FOREST PEST OCCURRENCE PREDICTION USING CA-MARKOV MODEL

Fangyi Xie [1], Xiaoli Zhang [1,*], Xiaoyan Chen [1]
[1] *Key Laboratory for Silviculture and Conservation, Ministry of Education, Beijing Forestry University, 100083, P. R. China*
[*] *Corresponding author, Address: Department of cartography and GIS, Beijing Forestry University, Beijing 100083.P.R.China,Tel:+86-010-62337830, Email: zhang-xl@263.net*

Abstract: Since the spatial pattern of forest pest occurrence is determined by biological characteristics and habitat conditions, this paper introduced construction of a cellular automaton model combined with Markov model to predicate the forest pest occurrence. Rules of the model includes the cell states rules, neighborhood rules and transition rules which are defined according to the factors from stand conditions, stand structures, climate and the influence of the factors on the state conversion. Coding for the model is also part of the implementations of the model. The participants were designed including attributes and operations of participants expressed with a UML diagram. Finally, the scale issues on forest pest occurrence prediction, of which the core are the prediction of element size and time interval, are partly discussed in this paper.

Key words: e-government, knowledge management, frameworks, e-governance

1. INTRODUNCTION

In cellular automata models, information from different sources can be translated into a set of transition rules which define the behavior of the system (Sergio A. Cannas et al., 1999). Thus, it was widely used in prediction and simulation of landscape dynamics combined with geographic models to increase their own realism, such as forest-landscape evolution(Cannas et al., 1999; Sprott et al., 2002), land use and cover changes (Li Xia et al., 2005), urban development simulation and analyzed

(Ke Chang Qing et al., 2006). While Markov models can show us a set of conditional probability states of different kinds of conversion with a transition matrix which is calculated by analyzing a pair of objects of different kinds. The CA-Markov model is a kind of enhanced cellular automata model, of which the knighthood rules and transition rules are combined with Markov models. Thus, the influence of subjective factors when modeling can be decreased to some extent.

From the perspective of insect population ecology, as an important property of population, spatial pattern is determined by biological characteristics and habitat conditions (Xu Lumei et al., 2005). Forest disease and pests prediction is not only quantitative calculation, but also spatial features analysis, such as location of the disease, spreading ranges and distribution patterns. So, with the research on the formation and change of forest pest spatial distribution, it's practicable to predict the forest pest occurrences from the perspective of spatial patterns. Obviously, the work would not be carried on without spatial models incorporated in. while it's just what CA-Markov can help us: transition factors from cellular automata and transition probability from Markov predicate the forest pest occurrence (Hou Xiyong et al., 2004).

In this paper, a CA-Markov model based on GIS was constructed to analyze the prediction of forest pest occurrence for GIS system. Here is something worth to noting that this model is not for some specific forest pest, such as Fall Webworm (Hyphantria cunea), but for major forest diseases and insect pests.

2. MODEL: RASTER-BASED CA-MARCOV

There are three key points when modeling with Raster-based CA-Markov: cell states rules, neighborhood rules and transition rules. Cell states represent the occurrence of forest pest, and transition rules express the likelihood of the cell state.

Five states should be included in this model: unchanged, healthy, slight, moderate and serious. If what the original cell state represent is not kind of vegetation, such as urban land, or kind of plants the pest never harm, the cell state should be unchanged, and should not change after the transition. While, other four states' determination depends on the probability calculated by transition rules integrated with Markov model's analyzing.

Generally, the occurrence of forest pest depends on these factors: date, climatic, stand conditions, stand structures; among those, climatic includes climatic type, temperature, humidity, rainfall; stand conditions includes slope, aspect, slope position and altitude; stand structures includes tree species (which and pure or mixed), stand layer structure, stand age, biodiversity and distance from the forest edge. Those factors are different from different kinds of forest pests, depended on the different biological characteristics (Zhang Xingyao et al., 2003).

The neighborhood rules should be made combined with those factors. Besides the four-neighbor method or eight-neighbor method, some distance, such as distance from the nearest road or from the nearest river, could also be used in the transition probability calculations, because some pest spread by transportation or water flow.

The core of CA models is how to define transition rules that control the conversion of states in the simulation. There are lots of methods to define the rules, such as the Multi-Criteria Evaluation (MCE), Logistic Regression, Principal Component Analysis (PCA) and Artificial Neural Network (ANN). Here we define the transition probability as

$$P = \sum_{i=1}^{m} \omega_i K_i + e \qquad (1)$$

Where P is the primary transition probability from the original state to the transform state, m is the number of the factors the model use, e is the error term. For each factor, i is the index, K_i is the value and ω_i is the weight.

This primary probability is calculated with the transition rules in cellular automata, which should be combined with the random probability from the matrix outputted by the Markov model's analyzing a pair of forest pest image. The two kinds of probability can be calculated as

$$PT = f(P, M) \qquad (2)$$

Where PT is the final conversion probability, while M is the random probability from Markov model's analyzing. The function f should be defined to decrease the influence of subjective factors.

In the constrained simulation, only the cells of which the conversion probability meets certain criteria can change those states.

3. CODING FOR THE MODEL

According to the designing of the model, we draw a UML diagram to descript the class structure. In Fig.1, it is shown that the member variables and member functions of classes and the relationship among the classes, which include Markov, StandCondition, StandStructure, DayClimate, MonthClimate, YearClimate, CAObject, CAMask, CAGroup and CommonFunction.

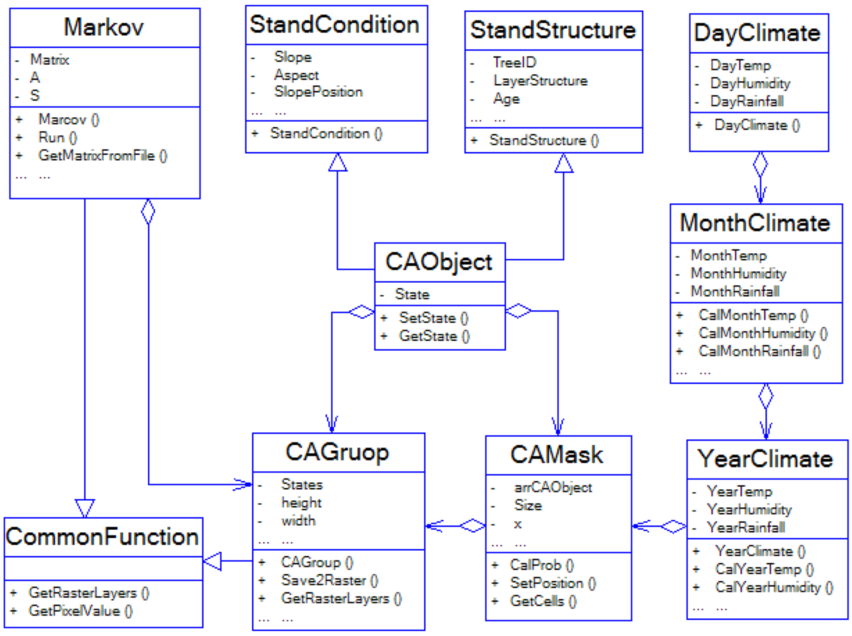

Figure 1. The UML diagram of CA-Markov model

Here are the Participants as follow:

Markov, which defines the matrix according to CAGroup, provides functions to calculate the transition probability matrix, and it also reads and saves the matrix from and to an existing file so that it's not necessary to calculate the transition probability matrix repeatability.

StandCondition defines the factors of stand conditions: slope, aspects, slope position, altitude.

StandStructure defines the factors of stand structures: tree species, stand layer structures, stand age, stand biodiversity and kinds of distance.

DayClimat, defines the factors of climate on the scale of the day: temp, humidity, rainfall.

MonthClimate provides functions to calculate the average, maybe maximum or minimum, of instances of DayClimate.

YearClimate provides functions like MonthClimate on the scale of the year.

CAObject which is the cell or element of the study scale, with the same size of the pixel size of raster layer data defines the state of the cell.

CAMask, which defines a gliding window to collect data around the core cell with the traversal of a pixel block, provides functions to calculate the primary probability according to the neighborhood rules.

CAGroup, which defines variables of CAObject instances, with the same size of a giving pixel block, provides functions to calculate the final probability, with the primary probability and the random probability, and determinate the convert states.

CommonFunction provides the common functions which could be used frequently, such as reading the raster layer data and getting value from a certain pixel from a raster layer data.

In all the calculating work, values of cells should be got from the raster data layers. It's not suggested to develop the analysis application only using advanced language. Using ArcGIS Engine, we can embed GIS functions into our existing applications for pre-processing vector and grid data. Especially, spatial extension, one of the extensions available for ArcGIS Engine Runtime, provides a set of functions to create, query, and analyze cell-based raster data. Factors such as slope and aspect can be easily calculated with the help of the extension.

The raster data should be calculated by pixel blocks, which conducive to the reading and writing data, using the interfaces such as IPixelBlock, IRasterCursor, IRasterEdit and so on. So that large raster can be divided into smaller pieces by creating a pixel block and reading portions of the raster sequentially.

4. DISCUSSIONS

Here is an aspect we could not neglect in the study of geological spatial analysis: scale. It's available to comprehend scale as the particularities of spatiotemporal characteristics of study objects (Zhang Tong et al., 2004). Generally, it can be repressed by spatial resolution and temporal sequence.

The scale issues, in which there are two core points, the prediction of element size and time interval, should be paid attention to in the study of forest pest occurrence prediction. It's more easily on the determination of time interval than that of element size, as most forest pest disaster occurs periodically because of the biological characteristics of the pest.

Obviously, the determination and acquisition of factors of neighborhood rules depends on the scale, and would change as the change of scale, so that it can adapt to the particularities. Factors act differently on the difference scale. The local scale incarnates the diversity of soil, plants and microclimate, which are influenced by biological effects, such as competitions. The regional scale incarnates the diversity of upper structures, which are influenced by geomorphic structures and soil combinations. And the global scale incarnates the diversity of macroclimate, which is influenced by terrestrial environment and atmosphere changes (Yu Xingxiao et al., 2005).

On the same scale, all the data should be raster data with the same cell size, not only the maps, but also the remote sensing data. All the data should be georeferenced to the same coordinate system. The choosing of the date should be paid attention to. The date of the data should be closed to that of the occurrence of forest pest, and the spatial resolution should not be lower than that of the cell size set in the model.

ACKNOWLEDEGMENTS

Funding for this research was provided by National Science and Technology Support Project in the 11the Five Years (Project No. 2006BAD23B04-4), and NCET (Programme for New Century Excellent Talents in University, Project No. NCET-06-0122).

REFERENCES

Cannas S A, Páez S A, Diana E M. Modeling plant spread in forest ecology using cellular automata. Computer Physics Communications, 1999, 121~122

Hou Xiyong, Change Bin, Yu Xinfang. Land use change in Hexi corridor based on CA-Markov methods. Transactions of the CSAE, 2004, 20(5):286~291

Ke Chang Qing, Ouyang Xiao Ying. The Advances in Modeling Urban Spatial Change Based on Cellular Automata. Journal of Nanjing University (Natural Sciences), 2006:42(1): 103~110(In Chinese)

Li Xia, Anthony Gar-On Yeh.. Celluar automata for simulating complex land use systems using neural networks. Geographical Research, 2005, 24(1): 19~27(in Chinese)

Liu Xiaoping, LI Xia. Retrieving CA Nonlinear Transition Rule from High-dimensional Feature Space.Acta Geographica Sinica

Sergio A. Cannas, Sergio A. Páez, Diana E. Marco. Modeling plant spread in forest ecology using cellular automata. Computer Physics Communications, 1999, 121(122):131~135

Sprott J C, Bolliger J, David J M. Self-organized criticality in forest-landscape evolution. Physics Letters A, 2002, 297: 267~271

Xu Lumei, Cheng Xinyue. Insert Population Ecology: Foundation and Front, 2005

Yu Xingxiao, Niu Shuzhi, Guan Wenlin, Feng Zhongke. Landscape Ecology, 2006

Zhang Tong, Cai Yongli. Scale in Ecological Research. Ecologic Science, 2004, 23(2):175~178

Zhang Xingyao, Luo Youqing. Major Forest Diseases and Insect Pests in China, 2003

Zhou Guo-Fa, XU Ru-Mei, LI Tian-Sheng. Spatio-temporal explicit interactive model for forecasting the pine caterpillar population dynamics. Acta Ecological Sinica, 1999, 19(4):464~467(in Chinese)