CrossMark

ORIGINAL ARTICLE

Numerical modelling spatial patterns of urban growth in Chandigarh and surrounding region (India) using multi-agent systems

Bhartendu Pandey[1,2] · P. K. Joshi[1,3]

Abstract This study explores application of multi agent system (MAS) to simulate spatial patterns of urban growth in Chandigarh and its surrounding region (India). A numerical simulation model is developed with MAS considering the dynamics of urban and rural population as the principal driver of urban growth. The model utilizes static and dynamic environment variables initialized using a logistic regression model. The logistic regression model uses pixel wise change/no-change information derived using Landsat TM data (1989–1999) as dependent variable and proximity, density, elevation and slope as independent variables. The optimum resolution of 90 m for modelling is decided using fractal analysis of series of transition probability surfaces generated using logistic regression from 30 to 240 m spatial resolution at 30 m interval. The model was finally calibrated using sensitivity analysis and behaviours space experiments with multiple simulation runs. A change to built-up area of 32.55 km^2 is observed during 1989–1999 and 113.51 km^2 in 1999–2009. The modelling shows a total 14.42 % disagreement between predicted map and reference map for the year 2009. The results were validated using ROC statistics and accuracy estimates with satellite data. The model was further used to predict urban growth for the year 2019. Diversity index was used to determine the potential of the model to capture overall spatial patterns of urban growth.

Introduction

Urban growth comprises of changes in physical and functional components of built environment resulting from transition of rural landscape to urban forms (Thapa and Murayama 2011). Such transformations often give rise to environmental problems such as declining ecosystem services (Su et al. 2012), net primary productivity (Xu et al. 2007), avian population (Green and Baker 2003), agricultural land (Seto et al. 2000), and increasing flood prone areas (Suriya and Mudgal 2012), solid waste (Vij 2012), human health risk (Moore et al. 2003), and others. In order to deal with such issues in social, economic and environmental dimensions, we first need to answer the questions how, why and where urban areas develop. The answers to these fundamental questions will also help in effective management of natural resources and sustainable built environment planning, while providing better infrastructure services and checking environmental degradation.

Increasing population is considered as the major driver for urban growth (Torrens 2006; Lagarias 2012). According to United Nations, the percentage of urban population in India has increased from 17 % in 1950 to 30.9 % in 2010 (United Nations 2010). By 2050, the percentage of urban population in India is expected to be 51.7 % (United Nations 2010). In order to deal with the pressures from

✉ P. K. Joshi
pkjoshi27@hotmail.com; pkjoshi@mail.jnu.ac.in

1 Department of Natural Resources, TERI University, New Delhi 110 070, India

2 Yale School of Forestry and Environmental Studies, New Haven, CT 06511, USA

3 School of Environmental Sciences, Jawaharlal Nehru University, New Delhi 110 067, India

rapid urbanization, planners and policy makers require information about growth rate, dynamics and patterns of urban areas and extents of urban growth (Sudhira et al. 2004; Taubenböck et al. 2009) for which remote sensing and GIS are widely used (Yeh and Li 2001; Sudhira et al. 2004; Taubenböck et al. 2009). Different approaches have been developed to understand urban dynamics using these spatial tools. These include empirical urban growth estimation (Sudhira et al. 2004; Hu and Lo 2007), fractal analysis (Ma et al. 2008), landscape metrics (Sudhira et al. 2004; Taubenböck et al. 2009), artificial neural network (Tayyebi et al. 2011), cellular automata (He et al. 2008; Vliet et al. 2009; Feng et al. 2011) and agent based models (ABMs) (Loibl and Toetzer 2003; Tian et al. 2011). Many of such approaches are not much successful (Torrens and Benenson 2005; Tian et al. 2011) becauase urban growth is a complex process with highly non-linear interactions among various social, economic and environmental components (Thapa and Murayama 2011). In such circumstances, multi-agent system (MAS) serves as a promising tool in decision making while embracing complexities associated with dynamic processes (Tian et al. 2011; An 2012). Adaptive agents in MAS interact with one another and with environment which causes varied influence on emerging results for the system (Tian et al. 2011). The models developed using MAS capture individual and/or group level interactions and emerging landscape change pattern. Such analysis includes prediction of changes and their processes (Rui and Ban 2010; Tian et al. 2011), assessment of alternative scenarios (Tian et al. 2011) and understanding of the system (Evans and Kelley 2004). Some of the related examples are, urban growth (Loibl and Toetzer 2003; Rui and Ban 2010; Tian et al. 2011), residential dynamics (Benenson 1998; Li and Liu 2007) and land-use and land-cover change (Evans and Kelley 2004; Valbuena et al. 2008). Apart from these capabilities it has tremendous potential to support and assist policy makers while addressing 'what if' scenario analysis (Valbuena et al. 2008; Tian et al. 2011).

While realising the increasing pressures of urbanization in developing countries towards land-use change, this study aims at development of a MAS model to simulate spatial patterns of urban growth in Chandigarh city and its surrounding region. The urban population of Chandigarh has increased by 340.32 % from 1971 to 2011 (see Fig. 1). The massive increase in the urban population among other factors has resulted in urban growth in Chandigarh and also in the surrounding areas. The surrounding region has witnessed encroachments on fertile agricultural land and vegetation cover due to development and expansion of urban areas in the recent past (Singh 2008). The city has a great potential to attract people from all income groups due to high quality lifestyle and high development rate (Sheffer
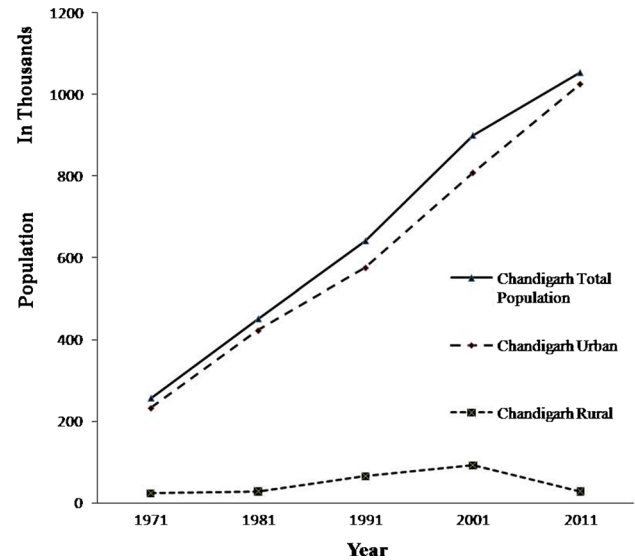

**Fig. 1** Population of Chandigarh from 1971 to 2011

and Levitt 2010). In 1952, the periphery control Act was passed which regulated all developments within 16 km of city limits. However due to pressures from skilled and unskilled labor of low income groups for housing, the government took various schemes for settlement of such people in 1975 (Sheffer and Levitt 2010). The overall mechanism of development of Chandigarh city and its surrounding region is complex and involves multi actor decision making. This study emphasizes on increasing population as a potential driver of urban expansion.

Study area and data

Study area

Chandigarh is located in the foothills of Shivalik range of Himalaya in Northwest India. The geographic location of the city is 30°45′N latitude and 76°47′E longitude coordinates. The city is surrounded by Rupnagar and Patiala district of Punjab, and Ambala district of Haryana. These districts include several satellite towns namely Kharar, Morinda, Kurali, Mohali (S. A. S Nagar), Zirakpur, Panchkula, Derra Bassi and few others. Apart from these populated areas, several villages are in close proximity to the city and the surrounding satellite towns. The majority of land use within a buffer of 30 km from the city cente are agricultural and waste lands with fragments of forest. The surrounding areas have witnessed tremendous increase in built up and reduced agricultural land. The high rates of urbanization have brought massive change in spatial structure of built-up area (Saini and Kaushik 2011). We defined our study area using point location of population area retrieved from GRUMP v1

Fig. 2 Study area (convex hull with 5 km buffer)
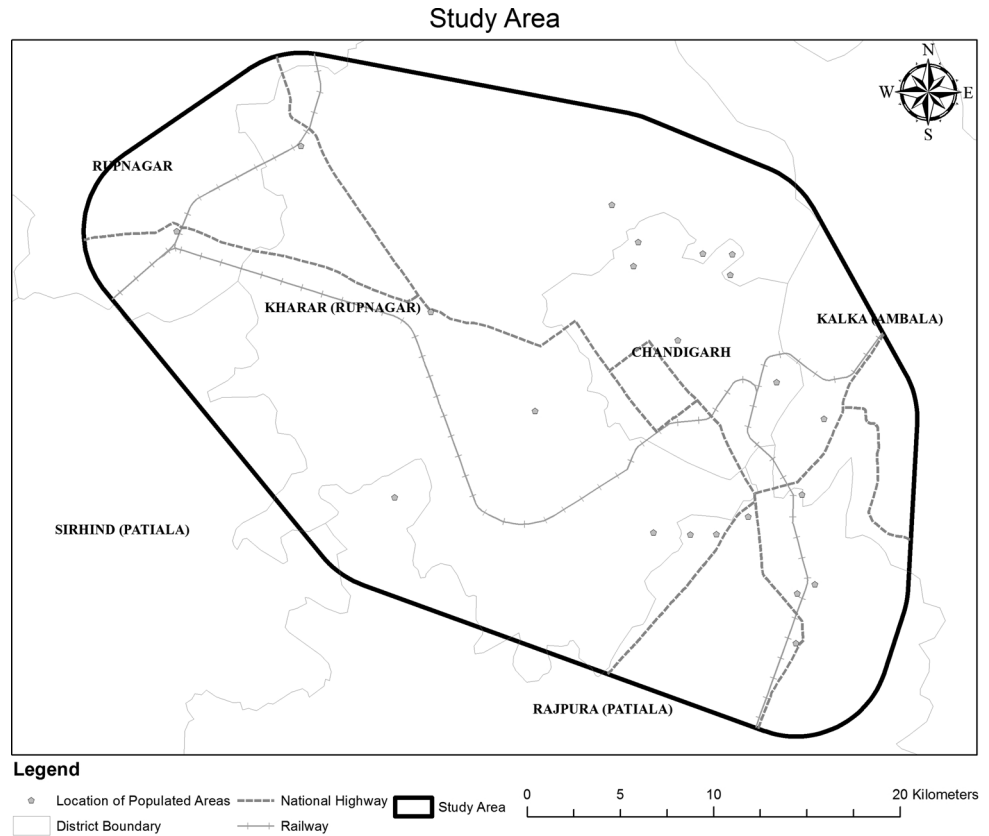
settlement dataset and Google Earth images while considering substantial urban growth that have occurred in and around the city. The settlement points were used to generate a convex hull with a buffer of 5 km. This approach enabled the potential populated areas to be included in the study while excluding the effects of nearby prominent cities like Rajpura, Ambala, Ludhiana and others (Fig. 2). The distribtion of area is 118 km^2 of Chandigarh city, 86.52 km^2 of Ambala and 892.78 km^2 of Rupnagar and Patiala districts. The total area under investigation is 1097.60 km^2.

Satellite and ancillary data

Landsat TM (Thematic Mapper)—5 data (1989, 1999 and 2009) were used to map built-up area for the study area. Survey of India open series maps (1:50000 scale) were used to generate GIS layers which were further utilized in the model. The scanned maps were georeferenced and mosaicked. The mosaicked map was digitized to obtain vector layers of features representing roads, national highways, railway, water bodies, reserved and protected forest boundaries, point location of villages, hospitals, dispensaries, post offices, railway stations, and educational institutions. Table 1 summarizes the data specifications. SRTM digital elevation model (DEM) was downloaded from www.cgiar-csi.org.

Methodology

Built-up area extraction

Support vector machine (SVM) classifier with radial basis function (RBF) kernel was used to extract four land use classes namely vegetation, built-up, barren land and water body for years 1989, 1999 and 2009. Binary images with classes, built-up and non-built-up were produced after masking vegetation, barren and water body. Accuracy assessment was carried out for 1999 and 2009 binary images using Google Earth images. Due to lack of reference for 1989, we used raw Landsat TM 5 image. The error matrix was generated using random stratified samples (n = 50) points, representing each class (Congalton 1991). This error matrix was further used to generate an unbiased population matrix (Pontius and Millones 2011). Quantity disagreement and allocation disagreement were calculated to quantify the errors (Pontius and Millones 2011).

Raster based environment variables

The model environment was specified using variables in the raster format using ESRI ArcGIS v9.3.1 software. These include proximity, density, elevation and slope. The proximity were calculated by finding out the Euclidean

Table 1 Satellite and ancillary data used

Satellite and Sensor	Path/row	Date of acquisition	Cloud cover
Landsat TM 5	147/39	25/10/1989; 27/03/1999; 06/03/2009	0 %
Open series map no.	Scale	Edition	Updated on
H43K14	1:50,000	1st Edition, 2010	2005–2006
H43K5; H43K9; H43K13	1:50,000	1st Edition, 2009	2005–2006
H43K10	1:50,000	1st Edition, 2008	2005–2006

distance from the nearest source feature. SRTM DEM of 90 m resolution was resampled to 30 m (used as elevation) and slope was calculated in degrees. The rasters obtained were divided by its respective maximum value and then subtracted from 1. The density was calculated using simple density function with search radius matching the grid extent. Further the raster was normalized to its respective maximum value. The variables thus obtained include (i) proximity to roads, national highways, railways, water bodies, reserved and protected forests, attractiveness, villages and existing urban and (ii) density of roads, national highways, railways, attractiveness and villages. It is noteworthy to mention that the proximity to attractiveness raster includes information on medical facilities including hospitals and dispensaries, education institutions including schools and universities and railway stations. Figure 3 shows all environment variables.

Logistic regression model

A logistic regression model was developed to associate driving factors with change in built-up between 1989 and 1999 and thus generate urban growth probability surface. The urban growth probability surface was further utilized to initialize ABM. Samples were taken with sample size calculated as described by Peduzzi et al. (1996) (Eq. 1).

$$N = \frac{10k}{p},\tag{1}$$

where, p is the smallest of the proportion of negative or positive cases in the population and k is the number of independent variables.

Backward stepwise logistic regression was carried out in "R" using glmulti package. The automated model development process is based on iteratively minimizing the AIC (Akaike Information Criterion) statistics (Calcagno and Mazancourt 2010) using Eq. 2.

$$AIC = 2k - 2\log(L),\tag{2}$$

where, k is number of parameters in the model and L is maximized value of the maximum likelihood function for an estimated model. A model with high AIC represents a poor fit. The step wise regression iteratively adds or drops the variables thereby finding the best set of variables explaining urban growth process i.e., transitions from non-

built-up to built-up. After the parameters for the model were iteratively determined, the regression beta values were used to generate the probability transition (P_t) surface. The P_t surface was calculated as given in Eq. 3.

$$P_t = \frac{1}{1 + e^{-(b_o + b_1 X_1 + b_2 x_2 + b_3 x_3 + \cdots + b_n x_n)}}\tag{3}$$

where, P_t is the probability of transition from non-built-up to built-up, x_1, x_2, $x_{3\ldots}x_n$ are the driving variables, b_1, b_2, $b_{3\ldots}b_n$ are the corresponding regression coefficients and b_o is constant.

Multi-scale modelling and fractal analysis

Fractal analysis is one means of identification of spatial scale at which a geographical process is operating (Lam and Quattrochi 1992). The fractal analysis was carried out on the P_t rasters calculated at different resolution i.e., 30 to 240 m at 30 m intervals, using stepwise logistic regression while generalizing raster datasets with aggregation using averaging method (Hu and Lo 2007). These were classified using natural groupings, into three classes i.e., low probability, medium probability and high probability. The high probability pixels were given a value 0 and the rest of the pixels as value 1. The binary images were imported into FRACTALYSE software to calculate fractal dimension using linear logarithmic regression. Relative operating characteristic (ROC) statistics for P_t rasters was calculated to check for consistency of result from the fractal analysis.

$$Log[N(e)] = D.\log(e) + c,\tag{4}$$

where N(e) is the number of cells representing high probability, e is the grid distance and D is the fractal dimension.

Urban growth simulation with MAS

We used Netlogo, a programmable multi agent modelling environment for the model implementation (Wilensky 1999). Our spatially explicit model briefly contains three procedures, (1) model initialization and parameter specification, (2) model simulation run, and (3) model termination and model outputs. The model initialization requires specifying the model parameters which includes, urban and rural population growth rates, percentage similarity, neighbourhood type, radius of spatial influence of the
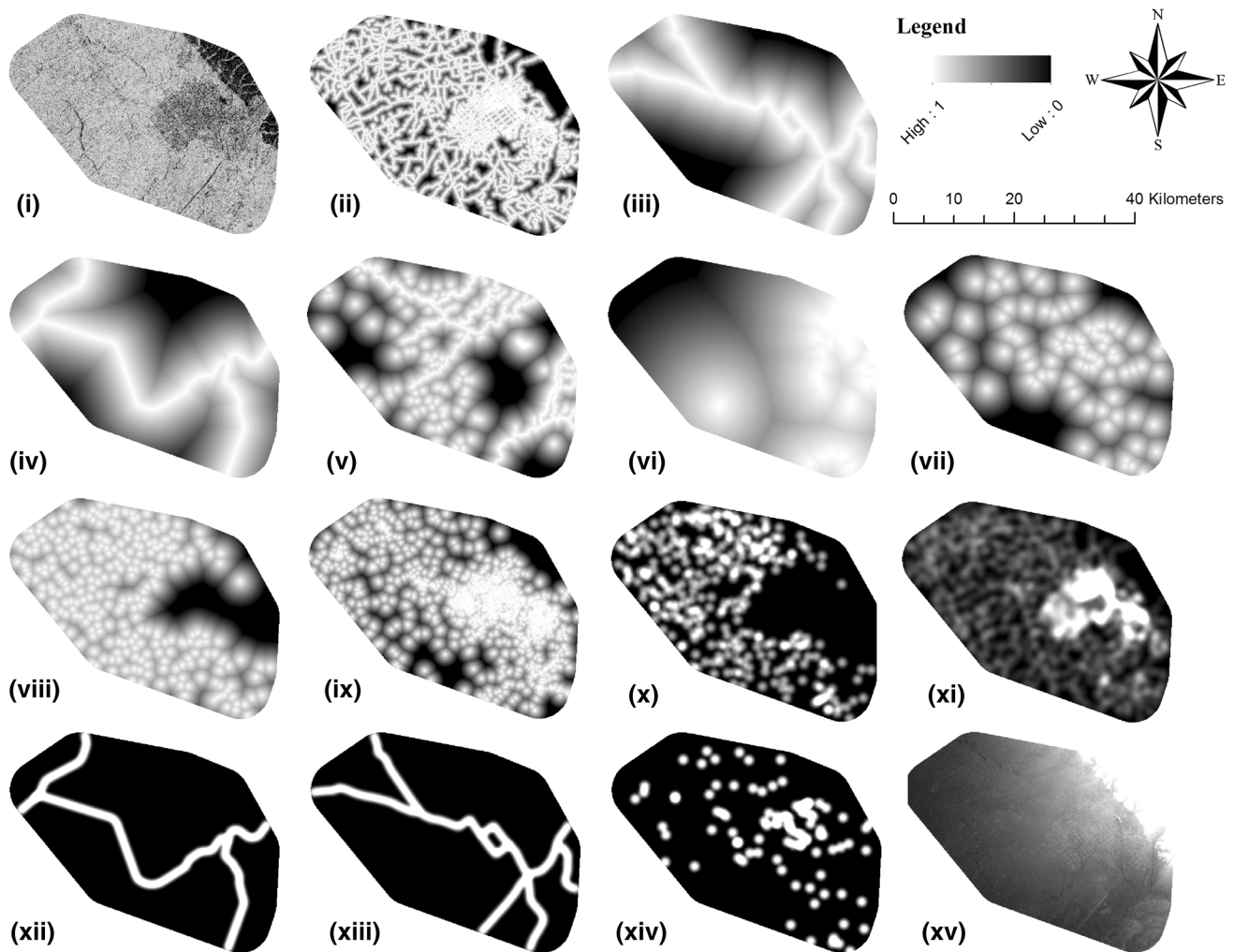
Fig. 3 Raster based environment variables (i) Slope, (ii) Proximity to roads, (iii) Proximity to national highways, (iv) Proximity to railways, (v) Proximity to water bodies, (vi) Proximity to forests, (vii) Proximity to villages, (viii) Proximity to infrastructure services, (ix) Proximity to urban areas, (x) Density of villages, (xi) Density of roads (xii) Density of railways, (xiii) Density of Highways, (xiv) Density of infrastructure services, and (xv) Elevation

agent, urban and rural agent's utility threshold and output year. During the model initialization all the environment variables were also loaded in the model so as to simulate the regional landscape. Netlogo uses a pseudo-random system to ensure that the model is reproducible. The random seed specified in the model ensures generation of random numbers in the model to be deterministic. After initialization, the model is run for specified number of iterations. The model terminates with an output grid showing the final utility value which corresponds to the potential of a cell for urban growth in a scale of zero to one. The final grid is subjected to thresholding with highest percentile values showing urban growth for the prediction period.

Population growth drives the changes in the landscape resulting in increase in built-up areas (Sudhira et al. 2004). We hypothesized that the spatial distribution and interaction of urban and rural population in a regional landscape with mixed urban and rural settings also has influence on the transition of landscape towards built environment. Such interplay of urban and rural population result a specific spatial pattern is difficult to model using empirical modeling techniques. MAS was thus utilized to model spatial distribution and interaction of population with one another and with the environment, resulting specific spatial patterns of urban growth. The main assumption of the model is that rural population residing in areas with limited infrastructure, utilities and other services restricts the potential of the area for urban growth while urban population promotes urban growth.

The model uses population growth as the global driver of urban growth in the study area. Urban and rural population growth rates were calculated using Census of India

data from 1971 to 2011. The weighted average was used to determine the population representing urban and rural population. The weights were determined according to the fraction of geographic area that each administrative region contributes. The rationale to follow this approach was that our study area does not correspond to specific administrative boundary and therefore there was little spatially explicit information of the urban and rural population. The population figures were utilized to calculate the growth rate of the population (Eq. 5).

$$Growth\ Rate\ =\ \left(\frac{P_t}{P_{t-1}}\right)^{\frac{1}{n}} - 1, \tag{5}$$

where n is the difference between the initial and final year and t is the final year.

Also, with the unavailability of sufficient census data to find out the exact population growth function, we assumed that the population growth in our study area follows an exponential function as given in Eq. 6 (Tobler 1970).

$$P_{t-1} = P_t \times e^x, \tag{6}$$

where, x is the population growth rate.

The two groups of the agents in our model: urban and rural agents were programmed to develop corresponding to the population figure reported by the system dynamics modeler. The model utilizes the probability surface obtained from logistic regression to initialize. Our further discussion pertaining to the action of the agents will correspond while taking example of a sample agent from urban population group. The iteration in our synchronous model involves the agent to evaluate the neighboring cells within the radius initially defined, for percentage of similar agents (agents of same group). In case of absence of minimum percentage of similar agents, the agent searches neighboring cells and migrate to another location. Secondly, as agent reaching a particular cell evaluates the utility value of the cell. The utility value of the cell is derived from a heuristic utility function (Eq. 7) defined distinctly for urban and rural population agents (Li and Liu

2007). Initially, the "value" variable in the utility function is derived from the transition probability raster, however during the simulation the agent's action changes "value" variable depending upon the agent and action type.

$$\begin{aligned} Utility\ value = &(W_1 \times PROX_{attractivness}) + (W_2 \times PROX_{forest}) \\ &+ (W_3 \times PROX_{national-raods}) \\ &+ (W_3 \times PROX_{railways}) + (W_4 \times PROX_{road}) \\ &+ (W_5 \times PROX_{urban}) + (W_6 \times PROX_{village}) \\ &+ (W_7 \times PROX_{water}) \\ &+ (W_8 \times PROX_{inverse-slope}) + (W_9 \times value). \end{aligned} \tag{7}$$

Weights (Table 2) for defining the utility function were derived using Saaty's pair wise comparison (Li and Liu 2007; Tian et al. 2011). Consistency ratio was examined to ensure consistency in the pairwise comparison matrix.

The urban agent increases the "value" varible only if the utility value of a cell under consideration is higher than its threshold value else the agenet searches neighbouring cells for higher utility value. The same procedure operates for the rural agents as well. But the rural agents instead of increasing, decreases the "value" variable. The value variable layer remain dynamic during the model run while the other environment variables remains static. The overall stucture of the model is summarized in Fig. 4.

ROC statistic was calculated for the model output considering real change/no-change in built-up areas during the period 1999–2009. Sensivity analysis was carried for each of the model parameters while nullifying the effect of other model parameters by modifying the model code. The model output for each of the simulation run during sensitivity analysis was compared with real change/no-change in built-up areas during 1999–2009. This approach enabled us to explore model dependency on individual parameters.

The resultant transition probability images were validated using the ROC statistics. This approach enabled to explore the model dependency on individual parameters.

Table 2 Distinct weights for urban and rural agent types

Environment variable	Weights (urban agents)	Weights (rural agents)
Proximity to attractiveness	0.0599	0.0654
Proximity to forest	0.0173	0.119
Proximity to national highways	0.0659	0.0318
Proximity to railways	0.034	0.0232
Proximity to roads	0.1626	0.0464
Proximity to existing urban	0.3001	0.0179
Proximity to village	0.0153	0.3623
Proximity to water bodies	0.0162	0.1905
Inverse slope	0.1673	0.1102
Dynamic environment	0.1616	0.0334

Fig. 4 Urban growth
simulation model
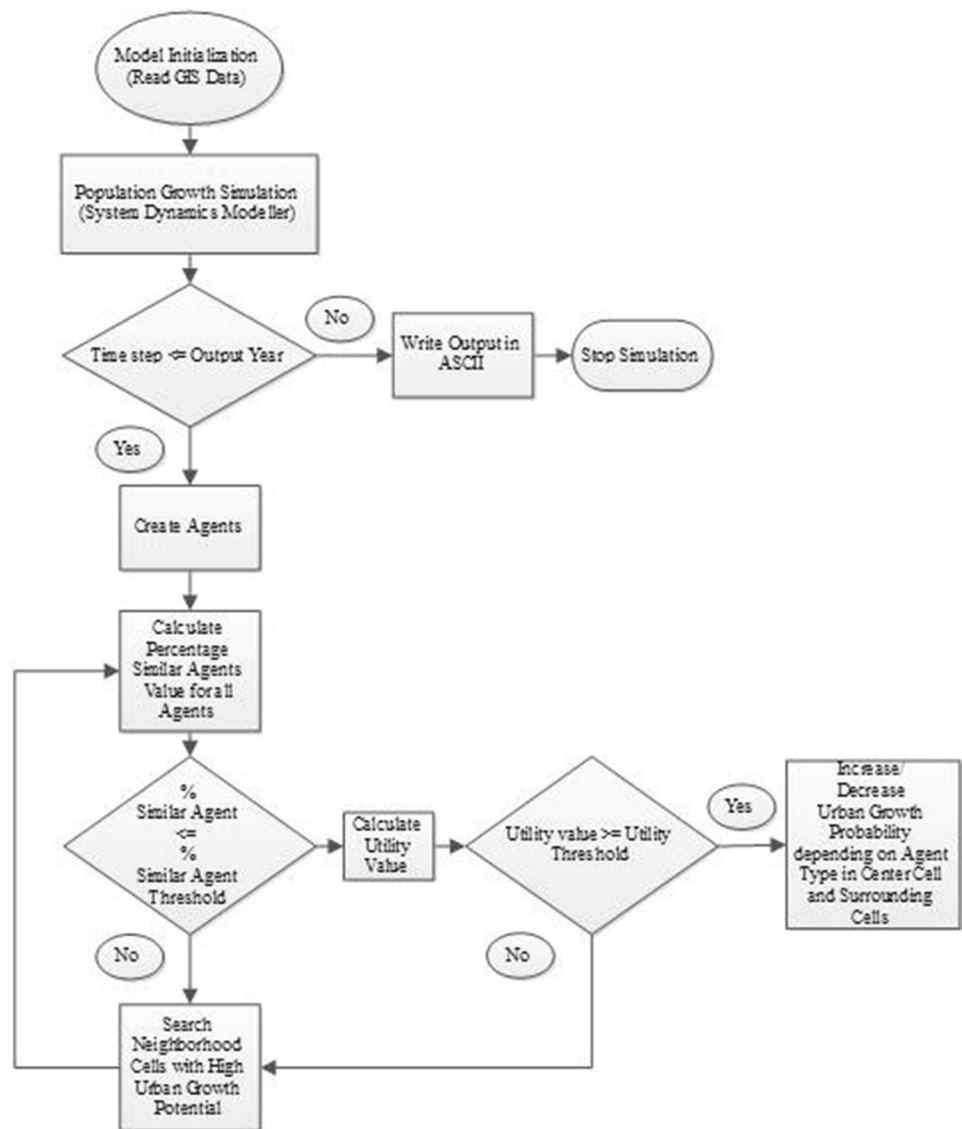
However, in an ABM different model parameters affect each other and thus produce significant variations in the model outcomes. Multiple simulations run were carried to perform behaviours space experiments to produce a calibrated model. All the model parameters i.e., neighborhood type, radius for the neighborhood, percentage similarity in the agents, urban and rural utility threshold, number of initial urban and rural agents, and the value by which the agents will increase or decrease the transition probability, were varied. Table 3 summarizes the induced changes in the model parameters for each run. A reporter function was defined to report cells with high transition probability values (highest percentile) and where real change has occurred during 1999–2009. This function when maximized reflects the optimum model parameter settings.

The calibrated model output in the form of transition probability was compared with real dataset to compute

ROC statistic. Thresholding was further done on the transition probability based on the highest percentile and quantity disagreement, and allocation disagreement was calculated against the real reference map for the year 2009. The model output for the year 2009 and real map of 1989,1999 and 2009 consisting of built up and non-built up classes were used to calculate Shannon's diversity index (SHDI) at landscape level (McGarigal and Marks 1995). SHDI is a relative index given in Eq. (8), used to assess the urban growth in terms of compactness or dispersion at landscape level. The values of SHDI varies from 0 to log n, where n is the total number of patches. A positive change in the value of SHDI represents dispersed growth (Yeh and Li 2001; Sudhira et al. 2004).

$$SHDI = -\sum_{i=1}^{m} P_i \times \ln P_i, \qquad (8)$$

Table 3 Change in model parameters for maximizing the reporter function

Model parameter	Extent of variation in model parameters	Magnitude of change in model parameters for each model run
Neighborhood type	Moore and Von-Neumann	–
Neighborhood radius	1–5	1 cell
Percentage similarity	0–100 %	25 %
Urban utility threshold	0–1	0.25 Unit
Rural utility threshold	0–1	0.25 Unit
Increase in probability	0.025–0.01	0.025 Units
Increase in probability in the surrounding cells	0.025–0.01	0.025 Units
Decrease in probability	0.025–0.01	0.025 Units
Decrease in probability in the surrounding cells	0.025–0.01	0.025 Units

where P_i is the proportion of the landscape occupied by a specific patch type i.

SHDI obtained for model output for the year 2009 was compared with the SHDI obtained from the 2009 binary reference raster. The calibrated model was also used to predict future urban growth for 2019 and, SHDI were calculated to assess the spatial patterns of predicted urban growth.

Results

Built-up and non-built up area extraction

The SVM classifier with RBF kernel produced the best results with γ parameter set to 0.7. The results indicate increase in built-up area from 62.78 km^2 in 1989 to 208.95 km^2 in 2009 (Fig. 5). The predominant change to built-up area of 113.51 km^2 is observed between 1999 and 2009. During 1989–1999 there was a marginal increase in built-up area of 32.55 km^2. Most of the growth is concentrated in the areas adjoining Chandigarh. The satellite cities and towns in close proximity of Chandigarh city namely, Zirakpur, Kharar, Panchkula, Naya Gaon and others showed recent urban growth. The small cities i.e., Kurali, Morinda and Dera Bassi, connected to Chandigarh city with the national highways also show significant increase in the built up area. Overall disagreement (allocation and quantity disagreement) in the built up/non-built up binary map was found to be 9.76, 3.38 and 3.04 % for the year 1989, 1999 and 2009, respectively.

Multi scale modelling and fractal analysis

Fractal analysis of the transition probability raster obtained at different resolutions from logistic regression pointed out 90 m spatial resolution as the optimal for modelling urban growth in the study area (Fig. 6). ROC statistic obtained from the probability surfaces confirmed the operational scale of urban growth process as indicated by fractal analysis. The probability surface generated at 90 m was therefore further used for modelling (Fig. 7). Table 4 shows the regression coefficients and standard errors of the variables at 90 m spatial resolution for urban growth.

Model calibration and validation

The initial sensitivity analysis revealed the dependency of individual model parameters on model outcomes (Fig. 8). The utility thresholds when kept high resulted decreasing ROC statistic. A lower utility threshold also causes immobility to the agents. In case of urban utility threshold a significant decrease in the ROC statistic was observed when the utility threshold increased from 0.8 to 1. It was also observed that with an increase in neighbourhood radius the ROC statistic decreased. The magnitude was observed to be higher in case of Moore neighbourhood as compared to Von-Neumann neighbourhood. In case of percentage similarity, a higher percentage similarity causes aggregation of similar agent types, where as lower percentage similarity causes disaggregated distribution of agent in the spatial domain. ROC is directly related to the percentage similarity and was found to increase with the corresponding increase in percentage similarity values.

It was interesting to observe the individual impacts of model parameters on model outcomes while nullifying the effect of other model parameters, however these relationships could not make significant contribution towards model calibration. The behavioursspace experiments enabled to determine the optimum parameter values for the model. The reporter function was found to be maximized with parameter settings as Von-Neumann neighbourhood with radius: 1, percentage similarity: 65 %, urban utility threshold: 0.75, rural utility threshold: 0.5, increase in transition probability in central cell and cell neighbourhood: 0.01 and 0.005 respectively, decrease in transition probability in central cell and cell neighbourhood: 0.01 and 0.005 respectively (Table 5). The calibrated model with the above stated parameter specifications was used to predict built-up

Fig. 5 Increase in built-up area from 1989 to 2009 in Chandigarh and surrounding areas

Change from 1989 to 2009
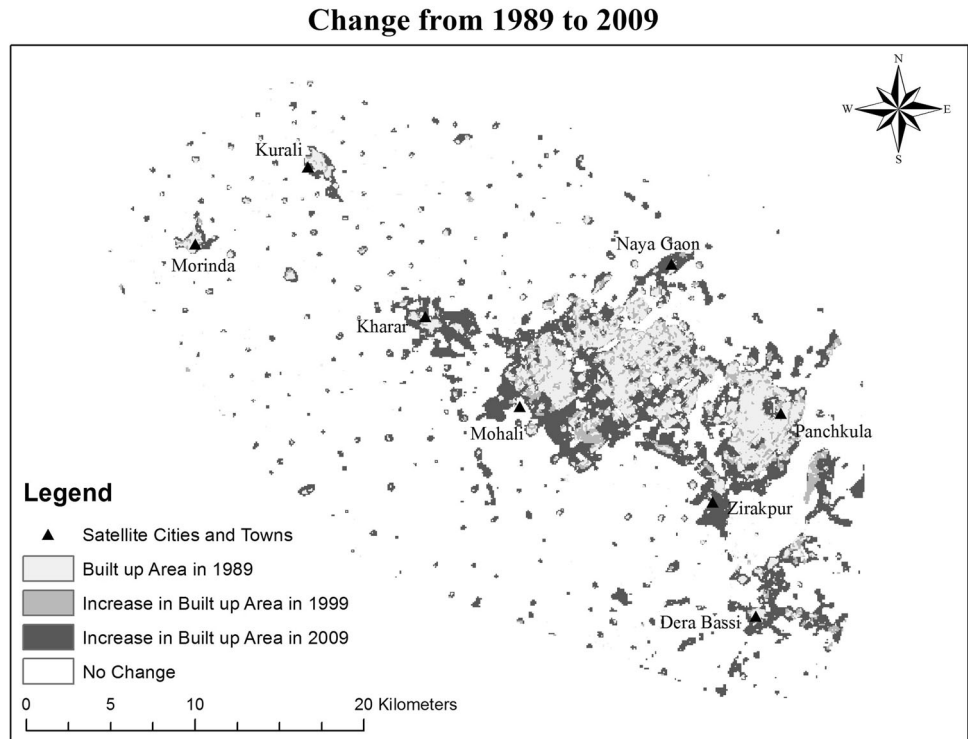
	1989	1999	2009
Built-up area	62.784 km^2	95.335 km^2	208.85 km^2
Non Built-up area	1034.816 km^2	1002.265 km^2	888.75 km^2

Fig. 6 Fractal dimension and relative operating characteristic values for probability surfaces generated using logistic regression plotted against spatial resolution
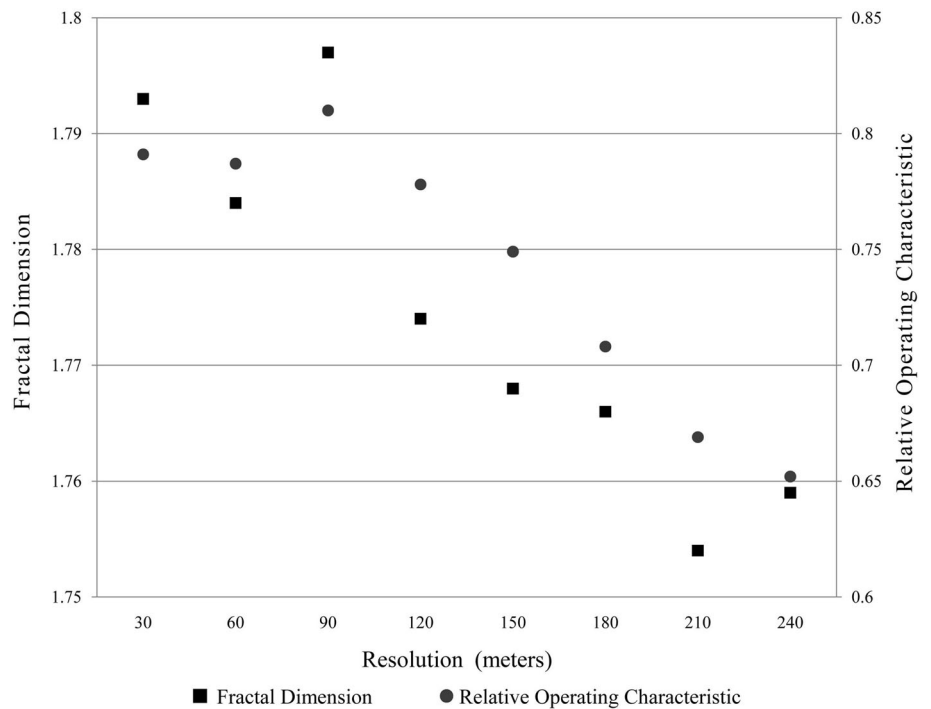
Fig. 7 Transition probability surface (90 m resolution) for non built-up to built-up
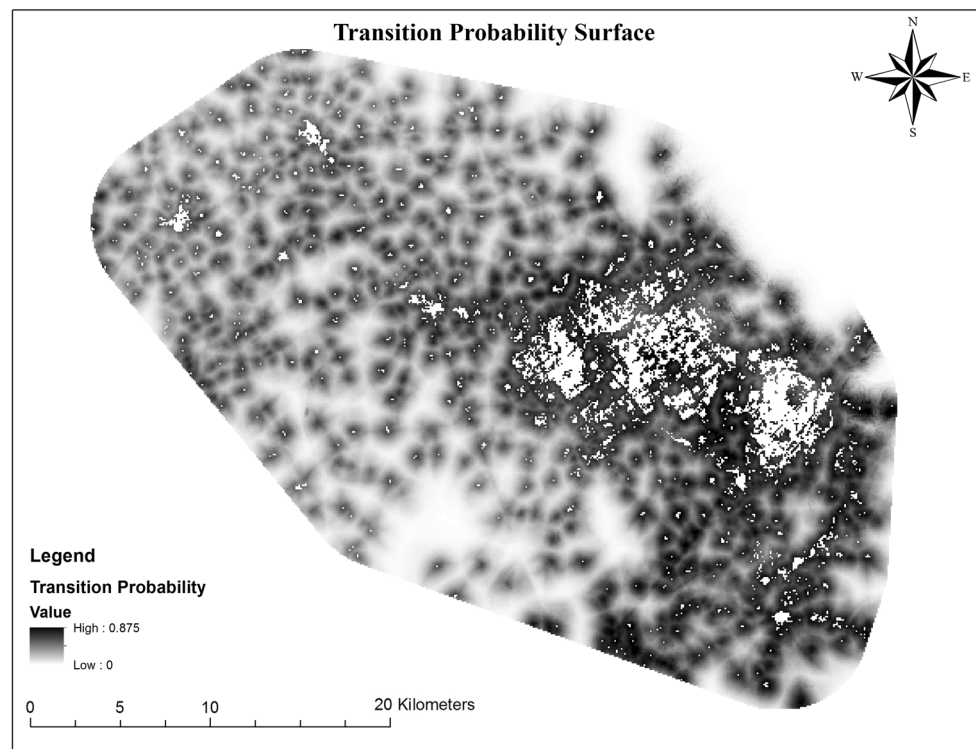
Table 4 The regression coefficients and standard errors of the variables (90 m spatial resolution) for urban growth patterns in Chandigarh and surrounding areas (statistically significant at 98 % confidence interval)

Driver variables	Regression β values	Standard error	2.75 %	97.5 %
Intercept	61.16	3.4311	54.5217	67.9718
Proximity to forest	−1.3297	0.1859	−1.6957	−0.9666
Proximity to national highway	−3.1128	0.4939	−4.0885	−2.1516
Inverse slope	−1.9165	0.8273	−3.5386	−0.2932
Density of national highway	0.5418	0.1857	0.1773	0.9055
Density of roads	0.9654	0.2974	0.3779	1.5446
Density of villages	0.7501	0.2341	0.2915	1.2097
Proximity to rail	1.5723	0.3463	0.8955	2.2534
Proximity to roads	−24.1049	3.0769	−30.1864	−18.1220
Proximity to existing urban	−37.65	2.2995	−42.2147	−33.1998
Proximity to water bodies	1.9413	0.6008	0.7644	3.1204
Cox and Snell R square	**0.39**			
Nagelkerke R square	**0.42**			

Cox and Snell R square is based on log-likelihood and takes the sample size into account. It is identity (an upper bound 1) for normal theory linear regression

Nagelkerke R Square means the power of explanation of the model. Its maximum value can be 1

areas for the year 2009. The ROC statistic for the model output was calculated to be 0.889. The new transition probability surface obtained was subjected to thresholding to obtain only highest percentile values as built-up areas and the rest as non built-up areas. Figure 9 shows the predicted urban growth for the year 2009.

The model output for the year 2009 shows that model is able to capture urban growth from 1999 to 2009 to some

extent. However there are areas where the model under predicted especially in case of urban growth in Chandigarh city peripheries in the south west direction. The allocation disagreement and quantity disagreement for the model output and reference map was calculated as 9.68 and 4.73 %, respectively. The total disagreement between the predicted map and the reference map for the year 2009 is 14.42 %. Figure 10 shows the model prediction for the
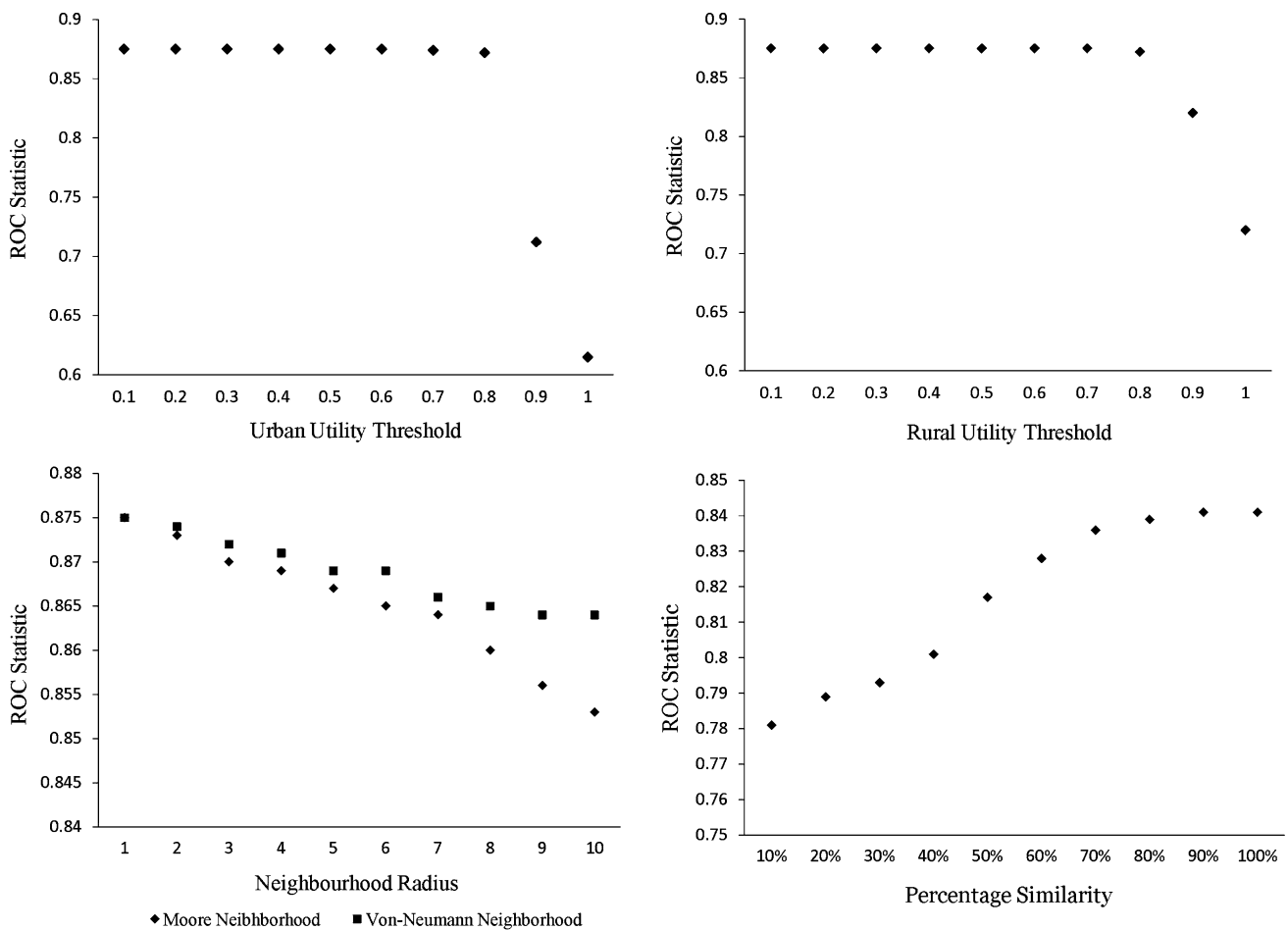
Fig. 8 ROC statistic plotted against major model parameters

Table 5 Cumulative effect of model parameters on model outcomes

Neighbourhood type	Radius	% Similarity	Urban utility threshold	Rural utility threshold	Increase in suitability value		Decrease in suitability value		Cells with high value and real change
					Central cell	Cell neighbourhood	Central cell	Cell neighbourhood	
Moore	1	75	0.75	0.75	0.01	0.005	0.005	0.0025	9385
Moore	1	75	0.75	0.5	0.01	0.005	0.005	0.0025	9928
Moore	1	75	1	0.5	0.01	0.005	0.005	0.0025	8976
Von– Neumann	1	75	0.75	0.75	0.01	0.005	0.005	0.0025	9913
Von–Neumann	1	75	0.75	0.5	0.01	0.005	0.005	0.0025	10,215
Von-Neumann	1	75	1	0.5	0.01	0.005	0.005	0.0025	9091
Von-Neumann	1	75	1	0.75	0.01	0.005	0.005	0.0025	9873
Von-Neumann	2	75	0.75	0.75	0.01	0.005	0.005	0.0025	9899

year 2019. The model prediction results show the emergence of new random built up patches along the roads, national highways and existing built-up area.

SHDI was calculated as 0.2202, 0.2951 and 0.4873 for the years 1989, 1999 and 2009, respectively. The distribution of built-up area was more compact in 1989 than in 2009, which showed more dispersed built-up area. SHDI calculated for the model output for the year 2009 was 0.5028. A very little difference in the value of SHDI (0.0155) for the model output and the reference data for the year 2009, revealed that the model captured similar characteristics of built up area as observed from the empirical data.

Discussion and conclusion

Urban growth is a highly complex, non-linear and heterogeneous process which involves multi actor decision making. Among all, human behaviours and decision making forms the basis of urban growth. Several different models exist which helps to understand the underlying

Fig. 9 Model prediction for the year 2009

Fig. 10 Model prediction for the year 2019

processes of urban growth. These models are largely empirical in nature and most of the time fails to incorporate population behaviours and decision making. Specifically, while modelling urbanization dynamics beyond the city boundaries it is important to understand the role of urban and rural population taken together in shaping the regional landscape. ABMs provide opportunities to understand a system which cannot be simply described using mathematical formula (Crooks 2012). Urban growth is determined by the spatial and temporal interactions of several causal factors. At a regional scale, the function of capital and population are the key factors that determine regional urban area demand (He et al. 2008). This study demonstrates complex interactions of urban and rural population. These complex interactions influence and shape the urban morphology (Rui and Ban 2010). In our study area, the fast rate of urban population growth among several other factors, have triggered urban growth in the surrounding satellite towns as well as in the Chandigarh city. Our simulation model takes into account the interactions of two groups of population i.e., rural and urban population and its influences on shaping the urban morphology. The two groups of population show spatial segregation in the regional landscape. Also, the dynamics of the two groups have a vast influence on urban growth. Our simulation model thus incorporates two fundament concepts of spatial segregation and population dynamics to simulate urban growth. The two types of agent groups interact with each other and with the environment dynamically in the whole
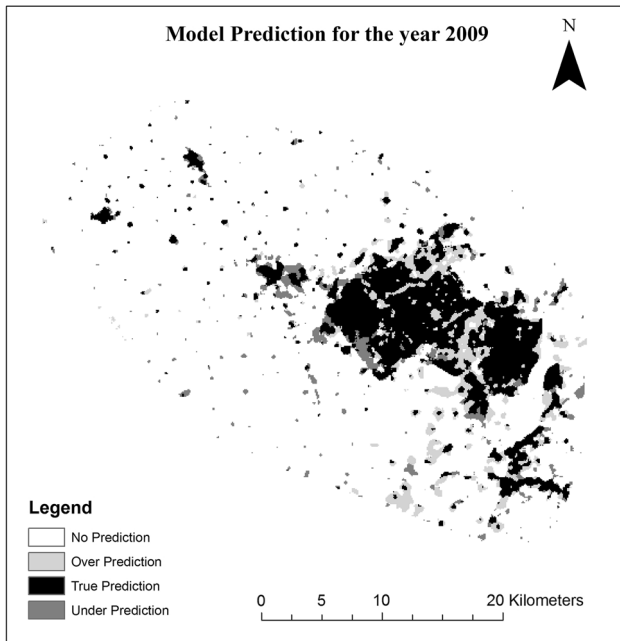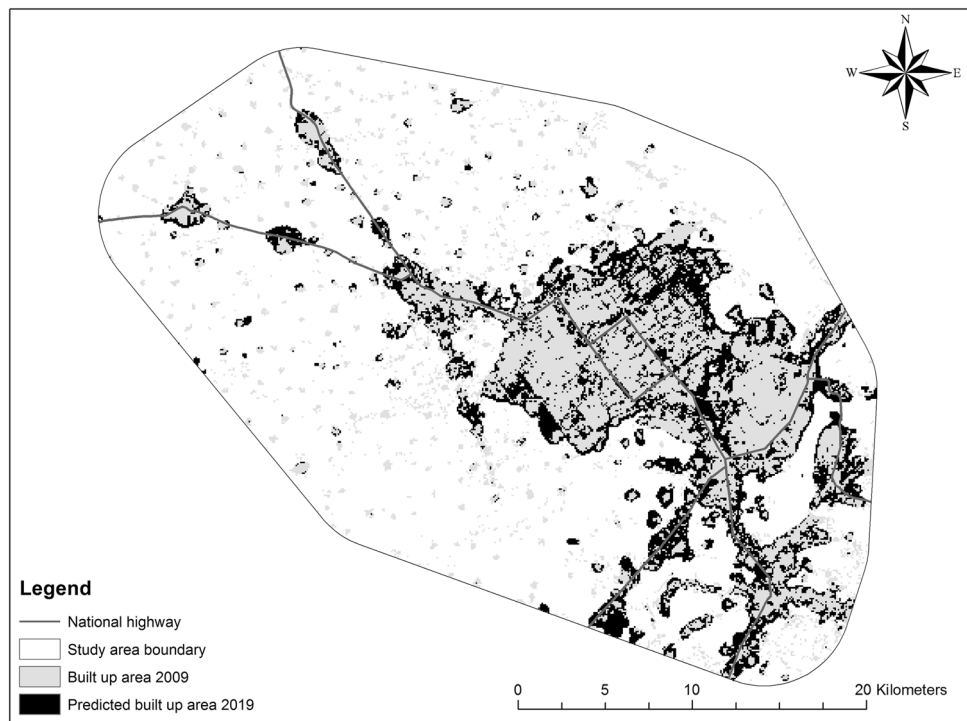
simulation process. These agents correspond to artificial life geospatial agents (Sengupta and Sieber 2007). During the simulation the dynamic environment also affects the behaviour of the agents. Our simulation result shows that the aggregated population level interactions can be manifested in simulating urban growth phenomenon in the regional landscape. The distinct preferences of urban and rural population, enabled to define the tendencies of these groups to reside in spatially distinct environments. Essentially, the main motive of giving different weightages to different variables was to simulate the aggregated level preferences of the population. Also, the dynamic variable defined in the utility function enabled the agents to react to the changing environment as they attempt to maximize their individual utility functions.

Logistic regression is used in several studies for prediction of probability of occurrence of an object, event, or phenomenon (Li et al. 1997; Hu and Lo 2007; Ozdemir 2011; Singh and Kushwaha 2011). If the nature of change of land use and land cover is binary, hence the probability of change from non built-up to built-up follows a logistic curve, described by the logistic function (Hu and Lo 2007). The changes in fractal dimension values across multiple spatial scales can be interpreted positively and the scale at which the highest fractal dimension is observed should be the one at which the most of the process operates (Goodchild and Mark 1987; Lam and Quattrochi 1992; Hu and Lo 2007). Since we calculated fractal dimensions using grid algorithm on binary images, it was imperative to find out if the model actually performs better at the resolution observed as turning point during fractal analysis. MAS proved to be a suitable technique to explore these interactions and address the spatial patterns and dynamics of urban areas within a conceptual modelling framework. The model can be further extended to include additional predominant factors for urban growth such as migration of population, employment opportunities in the region and others, to explore the interactions of rural and urban population and produce a realistic simulation of urban growth in a given region. Currently, the model does not incorporate the planning process and multi actor decision making, responsible for urban growth. A much greater effort is required in order to unravel the complete potential of MAS for modelling urban dynamics and provide support to policy makers.

References

An L (2012) Modeling human decisions in coupled human and natural systems: review of agent-based models. Ecol Model 229:25–36

Benenson I (1998) Multi-agent simulation of residential dynamics in the city. Comput Environ Urban Syst 22(1):25–42

Calcagno V, Mazancourt CD (2010) glmulti: an R package for easy automated model selection with (Generalized) linear models. J Stat Softw 34(12):1–29

Congalton RG (1991) A review of assessing the accuracy of classifications of remotely sensed data. Remote Sens Environ 37:35–46

Crooks AT (2012) The use of agent-based modelling for studying the social and physical environment of cities. In: De Roo G, Hiller J, Van Wezemael J (eds) Complexity and planning: systems. Assemblages and Simulations, Ashgate, Burlington, pp 385–408

Evans TP, Kelley H (2004) Multi-scale analysis of a household level agent-based model of landcover change. J Environ Manag 72:57–72

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

Goodchild MF, Mark DM (1987) The fractal nature of geographical phenomena. Ann Assoc Am Geogr 77(2):265–278

Green DM, Baker MG (2003) Urbanization impacts on habitat and bird communities in a Sonoran desert ecosystem. Landsc Urban Plan 63(5):225–239

He C, Okada N, Zhang Q, Shi P, Li J (2008) Modelling dynamic urban expansion processes incorporating a potential model with cellular automata. Landsc Urban Plan 86:79–91

Hu Z, Lo CP (2007) Modeling urban growth in Atlanta using logistic regression. Comput Environ Urban Syst 31:667–688

Lagarias A (2012) Urban sprawl simulation linking macro-scale processes to micro-dynamics through cellular automata, an application in Thessaloniki, Greece. Appl Geogr 34:146–160

Lam NSN, Quattrochi DA (1992) On the issues of scale, resolution, and fractal analysis in the mapping sciences. Prof Geogr 44(1):88–98

Li X, Liu X (2007) Defining agents' behaviors to simulate complex residential development using multicriteria evaluation. J Environ Manag 85:1063–1075

Li W, Wang Z, Ma Z, Tang H (1997) A regression model for the spatial distribution of red-crown crane in Yancheng Biosphere Reserve China. Ecol Model 103(2–3):115–121

Loibl W, Toetzer T (2003) Modeling growth and densification processes in suburban regions-simulation of landscape transition with spatial agents. Environ Model Softw 18(6):553–563

Ma R, Gu C, Pu Y, Ma X (2008) Mining the urban sprawl pattern: a case study on Sunan, China. Sensors 8:6371–6395

McGarigal K, Marks B. J. (1995) FRAGSTATS: spatial pattern analysis program for quantifying landscape structure. USDA Forest Service General Technical Report PNW-351

Moore M, Gould P, Keary BS (2003) Global urbanization and impact on health. Int J Hyg Environ Health 206(4–5):269–278

Ozdemir A (2011) Using a binary logistic regression method and GIS for evaluating and mapping the groundwater spring potential in the Sultan Mountains (Aksehir, Turkey). J Hydrol 405(1–2):123–136

Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR (1996) A simulation study of the number of events per variable in logistic regression analysis. J Clin Epidemiol 49:1373–1379

Pontius RG, Millones M (2011) Death to Kappa: birth of quantity disagreement and allocation disagreement for accuracy assessment. Int J Remote Sens 32(15):4407–4429

Rui Y, Ban Y. (2010) Multi-agent simulation for modeling urban sprawl in the greater toronto area. 13th AGILE International Conference on Geographic Information Science 2010. Guimarães

Saini SS, Kaushik SP (2011) Land use changes in haryana sub-region of chandigarh periphery controlled area: a spatio-temporal study. Inst Town Plan India J 8(3):96–106

Sengupta R, Sieber R (2007) Geospatial agents, agents everywhere. Trans GIS 11(4):483–506

Seto K, Kaufmann R, Woodcock C (2000) Landsat reveals China's farmland reserves, but they're vanishing fast. Nature 406:121

Sheffer D. A, Levitt R. E. (2010) The diffusion of energy saving technologies in the building industry: structural barriers and possible solutions. Collaboratory for Research on Global Projects. http://crgp.stanford.edu/publications/working_papers/Sheffer_Levitt_Diffusion_of_Energy_Saving_WP0057.pdf

Singh Y (2008) Land use and land cover related hydrological changes in different ecosystems of Inter State Chandigarh Region, NW India. Trans Inst Indian Geogr 30(1):69–84

Singh A, Kushwaha SPS (2011) Refining logistic regression models for wildlife habitat suitability modeling—a case study with muntjak and goral in the Central Himalayas, India. Ecol Model 222(8):1354–1366

Su S, Xiao R, Jiang Z, Zhang Y (2012) Characterizing landscape pattern and ecosystem service value changes for urbanization impacts at an eco-regional scale. Appli Geogr 34:295–305

Sudhira HS, Ramachandra TV, Jagadish KS (2004) Urban sprawl: metrics, dynamics and modelling using GIS. Int J Appl Earth Obs Geoinf 5:29–39

Suriya S, Mudgal BV (2012) Impact of urbanization on flooding: the Thirusoolam sub watershed—a case study. J Hydrol 412–413:210–219

Taubenböck H, Wegmann M, Roth A, Mehl H, Dech S (2009) Urbanization in India: spatiotemporal analysis using remote sensing data. Comput Environ Urban Syst 33:179–188

Tayyebi A, Pijanowski BC, Tayyebi AH (2011) An urban growth boundary model using neural networks, GIS and radial parameterization: an application to Tehran, Iran. Landsc Urban Plan 100:35–44

Thapa RB, Murayama Y (2011) Urban growth modeling of Kathmandu metropolitan region, Nepal. Comput Environ Urban Syst 35:25–34

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. Ecol Model 222:1129–1138

Tobler WR (1970) A computer movie simulating urban growth in the Detroit region. Econ Geogr 46:234–240

Torrens PM (2006) Simulating sprawl. Ann Assoc Am Geogr 96(2):248–275

Torrens PM, Benenson I (2005) Geographic automata systems. Int J Geogr Inf Sci 19:385–412

United Nations (2010) World urbanization prospects, 2009 revision. Population Division, Department of Economic and Social Affairs, United Nations, New York

Valbuena D, Verburg PH, Bregt AK (2008) A method to define a typology for agent-based analysis in regional land-use research. Agric Ecosyst Environ 128:27–36

Vij D (2012) Urbanization and solid waste management in India: present practices and future challenges. Procedia-Soc Behav Sci 37:437–447

Vliet JV, White R, Dragicevic S (2009) Modeling urban growth using a variable grid cellular automaton. Comput Environ Urban Syst 33(1):35–43

Wilensky U (1999) NetLogo. http://ccl.northwestern.edu/netlogo. Center for Connected Learning and Computer-Based Modeling. Northwestern University, Evanston

Xu C, Liu M, An S, Chen JM, Yan P (2007) Assessing the impact of urbanization on regional net primary productivity in Jiangyin County, China. J Environ Manag 85(3):597–606

Yeh AGO, Li X (2001) Measurement and monitoring of urban sprawl in a rapidly growing region using entropy. Photogramm Eng Remote Sens 67(1):83–90