Gorgan University of Agricultural Sciences and Natural Resources

Environmemtal Resources Research (ERR)

Print ISSN: 2783-4832
Online ISSN: 2783-4670

Modelling land use change using integrated Cellular Automata and Markov Chain Model (Case study; Azadshahr County)

Chooghi Bairam Komaki[1*], Hamidreza Asgari[2], Habib Nazarnejad[3], Mohammad Alinezhad[4]

[1] Assistant Professor, Department of Arid Zone management, Faculty of Rangeland and Watershed management, Gorgan University of Agricultural Sciences and Natural Resources, Gorgan, Iran

[2] Associate Professor, Department of Arid Zone management, Faculty of Rangeland and Watershed management, Gorgan University of Agricultural Sciences and Natural Resources, Gorgan, Iran

[3] Associate Professor, Department of Watershed management, Faculty of Rangeland and Watershed management, Gorgan University of Agricultural Sciences and Natural Resources, Gorgan, Iran

[4] Ph.D student, Department of Arid Zone management, Faculty of Rangeland and Watershed management, Gorgan University of Agricultural Sciences and Natural Resources, Gorgan, Iran

Article Info

Article type:
Research Article

Article history:
Received: September 2021
Accepted: January 2023

Corresponding author:
komaki@gau.ac.ir

Keywords:
Landsat
Object-Oriented Classification
Support-Vector Machine
Markov Chain
Azadshahr

Abstract

In this research, land-use changes from 1998 to 2009 were investigated for Azadshahr County, using Landsat 5 imageries and integration of the Markov chain and Cellular Automata methods. Using the object-based support-vector-machine image classification method, land-use maps were classified into three major categories, namely agriculture, forest and built-up areas for the years 1987, 1998 and 2009; with overall accuracies being 91.0% (1987), 91.0% (1998) and 88.8% (2009), and the respective Kappa values of 86.5% (1987), 86.5% (1998) and 83.2% (2009). The built-up areas showed the greatest change increasing 2.02% and 2.17% in the periods 1987-1998 (first period) and 1998-2009 (second period), respectively. During the first period, forest area shrunk by approximately -1.80%. However, as a result of the afforestation projects during 1998-2009, forest area increased 1.59%, while over the 22-year period, the total area of forest reduced by -0.21%. Agricultural areas on the one hand shrunk in favor of the built-up areas, and on the other, increased by conversion of forest, making a total reduction of -0.22 and -3.75% for the first and second period, respectively. The land-use pattern of 2020 was simulated using the MULOSCE extension of the QGIS software based on the integrated cellular automata and Markov chain technique. For this period, we encountered a 0.62% increase in built-up areas, and 0.48% and 0.15% reduction in agricultural fields and forest lands, respectively.

Cite this article: Komaki, Ch.B., Asgari, H.R., Nazarnejad, H., Alinezhad, M. 2023. Modelling land use change by an integrated Cellular Automata and Markov Chain Model (case study; Azadshahr County). *Environmental Resources Research*, 11 (1), 99-108.

DOI: 10.22069/IJERR.2023.13458.1218
Publisher: Gorgan University of Agricultural Sciences and Natural Resources

Introduction

There is a growing recognition of the significance of land use and land use changes for environmental sustainability, both regionally and globally. The intricate interplay of vegetation, geology, and land-use management defines climate and hydrological processes. These processes collectively impact the ecological capacities for food and water production (Saghafian et al., 2006). Furthermore, land use change serves as an indicator of the consequences of human activities on biodiversity reduction and land degradation. It encompasses interventions that directly influence services, conditions, and production (Badjana et al., 2015).

Accurate land use/cover assessments are crucial for understanding environmental characteristics and processes. Producing precise land use/cover maps is essential because any error at this stage can propagate into subsequent analyses, affecting the reliability of results (Deilmai et al., 2014). Land use classification serves as the initial step in land use assessment (Jensen, 1995). The integration of remote sensing and GIS has proven highly practical and reliable in providing spatiotemporal information for sustainable land and natural resources management. Recent advancements, such as global positioning systems, improved spatial analysis methods, more accurate satellite sensors, and advancements in spatiotemporal analysis theories, have further enhanced this integration (Jensen et al., 2009).

Land use classification in remote sensing is categorized into supervised, unsupervised, and object-based techniques. In supervised classification, cells are semi-automatically categorized based on spectral signatures derived from training samples and user expertise (Richards and Jia 2006). In contrast, unsupervised classification assigns spectral classes (clusters) in multispectral images to different classes without user input (Richards and Jia, 1999).

The supervised image classification techniques include the maximum likelihood and support vector machine (SVM) methods (Qixia et al., 2014). Numerous studies have indicated that the SVM method demonstrates higher capability and accuracy compared to the maximum likelihood method. The latter method faces limitations when dealing with complex land forms and covers, while the SVM method effectively handles such images (Qingsheng et al., 2014; Colgan et al., 2012; Motkan and Hajeb, 2013; Mahini et al., 2012; Alimohamadi et al., 2009; Ranjbar et al, 2014; Alavipanah et al., 2007; Komaki and Alavipanad 2005; Molamehr Alizadeh et al., 2005).

Apart from land use classification, it is now possible to simulate past land use change trends into the foreseeable future (Aspinall, 2004). Various simulation techniques, such as stochastic models, Markov chains, optimization models, dynamic simulation tools like cellular automata (CA), and factor-based simulation tools, offer ways to address this. CA, for instance, has been successfully applied to assess land-use changes by representing both spatial and temporal dynamics. However, CA models may not be well-suited for urban environments due to their oversimplified approach. To improve these models, integrations of CA and the Markov Chain have been proposed. The Markov Chain computes conversion probabilities of different classes in a map using transition matrices, while CA adds spatial elements and user knowledge to predict the geographic location of land use/cover classes (Khoshgoftar and Talei, 2010).

In this study, we considered a time frame from 1987 to 2009 with 11-year intervals and conducted land use classifications on Landsat 5 satellite images using the SVM method. We analyzed the trend in land use changes using the CA-Markov model and extrapolated the results to the year 2020. The research area, located in Azadshahr County, Golestan Province, Iran, features complex landforms and exhibits significant evidence of land conversions during the study period, making it an ideal case study.

Materials and Methods

Study area

The study area is part of the Azadshahr County of Golestan Province, Iran, which spans over the northern section of the county. Geographically, the study area spans over 55° 41′ 00″ to 55° 08′ 37 ″ eastern longitudes and 37° 02′ 13 ″ to 37° 05′ 39″ northern latitudes (Figure 1).

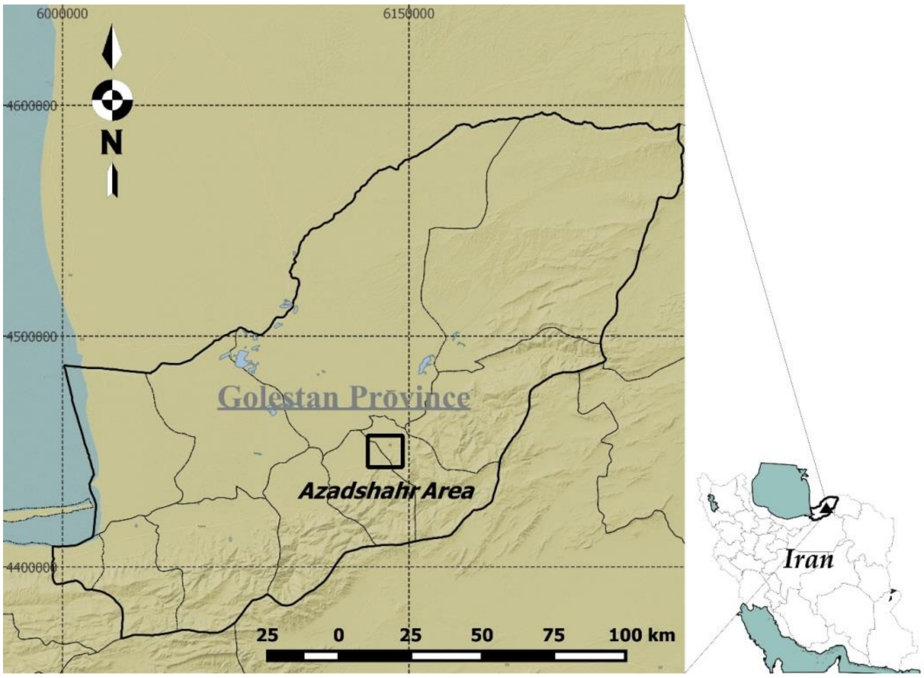

Figure 1. Location of the study area in relation to the Golestan Province and Iran boundaries

With the exception of urban areas, the northern plains predominantly serve agricultural purposes, while the southern regions consist mainly of mountainous terrain covered by forests. According to the Köppen climate classification, this area experiences a warm and temperate climate, characterized by an average annual temperature of 17.7°C and an annual average precipitation of 600 mm. The highest point is situated at an elevation of 1409 meters above sea level in the Elburz Mountains, while the lowest point is at 79 meters above sea level. As per the censuses conducted in 2006, 2011, and 2016, Azadshahr County has been inhabited by 82,251, 91,767, and 96,803 people, residing in 28,965 households. Over the past few decades, three significant historical events, including land reforms, an 8-year war, and development plans, have resulted in changes in land use. This county has experienced a relatively higher rate of population migration to urban areas compared to others. The primary economic activities encompass agriculture and animal husbandry in villages, while services dominate in towns. Widespread signs of soil degradation and erosion are evident, largely attributed to population pressure and inadequate agricultural management practices.

Input data

Landsat 5 TM data for 10 August 1987, 8 August 1998, and 21 July 2009 were acquired. These dates were selected as they were covered with less than 10% cloud. Cloudless images are hard to obtain over the northern part of Iran except for this month which offers more appropriate images to help classify ground features. At this time, most agriculture fields could be easily differentiated from forest lands.

Land use classification

To identify and classify land use classes, the object-oriented SVM classification was used; the procedure is illustrated in Figure 2.

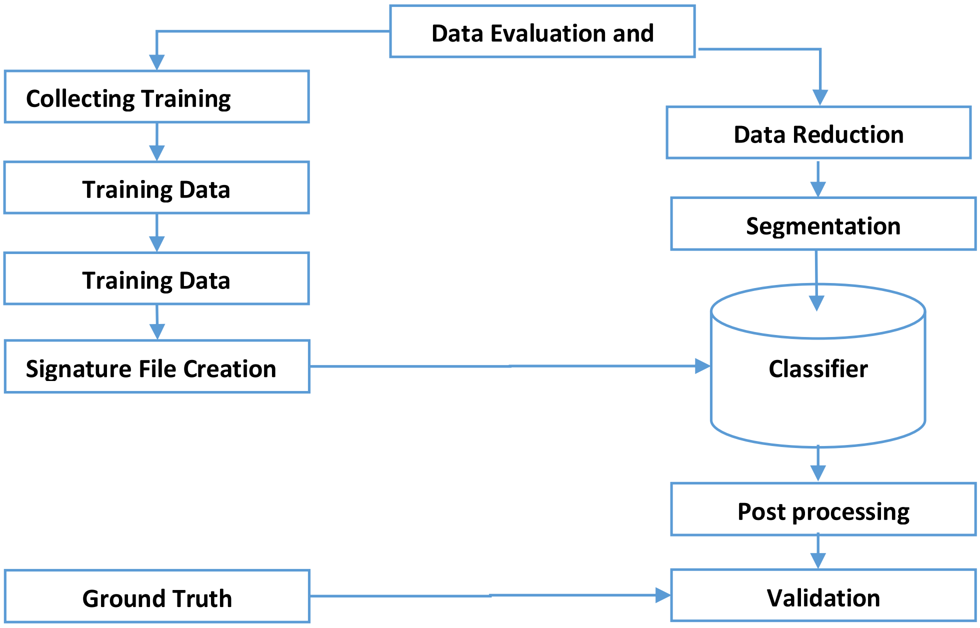

Figure 2. An overview of the supervised image classification method

The initial step in land use classification involves data evaluation and preprocessing. This process includes testing for normality in spectral bands and training samples using histograms. Principal Component Analysis (PCA) is employed to reduce data redundancy and achieve data compression. Subsequently, through image segmentation, pixels with similar spectral characteristics are grouped into distinct classes based on their spectral similarities, revealing various patterns based on pixel grouping shapes.

To identify classes and generate signatures, sets of training data are sampled. These training data sets are then assessed to ensure that the collected samples are sufficient. This assessment can be performed using distribution curves and various statistics such as the mean and variance. Once the classes have been edited and data adequacy is confirmed, a signature file is created to store this information.

In the following stage, different objects are distinguished and classified based on an array of features, including color, means, standard deviation, compactness, and rectangularity. Often, classification techniques result in a number of undesired cells and patches that need to be eliminated or merged into a dominant neighboring cluster.

Accuracy Assessment

To assess the overall accuracy of the classification technique, a confusion matrix is constructed using both the classified map and the ground truth data obtained through field surveys. In this study, a ground truth map was created by utilizing high-resolution satellite imagery, field surveys, and cross-referencing with Google Earth. An overall accuracy threshold of 85% is considered suitable, with lower accuracies deemed unacceptable (Mas, 1999). Overall accuracy measures the consistency between two maps: the classified map and the ground truth map.

To evaluate accuracy distribution among different classes, two indices, omission and commission errors, are employed, also referred to as producer's and user's accuracies. For this research, model performance verification involved generating a total of 500 samples using an equalized stratified randomized sampling approach, aiming to maintain an equal number of samples for each class. The study identified three land use classes: built-up areas, agricultural fields, and forest lands, with 167 samples generated for each class, totaling 501 samples. Finally, the Kappa coefficient was utilized to ensure that the achieved accuracy was not merely the result of chance.

CA-Markov model

The CA-Markov model represents an integrated approach that combines the Markov Chain and Cellular Automata techniques to monitor and simulate land-use changes. The Markov Chain Model is a fundamental simulation method for modeling stochastic processes, where the future state of a system is determined by its current state. Initially, the Markov Chain Model considers conversion probabilities for various land-use classes in the form of a probability matrix. However, the output of the Markov Model (land-use state change matrix) lacks spatial information and does not provide details about the geographic distribution of land-use classes. To predict the spatial allocation of land uses (land cover/use map) for future periods, the Cellular Automata technique has been integrated into the basic model. Essentially, Cellular Automata enhance the Markov Chain Model by incorporating spatial information (Khoshgoftar and Talei, 2010).

Land use data

Land use maps from 1987, 1998, and 2009 were utilized to calculate the transition matrix and project land use patterns for the year 2020. Additional factors, including proximity to roads, distance to rivers, and topography, were incorporated as key variables influencing land use changes. In this study, the Multi-Criteria Evaluation and Logistic Regression methods were employed to generate these layers.

Implementation of the CA-Markov Model

The MOLUSCE, as an open-source extension to the QGIS, was employed to simulate land use changes. After preparing the input layers, they were introduced to the modelling environment of the MOLUSCE. These included the land use maps of the years 1987, 1998 and 2009 with three distinct classes of built-up areas, agricultural fields and forest lands. The land use maps of 1998 and 2009 were respectively used to verify the outcome of the modelling tool. Next, the state change matrix and the leading variables were introduced to eventually predict future state of the land use pattern.

Results and Discussion

As mentioned earlier, the primary variables considered were distance to roads, distance to rivers, and the Topographic Ruggedness Index (TRI). The first two layers were generated using the Euclidean distance function in ArcGIS 10.4. The Topographic Ruggedness Index (TRI), developed by Riley and Eliot in 1999, quantifies terrain roughness based on elevation differences between individual pixels and their neighboring cells in a raster layer. It is computed by averaging the sum of squares of elevation differences between the central cell and its eight adjacent pixels, with the final TRI value obtained by taking the square root of the total. The correlation between these layers was assessed, resulting in correlations of 0.57 between distance to roads and distance to rivers, 0.55 between distance to roads and TRI, and 0.48 between distance to rivers and TRI.

Land use classification at different time periods

In total, three main land use categories were identified in the study area: built-up areas (with one subcategory), agricultural fields (with four subcategories), and forested land (with two subcategories). These land use categories were determined using the object-based SVM method applied to Landsat 5 satellite imagery. To validate the results, 501 sampling points from the ground truth map, generated through an equalized stratified randomized sampling procedure, were employed. High values of both producer's and user's accuracies indicated the reliability of the classification method. The overall accuracy and Kappa index for 1987 were 90.42% and 85.63%, respectively; for 1998, they were 91.02% and 86.53%; and for 2009, they were 88.82% and 83.23%. Changes in land use between 1987 and 1998, measured in hectares and as a percentage, are presented in Table 1. The spatial distribution of land use classes can be observed in Figure 3.

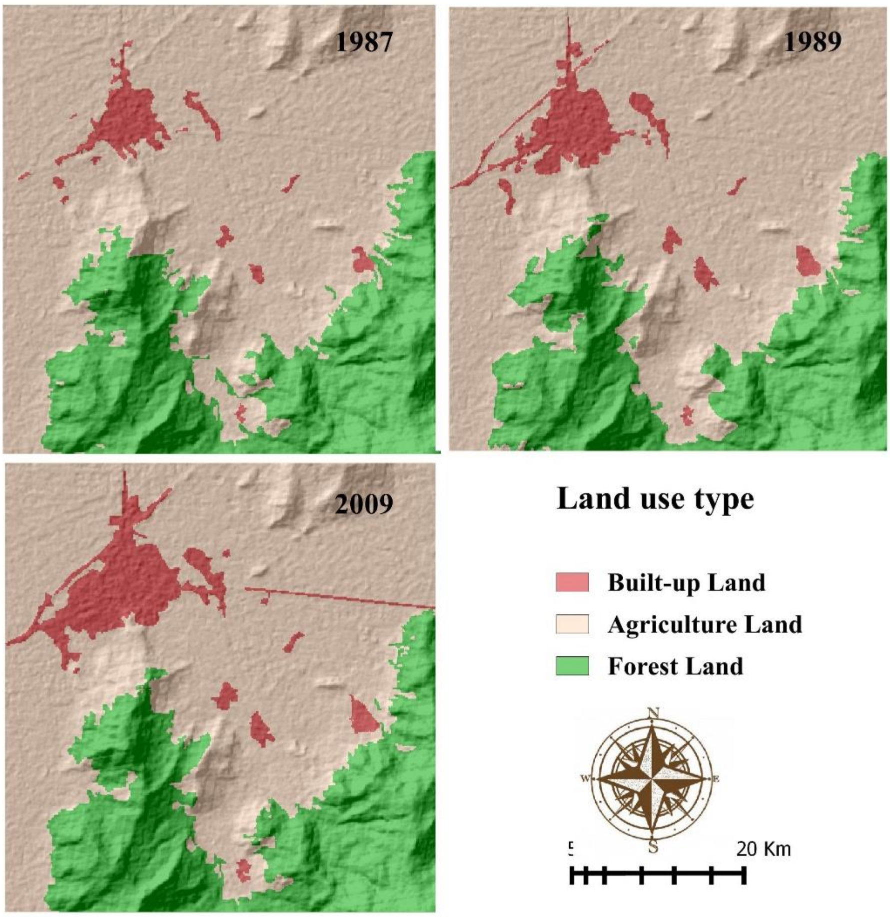

Figure 3. land use classes in 1987, 1998, and 2009 in Azadshahr area using the object-based SVM classification method

Table 1. Land use changes for the first (1987-1998), second (1998-2009) and to whole study period (1987-2009) in hectares and percentages

Land use	Change in area (ha)			Change in area (%)		
	1987-1998	1998-2009	Whole Period	1987-1998	1998-2009	Whole Period
Built-up area	184	197	381	2.02	2.17	4.19
Agriculture Fields	-20.34	-341.55	-361.89	-0.22	-3.75	-3.98
Forest land	-163.62	144.54	29.92	-1.80	1.59	-0.21

As can be seen in Table 1 and Figure 2, the change in built-up areas has been the largest with 4.19% 381 ha increase over the whole period. This increase amounts to 170.36% growth compared with the year 1987 (i.e. more than 2 times that of 1987). Based on the results, the greatest changes in forest land could be detected between 1987-1998 by about -1.8% reduction which amounts to a 164 ha since 1987. This reduction, however, is followed by 1.59% increase between 1987-1998, when several deforestation projects were undertaken to counter the decreasing trend in forest lands. However, over the whole period, approximately 0.21% of the forest area has decreased. There has been around 2-4% increase in built-up areas during the period at the expense of -1 to -4% reduction in agricultural fields and forest lands.

Land conversion and Markov transition

Based on Table 2, it is possible for agricultural fields to be converted into built-up areas, but the reverse transformation is less likely to occur. The probability of agricultural fields remaining unchanged over the 11-year period (1987-1998) is approximately 96.06%, with a 3.02% chance of conversion into built-up areas. During the same period, the likelihood of agricultural fields transitioning into forest lands was estimated at 0.92%. The probability of no conversion for forest land was estimated at 92.92%. However, there is an 8.08% probability that forest land could be converted into agricultural fields.

For the second period, the probabilities of agricultural fields converting into built-up areas and forest lands were estimated at 3.32% and 3.04%, respectively, with a no-conversion probability of approximately 93.64%. During this period, the likelihood of forest lands transitioning into built-up areas was minimal (below 0.01%), but it was comparatively higher for conversion into agricultural fields, at 1.59%. The no-conversion probability was also calculated at 98.40%. The values underlined in Table 2 represent the highest conversion probabilities during various time intervals.

As observed, between 1987-2009, the transition probability for agricultural fields into built-up areas has roughly doubled. Regarding forest land, during the first period, the conversion of forest lands into built-up areas was more probable. However, over time and with the expansion of agricultural fields, the chance of a forest patch being converted into a built-up area decreased by one-fourth. At the same time, the conversion probability of forest lands into agricultural fields increased by almost 4%. This suggests that over time, the conversion of forest lands into built-up areas has shifted from direct to indirect, with agricultural fields playing a significantly increased role in overall land use change. In summary, it appears that forest land must first be converted into agricultural fields before transitioning to urban areas.

CA-Markov Chain simulation results

The results of the land use change simulation for the year 2020, conducted using the CA-Markov chain model, are depicted in Figure 4. The transition probabilities were simulated utilizing the Artificial Neural Network function within the MOLUSCE extension in QGIS. These results indicate that adhering to the baseline land use change scenario will lead to a significant expansion of built-up areas, increasing from 648 to 705. However, this expansion comes at the cost of a substantial reduction in both agricultural fields and forest lands.

Table 2. Transition probability matrix of land use in different periods in Azadshahr area

Land use type	Year	Agriculture Land	Built-up Land	Forest Land
Agriculture Land	1987-1998	**0.9606**	0.0302	0.0092
	1998-2009	**0.9364**	0.0332	**0.0304**
	1987-2009	**0.9214**	**0.0631**	0.0155
Built-up Land	1987-1998	0.0000	**1.0000**	0.0000
	1998-2009	0.0000	**1.0000**	0.0000
	1987-2009	0.0000	**1.0000**	0.0000
Forest Land	1987-1998	0.0000	0.**0808**	**0.9292**
	1998-2009	0.0159	0.0001	**0.9840**
	1987-2009	**0.0417**	0.0002	**0.9581**

The summary of findings from this study for different land uses and conversions is presented in Table 3. It is evident that agriculture and forest land uses are experiencing a declining trend. While forested areas may witness a slight increase up to the year 2690, this change is expected to be minimal. There is a high likelihood of urban expansion, suggesting a significant probability of direct conversion of agricultural fields into urban areas. Additionally, in the proximity of villages and townships, there is a considerable risk of converting forested lands into urban areas. The reduction in agricultural land will undoubtedly have a detrimental impact on the local economy.

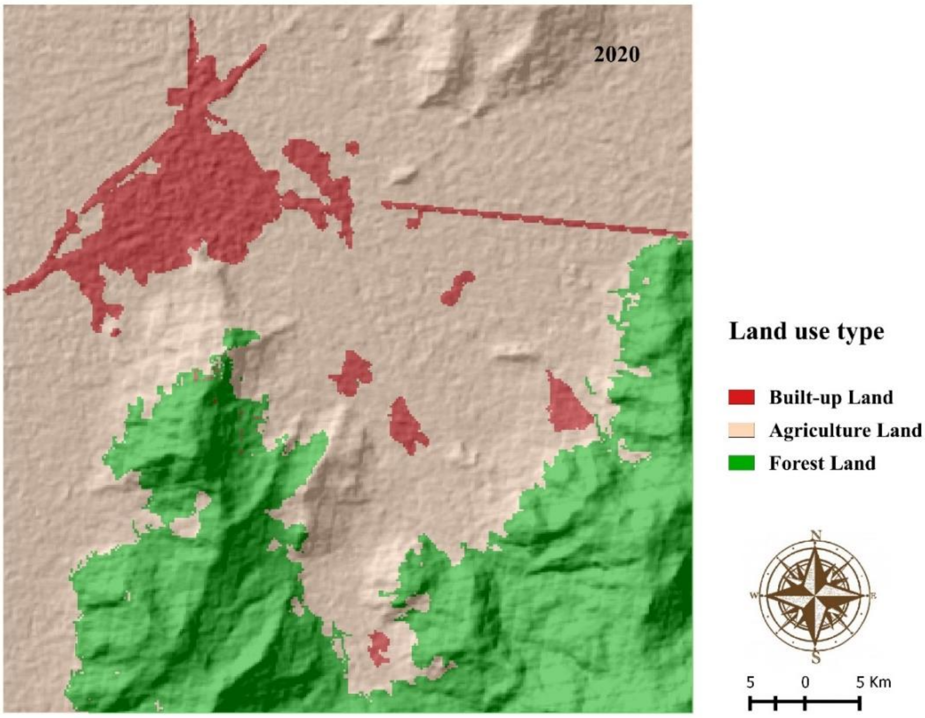

Figure 4. Simulated land use map of the year 2020

Furthermore, deforestation is likely to lead to land degradation and desertification, exacerbating the economic challenges already faced by the region. A noteworthy finding of this research is that the current management approach appears insufficient to address these looming threats. Consequently, it is strongly recommended that management scenarios and strategies be revised, with a focus on adopting mitigation or prevention measures to counteract these challenges.

Table 3. land use changes during 1987-2020 in Azadshahr area

Year	Agriculture Field	Forest Land	Built-up Area
1987	6116	2723	261
1998	6095	2559	446
2009	5748	2704	648
2020	5704	2690	705

Conclusion

The results suggest a significant shift in land use within the area, which is not surprising given the population growth and economic development that have placed considerable pressure on the land. Urban expansion and the proliferation of industrial facilities have led to the conversion of land, particularly in the agricultural sector. Furthermore, forested areas have been increasingly transformed into agricultural fields, driven by short-term economic gains. These conversions have resulted in reduced agricultural production and hindered overall economic growth. Notably, the results highlight a more pronounced trend of land conversion during the initial period, which was somewhat mitigated during the second period through the cultivation of needle-leaf species in previously deforested areas. However, these remedial actions have had adverse consequences for local fauna due to changes in their habitats.

The findings of this research affirm the effectiveness of the approach employed for land use identification. From a statistical perspective, an overall accuracy rate of 85% and a kappa index exceeding 80%

substantiate this claim. While the selection of images with suitable intervals is generally preferred, atmospheric corrections are of paramount importance. During wet years, the presence of substantial vegetation can make it more challenging to delineate boundaries between forested lands and agricultural fields. In summary, it should be noted that remote sensing and GIS are invaluable tools for data collection and processing for managers. The ability to produce accurate, large-scale maps rapidly is an immensely valuable resource. These data are not only verifiable but can also be updated with minimal time and effort.

References

Alavipanah, S.K., Ch. B Komaki, A. Goorabi., and Matinfar, H.R. 2007. Characterizing Land Cover Types and Surface Condition of Yardang Region in Lut Desert (Iran) Based upon Landsat Satellite Images. World Applied Sciences Journal. 2 (3), 212–228.

Alimohamadi, A., Motkan, A. A., Ziaeian, P., and Tabatabaei, H. 2009. Comparison of pixel-based classification, object-based classification methods and decision tree mapping for forest types mapping, using data from remote sensing (Case Study: Astara Forest). Applied research in geographic (Geographical Sciences). 10 (13), 7 -26.

Aspinall, R.. 2004. Modelling Land Use Change with Generalized Linear Models—a Multi-Model Analysis of Change between 1860 and 2000 in Gallatin Valley, Montana. Journal of Environmental Management, Modelling land use change and environmental impact. 72 (1–2), 91–103.

Azizi, M. 2014. Estimation of accuracy from a combination of cellular automata, SVM and ANFIS models in predicting the physical development of the city based on GIS Case Study: Shiraz - District Six. M.Sc Thesis, Hormozgan University. http://ganj.irandoc.ac.ir/articles/805458.

Badjana, H.M., Helmschrot, J., Selsam, P., Wala, K., Flügel, W.A., Afouda, A. and Akpagana, K., 2015. Land cover changes assessment using object-based image analysis in the Binah River watershed (Togo and Benin). Earth and Space Science. 2(10), 403-416.

Colgan, M. S., Baldeck, C. A., Féret, J. B., and Asner, G. P. 2012. Mapping Savanna Tree Species at Ecosystem Scales Using Support Vector Machine Classification and BRDF Correction on Airborne Hyperspectral and LiDAR Data. Remote Sensing. 4 (11), 3462–80. doi:10.3390/rs4113462.

Deilmai, B.R., Ahmad, B.B. and Zabihi, H. 2014. Comparison of two Classification methods (MLC and SVM) to extract land use and land cover in Johor Malaysia. In IOP Conference Series: Earth and Environmental Science. 20(1), 012052.

Jensen, J. R. 1995. Introductory Digital Image Processing: A Remote Sensing Perspective. 2nd ed. Prentice Hall.

Jensen, J.R., Im, J., Hardin, P. and Jensen, R.R. 2009. Image Classification. The SAGE Handbook of Remote Sensing. 269–281.

Khoshgoftar, M.M., and Talei, M. 2010. Urban growth modeling in Tehran using CA-Markov Model. Remote Sensing and GIS in Iran. 2 (26), 17-34.

Komaki, Ch.B., and Alavipanah, S.K. 2005. An assessment of the spectral resolution in data classes from the Lut Desert using satellite data. Geographical Research. 37 (54), 13-28.

Mahini, A. S., Nadali, A., Feghhi, J., and Raizi, B. 2012. Classification of forest areas in Golestan province using maximum likelihood method from ETM + satellite sensor in 2001. Journal of Environmental Science and Technology. 14 (3), 56-47.

Mas, J.F. 1999. Monitoring Land-Cover Changes: A Comparison of Change Detection Techniques. International Journal of Remote Sensing. 20 (1), 139–152.

Molamehr Alizadeh, F., Janati, M., and Shayan, S. 2005. Feasibility of remote sensing (RS) mapping of landforms and its role in environmental planning. Humanities teacher, 9 (443), 49-80.

Motkan, A.A., and Hajeb, M. 2014. Comparison between the ability of support vector machine, neural network and maximum likelihood classification methods for improving LiDAR data filtering algorithms. Iranian Journal of Remote Sensing and GIS. 5 (2), 1-16.

Qingsheng,L., Guo, Y., Liu, G., and Zhao, J. 2014. Classification of Landsat 8 OLI Image Using Support Vector Machine with Tasseled Cap Transformation. In 2014 10th International Conference on Natural Computation (ICNC), 665–69. Xiamen: Natural Computation (ICNC). doi:10.1109/ICNC.2014.6975915.

Qixia, M., Guo, H., Dong, P., Liu, G., and Shi, R. 2014. Support Vector Machines and Maximum Likelihood Classification for Obtaining Land Use Classification from Hyperspectral Imagery. In Proceedings of the International Geoscience and Remote Sensing Symposium (IGARSS), Quḗbec, Canada, 2878–2881. ftp://ftp.legos.obs-mip.fr/pub/tmp3m/IGARSS2014/pdfs/0002878.pdf.

Ranjbar, P., Shataei, Sh., and Mahini, A.S. 2014. The comparison of forest area changes detection methods using data from TM and ETM+ sensors. Wood Research Science and Technology.20 (3), 1-22.

Richards, J. A., and Jia, X. eds. 2006. Image Classification Methodologies. In Remote Sensing Digital Image Analysis: An Introduction, 295–332. Berlin, Heidelberg: Springer Berlin Heidelberg. DOI:10.1007/3-540-29711-1_11.

Richards, J.A. and Jia, X., 1999. The Interpretation of Digital Image Data. In Remote Sensing Digital Image Analysis: An Introduction, 67–82. Berlin, Heidelberg: Springer Berlin Heidelberg. DOI:10.1007/3-540-29711-1_3.

Riley, S.J., DeGloria, S.D., and Elliot, S.D. 1999. A Terrain Ruggedness Index That Quantifies Topographic Heterogeneity. Internation Journal of Science 5 (January), 23–27.

Saghafian, B., Farazjou, H., Sepehr, A., and Najafinejad, A. 2006. The effect of land use changes on flood in Golestan dam basin. Iran Water Resource Research. 2 (1), 18-28.