RESEARCH ON VEGETATION DYNAMIC CHANGE SIMULATION BASED ON SPATIAL DATA MINING OF ANN-CA MODEL USING TIME SERIES OF REMOTE SENSING IMAGES

Zhenyu Cai [1], Xiaohua Wang [1]

1 School of Economics and management, Hebei University of Engineering, Handan,Hebei, 056038,China

Abstract: Dynamic change of vegetation has become a very sensitive problem in China due to climate variability and human's disturbances in the Yellow river basin. Dynamic simulation and forecast of vegetation are regarded as an effective measure to decision support for local government. This paper presents a new method to support the local government's effort in ecological protection. In integrates cellular automata (CA) -artificial neural network (ANN) model with Geographical information system (GIS) and remote sensing. The proposed method includes three major steps: (1) to extract control factors; (2) to integrate CA and ANN models; (3) to simulate the selected area using CA-ANN model. The results indicted that the integrated approach can rapidly find condition of future vegetation cover that satisfy requirement of local relative department. It has demonstrated that the proposed method can provide valuable decision support for local government. the result indicts that NDVI of the vegetation has an increasing trend and   characteristics of distribution concentration trend, but the change rate is become lower from the year 2007 to 2014 compared with the changes from the year 2000 to 2007.

Keywords: CA-ANN model;   control factors;   Yellow river basin;   vegetation changes

1.   INTRODUCTION

Vegetation is the most significant natural resources, however, Yellow river basin, especially in arid and semi-arid regions, is facing various changes caused by human's activities and climate changes. These changes could have a significant impact on the carbon cycle, regional economy and

climate (Le Houerou, 1996; Angell and Mc Claran, 2001). The results of the eco-environmental vulnerability and its changes in the Yellow River Basin showed that the ecology become worse in some regions in the year from 1990 to 2000. There were many factors affecting the eco-environmental vulnerability, such as structural geology, meteorological conditions, sediment discharge, human activity, and forest degradation. The driving forces of the eco-environmental vulnerability changes were the rapid population growth, governmental policy, and vegetation degradation. Grassland degeneration, sandy desertation, and soil erosion had seriously affected the sustainable development of eco-environment. Study on changes simulation of future vegetation is undoubtedly helpful for resource management and environmental conservation.

At present, study on future vegetation changes is often based on statistical analysis data of several points, investigation and experiences of the planner (Xu et al., 2000). It is virtually impossible to effectively manage and monitor the resources using field investigation methods. In the past two decades, satellite remotely sensed data have been widely used for vegetation investigation because of their advantages of frequent revisit, global coverage and lower cost (Langley et al., 2001).

A cellular automata (CA) model is a dynamics model using local interactions to simulate the evolution of a large system, where time are considered as discrete units and space is represented as a regular lattice of two or more dimensions. A CA model is capable of representing those non-linear, spatial and stochastic processes (Wolfram, 1984). More recently, there are growing literatures on the applications of CA models in urban growth, land use change and vegetation dynamics (Balzter et al., 1998; Batty et al., 1999; White and Engelen, 2000). Li and Yeh (2000) demonstrated the potential of integrating a CA model and remote sensing techniques for agricultural protection area zoning and land use planning. About neural network, its an important characteristic is its capability to learn from the data being processed. The network weights are adjusted in the training process, which can be executed through a number of learning algorithms based on backpropagation learning (Ripley, 1996; Haykin, 1999; Zhou, 1999; Lee et al., 2004; Gomez and Kavzoglu, 2005; Yesilnacar and Topal, 2005). Because of complexity of underlying surface of yellow river basin,

Research on vegetation dynamic change simulation based on spatial data mining of ANN-CA model using time series of remote sensing images

3

tempro-spatial change simulation of vegetation using CA-ANN model according to spatial advantage of CA and learning ability of neural network.

2. METHODS

2.1 The study area

The study areas selected for this research are located individually in the up course, middle course and low course of Yellow river basin, and which mainly include vegetation of various growth condition and sorts, such as farm field, grassland, forest. The approximate latitudinal and longitudinal ranges of the study areas of up, middle and low courses are "N(35.48-37.32)，E(102.5-104.2)"，"N(37.71-39.20)，E(107.40-109.33)" and "N(35.31-37.02), E(102.54-104.43)" individually. Annual mean precipitation and an annual average temperature of the three study regions have very different. More than 90% of the land in the up course region is covered by vegetation, the land in the middle course region is covered by sparse vegetation of loess plateau and the land in the low course region is covered by farm field and plantation.

2.2 Data collection

Time series of MODIS image (the 16-day composite data) acquired from the year 2000 to 2007, In addition, essential ancillary data including topographic maps, land use, transportation and soil maps were also obtained from relative department of local government. All the data layers were registered to the same coordinate system and resampled to the same pixel resolution of 250 m.

2.3 The proposed approach

The proposed approach, to utilize the advantages of both ANN and CA models, is to simulate vegetation conditions using remote sensing techniques. Specifically, this approach includes three major steps (Fig. 2). The first step is to extract control factors from multi-source data. The second step is to

produce candidate CA-ANN model based on advantage of ANN and CA model. The third step is future vegetation simulation using CA-ANN model.
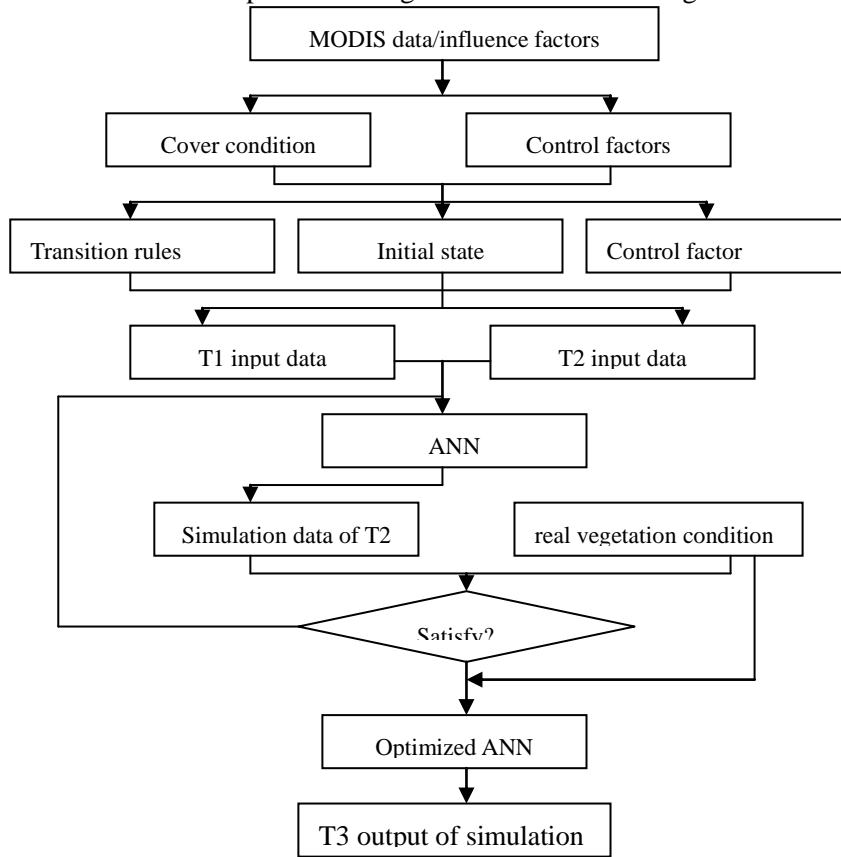

Fig.1   The flowchart of the proposed approach

2.3.1 Extracting control factors from multi-source data

   Vegetation indices are widely used in remote sensing of grassland since healthy vegetation has a high spectral response in near-infrared bands of remote sensing images (Langley et al., 2001). Vegetation indices generally involve the direct or modified forms of the ratio between the infrared band and the red band. Among vegetation indices, Normalized Differing Vegetation Index (NDVI) is regarded as one of the most useful indices to measure vegetation vigor. Coefficient of variation (CV) of time series of NDVI can reflect fluctuation of vegetation from 2000 to 2007 and which include also influence of some control factors such climate, slope of underlying surface. So CV is selected as control factors.

2.3.2 Establishing of CA-ANN model

    Cellular automata, a complex discrete and local grid dynamic system, is a representative model for simulating complex system. It has the characters such as the research frame of "from bottom to up", powerful compute function, intrinsic parallel calculate faculty, dynamic spatial and temporal concept, so it is suitable for complex system, to build cellular automata based model for complex Geo-Spatial system simulation is focused. About spatial simulation and prediction of vegetation condition, integrate the method of CA and artificial neural network, constructing artificial neural network prediction model using MATLAB software as development plot, simulating and predicting the dynamic changes of future vegetation condition. In the model, spatial cell database of vegetation condition and CV of NDVI from MODIS data from 2000 to 2007 as input data and result of the simulation of vegetation condition as output data.

2.3.3 Data of Simulating the selected area using CA-ANN model

(1) Processing of time series MODIS data

    NDVI data (mean value of DOY 177-273) of every study area is divide into 5 classes according to vegetation condition: up course is divided into level 1: 0.00-0.06, level 2: 0.06-0.24, level 3: 0.24-0.42, level 4: 0.42-0.60, level 5: 0.60-0.85; middle course: level 1: 0.00-0.06, level 2:0.06-0.15, level 3: 0.15-0.21, level 4: 0.21-0.3, level 5: 0.3-0.55; low course: level 1: 0.00-0.06, level 2: 0.06-0.25, level 3: 0.25-0.45, level 4: 0.45-0.65, level 5: 0.65-0.9.

(2) Setting up of input and output databases

    Training and prediction of artificial neural network need input and output data. The input data need CV data and cell data, and which compose of a piece of record. The cell data in record include self attribute (S0), 8 neighbor attributes (S1, S2, S3, S4…..S8) and spatial attribute (ID). The records of output data include self attribute (S0) and spatial attribute (ID). All these data and records can be implemented by the software MATLAB.

Tab 2-1 Structure of input data tables

ID	N0S	N1S	N2S	N3S	N4S	N5S	N6S	N7S	N8S	FACTOR

Tab 2-2 Structure of output data tables

ID	NS

(3) Process of simulation and prediction

1) Training, verification and predicting using ANN

After ANN studies using data tables of above input and aim data, input data of the verification database and aim database need to provide ANN, then we compare results of prediction with verification data tables, at last evaluate precision of simulation result, if the result can satisfy precision requirement, the ANN-CA model begins to execute the prediction of future vegetation condition, or the result can't satisfy precision requirement, we adjust control factors and relative parameters and retrain.

2) Structure of ANN and parameters

① Design of ANN structure

In input layer, regard attribute of original cell, its 8 neighbors and control factor CV as input data, the number of nerve cells of competitive layer is set N0.5 (N is the number of input samples). Nerve cells of output layer present 5 kinds of condition of future vegetation simulation.

② Set of ANN parameters

Step of ANN training is 10, and minimum mean square deviation is 0.01. When times of train reach 1000 or mean square deviation is less than 0.01, the training ends and the initial study efficiency is 0.5.

③ input vector and aim vector of ANN

The input data tables of vegetation condition correspond to input vectors of ANN, and aim data tables correspond to aim vectors of ANN.

3.   ANALYSIS OF THE RESULTS

According to above analysis and disign, we select time series MODIS data (mean value of NDVI (DOY from 177 to 273)) from 2000 to 2007 as study objects, and the result (figure 3,4,5 and table 3-1, 3-2, 3-3) indicts that NDVI of the vegetation has an increasing trend and characteristics of distribution concentration trend, but the change rate is become lower from the year 2007 to 2014 compared with the changes from the year 2000 to 2007.
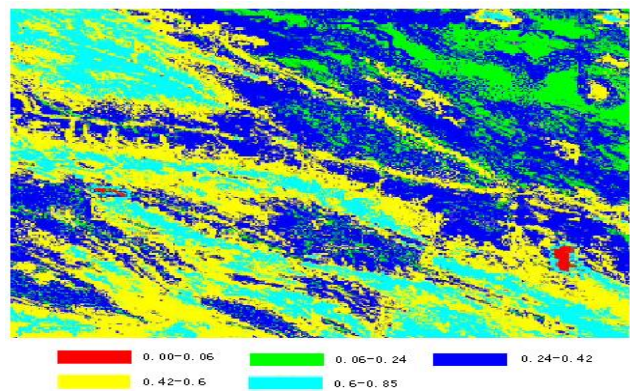
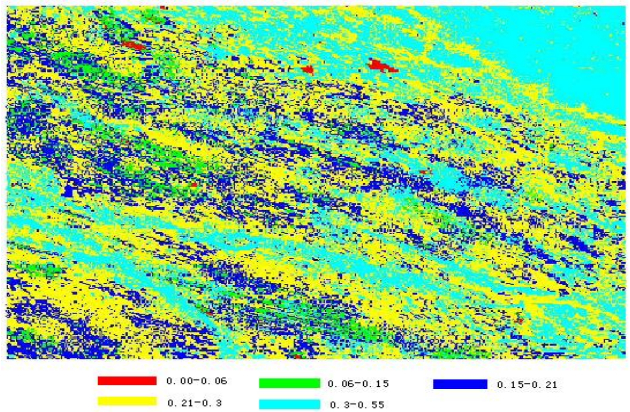*Fig 2 Vegetation condition simulation of up course in 2014*

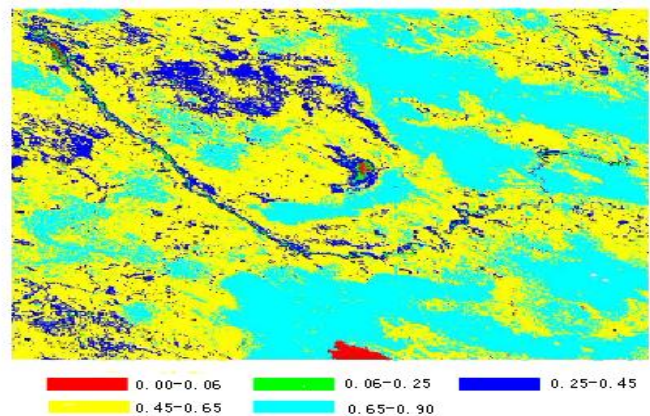Fig 3 Vegetation condition simulation of middle course in 2014

Fig 4 Vegetation condition simulation of middle course in 2014

Tab 3-1 Vegetation condition comparison of up course from 2000 to 2014

NDVI(level) / year	0-0.06（0）	0.06-0.24（1）	0.24-0.42（2）	0.42-0.60（3）	0.60-0.85（4）
2000	0.41%	40.53%	28.05%	22.73%	8.38%
2007	0.30%	22.23%	34.55%	29.64%	13.29%
2014	0.28%	12.13%	38.47%	33.06%	16.06%

Tab 3-2 Vegetation condition comparison of middle course from 2000 to 2014

NDVI(level) / year	0-0.06（0）	0.06-0.15（1）	0.15-021（2）	0.21-0.30（3）	0.30-0.55（4）
2000	0.27%	22.15%	49.97%	23.52%	4.22%
2007	0.30%	5.60%	25.01%	45.40%	23.70%
2014	0.25%	2.67%	14.46%	48.71%	33.81%

Tab 3-3 Vegetation condition comparison of low course from 2000 to 2014

NDVI(level) / year	0-0.06（0）	0.06-0.25（1）	0.25-0.45（2）	0.45-0.65（3）	0.65-0.9（4）
2000	0.433%	1.41%	29.19%	46.89%	22.08%
2007	0.437%	0.92%	11.03%	55.74%	31.87%
2014	0.46%	0.53%	3.81%	58.79%	36.41%

CONCLUSION AND EXPECTATION

The CA-ANN model is a kind of new method to research changes of future vegetation. It is proved effective to integrate advantage of CA and ANN for simulation of condition of future vegetation, but simulation of the changes of vegetation condition is very complex, and many collected factors need to consider and analyze.

REFERENCE

Le Houe rou, H.N., 1996. Climate change, drought and desertification. Journal of Arid Environment 34,133–185.

Angell, D.L., McClaran, M.P., 2001. Long-term influences of livestock management and a non-native grass on grass dynamics in the desert grassland. Journal of Arid Environments 49, 507–520.

Xu, Z.X., Zhao, M.L., Han, G.D., 2000. Eco-environmental deterioration and strategies for preventing it in Inner Mongolia. Grassland of China 5, 59–63.

Research on vegetation dynamic change simulation based on spatial data
mining of ANN-CA model using time series of remote sensing images

9

Langley, S.K., Cheshire, H.M., Humes, K.S., 2001. A comparison of single date and ultitemporal satellite image classification in a semi-arid grassland. Journal of Arid Environment 49, 401–411.

Wolfram, S., 1984. Universality and complexity in cellular automata. Physica D 10, 1–35.

Balzter, H., Braun, P.W., Hler, W.K., 1998. Cellular automata models for vegetation dynamics. Ecological Modeling 107, 113–125.

Batty, M., Xie, Y., Sun, Z., 1999. Modeling urban dynamics through GIS-based cellular automata. Computer, Environment and Urban Systems 23, 1–29.

White, R., Engelen, G., 2000. High-resolution integrated modeling of the spatial dynamics of urban and regional system. Computer, Environment and Urban System 24, 383–400.

Li, X., Yeh, G.O., 2000. Study on zoning of agricultural land protection using cellular automata model. China Environmental Science 20, 318–322.

Ripley, B., 1996. Pattern Recognition and Neural Networks. CambridgeUniv. Press, Cambridge.

Haykin, S., 1999. Neural Networks: A Comprehensive Foundation, Second edition. Prentice Hall, New Jersey.

Zhou, W., 1999. Verification of the nonparametric characteristics of backpropagation neural networks for image classification. IEEE Trans. Geosci. Remote Sens. 37, 771–779.

Lee, S., Ryu, J.,Won, J., Park, H., 2004. Determination and application of the weights for landslide susceptibility mapping using an artificial neural network. Eng. Geol. 71, 289–302.

Gomez, H., Kavzoglu, T., 2005. Assessment of shallow landslide susceptibility using artificial neural networks in Jabonosa River Basin, Venezuela. Eng. Geol. 78 (1–2), 11–27.

Yesilnacar, E., Topal, T., 2005. Landslide susceptibility mapping: a comparison of logistic regression and neural networks ethods in a medium scale study, Hendek region (Turkey). Eng. Geol. 79, 251–266.