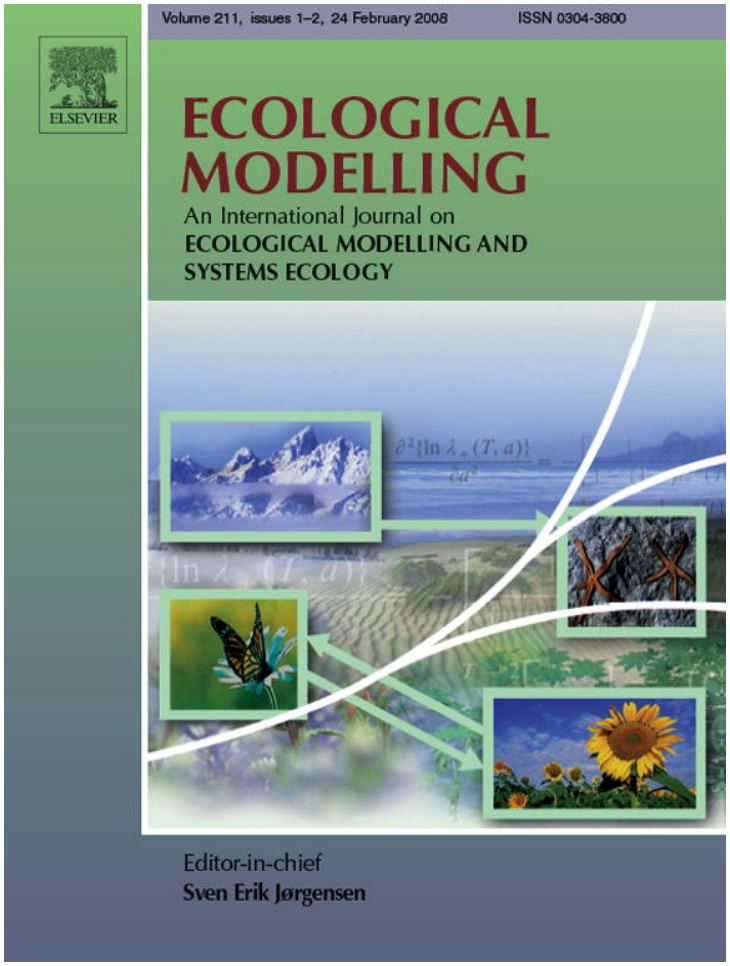
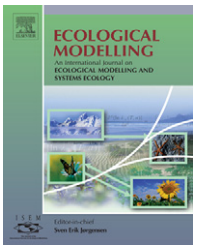


ECOLOGICAL MODELLING 211 (2008) 169–181

available at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/ecolmodel



Simulating complex urban development using kernel-based non-linear cellular automata☆

Xiaoping Liu[a], *Xia Li*[a,c,*], *Xun Shi*[b], *Shaokun Wu*[a], *Tao Liu*[a]

[a] *School of Geography and Planning, Sun Yat-sen University, Guangzhou 510275, PR China*
[b] *Department of Geography, Dartmouth College, Hanover, NH 03755, USA*
[c] *Guangzhou Institute of Geography, Guangzhou 510070, PR China*

ARTICLE INFO

Article history:
Received 11 April 2006
Received in revised form
17 August 2007
Accepted 30 August 2007
Published on line 22 October 2007

Keywords:
Cellular automata
Transition rules
Non-linear
Kernel-based
Urban systems

ABSTRACT

Cellular automata (CA) have been increasingly used to simulate complex urban systems. Empirical data can be used to calibrate CA models so that realistic urban patterns can be generated. Traditionally, the calibration procedure employs linear regression methods, e.g., multicriteria evaluation. However, the evolution of urban systems often manifests the complexity of non-linear features, for which the linear transition rules are insufficient. This paper proposes to use the kernel-based learning techniques to acquire non-linear transition rules for CA. The kernel-based approach maps the original data vectors to an implicit high-dimensional feature space, through which complex non-linear problems are translated into simple linear problems. Compared with a widely used non-linear method, neural network, the kernel method is more mathematically "transparent" and therefore the results are easier to be analyzed. A case study of simulating the expansion of Guangzhou, a fast growing city in China, shows that the kernel-based CA even achieves a slightly higher accuracy than a neural-network-based CA.

1. Introduction

Cellular automata (CA) were first proposed by Ulam in the 1940s and soon used by Von Neumann to investigate the logical nature of self-reproducible systems (White and Engelen, 1993). CA adopt a bottom-up approach, through which local individual behaviors can give rise to complex global patterns. An increasing number of recent studies have used CA to simulate complex geographical phenomena, such as population dynamics (Couclelis, 1988), wildfire spreading (Clarke et al., 1994; Karafyllidis and Thanailakis, 1997; Hargrove et al., 2000), epidemic propagation (Sirakoulis et al., 2000), forest dynamics (Lett et al., 1999), landscape changes (Wang and Zhang, 2001; Soares-Filho et al., 2002), urban growth (Batty and Xie, 1994; Couclelis, 1997; Clarke et al., 1997; Li and Yeh, 2000, 2002; Wu, 2002), and land-use changes (White and Engelen, 1993; Li and Yeh, 2002). Previous work shows that complex geographical phenomena can be effectively simulated using CA's simple local transition rules. This is in accordance with the complexity theory that a complex system may result from the interactions of simple subsystems (White and Engelen, 1993; Clarke et al., 1994).

Among all the geographical applications of CA, urban simulation may be the one that has been best explored and has most case studies (White and Engelen, 1993; Batty and Xie, 1994; Couclelis, 1997; Clarke and Gaydos,

☆ This paper proposes to use kernel-based learning techniques to acquire non-linear transition rules of cellular automata for urban simulation.

* *Corresponding author* at: School of Geography and Planning, Sun Yat-sen University, Guangzhou 510275, PR China. Tel.: +86 20 33047500; fax: +86 20 84115833.
E-mail address: lixia@mail.sysu.edu.cn (X. Li).

0304-3800/$ – see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.ecolmodel.2007.08.024

1998; Wu and Webster, 1998; Li and Yeh, 2004). Couclelis' (1985, 1997) early research shows that CA can be used as an analogue or metaphor to study how a variety of urban dynamics might arise; Batty and Xie (1994, 1997) carried out some pioneering work that simulates land use dynamics in the city of Buffalo, NY using CA and GIS; White et al. (1997) developed a CA model to simulate the land-use pattern of Cincinnati, Ohio; Clarke's (1997) simulation of urbanization of the San Francisco Bay Area; Li and Yeh (2000) were among the pioneers to integrate CA and GIS, and used the integrated system to simulate the rapid urban expansion in the Pearl River Delta, China.

Transition rules are essential in urban CA. These rules are represented by the parameters and weights associated with the spatial variables involved in the simulation. Usually, empirical data are used to derive and calibrate the values of these parameters and weights. Various approaches have been explored for the derivation and calibration. For example, Clarke et al. (1997) generates simulated scenarios using different combinations of parameter values, and then determines the most optimal combination through visually comparing the simulation results with the actual situation. While this approach is simple, it is not necessarily efficient and reliable when the possible combinations are numerous due to a large set of variables. Wu (1998) propose a more structured procedure based on the hierarchical analysis process (AHP) that defines the parameter values in a heuristic way. Later, logistic regression model was developed to obtain transition rules (Soares-Filho et al., 2002; Wu, 2002). Both the AHP and the logistical regression methods are inherently linear and consequently not good at dealing with complex relationships among a large number of spatial variables in urban land use dynamics. To overcome this problem, neural-network CA model was developed to handle the complex relationships and obtain parameter values automatically (Li and Yeh, 2002; Aitkenhead et al., 2004). Neural networks can significantly improve the CA's performance in urban simulation. However, they are difficult to be mathematically analyzed and tend to have problems of over-fitting. Exploration of other alternative non-linear methods for rule-generation in urban CA is then academically interesting may result in useful practical applications as well.

This study tests the performance of a kernel-based method in defining non-linear transition rules for urban CA. This method uses a kernel function to map data from the original feature space to a higher dimensional space, through which a non-linear problem is translated into a linear one and is to be solved in a linear way (Cawley and Talbot, 2003; Abdallah et al., 2004). Particularly for urban CA, the linear transition rules defined in the higher dimensional space can be used to represent the non-linear relationships among various spatial variables in the land use change. A kernel-based method can also reduce the number of parameters used by the classifier, since the parameters can be only about the most relevant samples. In this study, we developed a kernel-based CA model and applied it to the simulation of land use change in Guangzhou, a fast developing city in Southern China.

2. Retrieving transition rules of CA using kernel-based methods

Kernel-based learning methods were recently developed in the field of machine learning (Bousquet and Perez-Cruz, 2003), and have been quickly applied to non-linear problems in many applications (Cho et al., 2004). These methods are based on the Structural Risk Minimization principles in the computational learning theories (Vapnik, 1998). Compared with some widely used non-linear methods like neural network, they are relatively easy to be mathematically analyzed, because the classification is eventually a linear process in the high dimensional feature space (Klaus-Robert et al., 2001). In addition, some previous work demonstrates that the kernel-based methods are robust (Huang et al., 2004). Some representative methods in this category include support vector machines (SVM) (Scholkopf et al., 1997), kernel Fisher discriminant (KFD) (Mika et al., 1999) and kernel principal component analysis (Kernel PCA) (Scholkopf et al., 1998). Among these methods, KFD has been shown by many studies to have better performance than the others (Mika et al., 1999; Liu et al., 2004; Liang and Shi, 2004), and therefore was chosen in this study to develop the non-linear CA for urban simulation.

The traditional Fisher discriminant is linear and has proven successful in linear classification problems. Specifically, it seeks to find a linear transformation that maximizes the variances of between-class scatters and minimizes the variances of within-class scatters (Zhang and Huang, 2005). The criteria (called Fisher criteria) in this transformation are that the ratio of the mean between classes and the sum of variance in classes are maximized. Its basic process is to project d-dimensional data of j classes to a certain direction, on which the points of the same classes will be clustered together, and the points of different classes will be separated as much as possible. Fig. 1 shows a classification using the Fisher 2-dimensional linear discriminant. The data points in this figure are to be classified into two classes, Class 1 and Class 2. Their projections on either X_1 axis or X_2 axis are overlapped to some extent, and therefore they cannot be well classified directly based on values for X_1 or X_2. The linear Fisher discriminant

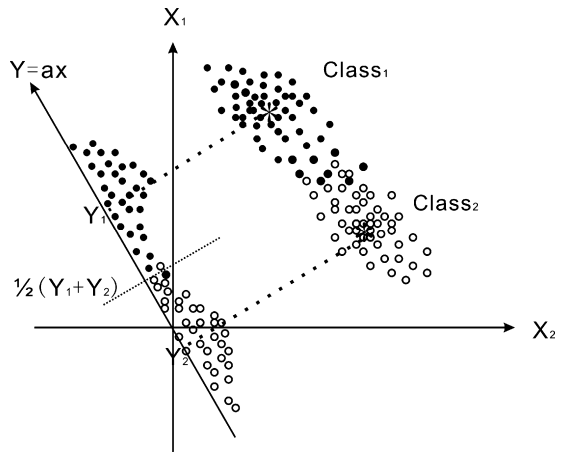


Fig. 1 – Classification using linear Fisher discriminant.

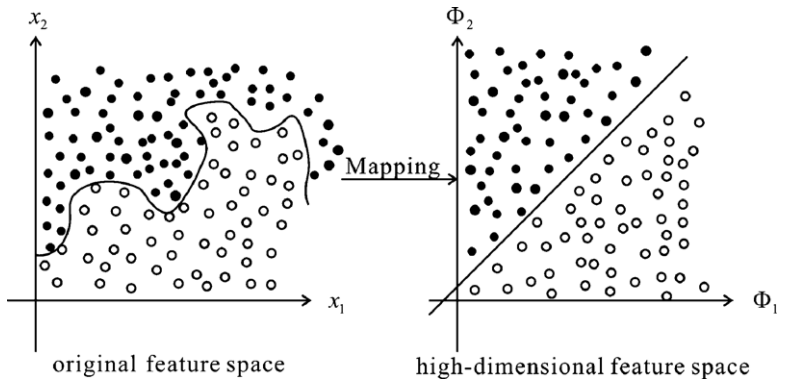


Fig. 2 – Mapping the input vectors into high dimensional feature space using kernel functions.

finds a beeline Y and projects all the data points onto it. The criteria for defining Y include maximizing the distance between the centers of the two clusters of the projected points (represented by Y_1 and Y_2, respectively), and minimizing the overlap of the two clusters. In this way, the perpendicular line passing through the midpoint between Y_1 and Y_2 will well classify the original data points.

For more complex classification problems, Mika et al. (1999) and Baudat and Anouar (2000) propose non-linear kernel Fisher discriminant (KFD). The KFD methods use certain kernel functions to project input vectors to high dimensional feature space, through which complex non-linear problems can be translated into simple linear problems (Fig. 2).

Formally, dataset X contains N d-dimensional samples, i.e., $X = \{x_1, x_2, \ldots, x_N\}$, and these samples belong to two classes, w_1 and w_2; The w_1 samples are denoted as $X_1 = \{x_1^1, x_2^1, \ldots, x_{N_1}^1\}$, and the w_2 samples are denoted as $X_2 = \{x_1^2, x_2^2, \ldots, x_{N_2}^2\}$; Samples $x \in R^d$ can be translated from the original feature space to a high-dimensional feature space F through a non-linear mapping function Φ, which can be represented as follows:

$$\Phi : R^d \to F, x \to \Phi(x) \tag{1}$$

After the translation, the traditional linear Fisher discriminant can be applied in the new feature space F. The Fisher criterion in F is then defined by:

$$J(w) = \frac{w^T S_b^\Phi w}{w^T S_w^\Phi w} (w \in F) \tag{2}$$

where w is the discriminant vector, S_b^Φ and S_w^Φ are the corresponding scatter matrices defined in F:

$$S_b^\Phi = (m_1^\Phi - m_2^\Phi)(m_1^\Phi - m_2^\Phi)^T \tag{3}$$

$$S_w^\Phi = \sum_{j=1}^{2} \sum_{i=1}^{N_j} (\Phi(x_i^j) - m_j^\Phi)(\Phi(x_i^j) - m_j^\Phi)^T \tag{4}$$

where $m_j^\Phi = \frac{1}{N_j} \sum_{i=1}^{N_j} \Phi(x_i^j), j = 1, 2.$

If the dimensions of F are very high, the linear Fisher discriminant may not be directly solvable. Kernel functions are introduced to overcome this difficulty. An attractive point of kernel functions is that the scalar product can be implicitly computed in the feature space, without explicitly using or even knowing the between-space mapping (Klaus-Robert et al., 2001; Scholkopf et al., 1999; Feng and Shi, 2004). Each (linear) algorithm that only uses scalar products can implicitly be executed by using kernels, i.e., one can construct a non-linear version of a linear algorithm. Commonly used kernel functions include Polynomial kernel, $K(x,y) = (xy + 1)^d$, and Radial Basis Function kernel (RBF), $K(x, y) = \exp(-|x - y|^2/\sigma^2)$.

The theory of reproducing kernels indicates that w can be written as follows (Aronszajn, 1950):

$$w = \sum_{i=1}^{N} a_i \Phi(x_i) \tag{5}$$

According to Eq. (5) and the definition of m_j^Φ, $w^T m_j^\Phi$ can be written as follows:

$$w^T m_j^\Phi = \frac{1}{N_j} \sum_{i=1}^{N} \sum_{k=1}^{N_j} a_i k(x_i, x_k^j) = a^T P_j \tag{6}$$

where $P_j = (P_{jk})_{k=1,2\ldots N}, P_{jk} = 1/N_j \sum_{i=1}^{N_j} K(x_k, x_i^j), \quad K(x_k, x_i^j) = \Phi(x_k)\Phi(x_i^j)$ is the kernel function. Now consider the numerator of Eq. (2), using the definition of S_b^Φ and Eq. (6) it can be rewritten as:

$$w^T S_b^\Phi w = a^T P a \tag{7}$$

where $P = (P_1 - P_2)(P_1 - P_2)^T$. Using Eq. (5), the definition of m_j^Φ, and a similar transformation as in Eq. (7), the denominator of Eq. (2) can be written as (Mika et al., 1999):

$$w^T S_w^\Phi w = a^T Q a \tag{8}$$

where $Q = \sum_{j=1}^{2} \sum_{i=1}^{N_j} (K_i^j - P_j)(K_i^j - P_j)^T$, $K_i^j = K(x_k, x_i^j)$, combining Eqs. (7) and (8), Fisher's linear discriminant in F is (Mika et

al., 1999):

$$J(a) = \frac{a^T P a}{a^T Q a} \tag{9}$$

Eq. (9) is the KFD function. When $J(a)$ reaches its maximum value, the eigenvector a_{opt} becomes the optimal mapping direction:

$$a_{opt} = arg\ \max(J(a)) = Q^{-1}(P_1 - P_2) \tag{10}$$

It is possible that Q in Eq. (10) returns a null value, resulting in an inability to solve the equation. The following modification can avoid this problem (Mika et al., 1999):

$$Q_\mu = Q + \mu I \tag{11}$$

where μ is a positive constant and I is a unit matrix. The projection of a test sample x onto the discriminant is computed by

$$G(x) = \sum_{i=1}^{N} a_{opt} K(x_i, x) \tag{12}$$

The decision function can then be provided as follows:

$$F(x) = \text{sign}(G(x) - b_0) \tag{13}$$

where $b = a_{opt}(P_1 + P_2)/2$. The decision rule can be represented by:

$$F(x) \begin{cases} > 0 \\ < 0 \end{cases} \rightarrow x \in \begin{cases} w_1(\text{class}) \\ w_2(\text{class}) \end{cases} \tag{14}$$

In this study we use KFD to derive transition rules in CA for simulating complex urban land use dynamics. The data of those spatial variables used in the CA modeling are from GIS and remote sensing. The general procedure of the kernel-based CA is illustrated in Fig. 3.

Determining the development probability using the transition rules is the most critical issue in urban CA. This probability, which is estimated based on a series of spatial variables (Batty and Xie, 1994; Wu and Webster, 1998; Li and Yeh, 2002), essentially represents the relationships between urban development and those variables. In this study, KFD is used to handle the highly complex relationships among those variables. The conventional KFD only generates Boolean boundaries, which are unsatisfactory for an application like urban simulation that is characterized by considerable uncertainty (Li and Yeh, 2004). Here we borrow a method developed by Wu (2002) to "soften" the clear-cut boundary using a logistic regression model. Many studies demonstrate that urban development fits the logistic pattern (Herbert and Thamas, 1990; Liu and Stuart, 2003). After incorporating the KFD function into a logistic model, the transition rules of CA can be written as follows:

$$P_{dev}(ij) = \frac{1}{a_0 + b_0\ \exp(-c_0(G(x_{ij}) - b))} \tag{15}$$

where a_0, b_0 and c_0 are the parameters of the logistic model. Corresponding to Eq. (12), $G(x_{ij})$ is the eigenvector value obtained on the optimal mapping direction a_{opt} for cell ij using the KFD method:

$$G(x_{ij}) = \sum_{h=1}^{N} a_{opt} K(x_h, x_{ij}) \tag{16}$$

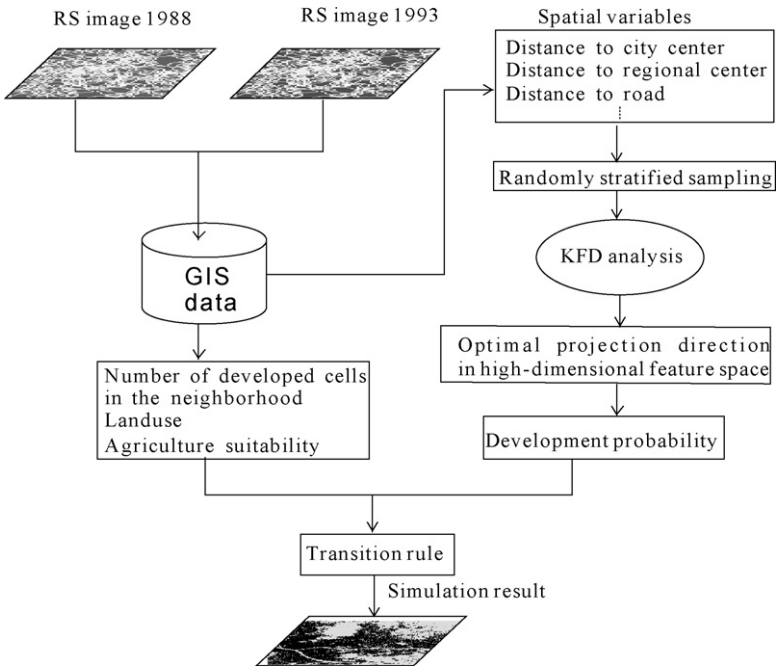


Fig. 3 – Flowchart of mining non-linear transition rules of CA using kernel-based method for urban simulation.

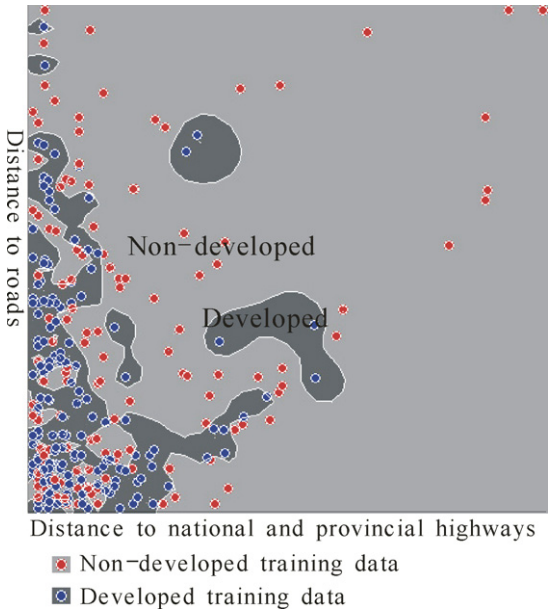


Fig. 4 – The kernel-based non-linear boundaries (transition rules) in a two-dimension space.

where $K(x_h,x_{ij})$ is a kernel function and x_h is the training data. The transition rules that were extracted by using this proposed method are non-linear, as shown in Eqs. (15) and (16). Fig. 4 illustrates the kernel-based, non-linear boundaries (transition rules) in a two-dimension space (defined by "distance to roads" and "distance to national and provincial highways"), and the output from the transition rules is the state of either "developed" or "non-developed".

Eq. (15) only accounts for proximity (distance) variables, but the development of a place is also influenced by its neighborhood situations. Therefore, in this study, the neighborhood influence is incorporated as an adjusting factor in the final estimation of the development probability. Specifically, it is assumed that the development probability increases when there are more developed cells in the neighborhood, and this neighborhood influence can be calculated as follows:

$$\Omega_{t-1}(ij) = \frac{\sum_{3\times3} N(\text{urban}(ij))}{3\times3-1} \quad (17)$$

where $\sum_{3\times3} N(\text{urban}(ij))$ is the total number of urbanized cells in a 3×3 window around the cell under test.

Furthermore, one expects to see low development probabilities in areas of rivers, steep slopes, and agricultural protection zones. These constraints can also be incorporated into the probability estimation (White et al., 1997; Li and Yeh, 2000).

The final development probability at time t is then estimated through integrating the classification using the KFD trained by empirical data, the incorporation of various constraints, and the inclusion of the neighborhood influence:

$$P_t(ij) = A * P_{dev}(ij) * \text{con}(\text{suit}(ij)) * \Omega_{t-1}(ij) \quad (18)$$

Table 1 – Spatial variables for simulating urban development using kernel-based CA model

Spatial distance variables						Local variables	Constrain condition		
Distance to cite proper	Distance to town centres	Distance to national and provincial highways	Distance to roads	Distance to railways	Distance to expressways	Number of developed cells in the neighborhood	Land elevation	Land use	Agriculture suitability

where A is a coefficient; con(suit(ij) is the total score of various constraints and varies between 0 and 1; and $\Omega_{t-1}(ij)$ is the neighborhood influence at time $t-1$.

In the CA simulation, when the development probability of a cell is greater than a threshold, the state of the cell will be converted to "developed". The high uncertainty in urban dynamics makes deterministically setting this threshold difficult. As a result, usually the threshold is decided in a Monte Carlo way. Specifically, the development probability of a cell is compared with a random number between 0 and 1 (Wu and Webster, 1998). If the development probability at a cell is greater than the random number, the state of the cell will be converted to "developed":

$$S_t(ij) = \begin{cases} \text{Developed}, & P_t(ij) > \gamma \quad \text{and} \quad \beta < 1/N \\ \text{Non-developed}, & P_t(ij) > \gamma \quad \text{and} \quad \beta < 1/N \\ \text{Non-developed}, & P_t(ij) \leq \gamma \end{cases} \tag{19}$$

where N is the number of iterations required for the simulation, γ and β are random variables ranging from 0 to 1.

3. Model implementation and simulation results

The kernel-based CA model was applied to the simulation of urban development of a fast-growing city, Guangzhou, in the Pearl River Delta of China. The actual urban areas in the years of 1988, 1993 and 2002 were identified from the TM satellite images. These empirical data were used to derive and calibrate transition rules. A series of spatial variables that are related to land development were prepared using GIS (Table 1 and Fig. 5). The distance variables were calculated using the *Eucdistance* function in ArcGIS. The number of developed cells in the 3×3 neighborhood was counted using the *Focal* function of ArcGIS. The agriculture suitability is calculated using the *raster calcu-*

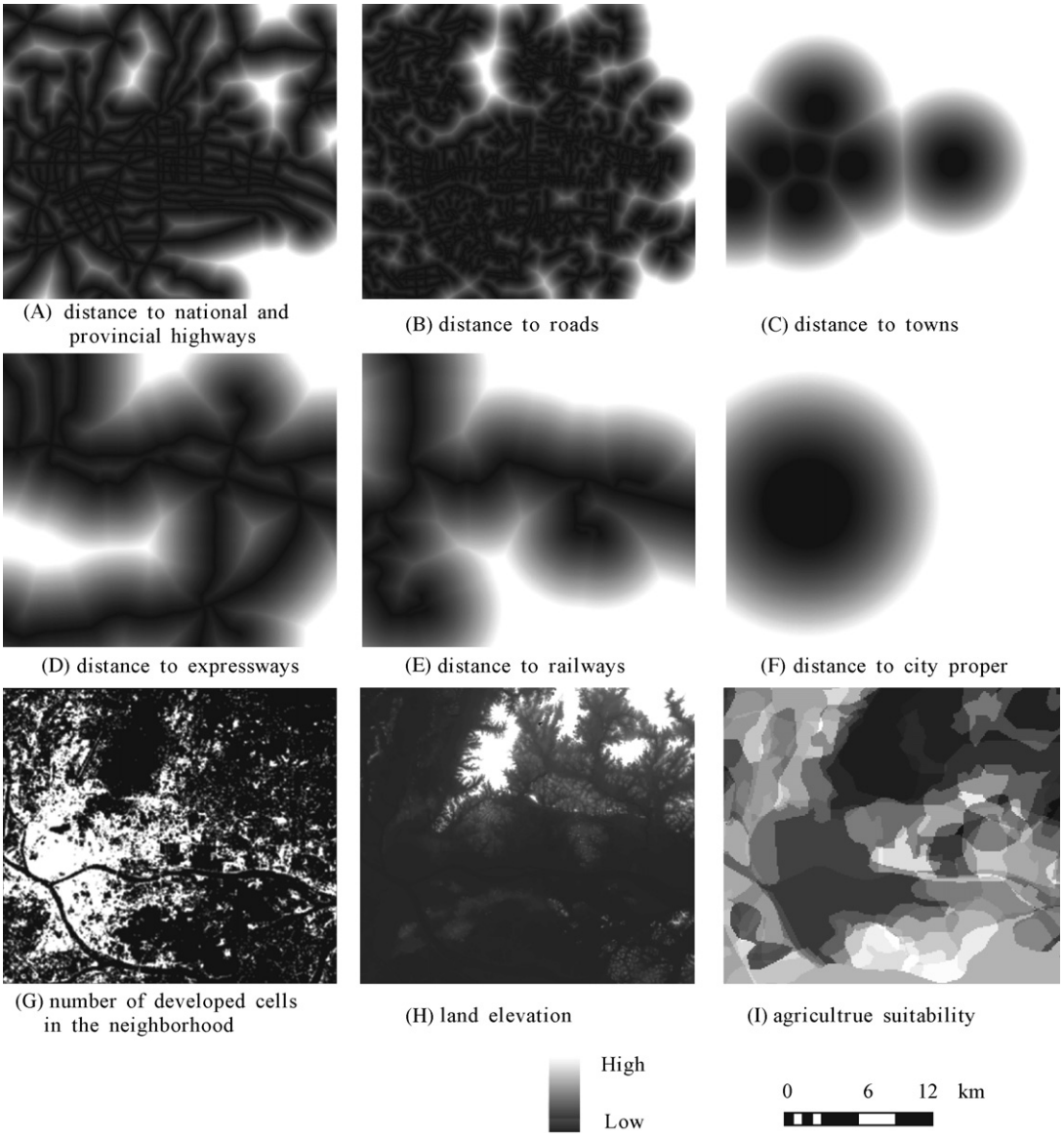


Fig. 5 – Various spatial variables prepared by a raster GIS.

Table 2 – Classification accuracies by using different kernel functions

Kernel function	Polynomial order (d)	Accuracy for training samples (%)	Accuracy for testing samples (%)
Polynomial	2	85.1	68.3
	3	87.0	69.5
	4	88.4	70.1
	Radial basis parameter (σ)	Accuracy for training samples (%)	Accuracy for testing samples (%)
Radial basis	0.5	100.0	58.4
	5	100.0	63.7
	10	99.1	67.2

lator function in ArcGIS. Land use data were acquired through classification of Landsat TM images.

The land use image for each year derived from the TM data was resampled to 50 m, which gives an image with size 396 × 450. While the image preserves the spatial details required for the analysis, its relatively high resolution may result in significant spatial autocorrelation and introduce considerable bias into the resulting rules. Moreover, directly working on such an image would be an overwhelmingly intensive computation process. Stratified random sampling on each original image was carried out to alleviate these problems (Li and Yeh, 2002). The final sample set contains data of 400 locations. These samples were divided into two equal groups, one as the training data to derive the transition rules and the other as the test data to verify the trained classifier.

The proposed model was implemented using Visual Basic 6.0, ArcObjects, and Matlab 7.1. ArcObjects provides access to spatial data, as well as tools for distance calculation and

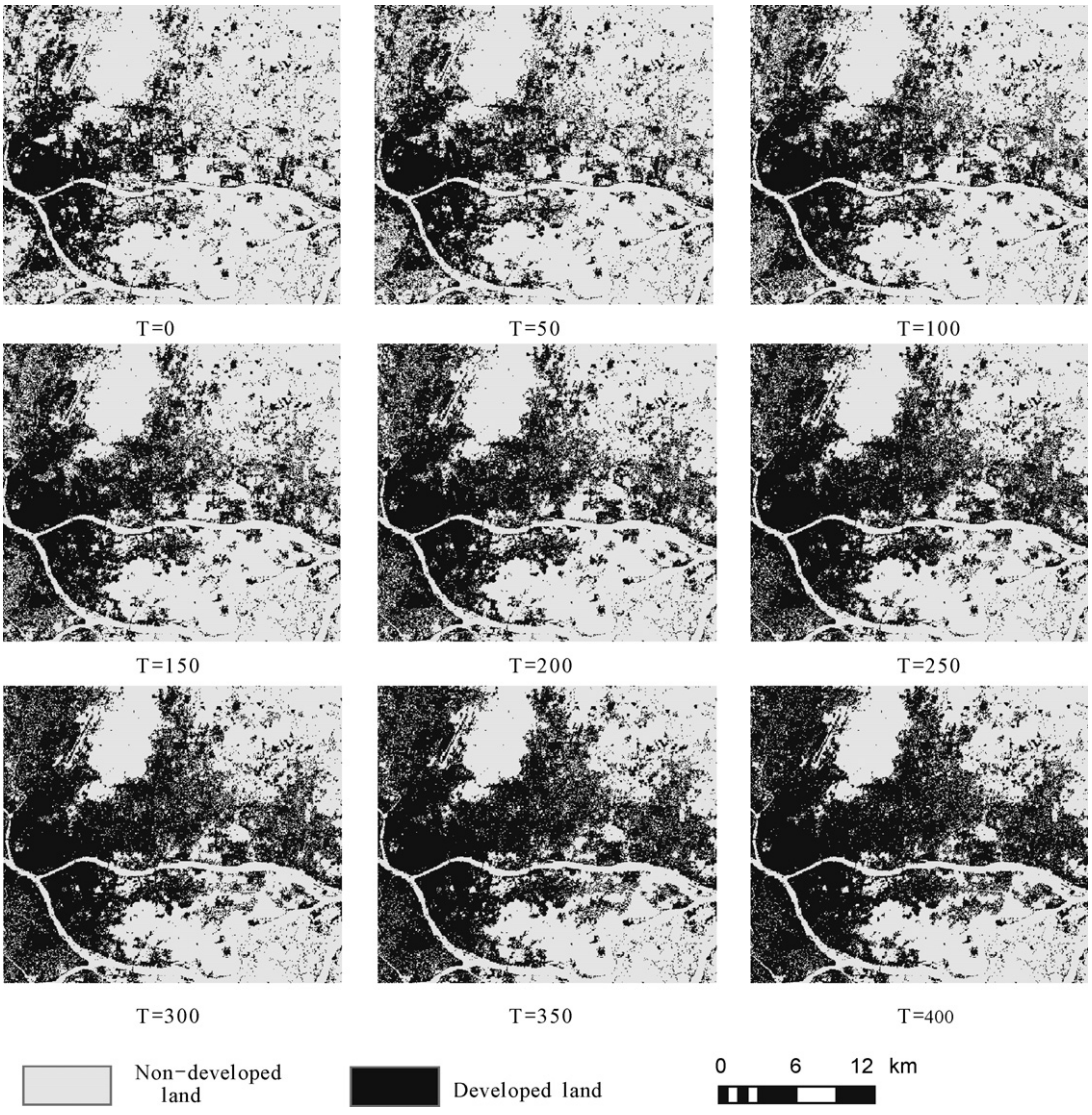


Fig. 6 – Simulation of urban dynamics of Guangzhou using kernel-based CA model.

focal operations. Matlab 7.1 was used to calculate the transition rules. Visual Basic was used to integrate the different components in the model.

Results of two different types of kernel functions: the polynomial kernel and the radial basis function (RBF) kernel were compared. Table 2 shows the comparison of these two functions on classifying the 200 samples in the test set. It is found that the RBF has a more serious over-fitting problem, which is the main reason for its poorer accuracy in classifying the test data. The table shows that the polynomial kernel is generally better than the RBF. According to the experiment, this kernel method yields the best results when $d = 4$. Therefore, $d = 4$ was used in the following estimation of the development probability.

The simulation was conducted in discrete temporal steps. A sufficient number of steps are required to reveal the effects of spatial interactions and produce details in the resulting spatial patterns. Although there is no consensus on the optimal number of steps, 100–200 iterations are common in practice (Wu, 2002; Li and Yeh, 2004). In this study, the urban development in 1993 and 2002 was obtained by running 200 iterations and 400 iterations respectively.

Fig. 6 shows that Guangzhou had an obvious urban expansion from 1988 ($T = 0$) to 2002 ($T = 400$). This urbanization process has altered the degree of fragmentation and structural complexity of the urban landscape. To quantify the changes in landscape structural complexity, we selected a set of metrics for measuring the entire landscape, including number of patches (NP), the largest patch index (LPI), edge density (ED), landscape shape index (LSI), and contagion (a measure of landscape configuration). These metrics were computed from the rasterized land-use maps using the Fragstats software package (McGarigal and Marks, 1995). These spatial pattern metrics capture ecologically relevant aspects of spatial pattern such as fragmentation (NP, LPI, ED, and contagion), patch shape (LSI), and amount of edges between contrasting patch types (ED and contagion) (Jenerette and Wu, 2001).

In terms of the landscape structure, from 1988 ($T = 0$) to 2002 ($T = 400$) the number of patches and the edge density increased, indicating an increase of fragmentation from 1988 to 2002 (Fig. 7A and C). Correspondingly, the largest patch index decreased (Fig. 7B). Urbanization also increased the complexity of patch shape (Fig. 7D). The contagion index tended to decrease, following the same trend of ED, which also indicates that fragmentation of landscape is increasing with urbanization (Fig. 7E).

4. Model validation

The assumption is that this model can be used to forecast the future land development if it is able to simulate the past trend quite well. Therefore, the simulated patterns for years 1993 and 2002 are validated by comparing them with the actual situations derived from the TM data. Following Clarke et al. (1997), White et al. (1997) and Ward et al. (2000), we conducted visual inspections and found that the simulated and actual patterns are very similar (Fig. 8).

Besides the visual inspection, quantitative comparisons were also performed to obtain more objective assessments. First, the comparison was conducted by a cell-on-cell overlay of two images. This method evaluates the coincidence of land use types at the exact locations between the simulated and the actual patterns (Li and Yeh, 2002). Table 3 shows the results of this comparison. The total accuracies are 79.0% and 74.1% for 1993 and 2002, respectively. The kappa coefficient, which better addresses the difference between the actual agreement and chance agreement (Fung and LeDrew, 1988; Congalton, 1991), are 57.0% and 48.2% for 1993 and 2002, respectively.

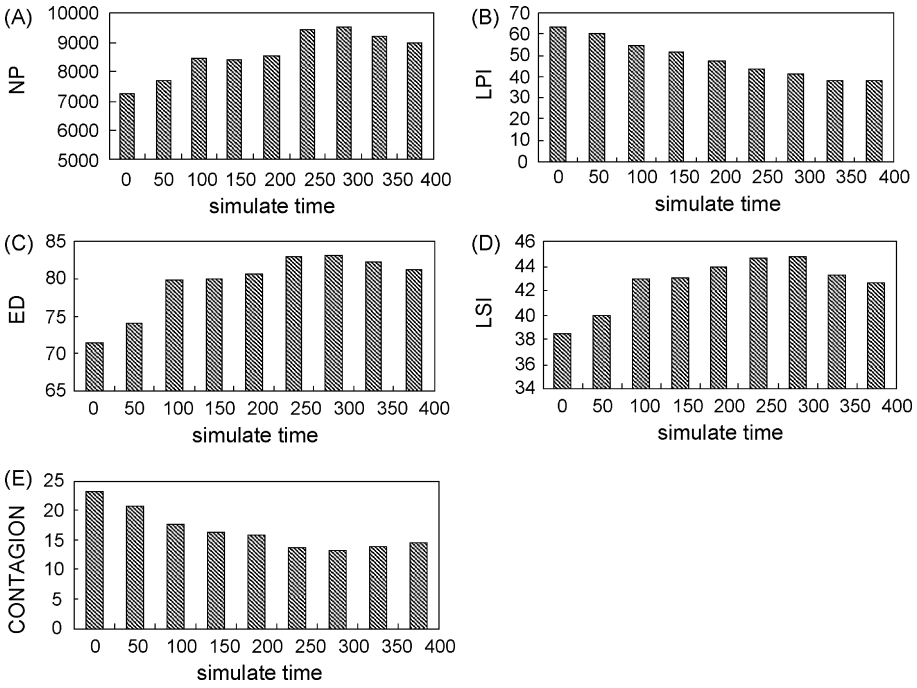


Fig. 7 – Landscape indices of simulated results by using kernel-based CA model.

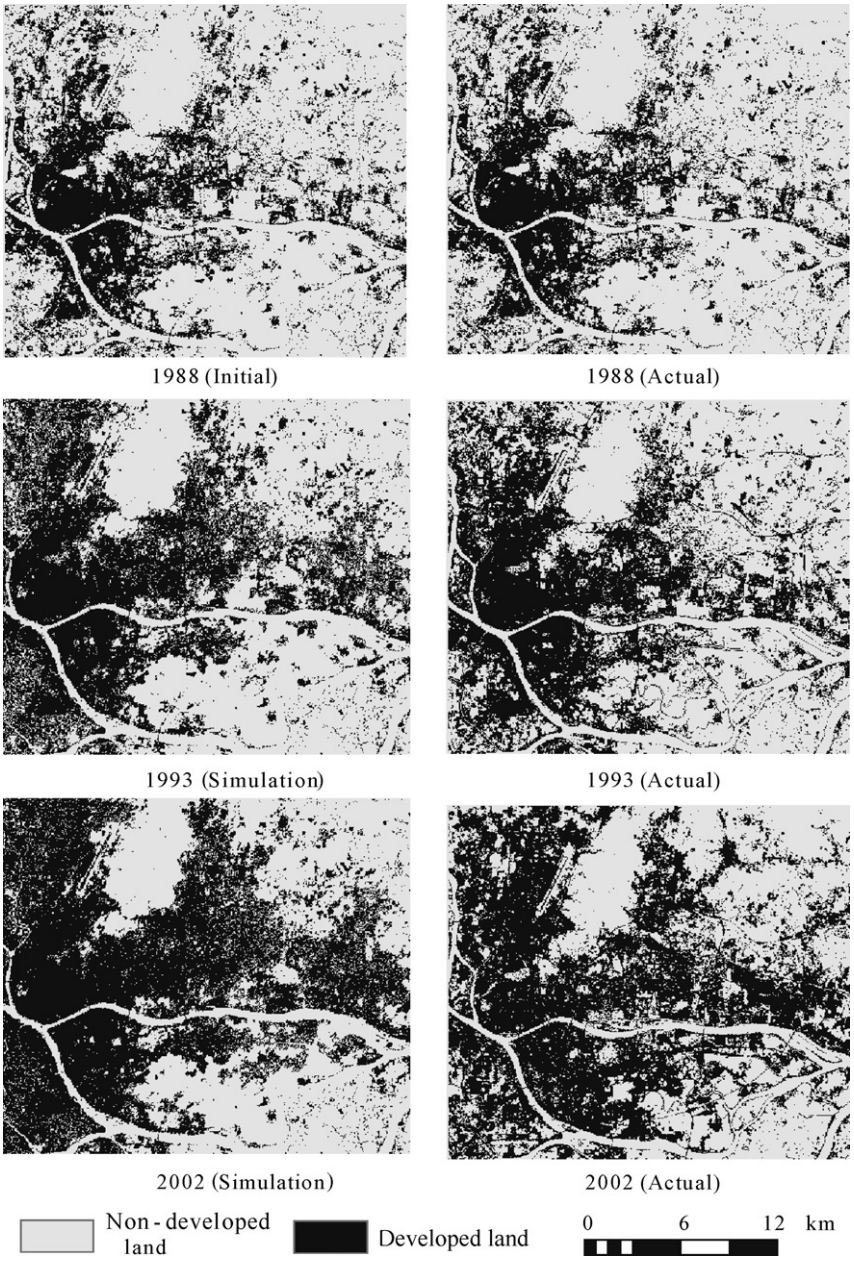


Fig. 8 – Simulated and actual development patterns of Guangzhou in 1988, 1993 and 2002.

Table 3 – Assessment of the goodness-of-fit for kernel-based CA model using the spatial overlay

	1988–1993 (cells)		
	Simulated non-urban	Simulated urban	Goodness-of-fit (%)
Actual non-urban	83653	21653	79.4
Actual urban	15834	57060	78.3
Total accuracy			79.0
Kappa coefficient			57.0
[0pt]	1993–2002 (cells)		
	Simulated non-urban	Simulated urban	Goodness-of-fit (%)
Actual non-urban	64595	24359	72.6
Actual urban	21961	67285	75.4
Total accuracy			74.1
Kappa coefficient			48.2

Table 4 – Landscape indices obtained for simulated patterns of kernel-based CA compared with actual patterns derived from TM data

	1988–1993				
	NP	LPI	ED	LSI	Contagion
Simulated	8526	47.49	80.63	43.83	15.92
Actual	8206	46.24	76.46	39.21	16.14
%Deviation	3.9	2.7	5.4	11.8	1.4
[0pt]	1993–2002				
	NP	LPI	ED	LSI	Contagion
Simulated	9001	38.26	81.17	42.59	14.58
Actual	8723	35.45	74.14	37.29	15.49
%Deviation	3.2	7.9	9.5	14.2	5.9

Table 5 – Assessment of the goodness-of-fit for neural networks-based CA model using the spatial overlay

	1988–1993 (cells)		
	Simulated non-urban	Simulated urban	Goodness-of-fit (%)
Actual non-urban	82144	23162	78.0
Actual urban	17521	55373	76.0
Total accuracy			77.2
Kappa coefficient			55.3
[0pt]	1993–2002 (cells)		
	Simulated non-urban	Simulated urban	Goodness-of-fit (%)
Actual non-urban	63732	25222	71.6
Actual urban	24349	64897	72.7
Total accuracy			72.2
Kappa coefficient			44.4

However, the overlay method cannot provide information about the morphology of the urban spatial structures, such as connectivity, fractals, and compactness. Landscape indices should be able to describe the characteristics of spatial patterns and provide useful insights about urban morphology. The validation is carried out by examining the differences between the simulated patterns and the actual ones (classified from remote sensing) on the landscape metrics (Table 4). The small differences on these metrics indicate a good conformity between the simulated and actual patterns in terms of landscape structure.

To establish a benchmark for evaluation of the proposed kernel-based CA model, we compared it with other CA models, the neural network and the logistic regression. The neural network CA (Li and Yeh, 2002; Aitkenhead et al., 2004) and the logistic-based CA (Soares-Filho et al., 2002; Wu, 2002) were applied to the same Guangzhou dataset.

Table 6 – Assessment of the goodness-of-fit for logistic-based CA model using the spatial overlay

	1988–1993 (cells)		
	Simulated non-urban	Simulated urban	Goodness-of-fit (%)
Actual non-urban	80325	24981	76.3
Actual urban	20834	52060	71.4
Total accuracy			74.3
Kappa coefficient			47.3
[0pt]	1993–2002 (cells)		
	Simulated non-urban	Simulated urban	Goodness-of-fit (%)
Actual non-urban	62472	26482	70.2
Actual urban	28098	61148	68.5
Total accuracy			69.4
Kappa coefficient			38.7

Table 7 – Landscape indices obtained for simulated patterns of neural networks CA compared with actual patterns derived from TM data

	1993				
	NP	LPI	ED	LSI	Contagion
Simulated	8583	48.13	81.04	42.64	15.67
Actual	8206	46.24	76.46	39.21	16.14
%Deviation	4.6	4.1	5.9	8.7	2.9
[0pt]	2002				
	NP	LPI	ED	LSI	Contagion
Simulated	9128	38.72	81.59	41.51	14.39
Actual	8723	35.45	74.14	37.29	15.49
%Deviation	4.6	9.2	10.0	11.3	7.1

Table 8 – Landscape indices obtained for simulated patterns of logistic-based ca compared with actual patterns derived from TM data

	1993				
	NP	LPI	ED	LSI	Contagion
Simulated	8534	48.74	81.58	42.92	15.41
Actual	8206	46.24	76.46	39.21	16.14
%Deviation	4.0	5.4	6.7	9.5	4.5
[0pt]	2002				
	NP	LPI	ED	LSI	Contagion
Simulated	9037	39.05	82.78	42.19	14.12
Actual	8723	35.45	74.14	37.29	15.49
%Deviation	3.5	10.1	11.7	13.3	8.8

Tables 5 and 6 show the accuracy evaluation using the cell-on-cell overlay for the neural network CA and the logistic-based CA. A comparison of Tables 3, 5 and 6 reveals that the proposed kernel-based CA model is slightly more accurate than the neural-network and the logistic regression. Tables 7 and 8 show the validation using landscape metrics for the neural-network CA and the logistic-based CA respectively. A comparison of Tables 4, 7 and 8 indicates that the kernel-based CA also performs better than the neural-network CA and the logistic-based CA in terms of structural conformity.

5. Conclusion

The simulation of urban systems using cellular automata (CA) may involve a large set of spatial variables, as well as their complex relationships. The traditional linear methods are inherently insufficient for capturing complex relationships in the simulation of complex urban dynamics. This paper presents a new method to derive non-linear transition rules for CA simulation of urban land use change. The method is based on the kernel Fisher discriminant (KFD), which projects the input vectors to a high dimensional feature space for reducing complex non-linear problems into simple linear problems.

The proposed method has advantages in dealing with complex relationships among a large number of spatial variables in urban land use dynamics, which can reflect the complexity of geographical phenomena much better than traditional linear approaches. Unlike neural networks, KFD is relatively easy to analyze mathematically, since it is based on simple dot product calculation and its final classification process is eventually linear. Based on the principles of structural risk minimization (SRM), KFD minimizes the upper bound of the expected generalization error, which leads to a global optimization. Furthermore, KFD allow us to interpret learning algorithms geometrically in the kernel space, thus combining statistics and geometry in an effective way.

The model has been successfully applied to the simulation of a fast growing region in southern China (accuracy = 79.0%). The proposed model has been validated using spatial overlay and landscape indices to evaluate the goodness-of-fit between the simulated patterns and the actual ones. It is found that there is a good conformity between the simulated and actual land development. Moreover, this proposed model has some improvement of the accuracy over the neural network method and the logistic regression. The KFD CA itself is generic and should be applicable to other regions, such as Europe, USA, or other parts of Asia. The variables included in the simulation, however, can be regional and should be considered by the user when applying the model to other regions.

Acknowledgements

This study was supported by the National Outstanding Youth Foundation of China (Grant No. 40525002), the National Natural Science Foundation of China (Grant No. 40471105), and the Hi-tech Research and Development Program of China (863 Program) (Grant No. 2006AA12Z206).

REFERENCES

Abdallah, F., Richard, C., Lengelle, R., 2004. An improved training algorithm for nonlinear kernel discriminants. IEEE T. Signal. Process. 52, 2798–2806.

Aitkenhead, M.J., Mustard, M.J., McDonald, A.J.S., 2004. Using neural networks to predict spatial structure in ecological systems. Ecol. Model. 179, 393–403.

Aronszajn, N., 1950. Theory of reproducing kernels. Trans. Am. Math. Soc. 68, 337–404.

Baudat, G., Anouar, F., 2000. Generalized discriminant analysis using a kernel approach. Neural Comput. 12, 2385–2404.

Bousquet, O., Perez-Cruz, F., 2003. Kernel methods and their application to signal processing. IEEE ICASS, P860–P863.

Batty, M., Xie, Y.C., 1994. From cells to cities. Environ. Plan. B 21, 531–548.

Batty, M., Xie, Y., 1997. Possible urban automata. Environ. Plan. B 24, 175–192.

Cawley, G.C., Talbot, L.C., 2003. Efficient leave-one-out cross-validation of kernel Fisher discriminant classifiers. Pattern Recogn. 36, 2585–2592.

Cho, J., Lee, D., Park, J., et al., 2004. Gene selection and classifications from microarray data using kernel machine. FEBS Lett. 571, 93–98.

Clarke, K.C., Brass, J.A., Riggan, P.J., 1994. A cellular automata model of wildfire propagation and extinction. Photogramm. Eng. Rem. S 60, 1355–1367.

Clarke, K.C., Hoppen, S., Gaydos, L., 1997. A self-modifying cellular automaton model of historical urbanization in the San Francisco Bay area. Environ. Plan. B: Plan. Des. 24, 247–261.

Clarke, K.C., Gaydos, L.J., 1998. Loose-coupling a cellular automaton model and GIS: long-term urban growth prediction for San Francisco and Washington/Baltimore. Int. J. Geogr. Inf. Sci. 12, 699–714.

Congalton, R.G., 1991. A review of assessing the accuracy of classification of remotely sensed data. Remote Sens. Environ. 37, 35–46.

Couclelis, H., 1985. Cellular worlds: a framework for modeling micro–macro dynamics. Environ. Plan. A 17, 585–596.

Couclelis, H., 1988. Of mice and men: what rodent populations can teach us about complex spatial dynamics. Environ. Plan. A 20, 99–109.

Couclelis, H., 1997. From cellular automata to urban models: New principles for model development and implementation. Environ. Plan. B: Plan. Des. 24, 165–174.

Feng, Y., Shi, P., 2004. Face detection based on kernel fisher discriminant analysis. In: Proceedings of the Sixth IEEE International Conference on Automatic Face and Gesture Recognition.

Fung, T., LeDrew, E., 1988. The determination of optimal threshold levels for change detection using various accuracy indices. Photogramm. Eng. Rem. S 54, 1449–1454.

Hargrove, W.W., Gardner, R.H., Turner, M.G., Romme, W.H., Despain, D.G., 2000. Simulating fire patterns in heterogeneous landscapes. Ecol. Model. 135, 243–263.

Herbert, D.T., Thamas, G.J., 1990. Cites in Space. City as Place. David Fulton Publishers, London.

Huang, Z., Chen, H., Hsu, C., et al., 2004. Credit rating analysis with support vector machines and neural networks: a market comparative study. Decision Support Syst. 37, 543– 558.

Jenerette, G.D., Wu, J.G., 2001. Analysis and simulation of land-use change in the central Arizona – Phoenix region, USA. Landsc. Ecol. 16, 611–626.

Karafyllidis, I., Thanailakis, A., 1997. A model for predicting forest fire spreading using cellular automata. Ecol. Model. 99, 87– 89.

Klaus-Robert, M., Sebastian, M., Gunnar, R., Koji, T., Bernhard, S., 2001. An introduction to kernel-based learning algorithms. IEEE T. Neural Networks 12, 181–201.

Lett, C., Silber, C., Barret, N., 1999. Comparison of a cellular automata network and an individual-based model for the simulation of forest dynamics. Ecol. Model. 121, 277–293.

Li, X., Yeh, A.G.O., 2000. Modelling sustainable urban development by the integration of constrained cellular automata and GIS. Int. J. Geogr. Inf. Sci. 14, 131–152.

Li, X., Yeh, A.G.O., 2002. Neural-network-based cellular automata for simulating multiple land use changes using GIS. GIS Int. J. Geogr. Inf. Sci. 16, 323–343.

Li, X., Yeh, A.G.O., 2004. Data mining of cellular automata's transition rules. GIS Int. J. Geogr. Inf. Sci. 18, 723–744.

Liang, Z., Shi, P., 2004. An efficient and effective method to solve kernel Fisher discriminant analysis. Neurocomputing 61, 485–493.

Liu, Q.S., Hanqing, L., Songde, M., 2004. Improving kernel fisher discriminant analysis for face recognition. IEEE T. Circ. Syst. Vid. 14, 42–49.

Liu, Y., Stuart, R., 2003. Modeling urban development with cellular automata incorporating fuzzy-set approaches. Comput. Environ. Urban Syst. 27, 637–658.

McGarigal, K., Marks, B., 1995. Fragstats: spatial pattern analysis program for quantifyinglandscape structure. USDA Forest Service—General Technical Report PNW-GTR-351.

Mika, S., Ratsch, G., Weston, J., 1999. Fisher Discriminant Analysis with Kernels. Neural Networks for Signal Processing IX. IEEE Press, New York, pp. 4–48.

Scholkopf, B., Sung, K., Burges, C.J., et al., 1997. Comparing support vector machines with gaussian kernels to radial basis function classifiers. IEEE T. Signal. Process. 45, 2758–2765.

Scholkopf, B., Smola, A., Muller, K.R., 1998. Nonlinear component analysis as a kernel eigenvalue problem. Neural Comput. 10, 1299–1319.

Scholkopf, B., Mika, S., Burges, et al., 1999. Input space versus feature space in kernel-based methods. IEEE T. Neural Network 10, 1000–1017.

Sirakoulis, G.C., Karafyllidis, I., Thanailakis, A., 2000. A cellular automaton model for the effects of population movement and vaccination on epidemic propagation. Ecol. Model. 133, 209–223.

Soares-Filho, B.S., Cerqueira, G.C., Pennachin, C.L., 2002. DINAMICA—a stochastic cellular automata model designed to simulate the landscape dynamics in an Amazonian colonization frontier. Ecol. Model. 154, 217–235.

Vapnik, V., 1998. Statistical Learning Theory. Wiley, New York.

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

Ward, D.P., Murray, A.T., Phinn, S.R., 2000. A stochastically constrained cellular automata model of urban growth. Comput. Environ. Urban Syst. 24, 539–558.

White, R., Engelen, G., 1993. Cellular automata and fractal urban form: a cellular modeling approach to the evolution of urban land-use patterns. Environ. Plan. A 25, 1175–1199.

White, R., Engelen, G., Uijee, I., 1997. The use of constrained cellular automata for high-resolution modelling of urban land use dynamics. Environ. Plan. B: Plan. Des. 24, 323–343.

Wu, F., Webster, C.J., 1998. Simulation of land development through the integration of cellular automata and multicriteria evaluation. Environ. Plan. B: Plan. Des. 25, 103–126.

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, 63–82.

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

Zhang, J., Huang, Z., 2005. Kernel fisher discriminant analysis for bearing fault diagnosis. In: Proceedings of the Fourth International Conference on Machine Learning and Cybernetics, Guangzhou, pp. 3216–3220.