The Research for the Evaluation of Cultivated Land Quality Based on Deep Belief Networks

Qiao-bing YUE\textsuperscript{1,2} and Xu-bing ZHANG\textsuperscript{1,2,*}

\textsuperscript{1}School of Public Administration, China University of Geosciences (Wuhan), Hubei, China, 430074
\textsuperscript{2}Geographic Environment and National Park Laboratory, China University of Geosciences (Wuhan), Hubei, China, 430074

*Corresponding author

Keywords: Cultivated land quality, Grading evaluation, Deep belief networks.

Abstract. Traditional evaluation methods of cultivated land quality are mainly on the basis of empirical judgments in the process of weight calculation and membership determination. In this paper, taking Enshi city as an example, we attempt to employ Deep Belief Networks (DBN) to estimate the cultivated land quality. We selected 8 evaluation indexes as the input data and the grade of cultivated land as the output data to construct an evaluation model and finally realized the classification of cultivated land quality grade. Compared with the results of the Supplement and Improvement of Farmland's Classification in Enshi city 2013 in quantity and spatial distribution, our experimental results show that accuracy of grade 11, grade 12 and grade 13 in quantity reach as high as 99.06\%, 90.04\% and 96.07\%, respectively. Besides, the spatial distribution is substantially in agreement with real distribution.

Introduction

Cultivated land is referred to as the most fundamental element of agricultural production, and therefore the objective evaluation of its quality plays an important role both in local agricultural development and economic development\textsuperscript{[1,2]}. With the rapid development of industrialization and urbanization, cultivated land in China has been declining year by year both in quantity and quality\textsuperscript{[3,4]}. Consequently, it is by no means a trivial task for us to propose a method to make a reasonable assessment of cultivated land quality. The traditional methods of evaluating farmland quality, such as factor comprehensive evaluation method, fuzzy comprehensive judgment method\textsuperscript{[5]} and principal component analysis (PCA), have become more sophisticated in recent years. However, most of these methods are generally susceptible to the interference of subjectivity of the evaluators. Since then, some scholars also apply more intelligent and informative methods, such as Back Propagation (BP)\textsuperscript{[6-8]}, to estimate cultivated land quality. Nevertheless, these methods cannot address the problem of nonlinear between evaluation factors and quality grades very well.

Deep learning, as a popular way of recognition and classification, has been widely used in image classification, speech recognition and language processing in recent years. Inspired by this, in this paper we attempt to extend the utilization of DBN to the evaluation of cultivated land quality. By considering fully the information of soil, natural environment and infrastructures, we make evaluation of cultivated land quality in Enshi city based on DBN. It should be indicated that we learn the weights of indicators automatically by the model, not any manual determination. Therefore, the evaluation results are more objective and reasonable.

Deep Belief Networks (DBN)

DBN is a probabilistic, generative neural network model\textsuperscript{[9]} proposed by Hinton in 2006. It is mainly composed of the multilayer unsupervised learning networks, which is called as Restricted Boltzmann Machine (RBM), and a supervised leaning network, namely Back Propagation\textsuperscript{[10,11]}. Each RBM consists of two layers, the hidden layer and the visible layer\textsuperscript{[12]}. The layer of hidden units are not connected to each other and have undirected, symmetrical connections to the layer of visible units (shown as Figure 1).
A joint configuration of the units has an energy given by

$$E(v,h|\theta) = -\sum_{i} a_i v_i - \sum_{j} b_j h_j - \sum_{i} \sum_{j} v_i w_{ij} h_j.$$  \hspace{1cm} (1)

Where $\theta = \{a_i, b_j, w_{ij}\}$ respects all parameters, $w_{ij}$ is the weight between connected visible unit $i$ and hidden unit $j$; $a_i$ and $b_j$ are bias terms of visible and hidden unit, respectively. $m$ and $n$ are the total number of visible units and hidden units. Due to the conditional independence between visible units and hidden units, we could obtain the conditional distribution of hidden units by activation function $\sigma$ while visible units are given.

$$P(h_j=1|v,\theta) = \sigma(b_j + \sum v_i w_{ij}).$$  \hspace{1cm} (2)

We also get the conditional distributions of visible units in the same way.

$$P(v_i=1|h,\theta) = \sigma(a_i + \sum h_j w_{ij}).$$  \hspace{1cm} (3)

Where $\sigma$ respects the tanh function. Using the Contrastive Divergence \cite{13} (CD) algorithm to quickly calculate the model’s expectation, this algorithm can estimate the model by iterating only a few times. And the rules of the updating parameters are given by as the following,

$$\Delta w_{ij} = \epsilon (v_i h_j >_{\text{data}} - v_i h_j >_{\text{recon}}).$$  \hspace{1cm} (4)

$$\Delta a_i = \epsilon (v_i >_{\text{data}} - v_i >_{\text{recon}}).$$  \hspace{1cm} (5)

$$\Delta b_j = \epsilon (h_j >_{\text{data}} - h_j >_{\text{recon}}).$$  \hspace{1cm} (6)

Where $\epsilon$ is the learning rate. Through the learning procedure, we can obtain the proper value of parameters. In additional, a multilayer DBN can be built by stacked multiple layers of RBMs trained in a layer by layer manner.

Experiments and Analysis

Determination of Study Area and Evaluating Units

We take Enshi city as the study area, which is located in the southwestern corner of Hubei province, with a geographic region of 29°50’33”~30°39’30” N and 109°4’48”~109°58’42” E. The city has a cultivated land area of 943.6km$^2$, occupying 23.75% of its total area, mainly including paddy field and dry land. There are 73909 patches in Enshi city, viewed as evaluating units according to the Cultivated Land Quality Updating Database of Enshi 2013.

Construction of Evaluation Index System

The impacts of cultivated land quality are different in various regions. In this paper, according to the practical experience, the evaluation index system (shown as Table 1) is consisted of 3 layers: target level, criteria level and index level. Target level is the quality grade of cultivated land; in the index level, we select 8 evaluation indicators, by considering fully the information of soil, natural environment and infrastructure, to describe the situation of criteria level. Furthermore, considering the diversities in evaluation indicator, for example, the pH value is always a quantitative data but
the texture is a qualitative data. We assign reasonable score to each evaluating indicator between 0 to 100 according to Table 2. In additional, we also make a normalization of each score as Eq.7.

\[ y = y_{\text{min}} + \frac{y_{\text{max}} - y_{\text{min}}}{x_{\text{max}} - x_{\text{min}}} \times (x - x_{\text{min}}). \]  

(7)

Table 1. Evaluation index system of cultivated land natural quality in Enshi city.

| Target level | Criteria level | Index level |
|--------------|----------------|-------------|
| Quality grade of cultivated land in Enshi city | Soil | Textural | PH Value | Thickness | Organic Matters | Rock Outcrop Degree | Soil Erosion |
| Natural environment | | | | | | |
| Infrastructure | | | | | | Irrigating Guarantee Rate |

Table 2. Evaluation information of cultivated land quality in Enshi city.

| Factors | Grade (Description) | Score |
|---------|---------------------|-------|
| Texture | R                   | 100   |
|         | N                   | 80    |
|         | S                   | 70    |
|         | L                   | 50    |
| Organic Matters [g/kg] | >=25 | 100 |
|         | 15-25              | 90    |
|         | 10-15              | 70    |
|         | <10                | 50    |
| pH Value | 6.5-7.0            | 100   |
|         | 6.0-6.5,7.0-7.5   | 90    |
|         | 5.0-6.0,7.5-8.0   | 70    |
|         | <5,>=8            | 50    |
| Irrigating Guarantee Rate | Level1 | 100 |
|         | Level 2           | 90    |
|         | Level 3           | 80    |
|         | Level 4           | 50    |
| Thickness | >=100            | 100   |
|         | 60-100            | 60    |
|         | <60               | 40    |
| Slope | T1                 | 100   |
|         | T2-T5             | 40    |
| Rock Outcrop Degree | 1             | 100   |
|         | 2                  | 70    |
|         | 3-5               | 40    |
| Soil Erosion | 1               | 100   |
|         | 2                  | 90    |
|         | 3                  | 80    |
|         | 4                  | 60    |
|         | 5                  | 50    |

Model Evaluation

The DBN model aims to provide more scientific basis for arable land quality evaluation. Shown as Figure 2 is the flowchart of arable land quality evaluation. The model contains three layers: input layer, hidden layer and output layer. After data processing, we take the evaluating indicators as input and the quality grade as output. We assume that hidden layers have the same number of units for simplicity, which is set up to 8, 10, 15 and 20. The network depth is set up to 2, 3, 4 similarly. Through lots of experiments, we can receive a promising result under the condition that the number of network depth and units number are set to 2 and 20, respectively. Other parameters are experimentally determined as follows: learning rate =0.01, training epoch =1,000; We divide the evaluating units into training data and test data by the means of random sampling, and ensure the uniform distribution of training data in spatial distribution. The number of training data and test data account for 10% and 90% of total patch number, respectively. In what follows, we have been
Training the model with large label samples till the precision of evaluation meets the presupposition. And then it is carried out to calculate the grade and make effective prediction of test samples.

Figure 2. Evaluation process of cultivated land quality based on DBN.

Comparative Analysis of Evaluation Results

Three grades are acquired by the evaluation models of DBN, in which the grade 11 is the best, the grade 12 is the middle and the grade 13 is the worst. In order to testify the effectiveness of the proposed method, the classification results based on DBN are compared with the results of the Supplement and Improvement of Farmland's Classification in Enshi city 2013 in the following two aspects of quantity and spatial distribution.

The Quantitative Comparison

As is shown in Table 3, the patch number of grade 11, grade 12 and grade 13 classified correctly are 4538, 32263, 24643 and corresponding accuracy rates are 99.06%, 90.04% and 96.07%, respectively.

The Spatial Distribution

Compared with the results of the Supplement and Improvement of Farmland's Classification in Enshi city 2013, the areas of grade 11 and 13 based on DBN are less 8.68 km², 24.04 km², and the area of grade 12 is more 32.86 km², respectively. And their areas are 88.08 km², 463.82 km² and 391.69 km². The comparison is shown as Figure 3.

Table 3. Total patch number of each grade summarized by DBN and traditional method.

| Grade | Traditional method | DBN method | Number of correct classifications | Accuracy (%) |
|-------|-------------------|------------|----------------------------------|--------------|
| Grade 11 | 4581 | 4550 | 4538 | 99.06 |
| Grade 12 | 35677 | 35779 | 32263 | 90.04 |
| Grade 13 | 25651 | 25580 | 24643 | 96.07 |

Note: Results of traditional method was derived from Supplement and Improvement of Farmland's Classification in Enshi city 2013, the same below.

Table 4. Total area of each grade summarized by DBN method and traditional method.

| Grade | Area (km²) | Area Difference(km²) |
|-------|------------|----------------------|
| Grade 11 | 79.43 | -8.68 |
| Grade 12 | 495.39 | 32.86 |
| Grade 13 | 368.78 | -24.04 |
| Total area | 943.6 | 0 |
Conclusion

(1) The evaluation model proposed here with a nonlinearity structure could avoid error caused by sensitive factors and make an enhancement of evaluation accuracy. In this paper, we attempt to employ DBN to make an empirical research on the quality evaluation of cultivated land in Enshi city. We carried out mainly two phases in the progress of evaluation. Firstly, we also have been realized the grade and normalization of evaluating indicators, in addition to export data from the cultivated land quality updating database of Enshi 2013 in the data processing. Then, we constructed evaluation model based on DBN, and make a successful attempt on the assessment of cultivated land quality. Furthermore, the proposed model is demonstrated to be effective for evaluation of farmland quality.

(2) In this study, the natural quality grade of cultivated land is divided into grade 11, grade 12, grade 13. The grade 12 has the largest area with 462.53 km$^2$, accounting for 49.02% of all cultivated land, distributed in most areas of Enshi city; the proportion of grade 11 is at least with an area of 88.11 km$^2$, only accounting for 9.34%, and mainly distributed in the southeastern of Enshi city; The grade 13 has an area of 392.92 km$^2$ with a ratio of 41.63%, and distributed in the northwest and northeastern regions of the city.

Acknowledgement

This work was financially supported by the Natural Science Foundation of Hubei Province, China (No.2015CFB433).

References

[1] Yu Zhenguo, Hu Xiaopin. Research on the relation of food security and cultivated land's quantity and quality in China, J. Geography and Geo-Information Science, 2003, 19(3):45-49.

[2] Fu Guozhen, Bai Wanqi. Advanced and prospects of evaluation cultivated land quality, J. Resource Science, 2015, 37(2):0226-0236.

[3] Zhang Chunlai, Li Xinhua. How to carry out farmland classification renewal evaluation, J. Territory and Natural Resources Study, 2015(2):52-53.

[4] Zhu Jinxia, Xu Baogen, BaoHaijun. Methods for assessing cultivated land quality at the county-level: A case study of Songyang County, Zhejiang province, J. Shanghai Land and Resources, 2015, 36(1):31-35.
[5] Pan Runqiu, Liu Jun, Song Danyu. Agricultural land classification based on fuzzy comprehensive analysis, J. Transaction of the Chinese Society of Agriculture Engineering, 2014, 30(18):257-265.

[6] Yuan Man, Liu Yaolin, Wei Xiaojian. Research on land use zoning based on fuzzy self-organizing map, J. Scientific and Technological Management of Land and Resources, 2013, 30(6):74-80.

[7] Qie Ruiqing, Guan Xia, Yan Xujiu, et al. Method and its application of natural quality evaluation of cultivated land based on self-organizing feature map neural network, J. Transaction of the Chinese Society of Agriculture Engineering, 2014, 30(23):298-305.

[8] Li Xiaogang, Ma Youhua, Zhang Yi, et al. Comparison of fuzzy evaluation and BP neural networks applied for cultivated land quality evaluation - a case study at FengTai county, J. Chinese Journal of Soil Science, 2015, 46(4):816-822.

[9] G. E. Hinton, S. Osindero. A fast learning algorithm for deep belief nets, J. Neural Computation, 2006, 18(7):1527-1554.

[10] G. E. Hinton, R. R. Salakhutdinov. Reducing the dimensionality of data with neural networks, J. Science, 2006, 313(5786): 504-507.

[11] Lv Qi, Dou Yong, Niu Xin, et al. Remote sensing image classification based on DBN model, J. Journal of Computer Research and Development, 2014, 51(9):1911-1918.

[12] Bengio Y. Learning deep architectures for AI, J. Foundations and Trends in Machine Learning, 2009, 2(1):1-127.

[13] G. E. Hinton. Training Product of Experts by Minimizing Contrastive Divergence, J. Neural Computation, 2002, 14(8):1771-1800.