Land Use Change Prediction using a Hybrid (CA-Markov) Model

Alireza Ildoromi¹ and Mahtab Safari Shad²*

¹ Associate Professor, Department of Range and Watershed Management, Malayer University, Malayer, Iran
² Ph.D. Student, Department Watershed Management, Faculty of Natural Resources, Sari University of Agriculture and Natural Resources, Sari, Iran

Received: 6 September 2016 / Accepted: 14 January 2017 / Published Online: 20 March 2017

ABSTRACT Landsat data for 1992, 2000, and 2013 land use changes for Ekbatan Dam watershed was simulated through CA-Markov model. Two classification methods were initially used, viz. the maximum likelihood (MAL) and support vector machine (SVM). Although both methods showed high overall accuracy and Kappa coefficient, visually MAL failed in separating land uses, particularly built up and dry lands. Therefore, the results of SVM were used for Markov Chain Model and “CA” filter to predict land use map for 2034. In order to assess the ability of “CA Markov” model, simulation for 2013 was performed. Results showed that simulated map was in agreement with the existing map for 2013 at 84% level. The land use map prediction showed that built up area of 0.8298 km² in 2013 will increase to 1.02113 km² in 2034. In contrast, irrigated agriculture will decrease from 17.33 km² to 17.16 km², and rain fed agriculture from 45.07 km² to 44.49 km². Results of this research proved the application of “CA Markov” model in simulating the land use changes.

Keywords: Cellular Automata model, MAL, SVM

1 INTRODUCTION
Currently, the intensity of land use changes has created unprecedented changes in ecosystem and ecological processes at local, regional and global levels (Muñoz-Rojas et al., 2015; Yalew et al., 2016). Changes in land use are a direct result of human activities (Jafarian Jeloudar et al., 2014; Debolini et al., 2015). In order to manage resources for sustainable development, the quantitative information about land use change is useful (Slam et al., 2016). Many works have been conducted on the effects of land use change (Sharma et al., 2011; Salazar et al., 2015). The research evidences show that land use change cause significant decrease in rainfall, increase in surface temperature, and have adverse effects on the adjacent areas. Land use change in this research, two methods of “maximum likelihood (MAL)” and the “support vector machine (SVM)” were used to classify images and preparation of the land use maps. The principle of SVM is linear classification of data that tries to select the line that has greater safety margin (Huang et al., 2006). The SVM is one of the supervised learning methods used for classification and has been found an effective method in many studies (Camps-Valls and Bruzzone, 2008; Mountrakis et al., 2011; Devadas et al., 2012; Saiful Bahari et al., 2016). It can also be used as a tool in predicting changes in land use and land conversion approaches by managers to identify and control the land use changes trend (Braimoh

*Corresponding author: Ph.D. Student, Department Watershed Management, Faculty of Natural Resources, Sari University of Agriculture and Natural Resources, Sari, Iran, Tel: + 98 918 270 9921, Email: mahtabsafari66@gmail.com
Since the predisposing factors of land use change are complex and human activities and natural factors in this process may be affected by different intensities and scales, its simulation is a difficult process (Verburg et al., 2002). For this reason, different models have been modified and developed by researchers to simulate the complexity and dynamics of lands (Singh, 2003).

“Markov chain” analysis is one of the statistical models that can predict the state of the modeled system at time of (t+1), based on its condition at the time “t”. The “transition probability matrix” is the most important of its output. This matrix is informative and does not include transfer of spatial distribution. It seems that models should have spatial components in order to provide the spatial pattern of distribution (Breckling et al., 2011). For this reason, “cellular automata” model is used in combination with “Markov chain”, which can present the dynamic of complex systems via simple rules. In this way, we can also simulate the future system spatially (Torrens and O’Sullivan, 2001).

Ekbatan Dam watershed is a sub watershed in Hamadan province, Iran. Due to the effects of light and shadow in this watershed as well its small size (being 221 km² and containing only 5 villages), preparation of a basic land use map through Landsat data is very difficult, especially for residential area. Therefore, the “MAL” and “SVM” methods were used to prepare the best land use map for the study area. Then by merging this map with “CA Markov model” the land use map for 2034 was predicted.

2 MATERIALS AND METHODS
2.1 Study area
Ekbatan dam watershed with an area about 221 km² is geographically located at 34° 36’ to 34° 45’ N and 48° 27’ to 48° 42’ E. It is a mountainous region north of Alvand Peak, 10 km southeast of Hamadan city at altitudes ranging from 1960 to 3572 m (Figure 1) (Hamadan Regional Water Authority, 2008).

2.2 Methodology
In order to achieve the objectives of the study, the land use maps for three periods were required. Due to the lack of these maps, three frames of non-clouded satellite images for 1992, 2000 and 2013 were used. Field data and GPS records were also applied in preparing the land use maps, for which the ArcGIS 9.3 and “IdrisiSelva” version were used. Furthermore, to predict land use map for the 2034, the “CA Markov” hybrid model was used.

2.3 Preparation of land use maps
First, radiometric quality was checked and geometric condition was controlled. The detectability of the images were enhanced by application of appropriate color composition that reflected the regional phenomenon, including bands 4, 3, and, 2 for TM and + ETM image sensor, and 5, 4 and 3 for the OLI image sensor. After visual inspection of the images, the two methods of supervised “MAL” and “SVM” were used to classify images. SVM has a sigmoid kernel function, through which the training data are planned in the form of nonlinear multidimensional space and separate them linearly in a data set that results in a linear classification of data (Vapnik, 1999). The main feature of this method is higher ability in the use of less training samples and achieves higher accuracy than other methods (Mantero, 2005). Thus in this research, the “MAL” (as a common method in land use classification) has been compared with the “SVM” method for the Ekbatan dam catchment. In order to assess the produced maps, the Kappa index, overall accuracy and visual interpretation were applied. Kappa index and overall accuracy equations are present bellow:

\[
\hat{K} = \frac{P_E - P_C}{1 - P_C}
\]
where $K$ is Kappa index; $P_o$ is overall accuracy; $P_c$ is chance agreement; $i$, $n$, and $n$ are row, column and total numbers of pixels in the classification table, respectively.

2.4 Implementation of “CA Markov” model

In order to use “CA Markov” model in prediction of land use maps, we first need to evaluate the ability of model to predict regional variations. So using “Markov chain” analysis and land use maps of 1992 and 2000, the “transition probability matrix” during this period was calculated. Then, using “cellular automata” model and result of “Markov” analysis, land use map was predicted for 2013. The forecast maps with the map that was produced using supervised classification for the same year was evaluated and the model was validated. After validating the model and based on the maps of the 2000 and 2013, land use map for 2034 was predicted.

![Figure 1: Geographical allocation of the study area in Iran](image-url)
3 RESULTS
3.1 Land use maps
Land use maps provided by the “MAL” and “SVM” methods are presented in Figure 2. Statistics error matrixes of the produced maps are presented in Table 1.

The results of the detection of past changes in the period 1992 to 2000, 2000 to 2013 and 1992 to 2013, and the area of land use classes is provided in Table 2.

Figure 2 The maps of land use classification with SVM and Max method

Figure 2 Continued

Table 1 Overall accuracy and Kappa coefficient of land use maps by the two classification methods

<table>
<thead>
<tr>
<th>Classification method</th>
<th>Overall accuracy (%)</th>
<th>Kappa coefficient (%)</th>
<th>Overall accuracy (%)</th>
<th>Kappa coefficient (%)</th>
<th>Overall accuracy (%)</th>
<th>Kappa coefficient (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAL</td>
<td>92.05</td>
<td>0.8242</td>
<td>90.01</td>
<td>93.6265</td>
<td>0.4851</td>
<td>75.8399</td>
</tr>
<tr>
<td>SVM</td>
<td>95.23</td>
<td>0.8922</td>
<td>96.5489</td>
<td>95.23</td>
<td>0.8344</td>
<td>87.7505</td>
</tr>
</tbody>
</table>

Table 2 The land cover changes in km² in 1992, 2000 and 2013

<table>
<thead>
<tr>
<th>Classes</th>
<th>1992 Area (km²)</th>
<th>2000 Area (km²)</th>
<th>2013 Area (km²)</th>
<th>1992-2000 change (km²)</th>
<th>2000-2013 change (km²)</th>
<th>1992-2013 change (km²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Irrigation farming</td>
<td>17.272647</td>
<td>18.232094</td>
<td>17.330215</td>
<td>-0.959447</td>
<td>-0.901879</td>
<td>0.057568</td>
</tr>
<tr>
<td>Dry farming</td>
<td>46.122504</td>
<td>36.616555</td>
<td>45.232129</td>
<td>-9.505949</td>
<td>8.45904</td>
<td>-1.046909</td>
</tr>
<tr>
<td>Rangeland</td>
<td>156.460231</td>
<td>164.926948</td>
<td>156.722129</td>
<td>8.466717</td>
<td>-8.204819</td>
<td>0.261898</td>
</tr>
<tr>
<td>Water bodies</td>
<td>0.22114</td>
<td>0.151481</td>
<td>0.465547</td>
<td>-0.069659</td>
<td>0.314066</td>
<td>0.244407</td>
</tr>
<tr>
<td>Built up</td>
<td>0.3474</td>
<td>0.4984</td>
<td>0.8298</td>
<td>0.151</td>
<td>0.3314</td>
<td>0.4842</td>
</tr>
<tr>
<td>Total</td>
<td>220.42</td>
<td>220.42</td>
<td>220.42</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
In 1992 to 2000, the dry farming had decreased by 20.6%, while residential area had increased by 43%. In 2000 to 2013, rangelands showed the highest reduction of about 5.24%, while residential area increased by 0.66%. In 1992-2013, dry farming had decreased by 2.27%, while the residential area had shown the highest growth (137%) compared with other land use classes.

Figure 3 Markov probability change of land use (1992-2000)

Figure 4 Map of 2034 land use: a result of “Cellular Automata-Markov”
Table 4 The land use changes in km² in 2013 and 2034

<table>
<thead>
<tr>
<th>Classes</th>
<th>2013 area (km²)</th>
<th>2034 area (km²)</th>
<th>2013-2034 change (km²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Irrigation farming</td>
<td>17.330215</td>
<td>17.169947</td>
<td>-0.160268</td>
</tr>
<tr>
<td>Dry farming</td>
<td>45.075595</td>
<td>44.493754</td>
<td>-0.581841</td>
</tr>
<tr>
<td>Rangeland</td>
<td>156.722129</td>
<td>157.209452</td>
<td>0.487323</td>
</tr>
<tr>
<td>Water bodies</td>
<td>0.465547</td>
<td>0.525717</td>
<td>0.06017</td>
</tr>
<tr>
<td>Built up</td>
<td>0.8298</td>
<td>1.02113</td>
<td>0.19133</td>
</tr>
<tr>
<td>Total</td>
<td>220.42</td>
<td>220.42</td>
<td></td>
</tr>
</tbody>
</table>

3.2 Land use change prediction
Matrix of “Markov Probability Change” between 1992 and 2000 in combination with “CA” model was used to produce land use map for 2013 (Figure 3). A comparison of actual land use map and predicted land use map by the “CA Markov” model for 2013 showed the degree of agreement between simulated and base maps was 0.8471, which indicated the success of this model to simulate changes in the study area. Map predicted for 2034 is presented in Figure 4.

Comparison of the predicted map of study area with the land use map of 2013 showed that during 2013-2034 the rangeland and built up classes would increase, while the agricultural land would decrease by about 0.74 km² (Table 4).

4 DISCUSSION
Since the detection of the land use changes is affected by limitations of spatial, temporal and spectral specifications of satellite data, there are many different ways to do this. Hence, selecting appropriate methods to detect changes is very important and, of course, very difficult. In order to prepare the land use maps, the supervised “MAL” and “SVM” methods were used in the present study. The overall accuracy and Kappa coefficient greater than 70% in terms of accuracy is very good and less than 40% is weak (Mirzaee Mosavand, 2011). According to Table 1, Kappa and overall accuracy of both methods for 2000 and 2013 are very good and for 1992 in terms of Kappa are very good but in terms of overall accuracy are appropriate. Although, according to Table 1, both methods have been successful; but reviewing and comparing the visual results of the analysis showed that “MAL” method allocated wrongly the irrigated agriculture pixels to build up pixels in some places and also had error in the simulation of dry lands. The results showed a better performance by the “SVM” than the “MAL” method, which is in correspondence with the results from other studies (Devadas et al., 2012; Rezaii Makhdoom et al., 2016; Saiful Bahari et al., 2016). Evaluation of changes in the region showed that the low gradients of agricultural land as well as its accessibility to built-up areas had resulted to its gradual destruction in due course of time. Prediction of the land use was done through “CA-Markov” model using the error matrix method. Verification of the result showed the predicted outcomes had the Kappa index of 0.84. Since Kappa coefficient involve both correct and incorrect predicted pixels, it is a more accurate criterion to evaluate the prediction accuracy (Rezaii Makhdoom et al., 2016). The “CA-Markov” model predicted the land use map for the year 2013, using the years 1992-2000 as the first period of “transition probability matrix”. Due to the differences in the effective economic and social factors, the land use change trends in the first and second period were not equal. Thus, it seems logical that the predicted changes by the model is somewhat different from the reality that has been also expressed in other studies (Falahatkar et al., 2009; Dezgham et al., 2015; Haibo et al., 2011. Despite the dynamic characteristic of the factors and land use change processes over
time in “CA Markov” model, one of the its weaknesses is the non-interference of the physical and socio economic factors affecting the land use changes. Since regional development policies, economic interests and demographic changes of temporal variables, especially in developing countries like Iran are not predictable for long periods, these factors have not been considered in predicting land use maps in the model, which may make the results not reliable in the long term. However, the aim of prediction of land use changes often is assessment of the consequences of different scenarios, especially the continuation of the existing trend. The results of these predictions, that give information about what will happen in the future, can be a warning for the future land use situation. Therefore, we recommend that the factors affecting land use changes should be involved in future studies in order to remove the disadvantage of the model partly and increase the reliability of the results.

5 CONCLUSION
Based on the results of this study, “SVM” is a good method to prepare land use maps and can succeed in classifications, even with fewer training samples. In addition, it shows a better performance even if classes show similar spectral behavior. The results of this study suggest that a combination of remote sensing and geographic information systems in implementing the spatiotemporal land use models. Information about the type and percentage rate of land use change in natural resources and other sectors are very useful tools in natural resource management and can help planners in management and comprehensive development.

6 ACKNOWLEDGMENTS
The authors of the present paper feel obliged to thank Hamadan Natural Resource Organization for providing us with the land use map and we should also express our gratitude to the NASA organization for supplying the land sat images to its users for free.

6 REFERENCES
Breckling, B., Pe’er, G. and Matisons, G. Modeling Complex Ecological Dynamics. Springer., 2011; 397 P.


Hamadan Regional Water Authority., Feasibility studies water and potential utilization of surface water Hamadan province, 2008; 69 P. (In Persian).


Singh, A.K. Modeling Land use/ Land cover Changes Using Cellular Automata in Geo-Spatial Environment, MSC Theses, Netherland.2003; 58 P.


پیش بینی تغییر کاربری با استفاده از مدل تلفیقی

کلمات کلیدی: روش حداکثر احتمال، ماسین بردار پشتیبان، مدل تلفیقی ساری

چکیده
با استفاده از توصیف ماهورهای لندسپ برای سال‌های 1992، 2000، 2013 و مدل CA- Markov، تغییرات کاربری اراضی برای جوهر اخیر سد اکائان شبه‌سازی شد. در اینجا دو روش طبقه‌بندی "حداکثر احتمال" و "ماسین بردار پشتیبان" استفاده شد. نتایج نشان داد اگرچه ضریب کاپا و درصد مثبت کلی دو روش بالاست، اما از نظر برخی روش حداکثر احتمال در تفکیک کاربری‌ها به خصوص مسکونی و اراضی دیم توسط ماسین بردار پشتیبان ناموفق بود. استفاده CA- Markov و فیلتر Markov به یک‌تایی نتایج حاصل از ماسین بردار پشتیبان برای مدل Markov منظر پیش‌ینی نشان داد کاربری اراضی سال 2034 استفاده شد. نتایج نشان داد نقشه شبیه‌سازی شده در سطح 84٪ با داده‌های موجود مطابقت دارد. همچنین، مناطق مسکونی از حدود 82٪ در سال 2034 به 100 کیلومتر مربع در سال 2013 به 44٪ کاهش می‌یابد. نتایج این تحقیق کاربردی بودن مدل CA- Markov برای شبیه‌سازی تغییر کاربری اراضی را نشان می‌دهد.

یافته‌ها
1- داشته‌اندگی، گروه مربوط به اخیرترین دانشگاه، دانشگاه منابع طبیعی، محیط زیست، دانشگاه، ملاصدرا، ایران
2- دانشجوی دکترای، گروه مخیداری، دانشگاه منابع طبیعی، دانشگاه علوم کشاورزی و منابع طبیعی ساری، ساری، ایران

