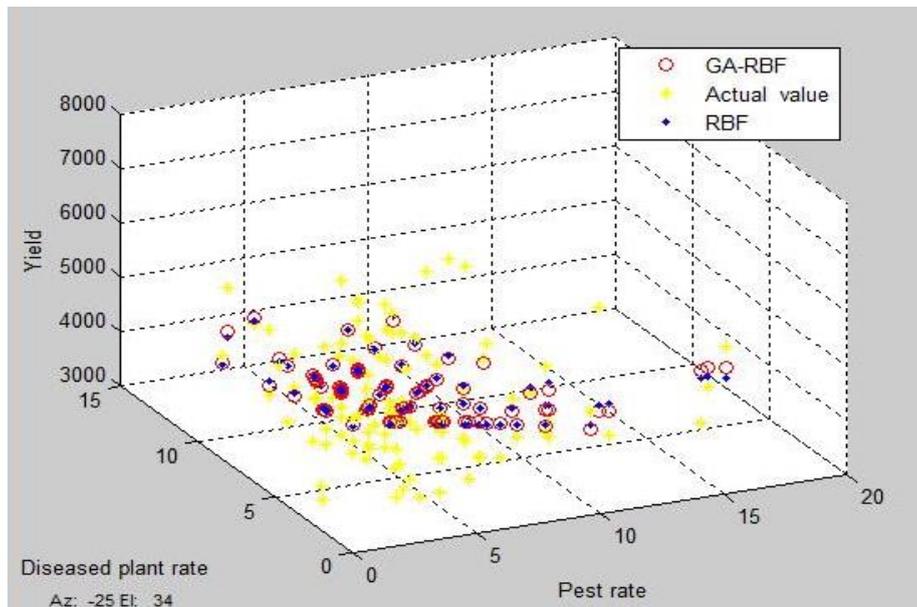


Fig.4. yield loss prediction of maize diseases and pests (Mutation rate 0.02)

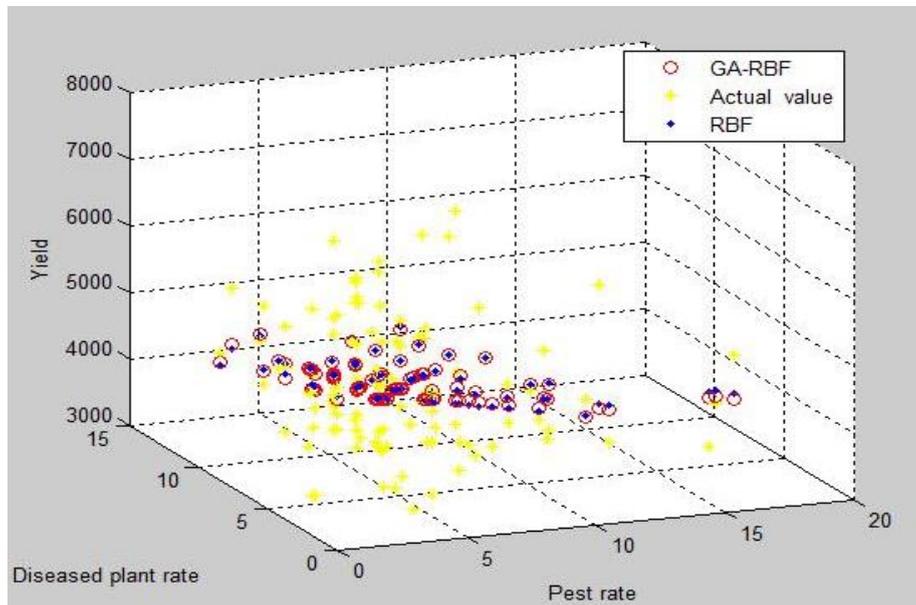


Fig.5. yield loss prediction of maize diseases and pests(Mutation rate 0.05)

Through the simulation calculation,when the mutation rate was 0.05, the prediction accuracy was higher than the 0.02.Therefore,the final prediction was made with the mutation rate of 0.05.

Rrom Figure 4,we can observe that the prediction relative error of GA-RBF fusion algorithm is smaller than that predicted by RBF neural network only,which was used to predict the yield loss of maize diseases and pests,and also the prediction accuracy has been improved.

4 Conclusion

For maize pests of yield loss prediction of a strong complexity and nonlinear characteristics,an optimization algorithm based on GA and RBF neural network fusion is presented,it is used to predict yield loss of maize diseases and pests in this paper.Research indicates:

(1)The RBF model based on GA is optimized, which can predict the yield loss of maize diseases and pests effectively.After the test and analysis of the forecast results,it is proved that the relative error of GA-RBF fusion algorithm in the prediction of yield loss of maize diseases and pests is lower than that of the traditional RBF neural network.

(2)By calculating and analyzing: after using RBF neural network to learning training sample data, the prediction error is 8.8%; And after GA - RBF fusion algorithm of prediction error is 6.4%, compared with before optimization error was reduced by 2.4%.Thus, after the optimization of the parameters of the GA-RBF fusion algorithm, maize yield loss prediction accuracy has been improved.

(3)Through many simulation experiments,when the number of training samples increases, compared with the traditional RBF network, the GA-RBF fusion optimization algorithm is improved in the convergence speed and the accuracy of the neural network approximation,and the global search ability of the network is better.and also achieved good effect in forecasting the yield loss of maize diseases and pests.

Acknowledgment

This work was funded by the China Spark Program.2015GA66004."Integration and demonstration of corn precise operation technology based on Internet of things".

References

1. Fu Baolong.Application Research of RBF Network in Forecasting Agricultural Disease and Insect Pest[J].Journal of Anhui Agricultural Sciences.(36):388-389.(2008)
2. Wang Zhimin, Fang Baoting.A review on theoretical models and development of yield analysis in crop production system[J].Journal of China Agricultural University. 14(1):1~7.(2009)
3. Luo Chengfeng.Estimating urban impervious surface percentage with BP neural network based on genetic algorithm[J].Science of Surveying and Mapping.36(1):48-49.(2011)
4. Liu Xuan,Tang Huiqiang,Xu Xiazhen,Xiang Ying.Prediction of Loss Affected by Pests Based on Neural Network[J]. Journal of Agricultural Mechanization Research21(4):13-16.(2009)
5. Jiang Yuying, Zeng Juan,Lu Minghong.Trend forecast of major crop pests and diseases in China in 2015[J].China Plant Protection.(2): 10-12.(2015)
6. Zhang Fangzhou,Hao Qinghui, Zhou Bo, Liu Qing, Han Dongyang.Application of RBF Neural Network of Genetic Algorithm in Calculation of Line Losses[J].Computer Technology And Development.24(6):192-199.(2014)
7. Nan Qu , Hong Mi , Bin Wang , Yulin Ren.Application of GA-RBF networks to the nondestructive determination of active component in pharmaceutical powder by NIR spectroscopy[J].Journal of the Taiwan Institute of Chemical Engineers 40:162–167. (2009)
8. Liu Jie, Jiang Yuying.Characteristic of Corn Diseases and Pest in 2012 and Cause Analysis[J]. Chinese Agricultural Science Bulletin.30(7):270-279.(2014)
9. Liu Qingjie, Chen Guiming, Liu Xiaofang.Genetic algorithm based SVM parameter composition optimization.[J].Computer Applications and Software. 29(4):94-100.(2012)
10. Shi Lun Zuo.Combining a Radial Basis Function Neural Network with Improved Genetical Gorithm for Vulcanizing Process Parameter Optimization[J].Applied Mechanics and Materials.433-437.(2013)
11. Sun Tiantian,Fan Zuowei, Liu Shuxia, Wu Haiyan.Different Maturity Maize Pest Occurrence in Jilin Province[J].Chinese Agricultural Science Bulletin.30(30):222-227.(2014)
12. Weikuan Jia, Dean Zhao, Ling Ding.An optimized RBF neural network algorithm based on partial least squares and genetic algorithm for classification of small sample[J].Applied soft Computing 48,373-384.(2016)
13. Kuang Ling,Research on China Grain Production Prediction Based on RBF Neural Network[J].Computer Simulation.28(11):189-192.(2011)
14. Ji Siqu,Wu Fang.Static Early-warning Models of Vegetable Diseases Based on Neural Network[J].Computer Knowledge and Technology.12(10):189-196.(2016)
15. Babita Majhi, Minakhi Rout , Vikas Baghel.On the development and performance evaluation of a multiobjective GA-based RBF adaptive model for the prediction of stock indices[J].Journal of King Saud University – Computer and Information Sciences,26, 319–331.(2014)
16. Ren Yanna, Xi Lei, Wang Jiang.Application and Simulation Research on Grain Yield Prediction Model[J].Computer Simulation. 28(4): 208-211.(2011)
17. Tang Zhong,Xie Tao.Application and Simulation of Matlab Neural Network Tool NNTool[J].Computer and Modernization. (12):44-47.(2012)
18. Guresen, Erkam, Kayakutlu Gulgun, Gulgun, Tugrul, U., 2011. Using artificial neural network models in stock market index prediction.Expert Syst. Appl. 38, 10389–10397.(2011)
19. Ling-Jing Kao, Chih-Chou Chiu, Chi-Jie Lu, Jung-Li Yang, 2012.Integration of nonlinear independent component analysis and support vector regression for stock price forecasting.

- Neurocomputing,<http://dx.doi.org/10.1016/j.neucom.2012.06.037>.(2012)
20. Qi Lingzi.Soybean defoliator hazard loss forecast model research[D].Jilin Agricultural University(2013)
 21. Hong Xiang Wang,Wen Xian Guo.Upgrading Water Distribution System Based on GA-RBF Neural Network Model[J].Advanced Materials Research.605-608.(2011)
 22. Zengyu.RBF Natural Network Based on Genetic Algorithm Used in Maximum Power Point Tracking of Photovoltaic System[D].Hunan University of Technology(2015)
 23. Zhang Ji.Grain Yield Combination Prediction Based On Neural Network Optimized By Genetic Algorithm[D].He'nan Normal University(2015).
 24. Zhou Weihua.Optimization Study of the Hidden Structure and Parameters in the RBF Neural Networks[D].East China University of Science and Technology(2013)