

A HYBRID APPROACH OF NEURAL NETWORK WITH PARTICLE SWARM OPTIMIZATION FOR TOBACCO PESTS PREDICTION

Jiake Lv^{1,2}, Xuan Wang^{1,2}, Deti Xie^{1,3}, Chaofu Wei^{1,3*}

¹ Chongqing Key laboratory of digital agriculture, Chongqing, 400716, P.R.China

² College of Computer and Information Science, Southwest University, Chongqing, 400716, P.R.China

³ College of Resources and Environment, Southwest University, Chongqing, 400716, P.R.China

* Corresponding author, Address: College of Resources and Environment, Southwest University, Chongqing, 400716, P.R.China, Tel: +86-23-13983377663, Email: lk@swu.edu.cn

Abstract: Forecasting pests emergence levels plays a significant role in regional crop planting and management. The accuracy, which is derived from the accuracy of the forecasting approach used, will determine the economics of the operation of the pests prediction. Conventional methods including time series, regression analysis or ARMA model entail exogenous input together with a number of assumptions. The use of neural networks has been shown to be a cost-effective technique. But their training, usually with back-propagation algorithm or other gradient algorithms, is featured with some drawbacks such as very slow convergence and easy entrapment in a local minimum. This paper presents a hybrid approach of neural network with particle swarm optimization for developing the accuracy of predictions. The approach is applied to forecast *Alternaria alternate* Keissl emergence level of the WuLong Country, one of the most important tobacco planting areas in Chongqing. Traditional ARMA model and BP neural network are investigated as comparison basis. The experimental results show that the proposed approach can achieve better prediction performance.

Keywords: tobacco pests prediction, particle swarm optimization, neural network, learning algorithm

1. INTRODUCTION

Alternaria alternata Keissl and *Ascochyta gossipii* are major pests in southwest china tobacco planting areas, causing damage to the tobacco by larval stages feeding on flower buds and berries (Wang, 2004). Accurate forecasts of the pests emergence in advance can help determining a timely treatment schedule. In general, there are some *complex and non-linear* relationships between emergence level and several environmental factors like rainfalls, solar radiation, temperature, air humid etc (Reichert and Omlin, 1997). Modeling such relationships with conventional techniques such as time series, regression analysis, ARMA model etc, has been attempted before with variable success (Shaffer and Gold, 1985; Roiditakis and Karandinos, 2001; Satake and Ohgushi et al., 2006). However, these models are based on statistical methods and show some deficiency in the presence of an abrupt change in environmental or sociological variables which are believed to affect pests emergence. In addition, the employed techniques for those models use a large number of complex relationships, require a long computational time, and may result in numerical instabilities, thus not achieving the desired accuracy.

During the past decade, neural networks and in particular, feed forward backward propagation perceptions, were widely applied in different fields (Hippert and Pedreira et al., 2001). The multi-layer perceptions could be trained with non-linear transfer to approximate and accurately generalize virtually any smooth, measurable function while taking no prior assumptions of the data distribution (Russell and Norvig, 2003). Several characteristics, including built-in dynamism in forecasting, data-error tolerance and lack of requirements of any exogenous input, make neural networks attractive for use in pests prediction. Park et.al used artificial neural network to predict Korean pine trees insect pests hazard rating (Park and Chung, 2006). Tourenq et.al adopted neural network to forecast rice crop damage of Camargue,France and the results show better performance than the traditional regression model(Tourenq and Aulaginer et al., 1999). A.Drake developed and used a multi-layer perception ANN to model *Alternaria sp* emergence in Queensland, Australia(Drake, 2001). Although the back propagation (BP) algorithm is commonly used in recent years to perform the training task, some drawbacks are often encountered in the use of this gradient-based method. They include: the training convergence speed is very slow and easy entrapment in a local minimum(Carpinteiro and Otavio et al., 2004).

Particle swarm optimization(PSO), with capability to optimize complex numerical functions, is initially developed as a tool for modeling social behavior(Kennedy, 1997).Unlike other evolutionary algorithms, PSO relies on *cooperation* rather than *competition* and can provide promising solution due to the ability of performing global search for best forecast model(Clerc,

1999; Shi and Eberhart, 1999). Moreover, it is ideally suited for solving discrete and/or combinatorial type optimization problems (Russell and Norvig, 2003). In this paper, a hybrid approach of neural network with particle swarm optimization is developed by adopting PSO to train multi-layer perceptrons. To forecast *Alternaria alternata* Keissldaily emergence level of the WuLong country, one of the most important tobacco planting areas in Chongqing. Comprised with Traditional ARMA model and BP neural network, the results show that the proposed approach can achieve better prediction performance.

This paper is organized as follows. Following the introduction, some preliminaries and notations are described in Section 2. Hybrid approach and detailed procedure used in this study are depicted in Section 3. In Section 4, experimental results and discussion are illustrated. Conclusions are finally made in Section 5.

2. PRELIMINARIES AND NOTATIONS

2.1 Particle swarm optimization

Particle swarm optimization (PSO) is a kind of algorithm to search for the best solution by simulating the movement and flocking of birds (Eberhart and Shi, 2000). The algorithm works by initializing a flock of birds randomly over the searching space, where every bird is called as a “particle” (Shi and Eberhart, 1999). These “particle” fly with a certain velocity vector, based on its momentum and the influence of its best position (p_b) as well as the best position of its neighbors (p_g), and then compute a new position that the “particle” is to fly to. Supporting the dimension for a searching space is D , the total number of particles is n , the position of the i th particle can be expressed as vector $X_i = (x_{i1}, x_{i2}, \dots, x_{iD})$; the best position of the i th particle being searching until now is denoted as $P_{ib} = (p_{i1}, p_{i2}, \dots, p_{iD})$; and the best position of the total particle swarm being searching until now is denoted as vector $V_i = (v_{i1}, v_{i2}, \dots, v_{iD})$. Then the PSO algorithm is described as:

$$v_{id}(t+1) = v_{id}(t) + c_1 * rand() * [p_{id}(t) - x_{id}(t)] + c_2 * rand() * [p_{gd}(t) - x_{id}(t)] \tag{1}$$

$$x_{id}(t+1) = x_{id}(t) + v_{id}(t+1) \quad 1 \leq i \leq n \quad 1 \leq d \leq D \tag{2}$$

Where c_1, c_2 are the acceleration constants with positive values; $rand()$ is a random number between 0 and 1; w is the inertia weight. In addition to

the parameters c_1, c_2 , the implementation of the original algorithm also requires placing a limit on the *velocity* (v_{\max}).

In the literature (Shi and Eberhart, 1998; Shi and Eberhart, 1999), several selection strategies of inertial weight w have been given. Generally, in the beginning stages of algorithm, the inertial weight w should be reduced rapidly, when around optimum, the inertial weight w should be reduced slowly. So in this paper, we adopted the following selection strategy:

$$w = \frac{w_s - w_e}{Max_{iter} - iter} + w_e \quad (3)$$

Where w_s the starting is weight, w_e is the final inertia, Max_{iter} is the maximum number of iterations and $iter$ is current iteration. After adjusting the parameters w and v_{\max} , the PSO can achieve the best search ability.

2.2 Neural networks

A neural network (NN) of arbitrary topology can be represented by a directed graph $N = (V, E, w)$ (Ahn and Cho et al., 2000). V is the set of nodes, which is divided into the subset V_I of input nodes, the subset V_H of hidden nodes and the subset V_O of output nodes. $E \subseteq V \times V$ is the set of connections. Each connection $(j, i) \in E$ is associated with a weight $w_{ij} \in \mathfrak{R}$. For each unit $i \in V$, let us define its “projective” field $P_i = \{j \in V | (i, j) \in E\}$ and its “receptive” field $R_i = \{j \in V | (j, i) \in E\}$. We denote by p_i and r_i the cardinality of P_i and R_i respectively.

Every non-input node $i \in V_H \cup V_O$ receives from its receptive field R_i a net input given by

$$u_i = \sum_{j \in R_i} w_{ij} y_j \quad (4)$$

Where y_j represents the output value of node j , and sends to its prospective field P_i an output equal to

$$y_i = f(u_i) \quad (5)$$

Where f is an arbitrary activation function

Input nodes do no computation: they just transmit an n -dimensional input pattern $x = (x_1, \dots, x_n)$. Thus, the output of input node $h \in V_I$ is the h attribute x_h of the input pattern.

3. HYBRID PSO-NEURAL APPROACH FOR PESTS PREDICTION

3.1 The structure of model

In this paper, the principle of hybrid approach is to utilize particle swarm optimization learning algorithm to train the weights of feed forward neural network, which not only can make use of neural network ability to handle the *non-linear* relationships between emergence level and the factors affecting it directly from historical data but also take advantage of strong global searching ability of PSO. Therefore, a multi-layer perception (MLP) neural network is firstly used. As a three-layer connected feed-forward network has been proved to approach any continuous function, in this study, we use a three-layer feed-forward network for pests' emergence level forecast. It is built as shown in Fig.1.

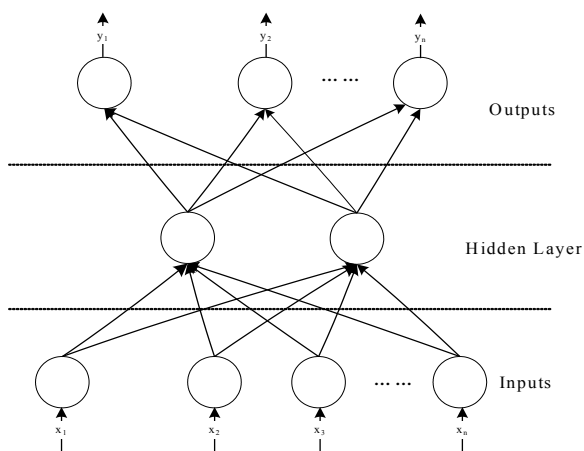


Fig.1. The structure of neural network for pests prediction

In this type of network the first layer is composed of the input variables such as rainfalls, solar radiation, temperature, air humid, and the last layer is composed of the output variables like emergence level of pests. The second layer consists of hidden nodes whose size plays a critical role for neural network *non-linear* mapping. In order to avoid over-fitting and can truly reflect the relationship between emergence level and the factors affecting, we use function in (WANG, 1999) to determine the number of hidden nodes. It is depicted as follows:

$$j = \sqrt{n + m} + a \tag{6}$$

Where j is a hidden node numbers, and n , m represents the number of input nodes and output nodes respectively, $a \in [1,10]$.

3.2 The procedure of hybrid approach

Once the network topology is specified, corresponding weights of each connecting nodes need to be trained. In the study, the total weights of the network will be thought of as the position vector of the particle in the PSO and the weights of each particle vector are initialized randomly and are evolved using the PSO algorithm.

The procedure for the hybrid approach can be summarized as follows:

Step 1: Assign the input variables, output variables and the number of hidden nodes for pests emergence level forecast model

Step 2: Initialize the positions and velocities of a group of particles randomly in the range of $[0,1]$

Step 3: Evaluate each initialized particle's fitness, and P_b is set as the position of the current particles, while P_g is set as the best position of the initialized particles.

Step 4: If the maximal interactive generations are arrived, go to *step 9*, else, go to *step 5*

Step 5: The best particle of the current particles is stored. The positions and velocities of all the particles are updated according to (1) and (2), then a group of new particles are generated, if a new particle flies beyond the boundary $[V_{\min}, V_{\max}]$, the new position will be set as x_{\min} or x_{\max} ; if a new velocity is beyond the boundary $[V_{\min}, V_{\max}]$, the new velocity will be set as V_{\min} or V_{\max}

Step 6: Evaluate each new particle's fitness value, and the worst particle is replaced by the stored best particle. If the i th particle's new position is better than P_{ib} , P_{ib} is set as the new position of the i th particle. If the best position of all new particles is better than P_g , then P_g is updated.

Step 7: Reduce the inertia weights w according to the selection strategy described in section 2.1 of section 2.

Step 8: If the current P_g is unchanged for ten generations, then go to *step 9*; else, go to *step 4*

Step 9: Output the output variables results.

3.3 Performance measure indicators

In order to evaluate the performance of solutions, root mean squared error (RMSE) and the absolute maximum error (MAXIMAL) measurements, in this paper are used as performance measure indicators.

The root mean squared error(RMSE) is expressed as

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N \left(Z_i - \hat{Z}_i \right)^2} \tag{7}$$

The absolute maximum error (MAXIMAL)is expressed as

$$MAXIMAL = \max_{i=1, \dots, N} \left| Z_i - \hat{Z}_i \right| \tag{8}$$

Where Z_i is the actual electrical load and \hat{Z}_i is the forecasted value, N is the number of samples.

4. EXPERIMENTAL RESULTS AND DISCUSSION

Alternaria alternate Keissl is the major insect pests in tobacco planting of southwest china region and emergence levels is greatly influenced by local climate factors. We choose data from WuLong Country Tobacco Corporation in Chongqing, China from May to October of 1996 to that of 2004 as training and testing. All environmental data are normalized into the range between -1 and 1 by using the maximum and minimum values of the variable, and the emergence is classified by 4 level according to china plant pests prediction assessment standard. The whole data set is shown as table 1.

Ten thousand training epochs are adopted as the stopping criteria. The sigmoid function is adopted at the hidden and output nodes. In the PSO training, the number of population is set to 100 while the maximum and minimum velocity value is 0.25 and -0.25 respectively. These values are obtained by trial and error. The w_s and w_e are set 1.0 and 0.4 respectively. By trial and error, it was found that the best values for c_1 and c_2 are 3.0 and 5.0.

For evaluating and comparing the performance of our proposed hybrid PSO-NN approach, traditional BP neural network and statistical autoregressive moving average method for the ARMA model are also utilized as comparison basis. For BP neural network, the learning rate is 0.09 and activation function is sigmoid. Ten thousand training epochs are also adopted as the stopping criteria. All procedures were implemented by using a specific MATLAB toolbox called netlab(Nabney, 2002).ARMA model is built by means of SPSS 14.0.

Fig.2 lists the prediction result using our hybrid PSO-NN approach with true data. BP-Neural network and ARMA model are also compared for the same data.

Table 1. Standardized data set

Year	Month	C ₁	C ₂	C ₃	C ₄	C ₅	D ₁
1996	5	0.12	0.11	-0.16	0.02	-0.19	0001
	6	-0.09	-0.09	-0.00	0.04	0.21	0001
	7	1.00	-0.53	0.36	0.25	0.10	0001
	8	0.43	-0.18	0.80	0.70	0.12	1000
	9	-0.86	0.96	0.90	0.93	-0.17	1000
	10	-0.67	0.28	0.32	0.32	-0.25	0010
1997	5	0.14	-0.06	-0.17	0.24	0.26	0001
	6	0.10	0.03	-0.32	-0.21	0.20	0001
	7	-0.41	-0.34	0.50	0.48	0.12	0001
	8	-0.07	0.12	0.87	0.90	0.15	0010
	9	-0.86	0.68	0.88	0.91	0.16	0010
	10	-0.67	-0.03	0.04	0.11	-0.23	0001
...
2004	5	-0.29	-0.25	-0.14	-0.09	0.17	0001
	6	-0.60	0.81	-0.0	-0.09	0.18	0010
	7	0.07	0.68	1.00	0.95	0.21	0100
	8	-0.39	0.28	0.92	0.86	-0.09	0100
	9	-0.46	0.15	0.01	0.09	0.13	0010
	10	-0.42	-0.56	-1.00	-0.98	0.16	0001

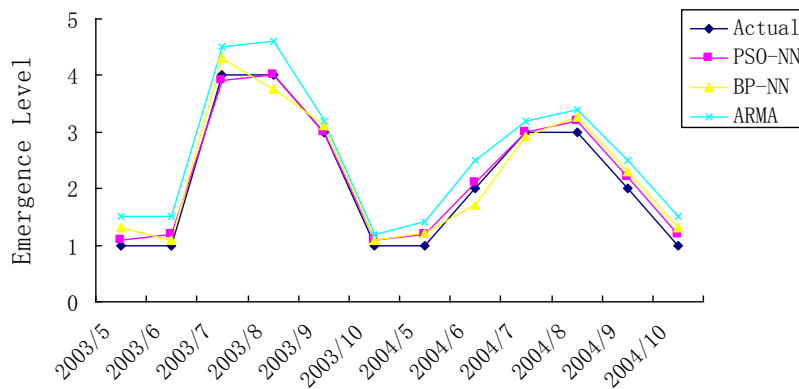


Fig.2 The actual and predictions about the *Alternaria alternata* Keissl emergence level

From Fig.2, it can be worked out that there is only one point at which BP performs a little better than PSO-NN on 2003/6 and none at which ARMA outperforms than PSO-NN. With reference to all individuals, the precision of predictions from PSO-NN is much better when we take the tendency as a whole.

Table 2 shows performance measure indicators values when we apply three different approaches. It can be observed that for both RMSE and MAXIMAL, PSO-NN exhibits better performance in the training process as well as better prediction ability in the validation process than those by BP-NN and ARMA.

Table 2 Performance measure indicators comparison for different forecasting approach

	Training set			Test set		
	PSO-NN	BP-NN	ARMA	PSO-NN	BP-NN	ARMA
RMSE	0.18	0.26	0.41	0.17	0.25	0.42
MAXIMAL	87.41	220.37	237.2	99.76	224.01	267.15

The above results indicate that the hybrid PSO-NN outperforms in pests emergence level forecasting and the prediction is more concise than that of BP-NN and ARMA. It is because PSO can effectively perform global search for better forecast model when it is utilized to train network learning, and avoid easy entrapment in a local minimum.

5. CONCLUSIONS

In this paper, we present a hybrid approach of neural network with practical swarm optimization learning algorithm for pests prediction. The approach is applied to tobacco *Alternaria alternata* Keissl emergence level and two performance measure indicators including RMSE and MAXIMAL are presented and compared. The results show that prediction is more accurate when compared with the traditionally used BP-based perception and ARMA model. Considering the factual circumstance of pests prediction, there are also some issues that need to be further discussed and improved. For example, some parameters like the maximum and minimum velocity values in PSO training initialization are now determined by trial and error, how to design a learning algorithm for the automated parameters selection will be an important issue in future research.

ACKNOWLEDGEMENTS

This paper was supported by the grants from National Science & Technology Pillar program in the Eleventh Five-year Plan Period (No.2006BAD10A01-02) and Chongqing Tobacco Corporation Development Foundation (NO.2006016).

REFERENCES

- Ahn, B. S. and S. S. Cho, et al. (2000). "The integrated methodology of rough set theory and artificial neural network for business failure prediction." *Expert Systems with Applications* 18 (2): 65-74.
- Carpinteiro and A. S. Otavio, et al. (2004). "A hierarchical neural model in short-term load forecasting." *Applied soft computing* 4 (4): 405-412.
- Clerc, M. (1999). The swarm and the queen: towards a deterministic and adaptive particle swarm optimization. *Proceeding Congress on Evolutionary Computation*, Washiongton DC.
- Drake, A. (2001). "Use of remote sensing and ANN in prediction of pests in Queensland." *Remote sensing of environment* 12 (4): 32-35.
- Eberhart, R. C. and Y. Shi (2000). Comparing Inertia Weights and Constriction Factors in Particle swarm optimization. *2000 congress on Evolutionary Computing*.
- Hippert, H. and C. Pedreira, et al. (2001). "Neural networks for short-term load forecasting: a review and evaluation." *IEEE Trans Power System* 16 (1): 45-55.
- Kennedy, J. (1997). The particle swarm: Social adaptation of knowledge. *Proceedings of the 1997 International Conference on Evolutionary Computation*, Indianapolis.
- Nabney, I. T. (2002). *NETLAB: Algorithms for pattern recognition*. London, Springer.
- Park, Y. and Y. Chung (2006). "Hazard rating of pine trees from a forest insect pest using artificial neural networks." *Forest ecology and management* 222 : 222-233.
- Reichert, P. and M. Omlin (1997). "On the usefulness of over parameterized ecological models." *Ecology modeling* 95 : 289-299.
- Roditakis, N. E. and M. G. Karandinons (2001). "Effects of photoperiod and temperature on pupal diapause induction of grape berry moth *lobelia botrana*." *Physiol Entomol* 26 : 329-340.
- Russell, S. and P. Norvig (2003). *Artificial intelligence: A modern approach*, Prentice-Hall International Inc.
- Satake, A. and T. Ohgushi, et al. (2006). "Modeling population dynamics of a tea pest with temperature-dependent development: predicting emergence timing and potential damage." *Ecology research* 21 : 107-116.
- Shaffer, P. L. and H. J. Gold (1985). "A simulation model of population dynamics of the coding moth *cydia pomonella*." *Ecology modeling* 30 : 247-274.
- Shi, Y. and R. C. Eberhart (1998). A modified particle sarm optimizer. *IEEE World Conf on Computation Intelligence*.
- Shi, Y. and R. C. Eberhart (1999). *Empirical study of Particle Swarm Optimization*. IEEE World Conference on Evolutionary Computation.
- Tourenq, C. and S. Aulaginer, et al. (1999). "Use of artificial neural networks for predicting rice crop damage by greater flamingos in the Camargue, France." *120 (2-3): 349-358*.
- WANG, N. (1999). "The research of hybrid optimization strategy in neural networks." *The Journal of Tsinghua University: 66-70*.
- Wang, G. (2004). "Advances and outlook for forecast work of tobacco disease and insect pests in China." *Journal of China tobacco Science* 1 : 44-46.