Taylor & Francis
Taylor & Francis Group

Research Article

Calibration of stochastic cellular automata: the application to rural-urban land conversions

FULONG WU

Department of Geography, University of Southampton, Southampton, SO17 1BJ, England, UK. e-mail: F.Wu@soton.ac.uk

Abstract. Despite the recognition of cellular automata (CA) as a flexible and powerful tool for urban growth simulation, the calibration of CA had been largely heuristic until recent efforts to incorporate multi-criteria evaluation and artificial neural network into rule definition. This study developed a stochastic CA model, which derives its initial probability of simulation from observed sequential land use data. Furthermore, this initial probability is updated dynamically through local rules based on the strength of neighbourhood development. Consequentially the integration of global (static) and local (dynamic) factors produces more realistic simulation results. The procedure of calibrated CA can be applied in other contexts with minimum modification. In this study we applied the procedure to simulate rural-urban land conversions in the city of Guangzhou, China. Moreover, the study suggests the need to examine the result of CA through spatial, tabular and structural validation.

1. Introduction

The definition of the CA rule remains a research issue, despite the emergence of CA as a powerful visualisation tool in urban growth simulation (Batty 1998). As an iterative computational procedure, CA characterise the change in geographical space as state changes and simulate the state changes through a neighbourhood that interconnects immediately neighbouring cells. Despite many appealing features of the CA approach, such as its abilities to simulate bottom-up dynamics and to capture self-organising processes, urban CA are developed largely through trial and error. The models are essentially *heuristic*. Various CA models have been successfully developed, especially in land use simulation (Batty and Xie 1994, Batty *et al.* 1999, Clarke *et al.* 1997, Clarke and Gaydos 1998, Li and Yeh 2000, White and Engelen 1993, White *et al.* 1997, Wu and Webster 1998, Wu 1998b). Nevertheless, calibration and validation of CA models had been two neglected issues until recent efforts to develop CA as a reliable procedure for the application of urban development simulation. Batty *et al.* (1999), for example, formalise the CA simulation by embedding CA rules in simulation software. The conceptual basis of the software is Xie's model of urban spatial extension (Xie 1996). Wu and Webster (1998) use multicriteria evaluation (MCE) to define the parameter values of CA models. Clarke *et al.* (1997)

validate the simulation results through visual tests. Clarke and Gaydos (1998) develop a more elaborated approach, which uses intensive computation to find the best fits out of numerous combinations of parameter values. Recently, Li and Yeh (2001) used neural networks (NN) to determine parameter values for CA simulation. Specifically the parameter values from the training of NN are imported into the CA models. The difficulty in finding the parameter value of CA simulation is partially due to the complexity of urban development (Batty *et al.* 1999). However, calibration and validation are two critical issues to be fully researched to develop CA as a reliable procedure for urban growth simulation.

Because CA are transparent, flexible and open procedures, model builders can define state transition rules that sound plausible. In essence the rule definition relies on an intuitive understanding of the process of cell state change, though some relationships between variables and cell states can be found through empirical studies. The number of transition rules is virtually unlimited. While it is interesting to see that simple rules, derived from uncoordinated local decision-making processes, can give rise to a structured global pattern, it is often difficult to identify such a rule among millions of alternative ones. In the *Game of Life,* originally developed by John Conway (see Gardner 1970, for an initial report and Batty 1997, for its relevance in urban simulation), if one changes the rules, the interesting pattern would disappear. Indeed it is notably difficult to propose a rule that can lead to a desirable urban form. For complicated urban development processes, this is particularly problematic, as there is no standard procedure to specify a rule. Various rules have been defined under the umbrella of CA (White and Engelen 1993, Batty and Xie 1994, Portugali and Benenson, 1995, Batty 1998, Clarke and Gaydos, 1998, Wu and Webster 1998, Li and Yeh 2000). Unlike a standard procedure used in multivariate regression, the procedure of rule definition in CA models is highly flexible and basically applies various microscopic rules to manipulate cell states in grid-based data.

The heuristic approach of urban CA is not wrong in itself. In fact, it is because the process of urban land development is so complicated and ill-defined that it is impossible to propose a universal law that would control the process in different places. This is in sharp contrast with many physical space-time dynamics such as the dispersion of plants, the predator-prey process, and surface runoff in catchments (Burrough 1998). However, the complexity of the built environment essentially requires a simulation approach because, without actually building a model, it is hard to test hypotheses about the complex behaviour of urban land development. The heuristic model is confronted with a severe computational constraint. For example, Clarke *et al.* (1997) applied various visual comparison methods to find the parameter values. In an automatic parameter calibration, Clarke and Gaydos (1998) use several hundred hours of high- performance workstation to search for the best combination of parameter values. Therefore it would be useful to use a procedure that can extract the initial values of some parameters. In a sense, the calibration of an initial probability surface for stochastic CA is important.

Classical formalisation of CA is proposed by Wolfram (1984). In the domain of geographical space, Takeyama and Couclelis (1997) proposed a geo-processing language to formalise CA simulation. Similar efforts have been made through the conceptualisation of the urban diffusion process (Batty and Longley 1994, Batty and Xie 1994), the generic simulation model (Xie 1996) and operational CA application software (Batty *et al.* 1999). Formalisation helps to develop a framework to address the diversity of rule definition. Moreover, it is useful to develop a calibrated method

to reduce the complexity of the model, as this will allow model builders to concentrate on characterising the process itself rather than constructing model structures.

At the core of urban CA is an understanding that urban development is neither the purely local process used by classical CA, *The Game of Life*, nor is it a purely global process modelled in classical urban models like the Lowry model. Couclelis (1997) elaborated a conceptual framework that extends universally applied local interaction to complex interaction across geographical space. The extended CA framework is more appropriate to modelling geographical phenomena, as 'geography' means exactly the heterogeneous space. The integration of simple abstraction of CA with heterogeneous geographical space requires calibration, which determines the parameter values from observed processes of state change.

The design of an appropriate simulation strategy of land development should consider that urban growth could be best articulated through the combination of global and local factors. This can be achieved through parameterisation of factors that affect the development. The advantage of parameterisation is that the balance of global and local processes can be explicitly addressed. The parameterisation of the urban cellular model, however, imposes an enormous challenge to the identification of appropriate parameters. So far, calibration has been achieved through intensive computation—for example, repetitive running of the same model with different combinations of parameter values (Clarke and Gaydos, 1998)—and through automatic training by NN (Li and Yeh, 2001). As Clarke and Gaydos (1998) demonstrate in their San Francisco and Washington/Baltimore model applications, the calibration is computationally intensive, which requires high performance computation facilities. Li and Yeh's (2001) NN-CA approach is interesting in that it can automatically retrieve the parameter values. However, the meaning of the parameter values might be difficult to interpret, which is not a problem of CAs but rather a feature of NNs.

Validation of the CA model is still a challenge to CA applications. Most of the CA models to date have used visual comparison to confirm the simulation results. The measure of model performance itself is a controversial issue. Because of the property of emergence in complex self-organising systems, CA models should be assessed on the basis of plausibility (Batty 1996) rather than one-to-one correspondence or correlation measures. This requires that the model should be validated in terms of whether the model can capture the basic features of urban land use, for example the structural similarity between simulated and actual land development. Clarke and Gaydos (1998) used four statistical measures to assess the model performance. These included a series of *r*-squared fits between actual and predicted development and between urban edges, and a modified Lee-Sallee shape index. The validation of the results of numerous combinations of parameters is, therefore, very time consuming, as all these performance indicators need to be calculated in each of the combinations. Li and Yeh (2001) use a conversion matrix to assess the accuracy of simulation, while Wu (1998a) explores the coefficients of the density function, Moran Index and fractal in model validation. However, considering that a new structure may occur in a CA simulation, the validation is dependent upon the purpose of the simulation. In other words, the measure of performance is related to the specific aspect that we wish to simulate. In the rest of the paper the application of calibrated stochastic CA in rural-urban land conversions is discussed.

2. Rural-urban land conversions

Rural-urban land conversions are by no means randomly distributed. The general characteristics of land use conversions can be revealed through a series of

development profiles such as the plotting of the quantity of land use conversion against distance from the city centre. Urban land use models clearly show that land development is constrained by location and geographical conditions. More precisely, urban economics establishes development propensity/probability through regression methods. Logistic regression or the multinomial logit model, for example, can be used to examine the relationship between land use changes and their locational characteristics (McMillen 1989).

However, the regression method is essentially static. While it reflects the global distribution of land use conversions in the metropolitan area, the method does not reveal the self-organising nature of land development, i.e. the clustering of land uses at a local scale or the level of development sites. Urban land development consists of two interrelated processes—that of spontaneous growth and that of self-organised growth. The former represents a process that is independent of sequential land use changes. Land conversions take place according to the demand and supply relationship through development propensity. The latter reflects a process that results from the previous development in the immediate neighbourhood. The chance of land conversion at any place is raised through the clustering of land development in the neighbourhood. In a CA simulation, the two processes should be appropriately addressed.

The relationship between development factors and urban spatial structure is conceptualised by the bid-rent theory in urban economics. The main determinant of urban land use change, according to the monocentric bid-rent theory, is the distance to the city centre (Alonso 1964). Along with the increase in the distance to the city centre, accessibility decreases and transport cost increases. Different land users have different utility functions, making trade-offs between land rent and transport cost. Their 'willingness to pay', namely the land bid rent, differs. In a fully functional land market, the highest bid will obtain the land. This is, of course, a simplified but elegant theoretical deduction. The actual land development propensity is more complicated. In a complex geographical context, land development is affected by various attributes, ranging from the physical characteristics of development sites to planning control and zoning. Practically, land development can be modelled through a discrete choice framework (Ben-Akiva and Lerman 1985). McMillen (1989) uses a multinomial logit model in land use changes. Land development thus can be imagined as particular land users choosing particular land plots. The joint probability of a particular type of development occurring at a particular site can be estimated through regressions. This probability should be estimated in different contexts.

The urban land model provides a clue for land development simulation. Rather than simulate land use changes based on heuristically plausible land development propensities, it is possible to derive or calibrate land development probability from the observation of land use changes. This means the rule of the simulation can be stated simply as a function of development variables, i.e. development probability = f (development factor 1, development factor 2, ...). The theory of discrete choice suggests that the specific form is dependent upon the statistical distribution of an error item and that in the context of a discrete choice it should be in the form of the logit model/multinomial logit model. In other words, the purpose of calibration is to establish the relationship between land use change and the factors that affect probability of land conversion. The regression can be seen as a process to extract the coefficients of the empirical relationships from observations, which is a critical step towards the development of more procedural and realistic urban CA simulation.

3. Calibration of development probability

The purpose of calibration is to extract the coefficients or parameter values of the rules from the observation of land use pattern at time t and $t+1$ (figure 1). Mathematically, this is generalised as the estimation of the probability of particular state transition y occurring at a particular location (i, j) through a function of development factors $(x_1, x_2, ...x_n)$. In the case of binary land use changes (being developed into urban use or remaining in the current state), a logistic model can be developed to calculate the probability of development. Specifically, the model assumes that the attractiveness of a site is a function of the independent variables such as the travel distance to the city, land elevation and slope. The dependent variable is a binary (categorical) one, namely whether the land has been developed or not in the observed period. Note that in this case both dependent and independent variables are grids derived in a GIS environment and subsequently exported to fit a regression model. According to the logistic model, the probability of a site experiencing land conversion can be computed as:

$$p_g(s_{ij} = urban) = \frac{\exp(\mathbf{z})}{1 + \exp(\mathbf{z})} = \frac{1}{1 + \exp(-\mathbf{z})} \tag{1}$$

Where, p_g is the observed *global* probability, s_{ij} is the state of the cell ij, $\mathbf{z}$ is a vector that describes the development features of the site:

$$\mathbf{z} = a + \sum_k b_k x_k$$

where a is a constant, b_k are coefficients of the regression model; x_k is a set of site attributes.

It must be noted that the probability is estimated globally and does not change according to simulation and local situations, it is thus without time denomination (although the probability is estimated from sequential data). The probability is estimated from the comparison of land use changes in a time period longer than one
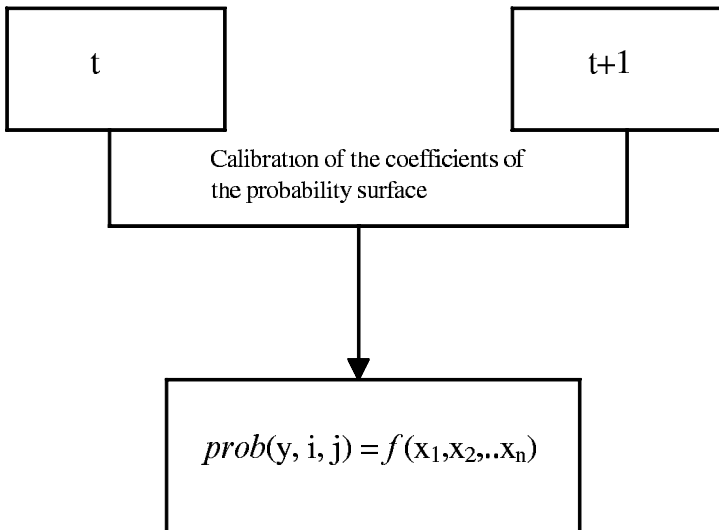

Figure 1. Calibration of the global probability surface from sequential land use data.

iteration used in CA. The direct application of the probability to cell state transition will be problematic, as it ignores the interlocking effect of land development. This is particularly important in microscopic simulation, as grid-based simulation differs from the zonal system in that the latter is based on aggregation of data. This difference can be seen in figure 2. In the zonal system, land development is represented as an aggregated 'density' indicator, while in a grid system, land development is defined as discrete conversions of individual cells. Both grids shown in the figure conform to the density requirement in the zonal system. But the left grid presents a totally unrealistic urban morphology, as studies of urban morphology show that land development follows a particular space-filling property, namely fractal (Batty and Longley 1994). It is more likely that the right grid shown in figure 2 below represents a more realistic morphology.

While it can be justified that the purpose of stochastic simulation is not to reproduce the exact pattern of urban structure, the urban morphology simulated from global probability by this random process is unrealistic (see the results and discussion later). The connectivity of developed sites and undeveloped areas is much lower in the simulation based on global probability. The problem of applying global statistical properties to a microscopic level is that this ignores the *path-dependent* and *self-organising* nature of land development. In other words, land development not only follows the static attractiveness measured as a geographic location and other physical attributes but is also conditioned by the sequence of development
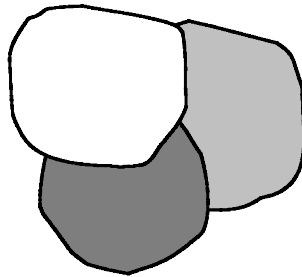
(a)
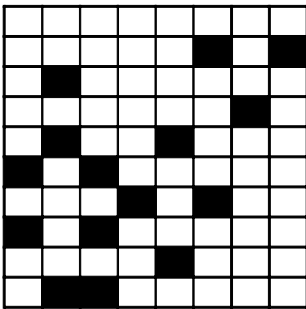
(b)
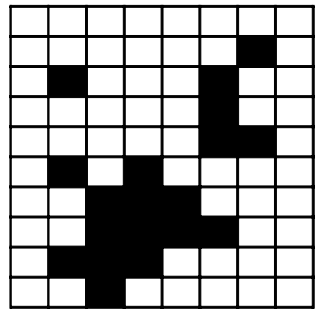
(c)

Figure 2. The zonal-based and grid-based spatial analysis units measured in density and morphology. (*a*) Zonal system measured in density; (*b*) unrealistic urban morphology in grid-based system; (*c*) more realistic urban morphology.

and the neighbourhood effect. This suggests that a local probability of development should be incorporated into the measure of development attractiveness. Theoretically it would be possible to include neighbourhood measures directly in the regression to estimate local probability if the land use conversion were observed on a fine time scale. However, this often proves to be difficult, as such data are often unavailable. Moreover, it is even more difficult to match the time scale of observation and simulation. In this application, the time lag between two land use coverages is 20 years while simulation proceeds on a year by year basis. To match the time scale of simulation with the time scale of development, we use the total quantity of land conversions. That is, we constrain the quantity of conversion at each iteration to the quantity of land conversion in one year.

In contrast to macroscopic land modelling, classical CA is a purely microscopic approach. The notion of microscopic characterisation in essence requires that the CA model be built upon the basic unit of behaviour. This is plausible because migration is the behaviour of a household rather than a census tract. Similarly, land development is the behaviour of a developer rather than a ward or district. The CA approach reasonably captures the connectivity between development sites/developers. However, the approach is mainly derived from natural phenomena (e.g. gas in physics, forest fire in biology) where the local effect may be strong. In land development, the influence of global development conditions is obvious and the geography of land development is *uneven*. This heterogeneous nature of develop- ment conditions is reflected in the global probability derived by land-use modelling. It is therefore plausible to compose a joint probability of land development from global and local conditions. In this study, the neighbourhood function is calculated in a conventional *ad hoc* way, i.e. through a 3×3 ($90 \, \text{m} \times 90 \, \text{m}$) kernel. The neighbourhood potentiality of cell transition is defined as:

$$\Omega_{ij}^t = \frac{\sum_{3 \times 3} con \; (s_{ij} = urban)}{3 \times 3 - 1} \tag{2}$$

where, Ω_{ij} is a neighbourhood evaluation function, here referring to the development density within the 3×3 neighbourhood, $con()$ is a conditional function which returns true if the state s_{ij} is urban land use. In this simulation, the neighbourhood is defined as 8 immediately neighbouring cells. It must be noted that Ω is denominated by the time t, which means that neighbourhood density changes along with the simulation.

The joint probability can be calculated as the product of global probability, cell constraint, and neighbourhood potentiality. Cell constraint refers to factors which exclude land development on the cells such as a body of water, a mountainous area and planning restriction zones. It is possible to use an evaluation score of land suitability instead of a binary one (suitable/unsuitable). The joint probability is stated as:

$$p_c^t = p_g \, con(s_{ij}^t = suitable)\Omega_{ij}^t \tag{3}$$

where $con( )$ converts the state of suitable land into a binary variable. Again, please note that the joint probability p_c is denoted with time t, indicating it changes along with iterations.

In sum, this study derives the initial/global probability surface of land develop- ment from calibrated logistic regression and constantly updates this probability surface using local conditions of development in a kernel throughout the simulation

iteration. This combination of global and local factors will generate better simulation results, as will be shown later.

3. Stochastic CA simulation

Based on the joint probability, a Monte Carlo process is launched to generate the result of simulation. However, updating the global probability with local conditions creates additional complications. Whereas the global probability, derived from land use data, would produce the quantity of land conversions which conform to the observed one, the joint probability may not be guaranteed to do so, because the constraining of global probability with the local one has changed the property of distribution. Intuitively, this means that the restriction of land use may eventually reduce the total land conversions. In contrast, the introduction of local agglomeration effect may also create too many potential sites. The problem is, therefore, to constrain the amount of cell transition according to land demand or projected land demand. This means that computationally the pseudo joint probability needs to be transformed into one that will produce the required quantity of state transitions.

As mentioned earlier, the joint probability is a dynamic one. Thus, the transformation requires the comparison of the probability at each iteration with the best site at that *particular* iteration. In essence, this means that the best site available at the particular time, rather than the ideal site, should be used as a benchmark. The form proposed in this simulation is an exponential distance-decay function, a non-linear one as the sites with higher evaluation scores are more likely to be developed. The probability of site conversion is transformed by comparing its value with the probability of the best site.

$$p_i^t = p_c^t \exp[-\delta(1 - p_c^t/\max(p_c^t))] \qquad (4)$$

where, max( ) returns the maximum value of p_c^t from the whole grid, δ is a dispersion parameter. The higher is the value of δ, the steeper the distance-decay gradient. The range of the dispersion parameter δ is usually from 1 to 10 (it has no units as it serves as a coefficient in the distance-decay function). The value 5 usually generates a quite stringent distance constraint. The dispersion parameter controls the shape of the skewed probability curve. The items inside the exponential function return the 'distance' between the scores of the best site and the site under evaluation. For the best site, $1-p_c^t/\max(p_c^t)$ gives a distance of zero.

The Monte Carlo simulation uses the method proposed by Hägerstrand (1965). First, a cell in the grid is randomly picked and its probability is compared with a random number uniformly distributed within 0 to 1. Because the probability is extremely small, it would take an extremely long time to identify a successful transition. To produce a large number of sites requires heavy computation. In other words, in the case of an $N \times N$ grid, the algorithm is very computationally *intensive*. Therefore, the algorithm must be adapted to a grid-based system. Alternatively, development sites can be evaluated simultaneously. This means multiple random number comparisons, which greatly speeds up the simulation. However, the adaptation does create a difficulty in the control of the total number of development in a specific time period, for example one year. Multiple comparisons mean each land parcel is evaluated individually and simultaneously without feedback from other sites. This could lead to a large number of sites developed at each iteration, exceeding the projected demand. In reality this would not happen because the competition among these sites will scale down the chance of development. Thus, the probability

of conversion needs to be further transformed if multiple evaluations take place at the same time. The probability that will lead to only one cell being converted is:

$$p_s^t(ij) = p_t^t / \sum_{ij} p_t^t(ij) \tag{5}$$

where p_s^t is the scaled probability that will lead to one expected conversion over the N × N grid.

The scaling of probability allows the expected quantity of development to be controlled according to the projected land demand. The expected total number of conversions in the grid when the p_s^t grid is simulated with a random grid of values in the range 0 to 1 is one. This transformation in fact constitutes an additional constraint to the joint probability. As a result, the scaled probability is composed of three probabilities: (1) the probability of development measured on global factors, (2) the probability of development measured on local factors, and (3) the probability of cell selection according to the projected land demand. The scaled probability that can generate the expected land conversion is specified by:

$$p_s^t(ij) = q p_t^t(ij) / \sum_{ij} p_t^t(ij) \tag{6}$$

where q is the number of cells to be converted according to projected land conversion at each iteration. The value of the right-hand should be limited at 1.0 to ensure that the probability is within the range of 0 to 1. The grid p_s^t is then directly compared with a random grid with uniform distribution from 0 to 1 to decide whether the cell is to be converted at time $t+1$:

$$s^{t+1}(ij) = \begin{cases} urban, & p_s^t(ij) > rand\ (ij) \\ rural, & p_s^t(ij) \leqslant rand\ (ij) \end{cases} \tag{7}$$

where $rand(ij)$ is a uniform 0–1 random distribution grid. The Monte Carlo process can be repeated many times. Each generates a different set of developed sites. These development sets may locate at different places in different simulations. But the distribution of development sites conforms the observed pattern and the total number of cells converted equals the number of expected developed sites.

In sum, the proposed CA model is a truly stochastic one as it follows the Monte Carlo procedure used in Hägerstrand (1965). By composing the final probability from three probabilities, namely global land development distribution, local probability due to neighbourhood agglomeration, and probability of site selection, this simulation ensures that the total quantity of development be constrained to the projected demand, thus bridging the macroscopic site selection with the microscopic CA dynamics.

5. Implementation in rural-urban land conversion

The above procedure was applied to simulate rural–urban land conversion in the city of Guangzhou in South China. Sequential land use data were used. Through the overlay of land use layers of 1973 and 1993 in ARC/INFO GIS, land use changes in this period were identified. The changes were then examined through spatial analysis, in particular the relationship between the distribution of developed sites and their physical and location characteristics. These characteristics are described by a number of 'contextual' development attributes, which include the travel distance to the edge of the city, the off-road distance to the nearest settlement, topographic elevation and slope, the variation of elevation and slope in a 90 m × 90 m kernel, and

dummy pre-development land use indicators (cultivated land, orchard, or wood). These attributes were measured by GIS operations and exported to a statistical package. The probabilities were then transformed and stochastically run through the Monte Carlo method. Different combinations of the probability and development constraints (local agglomeration) were tested. The results of the simulation were verified through profiles of land use distribution and urban morphology such as the extent to which different land uses are mixed and connected.

5.1. *Data*

The study area covers the Tianhe District, a suburban district of Guangzhou. Accelerated urban sprawl has caused a problem of urban land encroachment on the best quality agricultural land in the Pearl River Delta region (Yeh and Li 1997, 1999; Li and Yeh 2000). The main data consist of two land use coverages derived from 1973 and 1993 topographic maps and an elevation contour coverage. The coverage was digitised from 1:10 000 survey topographic maps. The digital detrain model (DEM) was then created from the digitised contours. From the triangulated irregular network (TIN), the slope grid was measured. The original land coverage records eight land use classes: arable land, orchard, wood, rural residential (settlements), urban area, industrial, transport, water body, and unused land. Since the purpose of this study is to test the procedure and it focuses on the rural–urban land conversion, these classes were generalised into three major categories: developed, and undeveloped land and water body. The land uses in 1973 and 1993 are shown in figures 3 and 4. The resolution of the grid is 30 m × 30 m, giving a total size of 546 rows and 629 columns.

The edge of the main urban area has been digitised and then rasterised. From the edge image, a diffusion function has been used to calculate the travel distance to the city. On-road speed (50 km hour^{-1}) and off-road speed (5 km hour^{-1}) are used to develop a travel time surface. A cost friction coefficient was added to the measurement of off-road travel according to the slopes (table 1). Similarly, the travel distance to the nearest rural settlements was calculated. To describe local variations of physical conditions, a 3 cell × 3 cell kernel was applied to the elevation and slope grids derived from DEM. The ranges of elevation and slope variations in the kernel were measured cell by cell. In order to assess the potential of undeveloped land, three dummy indicators are used to flag cultivated land, orchard, and wood. The development probability was then assessed by the comparison with that of non-agricultural land. In sum, the attributes used in the simulation are presented in table 2.

5.2. *Development probability and simulation*

Development probability was calibrated through a logistic regression using SPSS. The dependent variable is binary (whether the land use changed from 1973 to 1993) and independent variables are described by table 2. The result of the regression is presented in table 3. Two models, both significant at 0.001, show some regularity in land development. For example, land development probability decreases along with the increase in the distance to the city centre, as the sign of the coefficient of CITYDIST is negative. The full model (model II) is used to calculate the probability of land development. The function z is calculated as:

$$z = -8.1277 - 0.0437*CITYDIST - 0.0204*SETTDIST - 0.0032*SLOPE$$
$$- 0.0112*DEM - 0.0160*DEMRNG - 0.0075*SLOPERNG$$
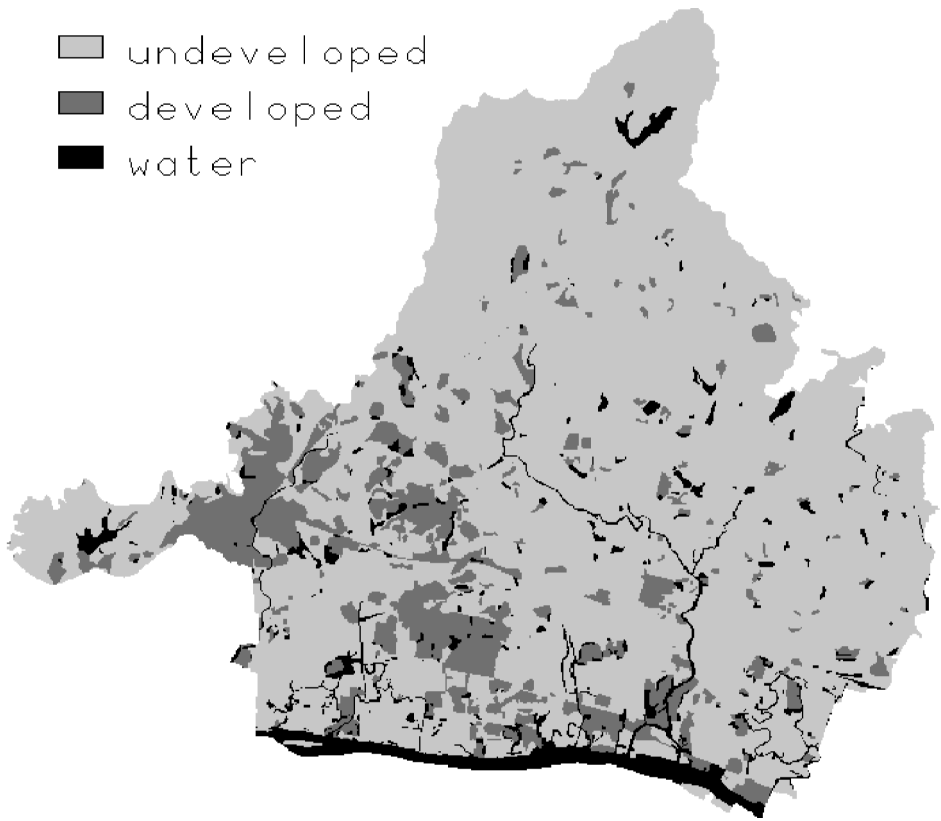$$+ 9.2319*USE* + 8.9502*USE2 + 9.99246*USE3 \tag{8}$$

Figure 3. The developed and undeveloped land use in the Tianhe district of Guangzhou in 1973.

The regression produces an overall classification accuracy of 79.4% and a highly significant model Chi-Square. The model should be regarded as an effective description of the land use conversions, as only a limited number of explanatory variables are used and land use conversions are usually distributed in a complicated way.

The probability calculated is shown in figure 5. The probability of development is higher at the immediate urban fringe, gradually fading away in the rural areas. The developed land (existing built-up area) has a probability of zero as further development is not modelled. The transformation probability must be specified with the particular rules used and the particular time. Figure 6 shows the transformation probability p_s^t at the time of 1985 in simulation 3 (see below).

The initial state of simulation starts from land use in the year 1973. Three simulation experiments were tested, each corresponding to a different combination of probabilities. Simulation 1 uses probability based on global conditions, thus $\Omega_{ij}^t = 1$; Simulation 2 uses probability based on local conditions, thus $p_g = 1$; Simulation 3 uses the combination of local and global conditions, thus p_g and Ω_{ij}^t are calculated accordingly. All three simulations were constrained by the dynamic land use change (i.e. developed land will no longer be transformed):

$$con(s_{ij}^t) = \begin{cases} 0 & if \ s = urban \ land \\ 1 & otherwise \end{cases} \tag{9}$$
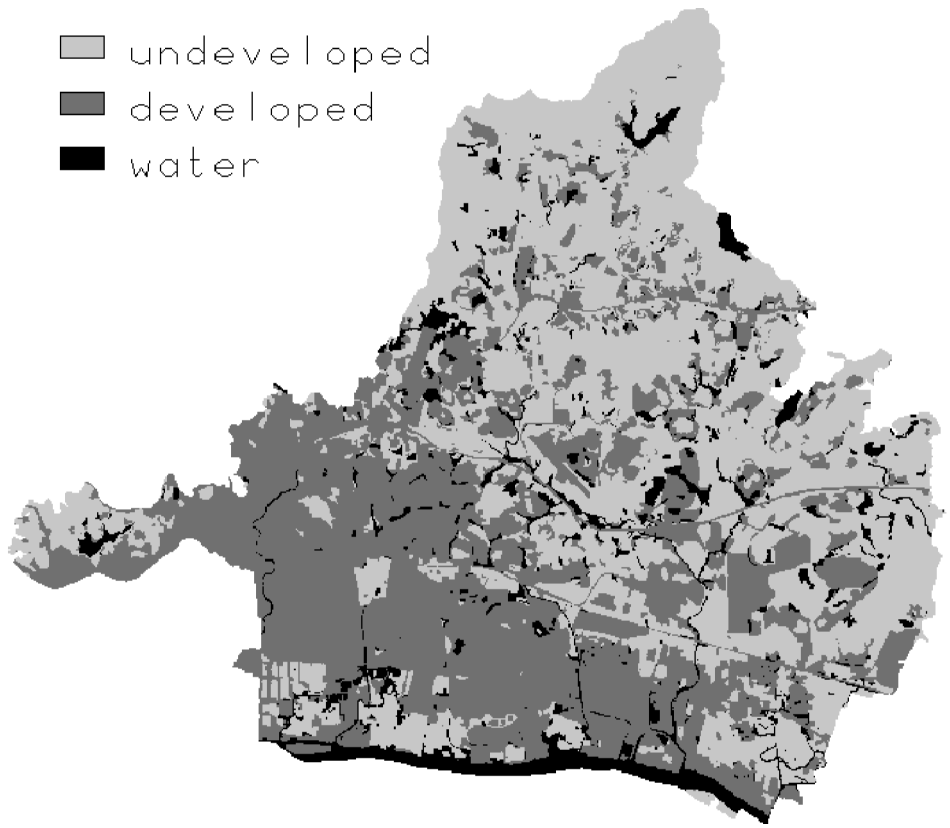
Figure 4. The developed and undeveloped land use in the Tianhe district of Guangzhou in 1993.

Table 1. The friction coefficient used in the measure of off-road travel distance.

Slope	Friction coefficient
0–6°	1.0
6–15°	1.5
15–30°	2.5
30–50°	3.0
> 50°	5.0
Water	10.0

Equation (9) means the model does not take redevelopment into account, i.e. developed land could not be redeveloped. However, there is no technical reason for not allowing redevelopment. The condition can be modified. For example, after a certain period of development redevelopment is allowed. The total area of the converted land use has been controlled according to the expected growth rate (2368 cells converted per year, equivalent to 213 ha). The simulation reproduced land uses in 1993. Simulated land use was then compared with the actual land use so as to assess the model's performance. The model can be re-run from the actual land use in 1993 to simulate land use in 2010 and can be projected according to estimated

Table 2. Development attributes used to compute land development probability.

Variable	Meaning	Measured in GIS
CITYDIST	Travel distance to the edge of the city, measured in minutes. This is composed by on-road and off-road travel speeds	Road network, land use, topographical features (slope), plus cost distance function
SETTDIST	Off-road distance to the nearest settlement (including villages and communities)	Topographical features plus cost distance function
DEM	Land elevation measured in meters	Built from TIN and in turn from topographical contour lines
SLOPE	The largest downward slope to adjacent land	Derived from DEM
DEMRNG	A contextual variable measuring the range of variation of elevation in a $30\,m \times 30\,m$ kernel	
SLOPERNG	A contextual variable measuring the range of slope variation in a $30\,m \times 30\,m$ kernel	GRID focal function
USE1	A dummy indicator if the land use is cultivated land	GRID reclassification
USE2	A dummy indicator if the land use is orchard	GRID reclassification
USE3	A dummy indicator if the land use is wood	GRID reclassification

The baseline land use is non-agricultural land use, coded by USE1=USE2=USE3=0. The meaning of coefficient of U1, U2, U3, is therefore that compared to non-agricultural land uses, the contribution to the likelihood of development from being cultivated land, orchard, and wood respectively. Despite similar coefficients, the model suggests that woodland is the most likely type to be developed, while cultivated is the least likely type.

Table 3. The probability of land use changes calibrated from the logistic regression models.

Variable	Model I B	S.E.	Model II B	S.E.
CITYDIST	-0.0363	0.0005	-0.0437	0.0006
SETDIST			-0.0204	0.0008
SLOPE	-0.0044	0.0007	-0.0032	0.0012
DEM	-0.0071	0.0003	-0.0112	0.0004
SLOPERNG			-0.0075	0.0009
DEMRNG			-0.0160	0.0021
USE1			9.2319	0.5384
USE2			8.9502	0.5390
USE3			9.9246	0.5385
Constant	0.2434	0.0112	-8.1277	0.5384
PCP	72.31		79.37	
$-2LL$	180 019.681**		139 739.556**	

Note: PCP=percentage correctly predicted; $-2LL=-2$ log likelihood at convergence; B=coefficient; S.E.=standard error; **=significant at 0.001.

land demand. The land development simulated from 1993 to 2010 is not reported here as the purpose of this paper is to see how the calibrated CA performs. The results of three experiments are shown in figures 7, 8 and 9 respectively.

Figure 5. The development probability calibrated from the logistic regression model (Model II).

5.3. *Analysis of model performance*

The model performance is assessed in three ways: namely, spatial overlay which generates a tabulation and visualisation, spatial statistics which measure connectivity and morphology, and structural measurement which assesses the goodness-of-fit according to the special domain of interest.

5.4. *Cross-tabulation*

The land use of 1993 generated by simulation was compared with the actual land use. This produces three cross-tabulation tables, from which the accuracy of each simulation can be assessed (see tables 4, 5, and 6).

The accuracy measured from cross-tabulation is as follows. Simulation 1 has an overall 72.9% correct prediction. The figure increased to 74.9% in Simulation 2, which is higher than expected. Intuitively, the transition rule that only takes account of the local probability should not be able to reproduce the land use pattern in 1993, as development factors are not considered. However, because the starting state of the simulation is based on the actual land use pattern of 1973, development factors were embedded in the distribution of land uses. The simulation allocated projected land demand to sites adjacent to existing urban areas through the neighbourhood function. Thus, the simulation produced a pattern very similar to the historical one.
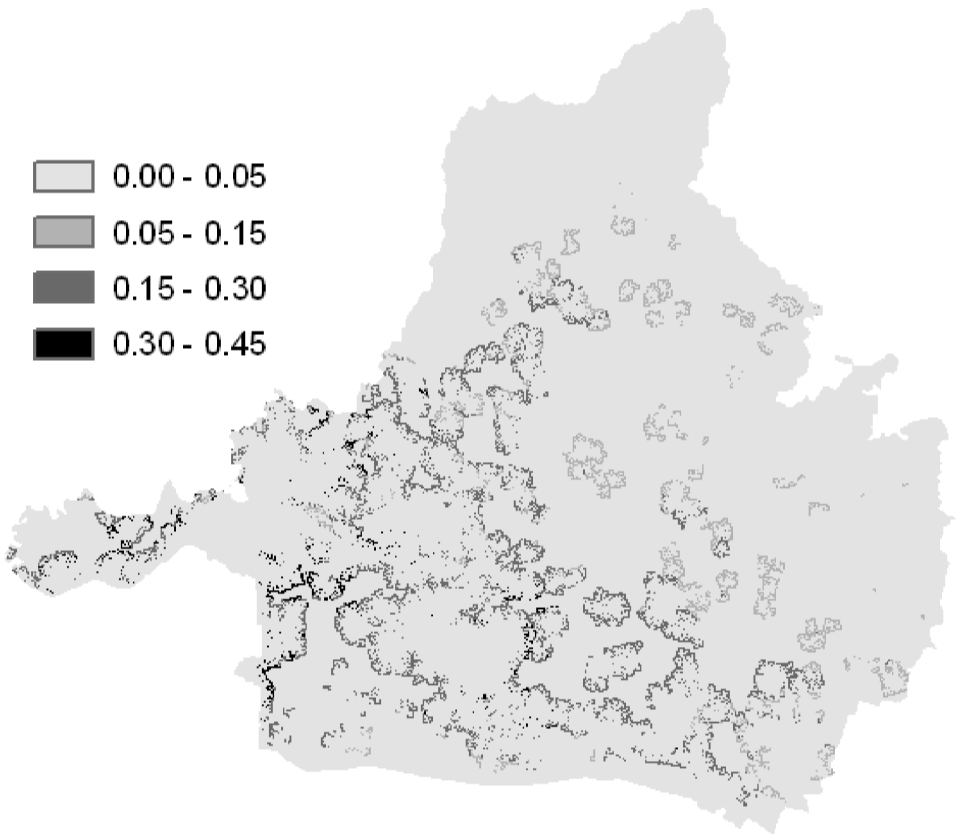
Figure 6. The transformed probability of conversion measured dynamically at the time of 1985 in Simulation 3 (see text for definition of Simulation 3).

Simulation 3 has the highest accuracy, which is 79.5%, among the three simulations. The simulation predicts developed and undeveloped land uses at an accuracy of 76.6% and 81.6%. Please note that cross-tabulation produces a stringent test of simulation as it measures on a cell-by-cell basis.

5.5. *Spatial overlay*

The cross-tabulation analysis discussed above compares the goodness-of-fit on a one-to-one cell basis. In other words, we assess the correspondence between the observed and simulated land use by individual cells at the same co-ordinates of grids, while their spatial distribution (such as how different land uses are connected) is ignored. Therefore, a high correspondence rate does not ensure that simulation produces a plausible urban morphology. On the other hand, a low correspondence rate does not necessarily mean poor spatial agreement because the simulated and observed land use patterns may be similar to each other in terms of spatial structures. The analysis of conformity should therefore be extended to the measurement of the spatial relationship between developed and simulated sites.

Visually, the simulated urban structure can be compared with the observed one through GIS overlay. For each pair of land use images, four categories have been produced. These include: undeveloped land predicted as undeveloped land, and
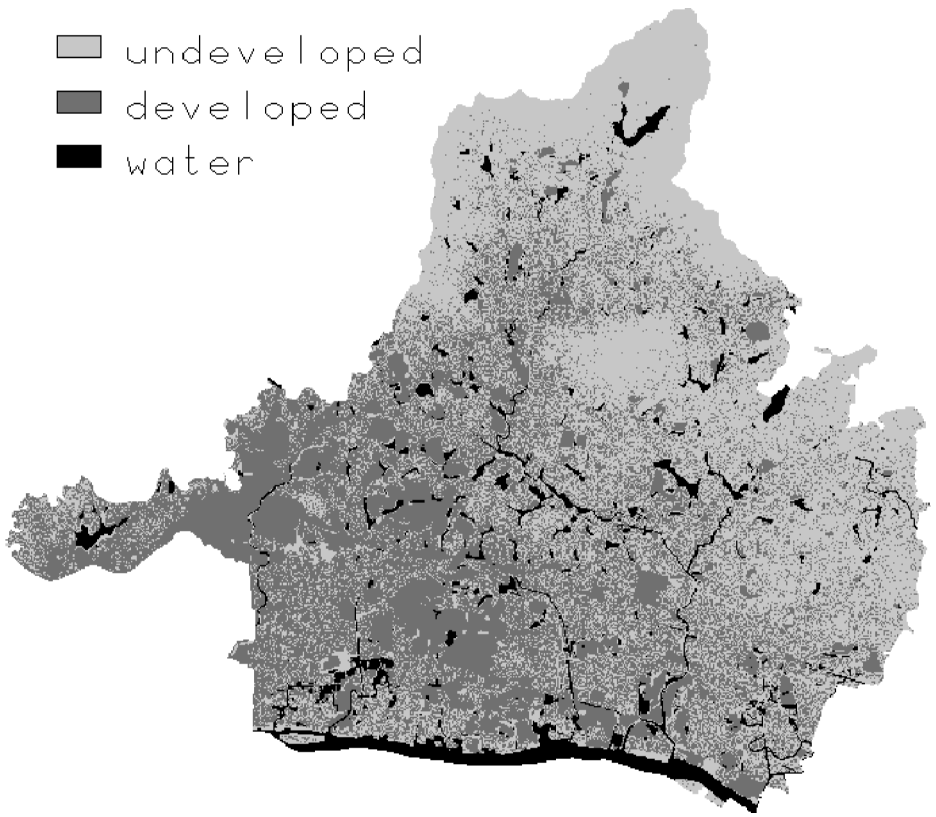
Figure 7.   The result of Simulation 1 (based on global development conditions).

developed land predicted as developed land, undeveloped land predicted as developed
land, developed land predicted as undeveloped land. The first two classes are correct
predictions, while the latter two are errors. The results of the three simulations are
shown in figures 7, 8 and 9. The error of Simulation 1 is distributed 'randomly' (the
black dots representing unpredicted developed sites). This is not surprising because
the calibrated land use model describes the trend of land use change without
considering unknown errors (residuals in regression). However, the morphology of
urban spatial structure simulated here is totally *unrealistic*. Such a scattered land
development pattern cannot be realised in the real world and is not possible even
in the form of urban sprawl. Development sites have to be connected in order to
develop infrastructure and service facilities. This suggests that land use modelling is
more appropriate at the aggregated scale than at microscopic scales. By using large
spatial units such as urban districts, wards and transport zones, the model does not
require an explicit description of interactions within units. Local interactions should
be introduced into the model if it is applied to the grid system.
  Simulations 2 and 3 introduced the local effect in land development simulation.
Both simulations produced realistic land use patterns. But Simulation 2 obviously
over-predicts land development (the dark grey shaded area) in remote rural areas.
This is because in Simulation 2 no global factor (such as distance to the city centre)
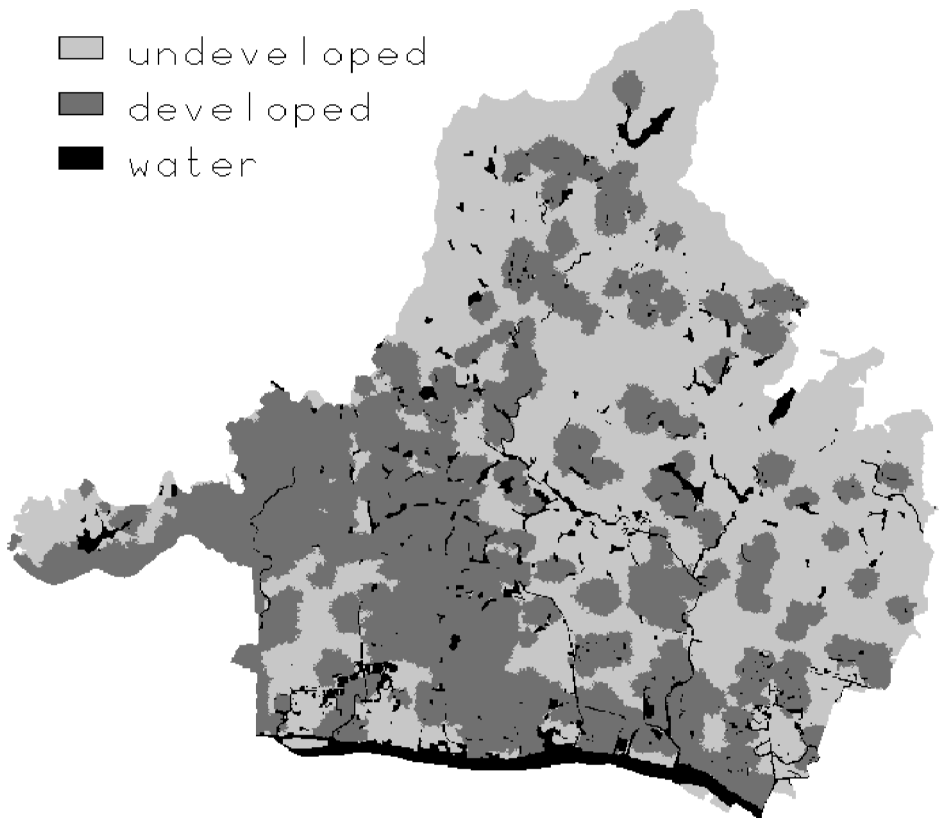has been taken into account. Therefore, a place near to a small rural settlement is

Figure 8.   The result of Simulation 2 (based on local development conditions).

seen as equally attractive as one close to the edge of the city. This is obviously not the case in urban development because accessibility as well as the physical character-istics of development sites (such as the slope), are important. As a result, Simulation 2 under-estimates development in areas near to the edge of the city and places with better development conditions.

The result of spatial overlay clearly indicates that Simulation 3 has produced the best spatial prediction of land development. Both simulations, however, did not predict the development of some large land parcels, because these large developments, supported by state investment in the era of the planned economy, broke away from the development trajectory. This is not unique to the socialist economy. In a market economy, extraordinary development events are inevitable and cannot be 'predicted' by modelling and planning.

5.6. *Spatial statistics*

The spatial overlay suggests that structural conformity is important in the assess-ment of simulation results. This conformity between observed and simulated land uses can be analysed through spatial statistics. Moran I, for example, is a spatial statistical indicator that reflects the degree of spatial autocorrelation (Goodchild 1986). The indicator is used to reveal the pattern of the clustering of the same type of use at adjacent cells and, therefore, the extent to which developed sites (coded as
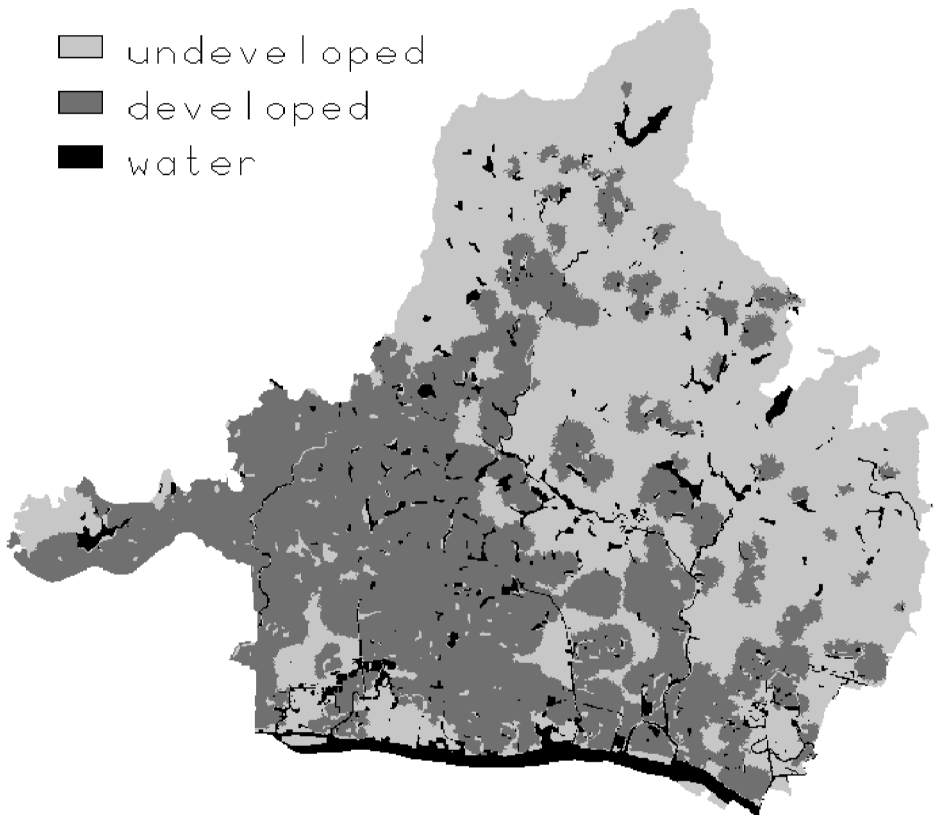
Figure 9. The result of Simulation 3 (based on the joint probability of global and local development conditions).

Table 4. The result of Simulation 1 (based on global development factors) compared with observed land development in number of cells.

	Simulated		
Observed	Developed	Not developed	% correct
Developed	47 356	21 348	68.9
Not developed	23 797	72 506	75.3
Overall	71 153	93 854	72.6

Table 5. The result of Simulation 2 (based on local conditions) compared with the observed land development in number of cells.

	Simulated		
Observed	Developed	Not developed	% correct
Developed	49 474	19 413	71.8
Not developed	22 097	74 842	77.2
Overall	71 571	94 255	74.9

Table 6. The result of Simulation 3 (based on the joint probability) compared with the observed land development in number of cells.

Observed	Simulated		
	Developed	Not developed	% correct
Developed	52 647	16 057	76.6
Not developed	17 720	78 583	81.6
Overall	70 367	94 640	79.5

1) and undeveloped sites (coded as 0) are mixed. The absolute concentration of land uses within a kernel generates a Moran I close to unity, while a more even distribution than can be expected by chance gives a value below zero. Although the absolute value of Moran I may not correspond to a fixed scale of spatial clustering, the indicator can be used to compare how close the simulated land use pattern is to the observed one. The Moran I function provided by Arc/Info measures immediately adjacent cells. The value thus only indicates land use patterns at the finest possible scale and does not describe the clustering of land uses as spatial objects or at a structural level. The more macroscopic assessment of spatial structures will be discussed later.

The Moran I function, which is applied to the observed pattern of land development, returned a value of 0.829. The values of Moran I for the three simulations are respectively 0.396, 0.898, and 0.894. This is fairly consistent with visual comparison. The second and third simulations are morphologically more similar to each other, while the first simulation stands out quite differently. While the difference is quite obvious in this case, spatial statistics is still useful in that it probably provides a quantified measure for more subtle cases. The quantified measurement is particularly important in the development of automated fine-tuning of a simulation model.

5.7. *Structural measurement*

As mentioned earlier, the structural measure is needed to assess the distribution of simulated land development in a larger area. Structural conformity has been measured through the profile of development that gives a meaningful indicator of how the simulated pattern matches the observed one in terms of spatial structures, for example land use distribution measured by travel time to the edge of the built-up area. The study has divided the whole simulation area into twenty-six 5-minute travel time zones. The total number of developed cells in each zone was counted. The percentage of development in each zone can be plotted against the travel time, producing a profile of development. The observed and simulated development profiles are shown in figure 13. The percentages can be cross-tabulated with the expected percentages to produce a χ^2 value. The indicator is used to assess the deviation of simulation from the expected frequency of observation. The χ^2 of three simulations are 0.21, 0.22 and 0.62. Clearly, Simulations 1 and 2 produced a rather similar pattern of distribution to the observation, while Simulation 3 slightly overestimates in the area near to the urban edge and underestimates in the rural areas far away from the city. Simulation 2 overestimates the development in the remote rural areas. It is not too surprising that Simulation 1, i.e. the simulation based on calibrated probability predicts better. This is because information about the urban structure has been fully taken into consideration without the disturbance of other factors.
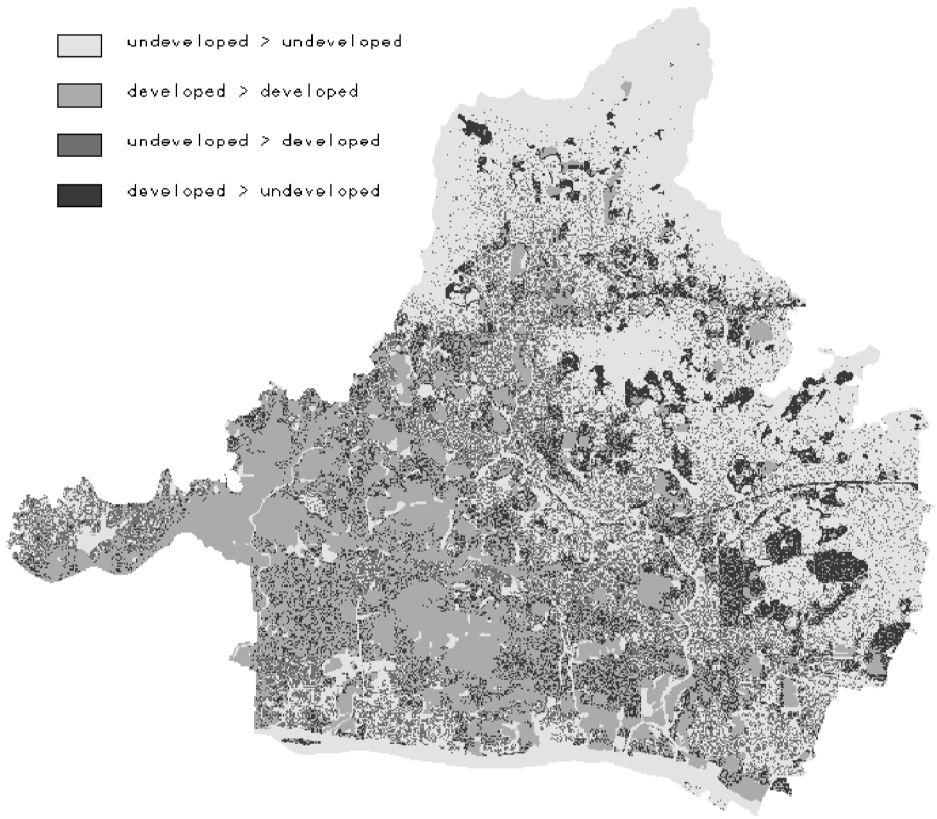
Figure 10.   The comparison of Simulation 1 with the observation through spatial overlay.

Similarly, other structural measures can be adopted, depending upon the purpose of the simulation.

6.   Conclusion

There have been persistent efforts to understand urban land use changes and urban spatial structure since the time of the Chicago School. However, conventional urban models are spatially aggregated, paying less attention to the interaction between land development factors at the local level. Recently there has been a surge of heuristic land development simulations. Most of these models rely on the capacity of computation and especially the spatial data analysis functionality of GIS. The process-based models can be divided into two major categories: analogy models which make an analogy of land development to a physical process such as diffusion limited aggregation (DLA) (Batty and Longley 1994) and the ecology of species diffusion (Burrough 1998), and behavioural models which use development rules that are deemed plausible (Batty and Xie 1994, Batty 1998, Clarke *et al.* 1997, Clarke and Gaydos 1998, Li and Yeh 2000, 2001, Portugali and Benenson 1995, White and Engelen 1993, 1997, Wu and Webster 1998). Both categories of modelling are confronted with a serious computational demand—unless the model adopts a simple mathematical form it would be very difficult to investigate all possible parameter ranges. Seemingly it is not a question of theoretical purity but rather a reliable
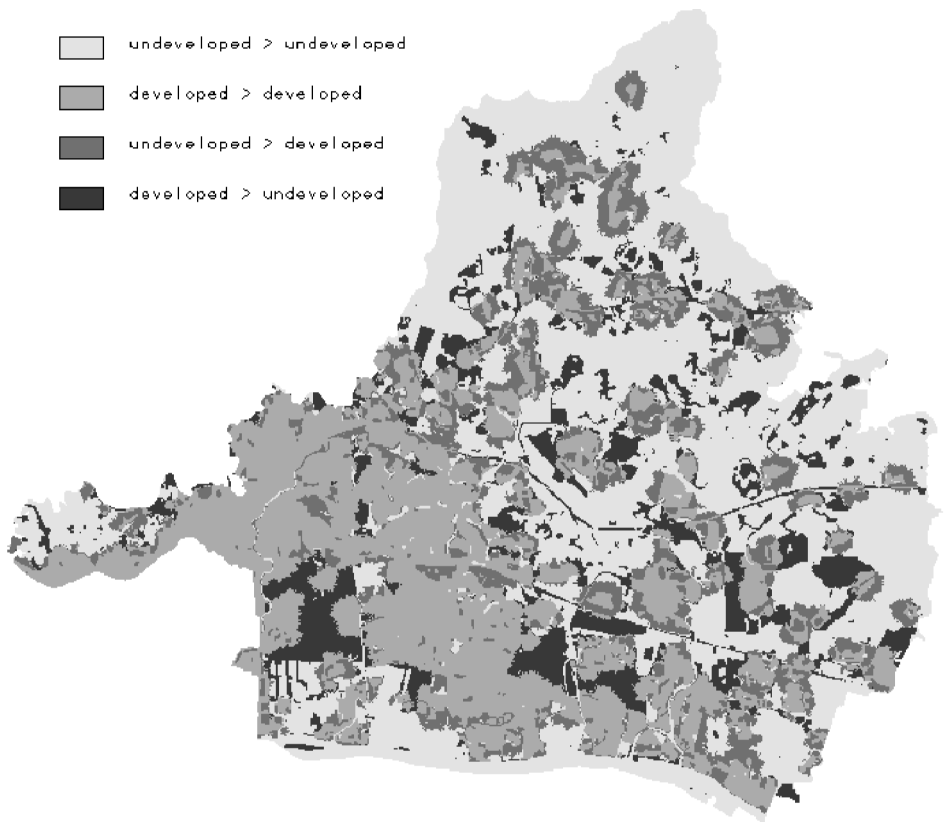
Figure 11.   The comparison of Simulation 2 with the observation through spatial overlay.

calibration method that matters. Neoclassical land bid-rent models and the Chicago model of social areas should be seen more as conceptual models rather than empirical ones. The need for a simulation method for empirical land use analysis is increasingly acute because of urban sprawl in both developed and developing countries.

This research is based on the development of a procedure which calibrates the initial global probability surface from sequential land use data and then modifies the global probability with the local probability that is updated at each simulation iteration. In addition, the probability of site selection is incorporated into the joint probability to constrain the quantity of simulated conversions to projected land demand. The result suggests that this joint probability has produced the best performance. This paper also emphasises the need to validate the model through both structural and cross-tabulation measures. In zone-based land use simulation, the accuracy issue is not a problem because aggregation of spatial units removes the internal differentiation. In grid-based microscopic simulation, the result needs to be presented in cells rather than in census tracts, enumeration districts or urban wards, thus imposing the need to validate simulation through spatial indicators. Moreover, the result of simulation should not be validated on a cell by cell basis. Land conversion involves interactions among actors who have bounded rationality and are affected by many unknown political, cultural and economic factors. However, the influence of these factors cannot be manifested in equations. As a result, the
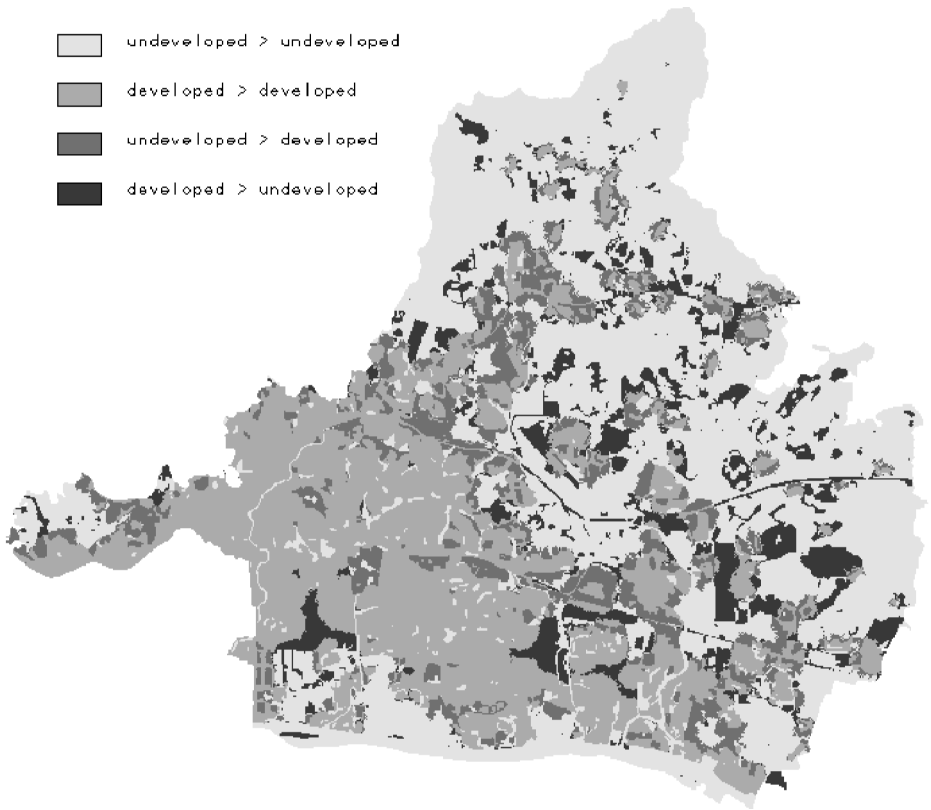
Figure 12.   The comparison of Simulation 3 with the observation through spatial overlay.
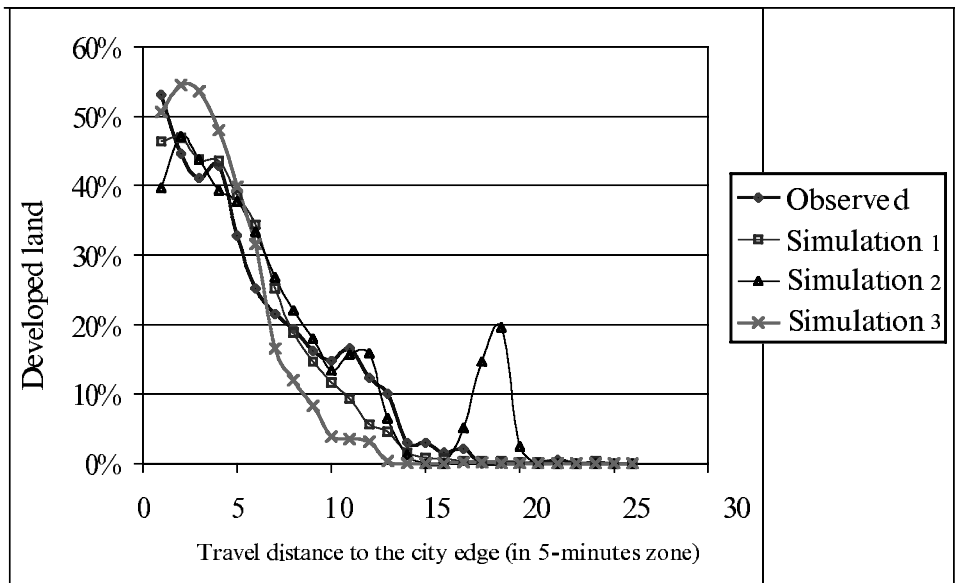

Figure 13.   The development profiles by the travel distance to the edge of the city.

distribution of land development should be better portrayed through observation. The development of initial global probability surface serves this purpose. The use-fulness of land use simulation lies in the prediction of general urban development, i.e. the trajectory of development. It has been suggested that the trajectory that involves both the global and local dynamics of land development is most likely to be captured by calibrated CA simulation. Moreover, if we 'calibrate' the CA simulation through the desired pattern of changes, then we may incorporate the desired relationship into future development evaluation to simulate alternative scenarios. By doing so, simulation becomes more than a simple projection based on past trends. Future research should be carried out to study the desired pattern of rural-urban land conversions.

Acknowledgments

References

ALONSO, W., 1964, *Location and Land Use* (Cambridge, Mass: Harvard University Press).

BATTY, M., 1998, Urban evolution on the desktop: simulation with the use of extended cellular automata. *Environment and Planning A*, **30**, 1943–1967.

BATTY, M., 1997. Cellular automata and urban form: a primer. *Journal of American Planning Association*, **63**, 266–274.

BATTY, M., and LONGLEY, P. A., 1994, *Fractal Cities* (London: Academic Press).

BATTY, M., and XIE, Y., 1994, From cells to cities. *Environment and Planning B*, **21**, 531–548.

BATTY, M., XIE, Y., and SUN, Z., 1999, Modeling urban dynamics through GIS-based cellular automata. *Computers, Environment and Urban Systems*, **23**, 205–233.

BEN-AKIVA, M., and LERMAN, S., 1985, Discrete Choice Analysis: Theory and Application to Travel Demand (Cambridge, MA: MIT Press).

BURROUGH, P. A., 1998, Dynamic modelling and geocomputation. In *Geocomputation: A Primer*, edited by P. A. Longley, S. M. Brooks, R. McDonnell and B. Macmillan. (Chichester: John Wiley & Sons), pp. 165–191.

CLARKE, K. C., HOPPEN, S., and GAYDOS, L. 1997, A self-modifying cellular automaton model of historical urbanization in the San Francisco Bay area. *Environment and Planning B*, **24**, 247–261.

CLARKE, K. C., and GAYDOS, L. J., 1998, Loose-coupling a cellular automaton model and GIS: long-term urban growth prediction for San Francisco and Washington/Baltimore. *International Journal of Geographical Information Science*, **12**, 699–714.

COUCLELIS, H., 1997, From cellular automata to urban models: new principles for urban development and implementation. *Environment and Planning B*, **24**, 165–174

GARDNER, M., 1970, The fantastic combinations of John Conway's new solitaire game 'Life'. *Scientific American*, **223**, 120–123.

GOODCHILD, M. F., 1986, *Spatial Autocorrelation: Concepts and Techniques in Modern Geography*, **47** (Norwich: Geo Books).

HÄGERSTRAND, T., 1965, A Monte Carlo approach to diffusion. *Archive of European Sociology*, **VI**, 43–67.

LI, X., and YEH, A. G. O., 2000, Modelling sustainable urban development by the integration of constrained cellular automata and GIS. *International Journal of Geographical Information Science*, **14**, 131–152.

LI, X., and YEH, A. G. O., 2001, Calibration of cellular automata by using neural networks for the simulation of complex urban systems. *Environment and Planning A*, **33**, 1445–1462.

MCMILLEN, D. P., 1989, An empirical model of urban fringe land use. *Land Economics*, **65**, 138–145.

PORTUGALI, J., and BENENSON, I., 1995, Artificial planning experience by means of a heuristic cell-space model: simulating international migration in the urban process. *Environment and Planning A*, **27**, 1647–1665.

TAKEYAMA, M., and COUCLELIS, H., 1997, Map dynamics: integrating cellular automata and GIS through geo-algebra. *International Journal of Geographical Information Science*, **11**, 73–92.

WHITE, R., and ENGELEN, G., 1993, Cellular automata and fractal urban form: a cellular modelling approach to the evolution of urban land-use patterns. *Environment and Planning A*, **25**, 1175–1189.

WHITE, R., ENGELEN, G., and UIJEE, I. 1997. The use of constrained cellular automata for high-resolution modeling of urban land use dynamics. *Environment and Planning B*, **24**, 323–343.

WOLFRAM, S., 1984, Universality and complexity in cellular automata. *Physica D*, **10**, 1–35.

WU, F., 1998a. An experiment on the generic polycentricity of urban growth in a cellular automatic city. *Environment and Planning B*, **25**, 731–752.

WU, F., 1998b, SimLand: a prototype to simulate land conversion through the integrated GIS and CA with AHP-derived transition rules. *International Journal of Geographical Information Science*, **12**, 63–82.

WU, F., and WEBSTER, C. J., 1998, Simulation of land development through the integration of cellular automata and multi-criteria evaluation. *Environment and Planning B*, **25**, 103–126.

XIE, Y., 1996, A generalized model for cellular urban dynamics. *Geographical Analysis*, **28**, 350–373.

YEH, A. G. O., and LI, X., 1997, An integrated remote sensing and GIS approach in the monitoring and evaluation of rapid urban growth for sustainable development in the Pearl River Delta. *International Planning Studies*, **2**, 193–210.

YEH, A. G. O., and LI, X., 1999, Sustainable land development model for rapid growth areas using GIS. *International Journal of Geographical Information Science*, **12**, 169–189.

ZHOU, Z. G., 1999, *Spatial and Temporal Aspects of Land Use in the Urban-Rural Fringe in China: A GIS Approach*, unpublished M.Sc. thesis, the Department of Geography, University of Durham.