# Approach to Land Cover Change Modelling Using Cellular Automata

Alexander Zamyatin, Nikolay Markov
Tomsk Polytechnic University
84, Sovetskaya street, Tomsk, 634034, Russia
zamyatin@tpu.ru

## SUMMARY

*Nowadays one of the most effective and widely used ways for the land cover change modelling is an application of cellular automata (CA). One of the key factors in influencing on the effectiveness of the CA application is defining the CA transition rules in each specific case.*

*The paper proposes a compositional probabilistic approach to defining CA transition rules with more complex nature compared to the traditional approach to the CA land cover change modelling. The preliminary investigation results show that proposed approach allows performing the modelling more effectively and its results could be more accurate compared to ones obtained by existing models of land cover change. The accuracy estimation of the proposed approach is conducted using model time series thematic maps and truth thematic maps obtained from the time series satellite images RESURS-O1 for the Uymon steppe area (Altay region, Russia) on 1998, 1999, 2000.*

**KEYWORDS:** *Cellular automata, land cover change, transition probabilities*

## INTRODUCTION

Land surface represents by itself a complex system, and a land cover change modelling is a complicated process, implicating a variety of driving forces (Lambin E., 1994). To represent a piece of the interesting land surface as a model it needs to obtain the image of the piece as a matrix. Each element of the matrix presents the specific land cover type of the investigated land surface. Actually the matrix is a raster thematic map of such land surface. Usually such thematic maps for land cover change modelling could be obtained by remote sensing (RS) methods and image processing, RS and mapping software.

Nowadays one of the most effective and widely used way of the land cover change modelling, taking into account the spatial interaction between the elements of such thematic map is an application of cellular automata (CA). One of the key factors influencing on the effectiveness of the CA application is defining the CA transition rules in each specific case.

The paper proposes a compositional approach to defining the CA transition rules differing more complex nature compared to traditional approach. Usually the CA transition rules in such traditional approach are defined by majority or (and) minority principles that in many cases does not allow to perform the effective land cover change modelling. The proposed approach allows to take into account the spatial interaction between the land cover elements and to get effective CA transition rules for each automaton used while modelling. All of this allows performing the land cover change modelling more effectively and the modelling results could be more accurate compared to results obtained by existing land cover change models.

## METHODOLOGY

The set of cellular automatons, which are used for modelling, is forming while consecutive scanning of the initial image. Each cellular automaton can be represented as square matrix $M_{CA} = [c_{ij}]$ with order of matrix $d$. The value of central element $c'^{kh}$ of each matrix depends in some way on all elements of the matrix $c'_{kh} = f_{CA}(c_{11}, c_{12},\ldots, c_{kh},\ldots, c_{dd})$. The obtained value $c'_{kh}$ is an element of the new resulting image. In the framework of the traditional approach to defining the CA transition rules the search of the function $f_{CA}$ often is based upon simple transition rules and the equal obtained function $f_{CA}$ is to be used for all automatons while modelling.

The proposed approach to land cover change modelling is developing the compositional way of forming function $f_{CA}$, where three probabilistic components together are used. It allows taking into account the features of the transition rules for the all the CA applied. Each component presents corresponding transition probability of land cover from state $\omega_i$ to state $\omega_j$ ($p_{ij}$). Let's consider the details of the every transition probability definition.

## Driving factors component

Nowadays for the land use change modelling the approach on the basis of urban growth model and so-called suitability maps are widely using (Clarke K.C. et. al., 1997). These suitability maps contain information about the probability of urbanized in a location of the investigated area depending on the different driving factors which influences on urbanization in a specific location. The examples of such driving factors are distance to roads, distance to markets, distance to water facilities, slope information etc. To apply such perspective approach with suitability maps to the land cover change modelling, it is needed to solve very complex and intractable task of suitability maps designing. The designing of the maps requires to formalize a lot of different data such as a variety of driving factors and the land cover types at the investigated area.

One of the most effective solutions, allowing to use such complex and non-formalized data and allowing to design the suitability maps for the land cover change modelling is the application of artificial neural networks (ANN). The widespread topology for solving such task is multilayer perceptron (MLP) with one hidden layer. Let's apply the technology of the MLP application described in (Yeh, A. G.-O. & X. Li., 2002) for urban growth modelling on the basis of ANN CA. The features of the ANN application for training and recall modes in the framework of the proposed approach are depicted in the Fig. 1.

In common case, applied MLP has $K = M{\cdot}F$ input and $M$ output neurons, where $M$ — the number of land cover types, $F$ — the number of driving factors. The process of getting by MLP the necessary probability information, considering revealed driving factors, consists in the idea that each $j$-th axon of output's neuron is interpreted as the transition probability of a state $\omega_i$ to a state $\omega_j$ of a pixel being analyzed in the initial image ($p_{ij}^{\mathrm{df}}$).

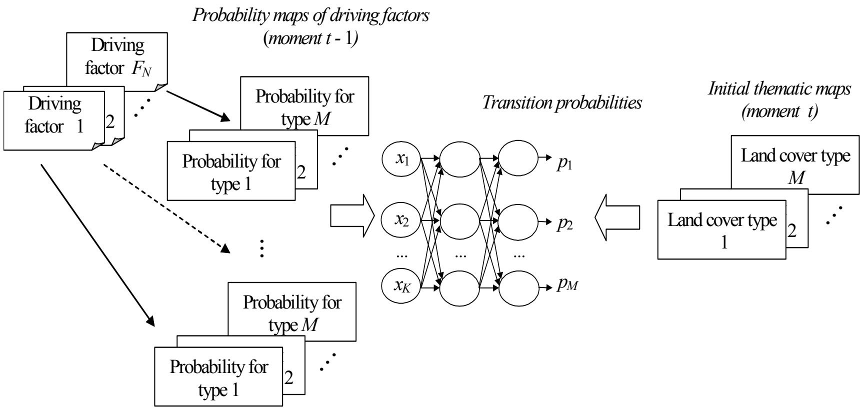
Let's consider in details the process of the revealed driving factors formalizing and obtaining the suitability maps for each land cover type. The process is based upon the technique of ANN application (Fig. 1). As initial data we use two thematic maps for the time moments $t-1$ and $t$.

For each revealed driving factor for the time moment $t-1$ by means of raster GIS functions and initial thematic map the $M$ probability suitability maps of driving factors are designed. These maps contain information about the occurrence probabilities in a location for the land cover types on the investigated area. It should be noticed if for some reason a probability suitability map could not be designed it is proposed to use at the map the equal non-zero value as *Probability for type i* (Fig. 1) in every location of such map in condition that

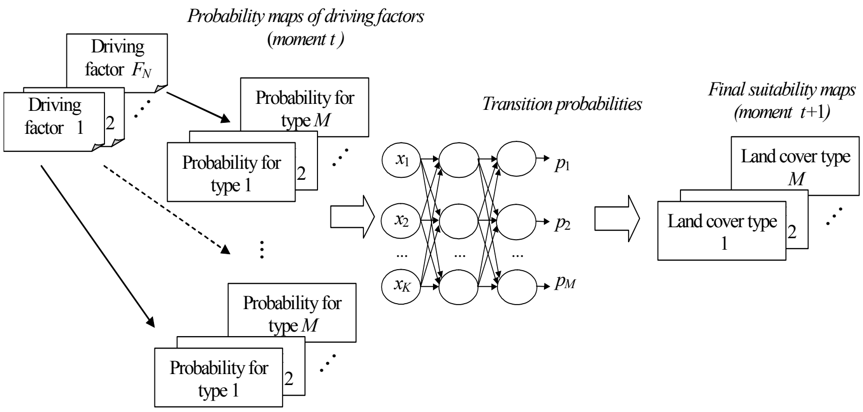
$$\sum_{i=1}^{M} \text{Probability for type} \quad i = 1.$$

When all input data is completed the ANN is to be trained. As the training data the criterion output values obtained from $M$ thematic maps for the time moment $t$ are to be used. Each map corresponds to its thematic class at that the pixel on the map contains the value 1 in the presence of the thematic class in a location and the value 0 — in the absence the thematic class. The ANN trains by the traditional algorithm of error backpropagation for root mean square (RMS) error minimization (Fig. 1a).

The ANN has been trained with the appropriate RMS error and after that it can be used for the designing the suitability maps for the time moment $t + 1$ (Fig. 1b). Before that the set of the probability maps for different driving factors for the time moment $t$ should be designed. The technique of designing such maps is also based upon raster spatial analysis functions (GIS functions) and it takes after the technique described above for the designing probability suitability maps for revealed driving factors for the time moment $t – 1$.




a — training mode




b — recall mode

*Figure 1:* Using neural network for driving factors formalization.

Thus, the all input patterns, actually it is the pixel values of the designed probability maps for the driving factors, are processed by the ANN. The output recall ANN patterns are to be used for designing the final probabilistic suitability maps for the time moment $t + 1$. These suitability maps contain the transition probabilities for each land cover type from a state $\omega_i$ to a state $\omega_j$ in all locations at the investigated area.

## Probabilistic component

The probability information in the land cover change modelling of land surface phenomena often is based upon Markov chains (Baker W. L., 1989). An important component of Markov chains is the transition probability matrix $\mathbf{P} = [p_{ij}]$, traditionally obtained by two time series thematic maps. Thus, defining the second component in the framework of the proposed approach is based upon the transition probability matrix elements. At that the transition probability of a state $\omega_i$ to a state $\omega_j$ in an analyzed neighborhood depends not only on the probability $p_{ij}$ but on the number of elements with the state $\omega_j$ in the neighborhood. For each state in the analyzed neighborhood, coinciding with the elements of a current cellular automaton, probability $p_{ij}^{\mathrm{prob}} = n_j \cdot p_{ij}$, is defined, $j = 1,2,\ldots, m$, $m$ — the number of states in the analyzed neighborhood, $n_j$ — the number of elements with state $\omega_j$ in the neighborhood.

## Spatial component

Third component of CA transition rules is based upon spatial characteristics of each land cover type. Spatial characteristics are defined on the basis of enrichment factor, proposed in (Verburg P.H. et al., 2003). The technique of the enrichment factor application consists in the following. Corresponding vector $\mathbf{F}_i^{\mathrm{enr}} = \{f_1, f_2,\ldots, f_M\}$ for each state $\omega_i$ is constructed, $i = 1..M$, $M$ — the number of land cover types. Vector $\mathbf{F}_i^{\mathrm{enr}}$ contains information about the enrichment of type $\omega_i$ on the whole image. Then for each pixel on the image the local enrichment vector $\mathbf{F}_i^{\mathrm{loc}}$ is calculated. After that the probability $p_{ij}^{\mathrm{spat}}$ is defined for every location as $p_{ij}^{\mathrm{spat}} \sim 1/d(\mathbf{F}_i^{\mathrm{enr}}, \mathbf{F}_i^{\mathrm{loc}})$, where $d$ — Euclidian distance between the vectors.

When all components are already defined the final transition probability of the initial image elements can be represented as $p_{ij}^{\mathrm{fin}} = f_{\mathrm{CA}}(p) \sim (p_{ij}^{\mathrm{prob}} + p_{ij}^{\mathrm{spat}} + p_{ij}^{\mathrm{df}})$.

An additional important moment for the land cover change modelling is the defining not only the transition probability $p_{ij}^{\mathrm{fin}}$, but also the change order of the elements of the resulting image (ranking). The ranking in the framework of the proposed approach is to be defined in the following way on the image. The elements with the greatest values of $p_{ij}^{\mathrm{fin}}$ are to be changed first and the elements with the least values $p_{ij}^{\mathrm{fin}}$ are to be transited last. The preliminary investigation results of the ranking show its usefulness. Moreover it allows to obtain significantly more adequate the land cover change modelling results.

# CASE STUDY

The accuracy estimation of the proposed approach using model time series thematic maps is conducted. The modelling performs on the basis of the first and the second images and the accuracy is estimated on the third image by confusion matrix and kappa index of agreement. Factually, we compare the obtained image and the actual image on the same time moment and create the confusion matrix between them. This matrix contains the number of pixels on obtained image which either coincident or do not coincident with the corresponding pixels of the actual image. To obtain more precise results of the accuracy estimation the experiments were conducted for many times. As the final accuracy estimation result the average value of all experiments is taken.

Some representative results for four triple images are depicted in Fig. 2. The notice «1,2–>3» on the abscissa means that images 1 and 2 were used for calculating parameters and obtaining forecast map, image 3 was used to compare it with the obtained forecast map and to estimate the accuracy by kappa index of agreement. The proposed approach to modelling, applied on the set of the model data, allows to reach the average accuracy about 93%. At the same time the traditional approach to CA modelling allows to get the accuracy that is not exceeded 75%.

Moreover the proposed approach to modelling the land cover change using CA has been applied for the efficiency investigation on truth thematic maps obtained from real time series aerospace images. As an example, the modelling is performed on the basis of the thematic maps, obtained from satellite images of Uymon steppe area (Altay region, Russia). Investigated area could be separated to six major land cover classes. There are flood-plain, vegetation, coniferous, high bushes, low bushes and water.

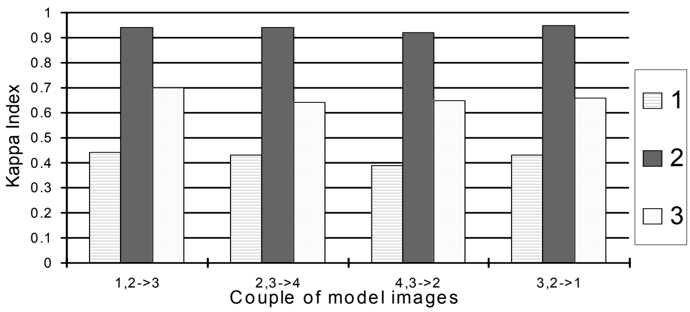



*Figure 2:* Accuracy estimation results obtained by the different CA modelling techniques: 1 — with probabilistic component only, 2 — proposed approach, 3 — traditional approach.

These satellite images with the spatial resolution about 30 m were obtained from RESURS-O1 satellite in 1998, 1999, 2000. The thematic interpretation of the satellite images has been performed by the original interpretation software, developed in the GIS laboratory of Tomsk Polytechnic University (A.V. Zamyatin & N.G. Markov, 2004).

On the Fig. 3 and Fig. 4 the modelled images for the time moments $t + 1$, $t + 2$ and $t + 3$ ($t$ — the time moment for 2000) are depicted. It should be noted that these images are obtained by the traditional approach (Fig. 3), which is based upon simple majority CA transition rules and by the proposed approach (Fig. 4). Even visually, it can be noticed that all modelled images in Fig. 4 have more legible borders of the thematic classes compared to images obtained by the traditional approach depicted in Fig. 3. Also it can be noticed that for modelled results on the time moments $t + 2$, $t + 3$ etc. the “salt and pepper” effect is becoming more significant that could be lead to decreasing the accuracy of the forecast maps on further time moments.

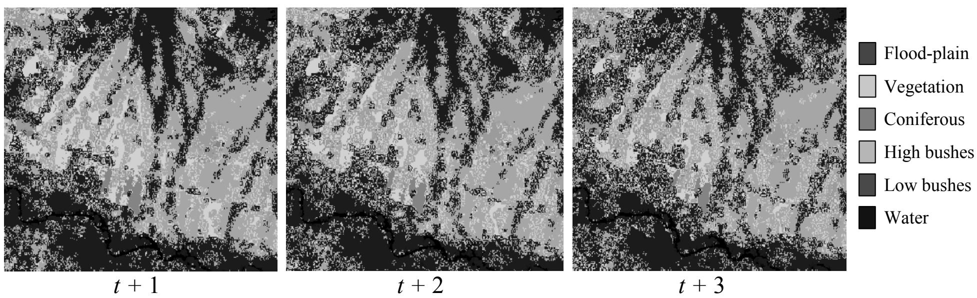
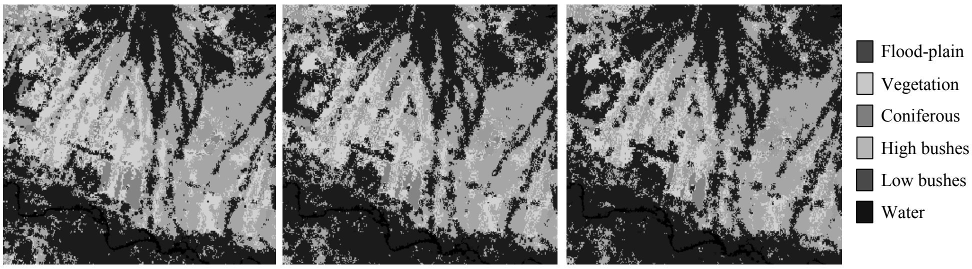



*Figure 3:* Example of land cover change modeling for the time moments $t + 1$, $t + 2$ и $t + 3$ with the traditional defining CA transition rules.

We perform the analysis of the land cover change in the study area. The analysis shows that flood-plain class was significantly increasing during the observed time period. For example, at 1998 it was 205,9 km$^2$ and at 2000 it became about 280,3 km$^2$. At the same time period the square of coniferous vise-versa lost 18 km$^2$ from 64 km$^2$ to 46 km$^2$. The water keep the same square during the all observed time period.




*t + 1* *t + 2* *t + 3*

*Figure 4:* Example of land cover change modeling for the time moments $t + 1$, $t + 2$ и $t + 3$ with the proposed defining CA transition rules.

The numerical accuracy estimation of the forecast map on 2000 is obtained by a comparison with the truth thematic map on 2000. The estimation demonstrates the accuracy of the obtained forecast map by the proposed approach with kappa index of agreement about 75%. The same forecast map, which is obtained by the traditional approach with the simple majority CA transition rules, allows to get the accuracy 61% only.

Thus, analyzing all the investigation results, which were obtained with the modelled data and real satellite data, we can promote the idea that the proposed approach including different probabilistic components for defining CA transition rules is significantly more effective compared to traditional approach.

## CONCLUSION

Considering the possibility of up-to-date ways to get the information of land cover dynamics including remote sensing methods the solving task of the land cover change modelling is very urgent. Nowadays one of the most effective facilities for the land cover change modelling is CA. The preliminary investigation results show the proposed probabilistic approach to defining the CA transition rules allows to perform the process of the land cover change modelling more effectively. Moreover such approach allows to make the design of the forecast information more accurate then it might me done by existing models of the land cover change modelling.

## ACKNOWLEDGMENTS

The authors are grateful for the support this research by the grant N 03-07-90124 from Russian Foundation for Basic Research and the grant N 432 from Federal Education Agency.

## BIBLIOGRAPHY

Baker, W.L. 1989. A review of models of landscape change. Landscape Ecology 2: 111-133.

Clarke K.C. et al., A self-modifying cellular automaton model of historical urbanization in the San Francisco Bay area, Environment and Planning B: Planning and Design 24:247-261, 1997.

Lambin, E. 1994. Modeling Deforestation Processes: A Review. European Commission publication EUR 15744 EN.

Verburg, P.H. et al., 2003. A method to analyse neighbourhood characteristics of land use patterns, Computers, Environment and Urban Systems, in press.

Yeh, A. G.-O. and X. Li. 2003. Simulation of development alternatives using neural networks, cellular automata, and GIS for urban planning. Photogrammetric Engineering and Remote Sensing 69(9): 1043-1052.

Zamyatin A.V. and Markov N.G., Designing Forecast Thematic Maps Using Time Series Remotely Sensed Images. IAPRS, part B2, Vol. XXXV, ISSN 1682-1750: 492-497, 2004.