CrossMark

Performance assessment of geospatial simulation models of land-use change—a landscape metric-based approach

Yousef Sakieh · Abdolrassoul Salmanmahiny

Abstract Performance evaluation is a critical step when developing land-use and cover change (LUCC) models. The present study proposes a spatially explicit model performance evaluation method, adopting a landscape metric-based approach. To quantify GEOMOD model performance, a set of composition- and configuration-based landscape metrics including number of patches, edge density, mean Euclidean nearest neighbor distance, largest patch index, class area, landscape shape index, and splitting index were employed. The model takes advantage of three decision rules including neighborhood effect, persistence of change direction, and urbanization suitability values. According to the results, while class area, largest patch index, and splitting indices demonstrated insignificant differences between spatial pattern of ground truth and simulated layers, there was a considerable inconsistency between simulation results and real dataset in terms of the remaining metrics. Specifically, simulation outputs were simplistic and the model tended to underestimate number of developed patches by producing a more compact landscape. Landscape-metric-based performance evaluation produces more detailed information (compared to conventional indices such as the Kappa index and overall accuracy) on the model's behavior in replicating spatial heterogeneity features of a landscape such as frequency, fragmentation, isolation, and density. Finally, as the main characteristic of the proposed method, landscape metrics employ the maximum potential of observed and simulated layers for a performance evaluation procedure, provide a basis for more robust interpretation of a calibration process, and also deepen modeler insight into the main strengths and pitfalls of a specific land-use change model when simulating a spatiotemporal phenomenon.

Keywords Performance evaluation · Landscape metrics · Spatial simulation · GEOMOD · Land-use and cover change

Introduction and problem definition

With the increased availability of spatial data from 3S technologies derived from GIS, RS, and GPS and development of functional computer software for land-use and cover change (LUCC) models (Al-shalabi et al. 2012), a significant increase has occurred in LUCC change modeling. This field of investigation has witnessed a large number of applications employing micro-simulation models such as agent-based and cellular automata (CA) simulation methods (Al-ahmadi et al. 2008; Feng et al. 2011; Wang et al. 2012; Jokar et al. 2013; Sakieh 2013; Dezhkam et al. 2014; Sakieh et al. 2015). Various spatially explicit models and simulation techniques based on different conceptual

Y. Sakieh (✉) · A. Salmanmahiny
College of the Environmental Sciences, Gorgan University of Agricultural Sciences and Natural Resources, Behzisti Avenue, Gorgan, Golestan Provice, Iran
e-mail: sakie.yusuf@gmail.com

A. Salmanmahiny
e-mail: rassoulmahiny@gmail.com

foundations are developed for simulating transitions between initial and ending time years, thereby enabling the modeler to generate a prediction map for a subsequent time. As a critical step, it is important to assess the performance of the model when simulating a dynamic phenomenon such as the LUCC process.

Due to their simplicity, potential for dynamic spatial simulation, capability for high resolution modeling, and compatibility with GIS and remotely sensed data (Sullivan and Torrens 2000), CA-based models are among the frequently used modeling methods. Recent computer programs based on CA modeling concepts include DINAMICA (Soares-Filho et al. 2002), SLEUTH (Clarke et al. 1997), iCity (Stevens and Dragicevic 2007), CLUE-S (Verburg et al. 2002), and IDRISI's CA-Markov and GEOMOD models (Pontius et al. 2001). GEOMOD has been used frequently to analyze baseline scenarios of deforestation for carbon offset projects at different scales and within various geographic regions. GEOMOD as a grid-based model is able to simulate the spatial arrangement of land change forwards and backwards in time. The model simulates transitions between exactly two categories denoted as developed and non-developed.

Performance of a model can be evaluated through multiple methods including intellectual, statistical, and spatial techniques. Each method has its own pitfalls and strengths that provide the user with information on agreement between quantity and location of modeling effort and the actual layer when simulating the LUCC phenomenon. Some of these methods are simple least squares regression (Silva and Clarke 2002), Kappa-based statistics derived from a contingency table (Pontius 2000; Pontius and Millones 2011), and area under the relative operating characteristic (ROC) curve (Pearce and Simon 2000; Pontius and Schneider 2001; Pontius and Batchu 2003; Pontius and Si 2014). One major pitfall of these indices is their incapability of providing information on morphological agreement between simulated and reference layers. Therefore, applying a method for quantifying the simulation accuracy based on spatial pattern agreement could provide the modeler with valuable information on the model performance.

Landscape metrics as descriptive measures of configuration attributes of a landscape pattern can be powerful tools in evaluating simulation success. These indices can address spatial heterogeneity and variation dimensions such as frequency, isolation, density, and distance from nearest neighbor of a spatial pattern. The utility of landscape metrics in describing spatiotemporal processes and dynamics of LUCC has been well-documented in the literature (Herold et al. 2003; Poelmans and Rompaey 2009; Onsted and Chowdhury 2014; Asgarian et al. 2014), but the number of studies applying these measures as model performance indices are limited. Herold et al. (2005) reported the efficacy of landscape metrics in assessing model performance in terms of location agreement between spatial distribution of urban development for both simulated and reference maps. Wu et al. (2009) evaluated performance of SLEUTH model through multiple methods, and the landscape metric-based performance assessment provided a more powerful interpretation of the model's ability concerning location agreement between simulation and the ground truth layers. Mas et al. (2010) believe that landscape metrics are powerful tools for consideration of spatial pattern other than location. Guan et al. (2011) developed an integrated CA-Markov chain model incorporated with socioeconomic factors in Saga, Japan. They highlighted that landscape metrics are spatially explicit measures of urban morphology and are of potential for providing spatial information concerning model performance when calibrating a model.

This paper specifically aims to answer the following question: *Do landscape metrics have the potential for more detailed performance evaluation of a spatial model compared to conventional indicators such as Kappa index and overall accuracy?*

Materials and methods

Study area

Karaj city is the capital of Alborz Province, spanning between latitudes 35° 67′–36° 14′ N and longitudes 50° 56′–51° 42′ E and covers a total area of 141 km². Alborz mountain bounds this city in the north and the elevation descends from north to south. The average elevation is 1320 m above the sea level and dominant wind direction is North-West. Annual rainfall is 246.3 mm, and the annual average temperature varies between 15 and16 °C. The total population of the city is 1,686,521 (Iranian Statistics Center 2012) (Fig. 1).

Following the designation of the area as new province of the country, dramatic population growth has occurred with accompanying unplanned expansion of
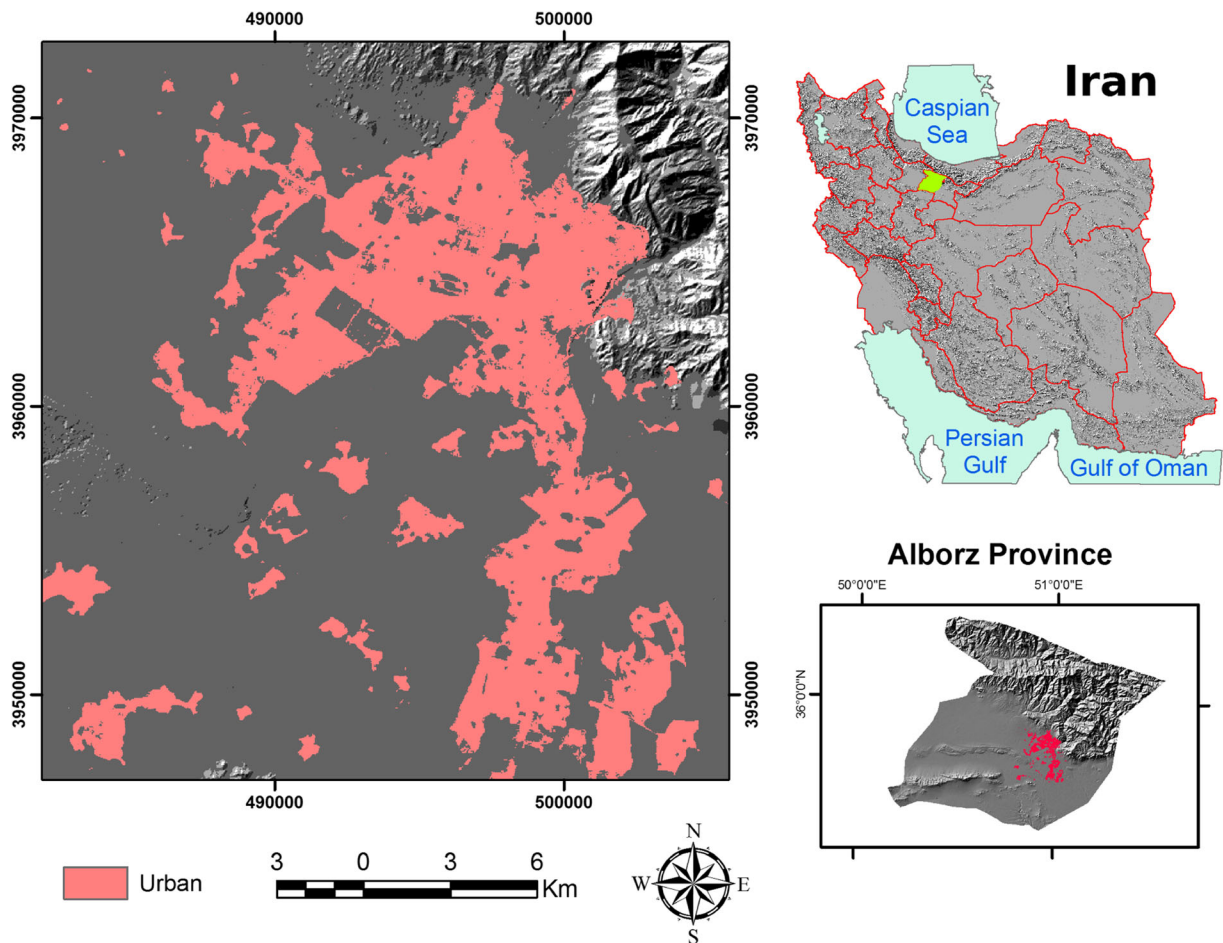
Fig. 1 Geographic location of study area across Alborz Province, Iran

urban and industrial sites that demands sustainable development plans be designed and employed to direct future changes (Sakieh et al. 2014b). Factors including proximity to Tehran city (capital of Iran), the surrounding environment, transportation system, affordable general facilities, and educational benefits have made the city as an attractive destination for immigrants from different parts of the country such that the Karaj city is now claimed to be one of the most culturally mixed regions in Iran. Accordingly, population growth associated with accelerated urbanization in the area has alarmed the necessity of spatial and regional planning efforts that support decisions made by policy makers. In addition, geospatial simulation models of LUCC are also considered to be potential tools that provide virtual environments for testing the possible outcomes of strategies adopted by authorities in the region (Sakieh et al. 2014a). Therefore, the area has been chosen since its

historical growth profile provides enough information regarding urban land-use change that makes a geospatial model to work properly.

Data processing

This study implements a GEOMOD model in a GIS environment and evaluates the simulation success of the model in terms of correspondence between spatial patterns of the simulated and ground truth maps. As spatially explicit indices, landscape metrics were calculated to address the level of similarity of local spatial patterns. Figure 2 illustrates our research flowchart.

Urban land-use layers were the most time-consuming input data to produce. Cloud-free Landsat TM images collected in 1984, 2000, 2006, and 2011 were processed for extracting developed lands. This historical profile is mainly selected to furnish the model with sufficient
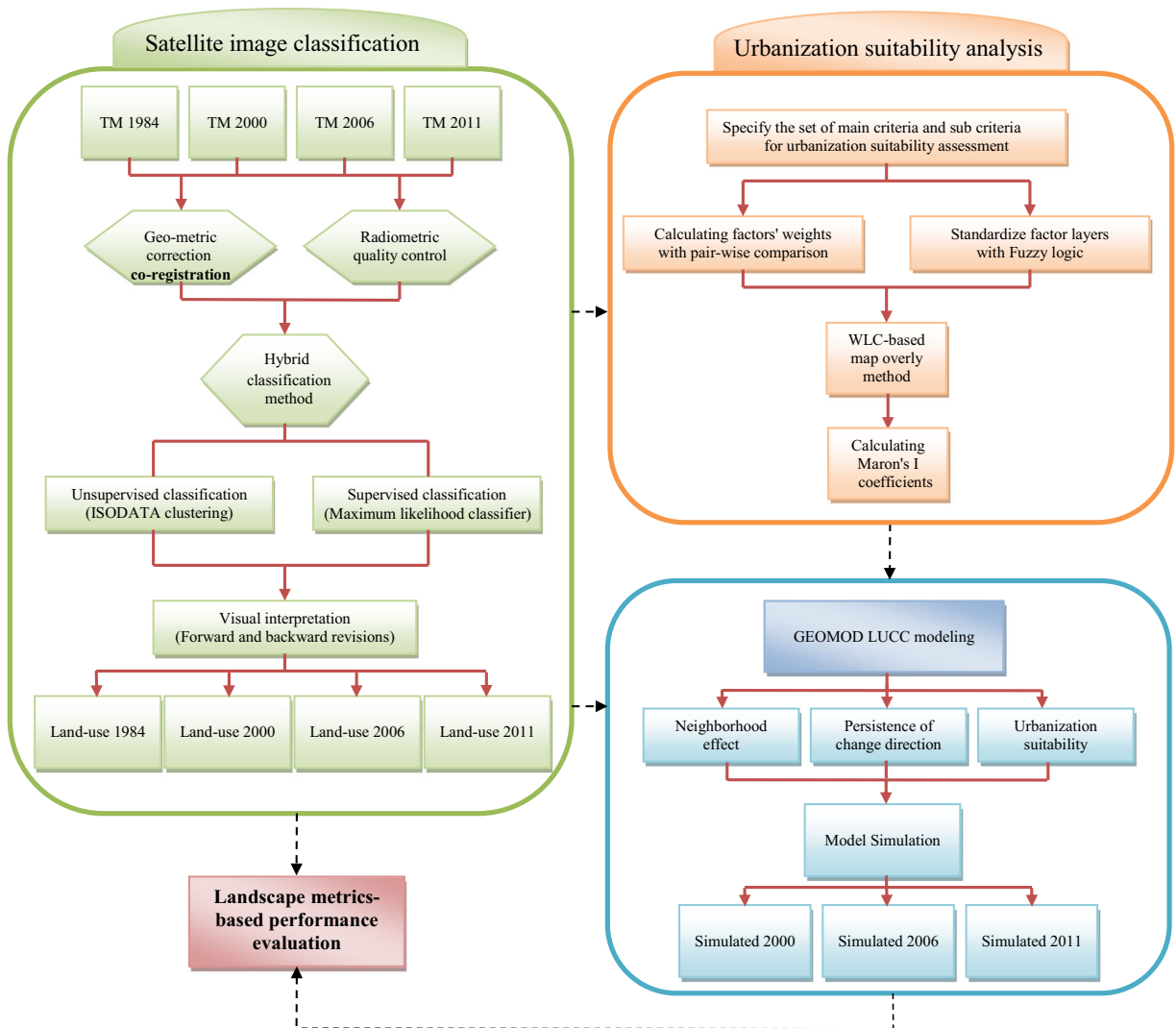
Fig. 2 Research flowchart

change in urban areas to make the model to work properly and to detect urban growth general patterns. The images were co-registered using the Universal Transverse Mercator (UTM) projection with acceptable RMSE (less than one pixel) using the nearest neighborhood algorithm. Quality control indicated that no noticeable distortions such as striping, banding, sweep error, duplicate pixels, or atmospheric error of clouds were found. The digital images had an acceptable radiometric quality. After geometric correction of the imagery, a hybrid method for image classification was employed including supervised (traditional maximum likelihood classifier), unsupervised (ISODATA clustering algorithm), and on-screen visual classifications.

Synthetic bands including NDVI- and PCA-derived layers were also calculated to further improve the classification accuracy.

We decided to conduct a hybrid classification process since the input data layers included only one category of developed lands. According to its historical profile in our study area, this category only increases through time and there is a very low amount of losses for this land feature. This characteristic allowed us to conduct forward and backward revisions through a visual control process. To be specific, for accuracy and consistency of classification process, the 1984 map layer was first totally checked to remove salt-pepper effect and to modify delineated borders of the developed lands (after

supervised and unsupervised classifications). Although very time-consuming, this process allowed us to produce a map with highest possible accuracy. In the next step, the 1984 and 2000 maps were cross-tabulated and revised. In this case, as the pervious step, boundaries were carefully delineated and the revised 2000 map was used for cross-tabulation with 2006 layer and the same process was also undertaken (forward revision for 2006 and 2011). Similarly, a backward revision was then carried out (delineated borders were carefully checked and modified to eliminate any possible inconsistency and misclassifications). In addition, accuracy evaluation was visually undertaken through pixel-by-pixel comparisons with various true and false color composites of the corresponding dates (1, 2, and 3; 4, 3, and 2; 3, 2, and 1; 2, 3, and 4; 1, 4, and 5), so that no apparent inconsistency was present in the maps in the end. The classification process explained here allowed us to produce highly accurate map layers since geospatial simulation models are sensitive to the accuracy of their input data. In addition, this classification method has found to be popular in areas where ground truth information is inaccessible due to physical or legal constraints (Mahiny and Clarke 2012, 2013; Sakieh et al. 2015; Jafarnezhad et al. 2015). The resultant raster layers of the years 1984, 2000, 2006, and 2011 included two categories of developed and non-developed coded as 1 and 2, respectively.

Urbanization suitability analysis

Applying the spatial multi-criteria evaluation (SMCE) method, suitability of the land for urban construction was calculated. SMCE is a grid-based analysis through which suitability of the land in response to a set of multiple digital input layers as evaluation criteria is determined. In order to reflect the relative importance of pixel's values within a raster and to make input layers comparable and ready to integrate as well as considering uncertainty embedded in datasets, fuzzy set theory (Zadeh 1965) was implemented for map standardization. Relative importance of each input data was calculated through pair-wise comparison of the analytic hierarchy process (AHP) method (Saaty 1980) (Table 1). In this case, a series of experts with knowledge on local growth patterns in the area were interviewed on the relative importance of the employed criteria. They were selected from scientific communities, who have conducted similar researches in our study area. In addition, authorities from Karaj city municipality were also

interviewed since they have more updated information regarding growth regulations. In the SMCE process, a collection of nine factors (elevation, slope, aspect, geology, distance to faults, distance to flood plain, proximity to urban edges, proximity to roads, and proximity to power lines) and six constraints (slopes higher than 20 %, elevations above 2300 m from the sea level, a series of buffer zones including 150 m distance from the main roads, 250 m distance from power lines, 300 m distance from flood potential areas, and 1000 m distance from faults) were selected. The criterion maps employed in this study were deemed to be the most important for urban environments in Iran (Makhdoum 2007; Mahiny and Clarke 2012) and were available for our study. The environmental attributes of the land implemented in our study site are decided with respect to the minimal risk of natural hazard (i.e., seismic activity and flood potential areas), terrain stability, construction expenses, environmental protection, conservation of riparian habitats, and accessibility to urban areas, main roads, and power lines. The layers were integrated using the following formula:

$$\text{SMCE for urbanization suitability} = \left(\sum_{i=1}^{n} W_i X_i \right) \prod C_i \tag{1}$$

where W_i is the relative weight of factor i, X_i is the fuzzified (standardized) factor i, $\prod$ is the multiplication operator, and C_i is constraint i. The SMCE process produces a single layer with values ranging between 0 and 255. The higher values indicate higher suitability for the land-use in question.

With the aim of brevity and saving space, classification of evaluation criteria, weighting scores, and fuzzification method are organized in Table 1. Figure 3 demonstrates the urbanization suitability surface of the study location derived from the SMCE process.

Finally, with respect to Tobler's first law of geography: "All things are related, but nearby things are more related than distant things" (Tobler 1969), the universal spatial auto-correlation Moran's I statistics (Rook's case) was calculated for studying spatial auto-correlation of the urbanization suitability map. The Moran I auto-correlation coefficient ranges from −1 to +1 and it approaches 1, when the spatial auto-correlation is high. The calculated statistics

Table 1 Classification of evaluation criteria, weighting scores, and fuzzification method

Criteria	Weight	Subcriteria	Weight	Indicator	Weight	Constraint	Factor
Ecological	0.800	Topography	0.224	Elevation	0.080	Elevations above 2300 m	Elevations extending between 400 and 1200 m were given the highest suitability to encourage urban sprawl in middle altitudes, where wind direction in daytime and nighttime is more suitable for air circulation in an urban environment.
				Slope	0.127	Slopes higher than 20 %	Slopes ranging between 0 and 20 % were linearly standardized such that slopes between 0 and 6 % were specified fuzzy membership value of 255, and the rest of the slope values extending between 6 and 20 % were standardized using a monotonically decreasing trend.
				Aspect	0.017	–	Regarding to semiarid environment of the study area, northern and eastern aspects were fuzzified with higher suitability values than western and southern aspects. The fuzzification schema for aspect layer was based on solar radiation received in built-up areas and architecture purposes for appropriate temperature in summer and winter.
		Geology	0.075	–	0.075	–	According to Makhdoum (2007), different rock formations comprising 11 categories were evaluated of their potential for urban construction and waste disposal. The resultant layer was fuzzified in a range between 0 and 255.
		Natural threats	0.501	Distance to faults	0.430	Buffer zone of 1000 m	Nearby lands to the faults were linearly standardized through which pixels with 5000-m distance received the highest suitability value.
				Distance to flood plain	0.072	Buffer zone of 300 m	The suitability of the immediate lands to the flood plains was gradually increased through a linear function and became constant at 3000 m (suitability value of 255).
Socioeconomic	0.200	LULC	0.150	Proximity to urban edges	0.150	–	To address higher suitability of the immediate lands to urban areas in terms of better service delivery and establishment of urban infrastructure (e.g., drainage systems and disposal management systems), lands within the l000-m distance from urban patches were specified the highest fuzzy membership value. The lands with 2500-m distance were assigned a value of 128, and the remaining pixels were fuzzified applying a monotonically decreasing function up to 6000 m, the maximum distance.
		Infrastructure	0.050	Proximity to roads	0.038	Buffer zone of 150 m	Areas with 150-m distance from the main roads were given the suitability values of 255, which remains constant up to 5000-m distance, and then, the values are decreased through a linear monotonically trend and become 0 at 15,000 m.

Table 1 (continued)

| Criteria	Weight	Subcriteria	Weight	Indicator	Weight	Constraint	Factor
				Proximity to power lines	0.013	Buffer zone of 250 m distance	The adjacent lands to the buffer zone were determined to possess the highest suitability (255) based on assumption that closer lands to power lines are more attractive for urbanization.

indicated that Moran's I index meets the acceptable threshold (0.9566) (Eastman 2009).

GEOMOD modeling

GEOMOD is a raster-based LUCC model that simulates the spatial arrangement of land change forwards or backwards in time. The model simulates transitions between exactly two land-use classes coded as 1 and 2, in our study referred to as non-developed and

developed. A map of a beginning time and information related to the number of grid cells of each class at an ending time is necessary. GEOMOD selects the location of the grid cells among the non-developed pixels that are most likely to be transformed into the developed category. Conversely, if there is a net increase in the number of non-developed pixels, the GEOMOD will explore among the developed areas to select the cells that are highly capable of being converted to the non-developed category between the time profiles. The minimum input

Fig. 3 Urbanization suitability surface of the study area generated through SMCE process
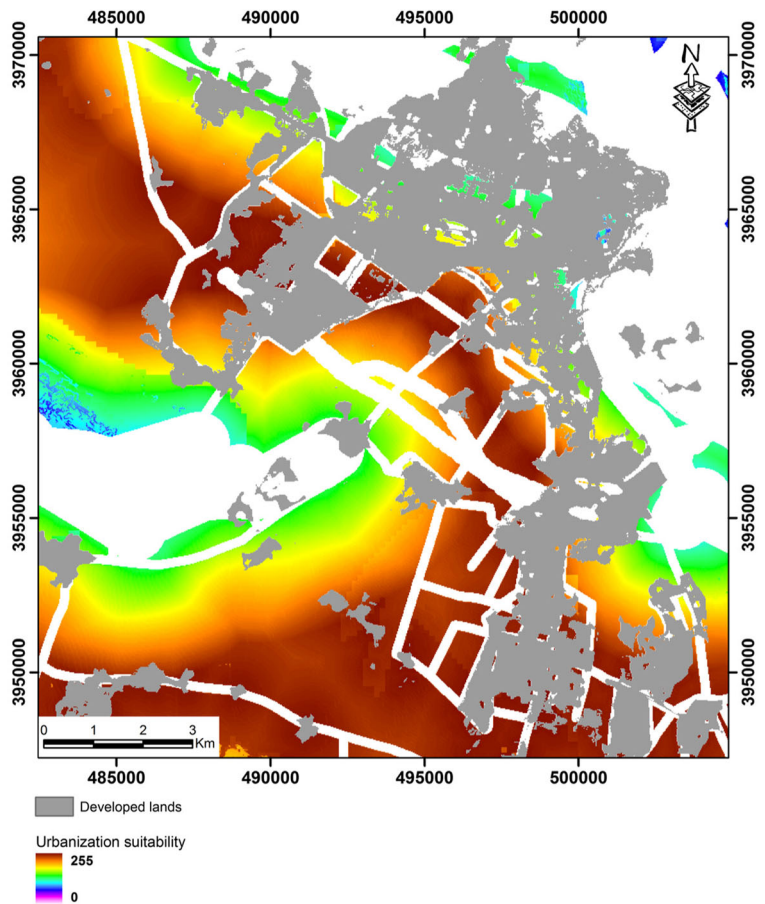
layers for model calibration are images of beginning and ending time as well as a suitability map of the targeted land-use. The most important output of the model is a map of the simulated landscape of the developed against non-developed pixels at the end time. Accordingly, using four reference map layers of the years 1984, 2000, 2006, and 2011 that depict the actual pattern of the developed lands, three simulation runs were executed to regenerate the landscape pattern of the years 2000, 2006, and 2011.

GEOMOD's decision rules to specify the location of change

The GEOMOD model assigns the location of land conversion according to three decision rules. The first decision rule addresses the persistence of change direction between land categories. In this paper, the model simulates a one-way change, from non-developed to developed lands. The second decision rule concerns the neighborhood effect. According to this rule, each cell is interactive with its adjacent neighbors (cells that are on the edge between developed and non-developed) and transformations between land categories are restricted within a small square window around each developed cells. This rule simulates the manner that new developed cells can grow out of previous development. The simulation process updates the definition of edge at every time step. If the model is required to transform more pixels than are available within the search radius of the window, all available cells within the search width are first converted and then the models start to convert cells within a wider search width to obtain enough number of cells for conversion. The third rule regards the suitability map, which depicts suitability values for the developed land category. Based on the third rule, GEOMOD simulates further development by exploring the landscape for areas of non-developed lands that have the highest suitability score. A model is said to be "dynamic" when the conditions at one time affect the transformation rules at the subsequent time. In this case, GEOMOD is not dynamic in the sense of the suitability map. On the contrary, referring to the neighborhood effect decision rule, the GEOMOD is dynamic such that the model recalculates for each year the adjacent cells on the edge between the developed and non-developed categories. Figure 4 illustrates reference maps versus their simulated versions.

Landscape metric-based performance evaluation

The correspondence between the simulated patterns of developed cells against their actual arrangement in reference maps was assessed using a series of landscape metrics. Although a wide variety of landscape metrics have been developed and applied in describing spatial composition and configuration of a landscape pattern (O'Neill et al. 1999; Turner et al. 2001; Herold et al. 2002, 2003; Dietzel et al. 2005; Dibari 2007; Weng 2007; Tang et al. 2008; Su et al. 2011; Tian et al. 2011), similar aspects of landscape patterns are measured by them due to overlap with each other. To reduce the data redundancy, a set of basic and most frequently applied metrics (Table 2) were selected (Luck and Wu 2002; Herold et al. 2003; Dietzel et al. 2005; Tang et al. 2008; Wu et al. 2009; Rafiee et al. 2009; Pham et al. 2011; Lechner et al. 2013; Asgarian et al. 2014; Jaafari et al. 2015; Sakieh et al. 2015; Jafarnezhad et al. 2015). They have then been calculated by FRAGSTATS mainly due to the following reasons (Botequila et al. 2006):

(1) These metrics quantify fundamental aspects of a landscape structure (number, area, size, distance, and shape of the patches), and the majority of the metrics are derived from these primary measures.
(2) They are easy to understand and interpret.
(3) They would be reliable when they are applied together, aiming to explain spatial complexity of a landscape in terms of patches, patch spatial distribution and patch shape complexity, and connectivity.

The metrics included number of patches (NP), edge density (ED), largest patch index (LPI), mean Euclidean nearest neighbor distance (ENN_MN), total class area (CA), splitting index (SPLIT), and landscape shape index (LSI).

In order to quantitatively address spatial similarity of the compared maps, the relative error (RE) index was calculated using Eq. (1):

$$RE = \left[\frac{(M_r - M_p)}{M_r} \right] * 100 \qquad (2)$$

where M_r indicates value of the landscape metric extracted from the reference map and M_p refers to the value of the landscape metric derived from the simulated map.
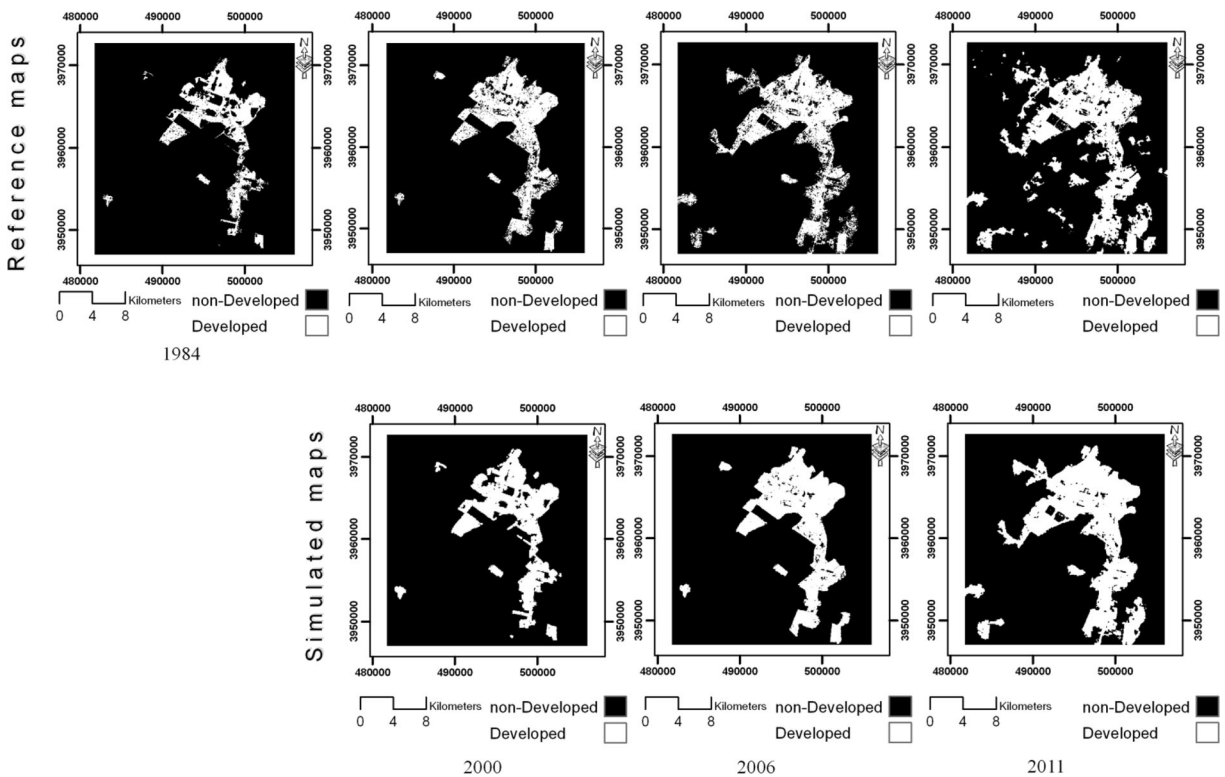
Fig. 4 Reference maps of the study area obtained through hybrid classification (*first row*) versus simulated maps of the developed lands of the years 2000, 2006, and 2011 (*second row*)

NP metric indicates total number of developed patches in the landscape. This metric is chosen mainly to explain model's ability in correct simulation of patch numbers. The formula is as follows:

$$NP = ni \qquad (3)$$

where n_i indicates number of patches in the landscape belonging to the class developed lands.

ED equals the number of developed patches divided by total landscape area (m^2) and multiplied by 10,000 (to convert to 100 ha). ED equals the sum of the lengths (m) of all edge segments involving the corresponding patch type, divided by the total landscape area (m^2), and multiplied by 10,000 (to convert to hectares). ED reports edge length on a per unit area basis that facilitates comparison

Table 2 Descriptions of landscape metrics used in the study (Mc Garigal and Marks 1995)

Landscape metrics	Abbreviation	Units	Range	Type of metric	Concept
Number of patches	NP	None	NP≥1, without limit	Landscape composition	Fragmentation
Edge density	ED	Meters per hectare	ED≥0, without limit	Landscape configuration	Density
Largest patch index	LPI	Percent	0 < LPI ≤ 100	Landscape composition	Dominance
Mean Euclidean nearest neighbor distance	ENN_MN	Meters	ENN > 0, without limit	Landscape configuration	Isolation
Total class area	CA	Hectare	CA > 0, without limit	Landscape composition	Area
Splitting index	SPLIT	None	1 ≤ SPLIT ≤ number of cells in the landscape area squared	Landscape configuration	Cumulative patch area distribution
Landscape shape index	LSI	None	LSI ≥ 1, without limit	Landscape configuration	Shape, aggregation index

among landscapes of varying size. The formula to calculate ED metric is the following:

$$ED = \frac{\sum_{k=1}^{m} e_{ik}}{A}(10,000) \qquad (4)$$

where e_{ik} denotes total length (m) of edge in landscape involving patch type developed lands; it includes landscape boundary and background segments involving the corresponding patch type and A that stands for total landscape area (m^2).

LPI corresponds to the area (m^2) of the largest patch of the respective patch type divided by total landscape area (m^2) and multiplied by 100 (to convert to a percentage). The LPI metric at the class level measures the percentage of total landscape area comprised by largest patch.

$$LPI = \max_{j=1}^{n}(a_{ij}) \, A(100) \qquad (5)$$

where a_{ij} means area (m^2) of patch ij and A is total landscape area (m^2).

ENN_MN metric equals the distance (m) to the nearest neighboring patch of the same type, based on shortest edge-to-edge distance. The ENN_MN metric is a measure of patch context and can be used to quantify patch isolation. The formula for this metric is as follows:

$$ENN\_MN \ = \ mean \ (hij) \qquad (6)$$

where *mean* refers to the sum, across all patches of the corresponding patch type, of the corresponding patch metric values, divided by the number of patches of the same type and *hij* is distance (m) from patch *ij* to nearest neighboring patch of the same type (class), based on patch edge-to-edge distance.

CA metric is a measure of landscape composition and equals the sum of the areas of developed land category (hectares). This metric is specifically selected to depict the simulation success in terms of correspondence between simulated and the actual area of land-use in question. The equation is as follows:

$$CA = \sum_{j=1}^{n} a_{ij}\left(\frac{1}{10,000}\right) \qquad (7)$$

where a_{ij} area (m^2) of developed lands is divided by 10,000 to convert to hectares.

SPLIT metric is the total landscape area (m^2) squared divided by the sum of patch area (m^2) squared, summed across all patches of the corresponding patch type. SPLIT equals 1 when the landscape consists of a single patch. The value for this metric increase as the main patch type is significantly fragmented into smaller patches. SPLIT is based on the cumulative patch area distribution and is interpreted as effective mesh number or number of patches with a constant patch size when the corresponding patch type is subdivided into S patches, where S is the value of splitting index. The formula to calculate the SPLIT metric is as foolows:

$$SPLIT = \frac{A^2}{\sum_{j=1}^{a} a_{ij}^2} \qquad (8)$$

where a_{ij} is area (m^2) of developed areas and A denotes total landscape area (m^2).

LSI metric provides a standardized measure of total edge or edge density that adjusts for the size of the landscape. The metric approaches 1 when landscape shape of a particular type is almost square. LSI increases without limit as landscape shape becomes more irregular and/or as the length of edge within the landscape increases. This metric mirrors the similarity between the actual shape of particular patch type versus its simulated shape. The formula for calculating LSI metric is as follows:

$$LSI = \frac{e_i}{\min e_i} \qquad (9)$$

where e_i denotes the perimeter of the developed land category and $\min e_i$ means minimum perimeter of this class (Mc Garigal and Marks 1995).

Results and discussion

The SMCE process was implemented to weight, standardize, and integrate several environmental parameters that are deemed to have considerable influence on urbanization suitability in our study location. As Table 1 shows, ecological criteria gained higher levels of relative importance. This is attributed to the fact that natural hazards and geomorphologic factors are the first priority

in terms of legal restrictions in the study area. After years of unplanned urban expansion, city authorities strictly prevent any further establishment of new urban centers without detailed studies of environmental sustainability. On the other hand, factors including aspect and proximity to power lines received the lowest weight.

The northern part of the study area is totally unsuitable for urban construction, which is consistent with historical urban growth in the area. Current urban boundaries are greatly associated with agricultural fields in the periphery of the developed lands. These locations possess suitable slope, high accessibility and are less influenced by natural hazards and therefore have more potential for urbanization. Visual interpretation of urban patches distribution reveals that there are some urban centers in the eastern part, which are located in totally unsuitable lands (with proximity to active faults). Generally, the southern and eastern parts are more suitable for urbanization, while linear and scattered urban development was the dominant growth form in these directions. Since proximity to urban edges was scored as the second important factor in our study, dispersed urban patches with small physical size and less accessibility to the transportation network also regulated local distribution of suitability values in southern and eastern directions.

Detailed and quantitative measures of model performance in terms of agreement between spatial arrangements of simulated and reference maps are provided in Table 3 and Fig. 5. There are some interesting results. NP, ED, LPI, ENN_MN, CA, SPLIT, and LSI have successfully reflected the level of simulation success in capturing spatial pattern of the developed lands.

The NP metric depicted considerable difference between simulation and ground truth layers and where the landscape becomes more complex with increasing numbers of patches, the model tends to simulate less accurately (with RE values of 59, 82, and 82 % for 2000, 2006, and 2011, respectively). In fact, there is a tendency to underestimate the NP metric with the highest difference recorded for the year 2011. Visual interpretation of the simulation outputs implies that spatial arrangement of small and scattered patches of developed lands were less successfully simulated. On the other hand, referring to the ability of the model in regenerating the real pattern of the biggest patch of the developed land, there was a high level of similarity between CA metrics derived from reference and the simulated layers

Table 3 Landscape metric values derived from reference and simulated maps of the years 2000, 2006, and 2011 and their difference percent as RE index

| RE of metrics/developed lands	NP			ED			LPI			ENN_MN			CA			SPLIT			LSI		
	2000	2006	2011	2000	2006	2011	2000	2006	2011	2000	2006	2011	2000	2006	2011	2000	2006	2011	2000	2006	2011
Metric value derived from reference map	59	176	185	7.45	11.95	13.52	11.76	14.51	11.88	180.63	104.57	212.66	8471	10,522	14,101	71.89	47.33	30.99	12.60	18.19	17.80
Metric value derived from simulated map	24	30	33	6.01	6.20	6.30	12.21	15.30	19.87	792.26	525.80	215.99	8247	10,347	14,165	66.93	42.60	25.16	10.31	9.47	8.29
Difference percent (RE)	59	82	82	19	48	53	4	5	11	33	40	2	2	1	0	6	10	18	18	47	53
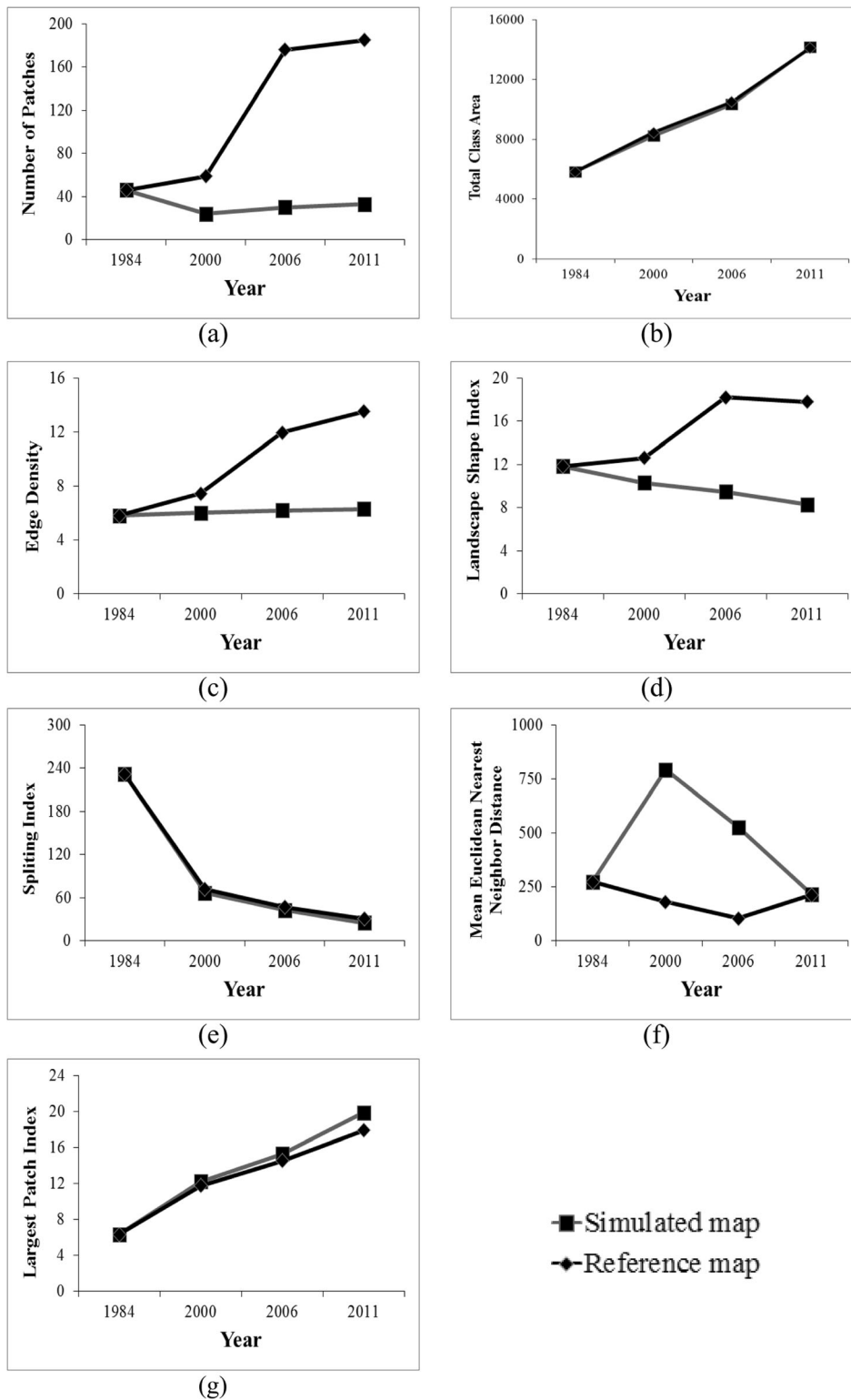
Fig. 5 Calculated metrics for reference and simulated maps of the years 2000, 2006, and 2011. Number of patches (NP) (**a**), total class area (CA, hectare) (**b**), edge density (ED, meters per hectare) (**c**), landscape shape index (LSI) (**d**), splitting index (SPLIT) (**e**), mean Euclidean nearest neighbor distance (ENN_MN, meter) (**f**), and largest patch index (LPI, percent) (**g**)

(with RE values of 2, 1, and 0 % for 2000, 2006, and 2011, respectively). This explains that as the landscape becomes more compact and less complex in shape, the model records a convincing level of agreement between actual and simulation maps for the CA metric. Similarly, the model's outputs for the LPI metric are close to those of derived from their corresponding actual layers but accuracy of the model tends to decrease more dramatically (especially for the year 2011 with RE values of 4, 5, and 11 %). This behavior for the LPI metric indicates that accuracy of the model for simulating main urban core in the area tends to decrease through time. The ED metric demonstrates noticeable differences between the real data and the modeling effort (with RE values of 19, 48, and 53 %). In this case, as the number of the developed patches increased and the landscape became more fragmented, the ability of the model to simulate correctly more edge densities decreased. This may be attributed to the fact that landscapes with more small and dispersed patches produce more complicated edges, which is difficult to model. Referring to the SPLIT metric, model performance was best for the year 2000 and worst for 2011 (with RE values of 6 and 18 %, respectively). This metric indicates that as the focal patch type is increasingly reduced in area and subdivided into smaller patches, the model's ability significantly decreased in capturing the real landscape pattern of the developed lands. Regarding the EMM_MN metric, there is also considerable RE values for years 2000 (33 %) and 2006 (40 %) but the performance of the model increases for the year 2011 (2 %). This is because the model tends to create more values for NP and CA over time that can improve performance of the model for this metric. Finally, the LSI metric shows an increasing trend in difference between the simulated and the reference maps (with RE values of 18, 47, and 53 %). This process mirrors the fact that as the landscape becomes more fragmented in structure and more complex in shape, the simulation success decreases and the outputs tend to be more compact and simpler.

Taking the total set of landscape metrics into account, the GEOMOD model tends to produce more compacted landscape and simplistic simulation results. Although there is a convincing level of similarity between simulation and the ground truth layers in terms of CA, LPI, and SPLIT metrics, NP, ENN_MN, LSI, and ED demonstrated considerable differences. This may clarify that quantity-based agreement between modeled outputs and the real datasets does not guarantee the high accuracy for spatial pattern indices.

As shown in Fig. 3, southern and eastern directions of the study area retain larger areas with higher urbanization suitability values. To some extent, the SMCE-derived urbanization suitability layer can explain the historical growth profile of Karaj city. Pixels with highest suitability values are considerably located in the periphery of existing developed lands. This is consistent with the GEOMOD simulation behavior in our study site. In this case, main urban core of Karaj city with highest area is better simulated for different metrics; while on the other hand, the model is not very successful in replicating the growth patterns of those patches located in areas with less urbanization suitability values and more distance from main urban center. In this study, factors including distance to faults, proximity to urban edges, and slope gained greater relative importance in terms of weighting score (0.430, 0.150, and 0.127, respectively). Therefore, distribution of urbanization suitability values is more correlated with distribution of their corresponding cells in factor layers with higher fuzzified scores. This matter provides some important implications for a more robust modeling effort. In this case, for realistic simulations, the modeler needs to include a combination of socioeconomic and ecological parameters that inform the model with information on human decisions (actors) and environmental elements (factors). In addition, locations with better urban land-use planning and sustainable development strategy can be better modeled and simulated. This means that by guiding urban growth directions in a more sustainable trajectory, the model can detect a pattern which is consistent with environmental capability and human needs. Since historical growth of Karaj city depicted a scattered and unplanned pattern of urban development, a large number of urban patches with small physical size have appeared with no connection neither to each other nor the main urban core. Therefore, the model faced a very complex growth pattern, which resulted in the model's underestimation for the number of patches and complexity of the landscape. Accordingly, future predictions of the LUCC in this area would be either highly uncertain or very data-demanding and time-consuming.

Spatially explicit predictive models require performance evaluation in terms of spatial dimensions of the land-use and land-cover arrangements. The processes and functions of a landscape are altered by evolving its structure, composition, and configuration (Forman and

Godron 1986; Farina 2006; Turner et al. 2001; Botequila et al. 2006; Asgarian et al. 2014; Hasani Sangani et al. 2014). In this study, a procedure was implemented for spatially explicit performance evaluation of the GEOMOD land-use change model based on landscape pattern analysis. The method described here is of potential for examining agreement between simulated and ground truth layers considering spatial arrangement of land-use categories. In this process, the number of patches, edge density, and shape of the patches as well as their isolation are considered which provide the modeler with valuable information on the model's behavior. Similar studies conducted by Herold et al. (2005), Wu et al. (2009), Mas et al. (2010) Guan et al. (2011), and Jafarnezhad et al. (2015) confirm the utility of the methodology explained in this paper. As Fig. 4 illustrates, the GEOMOD model tends to underestimate the complexity of the landscape by simulating more aggregated landscape with fewer number of patches of developed lands (referring to NP, ENN_MN, ED, and LSI metrics). In contrast, based on visual interpretation of the results, the general structure of the landscape is well-simulated as is reflected in the value of CA, LPI, and SPLIT metrics.

Referring to urbanization suitability map of the study area, where some of urban patches are located in totally unsuitable lands, model inability in realistic simulation of number of patches and urban edge densities could be explained. Since the GEOMOD model can dynamically adopts itself to new conditions in terms of the neighborhood effect rule, this characteristic of the model could lightly compensate for static influence of urbanization suitability map.

The GEOMOD model takes advantage of three decision rules including neighborhood effect, persistence of change direction, and suitability values. There are other rules such as regional stratification and land price that could be embedded in the model formulation when more location-based accuracy of model performance is desired. Specifically, when urban patches are located on unsuitable lands, as shown in this study, incorporation of other dynamic or static transition rules is of more importance.

Referring to sensitivity of the landscape metrics to classification accuracy and classification approach (pixel-based classification or segmentation), it is necessary that reference maps retain high spatial accuracy. Accordingly, an accurate process (e.g., hybrid classification method) and careful attention (e.g., visual control) should be undertaken to produce high quality land-use and land-cover maps as they are input data to land change simulation models. This might ensure accuracy and consistency of data preparation, and therefore, validity of metrics.

Conclusions

Simulation models of LUCC transformations as innovative planning tools are of noticeable interest for monitoring, modeling, and analyzing landscape change. Accordingly, assessing the spatial performance of such models requires including basic elements of the landscape structure, which provides additional level of knowledge on simulation success of LUCC change model. Numerous landscape metrics have been designed and implemented in various locations during the last decades; however, a few number of researches have employed landscape indices as spatial indicators of model performance. The present document embraced a different approach to measure the simulation success of the GEOMOD model, which is reflected in RE values of metric calculations. According to the results, this method has potential to evaluate model performance by considering morphological and structural features of LUCC patches such as size, number, distance, shape, and complexity. In addition, this practice has potential to simultaneously quantify simulation success at three levels including a specific patch, class, and landscape, which merits further research. Therefore, the LUCC modelers and practitioners can compare different models in a spatially explicit manner and decide on best-performing simulation method in terms of replicating a landscape structure. This property would be interesting for ecologists that tend to measure the effects of structural changes on a specific landscape function (e.g., carbon sequestration, nutrient cycling).

There is no universally accepted method for selecting the best-performing landscape metrics for a specific study. In this matter, the relevancy between research questions and metric types can guide the modeler to decide the best indices. The employed metrics in this study were selected based on their potential in measuring the simulation success of a LUCC change model. Therefore, other spatial simulation models in different study areas would require different set of the metrics to prevent information on missing or exaggerated results.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

Al-ahmadi, K., See, L., Heppenstall, A., & Hogg, J. (2008). Calibration of a fuzzy cellular automata model of urban dynamics in Saudi Arabia. *Ecological Complexity, 6*(2), 80–101.

Al-shalabi, L., Billa, L., Pradhan, B., Mansor, S., & Al-sharif, A. A. A. (2012). Modelling urban growth evolution and land-use changes using GIS based cellular automata and SLEUTH models: the case of Sana'a metropolitan city, Yemen. *Environmental Earth Sciences, 70*(1), 425–437.

Asgarian, A., Amiri, B. J., & Sakieh, Y. (2014). Assessing the effect of green cover spatial patterns on urban land surface temperature using landscape metrics approach. *Urban Ecosystems, 18*(1), 209–222.

Botequila, A. l., Miller, J., Ahem, J., & McGarigal, K. (2006). *Measuring landscapes: a planner's handbook*. Washington: Island Press. 272p.

Clarke, K. C., Hoppen, S., & Gaydos, L. (1997). A self-modifying cellular automaton model of historical urbanization in the San Francisco Bay area. *Journal of Environment and Planning B: Planning & Design, 24*(2), 247–261.

Dezhkam, S., Amiri, B. J., Darvishsefat, A. A., & Sakieh, Y. (2014). Simulating urban growth dimensions and scenario prediction: a case study of Rasht County, Guilan, Iran. *Geojournal, 79*(5), 591–604.

Dibari, J. N. (2007). Evaluation of five landscape-level metrics for measuring the effects of urbanization on landscape structure: the case of Tucson, Arizona, USA. *Landscape and Urban Planning, 79*(3–4), 308–313.

Dietzel, C., Herold, M., Hemphill, J. J., & Clarke, K. C. (2005). Spatio-temporal dynamics in California's Central Valley: empirical links to urban theory. *International Journal of Geographic Information Science, 19*(2), 175–195.

Eastman, R. (2009). *Idrisi Taiga Version* 16.01, Clark Laboratories, Clark University, Worcester, MA.

Farina, A. (2006). *Principles and methods in landscape ecology: toward a science of landscape*. Netherland: Springer publication. 412p.

Feng, Y., Liu, Y., Tong, X., Liu, M., & Deng, S. (2011). Modeling dynamic urban growth using cellular automata and particle swarm optimization rules. *Landscape and Urban Planning, 102*(3), 188–196.

Forman, K. E., & Godron, M. (1986). *Landscape ecology*. New York: Wiley. 432p.

Guan, D., Li, H. F., Inohae, T., Su, W., Nagaie, T., & Hokao, K. (2011). Modeling urban land use change by integration of cellular automata and Markov model. *Ecological Modeling, 222*(20–22), 3761–3772.

Hasani Sangani, M., Amiri, B. J., Alizadeh Shabani, A., Sakieh, Y., & Ashrafi, S. (2014). Modeling relationships between catchment attributes and river water quality in southern catchments of the Caspian Sea. *Environmental Science and Pollution Research, 22*(7), 4985–5002.

Herold, M., Couclelis, H., & Clarke, K. C. (2005). The role of spatial metrics in the analysis and modeling of urban land use change. *Computers, Environment and Urban Systems, 29*(4), 369–399.

Herold, M., Goldstein, N. C., & Clarke, K. C. (2003). The spatio-temporal form of urban growth: measurement, analysis and modeling. *Remote Sensing of Environment, 86*(3), 286–302.

Herold, M., Scepan, J., & Clarke, K. C. (2002). The use of remote sensing and urban landscape metrics to describe structures and changes in urban land uses. *Environment and planning B: planning and design, 34*(8), 1443–1458.

Iranian Statistics Center. (2012). General census of population and housing of Karaj City.

Jaafari, S., Sakieh, Y., Shabani, A. A., Danehkar, A., & Nazarisamani, A. (2015). Landscape change assessment of reservation areas using remote sensing and landscape metrics (case study: Jajroud reservation, Iran). *Environment, Development and Sustainability*. doi:10.1007/s10668-015-9712-4.

Jafarnezhad, J., Salmanmahiny, A., & Sakieh, Y. (2015). Subjectivity versus objectivity—a comparative study between Brute Force method and Genetic Algorithm for calibrating the SLEUTH urban growth model. *Urban Planning and development*. doi:10.1061/(ASCE)UP.1943-5444.0000307.

Jokar, J., Helbich, A., & Noronha, E. (2013). Spatiotemporal simulation of urban growth patterns using agent-based modeling: the case of Tehran. *Cities, 32*, 33–42.

Luck, M., & Wu, J. G. (2002). A gradient analysis of urban landscape pattern: a case study from the Phoenix metropolitan region, Arizona, USA. *Landscape Ecology, 17*(4), 327–339.

Lechner, A. M., Reinke, K. J., Wang, Y., & Bastin, L. (2013). Interactions between landcover pattern and geospatial processing methods: effects on landscape metrics and classification accuracy. *Ecological Complexity, 15*, 71–82.

Makhdoum, F. M. (2007). *Fundamental of land use planning*. Tehran: University of Tehran press.

Mahiny, A. S., & Clarke, K. C. (2012). Guiding SLEUTH land-use/land-cover change modeling using multicriteria evaluation: towards dynamic sustainable land-use planning. *Environment and planning B: planning and design, 39*(5), 925–944.

Mahiny, A. S., & Clarke, K. C. (2013). Simulating hydrologic impacts of urban growth using SLEUTH, multi criteria evaluation and runoff modeling. *Journal of Environmental Informatics, 22*(1), 27–38.

Mas, J. F., Vega, A. P., & Clark, C. (2010). Assessing "spatially explicit" land use/cover change models. Proceedings of the IUFRO Landscape Ecology Working Group International Conference, Bragança, Portugal.

Mc Garigal, K., & Marks, B. J. (1995). *FRAGSTATS*: spatial pattern analysis program for quantifying landscape structure, USDA Forest Service.

O'Neill, R. V., Riitters, K. H., Wickham, J. D., & Bruce Jones, K. (1999). Landscape pattern metrics and regional assessment. *Ecosystem Health, 5*(4), 225–233.

Onsted, J. A., & Chowdhury, R. R. (2014). Dose zoning matter? A comparative analysis of landscape change in Redland, Florida using cellular automata. *Landscape and Urban Planning, 121*, 1–18.

Pearce, J., & Simon, F. (2000). Evaluating the predictive performance of habitat models developed using logistic regression. *Ecological Modelling, 133*(3), 225–245.

Pham, H. M., Yamaguchi, Y., & Bui, T. Q. (2011). A case study on the relation between city planning and urban growth using remote sensing and spatial metrics. *Landscape and Urban Planning, 100*(3), 223–230.

Poelmans, L., & Rompaey, A. (2009). Detecting and modelling spatial patterns of urban sprawl in highly fragmented areas: a case study in the Flunders-Brussels region. *Landscape and Urban Planning, 93*(1), 10–19.

Pontius, R. G., Jr. (2000). Quantification error versus location error in comparison of categorical maps. *Photogrammetric Engineering and Remote Sensing, 66*(8), 1011–1016.

Pontius, R. G., Jr., & Batchu, K. (2003). Using the relative operating characteristic to quantify certainty in prediction of location of land cover change in India. *Transactions in GIS, 7*(4), 467–484.

Pontius, R. G., Jr., & 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.

Pontius, R. G., Jr., & Schneider, L. C. (2001). Land-cover change model validation by an ROC method for the Ipswich watershed, Massachusetts, USA. *Agriculture, Ecosystems and Environment, 85*(1–3), 239–248.

Pontius, R. G., Jr., & Si, K. (2014). The total operating characteristic to measure diagnostic ability for multiple thresholds. *International Journal of Geographical Information Science, 28*(3), 570–583.

Pontius, R. G., Jr., Cornell, D. C., & Hall, C. A. S. (2001). Modeling the spatial pattern of land-use change with GEOMOD2: application and validation for Costa Rica. *Agriculture, Ecosystems and Environment, 85*(1–3), 191–203.

Rafiee, R., Mahini, A. S., & Khorasani, N. (2009). Assessment of changes in urban green spaces of Mashad city using satellite data. *International Journal of Applied Earth Observation and Geoinformation, 11*(6), 431–438.

Saaty, T. L. (1980). *The analytic hierarchy process: planning, priority setting, resource allocation*. New York: McGraw Hill.

Sakieh, Y. (2013). *Urban sustainability analysis through the SLEUTH urban growth model and multi criteria evaluation: a case study of Karaj City*. Dissertation, University of Tehran.

Sakieh, Y., Amiri, B. J., Danekar, A., Feghhi, J., & Dezhkam, S. (2014a). Scenario-based evaluation of urban development sustainability: An integrative modeling approach to compromise between urbanization suitability index and landscape pattern. *Environment, Development and Sustainability, 17*(6), 1343–1365.

Sakieh, Y., Amiri, B. J., Danekar, A., Feghhi, J., & Dezhkam, S. (2014b). Simulating urban expansion and scenario prediction using a cellular automata urban growth model, SLEUTH, through a case study of Karaj City, Iran. *Housing and Built Environment, 30*(4), 591–611.

Sakieh, Y., Salmanmahiny, A., Jafarnezhad, J., Mehri, A., Kamyab, H., & Galdavi, S. (2015). Evaluating the strategy of decentralized urban land-use planning in a developing region. *Land Use Policy, 48*, 534–551.

Silva, E. A., & Clarke, K. C. (2002). Calibration of the SLEUTH urban growth model for Lisbon and Porto, Portugal. *Computers, Environment and Urban Systems, 26*(6), 525–552.

Soares-Filho, B. S., Cerqueira, G. C., & Pennachin, C. L. (2002). DINAMICA—a stochastic cellular automata model designed to simulate the landscape dynamics in an Amazonian colonization frontier. *Ecological Modelling, 154*(3), 217–235.

Stevens, D., & Dragicevic, S. (2007). A GIS-based irregular cellular automata model of land-use change. *Environmental and Planning B: Planning and Design, 34*(4), 708–724.

Sullivan, D. O., & Torrens, P. M. (2000). Cellular models of urban systems, CASA working paper series, paper 22, available online at: www.casa.ucl.uk.

Su, S., Xiao, R., & Zhang, Y. (2011). Multi-scale analysis of spatially varying relationships between agricultural landscape patterns and urbanization using geographically weighted regression. *Applied Geography, 32*(2), 360–375.

Tang, J., Wang, L., & Yao, Z. (2008). Analyses of urban landscape dynamics using multi-temporal satellite images: a comparison of two petroleum-oriented cities. *Landscape and Urban Planning, 87*(4), 269–278.

Tian, G., Ouyang, Y., Quan, Q., & Wu, J. (2011). Simulating spatiotemporal dynamics of urbanization with multi-agent systems—a case study of the Phoenix metropolitan region, USA. *Ecological Modelling, 222*(5), 1129–1138.

Tobler, W. R. (1969). Geographical filters and their inverses. *Geographical Analysis, 1*(3), 234–253.

Turner, M. G., Gardner, R. H., & O'Neill, R. V. (2001). *Landscape ecology in theory and practice: pattern and process*. New York: Springer.

Verburg, P. H., Soepboer, W., Veldkamp, A., Limpiada, R., Espaldon, V., & Mastura, S. S. A. (2002). Modeling the spatial dynamics of regional land use: the CLUE-S model. *Environmental Management, 30*(3), 391–405.

Wang, H., He, S., Liu, X., Dai, L., Pan, P., Hong, S., & Zhang, W. (2012). Simulating urban expansion using a cloud-based cellular automata model: a case study of Jiangxia, Wuhan, China. *Landscape and Urban Planning, 110*, 99–112.

Weng, Y. C. (2007). Spatiotemporal changes of landscape pattern in response to urbanization. *Landscape and Urban Planning, 81*(4), 341–35.

Wu, X., Hu, Y., He, H. S., Bu, R., Onsted, J., & Xi, F. (2009). Performance evaluation of the SLEUTH model in the Shenyang metropolitan area of northeastern China. *Environmental Modeling and Assessment, 14*(2), 221–230.

Zadeh, L. A. (1965). Fuzzy sets. *Information and Control, 8*(3), 338–353.