ORIGINAL ARTICLE

Predicting Spatial and Decadal LULC Changes Through Cellular Automata Markov Chain Models Using Earth Observation Datasets and Geo-information

Sudhir Kumar Singh · Sk. Mustak ·
Prashant K. Srivastava · Szilárd Szabó · Tanvir Islam

Abstract Remote sensing and GIS are important tools for studying land use/land cover (LULC) change and integrating the associated driving factors for deriving useful outputs. This study is based on utilization of Earth observation datasets over the highly urbanized Allahabad district in India. Allahabad district has experienced intense change in LULC in the last few decades. To monitor the changes, advanced techniques in remote sensing and GIS, such as Cellular Automata (CA)-Markov Chain Model (CAMCM) were used to identify the spatial and temporal changes that have occurred in LULC in this area. Two images, 1990 and 2000, were used for calibration and optimization of the Markovian algorithm, while 2010 was used for validating the predictions of CA-Markov using the ground based land cover image. After validating the model, plausible future LULC changes for 2020 were predicted using the CAMCM. Analysis of the LULC pattern maps, achieved through classification of multi-temporal satellite datasets, indicated that the socio-economic and biophysical factors have greatly influenced the growth of agricultural lands and settlements in the area. The two

S. K. Singh (✉)
K. Banerjee Centre of Atmospheric and Ocean Studies, IIDS, Nehru Science Centre, University of Allahabad, Allahabad 211002, India
e-mail: sudhirinjnu@gmail.com

S. Mustak
School of Studies in Geography, Pt. Ravishankar Shukla University, Raipur, Chhattisgarh, India

P. K. Srivastava
Hydrological Sciences, NASA Goddard Space Flight Center, Greenbelt, MD, USA

P. K. Srivastava
Earth System Science Interdisciplinary Center, University of Maryland, College Park, MD, USA

S. Szabó
Department of Physical Geography and Geoinformatics, University of Debrecen, Debrecen, Hungary

T. Islam
NOAA/NESDIS Center for Satellite Applications and Research, College Park, MD, USA

T. Islam
Cooperative Institute for Research in the Atmosphere, Colorado State University, Fort Collins, CO, USA

urbanization indicators calculated in this study viz. Land Consumption Ratio (LCR) and Land Absorption Coefficient (LAC) were also used, which indicated a drastic change in the area in terms of urbanization. The predicted LULC scenario for year 2020 provides useful inputs to the LULC planners for effective and pragmatic management of the district and a direction for an effective land use policy making. Further suggestions for an effective policy making are also provided which can be used by government officials to protect this important land resource.

Keywords LULC · Cellular automata · Markov chain analysis · Remote sensing and GIS · Predictive modeling · India

1 Introduction

Humans have largely influenced the earth environment by changing the land use/land cover (LULC) dynamics. LULC modeling is a thrust area of research in recent decades for solving the problem that arises due to the modification and conversion of LULC (Lambin et al. 2001). Effective analysis and monitoring of land cover require a substantial amount of data about the Earth surface and the living habitats (Schneider and Woodcock 2008). Most of the time, anthropogenic effects initiate the development of LULC change (LULCC) followed by natural processes (Niemelä et al. 2000; Srivastava et al. 2012b). Urbanization is a global trend that has been accelerated significantly since the 20th century (Srivastava et al. 2010) which acts as a major driver of change in landscape structure and functions (Antrop 2004; Banerjee and Srivastava 2013). In contrast, LULCC is one of the most important variables that decide most of the resource planning and control measures (Weng 2007), as well as pollution of many natural resources such as water and soil (Grimm et al. 2008; Srivastava et al. 2012a). Several landscape pattern scenarios, considering changing environmental conditions for, e.g., climate change, land use change, establishing new road networks etc., have also been found responsible for urbanization by several researchers (Lambin et al. 2001; Satterthwaite 2009; Csorba and Szabó 2012; Srivastava et al. 2012c; Vaz et al. 2012; Patel and Srivastava 2013).

Remote sensing (RS) provides synoptic and continuous data used by many researchers in the field of LULCC studies (Gupta and Srivastava 2010; Szabó et al. 2012). Satellite datasets, like Landsat, IRS and IKONOS, provide valuable information that can be used as an input for many prediction studies. Many researchers have applied geospatial techniques for natural resource management and planning purposes (Singh et al. 2010; Túri and Szabó 2008; Patel et al. 2012; Patel et al. 2013; Singh et al. 2013a; Srivastava et al. 2013b). Cellular Automata (CA) is a popular technique which works on a uniform grid-based principle, and has been utilized in urban growth modeling for simulating spatial processes (Wu and Webster 2000; O'Sullivan 2001; Wu 2002). The merit of CA is its ability to represent any complex systems through a small set of simple rules and states, which makes it useful for urbanization studies (Li and Yeh 2000; Wu and Webster 2000). The CA works on what-if scenarios; therefore, it can be utilized for planning related activities (Irwin et al. 2009; Araya and Pedro Cabral 2010).

The coupling of CA with Markov Chain Model (CAMCM) provides a robust approach in spatial and temporal dynamics modeling of LULCC, because RS and GIS data can be efficiently incorporated (Kamusoko et al. 2009; Steeb 2011) and provide a more detailed information on a synoptic scale. In the CAMCM, the MCM controls the temporal dynamics among LULC classes relying on transition probabilities (Kamusoko et al. 2009), while CA helps in determining the local rules either by the CA spatial filter or transition potential maps (Sang et al. 2011). The potential application of the CAMCM in an urban environment has been

recognized by many researchers by combining biophysical and socioeconomic data for simulation of accurate LULC in plausible future (Chen 2006; Kamusoko et al. 2009; Guan et al. 2011; Wang and Zhang 2001; Guan et al. 2011; Jokar Arsanjani et al. 2011; Jokar Arsanjani et al. 2013; Yang et al. 2014).

In this study, an integrated approach was applied to reach the following objectives: (i) to analyze the temporal and spatial changes in the study area in 1990–2000–2010; (ii) to simulate and predict land use change for years 2010 and 2020 based on CAMCM to preserve the unique natural characteristics; (iii) to provide values for two urban indicators, LCR and LAC. The outcome of the study will be highly useful to urban planners, resource managers and policy makers and could be utilized in other geographical locations. This paper is divided into four sections. The first section, i.e., introduction provides a brief background of other related studies while the second section deals with the materials and methodology, which includes description of study area, socio-economic status, Earth Observation datasets and algorithm structure of CAMCM. The third section, results and discussion, provides outcomes of image classification, accuracies and the performance of CAMCM model followed by some suggestions and recommendations which can be utilized for protection and conservation of land resources. The fourth section provides the final remarks and conclusions of this study.

2 Materials and methods

2.1 Study area

The study area, Allahabad district of Uttar Pradesh State, covers 5246 km^2; it is located between north latitudes 24°47′ to 25°47′ and east longitudes 81°09′ and 82°21′. The area is surrounded by the Ganga and Yamuna Rivers (Fig. 1). The district climate is generally characterized by cold and hot seasons. A variety of seasons are recognized: generally, the period from November to February is the winter season, while the summer starts from May and is followed by monsoon which continues up to the late September. The monsoon season is responsible for a high annual rainfall (1027 mm) in the area.

2.2 Population status

Allahabad is a district in the Uttar Pradesh State. The district has a total population of 5,954,391 people with 3,131,807 males and 2,822,584 females (Census of India 2011). A high increase in population is also observed in rural pockets. From the demographical statistics of 2011, each year about a total of one million people were added both in the urban and rural area of the district. The decadal population growth rate of the study area is given in Table 1. In year 2001, the population density was 901 persons per km^2 while in year 2011 it reached to 1087 persons per km^2. The average literacy rate in year 2001 was 62.11 and increased to 72.32 % in year 2011.

2.3 Satellite datasets

In the context of this study, multi-date and multi-sensor satellite images were collected along with field investigation and socioeconomic statistical data since 1990. The Landsat satellite data were selected because of their availability with the medium spatial and high temporal
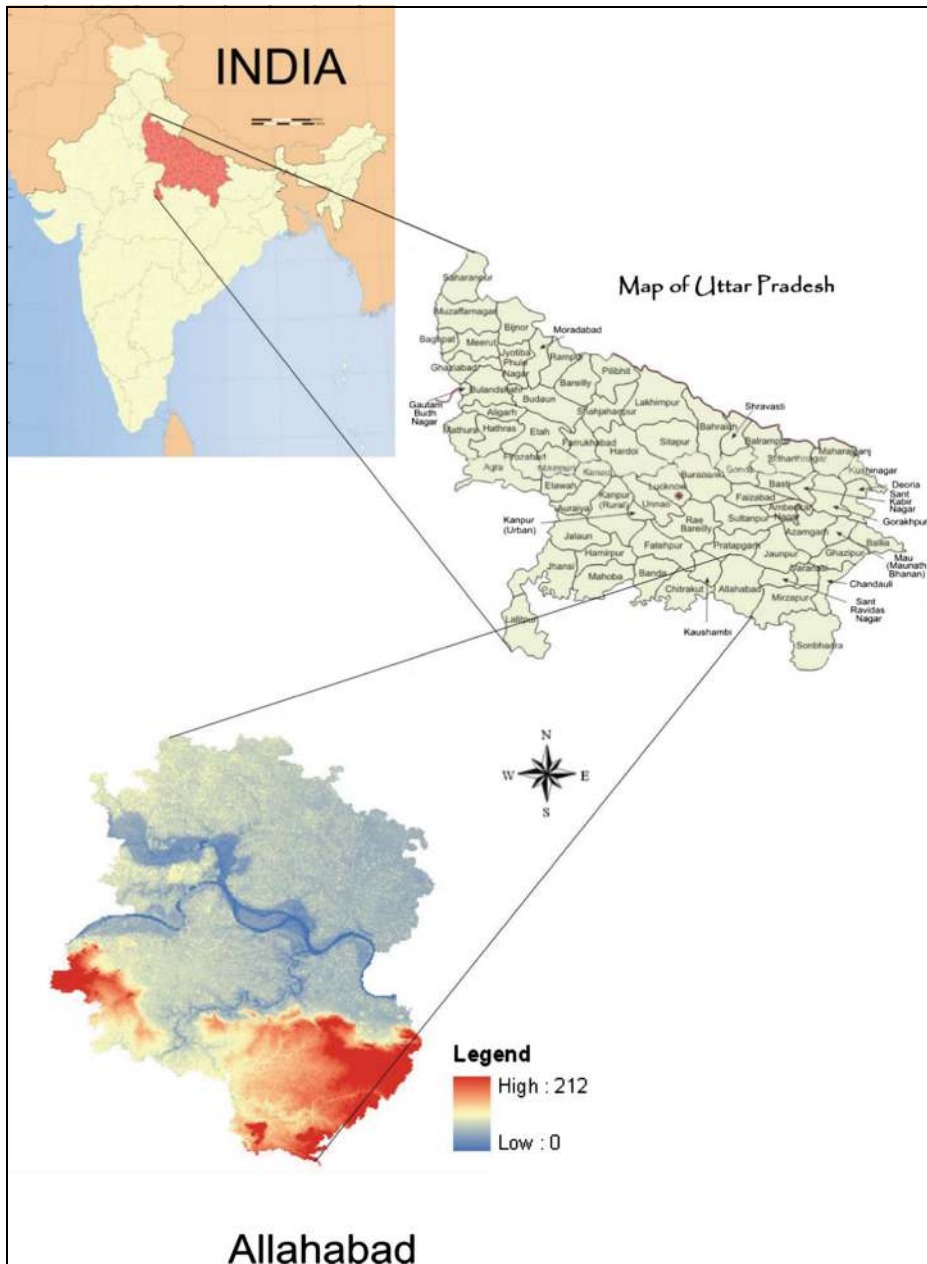
Fig. 1 Geographical location of the study area

resolutions. It has been acquired for three distinct years 1990, 2000, and 2010 for land use mapping purposes (Table 2). The acquired satellite data for years 1990, 2000 and 2010 were imported in ERDAS 9.1 for geometric correction. The collateral and auxiliary datasets, such as Digital Elevation Model (DEM) of Shuttle Radar Topography Mission (SRTM) were obtained from the United States Geological Survey (USGS) using the website: http://srtm.usgs.gov/,

Table 1 Decadal growth rate of population in district Allahabad

Years	Rural	Urban
1901–1911	−1	−8
1911–1921	−4	−7
1921–1931	5	13
1931–1941	18	40
1941–1951	11	22
1951–1961	19	21
1961–1971	20	22
1971–1981	26	43
1981–1991	29	32
1991–2001	26.94	28.22
2001–2011	20.18	22.36

while the drainage and major road networks were obtained after digitization of Survey of India topographical sheets. The field observation and surveys were made to collect ground truth datasets. The LULC classification was performed using the unsupervised classification technique for years 1990, 2000 and 2010 of satellite data. The classification accuracy assessment was performed for each LULC map based on the collected ground control points using Garmin eTrex$^@$ 10 GPS device with ±15 m positional accuracy and additional information. The thematic layers of supporting database, including different demographic and biophysical drivers of LULC change, were collected from statistical handbook of Census of India and from various other central and state government agencies.

2.4 CA-Markov chain model description

The CAMCM has the capability to simulate changes and predict decadal variations using satellite images (Kamusoko et al. 2009; Jokar Arsanjani et al. 2013). CAMCM is developed combining two rationales- Cellular automata and a transition probability matrix based on land cover changes generated by the cross tabulation of the two images adjusted by the proportional errors. A cellular automation is a cellular entity and is based on proximity concept, which indicates that the regions which are closer to the existing areas of the same class are more probable to change to a different class, conditioned by Markov transition rule and adjacent neighbors. The transition probability matrix determines the likelihood that a cell or pixel will move from a land use category or class to every other category (Schweitzer 1968). The MCM

Table 2 Specification of the satellite data sets used

Year	Satellite/Sensor	Spatial resolution (m)	Path/row	Available band combination	Date of acquisition
1990	Landsat, TM	30	p143 r042 and p143 r043	1,2,3,4,5,6,7	17 November 1990/16 October 1990
2000	Landsat, ETM$^+$	30	p143 r042 and p143 r043	1,2,3,4,5,6, 7	20 November 2000/20 November 2000
2010	Landsat, ETM$^+$	30	p143 r042 and p143 r043	1,2,3,4,5,6,7	18 October 2010/18 October 2010

analysis describes the probability of land cover change from one period (t1) to another (t2) (Bruzzone and Serpico 1997). In this study, a model was developed from transition probability matrix for each of the land cover classes representing the changes from the periods 1990 and 2000. The Markov chain process can be expressed as given in Eq. (1) (Memarian et al. 2012):

$$F_x\left(X\left(t_{n+1}\right) \le x_{n+1} \middle| X\left(t_n\right) = x_n, X\left(t_{n-1}\right) = x_{n-1}, ---, X\left(t_1\right) = x_1\right)$$

$$= \left(X\left(t_{n+1}\right) \le x_{n+1} \middle| X\left(t_n\right) = x_n\right) \tag{1}$$

where: $X(t)$ represents the Markovian chain process for a particular time (t), t_n defines the present time while the time for changes in the future is defined by t_{n+1}. Similarly, the notation t_{n-1} is used to denote the previous changes.

Equation (2), as given in Memarian et al. (2012), denotes the transition probability from state i to state j, while $X[k]$ represents the states $\{x1, x2, x3, \cdots\}$. However, many times Markov chain follow a finite number of states (n); in the latter case the transition probability matrix can be defined by Eq. (3) (Memarian et al. 2012):

$$P_{i,j} = P_r\left(X[k+1] = j \middle| X[k] = i\right) \tag{2}$$

$$\begin{bmatrix} P_{1,1} & P_{1,2} & - & - & - & - & - & P_{1,n} \\ P_{2,1} & P_{2,2} & - & - & - & - & - & P_{2,n} \\ - & - & - & - & - & - & - & - \\ - & - & - & - & - & - & - & - \\ P_{n,1} & P_{n,2} & - & - & - & - & - & P_{n,n} \end{bmatrix} \tag{3}$$

2.5 Implementation and accuracy assessment of CAMCM

The suitability maps were created by applying the Multicriteria Evaluation (MCE) module (Weight Liner Combination Method, WLC) and two constraints of built-up and water bodies using MCE module (Boolean intersection method, BIM). The criteria of the suitability maps were determined based on the behavioral pattern of past land transformation scenario in the landscape. MCE is matrix based optimization linear algorithm which allows data set of different criteria and constraints based on their weight derived from decision making process. MCE is a decision making tool embedded in IDRISI Kilimanjaro (Figs. 2 and 3).

Model calibration and validation are important steps in the modeling process especially when one is predicting future decadal changes, where no datasets are available for accuracy of prediction (Srivastava et al. 2013a). One way to evaluate the predictive power of the model is to compare the result of the simulation with the present day changes using Kappa statistics such as Kappa for no information (Kno), Kappa for location (Klocation) and Kappa for quantity (Kquantity), and, once the model is optimized with satisfactory performances, then use it for predicting the future changes. The overall accuracy of simulation run is defined by Kno, which is the variation of the standard kappa index of agreement. The Klocation validates the ability of simulations to predict the location while Kquantity predicts the quantity (Pontius and Schneider 2001). When the value of all three Kno, Klocation and Kquantity indices are equal to 1 then the simulation is defined as perfect, and if it is 0 then the simulation is considered as imperfect (Pontius and Schneider 2001; Pontius 2002). The predictive power of CACAM is stronger, when the efficiency
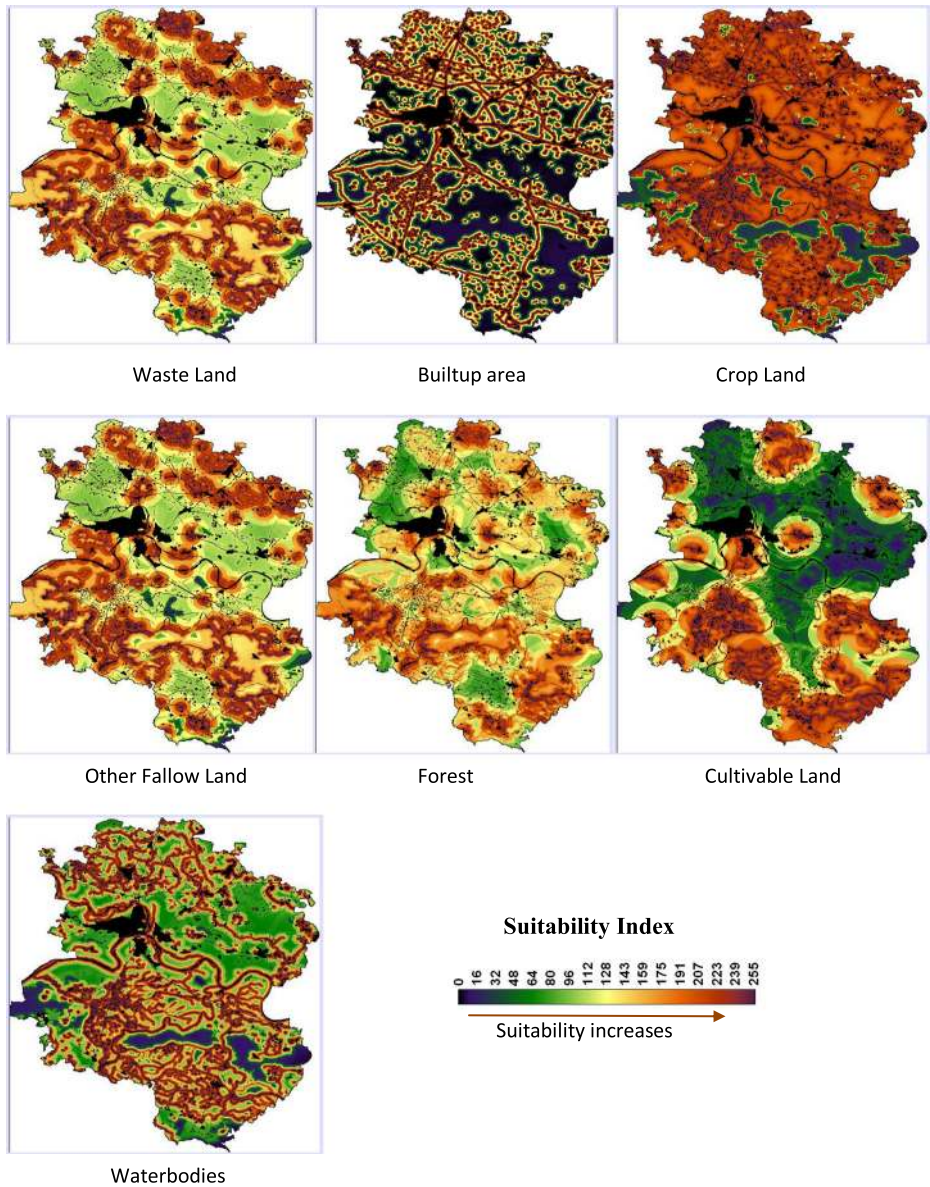
Waste Land   Builtup area   Crop Land

Other Fallow Land   Forest   Cultivable Land

Waterbodies

Suitability Index

0 16 32 48 64 80 96 112 128 143 159 175 191 207 223 239 255

Suitability increases

Fig. 2 Transitional suitability maps for different land use/land cover classes using MCE module (Weight Leaner Combination Method, WLC)

reaches≥80 % for initial year images (2000 and 2010) then it will be reasonably good to make plausible future projections.

In this study, the prediction accuracy of the validation images are also presented through agreement and disagreement statistics between the simulated map and the reference map of 2010. The precision of simulation or classification image results, pixel-by-pixel, is accessed via the kappa accuracy index. This statistic measures the goodness of fit between model
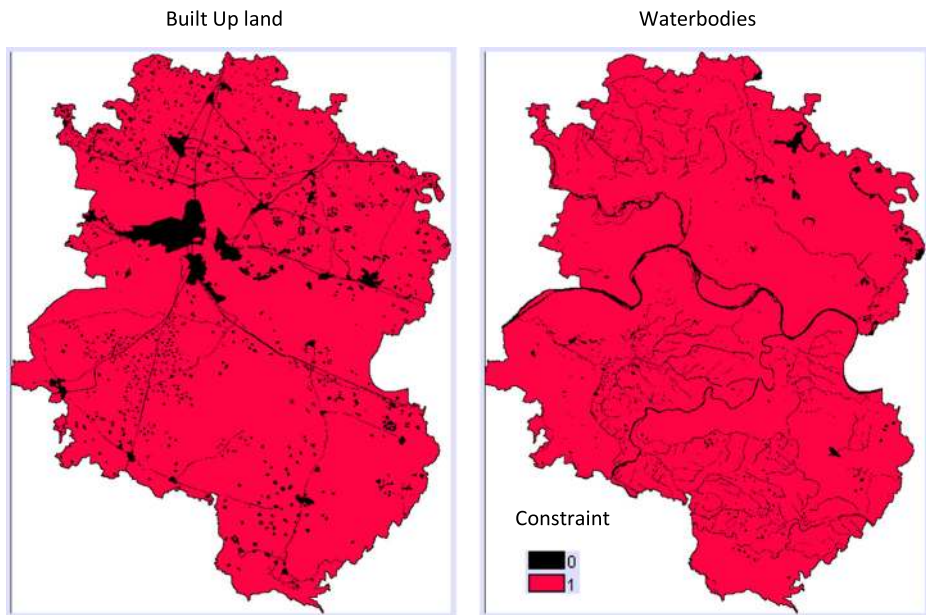
Built Up land   Waterbodies

Constraint

■ 0
■ 1

Fig. 3 Constraint layer of built-up and water bodies as obtained using MCE module (Boolean Intersection Method, BIM)

predictions and reality, corrected for accuracy by chance (Mukherjee et al. 2009). According to Omar et al. (2014) for the Kappa variations, the expressions used for the summary statistics are listed as Eqs. (4–7):

$$Kno = (M(m) N(n)) / (P(p) - N(n)) \tag{4}$$

$$Klocation = (M(m) N(m)) / (P(m) - N(m)) \tag{5}$$

$$Kquantity = (M(m) H(m)) / (K(m) - H(m)) \tag{6}$$

$$Kstandard = (M(m) N(n)) / (P(p) - N(n)) \tag{7}$$

where no information is defined by $N(n)$, medium stratum level information by $H(m)$, medium grid cell level information by $M(m)$, perfect grid cell-level information given imperfect stratum-level information by $K(m)$ mean, and perfect grid cell-level information across the landscape by $P(p)$.

2.6 Indicators of urbanization LAC and LCR

The expansion of urban area is generally related in search of better infrastructural facility. The forces and drivers of urban expansion are different in each region. The change in urban area is defined by the expansion index which indicates the change in the urban area in terms of population and spatial expansion. Two different indices are used for measuring urban expansions: LCR and LAC (Sharma et al. 2012; Pandey et al. 2012; Kumar et al. 2013). The value of index LCR indicates a progressive spatial expansion of the land cover class and measures the

compactness, while LAC is directly related to population expansion. The following Eqs. (8, 9, 10) are employed for the estimation of these indices.

$$LCR = {^A}/{_P} \tag{8}$$

where: A is the areal extent of the land cover class in hectares; and P is the population.

$$LAC = \frac{(A_2 - A_1)}{(P_2 - P_1)} \tag{9}$$

where: A_1 and A_2 are the areal extents (ha); P1 and P2 are the population figures for the early and later years respectively.

The predicted population of the desired time period can be calculated by using the following formula (Siegel and Swanson 2004):

$$P = P_b \left(1 + R \big/ 100 \right)^n \tag{10}$$

where: P is population of the desired time period; P_b is the Population of base year; R is the Rate of growth of population; and n is the number the years.

3 Results and discussion

3.1 Analysis of land use/ land cover change

After the classification of satellite data, the reliability of results depends on the overall accuracies of the classified images. The result of this process indicates whether the LULC changes have been accurately identified and extracted. In this study, the accuracy assessment of the classified data of 1990, 2000 and 2010 confirmed that the results are reasonable and satisfactory for any applications (Table 3). According to Anderson (1976), approving the reliability of classified images is through estimation of overall accuracies. The overall accuracy should clearly exceed the minimum acceptable standard of ≥85 % stipulated by the USGS classification scheme. The overall classification accuracies meet this standard criterion in our study. The results of LULC category achieved through classification are presented in Table 4. The historical land use pattern of the district has been changing steadily at a slow pace which also brings adverse effects on natural resources of the region (Singh et al. 2013b, c; Singh et al. 2014). The area has witnessed increase in urbanization which brought on the fundamental land use change during the last two decades; this is proved by the rapid increase of urban and rural population. The overall results of LULC distribution for years 1990, 2000, 2010, and 2020 showed that the agricultural land was the primary dominant land cover category (Table 4); followed by the built up area. Similar trend is also prevalent in 2020 classified imagery. The built up area showed an overall increasing trend, while the waste land and other fallow land showed a decreasing trend. From the analysis of probability matrix (Table 5) and transition area matrix (Table 6), we estimated the future LULCC such as for years 2010 and 2020. Specifically, the built up area increased from 555.89 km^2 in 1990 to 744.16 km^2 in 2010 with a percentage increase of 33.9 %. This increase probably took place due to migration of population towards the district Allahabad, which offers better education activities, and business and job opportunities. The total cultivable land area decreased from 246.19 km^2 in 1990

Table 3 Accuracy assessment results of different land use/land cover classes of different years

LULC Class	1990		2000		2010	
	Producer's accuracy (%)	User's accuracy (%)	Producer's accuracy (%)	User's accuracy (%)	Producer's accuracy (%)	User's accuracy (%)
Built up area	85.71	100	86.92	100	87.5	100
Agriculture	93.33	82.35	96.23	82.30	100	81.82
Forest	62.5	83.33	63.01	84.21	89	92
Water body	83.33	100	93.24	100	100	100
Cultivable land	100	50	87.21	79	87.5	70
Other fallow land	88	89	82.01	89	88.89	100
Waste land	87	85.02	89	88	89	90
Overall accuracy	80.39		85.36		88.10	
Kappa coefficient	0.7599		0.83211		0.8522	

to 116.97 km^2 in 2010, i.e., by 52.5 %. The decrease is due to expansion in urban area. The area of other fallow land decreased from 186.04 km^2 in 1990 to 179.91 km^2 in 2010, i.e., by 32.4 %, which can be attributed to urbanization, some reforestation and industrial activities. Some change in water body class was also seen, which decreased from 328.08 km^2 in 1990 to 298.00 km^2 in 2010, i.e., by 9.2 %. This decrease in water body was due to improper management and over-exploitation of water. The forest area, which was 335.87 km^2 in 1990 has decreased to 311.92 km^2 in 2010, i.e., 7.1 % decrease; this can be attributed to deforestation activities for various purposes like agriculture, urbanization etc. There was a continuous change in the agricultural area which increased from 3638.40 km^2 in 1990 to 3793.25 km^2 in 2010, i.e., by 4.25 %. It can be attributed to decrease in fallow and waste land area.

3.2 Analysis of LULC transition probabilities matrix

For this purpose, two different classified images of the years 1990 and 2000 are used to predict the LULCC for year 2010 (simulated image) using CAMCM. Further, it is validated with the

Table 4 Temporal distribution of land use/land cover distribution by years

Land use land cover distribution (1990,2000,2010, 2020)

Land use/Land Cover Categories	1990		2000		2010		2020		
	Area (km^2)	Area (%)	Area (km^2)	Area (%)	Area (km^2)	Area (%)	Area (km^2)	Area (%)	No. Pixel
Forest	335.87	6.03	248.98	4.47	311.92	5.6	293.33	5.26	123,702
Waste land	279.62	5.02	234.50	4.21	179.91	3.23	222.45	3.99	547,172
Built up area	555.89	9.98	647.80	11.63	744.16	13.36	968.64	17.39	1,076,275
Cultivable land	246.19	4.42	193.28	3.47	116.97	2.1	23.47	0.42	26,087
Agriculture	3638.40	65.32	3739.22	67.13	3793.25	68.1	3381.79	60.72	3,679,774
Water body	328.08	5.89	324.18	5.82	298.00	5.35	423.85	7.61	470,947
Other fallow land	186.04	3.34	182.14	3.27	125.88	2.26	255.80	4.59	506,454
Total	5570.09	100	5570.09	100	5570.09	100	5570.09	100	

Table 5 Probability of LULC changes in 2010 by percentage

1990/2000	Water body	Cultivable land	Built up area	Agriculture	Forest	Other fallow land	Waste land
Water body	0.4578	0.0076	0.0201	0.4842	0.0000	0.0235	0.0067
Cultivable land	0.0017	0.0001	0.0160	0.0441	0.0020	0.9365	0.0017
Built up area	0.0010	0.0027	0.9782	0.0000	0.0020	0.0106	0.0085
Agriculture	0.0955	0.0120	0.0710	0.7419	0.0001	0.0382	0.0413
Forest	0.0076	0.0044	0.0071	0.2483	0.3001	0.3940	0.0386
Other fallow land	0.0230	0.1434	0.0094	0.3485	0.0010	0.2830	0.1927
Waste land	0.1415	0.0094	0.3585	0.4811	0.0010	0.0094	0.0002

classified image of year 2010 (reference imagery). Similarly for years 2000 and 2010 classified images are used for the prediction of 2020 LULC classes. To estimate the accuracy of the classified and predicted images of 1990, 2000 and 2010, a detailed accuracy analysis was conducted. After a satisfactory performance obtained during the trial periods through CAMC M, attempts were made for generating 2020 LULC scenarios. The changes were depicted through the 7×7 LULC class matrix table, in which rows represent the previous land cover categories during the time (t1), while the column represents the later land cover categories (t2). For example, the row in 1990 represents actual LULC classes while the column (year 2000) categories, and helps in simulation of LULC for year 2010 (Table 5). Similarly, rows represent year 2000 and column represents year 2010, and predicts the LULC of year 2020. Further, this matrix represents the prior probabilities in LULC classes, which is then used to predict the LULC of 2010. After validating the predicted 2010 image through the 2010 (real image), the 2020 future LULC scenario is generated.

The information generated from the model prediction is presented in Table 5 (probability in %), and indicates that in year 2000 the water body class has a probability of 48.4 % to change into agriculture. This may be attributed to change of wetland for agricultural activities. The cultivable land indicates a highest transformation, with probability 0 % to remain as cultivable land, which means it may be completely transformed into other classes such as other fallow land (93.7 %) and agriculture (4.4 %). The built up area class indicates a highest stability to remain as built up area with a probability 97.8 %. Similarly, the agriculture class was also found stable with probability 74.2 % to remain as agriculture class. The forest class has 30 % probability to remain as forest. It has 24.83 % probability to change to agriculture class and 39.4 % to other fallow land class, which can be attributed to increased deforestation activities

Table 6 Expected transition in 2010 by number of pixels

| 1990/2000	Water body	Cultivable land	Built up area	Agriculture	Forest	Other fallow land	Waste land
Water body	192,473	3205	8431	203,578	113	9888	2818
Cultivable land	253	105	2362	6495	240	138,084	253
Built up area	120	900	330,263	203	123	3577	2873
Agriculture	490,597	61,385	364,875	3,810,333	94	196,352	211,972
Forest	160	250	213	187	133,702	126	153
Other fallow land	9257	57,779	3806	140,450	145	114,046	77,648
Waste land	37,520	2501	95,052	127,569	123	2501	170
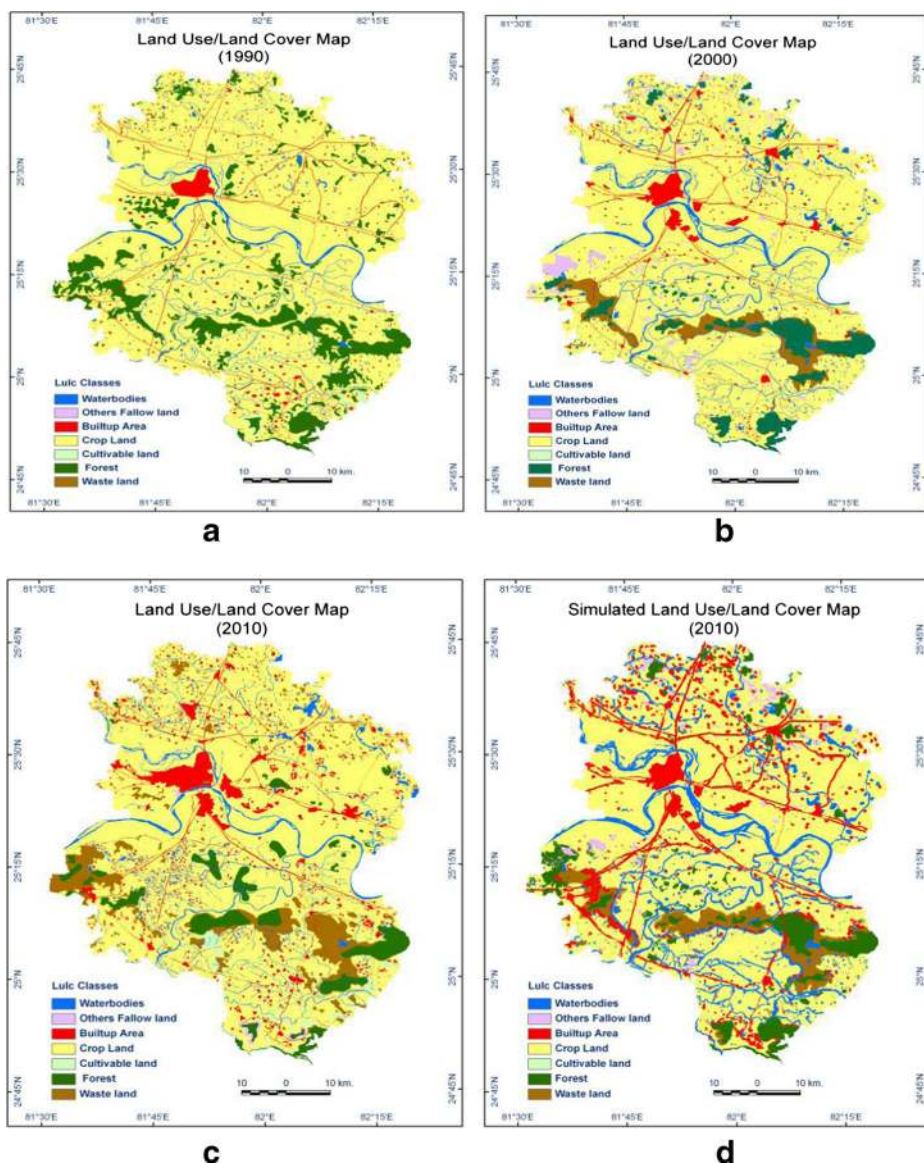
Fig. 4 Land use/land cover map by years **a** 1990, **b** 2000, **c** 2010 real, **d** 2010 simulated and **e** 2020 simulated

pertaining increasing urban and rural population. The class other fallow land also shows instability with a probability of 19.27 % to remain in the same class while having a probability 34.85 % to change into agriculture class. It means the other fallow land is converting into agricultural field. The waste land class has 0 % probability to remain as waste land. It indicates that it has been transformed into agriculture class with a probability 48.11 % and waste land has a probability 35.85 % to convert into built up area. It is due to development of few thermal power plants. The analysis indicates that the transformation of the different land use classes into forest class has lower probability as compared to transformation into agriculture class. It
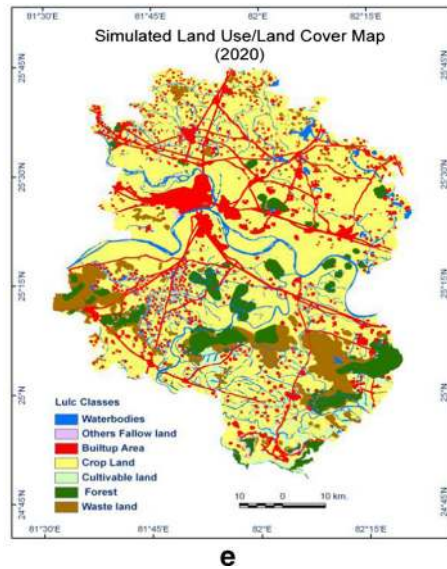
Fig. 4 (continued)

may be due to growing population and urbanization and development of small scale industries in the region (Singh et al. 2013c).

3.3 LULC changes for 2020 simulated map

The changes in LULC categories from 1990 to 2020 can be seen in Fig. 4e and Table 4. The probability for LULCC in 2020 is presented in Table 7. The water body has 66.91 % probability to change into agriculture land use class while it has 23.90 % to remain as water body. The cultivable land has 51.96 % probability to convert into waste land, and 36.29 % probability to change into agriculture class. The built up area has 97.60 % probability to remain as built up area. The agriculture class has 70.99 % probability to remain as agriculture category and 10.82 % probability to change into built up area. The forest has 70.47 % probability to change into agriculture class and 14.99 % probability to remain as forest. The other fallow has 31.36 % probability to remain as other fallow land, and 35.05 % probability to change into waste land and 23.12 % to change into agriculture class. The waste land has

Table 7 Probability of LULC changes in 2020 by percentage

2000/2010	Water body	Cultivable land	Built up area	Agriculture	Forest	Other fallow land	Waste land
Water body	0.2390	0.0036	0.0341	0.6691	0.0058	0.0152	0.0331
Cultivable land	0.0176	0.0001	0.0249	0.3629	0.0201	0.0549	0.5196
Built up area	0.0000	0.0003	0.9760	0.0002	0.0010	0.0104	0.0124
Agriculture	0.0682	0.0015	0.1082	0.7099	0.0175	0.0473	0.0473
Forest	0.0254	0.0002	0.0450	0.7047	0.1499	0.0300	0.0450
Other fallow land	0.0366	0.0398	0.0116	0.2312	0.0165	0.3136	0.3505
Waste land	0.0097	0.0046	0.0150	0.3264	0.0146	0.3045	0.3252

32.64 % probability to change into agriculture class. The information given in Table 8 shows that about 81,713 pixels of water body class will remain unchanged. It also indicates that approx. 228,728 pixels are expected to change into agriculture class. The change could be due to urbanization as well as overexploitation of natural resources, such as water bodies and forest. There could be some uncertainty in prediction, which might be due to imagery quality, classification error and resolution of the sensors.

The analysis indicates that cultivable land will change the most, and will transform to primarily three classes that are waste land, agriculture and built up. The major transformation is of about 11,368 pixels in waste land, about 7940 pixels to agriculture, 544 pixels into built up area, and 368 into water body. The built up class showed a maximum stability, with approximately 519,147 pixels of this class remaining unchanged. This seemed to be a good projection, although a slight transformation of this class is predicted to become waste land (6577 pixels) and other fallow land (5558 pixels), as seen in Table 8. This may be due to the migration of population from their old habitat to a new habitat (Singh et al. 2014). The class agriculture, which remains the major class of the study area with 3,484,049 pixels, seems to be highly transformed to other classes of the study area with major transformation towards built up class (530,864 pixels). The overall analysis indicates satisfactory results, as the urban area and the total population are also increasing. Furthermore, the prediction signed that classes of waste land and fallow land is about to turn into the agricultural class.

The forest class indicates a major transformation with only 9955 pixels remaining in the same class while a major portion will transform into other classes. For example, around 1685 pixels will change into water body, 46,790 pixels into agriculture class. It may be ascribed to increased deforestation activities in the future. The other fallow land class will also change, in which about 132,458 pixels will transform into waste land class, 87,374 pixels into agriculture class, 15,056 pixels into other fallow land, and 13,839 pixels into water body. The waste land class will also experience a transformation in which 150,522 pixels will be converted into agricultural class and about 140,951 pixels into other fallow land class.

3.4 Modeling results and validation

Analysis of the modeling results showed that the reference and simulated maps of year 2010 were reasonably very much similar. Further, for the justification of this, a more detailed analysis was performed using the Kappa variations. When the values of these indices reach 100 %, a stronger agreement between real and simulated maps is indicated. The value of Kno gives the overall accuracy of the simulated map (88.5 %). The other model indices performed relatively well in the ability to correctly specify location (Klocation Strata: 96.5 %; Kstandard:

Table 8 Expected transition in 2020 by number of pixels

2000/2010	Water body	Cultivable land	Built up area	Agriculture	Forest	Other fallow land	Waste land
Water body	81,713	1222	11,672	228,728	1996	5209	11,310
Cultivable land	386	105	544	7940	439	1202	11,368
Built up area	140	136	519,147	130	507	5558	6577
Agriculture	334,537	7556	530,864	3,484,049	86,072	232,063	232,361
Forest	1685	120	2987	46,790	9955	1991	2987
Other fallow land	13,839	15,056	4398	87,374	6234	118,515	132,458
Waste land	4503	2108	6965	151,070	6776	140,951	150,522

Table 9 Value of annual population growth and estimated population for 1990, 2000, 2011, and 2021 of Allahabad

Year	1990	2000	2011	2021
Annual population growth	112,392	104,540	102,226	113,236
Population of desired time period	4,920,954	4,936,020	5,958,372	7,092,159

84.8 %; and Kquantity: 83.0 %). Still the intrinsic discrepancies remain for the simulated map of 2010. The probable reasons of these are inadequately suitable maps for modeling the correct phenomenon, generalizations applied in image classification to achieve better results, and the shape of the contiguity filter used. The role of suitability maps is to define the rules in the modeling process, which has great influence on the result. Further, constraints and factors also have effects on results. After defining the parameters used for the calibration and modeling, and assessing the validity, it is interesting to examine the pattern and tendency of change in a long term simulation. For examining the pattern and tendency of change in a long term simulation, the method needs calibration and assessment of validity. Based on this, the land cover projection was performed in a similar way for 2020 (Fig. 4e). CAMCM has a potential to simulate the transition among any number of classes, and its nature of simulation is bidirectional. The expected changes in the number of pixels from year 2010 to year 2020 are shown in Table 4. The model showed poor results for the waste land and water body classes, therefore, these class matrices did not predict any changes for 2020. The matrix has two components; the diagonal of the matrix indicates the number of pixels that have persisted during the simulation process, while the off-diagonal component of the matrix shows the number of pixels that have changed class during the simulation.

3.5 LCR and LAC with urbanization

The annual population growth (n) and estimated population (P_n) was also calculated (Table 9). From Table 9, the value of n changes from 112,392 to 104,540 and 104,540 to 102,226 in periods 1990–2000 and 2000–2011, respectively, whereas P_n changes from 4,920,954 to 4,936,020 and 4,936,020 to 5,958,372 in periods 1990–2000 and 2000–2011, respectively. The changes in the LAC and LCR values are shown in Table 10. LAC changed from 0.00621 to 0.00009 and from 0.00009 to 0.00018 between periods 1990–2000 and 2000–2010, respectively (with respect to the built up area), whereas LCR changed from 0.00011 to 0.00013 and from 0.00013 to 0.00012 between periods 1990–2000 and 2000–2010, respectively (with respect to the built up area), which shows how changes are governing in the study area. Table 10 deciphers how cultivable land decreased with increasing population in the same years (Table 10). The LCR values are 0.00012, 0.00011, 0.00013 (built up area) during 1990 to 2011, and on the basis of estimated population for year 2021, we can give an idea of LCR value of the built up area of Allahabad. Other classes can also be predicted on the basis of this population.

Table 10 LAC and LCR for built up area of Allahabad district

Year	1990/2000	2000/2010	1990/2010
LAC=(A_2-A_1/P_2-P_1) (Built up area)	+0.00621	0.00009	+0.00018
LCR=A/P (Built up area)	0.00011	0.00013	0.00012

(+) sign stands decrease and (−) stand for increase

3.6 Policy development for sustainable land use

The prominent features observed in the area are the growth of urban population which can be related to urbanization. The increase in urban population due to people migration in search of better infrastructural, medical and education facilities, leads to an increase of the number of residential houses in the Allahabad district. The share of urban population has also seen an increase, as expected, from 18.12 % in 1990 to 24.74 % in 2011. During this 21 year period, in general an increase in population was seen in the district which also has been affecting the land surface area on a day-by-day basis. These trends showed that there is a need for proper planning for housing, in both urban and rural areas. It is imperative to take immediate measures for conservation taking into consideration: afforestation programs by various governments, public and private agencies, including several NGOs; limited groundwater use in a controlled way coupled with grazing management; understanding the landscape and its bearings with the local climate and environment to reduce human stress in the affected areas; restoring crop production, food security and biodiversity. Moreover, different systems of agricultural practices should be introduced following sustainable land management policies, attuned to the local climatic factors to mitigate the imminent problems of arable soil degradation and reduced soil fertility.

4 Conclusions

Remote sensing and GIS are very promising tools for management of natural resources on a large spatial and temporal scale. In this study, satellite derived LULC maps were used in association with biophysical and socioeconomic data to estimate the past and present changes, and simulate future possibilities. Sophisticated algorithms such as CAMCM were used for the prediction of the expansion or loss of different land use categories with time for 2020. Extensive field visits and incorporation of local area knowledge were used along with CAMCM to simulate the LULCC through GIS based multi-criteria analysis. The results after validation seem quite satisfactory, which indicates that this methodology can be used for simulation of future LULC change. Such studies are useful for developing countries, where LULC changes occur at much faster rates. These developments have definite impacts on the environment. This study provides a strategic guide to both urban and rural land use planners for an effective land management. Our future objective is to embed the biophysical and socioeconomic factors, and role of policies that influence the urbanization.

References

Anderson JR (1976) A land use and land cover classification system for use with remote sensor data: US Government Printing Office

Antrop M (2004) Landscape change and the urbanization process in Europe. Landsc Urban Plan 67(1):9–26

Araya YH, Cabral P (2010) Analysis and modeling of urban land cover change in Setúbal and Sesimbra, Portugal. Remote Sens 2:1549–1563. doi:10.3390/rs2061549

Banerjee R, Srivastava PK (2013) Reconstruction of contested landscape: detecting land cover transformation hosting cultural heritage sites from central India using remote sensing. Land Use Policy 34:193–203

Bruzzone L, Serpico SB (1997) An iterative technique for the detection of land-cover transitions in multitemporal remote-sensing images. IEEE Trans Geosci Remote Sens 35(4):858–867

Census of India (2011) Provisional population totals, paper 1 of 2011 India, Series-1, Office of the Registrar General & Census Commissioner, New Delhi

Chen Q (2006) Cellular automata and artificial intelligence in ecohydraulics modelling. Balkema, Taylor & Francis Group, UK

Csorba P, Szabó S (2012) The application of landscape indices in landscape ecology. In: Tiefebacher J (ed) Perspectives on nature conservation–paterns, pressures and prospects. InTech, Rijeka, pp 121–140

Grimm NB, Foster D, Groffman P, Grove JM, Hopkinson CS, Nadelhoffer KJ, Pataki DE, Peters DP (2008) The changing landscape: ecosystem responses to urbanization and pollution across climatic and societal gradients. Front Ecol Environ 6(5):264–272

Guan D, Li H, Inohae T, Su W, Nagaie T, Hokao K (2011) Modeling urban land use change by the integration of cellular automaton and Markov model. Ecol Model 222(20):3761–3772

Gupta M, Srivastava PK (2010) Integrating GIS and remote sensing for identification of groundwater potential zones in the hilly terrain of Pavagarh, Gujarat, India. Water Int 35(2):233–245

Irwin EG, Jayaprakash C, Munroe DK (2009) Towards a comprehensive framework for modeling urban spatial dynamics. Landsc Ecol 24(9):1223–1236

Jokar Arsanjani J, Kainz W, Mousivand AJ (2011) Tracking dynamic land-use change using spatially explicit Markov Chain based on cellular automata: the case of Tehran. Int J Image Data Fusion 2(4):329–345

Jokar Arsanjani J, Helbich M, Kainz W, Darvishi Boloorani A (2013) Integration of logistic regression, Markov chain and cellular automata models to simulate urban expansion. Int J Appl Earth Obs Geoinf 21:265–275

Kamusoko C, Aniya M, Adi B, Manjoro M (2009) Rural sustainability under threat in Zimbabwe–simulation of future land use/cover changes in the Bindura District based on the markov-cellular automata model. Appl Geogr 29(3):435–447

Kumar P, Singh BK, Rani M (2013) An efficient hybrid classification approach for land use/land cover analysis in a semi-desert area using and LISS-III sensor. IEEE Sensors J 13(6):2161–2165

Lambin EF, Turner BL, Geist HJ, Agbola SB, Angelsen A, Bruce JW, Coomes OT, Dirzo R, Fischer G, Folke C (2001) The causes of land-use and land-cover change: moving beyond the myths. Glob Environ Chang 11(4):261–269

Li X, Yeh AGO (2000) Modelling sustainable urban development by the integration of constrained cellular automata and GIS. Int J Geogr Inf Sci 14(2):131–152

Memarian H, Balasundram SK, Talib JB, Sung CTB, Sood AM, Abbaspour K (2012) Validation of CA-Markov for simulation of land use and cover change in the Langat Basin, Malaysia. J Geogr Inf Syst:542–554. doi.org/10.4236/jgis.2012.46059

Mukherjee S, Shashtri S, Singh C, Srivastava P, Gupta M (2009) Effect of canal on LULC using remote sensing and GIS. J Indian Soc Remote Sens 37(3):527–537

Niemelä J, Kotze J, Ashworth A, Brandmayr P, Desender K, New T, Penev L, Samways M, Spence J (2000) The search for common anthropogenic impacts on biodiversity: a global network. J Insect Conserv 4(1):3–9

O'Sullivan D (2001) Graph-cellular automata: a generalised discrete urban and regional model. Environ Plann B 28(5):687–706

Omar NQ, Ahamad MSS, Hussin WMAW, Samat N, Ahmad SZB (2014) Markov CA, multi regression, and multiple decision making for modeling historical changes in Kirkuk City, Iraq. J Indian Soc Remote Sens 42(1):165–178. doi:10.1007/s12524-013-0311-2

Pandey PC, Sharma LK, Nathawat MS (2012) Geospatial strategy for sustainable management of municipal solid waste for growing urban environment. Environ Monit Assess 184(4):2419–2431

Patel DP, Srivastava PK (2013) Flood hazards mitigation analysis using remote sensing and GIS: Correspondence with town planning scheme. Water Resour Manag 27(7):2353–2368. doi:10.1007/s11269-013-0291-6

Patel DP, Dholakia MB, Naresh N, Srivastava PK (2012) Water harvesting structure positioning by using geo-visualization concept and prioritization of mini-watersheds through morphometric analysis in the lower Tapi basin. J Indian Soc Remote Sens 40(2):299–312

Patel DP, Gajjar CA, Srivastava PK (2013) Prioritization of Malesari mini-watersheds through morphometric analysis: A remote sensing and GIS perspective. Environ Earth Sci 69(8):2643–2656

Pontius RG Jr (2002) Statistical methods to partition effects of quantity and location during comparison of categorical maps at multiple resolutions. Photogramm Eng Remote Sens 68(10):1041–1050

Pontius RG Jr, Schneider LC (2001) Land-cover change model validation by an roc method for the Ipswich watershed, Massachusetts, USA. Agr Ecosyst Environ 85(1):239–248

Sang L, Zhang C, Yang J, Zhu D, Yun W (2011) Simulation of land use spatial pattern of towns and villages based on CA–Markov model. Math Comput Model 54(3):938–943

Satterthwaite D (2009) The implications of population growth and urbanization for climate change. Environ Urban 21(2):545–567

Schneider A, Woodcock CE (2008) Compact, dispersed, fragmented, extensive? A comparison of urban growth in twenty-five global cities using remotely sensed data, pattern metrics and census information. Urban Stud 45(3):659–692

Schweitzer PJ (1968) Perturbation theory and finite Markov chains. J Appl Probab:401-413

Sharma L, Pandey PC, Nathawat M (2012) Assessment of land consumption rate with urban dynamics change using geospatial techniques. J Land Use Sci 7(2):135–148

Siegel J, Swanson D (2004) The methods and materials of demography (2nd edition) Elsevier Academic Press, London

Singh SK, Singh CK, Mukherjee S (2010) Impact of land-use and land-cover change on groundwater quality in the Lower Shiwalik hills: a remote sensing and GIS based approach. Cent Eur J Geosci 2(2):124–131. doi: 10.2478/v10085-010-0003-x

Singh SK, Srivastava PK, Gupta M, Thakur JK, Mukherjee S (2013a) Appraisal of LULC of mangrove forest ecosystem using support vector machine. Environ Earth Sci, 1–11

Singh S, Srivastava PK, Pandey A (2013b) Fluoride contamination mapping of groundwater in Northern India integrated with geochemical indicators and GIS. Water Sci Technol Water Supply 13:1513–1523

Singh SK, Srivastava PK, Pandey AC, Gautam SK (2013c) Integrated assessment of groundwater influenced by a confluence river system: concurrence with remote sensing and geochemical modelling. Water Resour Manag 27:4291–4313

Singh SK, Srivastava PK, Singh D, Han D, Gautam SK, Pandey AC (2014) Modeling groundwater quality over a humid subtropical region using numerical indices, earth observation datasets, and X-ray diffraction technique: a case study of Allahabad district. India Environ Geochem Health. doi:10.1007/s10653-014-9638-z

Srivastava PK, Mukherjee S, Gupta M (2010) Impact of urbanization on LULC change using remote sensing and GIS: A case study. Int J Ecol Econ Stat 18(S10):106–117

Srivastava PK, Gupta M, Mukherjee S (2012a) Mapping spatial distribution of pollutants in groundwater of a tropical area of India using remote sensing and GIS. Applied Geomatics 4(1):21–32

Srivastava PK, Han D, Gupta M, Mukherjee S (2012b) Integrated framework for monitoring groundwater pollution using a geographical information system and multivariate analysis. Hydrol Sci J 57(7):1453–1472

Srivastava PK, Han D, Rico-Ramirez MA, Bray M, Islam T (2012c) Selection of classification techniques for LULC change investigation. Adv Space Res 50(9):1250–1265

Srivastava PK, Han D, Rico-Ramirez MA, Islam T (2013a) Sensitivity and uncertainty analysis of mesoscale model downscaled hydro-meteorological variables for discharge prediction. Hydrol Process. doi:10.1002/hyp.9946

Srivastava PK, Singh SK, Gupta M, Thakur JK, Mukherjee S (2013b) Modeling impact of land use change trajectories on groundwater quality using remote sensing and GIS. Environ Eng Manag J 12(12):2343–2355

Steeb WH (2011) The nonlinear workbook: Chaos, Fractals, Cellular Automata, Neural Networks, Genetic Algorithms, Gene Expression Programming, Support Vector Machine, Wavelets, Hidden Markov Models, Fuzzy Logic With C++, Java And Symbolicc++ Programs: World Scientific

Szabó S, Csorba P, Szilassi P (2012) Tools for landscape ecological planning—scale, and aggregation sensitivity of the contagion type landscape metric indices. Carpathian J Earth Environ Sci 7:127–136

Túri Z, Szabó S (2008) The role of resolution on landscape metrics based analysis. Acta Geographica Silesiana 4(1):47–52

Vaz E, Caetano M, Nijkamp P, Painho M (2012) A multi-scenario forecast of urban change: a study on urban growth in the Algarve. Landsc Urban Plan 104(2):201–211

Wang Y, Zhang X (2001) A dynamic modeling approach to simulating socioeconomic effects on landscape changes. Ecol Model 140(1):141–162

Weng YC (2007) Spatiotemporal changes of landscape pattern in response to urbanization. Landsc Urban Plan 81(4):341–353

Wu F (2002) Calibration of stochastic cellular automata: the application to rural–urban land conversions. Int J Geogr Inf Sci 16(8):795–818

Wu F, Webster CJ (2000) Simulating artificial cities in a GIS environment: urban growth under alternative regulation regimes. Int J Geogr Inf Sci 14(7):625–648

Yang X, Zheng XQ, Chen R (2014) A land use change model: Integrating landscape pattern indexes and Markov-CA. Ecol Model 283:1–7