The Application of BP Networks to Land Suitability Evaluation

LIU Yanfang  JIAO Limin

1 Introduction

Artificial Neural Network (ANN) is a new kind of information transaction and computing system. It is primarily used for some complex problems, such as pattern recognition, evaluation, decision, etc. BP network is a kind of feedforward networks.

Land suitability evaluation is an analyzing process of the land suitability degree for a special use, based on study of land factors (natural, economic and social), aiming at land optimal utilization. The land suitability evaluation needs to select land factors, and to determine the style and degree of influence from each factor; all of these are difficult to do exactly.

Ordinary evaluation methods, such as regression model, empirical model, Fuzzy Rule-based Systems, gray system theory, etc., are too simple or formalized. Such systems will generate remarkable error once an input is beyond the predetermined rule base. The limitations of these methods mainly include:

1) They depend strongly on experience and knowledge. This kind of system has not the self-learning ability and self-adaptability. It is necessary to build rule base or knowledge base.

2) The problem is over-simplified because certain mathematical functions employed as evaluation function are too simple.

3) They cannot be used universally and have not fault tolerance ability. One evaluation system is only for a special region at a special time.

To solve these problems, domestic and international researchers all concentrate on investigating a new method of land evaluation, which should overcome the defect of ordinary methods, and have the self-learning ability, self-adaptability and universal applicability.

In recent years, ANNs develop fast and have been used successfully in pattern recognition, autocontrol, etc. The study of
application in decision-making also makes great progress, such as comprehensive appraisal of urban environment\textsuperscript{11}, analysis of urban anti-disaster ability\textsuperscript{[21]}. The achievements of application show that ANN is not only “easy and practical”, but also “objective and universal”\textsuperscript{[22]}. For appraisal or evaluation, ANN can “discover” and “learn” hidden knowledge and rules from actual samples. After training and test, the network model can be used to evaluate.

2 Application of ANN to the land suitability evaluation

2.1 ANN

ANN is a kind of simulation of the biology neural networks. In fact it is a parallel and distributed information process system constructed by many artificial neurons connected with each other.

The artificial neuron is a kind of simplification or abstraction of the cerebral nerve cell. Each neuron has a single-value output, which connects with many other neurons, and may have many inputs. Each input connection has a weight. Fig. 1 is a simple model of artificial neuron.

As a part of network structure, connection coefficients (weights) usually have no actual meaning, only for computing.

Activation functions in neuron model may have different forms, here are some examples used widely: identity function, bias functions, piecewise linear functions, sigmoid functions, etc. In the neural networks used in land suitability evaluation, linear functions with bias and sigmoid functions are employed. The expression of a usual sigmoid function is shown as follows:

$$f(v) = \frac{1}{1 + \exp(-a \times v)}$$

and the graph is shown in Fig. 2

![Sigmoid function](image)

Fig. 2 Sigmoid function

This is a continuous, bounded, nonlinear function. Parameter $a$ can control the function’s slope. This function is usually used for hidden neurons, and linear functions for output neurons.

An ANN is constructed by many artificial neurons connected with each other. A typical ANN usually contains three layers: input layer, one or more hidden layers, and output layer.

For the land suitability evaluation, each input neuron represents a land factor; the hidden neurons are just for computing, without actual meaning; the output neuron represents the suitability of a parcel aiming at certain use.

There are two kinds of ANNs: feedforward networks and feedback networks. Among these nets, BP networks, error back propagation neural networks, are feedforward, which are researched often, used widely, and known comparatively clearly. In this paper, the ANN used in the land suitability evaluation is a BP network.

2.2 ANN's structures used in the land suitability evaluation

The structure is crucial for the generation ability of a network. It is a common principle that the minimum network is the best if it can be appropriate for samples when there is no other experience. It is equal to using the smoothest function to approximate
an unknown nonlinear map under tolerated error condition.

Because three-layer feedforward BP networks have the capability of approximating continuous maps, they are suitable to be employed in the land suitability evaluation. Input layer and output layer of the neural networks usually have certain actual meaning. Hidden layers are set according to the request of model and the complexity of problem. Input layer and output layer must be first determined, then hidden layer.

In this paper, the evaluation for paddyfield use of Qionghai city, Hainan province, will be used as an example to show how to use BP model in land suitability evaluation.

2.2 Determination of input layer

Land suitability class aims at different land use. The influence of factors on land suitability is different for different land use. In ordinary evaluation methods, factors system of land suitability evaluation, influence factors and weights of each kind of land use, must be determined first. In theory, it is unnecessary to determine factors system when using ANN since ANNs have self-learning ability. ANN will get the influence rule of factors by learning from samples. But, if the same model structure is adopted for each kind of land use, having the same number of input neurons, hidden neurons and output neurons, the model will be more complicated. More time will be spent to train the network. Model accuracy will decrease because of many unnecessary inputs. So influence factors of each kind of land use should be first selected to construct model, unnecessary to determine weights.

In the evaluation for paddyfield use, nine factors are used, which are: soil organic content, soil texture, water conservancy condition, slope grade, plough layer thickness, potential of hydrogen, nitrogen content, phosphor content and kalium content. Therefore, the BP network has nine neurons in input layer, each neuron for a factor.

2.2.2 Determination of output layer

The result of land evaluation is land suitability, so the output of ANN should reflect this result. Output layers are different according to different design ideas. The relation between expense and yield determines Land suitability. Profit, yield subtracted by expense, is used to represent land suitability. In some region, for the same land utilization, the expense is almost the same, so yield can be used to represent suitability.

Usually, we use a real number that varies continuously within a certain range (such as 0-1, standardization of profit or yield) to represent land suitability. So one neuron is enough for output layer.

2.2.3 Determination of hidden layer

The hidden neurons are only for computing, whose number is not close restraint. Increasing the number of hidden neurons means increasing the dimensions of optimize surface, the network will distinguish more kinds of samples, but computational complexity and computer memory consumption will also increase. Although the network error on training set maybe decreases along with more and more training, error on test set maybe increases, and generalization ability droops. According to relevant references and our experiments, considering accuracy and efficiency, the number of hidden neurons should satisfy the following two formula simultaneously:

\[ 2^m > n \]

\[ m = \sqrt{w + n + R(10)} \]

where \( m \) is the number of hidden neurons; \( n \) is the number of input neurons; \( w \) is the number of output neurons; and \( R(10) \) represents an integer from 0 to 10.

In practice, the most appropriate size of hidden layer can be determined by trying. The most optimal number of hidden neurons of the BP network used in evaluation for paddyfield use is 10 (given in Section 3.4 of this paper).

According to these conclusions, the structure of BP network used in evaluation for paddyfield use is shown in Fig. 3, in which each circle represents an artificial neuron.

Each neuron except input neurons has the structure shown in Fig. 1. Input neurons do not make any change on input information. There are no connections between neurons in the same layer, and hetero-layer neurons connect forward. Sigmoid function is employed in hidden layer, and linear function in output layer. It has been proved that this kind of networks can approximate any map from...
\( n \)-dimensional spaces to \( \omega \)-dimensional spaces under the condition of sufficient samples \((n\) represents the number of input neurons, and \( \omega \) represents the number of output neurons\(\textsuperscript{[3]} \)).

2.3 Learning algorithm of the BP networks

BP algorithm includes forward propagation and back propagation. In the forward propagation, the input information was processed layer by layer and passed to the output layer. If the output layer cannot generate the expectative output, then the back propagation begins. The error between output and expectation will be passed back by original connections. By modifying each connection and bias, the output error will decrease, and then the forward propagation begins again. Reiteration will continue until the output error is less than the given value.

For the BP network used in land evaluation (in Fig. 3), there are some points to be paid more attention to in the algorithm.

First, hidden neurons are sigmoid and output neurons are linear. The output of hidden neuron is given by

\[
y_j = \frac{1}{1 + \exp(-sy_i + q_1j)}
\]

where \( q_1j \) is the bias of hidden neuron \( j \). The output of output neuron, i.e. land suitability, is given by

\[
z_k = s2_k + q2_k
\]

where \( q2_k \) is the bias.

Second, momentum item is used to accelerate convergence. Therefore, corrections of the weight and bias of output layer are given by

\[
W_{2p+1}^p = W_{2p}^p + \Delta W_{2p}^p + \mu \times \Delta W_{2p-1}^p
\]

\[
q_{2p+1}^p = q_{2p}^p + \alpha \times e_p + \mu \times (q_{2p}^p - q_{2p-1}^p)
\]

where

\[
e_p = z_{0p} - z_p
\]

\[
\Delta W_{2p}^p = \alpha \times e_p \times y_j
\]

and \( e_p \) is the error of output neuron; \( \Delta W_{2p}^p \) is the weight correction of output neuron of sample \( p \); \( \Delta W_{2p-1}^p \) is the weight correction of output neuron of sample \( p - 1 \); \( W_{2p-1}^p \) is the weight of sample \( p + 1 \); \( q_{2p}^p \) is the bias of sample \( p - 1 \); \( q_{2p}^p \) is the bias of sample \( p \); \( q_{2p+1}^p \) is the bias of sample \( p + 1 \); \( \alpha \) is the learning rate; \( \mu \) is a positive real number less than \( 1 \), which is called momentum coefficient. Like output layer, calculate corrections of the weight and bias of hidden layer as follows:

\[
W_{1p}^p = \alpha \times e_j \times x_i
\]

\[
e_j = y_j(1 - y_j) \times (\sum_{k=1}^{w} W_{2k}^p \times e_k)
\]

\[
W_{1p+1}^p = W_{1p}^p + \Delta W_{1p}^p + \mu \times \Delta W_{1p-1}^p
\]

\[
q_{1p+1}^p = q_{1p}^p + \alpha \times e_j + \mu \times (q_{1p}^p - q_{1p-1}^p)
\]

where \( \Delta W_{1p}^p \) is the weight correction of hidden neuron of sample \( p \); \( \Delta W_{1p-1}^p \) is the weight correction of hidden neuron of sample \( p - 1 \); \( W_{1p-1}^p \) is the weight of sample \( p + 1 \); \( q_{1p}^p \) is the bias of sample \( p - 1 \); \( q_{1p}^p \) is the bias of sample \( p \); \( q_{1p+1}^p \) is the bias of sample \( p + 1 \); \( \alpha \) and \( \mu \) have the same meaning mentioned above.

Finally, the mean square error of all samples in the training set is given by

\[
R = \frac{1}{L} \sum_{p=1}^{L} \sum_{i=1}^{w} (Z_{0i} - Z_{i})^2
\]

where \( L \) is the number of the samples in the training set. Learning is completed when \( R \) is less than the given value.
3 Case study

3.1 Summary

Qionghai city lies in the eastern part of Hainan province. The total area is 1710.3 km².

The partition of land evaluation units is mainly based on land utilization type, and the total number is more than 30,000.

Land suitability evaluation in Qionghai city is multi-aiming evaluation. This evaluation aims at seven kinds of land utilization, which are: paddy field, dry farm, forest, grassland, aquiculture, commercial herbage, and commercial woody crop.

3.2 Data process

11 factors were collected and processed in the land evaluation in Qionghai city, such as soil organic content, soil texture, slope grade, water conservancy condition, etc. For the convenience of calculation, qualitative data are quantified first, and then all factors data were standardized as a real number in 0-1. Two examples are provided to show the standardization. One is soil organic content, a quantitative factor. The other is water conservancy condition, a qualitative factor.

Statistics of soil organic content state that the maximum value is 4.5%. So the standard value is given by

\[ x = \frac{x_0}{0.045} \]

where \( x_0 \) represents the original value.

Water conservancy condition is a qualitative factor, indicated by “best”, “good”, “ordinary”, and “poor”. The standardization results are listed in Table 1.

| Qualitative factor | Best | Good | Ordinary | Poor |
|--------------------|------|------|----------|------|
| Standardization    | 0.8  | 0.6  | 0.4      | 0.2  |

Factor values of each evaluation unit were got by spatial analysis using Arc/Info. Factors were selected from different uses. Then multi-purpose evaluation was done using a special method.

Data processing of land suitability evaluation using BP networks is shown in Fig. 4.
3.3 Sample data

ANN must "learn" from sufficient samples first, and then get necessary knowledge. These samples are called training set. Accordingly, test set is also required to test the reliability of the network trained. Since ANN learns from samples, the quality of samples is crucial in practice. The samples must be representative and exact.

In consideration of samples' representation, all training and test samples were collected mainly by random, under the condition of rough balance of local distribution, utilization, and suitability. Since factors of each parcel had been got by spatial analysis, just use and corresponding suitability degree were required in survey. In practice, land suitability can be represented by a real number between 0-1, which is the standardization of yield or economic benefit of crop. For example, if the maximum yield of a kind of crop in the evaluation area is about 10 000 kg/hm², the yield of a parcel is about 8 000 kg/hm², then the suitability of this parcel can be quantified as 0.8.

3.4 Application of networks

The network in evaluation for paddy field use has the structure shown in Fig. 3, nine input neurons, and one output neuron. The number of training samples is 400 and the number of test samples is 210. To determine the most appropriate number of hidden neurons, the authors tried many times in a predetermined range. Some experiments were shown in Table 2.

After the comparison, the number of hidden neurons was set to 10. When the network has 14 neurons, the error on training set reduces, but that on test set increases. This indicates that the network has too many dimensions and generation ability drops.

| Number of hidden neurons | Training times when MSE<0.001 | MSE when converge/10⁻⁴ | Error on test set/10⁻⁴ |
|--------------------------|-------------------------------|------------------------|-----------------------|
| 6                        | 381                           | 4.9                    | 5.6                   |
| 8                        | 367                           | 4.1                    | 4.9                   |
| 10                       | 295                           | 3.2                    | 3.9                   |
| 12                       | 311                           | 3.2                    | 4.2                   |
| 14                       | 365                           | 2.5                    | 4.3                   |

The network trained successfully was tested on the test set, and the accuracy rate is 93.4%. 14 samples did not get the expectative results. It can be seen from test result that the application of networks is successful. Because some special value-combinations seldom appear in training set, some parcels did not get expectations. These special samples can be added to training set, and then retrain the network to improve recognition ability. The network learns new knowledge and its performance has been improved.

3.5 Comparison between methods

Fuzzy comprehensive assessment is a relatively famous and perfect method in land suitability evaluation. Therefore, based on the factors system determined by experts' argumentation, fuzzy comprehensive assessment method is also employed in this land evaluation. The part of the factors system about paddy field is shown in Table 3.

| Table 3 Factors system of the evaluation for paddy field use |
|-----------------|-----------------|-----------------|-----------------|----------------|
| Factor          | Weight          | I               | II              | III             | IV              |
| Organic content/% | 0.2             | >3              | 3 - 2           | 2 - 1           | <1              |
| Soil texture    | 0.2             | Loamy soil      | Light soil      | Silty soil      | Sand, clay soil |
| Water conservancy condition | 0.25 | Best | Good | Ordinary | Poor |
| Plough layer thickness/cm | 0.1 | >20 | 20 - 15 | 15 - 10 | <10 |
| slope grade (/') | 0.1             | <1              | 1 - 2           | 2 - 4           | >4              |
| Nitrogen content | 0.04            | >0.15           | 0.1 - 0.15     | 0.1 - 0.05     | <0.05           |
| Phosphor content | 0.03            | >0.18           | 0.18 - 0.09    | 0.09 - 0.04    | <0.04           |
| Potassium content | 0.03            | >1.8            | 1.8 - 0.9      | 0.9 - 0.6      | <0.6            |
| pH              | 0.05            | <5.5            | 5.5 - 7.0      | 7.0 - 8.2      | >8.2            |
Table 4 Comparison of ANN and fuzzy comprehensive assessment

| Need to determine weights? | ANN   | fuzzy comprehensive assessment |
|----------------------------|-------|-------------------------------|
| Need samples?             | Yes   | No                            |
| Evaluation rule           | Learning from samples | Determining artificially |
| Accuracy on test set      | 93.4% | 90.6% |

The accuracy of fuzzy comprehensive assessment method on test samples above is 90.6%, a little poorer than that of ANN. Factors influencing the accuracy of fuzzy method include and factors system, membership function, etc.

Compared with fuzzy comprehensive assessment on the spatial distribution of evaluation result, ANN can always give more reasonable evaluation result, which inoculates with factors better, specially in regions where factors change obviously, such as watershed, etc. By learning from field surveying samples, ANN can reflect the influence of factors on land suitability more exactly. Factors system, required by ordinary methods, whose grading rules and weights cannot match the fact accurately, influences the accuracy of evaluation result remarkably.

4 Conclusion

The application of ANN in land suitability evaluation is successful. Compared with ordinary methods, ANNs have many advantages such as:

1) Weights of factors are not required, so human influence is lessened and the result is more objective.

2) Unnecessary to determine evaluation function, the error generated by the unconformity between the model and the fact decreases, and the evaluation system becomes universal.

3) In practice, when the situation changes, collect new samples and retrain the network. For different regions or different times in the same area, the only need is to collect or update samples.

Some problems remains to be solved, e.g. interpretation of neural networks. Generation ability of neural networks has been proved, and application in land evaluation is also successful, but how to exactly interpret the meaning represented by the computing process of networks is a difficult problem. In addition, some other problems, such as how to combine neural networks with ordinary methods, importing fuzzy theory into neural networks, should be studied further. These researches will improve ANNs.

References

1. Li Z Y (1995) B-P network model of city environment quality. *Systems Engineering*. (in Chinese)
2. Liu Y L (1996) Study on the model of B-P network of 2-levels for urban anti-disaster ability. *Journal of Wuhan Technical university of Surveying and Mapping*, 21 (Suppl.): 20-26 (in Chinese)
3. Zhang N Y, Yan P F (1998) Neural networks and fuzzy control. *Beijing: Tsinghua University Press*. (in Chinese)
4. Liu Z L, Liu Y C (1996) Fuzzy logic and neural networks. *Beijing: Beijing University of Aviation and Aerospace Press*. (in Chinese)
5. Sun Z Q (1997) Theory and technique of intelligent control. *Beijing: Tsinghua University Press*. (in Chinese)
6. Ni S X (1999) Land types and land evaluation. *Beijing: Higher Education Press*. (in Chinese)
7. Wang L, Zheng D Z (2000) Training algorithm for FNN based on a class of nonlinear property. *Control and Decision*. (in Chinese)
8. Castellano G, Maria A, Pelillo M (1977) An iterative pruning algorithm for feedforward NN. *IEEE, Trans. NN*, 8: 519-537
9. Benitez J. M., Castro J. L., Requena I. (1997) Are artificial neural networks black box? *IEEE, Trans. NN*, 8:1: 156-1 164
10. Cho Sungzoon, Cho Yongjung, Yoon Sungchul (1997) Reliable roll force prediction in cold mill using multiple neural networks. *IEEE, Trans. NN*, 8:1: 874-882