

APPLICATION OF CA-MARKOV MODEL AND LAND USE/LAND COVER CHANGES IN MALACCA RIVER WATERSHED, MALAYSIA

HUA, A. K.

*Department of Environmental Sciences, Faculty of Environmental Studies, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia
e-mail: angkeanhua@yahoo.com*

(Received 28th Mar 2017; accepted 6th Jun 2017)

Abstract. Modeling the effects of past and current land use composition on surface water quality of the Malacca River watershed provides valuable information for environmental and land planning. The Land use land cover (LULC) spatio-temporal changes in watershed are classified as non-industrial area, industrial area, vegetation area, open space area, water bodies, as well as farming area. This is compared with the future spatial pattern simulated using CA-Markov model to evaluate qualitative and quantitative changes of LULC over time (between 2001, 2008, 2015, and 2029). This analysis consisted of the entire Malacca River watershed area. Changes may provide information on determinant of river water quality in future. Based on LULC monitoring and future simulation approach, the results may be used in an early warning system by demonstrating trend and consequences of changes in watershed. Continuous downward trends for vegetation area, open space area, and water bodies are due to expansion in non-industrial area, industrial area, and farming areas that increase contamination in the Malacca River. The findings demonstrate the need for more regulation in land use policy design, planning, and development for the protection of river water quality.

Keywords: *land use land cover changes, CA-Markov model, land-sat, quantitative and qualitative simulation*

Introduction

Anthropogenic activities negatively affect the conditions of the world's water. Various studies have estimated that half of the world's water quality has decreased due to human activities (Aris et al., 2013; Najar and Khan, 2012), and this condition has continuously worsened until today (Hua, 2017; Rosli et al., 2015). According to Hua et al. (2016), the main characteristic to cause contamination and detected as pollutant sources are residential, industrial, and agricultural activities. This matter is confidently supported by a majority of researchers (Aris et al., 2013; Hua, 2017; Hua et al., 2016; Rosli et al., 2015; Wan et al., 2014). Human alteration of cities engineered toward the land use land cover (LULC) change drastically results in negative impact on water quality, especially in rivers (Wan et al., 2014; Yang et al., 2012). Based on the evidence, contaminated rivers could threaten aquatic animal, reduce the availability of quality water. This will affect the flora and fauna, as well as affecting the climate at a regional scale (Carey and Fulweiler, 2012; Wan et al., 2014). Hence, it has become essential to obtain an understanding of how recent trajectories of land use will manifest in the future (Parsa et al., 2016; Subedi et al., 2013).

LULC change models are usually applied to detect location of the changes occurred or will potentially occur (Halmy et al., 2015; Yagoub and Al Bizreh, 2014). By determining the factors of the changes, the models benefit through providing probabilistic prediction of possible changes may occur (Halmy et al., 2015). Technically, change analysis is carried out

using historical land use data in which the past land transformations and transitions were assessed. The transition trend is incorporated with environmental variables to provide an estimate of future scenarios (Behera et al., 2012; Yagoub and Al Bizreh, 2014). Prediction of LULC changes is important to understand and highlight the potential modifications and alterations that might be happen to the landscape in the near future. Typically, changes occur due to increasing population, distance to roads or other facilities, type of the soil, environmental issues, and so on. LULC models are usually used to assess the cumulative impact of land use changes and develop the future scenarios (Behera et al., 2012; Halmy et al., 2015), which is important in providing support and help to land use planners, resource managers, and conservation practitioners in making decisions (Halmy et al., 2015; Memarian, et al., 2013; Parsa et al., 2016). Simultaneously, prediction LULC changes model provide advantages in different applications, such as urban planning through modeling rural development and urban growth (Arsanjani et al., 2013); selecting conservation priority areas and setting alternative conservation measures (Subedi et al., 2013); studying dynamics of shifting cultivation (Halmy et al., 2015); as well as simulating the land use range dynamics under different watershed area (Behera et al., 2012; Memarian, et al., 2013).

Markov chain analysis works on the assumptions of physics in which the probability of a system being a certain state at certain time can be determined if the earlier time state is known (Arsanjani et al., 2013; Behera et al., 2012; Halmy et al., 2015). In other words, Markov chain analysis uses statistical spatially dependent land use data required by logistic regression (Arsanjani et al., 2013). This method is based on developing a transition probability matrix of land use change between two different dates derived from observation, which will be used to provide estimations of the probability that each pixel of certain LULC classes will be transformed to another class or remain in the same class (Yagoub and Al Bizreh, 2014). Nevertheless, Markov chain analysis has several issues as the methods only provide short-term projection (Sinha and Kumar, 2013) and are not spatially explicit (Halmy et al., 2015), as it does not provide spatial distribution of the changes that might occur to understand more about the potential impact of the projected changes. These issues can be overcome through suggestion of integrated modeling approaches with other different dynamic and empirical methods considered suitable for modeling land use changes processes (Halmy et al., 2015; Sinha and Kumar, 2013).

The Markov chain model can be integrated with the cellular automata model (or CA-Markov model), and these models have been widely used in different scale especially involve with modeling and predicting the land use change (Arsanjani et al., 2013; Subedi et al., 2013). The CA-Markov model is applied in this for the advantages of the combination of the stochastic aspatial Markov techniques with the stochastic spatial cellular automata method (Arsanjani et al., 2013). These include the prediction of two-way transitions among the availability of LULC classes rather than the Geomod technique that only concentrates on prediction of one way loss or gain in one classes (Halmy et al., 2015). Hence, transition-based models are the integration between aspatial Markov model with spatial cellular automata model to performed the regression-based models in predicting the land use change. This study has been carried out to provide information on the presence of LULC patterns in the Malacca River watershed and aims to analyze the changes of landscape trend pattern using the past LULC, the presented LULC, and future LULC using simulation modeling for the Malacca River watershed.

Materials and Methods

Study Area

Malacca state is located in southwest Peninsular Malaysia at 2°23'16.08"N to 2°24'52.27"N latitude and 102°10'36.45"E to 102°29'17.68"E for longitude. Malacca state has a catchment area of approximately 670 km² and contains an 80 km length of Malacca River flowing through Alor Gajah and Malacca Central (*Figure 1*). Malacca state also has a reservoir located between Alor Gajah and Malacca Central called the Durian Tunggal Reservoir; with a catchment of 20 km². This reservoir supplies water to Malacca residents. Increasing local population has led to increasing public facilities such as transport, healthcare, accommodation, sewage and water supply services (Hua, 2017; Rosli et al., 2015). Due to drastic population growth, rapid urban development in Malacca state also increased; especially demanding from the land use perspective. A majority of residential activities are centralized in the city, which extends about 10 km to the west, the east, as well as the north for 20 km. Land use changes is continuously develop until today. Indirectly, these actions have contributed to economic growth and social relationships, including impacting the environmental quality of the water in the Malacca River.

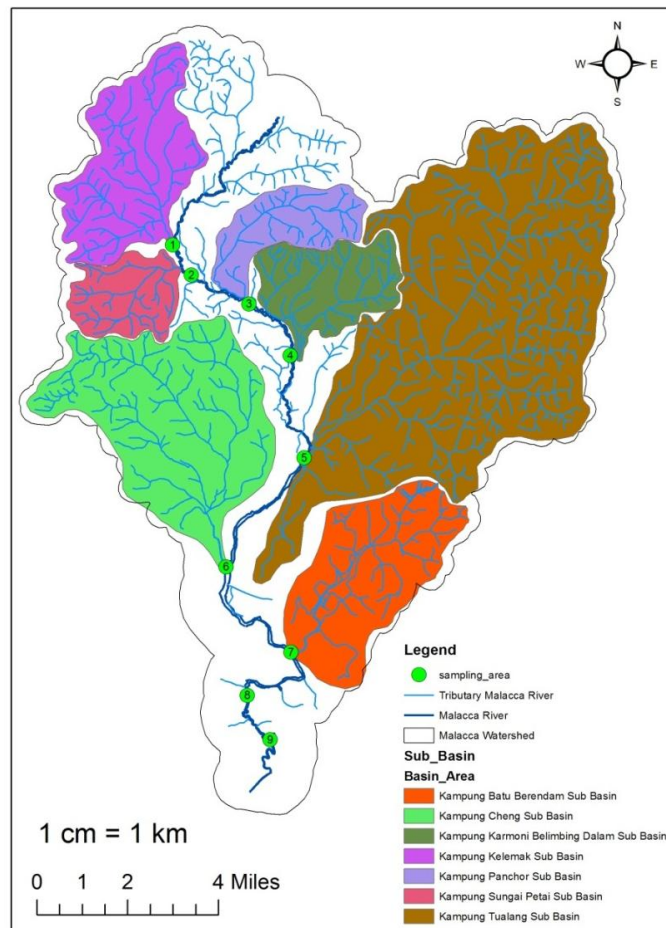


Figure 1. The catchment area with the river flow in Malacca state

Data Collection

Remotely sensed data have been widely used as a very cost effective mean to obtain geo-referred data and maps for evaluation and monitoring LULC. Three LULC data sets of the Malacca River watershed were obtain from USGS Earth Explorer dated 2001, 2008, and 2015 and used in this study. These three data sets are considered suitable with cloud-free spatial coverage and relatively high spatial and spectral resolution of satellite images. Remote sensing images of Landsat satellites were selected for the investigation of long-term variations in LULC types in the study area (*Table 1*). After the limitations and constraints regarding of the data were taken into account, a 14-years period or time-span was found appropriate for monitoring and evaluation of LULC dynamic. The steps of the research study are shown in *Figure 2*.

Table 1. Landsat data sources

Satellite	Sensor	Acquisition	Band Combination
Landsat 4-5	TM	2001	1,2,3
Landsat 7	ETM+	2008	1,2,3
Landsat 8	OLI	2015	2,3,4

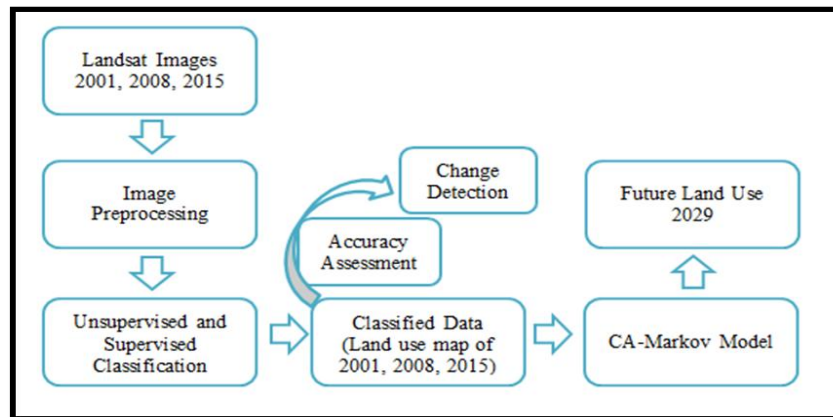


Figure 2. Demonstration of research methodology as a flowchart. The 2001 and 2008 maps used to simulated 2015 maps and the simulated 2015 map was used to validate the model. Then 2008 and 2015 maps were use as the input to produce simulated 2029 map

Data Pre-Processing and Image Classification

Preprocessing of satellite images is essential and aims at the unique goal of establishing a more direct linkage between data and the biophysical phenomena it represents (Parsa et al., 2016). Pre-processing is accomplished using ArcGIS version 10.0 for geo-referencing, mosaicking and sub setting of the image for the Area of Interest (AOI). Landsat 8 image underwent spatial sharpening using the panchromatic bands, which resulted in a 15m resolution. Meanwhile, Landsat 5 TM and Landsat 7 ETM+ images for 2001 and 2008 were in original 30m resolution. Further image processing analysis was carried out using

ENVI 4.0. The image was displayed in natural color composite using a band combination of 3, 2, 1 for Landsat 5 TM and 4, 3, 2 for Landsat 8. Maximum Likelihood supervised classification was performed using several selected regions, and Regions of Interest (ROI) were based on delineated classes of vegetation area, industrial and non-industrial area, water, open space area and farming area (Table 2).

Table 2. Classes delineated on the basis of supervised classification

Class Name	Description
Vegetation Area	Includes all agricultural lands and forest fields.
Non-Industrial Area	Includes all residential, commercial, administration, cemetery and transportation, as well as sewage treatment plant (include individual septic tank).
Industrial Area	Includes all industrial area.
Water Bodies	Includes all water bodies (river, lakes, gravels, stream, canals, and reservoirs).
Open Space Area	Includes all land areas that exposed soil and barren area influenced by human.
Farming Area	Includes all animal husbandry (or farming) area.

Accuracy Assessment

Accuracy assessments for the 2001, 2008, and 2015 images were carried out to determine the quality of information provided from the data. If the data are to be used for change detection analysis, it is important to conduct accuracy assessment for individual classification (Behera et al., 2012). Kappa tests are used to measuring the accuracy of classification as the test is able to account all elements in confusion matrix including diagonal elements (Halmy et al., 2015). The Kappa test is a measure between predefined producer rating and user assigned rating, which can be expressed in the formula as:

$$K = \frac{P(A) - P(E)}{1 - P(E)} \quad (\text{Eq. 1})$$

where $P(A)$ is the number of time the k raters agree, and $P(E)$ is the number of time the k raters are expected to agree only by chance (El-Kawy et al., 2011; Pontius and Millones, 2011). Meanwhile, user accuracy can be defined as the probability of a pixel on the image actually representing a class on the ground. Producer's accuracy indicates the probability a pixel being correctly classified and is mainly used to determine how well an area can be classified (Pontius and Millones, 2011). As described earlier, the six categories of classes which include vegetation area, non-industrial area, industrial area, water bodies, open space

area, and farming area, that have been delineated should have a minimum of 50 points for each considered category to increase the percentage of accuracy assessment (El-Kawy et al., 2011). Therefore, the accuracies of classification for 2001, 2008 and 2015 are 89.51%, 88.49%, and 92.21%, which have kappa statistics of 0.87, 0.85 and 0.90 respectively. According to Weng (2010), the minimum level for accuracy assessment in identification of LULC categories in remote sensing data should be at least 85%. Afterwards, the data is exported into an ASCII text file to enable for further analysis in ArcGIS version 10.

LULC Change Detection Analysis

In performing LULC change detection, the post-classification detection method is applied in the IDRISI Selva environment v.17, which involves two classified images to make a comparison to produce change information on a pixel basis. In other words, the interpretation between two image provide will provide changes “-from, -to” information. Classified images from two different data sets are compared using cross-tabulation in determining qualitative and quantitative aspects of changes for periods from 2001 to 2015. The magnitude of change and percentage of changes can be expressed in a simple formula as follows:

$$K = F - I \quad (\text{Eq. 2})$$

$$A = \frac{(F - I)}{I} \times 100 \quad (\text{Eq. 3})$$

where K is magnitude of changes, A is percentage of changes, F is first data, and I is reference data (Mahmud and Achide, 2012). Additional, prediction or estimation of LULC changes for 2029 will also use IDRISI Selva environment v.17. This research study uses LULC techniques in remote sensing to determine differences and define the percentage of land use changes within that time, as well as estimation for the next 14 years.

Markov Chain Model Analysis

The Markov chain model was presented by a Russian mathematician named Andrei A. Markov in 1970. This model was first used by Burnham for land use modeling (Mishra and Rai, 2016; Parsa et al., 2016). Markov chains are stochastic processes (Halmy et al., 2015; Subedi et al., 2013) and the matrices to show changes between land use categories (based on the basic core principle of continuation of historical development) (Koomen and Borsboom-van Beurden, 2011) and are often used in modeling and simulation changes and trends of LULC (Halmy et al., 2015; Mishra and Rai, 2016; Parsa et al., 2016). The homogeneous Markov model for prediction of land use changes can be mathematically presented as follows (Subedi et al., 2013):

$$L_{(t+1)} = P_{ij} \times L_{(t)} \quad (\text{Eq. 4})$$

$$P_{ij} = \begin{bmatrix} P_{11} & P_{12} & P_{1m} \\ P_{21} & P_{22} & P_{2m} \\ P_{m1} & P_{m2} & P_{mm} \end{bmatrix} \quad (\text{Eq. 5})$$

where $L_{(t)}$ and $L_{(t+1)}$ represent land use status at time t and $t + 1$ respectively. Including $\sum_{j=1}^m P_{ij} = 1 (i, j = 1, 1, 2, \dots, m)$ is the transition probability matrix in a given state. Practically in IDRISI Selva v.17, this study uses 2001 and 2015 map into Markov chain model to produce the transition matrix changes between the current 14 years, and the process is repeated onto map 2015 and 2029 for future land use to derive the transition matrix changes.

Cellular Automata (CA)

Cellular Automata (CA) was developed by Ulam in 1940 (Mishra and Rai, 2016; Parsa et al., 2016) for application in land use changes conceptually. Afterwards, CA was used by Tobler in geographical modeling (Arsanjani et al., 2013), and widely used in spatial model (Halmy et al., 2015) for forecasting future land use. CA consists of a grid or a raster space, a set of states characterizing the grid cells and a definition for the neighborhood arrangement of cells, a set of transition rules determine the state transitions for each of the cells as a function of the position of neighboring cells and a sequence of discrete time steps then updates composition and configuration of all the cell simultaneously (Arsanjani et al., 2013; Mishra and Rai, 2016; Parsa et al., 2016). The basic principle of CA is that the land use changes for any location (cells) can be explained by the current state and changes in neighboring cells (Koomen and Borsboom-van Beurden, 2011).

CA-Markov Chain Model

The ‘Cellular Automata’ and ‘Markov Chain’ models are considered to be advantageous for modeling land use changes (Mishra and Rai, 2016; Parsa et al., 2016). The issues involved when a Markov chain model lacks spatially referred output and transition probabilities may be accurate on a categorical basis, there are no specifications on spatial distribution of each land use category occurrence (Arsanjani et al., 2013). Cellular automata added into a Markov model lead to probable spatial transitions occurring in particular area over a period time (Subedi et al., 2013). In other words, the quantity of changes from the Markov Chain model then are made geo-referred and spatial through cellular automata (Mishra and Rai, 2016). The CA-Markov model uses Markov Chain analysis outputs, particularly the Transition Area file, to apply a contiguity filter to enable the development of other land use characteristics from time two into a later time period (Parsa et al., 2016). The CA is able to develop a weighting spatial on the particular areas which have approximately the same to the existing land use based on classes, and it is not random (Subedi et al., 2013). Hence, the CA-Markov model is considered a robust approach because of the quantitative estimation and the spatial and temporal dynamic it has for modeling the LULC dynamic (Arsanjani et al., 2013; Mishra and Rai, 2016; Parsa et al.,

2016; Subedi et al., 2013). GIS and remote sensing is easy to fit with CA-Markov model to facilitate tasks and reduce the cost and time needed (Mishra and Rai, 2016; Parsa et al., 2016). In this study, IDRISI Selva v.17 is used to predict the future LULC of study area on CA-Markov model. In other words, CA-Markov will use the 2001 and 2008 maps to produce a simulated 2015 map, which is important to validate with actual LULC of 2015 map through KIA (Kappa Agreement of Index) approach (Mishra and Rai, 2016; Parsa et al., 2016). Afterwards, the techniques are repeated using the 2008 and 2015 map to produce a simulated 2029 map.

Validating LULC Prediction Model

In order to avoid miscalculation, investigation between actual image and simulated image will be carried out, where the model's output was compared to a present or actual land use map. Comparing the predicted LULC map representing the 2015 LULC with actual LULC (map of 2015) was based on Kappa Index of Agreement (KIA) approach, which is widely used in validate LULC change predictions (Mishra and Rai, 2016; Parsa et al., 2016; Subedi et al., 2013). An additional map of Landsat TM 2008 will be used to help in validating process, before CA-Markov model can be applied for estimation of the next 14 years. The validation module is available in IDRISI Selva environment v.17 for this purpose.

Results

LULC maps assessment

The images show that the LULC maps significantly change from one class to another class (*Figure 3*). Based on general observation, drastic changes can be detected on water bodies between 2001 to 2015 [*Figure 3 (a) and (c)*], as most of these classes change from blue into green color and the area becomes smaller-scale. In other words, the water bodies coverage is suspected to transform into vegetation area, which refers as agriculture activities that carried out within the reservoir. Meanwhile, the non-industrial area is suspected to increase by a big margin, especially in the urban area. Majority classes that converted into non-industrial area are vegetation area and open space area. Only the small scale is subjected to the changes that occur in sub-urban and rural areas. At the same time, small scale shows an increase in industrial area, which is likely to be converted from open space area.

Continuously, the LULC map changes from 2015 to 2029 indicate that majority activities are converted from vegetation area into non-industrial and industrial area. At the same time, several areas on open space classes are also having converted into non-industrial area. The non-industrial area having big margin in conversion of development, followed by industrial area that have partially big scale in increasing the area for development purposes. Most of the development is occur in sub-urban and rural area. This rapid development shows positive impact towards the urbanization and modernization in Malacca state.

Validating LULC prediction model

In order to validate the LULC prediction given by the CA-Markov model, the simulated land use areas were used to compare the actual present land use areas. The LULC for 2015

was predicted through the model to produce simulated map and make comparison with the actual LULC map of 2015. Comparison of simulated and classified map for the year 2015 can be shown in *Table 3*. Visual analysis shows that simulated LULC map and actual map have relatively close resemblances.

Table 3. Comparison of actual and projected LULC types in 2015

LULC type	Area (ha)	
	Actual map	Simulated map
Vegetation Area	39650	39873
Non-Industrial Area	21370	20852
Industrial Area	2660	2628
Open Space Area	1348	1689
Water Bodies	1103	1104
Farming Area	869	854

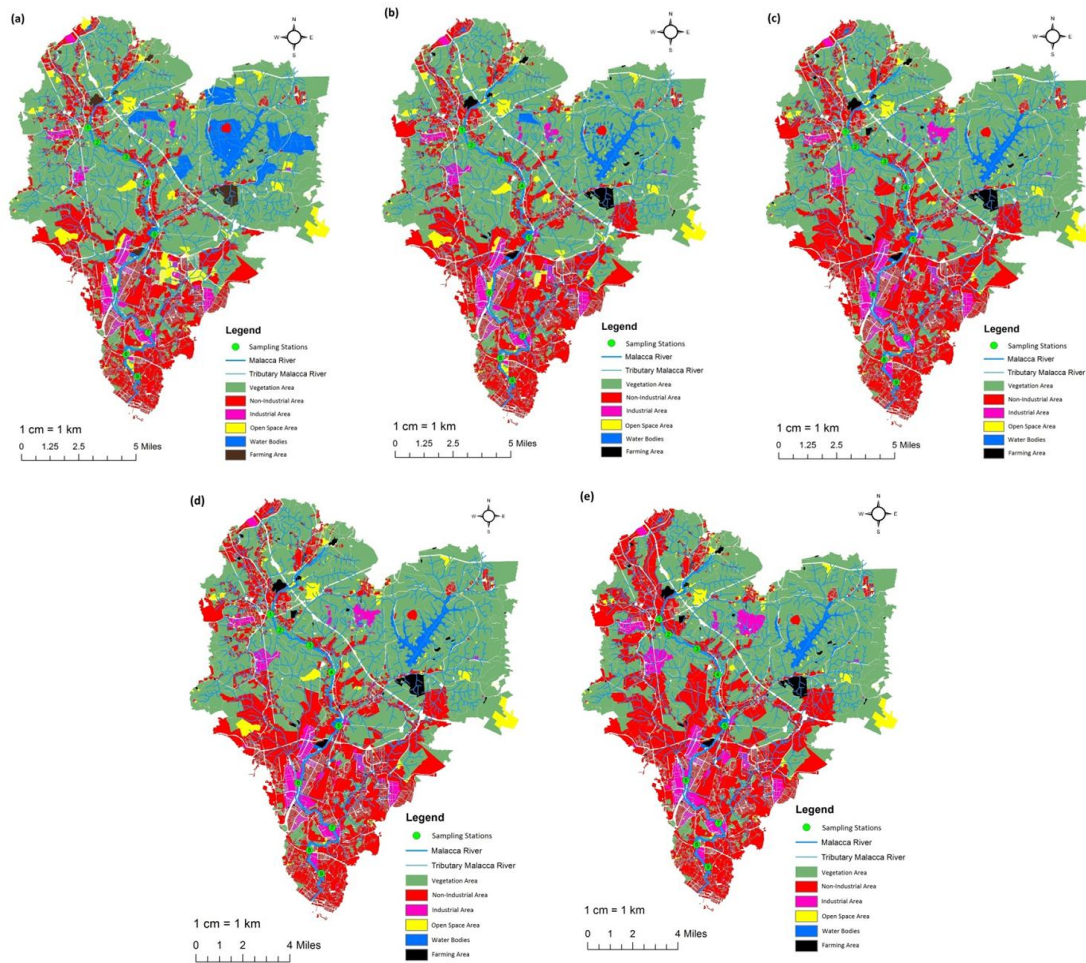


Figure 3. LULC actual map of (a) 2001, (b) 2008 and (c) 2015. LULC simulated map of (d) 2015 and (e) 2029

Table 3 indicates that industrial area, water bodies, and farming area have the best agreement. The simulated areas are 2628 ha, 1104 ha and 854 ha, while the actual area is 2660 ha, 1103 ha and 869 ha respectively. The simulated LULC map shows that non-industrial area, industrial area, and farming area are underestimated; while vegetation area, open space area, and water bodies are overestimated. Hence, the detailed statistical analysis based on the Kappa coefficient is used to measure the overall agreement of matrix, the ratio diagonal values summation versus total number of pixel counts within matrix, and the non-diagonal elements will be the best approach to assess the model accuracy (Arsanjani et al., 2013).

A kappa value of 0 illustrates the agreement between actual and reference map (equals chance agreement), the upper and lower limit of kappa is +1.00 (its occur when the is total agreement) and -1.00 (its happen when agreement is less chance (Arsanjani et al., 2013). The accuracy assessment process was done using VALIDATE module in IDRISI Selva environment v.17. The results indicate K values (Kno = 0.878; Klocation = 0.863; KlocationStrata = 0.863; Kstandard = 0.843) above 0.8 showing satisfactory level of accuracy. According to Viera and Garrett (2005), if the results are greater than 0.8 for each kappa index agreement, then the K statistics are considered accurate. Therefore, CA-Markov modeling is suitable for accurate prediction of future LULCs. This may be useful in this study for environmental management decision making and planning which involve the water quality of Malacca River.

LULC change detection

LULC change can be described in *Table 4*, where the activities carry out in Malacca River basin can be classes into vegetation area, non-industrial area, industrial area, open space area, water bodies, and farming area. According to LULC changes between 2001 to 2008 (*Table 5*), vegetation area and open space area are likely to be transform into non-industrial area (with 22.36% and 24.31%) and industrial area (with 3.13% and 2.04%) respectively. Water bodies is suspected to be decline about 66.57% and change into vegetation area, where this situation happen due to the sufficient water supply and fertile soil leads to the agriculture activities could be carried out. Based on the *Figure 3 (a)* and *(b)*, the agriculture activities are likely to occur in north-east area especially surrounding the reservoir. Meanwhile, open space area is also subjected to transform into vegetation area with 7.07%. Only a small percentage is detected to convert from vegetation area into farming area with 0.05%, indicating that a minority of animal husbandry is carried out on a small scale and mostly being conducted by local residents.

Table 4. Area (ha) of LULC type in Malacca River basin for 2001, 2008, 2015, and 2029

LULC type	Year			
	2001	2008	2015	2029
Vegetation Area	41841	41391	39650	35755
Non-Industrial Area	15829	18915	21370	24759
Industrial Area	1763	2173	2660	3313
Open Space Area	2695	2115	1348	1237
Water Bodies	4075	1603	1103	1025
Farming Area	797	803	869	911

Table 5. Transition probability of area and matrix calculated using land-use maps of 2001-2008

Category of LULC		2001-2008					
		VA	NIA	IA	OSA	WB	FA
Vegetation Area	F	30987274	93055338	1302581	0	0	20808
	P	74.46	22.36	3.13	0.00	0.00	0.05
Non-Industrial Area	F	521160	14766200	521160	521160	521160	521160
	P	3.00	85.00	3.00	3.00	3.00	3.00
Industrial Area	F	0	0	1672800	295200	0	0
	P	0.00	0.00	85.00	15.00	0.00	0.00
Open Space Area	F	170033.5	584655.5	49062	1601249	0	0
	P	7.07	24.31	2.04	66.58	0.00	0.00
Water Bodies	F	1889922	0	0	0	949077.7	0
	P	66.57	0.00	0.00	0.00	33.43	0.00
Farming Area	F	24000	24000	24000	24000	24000	680000
	P	3.00	3.00	3.00	3.00	3.00	85.00

VA means Vegetation Area; NIA means Non-Industrial Area; IA means Industrial Area; OSA means Open Space Area; WB means Water Bodies; FA means Farming Area; F means Frequency (ha); P means Percentage

The demand for development is detected to increase for the next 7 years, from 2008 to 2015. This situation spurred by the recognition of Malacca state as a historical tourism by UNESCO in 2007, where the condition increasing the development activities especially involve with non-industrial area and industrial area [Table 6 and Figure 3(c)]. For example, commercial activities such as restaurants, shopping center, and hotels will be built up in a big margin to provide adequate services for tourists. Directly, vegetation area will continuously decline into non-industrial area (with 7613802 ha), industrial area (with 1430374 ha), and open space area (with 4052.05 ha). At the same time, the open space area will also become a ‘victim’ of development by transforming into non-industrial area (with 634941 ha) and industrial area (with 158432.3 ha). The development of industrial and commercial activities provides job opportunities, attracting the people to centralize in urban areas of Malacca state. Hence, migration of local residents from rural areas and non-local residents into the Malacca state has increased the development activities that boosted into sub-urban area in providing residential area to the people. Indirectly, water bodies coverage will continue being pressure to be decline in quantity in supplying area to vegetation with 560683.2 ha; while farming area is subjected to be increase equivalent with non-industrial and industrial area, which converted from vegetation area for 283643.5 ha. Most of the farming area is referring to the animal husbandry activities.

Table 6. Transition probability of area and matrix calculated using land-use maps of 2008-2015

Category of LULC		2008-2015					
		VA	NIA	IA	OSA	WB	FA
Vegetation Area	F	31188629	7613802	1430374	4052.05	0	283643.5
	P	76.97	18.79	3.53	0.01	0.00	0.70
Non-Industrial Area	F	604275	17121125	604275	604275	604275	604275
	P	3.00	85.00	3.00	3.00	3.00	3.00
Industrial Area	F	0	387848	2028652	0	0	0
	P	0.00	16.05	83.95	0.00	0.00	0.00
Open Space Area	F	0	634941	158432.3	938126.7	0	0
	P	0.00	36.67	9.15	54.18	0.00	0.00
Water Bodies	F	560683.2	0	0	0	792316.8	0
	P	41.44	0.00	0.00	0.00	58.56	0.00
Farming Area	F	127155.6	0	0	0	0	708844.4
	P	15.21	0.00	0.00	0.00	0.00	84.79

VA means Vegetation Area; NIA means Non-Industrial Area; IA means Industrial Area; OSA means Open Space Area; WB means Water Bodies; FA means Farming Area; F means Frequency (ha); P means Percentage

Overall, the continuous development that occur for 14 years began from 2001 to 2015 is non-industrial and industrial area which converted from vegetation area into 11009434.1 ha (27.02%) and 1752056.5 ha (4.30%) respectively; while reducing coverage area is detected from the change of water bodies (2589000 ha) into vegetation area for 1992494.4 ha (76.96%), and open space area (2021500 ha) transform into non-industrial area with 894513.75 ha (44.25%), industrial area with 162528.6 ha (8.04%), and vegetation area with 106533.05 ha (5.27%) (Table 7). Meanwhile, the simulated future LULC changes for 14 years from 2015 to 2029 indicate that the statistics for non-industrial area, industrial area, and farming area continue to climb (Table 8). The vegetation activities keep ‘donating’ area into non-industrial area with 25.97%, industrial area with 4.13%, and farming area with 0.56%. At the same time, open space activities are transforming into non-industrial area at 17.67% and industrial area at 4.24%. Water body coverage will be pressured by vegetation area for 34.36% or 365590 ha to provide agricultural activities for local residents.

Table 7. Transition probability of area and matrix calculated using land-use maps of 2001-2015

Category of LULC		2001-2015					
		VA	NIA	IA	OSA	WB	FA
Vegetation Area	F	27800654.65	11009434.1	1752056.5	0	0	183354.75
	P	68.23	27.02	4.30	0.00	0.00	0.45
Non-Industrial Area	F	557985	15809575	557985	557985	557985	557985
	P	3.00	85.00	3.00	3.00	3.00	3.00
Industrial Area	F	0	359147.6	1852352.4	0	0	0
	P	0.00	16.24	83.76	0.00	0.00	0.00

Open Space Area	F	106533.05	894513.75	162528.6	857924.6	0	0
	P	5.27	44.25	8.04	42.44	0.00	0.00
Water Bodies	F	1992494.4	0	0	0	596505.6	0
	P	76.96	0.00	0.00	0.00	23.04	0.00
Farming Area	F	126782.6	0	0	0	0	706217.4
	P	15.22	0.00	0.00	0.00	0.00	84.78

VA means Vegetation Area; NIA means Non-Industrial Area; IA means Industrial Area;
 OSA means Open Space Area; WB means Water Bodies; FA means Farming Area;
 F means Frequency (ha); P means Percentage

Table 8. Transition probability of area and matrix calculated using land-use maps of 2015-2029

Category of LULC		2015-2029					
		VA	NIA	IA	OSA	WB	FA
Vegetation Area	F	26142914	9791339	1557113	0	0	211134
	P	69.34	25.97	4.13	0.00	0.00	0.56
Non-Industrial Area	F	0	19531019	3533481	0	0	0
	P	0.00	84.68	15.32	0.00	0.00	0.00
Industrial Area	F	89595	89595	2538525	89595	89595	89595
	P	3.00	3.00	85.00	3.00	3.00	3.00
Open Space Area	F	0	228385	54802	1009313	0	0
	P	0.00	17.67	4.24	78.09	0.00	0.00
Water Bodies	F	365590	0	0	0	698410	0
	P	34.36	0.00	0.00	0.00	65.64	0.00
Farming Area	F	111517	0	0	0	0	778483
	P	12.53	0.00	0.00	0.00	0.00	87.47

VA means Vegetation Area; NIA means Non-Industrial Area; IA means Industrial Area;
 OSA means Open Space Area; WB means Water Bodies; FA means Farming Area;
 F means Frequency (ha); P means Percentage

LULC changes trends can be described as the vegetation area transforming into open space area before it converted into non-industrial area, industrial area, and farming area. Meanwhile, water bodies are likely to change into vegetation area. Therefore, the LULC changes trends can be express in the net changes for 2001 to 2015, which indicate that non-industrial area, industrial area, and farming area have high positive values of 9473170.45 ha, 2113422.5 ha, and 614557.15 ha respectively; while vegetation area, open space area, and water bodies have high negative values with -10161050.3 ha, -605590.4 ha, and -1434509.4 ha respectively [(Table 9 (i)]. For the next 14 years between 2015 to 2029 [(Table 9 (ii))], only industrial area has high increment with 2583998.5 ha to resulted as 4697421 ha, while non-industrial area and farming area remain positive for 6575838 ha and 189212 ha respectively. Lastly, the net changes for open space area and water bodies remain negative by -193592 ha and -275995 ha, while vegetation area continue to decrease by -10992884 ha.

Table 9 (i). Rate of losses, gains, and net changes of LULC areas (ha) for 2001 to 2015

LULC types	2001-2015		
	L	G	NC
VA	-12944845.35	2783795.05	-10161050.3
NIA	-2789925	12263095.45	9473170.45
IA	-359147.6	2472570.1	2113422.5
OSA	-1163575.4	557985	-605590.4
WB	-1992494.4	557985	-1434509.4
FA	-126782.6	741339.75	614557.15
Total	-19376770.35	19376770.35	0

L means Losses; G means Gains; NC means Net Changes; VA means Vegetation Area;
 NIA means Non-Industrial Area; IA means Industrial Area; OSA means Open Space Area;
 WB means Water Bodies; FA means Farming Area

Table 9 (ii). Rate of losses, gains and net changes of LULC areas (ha) for 2015 to 2029

LULC types	2015-2029		
	L	G	NC
VA	-11559586	566702	-10992884
NIA	-3533481	10109319	6575838
IA	-447975	5145396	4697421
OSA	-283187	89595	-193592
WB	-365590	89595	-275995
FA	-111517	300729	189212
Total	-16301336	16301336	0

L means Losses; G means Gains; NC means Net Changes; VA means Vegetation Area;
 NIA means Non-Industrial Area; IA means Industrial Area; OSA means Open Space Area;
 WB means Water Bodies; FA means Farming Area

Discussion

The Markov chain model becomes an important model to describe probability movements of an individual in a system comprised of discrete states. When applying in land used perspective, Markov chains are often used to specify both time and a finite set of states as discrete values (Subedi et al., 2013). Transitions between the states of the system are recorded in the forms of a transition matrix that recorded the probability of moving from one state to another state (Behera et al., 2012). Applications of Markov chains in urban and non-urban land use began in the 1970s as an alternative to the use of a large-scale area in simulation models for land use forecasting (Arsanjani et al., 2013). Advantage in determining the development trend and benefit in predicting the state of future leading the Markov chain model are continuously applied in various research studies. As proved, the model's applicability and feasibility had been testified in many research papers and its results approximately conform to observed results (Arsanjani et al., 2013; Behera et al., 2012; Sinha and Kumar, 2013). Therefore, the Markov model result is a transition matrix which shows the probability of changes from each class of LULC to another class in the

future, which can be described in *Table 8* as transitional probability matrices and as a prediction of LULC statistic for year 2029.

Built-Up Area

Previous studies on water resources through statistical analysis indicate the source of pollutants for river water contamination can involve with residential activities, septic tank and sewage treatment plant activities, farming activities (or animal husbandry activities), and industrial activities; which can be grouped together to classified as built-up area to bring negative impact to the Malacca River (Hua, 2017; Hua et al., 2016). According to Hua (2017) and Rosli et al (2015) stated that Malacca River water quality is detected pollution by physico-chemical and biological parameters, which occur due to major source of pollutants from residential area, industrial area, and agricultural activities. These activities are carried out in urban and sub-urban area. On the other hands, LULC changes from the year 2001 to 2008 and 2008 to 2015 signify the demands towards built-up area category are continuously increasing. In other words, larger scale usage of the LULC area for human activities could increase the percentage of river pollution through the contribution of pollutant sources. Indirectly, increasing demand for land used will develop various categories of activities, which will increase the contamination in the Malacca River. Simultaneously, future land used statistic indicate 11462471 ha (70.32%) in continue 'add-on' for built-up area will estimate the possibilities of river water pollution to changes from clean to slightly polluted or slightly polluted into polluted condition.

Agriculture Land

According to the statistical analysis, agriculture activities are confirmed to be categorized as one element to contribute as a pollutant source. As additional evidence, although LULC changes in agriculture perspective are decreasing from 2001 to 2008 and increasing in a small percentage from 2008 to 2015 (*Table 5* and *Table 6*); however, larger fertilizer and chemical pesticide usage in agriculture activities will contribute non-point source pollution through surface runoff towards the Malacca River. The main focus on development of agricultural land happened in 2008; majority of water category was converted into open space as a preparation for agricultural activities. When the prediction of land used in 2015 to 2029 is estimated, possibility agricultural lands will conquest the water coverage area due to the nature of fertile soil. Moreover, the demands for built-up activities development in 'north-west' area are higher percentage, which may cause the agriculture land to shift from 'north-west' area into 'north-east' area. Nevertheless, even agriculture activities are estimated to reduce in land used area for about -11559586 ha (70.91%), it is not possible to prevent further contamination to occur in the Malacca River.

Water Coverage and Open Space Area

Continuously decrement of water coverage indicates that the area are being converted into an open space or agricultural land as seen in changes from the year 2001 to 2008 with 1889922 ha (66.57%) and 2008 to 2015 with 560683.2 ha (41.44%). As described previously, the water category is being transformed into other activities due to the higher fertility soil for farming, potentially for building construction, widen facilities and services,

and so on. Nevertheless, the reduction in percentage of water areas not only bring negative impacts; as the means of fresh water supplies and food source, but also affects the ecosystem, especially aquatic animals and contribute contamination through human activities of the point source and non-point source pollution. Future prediction of LULC classes estimates water coverage area is probably negative about 365590 ha (2.24%), which may possible to increase river water pollution and threatened the ecosystem survivor through endangered the environment.

Compared to open space category in LULC, the increment area for 7 years starting 2001 about 803751 (33.42%), but the total area are not remain static which start to reduce about 793373.3 ha (45.82%) for the next 7 years. The possible reason that caused the LULC change between 2001 to 2015 are the majority open space being transform from the area of water coverage and agriculture land into built up area. The main reason to cause this condition to happen is the human population increase drastically causing demand towards residences, job opportunities, food sources, facilities and services to increase. Indirectly, open space activities would also bring non-point source pollution through rapid surface runoff from soil erosion; and oil, grease and toxic chemicals from driveways and roads. Although there is probability to reduce pollution by reducing open space category between the year 2008 to 2015, but the percentage to increase pollutant sources through various human activities will still continue to happen. This is due to the number of land areas converted into built-up areas. Unfortunately, future prediction for the open space category are continue to be negative about 283187 (1.73%), which contribute to the land used activities or build-up area that serve the interest of the people without awareness of environment that would probably continuous to contaminate the Malacca River.

Conclusion

This study has shown the important role of LULC in Malacca River watershed in providing proper information in time for decision making concerning the degradation of river water quality. The above-mentioned method for extraction of LULC maps (2001, 2008, and 2015), and the simulation used to predict the future LULC (2029) aided by the CA-Markov modeling, have allowed the researcher to obtain information from analysis of LULCs in respect to the type and the extent of each LULC, while illustrating LULC conversions for the whole watershed in the study area. Generally, the results of this study indicate that supervised classification of remote sensing images is a robust mean of extracting appropriate LULC maps. The ability of simulation model in prediction proved to be good, while the CA-Markov model is a useful tool for LULC prediction (especially in quantitative and qualitative). Therefore, the CA-Markov model is important to land use policy design and planning especially involve with LULC development, which requires a framework for archiving the goals and objectives of sustainable land use development.

The Malacca River has experienced contamination due to the increasing of population, industrial development, agricultural expansion and forest degradation, and this has resulted to uncontrolled and unmanageable of LULC development. Therefore, continuous development has led to increasing pollution and aquatic ecosystem degradation, biodiversity resources destruction, as well as environmental damage in the watershed.

Hence, future land use change maps can be used as an early warning system for proper land use development to protect undisturbed area of natural ecosystem and biodiversity from human activities. Simultaneously, a future map could help for planning and management in reducing river contamination by selecting only the land suitable to be developed for economic purposes without disturbing the importance of the environmental perspective.

Acknowledgement. The author would like to thank the Department of Environment (DOE) Malaysia, Department of Irrigation and Drainage (JPS) Malaysia, Department of Town and Country Planning (JPBD) Malaysia, Malaysian Remote Sensing Agency (ARSM), and United States Geological Survey (USGS) for providing the base data for water quality data, river, GIS-map based maps including land use activities, and remote sensing imageries.

REFERENCES

- [1] Aris, A. Z., Praveena, S. M., Isa, N. M., Lim, W. Y., Juahir, H., Yusoff, M. K., Mustapha, A. (2013): Application of environmetric methods to surface water quality assessment of Langkawi Geopark (Malaysia). – *Environmental Forensics* 14 (3): 230-239.
- [2] Arsanjani, J. J., Helbich, M., Kainz, W., Bolorani, A. D. (2013): Integration of logistic regression, Markov chain and cellular automata models to simulate urban expansion. – *International Journal of Applied Earth Observation and Geoinformation* 21: 265-275.
- [3] Behera, D. M., Borate, S. N., Panda, S. N., Behera, P. R., Roy, P. S. (2012): Modelling and analyzing the watershed dynamics using Cellular Automata (CA)–Markov model–A geoinformation based approach. – *Journal of earth system science* 121(4): 1011-1024.
- [4] Carey, J. C., Fulweiler, R. W. (2012): Human activities directly alter watershed dissolved silica fluxes. – *Biogeochemistry* 111 (1-3): 125-138.
- [5] El-Kawy, O. A., Rød, J. K., Ismail, H. A., Suliman, A. S. (2011): Land use and land cover change detection in the western Nile delta of Egypt using remote sensing data. – *Applied Geography* 31 (2): 483-494.
- [6] Halmy, M. W. A., Gessler, P. E., Hicke, J. A., Salem, B. B. (2015): Land use/land cover change detection and prediction in the north-western coastal desert of Egypt using Markov-CA. – *Applied Geography* 63: 101-112.
- [7] Hua, A. K. (2017): Analytical and Detection Sources of Pollution Based Environmetric Techniques in Malacca River, Malaysia. – *Applied Ecology and Environmental Research* 15 (1): 485-499.
- [8] Hua, A. K., Kusin, F. M., Praveena, S. M. (2016): Spatial Variation Assessment of River Water Quality Using Environmetric Techniques. – *Polish Journal of Environmental Studies* 25 (6): 2411-2421.
- [9] Koomen, E., Borsboom-van Beurden, J. (2011): Land-use modelling in planning practice. – *Springer*, XVI-214.
- [10] Mahmud, A., Achide, A. S. (2012): Analysis of Land Use/Land Cover Changes to Monitor Urban Sprawl in Keffi-Nigeria. – *Environmental Research Journal* 6 (2): 129-134.
- [11] Memarian, H., Balasundram, S. K., Talib, J. B., Teh Boon Sung, C., Mohd Sood, A., Abbaspour, K. C. (2013). KINEROS2 application for land use/cover change impact analysis at the Hulu Langat Basin, Malaysia. – *Water and Environment Journal* 27 (4): 549-560.
- [12] Mishra, V. N., Rai, P. K. (2016): A remote sensing aided multi-layer perceptron Markov chain analysis for land use and land cover change prediction in Patna district (Bihar), India. – *Arabian Journal of Geosciences* 9 (4): 1-18.

- [13] Najar, I. A., Khan, A. B. (2012): Assessment of water quality and identification of pollution sources of three lakes in Kashmir, India, using multivariate analysis. – *Environmental Earth Sciences* 66 (8): 2367-2378.
- [14] Parsa, V. A., Yavari, A., Nejadi, A. (2016): Spatio-temporal analysis of land use/land cover pattern changes in Arasbaran Biosphere Reserve: Iran. – *Modeling Earth Systems and Environment* 2 (4): 178.
- [15] Pontius Jr, R. G., Millones, M. (2011): Death to Kappa: birth of quantity disagreement and allocation disagreement for accuracy assessment. – *International Journal of Remote Sensing* 32 (15): 4407-4429.
- [16] Rosli, S. N., Aris, A. Z., Majid, N. M. (2015): Spatial variation assessment of Malacca River water quality using multivariate statistical analysis. – *Malaysian Applied Biology* 44 (1): 13-18.
- [17] Sinha, P., Kumar, L. (2013): Markov land cover change modeling using pairs of time series satellite images. – *Photogramm. Eng. Remote Sens.* 79 (11): 11.
- [18] Subedi, P., Subedi, K., Thapa, B. (2013): Application of a hybrid cellular automaton Markov (CA-Markov) Model in land-use change prediction: a case study of saddle creek drainage Basin, Florida. – *Applied Ecology and Environmental Sciences* 1(6): 126-132.
- [19] Viera, A. J., Garrett, J. M. (2005): Understanding interobserver agreement: the kappa statistic. – *Fam Med.* 37 (5): 360-363.
- [20] Wan, R., Cai, S., Li, H., Yang, G., Li, Z., Nie, X. (2014): Inferring land use and landcover impact on stream water quality using a Bayesian hierarchical modeling approach in the Xitiao River Watershed, China. – *Journal of environmental management* 133: 1-11.
- [21] Weng, Q.H. (2010): *Remote Sensing and GIS integration*. – McGraw-Hill, New York, 424.
- [22] Yagoub, M. M., Al Bizreh, A. A. (2014): Prediction of land cover change using Markov and cellular automata models: case of Al-Ain, UAE, 1992-2030. – *Journal of the Indian Society of Remote Sensing*,42(3): 665-671.
- [23] Yang, X., Ren, L., Singh, V. P., Liu, X., Yuan, F., Jiang, S., Yong, B. (2012): Impacts of land use and land cover changes on evapotranspiration and runoff at Shalamulun River watershed, China. – *Hydrology Research* 43 (1-2): 23-37.