Stoch Environ Res Risk Assess
DOI 10.1007/s00477-015-1128-z

ORIGINAL PAPER

Modeling urban growth with GIS based cellular automata and least squares SVM rules: a case study in Qingpu–Songjiang area of Shanghai, China

Yongjiu Feng[1,2] · Yan Liu[3] · Michael Batty[4]

Abstract A critical issue in urban cellular automata (CA) modeling concerns the identification of transition rules that generate realistic urban land use patterns. Recent studies have demonstrated that linear methods cannot sufficiently delineate the extraordinary complex boundaries between urban and non-urban areas and as most urban CA models simulate transitions across these boundaries, there is an urgent need for good methods to facilitate such delineations. This paper presents a machine learning CA model (termed MachCA) with nonlinear transition rules based on least squares support vector machines (LS-SVM) to simulate such urban growth. By projecting the input dataset into a high dimensional space using the LS-SVM method, an optimal hyper-plane is constructed to separate the complex boundaries between urban and nonurban land, thus enabling the retrieval of nonlinear CA transition rules. In the MachCA model, the transition rules are yes–no decisions on whether a cell changes its state or not, the rules being dynamically updated for each iteration of the model implementation. The application of the MachCA for simulating urban growth in the Shanghai Qingpu–Songjiang area in China reveals that the spatial configurations of rural–urban patterns can be modeled. A comparison of the MachCA model with a conventional CA model fitted by logarithmic regression (termed LogCA) shows that the MachCA model produces more hits and less misses and false alarms due to its capability for capturing the spatial complexity of urban dynamics. This results in improved simulation accuracies, although with only less than 1 % deviation between the overall errors produced by the MachCA and LogCA models. Nevertheless, the way MachCA model use in retrieving the transition rules provides a new method for simulating the dynamic process of urban growth.

Keywords Cellular automata (CA) · Nonlinear transition rules · Least squares support vector machines (LS-SVM) · Urban growth · The Shanghai Qingpu–Songjiang area

1 Introduction

As a result of the rapid population explosion, urban growth has increasingly become a major challenge in many suburban areas (Ji et al. 2006), particularly in developing countries such as China (Ding et al. 2013; Li et al. 2014; Zhang et al. 2015). There is a pressing need to understand how suburban areas expand across space and over time under the conditions of rapid population growth and socio-economic development which is dominated by physical constraints. It is also important to develop state-of-the-art methods and tools to detect the spatio-temporal patterns and changes in urban and regional development (Ye and Dang 2013). Cellular automata (CA) based urban models are capable of catching the dynamics of urban growth and land use change, in comparison with more conventional

✉ Yan Liu
yan.liu@uq.edu.au

1 College of Marine Sciences, Shanghai Ocean University, 999 Huchenghuan Road, Shanghai 201306, China

2 The Key Laboratory of Sustainable Exploitation of Oceanic Fisheries Resources (Shanghai Ocean University), Ministry of Education, Shanghai 201306, China

3 School of Geography, Planning and Environmental Management, The University of Queensland, Brisbane, QLD 4072, Australia

4 Centre for Advanced Spatial Analysis, University College London, 90 Tottenham Court Road, London W1N 6TR, UK

Published online: 23 July 2015

land use transportation models which tend to simulate urban structure at only one point in time (Batty et al. 1999; Clarke and Gaydos 1998; Couclelis 1997; Guan et al. 2011; Liu 2008; Verburg et al. 2002; Wu 1998).

Substantial progress has been made on CA models for understanding geographical systems and urban dynamics over the last two decades (Feng and Liu 2013a, b; Lau and Kam 2005; Liao et al. 2014; Liu et al. 2008; Liu and Phinn 2003; Wu and Webster 1998; Wu et al. 2010). Specifically, CA models have been applied to simulate urban growth and land use change (Al-shalabi et al. 2013; Cheng and Masser 2003; Deng et al. 2014; Mitsova et al. 2011; Nourqolipour et al. 2014), project future scenarios under different planning conditions (Barredo et al. 2003; He et al. 2011; Maeda et al. 2011; Yin et al. 2011) and investigate ecological security patterns (Gong et al. 2009; Mao et al. 2013). Existing studies have demonstrated that the global spatial pattern of urban growth is the result of local circumstances (Lau and Kam 2005). As more and more remotely sensed imagery with high spatial resolution has become readily available, spatially related driving factors affecting urban growth can be extracted from these images using professional data mining software, and thence used to construct CA models. However, assessing the extent these driving factors impact on urban growth from which the probability a land parcel (or cell) changing state from non-urban to urban still remains a challenging issue for most CA model-building.

Many spatial statistical methods have been reported for quantifying the impact of driving factors, typically considered as CA parameters. Such methods usually include the analytical hierarchy process devised as a tool for eliciting weights associated with driving factors (Wu 1998), multi-criterion evaluation of the importance of these same factors (Wu and Webster 1998) and logistic regression which determines the weights of such factors (Arsanjani et al. 2013; Liu and Feng 2012; Wu 2002). Although explicit CA parameters can be obtained by using spatial statistical methods, the multi-collinearity effects of the driving factors cannot be effectively removed, and hence the nonlinear complexity of the fundamental interactions of urban dynamics cannot be sufficiently reflected (Feng and Liu 2013a).

With the rapid development of evolutionary computation, a number of optimization methods have been used to construct CA models such as genetic algorithms (Cao et al. 2014; Liu and Feng 2012; Liu et al. 2008, 2014), particle swarm optimization (Feng et al. 2011) and simulated annealing (Feng and Liu 2013b). Optimized CA transition rules have been discovered from spatial databases by minimizing the differences between the simulated urban patterns generated from a conventional CA model (which typically employ statistical methods such as logistic regression or the analytic hierarchy process) and the actual patterns classified from remote sensing images. CA models incorporating evolutionary computation algorithms are able to produce simulation results with higher accuracy than those without optimization (Feng et al. 2011; Feng and Liu 2013b; Liao et al. 2014; Liu et al. 2014). In addition, fuzzy-logic-control and approaches based on rough sets have been reported as defining CA transition rules in the form of "IF…THEN…" or "WHAT…IF…" statements (Liu 2008, 2012; Liu and Phinn 2003; Wang et al. 2011). Moreover, Markov chain-based CA models have also been developed to model land use change and urban growth processes (Arsanjani et al. 2013; Guan et al. 2011; Kamusoko et al. 2009; Shafizadeh Moghadam and Helbich 2013).

For most applications to date, the essential nature of CA modeling is to identify the complex nonlinear boundaries between urban and non-urban rural areas and how these boundaries evolve over time (Feng and Liu 2013a; Huang et al. 2009; Liu et al. 2008; Yang et al. 2008). This issue can be transformed to a binary classification problem. Consequently, kernel methods have been used to retrieve CA transition rules by mapping the original data into a high dimensional feature space (Feng and Liu 2013a; Huang et al. 2009; Liu et al. 2008; Yang et al. 2008). For instance, Feng and Liu (2013a) used kernel principal component analysis while Liu et al. (2008) used kernel Fisher discriminant analysis to retrieve the nonlinear CA transition rules, both of which were inspired by Schölkopf et al. (1998) method. Another type of kernel method-support vector machines (SVMs) initially proposed by Vapnik (1998, 2000)—has also been applied in for land use change modeling (Huang et al. 2009; Yang et al. 2008). Huang et al. (2009) have developed a modeling framework based on unbalanced SVMs to model and analyze urban land-use change in relation to various factors such as population, distance to roads and facilities, and surrounding land uses. Although not strictly a CA modeling approach, this framework does report a plausible performance for land-use-change modeling. Yang et al. (2008) has presented a standard SVMs method to establish the nonlinear CA transition rules for simulating urban development in the Shenzhen City, China. Their result has demonstrated that the standard SVMs based CA model can overcome some limitations (e.g. harmful effects of inter-correlations between different driving factors) of existing CA models and thus obtain higher accuracy in simulating complex urban systems. Most of the above-mentioned CA models appear to use probability to define the transition rules which are static in nature. These static transition rules are limited in their reflection of the dynamic processes of urban growth. On the other hand, a threshold must be selected for the probability-based transition rules to determine whether or not a non-urban cell changes its state However, the determination of the threshold is largely based on trial-and-error which also affect simulation results. Therefore, there is a pressing need to explore a CA model which can update its transition rules dynamically and does not require an arbitrary definition of such a threshold.

This paper presents a nonlinear CA model based on a typical machine learning method-least squares support vector machines (LS-SVM)—to simulate the process of urban growth. As a modified version of support vector machines (SVMs), LS-SVM is able to generate a direct solution by solving a set of linear equations instead of representing the optimization problem as one of quadratic programming (Suykens et al. 2002; Suykens and Vandewalle 1999; Ye and Xiong 2007).

Using this approach, the paper aims to discover the nonlinear CA transition rules by constructing a LS-SVM based CA model (named MachCA, meaning machine learning based CA) to simulate the spatio-temporal processes of urban growth. Previously, Yang et al. (2008) used SVMs to retrieve the land conversion probability and by using LS-SVM, this research will advance this line of research to directly determine the land use transition rules. Most importantly, the transition rules are dynamically updated for each iteration of the model through the LS-SVM which also result in a crisp yes–no decision on whether a land cell changes its state or not. The great advantage of the method is thus that it does not need any arbitrary definition of a transition probability threshold. Compared with the CA models with probability based transition rules, the proposed CA model offers a much more explicit way of simulating and understanding the dynamics of urban growth.

The rest of this paper is organized as follows. Following the introduction, Sect. 2 presents the location of the study area and the data sources used. Section 3 presents the LS-SVM method which is able to discover the nonlinear CA transition rules and then construct the MachCA model. In Sect. 4, the implementation of the proposed MachCA model in the study area is presented, and its performance is evaluated by comparing it with a standard LogCA model. Finally, the conclusions are presented in Sect. 5.

2 The study area and its data

A fast growing suburban region, the Qingpu–Songjiang area of western Shanghai, China as shown in Fig. 1, was chosen as the study area which we used to examine the validity of the MachCA model. The Qingpu–Songjiang area includes two administrative districts, Qingpu District and Songjiang District, which have a total area of 1275 km^2. Over the last two decades, the study area has witnessed rapid development with dramatic land use and land cover change, considerable expansion of its urban extent, a population explosion, and a massive increase in its productive economy. For instance, the population in the study area grew from 150,000 in 1985 to 1.83 million by the end of 2008 with population density increasing to 1435 per square kilometer land (Shanghai Municipal Statistics Bureau 1985, 2008). With this high rate of population growth, an explosive growth of the urban extent has been observed in this area. Therefore, Qingpu–Songjiang is a representative study area in which to explore the rapid development of suburban regions in China (Hu and Lo 2007; Ji et al. 2006; Yue et al. 2012), and tracking of urban development in such rapidly developing metropolitan cities are important for our generic understanding of urban dynamics (Feng and Liu 2013a, b; Yue et al. 2014).

In this study, two Landsat images covering the study area acquired on 18 July 1992 and 24 March 2008 were used. The original spatial resolution of the images is 30 m. In addition, a digital topographic map at a scale of 1:50,000 was obtained as the reference data for georectification. A total of 30 ground control points were chosen from both the digital topographic map and two Landsat images. The images were georectified using the polynomial method ENVI 4.5, and the root mean squared errors for the geometric correction were all less than 0.52 pixel. Based on the georectified images in 1992 and 2008, a land change map shown in Fig. 2 was produced using the default spectral angle mapper classifier of ENVI 4.5 and overlay analysis in ArcGIS 10.1. The excluded land use type shown in Fig. 2 contains water bodies and wetlands that are resistant to urban growth. The non-urban land use type includes public green spaces, parks, agricultural lands (including basic protective farmland, or BPF), and aquaculture farms.

The 1992 classified land use map has served as the initial and reference state for modeling urban growth with the MachCA model. Figure 2 shows that apart from the water bodies, the urban area only constituted 6.5 % of the land in 1992 while the rest were non-urban. However, by 2008, an additional 42.9 % of the candidate region had been converted into urban land. This is an exemplar of a typical very fast growing urban area in eastern and coastal China, characterized by very complex nonlinear boundaries between urban and non-urban rural areas. The Qingpu–Songjiang area is thus extremely suitable for testing the MachCA model, which aims at separating the complex boundaries between urban and non-urban land.

3 Methodology

3.1 The MachCA model flowchart

The MachCA model consists of three modules: the LS-SVM Model Training, the Land Use Change Decision

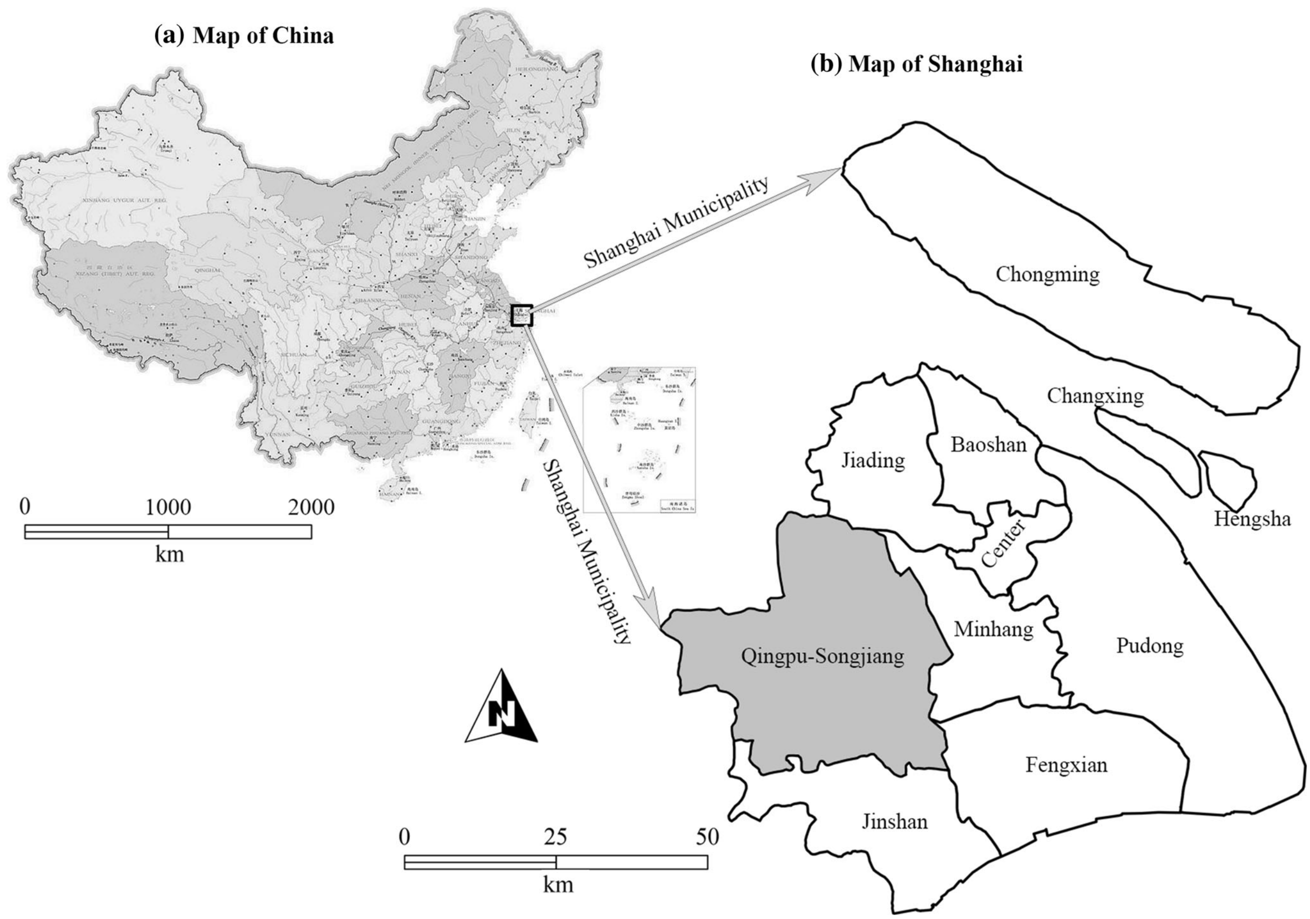


Fig. 1 The Qingpu–Songjiang area of Shanghai, China

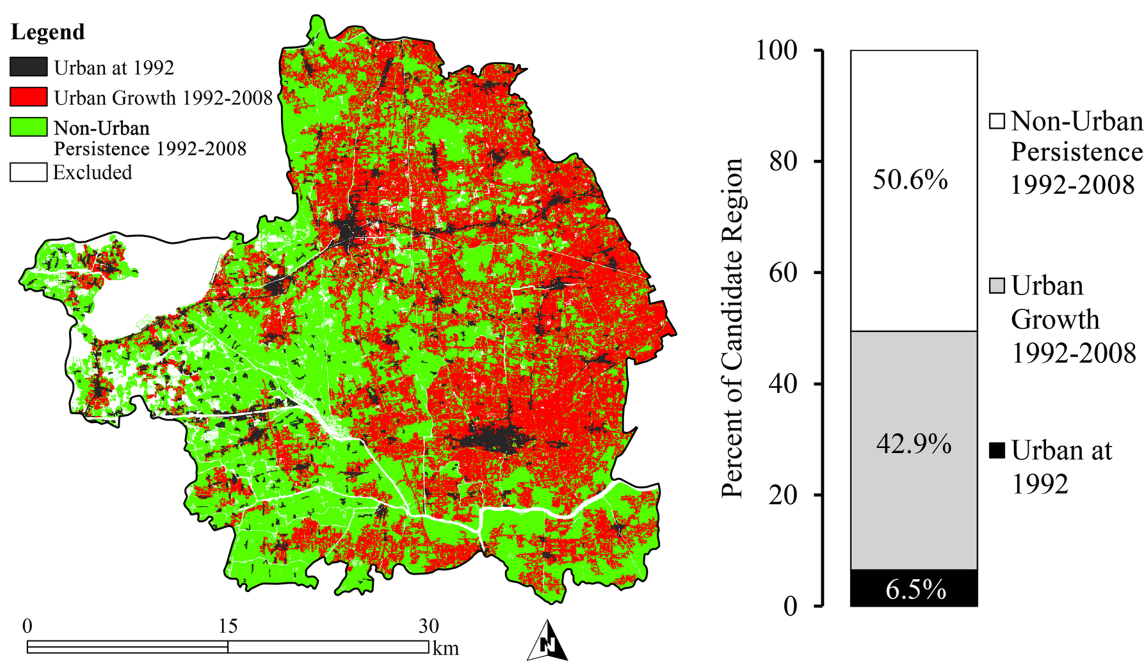


Fig. 2 Observed urban growth from 1992 to 2008 in Qingpu–Songjiang area, Shanghai. The excluded land use type contains water bodies and wetlands, a persistent land use category in urban growth in this study

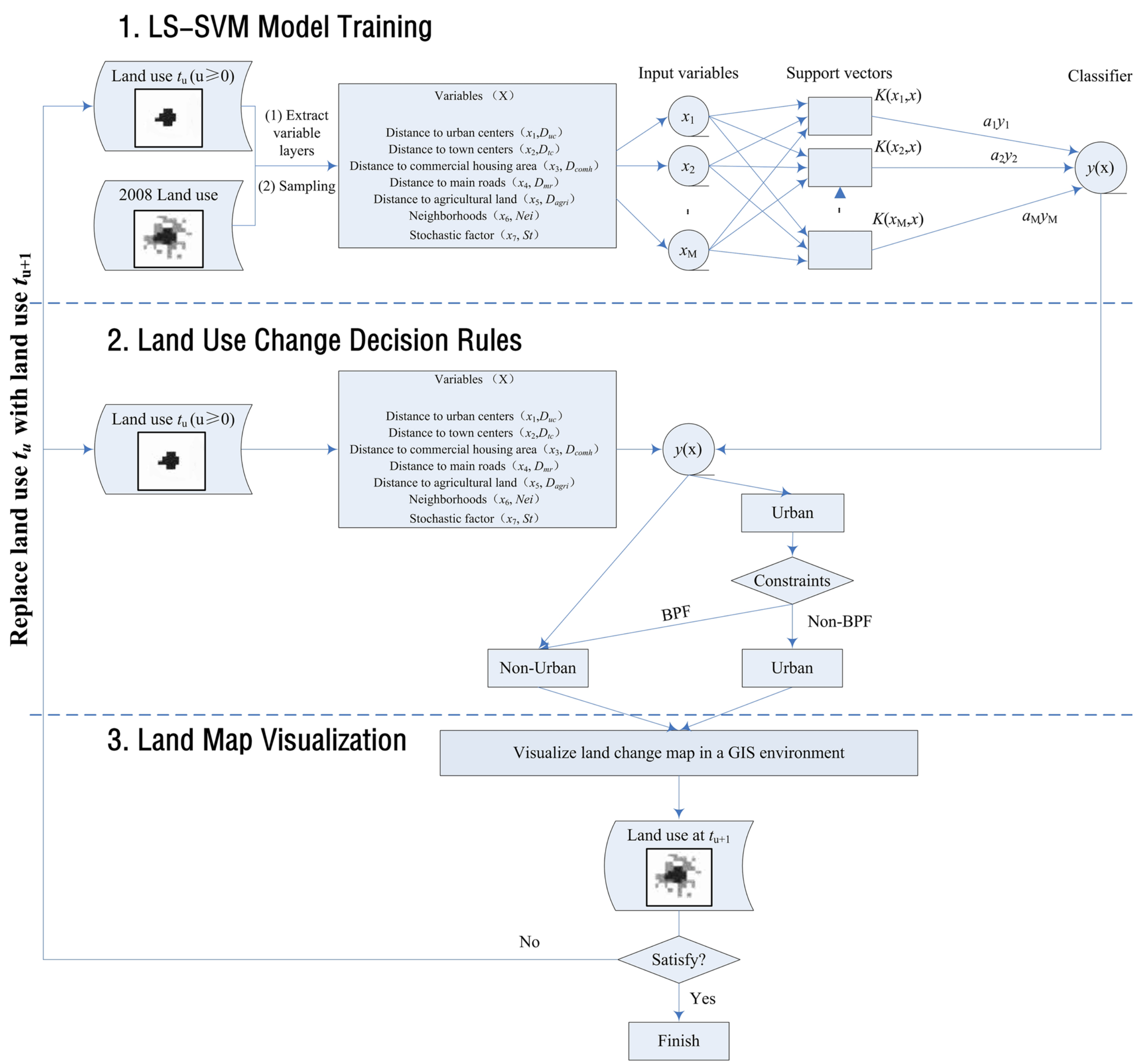


Fig. 3 Flowchart of the MachCA model for urban growth simulation. BPF represents Basic Protective Farmland

Rules, and Map Visualization which we show as a three stage work flow in Fig. 3. The first two modules were realized using LS-SVM lab v1.8 released by Suykens et al. (2002) while the map visualization module was realized in ArcGIS. Each of the three modules is discussed below.

In Fig. 3, t_u indicates the time at iteration u; M is the number of variables for model training; x_1, x_2,…, x_M are the input variables; K(x_1, x), K(x_2, x),…, K(x_M, x) are the support vectors; and $y(x)$ is the classifier. The land use map at time t_u ($u = 0$) is based on observed land use in 1992 classified from the relevant remote sensing image; land use maps at future time instants t_u (u ranges from 1 to n) are simulated results using the MachCA model. The minimum number of iterations of the model is 1, however, the total number of iterations (n) are not related to the temporal process but rather the calibration procedure. The LS-SVM Model Training module learns the CA transition rules using the LS-SVM method. Samples for model training were obtained from land use map at t_u and used to simulate the land use change at t_{u+1}. Therefore, the transition rules were dynamically reconstructed for each iteration of the model implementation. Different from traditional CA modelling methods, this LS-SVM model takes the neighbourhood as a variable (as seen in Table 1). The Land Use Change Decision Rules module determines whether a non-urban cell will be converted to an urban cell or not, based

Table 1 Variables use in the MachCA model for the generic simulation of urban growth in the Qingpu–Songjiang area of Shanghai, China

Variable	Meaning	Interpretation	Role
y	Conversion label	y is assigned to 1 if a cell is transformed from non-urban state to urban state, whereas y is assigned to -1 if a cell keeps its state over time	Training LS-SVM model and constructing final CA rules
D_{uc}	Distance to urban center	An Euclidean distance from a cell to the city center, that is usually the administrative center in China; a map at 1992 was used and the locations of the urban centers of the study area did not change since 1992	
D_{tc}	Distance to town center	An Euclidean distance from a cell to the town center; a map at 1992 was used and the locations of the town centers of the study area did not change since 1992	
D_{comh}	Distance to commercial housing area	An Euclidean distance from a cell to commercial housing area launched by local government of Shanghai; a map at 1992 was used and the locations of the commercial housing areas of the study area did not change since 1992	
D_{mr}	Distance to main roads	An Euclidean distance from a cell to main roads. The length of the main roads of the Qingpu–Songjiang area has experienced a rapid growth from 1992 to 2008 and a map of main roads at the middle time point (i.e. 2000) was used	
D_{agri}	Distance to agricultural land	An Euclidean distance from a cell to agricultural land; a map at 1992 was used and the protective agricultural land did not change since 1992	
Nei	Neighborhoods	5×5 cells of the immediate neighborhoods of a central cell which was given by: $Nei_h = \frac{\sum_{j=1}^{5\times5}(S_k = Urban)(h \neq k)}{5\times5-1}$	
St	Stochastic	A stochastic factor imitates the impact of an emergent development due to institutional policy, which was given by: $St = 1 + (-\ln r)^{\theta}$ where $r = 0.5$ and $\theta = 5$ were selected for the MachCA model according to previous research (Feng and Liu 2013a, b; Feng et al. 2011)	
LC	Local constraints	Local constraints on urban growth are regions unavailable for development such as basic protective farmland (BPF)	Constructing final CA rules

on the result from the LS-SVM model and subject to BPF (Basic Protective Farmland) constraints. The Map Visualization module uses GIS to visualize land change. The three modules iterate over the entire modelling process. If, at any time (e.g. t_u), the model generates a result with a low simulation accuracy (e.g. lower than the simulation accuracies generated in other areas of Shanghai (Feng and Liu 2013a, b; Feng et al. 2011)), then the entire MachCA model would be reconstructed at the next time step (i.e., t_{u+1}) based on resampling from the simulated results at t_u. Subsequently, the regenerated MachCA model would be used to simulate the land use map at time t_{u+1}.

3.2 Input data

In China, urban development is a complex process which is affected by many physical, institutional and economic factors such as globalization, rural industrialization, transportation and infrastructural development, and land management systems (Wei and Ye 2014). However, data reflecting institutional and economic factors are largely unavailable to researchers. Therefore, similar to other studies (Al-shalabi et al. 2013; Feng and Liu 2013a; Yang et al. 2008; Zhang et al. 2015), our model of urban growth focuses on understanding the physical dimension of the driving factors. A total of eight driving factors were chosen for simulating urban growth in the Qingpu–Songjiang area from 1992 to 2008. These driving factors were selected by considering urban growth dynamics in Shanghai (Feng and Liu 2013a, b) which include the distance-based variables, neighbourhoods, constraints, and a stochastic factor which we show in Table 1.

As identified in the spatial analysis of urban growth, spatial variables adopted in CA models should closely relate to urban development and land use changes (Rezayan et al. 2010; Stanilov and Batty 2011; Wu and Webster 1998). With regard to spatial variables used in the MachCA model, some of them (e.g. distance to urban centres) clearly meet such requirements as revealed by existing research (Feng et al. 2011; Wu 1998, 2002; Wu and Webster 1998). Other spatial variables, such as distance to commercial housing areas, were chosen to reflect the fact that these factors play important roles in the urban growth of Chinese cities. On the other hand, some of the widely

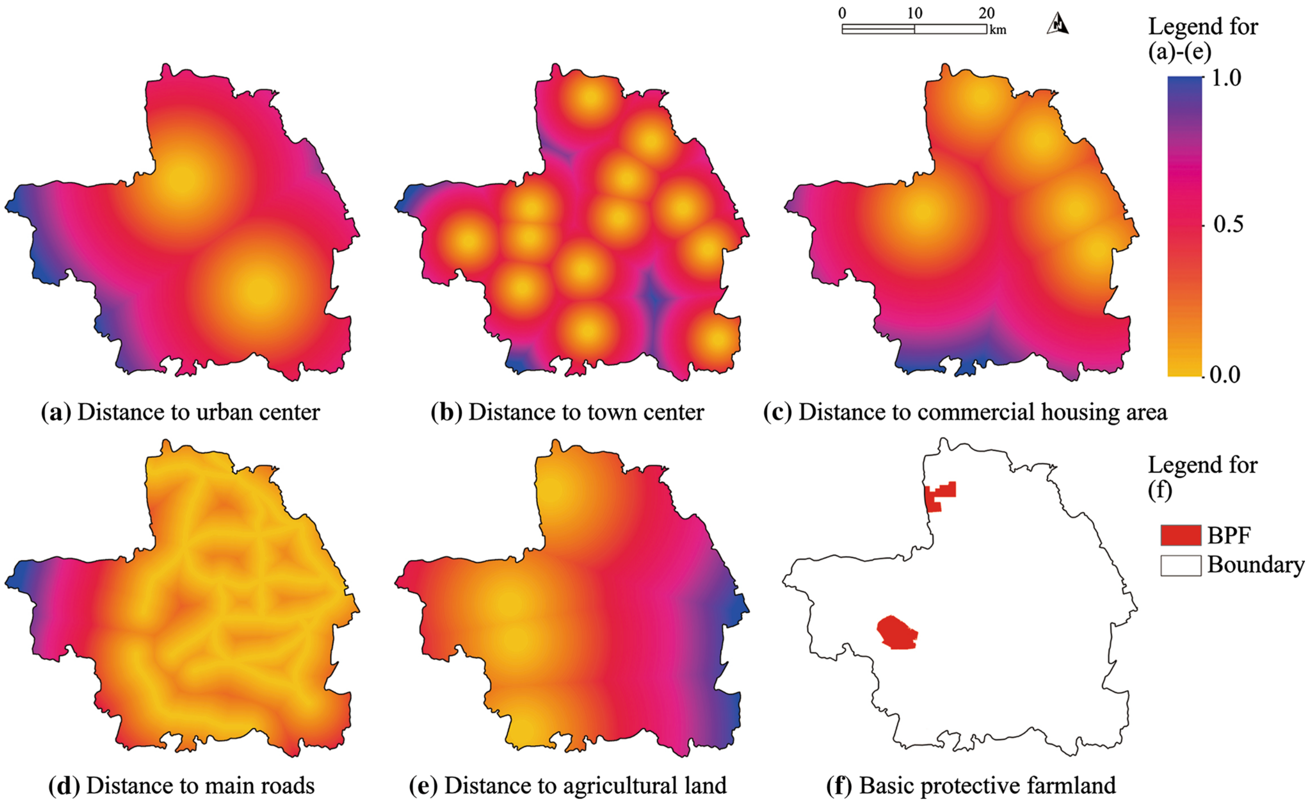


Fig. 4 Distribution of distance variables and Basic Protective Farmland (PBF) as a local constraint to urban growth

used variables, such as slope or digital elevation models, were excluded from this study as the terrain of Qingpu–Songjiang area is pretty flat. In addition, the process of urban growth is considered as being impacted by many factors that are unclear and uncertain and might not follow any well-defined growth trajectories (Feng et al. 2011; White and Engelen 2000; Wu and Webster 1998). This enables us to introduce a stochastic factor in the MachCA model. At the same time, water bodies and total areas available for development were considered as the local and global constraints, respectively. The total areas available for development include all areas excluding the urban area at 1992, water bodies, wetlands, and the BPF. Amongst such factors, spatial variables and neighbourhoods reflect the agglomeration effect of urban development and the attractive power of infrastructure compared with the stochastic and constraint factors (Feng and Liu 2013a, b).

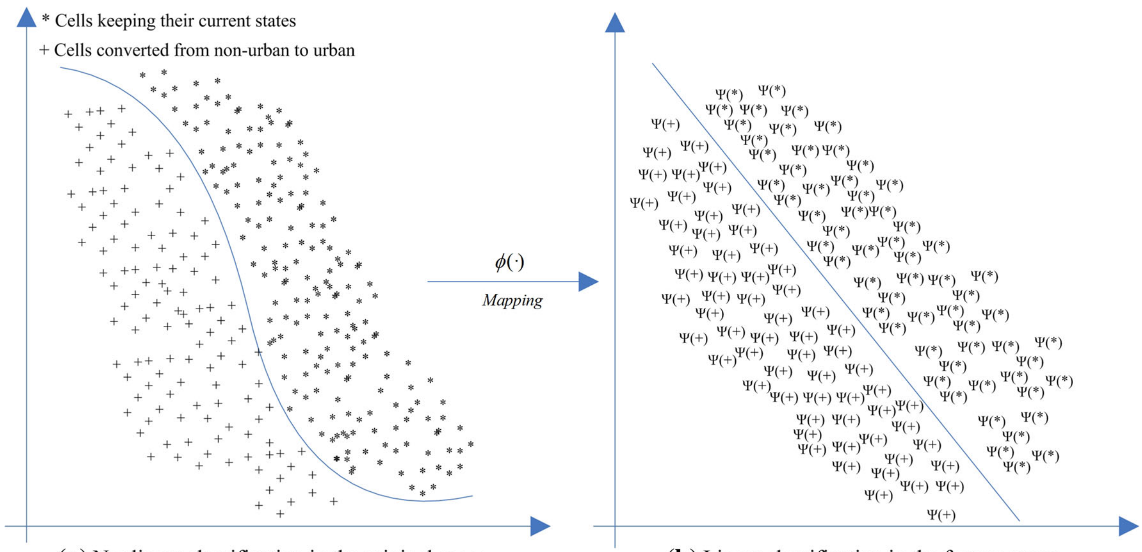
All distance variables used in this study were normalised by (Feng and Liu 2013a):

$$D_{norm} = \frac{D_{orig}}{D_{max} - D_{min}}, \quad (1)$$

where D_{norm} is the normalised value of each distance variable which ranges from 0 to 1; D_{orig} is the original value of each spatial variable; D_{max} and D_{min} are the maximum and minimum values of each spatial variable, respectively. Figure 4 illustrates these distance variables and the distribution of BPF as local constraint to urban growth.

3.3 Least squares support vector machines

LS-SVM method is a new development of standard SVMs which leads to direct optimization of the objective function that defines the hyperplane separating the two regions of the higher dimensional space by replacing the inequality constraint conditions in standard SVMs by equality constraint conditions (Suykens et al. 2002; Suykens and Vandewalle 1999). The LS-SVM method inherits the essential idea of the standard SVMs on searching the optimal classification hyper-plane in a higher dimensional space, but has lower computational complexities and less memory requirements than the standard SVMs (Ye and Xiong 2007). In this sense, it relaxes the conditions for optimality posed by original SVMs in by replacing the quadratic optimization function with a set of linear equations that can be solved simultaneously. Readers who are unfamiliar with these kinds of multivariate optimization methods are referred to the literature, particularly to Suykens and Vandewalle (1999).



(a) Nonlinear classification in the original space **(b)** Linear classification in the feature space

Fig. 5 Transferring the two-dimensional nonlinear classification in the original space into a linear classification in the feature space by using the LS-SVM method. *Left* Nonlinear classification in the original space, and *Right* Linear classification in the feature space. (Feng and Liu 2013a)

When modeling urban growth, it is somewhat complicated to determine the classification boundary for converting a non-urban cell to an urban cell (which we show schematically in Fig. 5a). Traditional linear methods such as logistic regression seem insufficient to find the classification boundary (Feng and Liu 2013a; Yang et al. 2008). In contrast, the LS-SVM method transfers the two-dimensional nonlinear classification (i.e. non-urban or urban classification) in the original space into a linear classification in the feature space, and in this way, nonlinear CA transition rules can be established (as implied in Fig. 5b).

To briefly introduce the LS-SVM method, we suppose that there are M training sample pairs $\{x_i, y_i\}_{i=1}^{M}$, where $x \in R^M$ is the ith input pattern with known output pattern and $y_i \in \{-1, +1\}$. Thus a decision function of the binary classifier can be constructed as follows:

$$f(\boldsymbol{x}) = \operatorname{sgn}[\boldsymbol{W}^T \varphi(x) + \beta] \tag{2}$$

where sgn is a decision rule, $\boldsymbol{W} = (w_1, w_2, \ldots, w_n)$ and β are the weight vectors and bias to the hyper-plane $< W \cdot \boldsymbol{x} > + \beta = 0$, respectively. In the LS-SVM, the formulation of an optimization problem is given by Suykens and Vandewalle (1999) and we follow their treatment and notation in the following. Then:

$$\min_{\boldsymbol{W},\beta,\varepsilon} J(\boldsymbol{W}, \beta, \varepsilon) = \frac{1}{2}\boldsymbol{W}^{\mathrm{T}}\boldsymbol{W} + \frac{1}{2} C \sum_{i=1}^{M} \varepsilon_i \tag{3}$$

$$s.t. \quad y_i\left[\boldsymbol{W}^{\mathrm{T}}\varphi(x_i) + \beta\right] = 1 - \varepsilon_i, \quad i = 1, \ldots, M \tag{4}$$

where $\varphi(\cdot)$ is a nonlinear function that maps the input space into a higher dimensional space; ε_i are variables (dummy error terms) created to violate constraints, i.e.,$y_i\left[\boldsymbol{W}^{\mathrm{T}}\varphi(x_i) + \beta\right] \geq 1$, which is also called as the regularization parameter that determines the training error and generalization potential of the LS-SVM model (Suykens and Vandewalle 1999). Furthermore, the Lagrange function for the above optimization problem is defined as:

$$\begin{aligned} L(\boldsymbol{W}, \beta, \varepsilon_i; \alpha_i) &= J(\boldsymbol{W}, \varepsilon_i) \\ &\quad - \sum_{i=1}^{M} \alpha_i \left\{ y_i \left[\boldsymbol{W}^{\mathrm{T}}\varphi(x_i) + \beta\right] - 1 + \varepsilon_i \right\} \end{aligned} \tag{5}$$

where α_i are Lagrange multipliers that can be either positive or negative.

Consequently, the conditions for optimality are derived by Suykens and Vandewalle (1999) as

$$\begin{cases} \dfrac{\partial L}{\partial \boldsymbol{W}} = 0 \rightarrow W = \displaystyle\sum_{i=1}^{M} \alpha_i y_i \varphi(x_i) \\ \dfrac{\partial L}{\partial \beta} = 0 \rightarrow \displaystyle\sum_{i=1}^{M} \alpha_i y_i = 0 \\ \dfrac{\partial L}{\partial \varepsilon_i} = 0 \rightarrow \alpha_i = C\varepsilon_i \\ \dfrac{\partial L}{\partial \alpha_i} = 0 \rightarrow y_i\left[\boldsymbol{W}^{\mathrm{T}}\varphi(x_i) + \beta\right] - 1 + \varepsilon_i = 0 \end{cases} \tag{6}$$

By eliminating ε_i and $\boldsymbol{W}$, the optimization problem can be written immediately as the solution to the following set of linear equations:

$$\begin{bmatrix} 0 & \boldsymbol{y}^{\mathrm{T}} \\ \boldsymbol{y} & \boldsymbol{\Omega} + C^{-1}I \end{bmatrix} \begin{bmatrix} \beta \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ 1_M \end{bmatrix} \tag{7}$$

where $\boldsymbol{\Omega} = \boldsymbol{Z}\boldsymbol{Z}^{\mathrm{T}}$, $\boldsymbol{Z} = [\varphi(x_1), \varphi(x_2), \ldots, \varphi(x_M)]^{\mathrm{T}}$, $\boldsymbol{y} = [y_1, y_2, \ldots, y_M]^{\mathrm{T}}$, $1_N = [1, 1, \ldots, 1]$, and $1_M = [\alpha_1, \alpha_2, \ldots,$

α_M]. According to Mercer's condition defining a positive-definite matrix, the matrix $\mathbf{\Omega} = \mathbf{Z}\mathbf{Z}^{\mathrm{T}}$ can be taken as a kernel:

$$\Omega_{ij} = y_i y_j \varphi(x_{i2})^{\mathrm{T}} \varphi(x_j) = y_i y_j K(x_i, x_j) \tag{8}$$

Among the choices presented by Vapnik (1998), the Gaussian radial basis function (RBF) is selected for modeling urban growth on the basis of to the related works (Feng and Liu 2013a; Huang et al. 2009; Liu et al. 2008; Yang et al. 2008) and this is given by:

$$K(\boldsymbol{x}, x_i) = \exp\left\{-||x - x_i||^2 / 2\sigma^2\right\} \tag{9}$$

where σ is a constant reflecting data distribution properties.

Accordingly, Eq. (1) is achieved as the objective binary classifier by solving the linear set of Eqs. (6) and (7), and thence enabling us to categorize the linearly non-separable urban boundaries. The decision function of the binary classifier is thus given by:

$$f(\boldsymbol{x}) = sgn\left[\sum_{i=1}^{M} \alpha_i y_i K(\boldsymbol{x}, x_i) + \beta\right] \tag{10}$$

and this enables us to refine the transition rules.

3.4 Transition rules

Based on the binary classifier of the LS-SVM model, whether a non-urban cell would be converted to urban cell or remain its current state at time t_{u+1} can be determined. The final CA transition rules were given by:

$$S_{t_{u+1}} = \begin{cases} 1, & s.t.\left(\sum_{i=1}^{M} \alpha_i y_i K(\boldsymbol{x}, x_i) + \beta\right)_{t_i} \geq 0 \quad \text{and} \quad LC \neq BPF \\ -1, & s.t.\left(\sum_{i=1}^{M} \alpha_i y_i K(\boldsymbol{x}, x_i) + \beta\right)_{t_i} < 0 \end{cases} \tag{11}$$

where 1 indicates the central non-urban cell will be converted to urban state while -1 indicates that the state of the central non-urban cell will remains unchanged. Note that the vectors $\boldsymbol{x}$ represent the heart of the model in that these contain the various input data that are defined in Table 1, and thus these are intimately related to the transition from non-urban -1 to urban $+1$.

3.5 LogCA model for comparison

As a comparison, a conventional logistic regression-based CA model (LogCA) was also constructed using the same input variables in the same study area to generate a simulation which is comparable to that of the MachCA model. The transition rule for each cell h in the LogCA was given by:

$$\begin{aligned} P_h^{t_u} = {} & (1/(1 + \exp(1.109 - 0.815D_{uc} - 0.766D_{tc} \\ & -0.587D_{comh} - 0.749D_{mr} + 0.532D_{agri})) \\ & \times Nei_h \times St \times LC \end{aligned} \tag{12}$$

where P_h^t is the land use conversion probability of a cell h at time t_u, Nei_{jh} is the neighborhood effect, St is the stochastic factor, and LC is the local constraint defined in Table 1. The distances D_* are also defined in this table. By comparing this conversion probability with a pre-defined threshold value (Liu and Feng 2012; Wu 2002), if the conversion probability of the cell at time t_u is larger than this pre-defined threshold (e.g. >0.5), the cell will be converted to an urban state in the subsequent time step, that is $y_h = +1$. Otherwise, the state of the cell will remain unchanged as $y_h = -1$.

4 Results and discussion

4.1 Simulation results

Using the variables listed in Table 1, a MachCA model which focuses on the transition from non-urban to urban use has been constructed to simulate the process of urban growth in Shanghai's Qingpu–Songjiang area from 1992 to 2008. In the MachCA model, the regularization parameter γ in the decision function and the coefficient σ^2 (sig2) in the radial basis function were defined as $\gamma = 10$ and sig2 $= 0.4$, respectively. Based on the training results with 12,470 samples, 12,470 support vectors were acquired with a bias $\beta = -0.8051$. The MachCA model was operated for 20 iterations, noting that these are not related to the temporal process but rather the calibration procedure. The simulation accuracies were calculated using a cell-by-cell comparison method (i.e. a confusion matrix) based on the 2008 observed land use map with the MachCA model generating the simulated map results. At the start of the simulation, the 1992 observed land use map was considered as the simulated result without a CA model (the NULL model) and was compared with the 2008 observed land use map with such a NULL model being initially proposed by Pontius and his colleagues (Pontius et al. 2004; Pontius and Spencer 2005). In the other case, the simulated results from 1 to 20 iterations were compared with the 2008 observed land use map and this generated the simulation accuracies which are shown in Fig. 6.

Figure 6 shows that the accuracy of the NULL model is 57.2 % (with respect to the 1992 vs. 2008 observations). The 1st iteration of the MachCA model generated an accuracy of 63.8 % with the accuracy increasing consistently for the first 16 iterations until it reaches its highest accuracy of 81.2 % at the 16th iteration. However, the

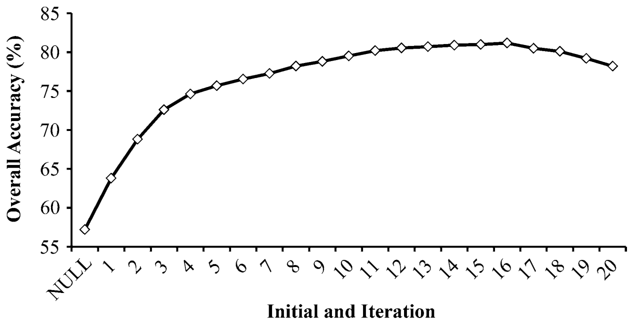


Fig. 6 The change in the accuracies through the iterations of the MachCA model

accuracy decreased slightly from 80.5 to 78.2 % in the last four iterations from 17 to 20 and the tendency over this latter range was towards stability. Therefore, the MachCA model was finally implemented with its 16th iteration. As a result, the urban growth simulations were conducted with both the LogCA and proposed MachCA models for 2008 in Qingpu–Songjiang area of Shanghai as we illustrate in Fig. 7a, b.

Figure 7 shows very strong visual similarities between the simulation maps generated by the LogCA and MachCA models for the end year 2008. Visual comparison is important because the human mind can quickly detect key patterns between different maps in contrast to spatial statistics which in general are not sufficiently indicative of subtle differences in pattern that only the mind's eye can discern (Pontius et al. 2004). It was observed from Fig. 7 that urban growth in Qingpu–Songjiang area has mainly occurred around the position of clusters of urbanized cells which existed at previous times. Using the method of Chen and Pontius (2010), simulated urban growth during 1992–2008 produced by both the LogCA and MachCA models were compared with the reference map of urban growth during this period as illustrated in Fig. 2, the map which was generated from remote sensing images. Figure 7 shows that the LogCA model missed 11.1 % of the total landscape that changed the state from non-urban to urban, incorrectly simulated 8.3 % of the total landscape that did not change the state, correctly simulated 31.6 % that changed the state, and correctly rejected 42.5 % of the total landscape that did not change the state from non-urban to urban. Hence the overall accuracy of the LogCA model is 80.6 % which include the hit (H), correct rejection of non-urban areas (CR) and the correct rejection of urban areas at 1992 (U).

Compared to the LogCA model, the MachCA model had a lower proportion of misses (10.4 vs. 11.1 %) and higher hits (32.4 vs. 31.6 %), while it also had lower correct rejections (42.3 vs. 42.6 %) and lower false alarms (8.3 vs. 8.4 %). Therefore, the overall accuracy is 81.2 % (Fig. 7). Although the MachCA model has generated only less than 1 percentage point better simulation accuracies than the LogCA model, the consistent improvement in most of the accuracy measurements still demonstrates a somewhat improved performance over the LogCA model.

4.2 Quantity and allocation errors

Apart from the misses, false alarms and correct rejections identified, the simulation errors were collapsed into a budget of quantity and allocation errors using the method proposed by Pontius and his colleagues (Pontius et al. 2004; Pontius and Malanson 2005; Pontius and Millones 2011; Pontius and Spencer 2005). This error budget reveals the proportions of components of agreement and disagreement according to the comparison of map pairs. The error budget is derived from comparison of the simulation and reference maps based on two measurements: quantity errors (Q) and allocation errors (A). Using the error budget method, the quantity and allocation errors were computed for the 2008 simulation maps produced by both the LogCA and MachCA models respectively as we illustrate in Fig. 8.

Figure 8 shows, for the LogCA model, amongst the 19.4 % simulation errors, 2.8 % is due to quantity error and 16.6 % to allocation error. For the MachCA model, amongst the 18.8 % simulation errors, 2.0 % is due to quantity error and 16.8 % due to allocation error. The quantity error is the absolute difference between misses and false alarms, which derives from the fact that the model missed some location of the observed urban growth and simulated the wrong location. If a model simulates less than the reference urban gain, then the false alarms are fewer than the misses. If a model simulates more than the reference urban gain, then the false alarms are more than the misses. The logCA model simulated less urban gains than the reference urban gains by 2.8 %, while the MachCA model simulated less urban gain by 2.0 %. In fact, more non-urban cells could be converted to urban cells with more model iterations (e.g. more than 16 iterations). Therefore, the MachCA model could miss less reference growth with more iterations, or it could simulate more wrong locations and suffer a lower overall agreement. Hence, the MachCA model simulated less change than the reference change to avoid the false alarms and minimize the errors.

Obviously the hits (>31 %) produced by the both CA models applied to the Qingpu–Songjiang area of Shanghai are much higher than that of a relatively steadily growing area such as Logan City, Australia with less than 8 % hits (Liu et al. 2014). This may be due largely to the increase of urban areas in the Qingpu–Songjiang area over the modeled period compared to Logan City, rather than any particular features of the model used. It is speculated that it is easier for the same CA model to get a higher number of

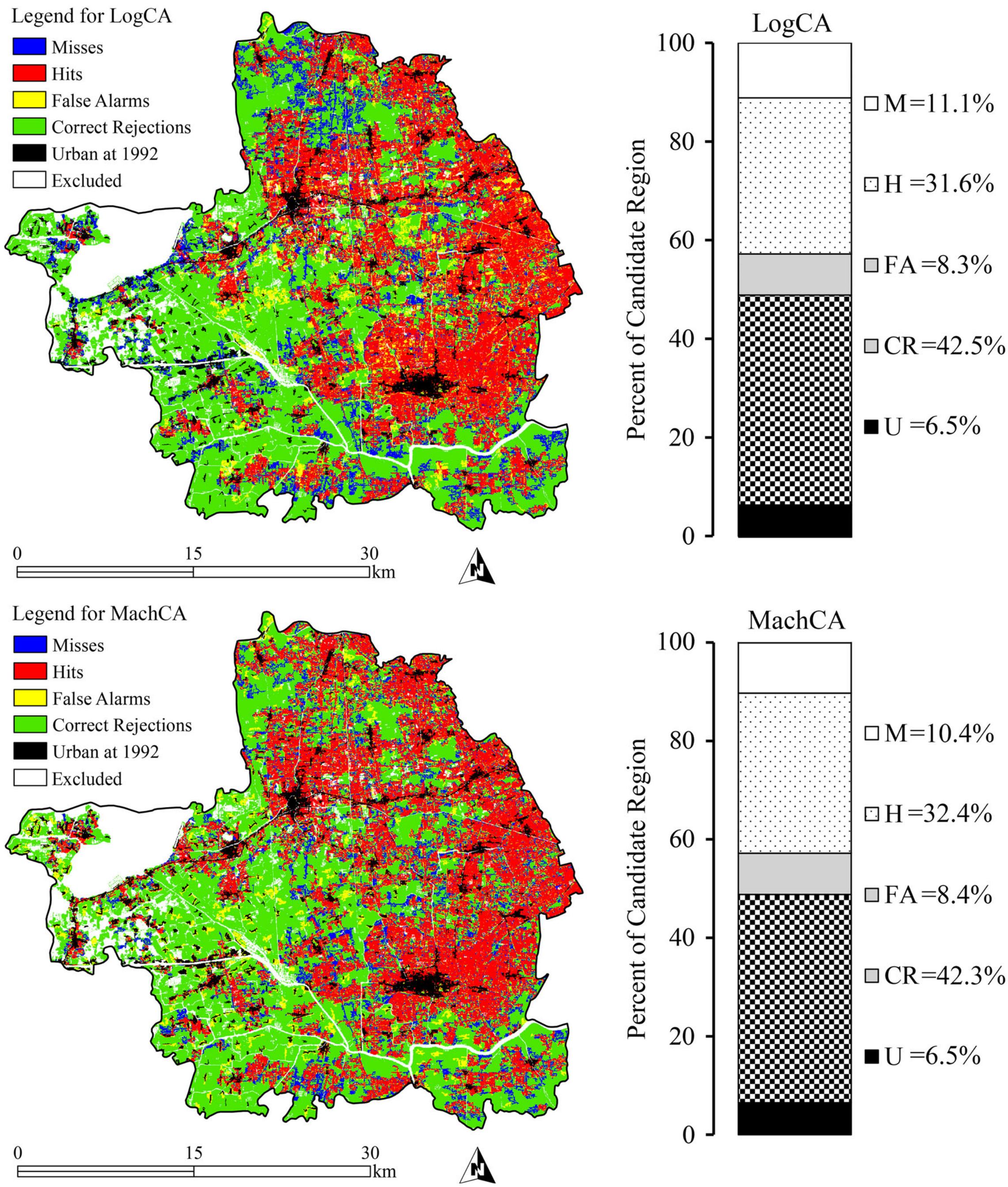


Fig. 7 The 2008 simulated urban growth in the Qingpu–Songjiang area using both the LogCA and MachCA models. Hits (H) indicate that observed growth was simulated as growth, misses (M) indicate that observed growth was simulated as persistence, false alarms (FA) indicate that observed persistence was simulated as growth, and correct rejections (CR) indicate that observed persistence was simulated as persistence (after Chen and Pontius 2010). The land use titled 'excluded' indicates water areas and wetlands

hits in a fast growing area (Pontius et al. 2008). However, the MachCA model obtained more hits and less quantity error in simulating the same study area compared with the LogCA model, but the allocation error of the MachCA model is slightly higher.

Although MachCA does not perform better in all aspects compared to the LogCA, the model does prove to be capable of capturing urban dynamics associated with the complex boundaries of the urban-nonurban fringe. Most importantly, the MachCA model is different from conventional CA models in retrieving the transition rules as it extracts these rules from the location of the input data in high dimensional feature spaces (Feng and Liu 2013a; Liu 2008; Yang et al. 2008). It also differs from the standard

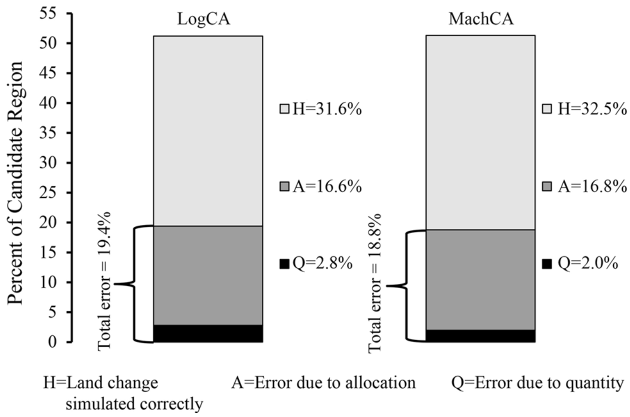


Fig. 8 Quantity error, allocation error and hits in the non-urban area of 1992

SVM based CA model (Yang et al. 2008) in that the model directly retrieves a yes–no decision with respect to the transition rules, rather than relying on the definition of an arbitrary probability threshold value through trial-and-error.

5 Conclusion

This research has demonstrated that various nonlinear features of urban dynamics can be captured by the LS-SVM method. Based on the extracted spatial variables and the classification of initial states, the nonlinear CA transition rules can be constructed in a higher dimensional space implied by the LS-SVM approach. With these nonlinear transition rules, the MachCA model identifies the complex boundaries of urban extent which cannot be sufficiently discriminated by conventional methods such as the logistic regression and principal component analysis.

The MachCA model was applied in a very rapidly growing area of one of the largest Chinese city regions using data from 1992 to 2008. The error budget conducted between the reference and simulation maps shows that, both in 2001 and 2008, the MachCA model reduces by less than a percentage point the quantity error in comparison with its slightly simpler cousin, the LogCA model.

As a novel urban modeling technique, MachCA enriches the theories and methods of geographical CA by addressing the complex boundaries of urban extent. However, limitations of this method also exist because the LS-SVM method is relatively complex in its theory and implementation. Therefore, it requires an understanding of the mechanisms of urban dynamics as well as mastery over the requisite mathematical and computer knowledge for its application.

Acknowledgments This study was supported by National Natural Science Foundation of China (41406146), Natural Science Foundation of Shanghai Municipality (13ZR1419300) and Research Fund for the Doctoral Program of Higher Education of China (20123104120002).

References

Al-shalabi M, Billa L, Pradhan B, Mansor S, Al-Sharif AAA (2013) Modelling urban growth evolution and land-use changes using GIS based cellular automata and SLEUTH models: the case of Sana'a metropolitan city Yemen. Environ Earth Sci 70(1):425–437

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

Barredo JI, Kasanko M, McCormick N, Lavalle C (2003) Modeling dynamic spatial processes: simulation of urban future scenarios through cellular automata. Landsc Urban Plan 64:145–160

Batty M, Xie Y, Sun Z (1999) Modeling urban dynamics through GIS-based cellular automata. Comput Environ Urban Syst 23:205–233

Cao K, Huang B, Li M, Li W (2014) Calibrating a cellular automata model for understanding rural–urban land conversion: a Pareto front-based multi-objective optimization approach. Int J Geogr Inf Sci 28(5):1028–1046

Chen H, Pontius RG (2010) Diagnostic tools to evaluate a spatial land change projection along a gradient of an explanatory variable. Landsc Ecol 25:1319–1331

Cheng J, Masser I (2003) Urban growth pattern modeling: a case study of Wuhan city, PR China. Landsc Urban Plan 62:199–217

Clarke KC, Gaydos LJ (1998) Loose-coupling of a cellular automaton model and GIS: long-term urban growth prediction for San Francisco and Washington/Baltimore. Int J Geogr Inf Sci 12(7):699–714

Couclelis H (1997) From cellular automata to urban models: new principles for model development and implementation. Environ Plann B 24(2):165–174

Deng Z, Zhang X, Li D, Pan G (2014) Simulation of land use/land cover change and its effects on the hydrological characteristics of the upper reaches of the Hanjiang Basin. Environ Earth Sci. doi:10.1007/s12665-014-3465-5

Ding W, Wang R, Wu D, Liu J (2013) Cellular automata model as an intuitive approach to simulate complex land-use changes: an evaluation of two multi-state land-use models in the Yellow River Delta. Stoch Environ Res Risk Assess 27(4):1–9

Feng Y, Liu Y (2013a) A cellular automata model based on nonlinear kernel principal component analysis for urban growth simulation. Environ Plann B 40(1):116–134

Feng Y, Liu Y (2013b) A heuristic cellular automata approach for modeling urban land-use change based on simulated annealing. Int J Geogr Inf Sci 27(3):449–466

Feng Y, Liu 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

Gong JZ, Liu YS, Xia BC, Zhao GW (2009) Urban ecological security assessment and forecasting, based on a cellular automata model: a case study of Guangzhou China. Ecol Model 220(24):3612–3620

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(20–22):3761–3772

He C, Wei A, Shi P, Zhang Q, Zhao Y (2011) Detecting land-use/land-cover change in rural–urban fringe areas using extended change-vector analysis. Int J Appl Earth Obs 13:572–585

Hu Z, Lo C (2007) Modeling urban growth in Atlanta using logistic regression. Comput Environ Urban Syst 31(6):667–688
Huang B, Xie C, Tay R, Wu B (2009) Land-use-change modeling using unbalanced support-vector machines. Environ Plann B 36:398–416
Ji W, Ma J, Twibell RW, Underhill K (2006) Characterizing urban sprawl using multi-stage remote sensing images and landscape metrics. Comput Environ Urban Syst 30:861–879
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
Lau KH, Kam BH (2005) A cellular automata model for urban land-use simulation. Environ Plann B 32:247–263
Li W, Wu C, Zang S (2014) Modeling urban land use conversion of Daqing City, China: a comparative analysis of "top-down" and "bottom-up" approaches. Stoch Environ Res Risk Assess 28(4):817–828
Liao J, Tang L, Shao G, Qiu Q, Wang C, Zheng S, Su X (2014) A neighbor decay cellular automata approach for simulating urban expansion based on particle swarm intelligence. Int J Geogr Inf Sci 28(4):720–738
Liu Y (2008) Modelling urban development with geographical information systems and cellular automata. CRC Press, Taylor & Francis Group, New York
Liu Y (2012) Modelling sustainable urban growth in a rapidly urbanizing region using a fuzzy constrained cellular automata approach. Int J Geogr Inf Sci 26(1):151–167
Liu Y, Feng Y (2012) A logistic based cellular automata model for continuous urban growth simulation: a case study of the Gold Coast city, Australia. In: Heppenstall AJ, Crooks AT, See LM, Batty M (eds) Agent-based models of geographical systems. Springer, Dordrecht, pp 643–662
Liu Y, Phinn SR (2003) Modelling urban development with cellular automata incorporating fuzzy-set approaches. Comput Environ Urban Syst 27:637–658
Liu X, Li X, Shi X, Wu S, Liu T (2008) Simulating complex urban development using kernel-based non-linear cellular automata. Ecol Model 211:169–181
Liu Y, Feng Y, Pontius RG (2014) Spatially-explicit simulation of urban growth through self-adaptive genetic algorithm and cellular automata modeling. Land 3(3):719–738
Maeda EE, Almeida CM, Ximenes AC, Formaggio AR, Shimabukuro YE, Pellikka P (2011) Dynamic modeling of forest conversion: simulation of past and future scenarios of rural activities expansion in the fringes of the Xingu National Park, Brazilian Amazon. Int J Appl Earth Obs 13:435–446
Mao X, Meng J, Xiang Y (2013) Cellular automata-based model for developing land use ecological security patterns in semi-arid areas: a case study of Ordos, Inner Mongolia China. Environ Earth Sci 70(1):269–279
Mitsova D, Shuster W, Wang X (2011) A cellular automata model of land cover change to integrate urban growth with open space conservation. Landsc Urban Plan 99(2):141–153
Nourqolipour R, Shariff ARBM, Balasundram SK, Ahmad NB, Sood AM, Buyong T, Amiri F (2014) A GIS-based model to analyze the spatial and temporal development of oil palm land use in Kuala Langat district. Environ Earth Sci, Malaysia. doi:10.1007/s12665-014-3521-1
Pontius RG, Malanson J (2005) Comparison of the structure and accuracy of two land change models. Int J Geogr Inf Sci 19:243–265
Pontius RG, Millones M (2011) Death to Kappa: birth of quantity disagreement and allocation disagreement for accuracy assessment. Int J Remote Sens 32:4407–4429
Pontius RG, Spencer J (2005) Uncertainty in extrapolations of predictive land-change models. Environ Plann B 32:211–230
Pontius RG, Huffaker D, Denman K (2004a) Useful techniques of validation for spatially-explicit land-change models. Ecol Model 179(4):445–461
Pontius RG, Shusas E, McEachern M (2004b) Detecting important categorical land changes while accounting for persistence. Agric Ecosyst Environ 101:251–268
Pontius RG, Boersma W, Castella JC, Clarke K, Nijs T, Dietzel C, Duan Z, Fotsing E, Goldstein N, Kok K, Koomen E, Lippitt CD, McConnell W, Sood AM, Pijanowski B, Pithadia S, Sweeney S, Trung TN, Veldkamp AM, Verburg PH (2008) Comparing the input, output, and validation maps for several models of land change. Ann Reg Sci 42(1):11–47
Rezayan H, Delavar MR, Frank AU, Mansouri A (2010) Spatial rules that generate urban patterns: emergence of the power law in the distribution of axial line length. Int J Appl Earth Obs 12:317–330
Schölkopf B, Smola A, Müller K-R (1998) Nonlinear component analysis as a kernel eigenvalue problem. Neural Comput 10(5):1299–1319
Shafizadeh Moghadam H, Helbich M (2013) Spatiotemporal urbanization processes in the megacity of Mumbai, India: a Markov chains-cellular automata urban growth model. Appl Geogr 40:140–149
Shanghai Municipal Statistics Bureau (1985) Shanghai statistical yearbook. China Statistics Press, Beijing
Shanghai Municipal Statistics Bureau (2008) Shanghai statistical yearbook. China Statistics Press, Beijing
Stanilov K, Batty M (2011) Exploring the historical determinants of urban growth patterns through cellular automata. Trans GIS 15(3):253–271
Suykens JAK, Vandewalle J (1999) Least squares support vector machine classifiers. Neural Process Lett 9(3):293–300
Suykens JAK, Brabanter JD, Lukas L, Vandewalle J (2002) Weighted least squares support vector machines: robustness and sparse approximation. Neurocomputing 48:85–105
Vapnik V (1998) Statistical learning theory. Wiley, New York
Vapnik V (2000) The nature of statistical learning theory, 2nd edn. Springer, New York
Verburg PH, Soepboer W, Veldkamp A, Limpiada R, Espaldon V, Mastura S (2002) Modeling the spatial dynamics of regional land use: the CLUE-S model. Environ Manag 30(3):391–405
Wang F, Hasbani JB, Wang X, Marceau DJ (2011) Identifying dominant factors for the calibration of a land-use cellular automata model using Rough Set Theory. Comput Environ Urban Syst 35(2):116–125
Wei YD, Ye X (2014) Urbanization, urban land expansion and environmental change in China. Stoch Environ Res Risk Assess 28(4):757–765
White R, Engelen G (2000) High resolution modeling of the spatial dynamics of urban and regional systems. Comput Environ Urban Syst 24:383–400
Wu F (1998) SimLand: a prototype to simulate land conversion through the integrated GIS and CA with AHP-derived transition rules. Int J Geogr Inf Sci 12(1):63–82
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 (1998) Simulation of land development through the integration of cellular automata and multicriteria evaluation. Environ Plann B 25:103–126
Wu D, Liu J, Wang S, Wang R (2010) Simulating urban expansion by coupling a stochastic cellular automata model and socioeconomic indicators. Stoch Environ Res Risk Assess 24(2):235–245

Yang Q, Li X, Shi X (2008) Cellular automata for simulating land use changes based on support vector machines. Comput Geosci 34:592–602

Ye X, Dang A (eds.) (2013) Spatial analysis and modeling on urban and regional development. Ann GIS 19(3): 129

Ye J, Xiong T (2007) SVM versus least squares SVM. In Proceeding In: International conference on artificial intelligence and statistics, pp 640–647

Yin J, Yin Z, Hu X, Xu S, Wang J, Li Z, Zhong H, Gan F (2011) Multiple scenario analyses forecasting the confounding impacts of sea level rise and tides from storm induced coastal flooding in the city of Shanghai China. Environ Earth Sci 63(2):407–414

Yue W, Liu Y, Fan P, Ye X, Wu C (2012) Assessing spatial pattern of urban thermal environment in Shanghai China. Stoch Environ Res Risk Assess 26(7):899–911

Yue W, Ye X, Xu J, Xu L, Lee J (2014) A brightness–darkness–greenness model for monitoring urban landscape evolution in a developing country–A case study of Shanghai. Landsc Urban Plan 127:13–17

Zhang H, Jin X, Wang L, Zhou Y, Shu B (2015) Multi-agent based modeling of spatiotemporal dynamical urban growth in developing countries: simulating future scenarios of Lianyungang city, China. Stoch Environ Res Risk Assess 29:63–78