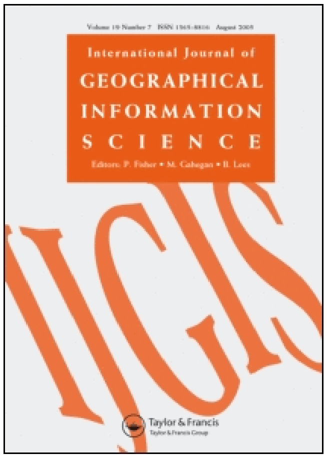

International Journal of Geographical Information Science
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713599799

Embedding sustainable development strategies in agent-based models for use as a planning tool

Xia Li [a]; Xiaoping Liu [a]
[a] School of Geography and Planning, Sun Yat-sen University, Guangzhou, PR China

Online Publication Date: 01 January 2008
To cite this Article: Li, Xia and Liu, Xiaoping (2008) 'Embedding sustainable development strategies in agent-based models for use as a planning tool', International Journal of Geographical Information Science, 22:1, 21 - 45
To link to this article: DOI: 10.1080/13658810701228686
URL: http://dx.doi.org/10.1080/13658810701228686

International Journal of Geographical Information Science
Vol. 22, No. 1, January 2008, 21–45

Research Article

Embedding sustainable development strategies in agent-based models for use as a planning tool

XIA LI* and XIAOPING LIU
School of Geography and Planning, Sun Yat-sen University, Guangzhou, 510275, PR China

(*Received 22 June 2006; in final form 11 January 2007*)

Rapid land development in rapidly growing countries has created a series of land-use problems. The implementation of sustainable land use can alleviate some of these problems. It needs a set of tools for the exploration, design, modification, illustration, and evaluation of alternative planning scenarios. This paper demonstrates that the integration of cellular automata and agent-based modelling can provide a spatial exploratory tool for generating alternative development patterns. Sustainable development strategies are embedded in the modelling to regulate agents' behaviours. The use of agents can help to represent human–environment interactions in solving complex land-use problems. It is able to examine the effects of different stakeholders in influencing the process of land development. The proposed model has been applied to the simulation of planning scenarios for residential development in a rapidly expanding city in the Pearl River Delta.

Keywords: Agent-based modelling; Cellular automata; GIS; Urban planning

1. Introduction

Rapid urban expansion in rapidly growing countries has created a major concern for sustainable land use in these regions (Li and Yeh 2001). Massive conversion of non-urban land into urban land has created a series of land-use problems, such as a decrease in food production, destruction of sensitive ecosystems, water and air pollution, and deprivation of future land supply (Yeh and Li 1999, Jantz *et al.* 2005). Sustainable land use, which should also coordinate the land-use demands from multiple aspects and different interest groups, can provide a useful tool to alleviate some of these land-use problems. The implementation of sustainable land use is quite complex because it involves social, economic, and environmental factors. Modelling systems can be developed to provide assistance in implementing the initiatives of sustainable land use (Zandera and Kächele 1999). These models are useful for carrying out a scenario analysis which is a promising and interesting planning tool for investigating future possibilities in a changing environment (Nijkamp *et al.* 1997).

Cellular automata (CA), a type of bottom-up approach, have been used to investigate the 'business-as-usual' scenario, that is, further development of present conditions. These models have been widely used to simulate complex geographical phenomena which have nonlinear and emergent features (White and Engelen 1993,

*Corresponding author. Email: lixia@mail.sysu.edu.cn

International Journal of Geographical Information Science
ISSN 1365-8816 print/ISSN 1362-3087 online © 2008 Taylor & Francis
http://www.tandf.co.uk/journals
DOI: 10.1080/13658810701228686

Batty and Xie 1994, Li and Yeh 2000). However, CA have limitations reflecting the decision behaviours of individuals, such as governments and investors, in shaping urban growth. The influence of human factors is difficult to include in traditional CA (Torrens and Benenson 2005).

Agent-based modelling (ABM) can be used as a tool for analysing complex natural systems (Courdier *et al.* 2002). A major feature and advantage of ABM is the ability to produce nonlinear and emergent phenomena based on behaviour of individuals. In the past, ABM have been used mostly in purely social contexts (Gilbert and Conte 1995). They were used to validate or illustrate social theories (e.g. biological, economic, and political theories) or to predict the behaviour of interacting social entities (e.g. actors in financial markets and consumer behaviour) (Basu and Pryor 1997). However, this type of model does not make use of spatial information central to geographical analyses.

Both CA and ABM are limited in their geographic functionality when considered in isolation (Torrens and Benenson 2005). However, increasingly researchers are turning to the integration of CA with ABM to produce better simulation results. Research has indicated that not only the neighbourhoods (or cell states of CA) but interactions between local actors and their environment must be considered in order to forecast landscape transition with a higher accuracy (Loibl and Toetzer 2003). The integration of CA with ABM promises to provide a powerful spatial approach to the modelling of complex geographic systems that are affected by physical factors (e.g. land use and accessibility) and individuals (e.g. organizations and human objects) (Torrens and Benenson 2005).

There are as yet no published studies on the integration of both techniques as a planning tool for implementing the initiative of sustainable land use. Sustainable land-use planning generally requires the analysis of a vast array of spatial data. It needs a set of tools for the exploration, design, modification, illustration, and evaluation of alternative planning scenarios (Henton and Studwell 2000).

This paper will examine the integration of CA and ABM as a planning tool for managing residential development. The crucial part of this model is to define agents' behaviours based on sustainable development strategies. The efficiency criteria in using land resources are adopted to alleviate land-use conflicts in rapidly growing cities. This bottom-up approach is well adapted to the simulation of the interactions and negotiations of different stakeholders. It can provide a useful spatial exploratory tool for comparing various development options and evaluating the potential impacts of implementing certain land-use policies.

2. Study area and data

The study area is situated in the Haizhu district of Guangzhou, a rapidly growing city in the Pearl River Delta, China. Unprecedented land-use changes have been witnessed in the region in the last two decades (Li and Yeh 2004). The land-use changes are associated with many environmental problems, such as agricultural land loss, urban sprawl, and soil erosion (Yeh and Li 1999). In particular, urban expansion has triggered the loss of a large amount of agricultural land in the Pearl River Delta.

The spatial information for the proposed integrated CA and ABM model is obtained using remote sensing and GIS data. The common land-use types in this study area include urban land, farmland, forest, orchard, and water. GIS are used to provide the spatial information related to land-use changes. This type of spatial information includes the maps of planning schemes, land price, land use, and public facilities (e.g. hospitals, schools, and parks) (figure 1). Additional information (e.g. age and income) is also obtained from the statistical yearbooks of Guangzhou and the *Fifth National Censu*. The above information is used as the inputs to the modelling and the basis to define agents' properties.

3. Integrated CA and ABM planning model

The proposed model consists of three components, GIS, CA, and ABM, for simulating planning options related to residential development (figure 2). The GIS component is used to provide the inputs to simulation and model calibration. The CA component is to reflect neighbourhood influences of physical factors. The ABM component provides a flexible tool to address the interactions between various stakeholders that affect residential development. The following sections describe the detailed procedures in implementing this planning model.

3.1 *Retrieving physical factors using GIS*

3.1.1 Land use. Land use is one of the important factors in urban simulation. Agents have different decision behaviours with regard to land-use types. For

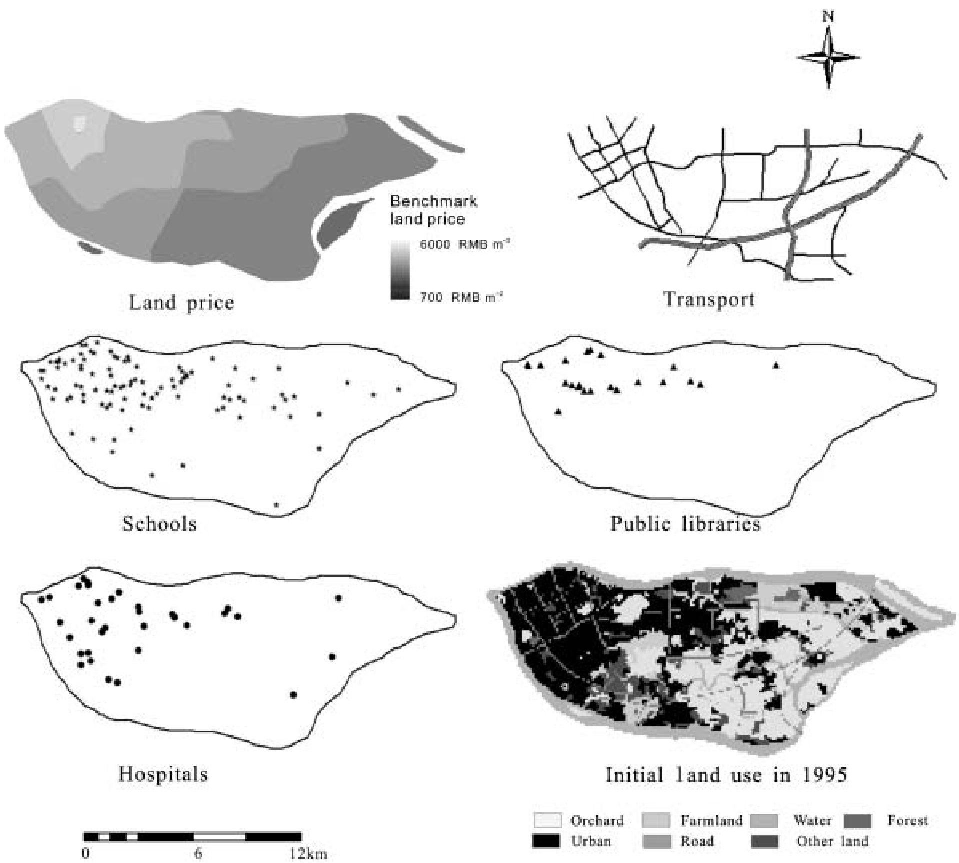

Figure 1. Spatial information as the inputs to the simulation.

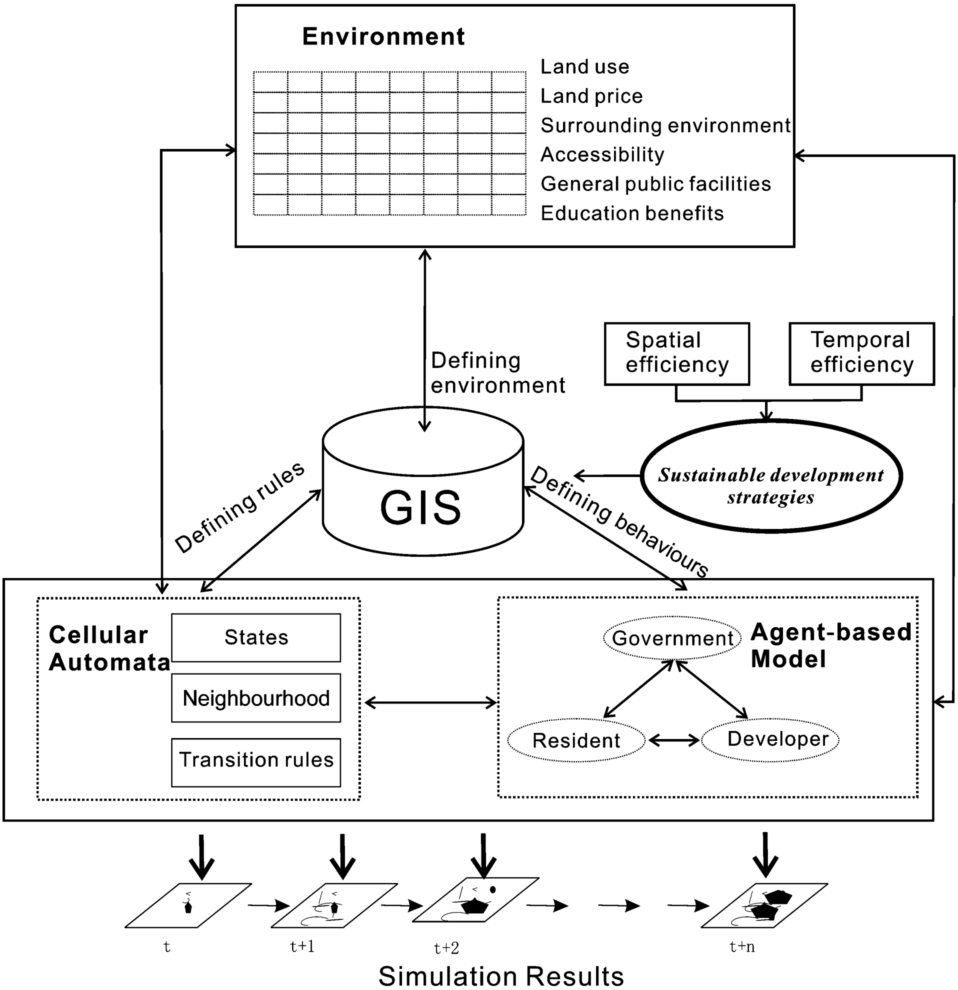

Figure 2. Planning model by the integration of cellular automata, agent-based modelling and GIS.

example, a resident agent has a preference to live in the sites surrounded by a large area of green land (e.g. forest and orchard) and water, instead of densely developed land.

3.1.2 Land price. Land price plays a key role in affecting urban development, especially residential development. Land price is correlated to housing price, which is a major concern for a potential home buyer. Residents' financial status determines their location preferences in buying a home. High-income residents choose locations of high housing prices to live, while low-income residents choose places of low housing prices.

3.1.3 Surrounding environment. The attraction of a site for urban development is related to its surrounding living environment. The surrounding environment is measured using two indicators, the percentage of green land and the percentage of water in the neighbourhood. These are calculated using a moving 9×9 window in classified satellite images. Finally, the utility (attraction) of a site related to this amenity is obtained using the following equation:

$$B_{\text{env}}(i) = \tfrac{1}{2} G_{\text{percent}}(i) + \tfrac{1}{2} W_{\text{percent}}(i) \quad 0 \le G_{\text{percent}}(i) + W_{\text{percent}}(i) \le 1 \tag{1}$$

where $B_{\text{env}}(i)$ is the utility of the surrounding environment, and $G_{\text{percent}}(i)$ and $W_{\text{percent}}(i)$ are the percentages of green land and water at location i, respectively. These two variables are treated with equal importance, since there is no prior knowledge.

3.1.4 Accessibility. Accessibility is related to its geographical location (e.g. distance to roads and town centres) and the conditions of road networks. A site will be more likely to develop if it is easily accessed. The utility (benefits) of a site related to the accessibility is represented as follows:

$$B_{\text{access}}(i) = \frac{1}{3} e^{-b_1 \cdot D_{\text{road}}(i)} + \frac{1}{3} e^{-b_2 \cdot D_{\text{express}}(i)} + \frac{1}{3} e^{-b_3 \cdot D_{\text{centre}}(i)} \tag{2}$$

where $B_{\text{access}}(i)$ is the utility related to accessibility at location i; the variables $D_{\text{road}}(i)$, $D_{\text{express}}(i)$, and $D_{\text{centre}}(i)$ are the Euclidean distances to roads, expressways, and urban centres, respectively; and b_1, b_2, and b_3 are the decay coefficients for these variables. The same weight (1/3) is also applied to all these variables for simplicity.

3.1.5 General public facilities. A site will be more likely to develop if it is closer to facilities, such as hospitals, gardens, commercial centres, and entertainment centers. Therefore, the utility of a site in terms of facility provision can be represented as follows:

$$B_{\text{facil}}(i) = \frac{1}{4} e^{-b_1 \cdot D_{\text{hospital}}(i)} + \frac{1}{4} e^{-b_1 \cdot D_{\text{garden}}(i)} + \frac{1}{4} e^{-b_1 \cdot D_{\text{commercial}}(i)} + \frac{1}{4} e^{-b_1 \cdot D_{\text{entertainment}}(i)} \tag{3}$$

where $B_{\text{facil}}(i)$ is the utility related to the provision of public facilities at location i, such as hospitals, gardens, commercial centres and entertainment; and the variables $D_{\text{hospital}}(i)$, $D_{\text{garden}}(i)$, $D_{\text{commercial}}(i)$, and $D_{\text{entertainment}}(i)$ are the Euclidean distances to these facilities, respectively. The same decay coefficient of b_1 in equation (2) is used, since these facilities are mainly accessed by roads. All these variables are treated with the same weight (1/4) in the calculation.

3.1.6 Education benefits. Education is an important attraction factor to home buying. A Euclidean distance function can also be used to represent the accessibility of a location to education facilities (e.g. schools and libraries). More education benefits can be achieved if the location is closer to these facilities. This utility is estimated as follows:

$$B_{\text{edu}}(i) = \frac{1}{2} e^{-b_1 \cdot D_{\text{school}}(i)} + \frac{1}{2} e^{-b_1 \cdot D_{\text{library}}(i)} \tag{4}$$

where $B_{\text{edu}}(i)$ is the utility related to the provision of educational facilities in terms of schools and public libraries at location i; and the variables $D_{\text{school}}(i)$ and $D_{\text{library}}(i)$ are the Euclidean distances to these facilities, respectively. The same decay coefficient of b_1 in equation (2) is used, since these facilities are mainly accessed by roads. The same weight (1/2) is also used for these two variables.

3.2 *ABM component*

This study assumes that land-development patterns are affected by three types of agents—government agents, developer agents, and resident agents. Government agents have no location attributes, since their influences are uniform for the whole region. It is also difficult to define the exact locations for developer agents. The main objective of developer agents is to make the profit as high as possible. Resident agents are movable, and their decisions to reside in a place can influence land-development patterns. The resident agents are randomly located in the initial stage. They can move into a place for residency according to their financial status and the site attributes. However, they do not actually move around the landscape with every time step for reducing computation time.

3.2.1 Implementing the initiatives of sustainable development by government agents. The strategies of sustainable development can help to develop methods on how to grow with harmony with the environment (Markandya and Richardson 1992). Some principles related to sustainable development can be incorporated in formulating land-development plans. In this model, these principles are defined as follows:

- Land demand is a factor for promoting regional economic development. However, a mechanism is required to ensure the proper distribution of land consumption at different planning stages.
- Land development should avoid the use of good-quality agricultural land as much as possible. This can be realized by incorporating the criterion of spatial efficiency.
- Negotiations are necessary to achieve practical solutions to land-use conflicts.

In this model, government agents will consider spatial and temporal efficiencies in using land resources. The first step is to incorporate the criterion of spatial efficiency for government agents. Government agents will decide if an application for land development is successful or not, according to a number of factors. Existing land use is a major factor in determining land-use conversion. Different land uses will have different values of approval probability for land development. For example, land development is not allowed in ecological sensitive areas. The probability for land development in wetland areas or mountainous areas is much lower. The approval probability is also related to existing plan schemes. It is more likely that an application can be approved if there are no conflicts with existing land-use plans. In this study, the approval probability for government agents is defined to represent various planning objectives.

The second step is to implement the equity of using land resources in a temporal dimension by government agents. The temporal efficiency criterion is to produce the maximum benefits from the use of land resources across generations. Tietenberg (1992) proposes a method to realize efficient allocation of depletable resources and maintain the equity between generations in a time dimension. It assumes that the demand curve for a depletable resource is linear and stable over time (figure 3). Thus, the inverse demand curve in year t can be written as follows:

$$D_t = a - bq_t \tag{5}$$

where a and b are the intersect and slope of the curve of the marginal benefit, respectively, and q_t is the proposed amount of resource consumed in each period t.

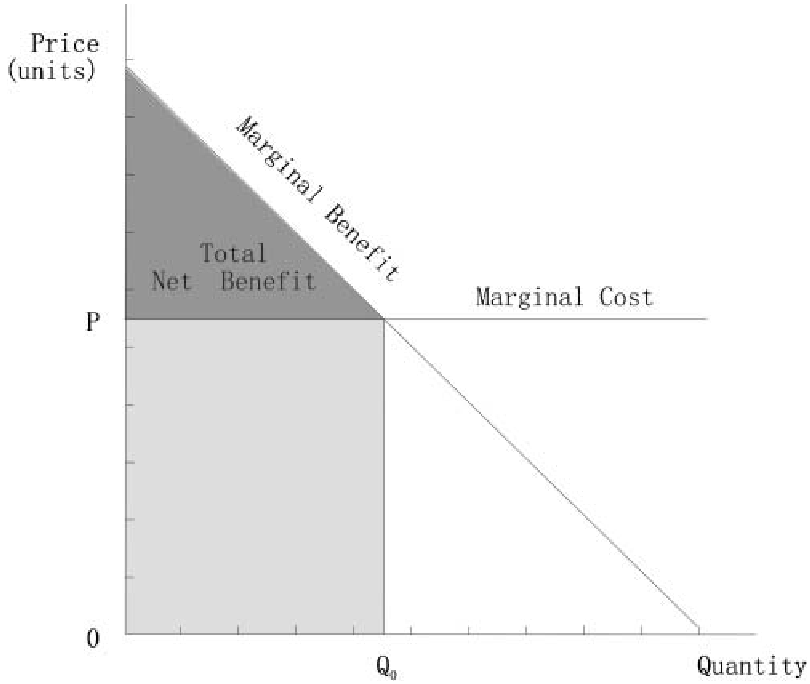

Figure 3. Maximizing the total net benefit derived from the use of land resources.

Then, the total benefit B_T from extracting an amount q_t in year t is the integral of equation (5):

$$\begin{aligned} B_T &= I(a-bq_t)\,\mathrm{d}q_t \\ &= aq_t - bq_t^2/2. \end{aligned} \tag{6}$$

The marginal cost of extracting that resource is further assumed to be a constant c. The total cost C_T of extracting the amount q_t is:

$$C_T = cq_t \tag{7}$$

where c is a constant.

Then, the efficient allocation of a resource over n years should satisfy the following maximization condition (Tietenberg 1992):

$$\operatorname*{Max}_{q_t} \sum_{t=1}^{n} \left(aq_t - bq_t^2/2 - cq_t\right) \Big/ (1+r)^{t-1} + \lambda \left[Q - \sum_{t=1}^{n} q_t \right] \tag{8}$$

where Q is the total available amount of the resource supplied, and r is the discount rate.

When the factor of population growth is considered, the maximization is revised by solving the following equations (Yeh and Li 1998):

$$\begin{aligned} &(a - bq_t/P_{ta} - c) \Big/ (1+r)^{t-1} - \lambda = 0 \quad t = 1, \cdots, n \\ &Q - \sum_{T=1}^{n} q_t = 0 \end{aligned} \tag{9}$$

where Q is the total available amount of land resource supplied. P_{ta} is the projected additional population in period t.

3.2.2 Making profits for developer agents. The main objective of property developers is to achieve a certain amount of profit above expectations. The following equation is used for the assessment of development potentials:

$$D^t_{\text{profit}}(i) = H^t_{\text{price}}(i) - L^t_{\text{price}}(i) - D^t_{\text{cost}}(i) \tag{10}$$

where $D^t_{\text{profit}}(i)$ represents the investment profit at location i, $H^t_{\text{price}}(i)$ is the housing price, $L^t_{\text{price}}(i)$ is the land price, and $D^t_{\text{cost}}(i)$ is the development cost.

The development probability related to developer agents can thus be represented as follows:

$$P^t_{\text{developer}}(k,i) = \frac{D^t_{\text{profit}}(i) - D_{\text{tprofit}}}{D_{\text{mprofit}} - D_{\text{tprofit}}} \tag{11}$$

where $P^t_{\text{developer}}(k,i)$ is the development probability related to developer agents, D_{tprofit} is a threshold value, and D_{mprofit} is the maximum value of the investment profit.

3.2.3 Location choice by resident agents. The behaviours of resident agents are determined by two types of factors: the location factors and agents' status factors (e.g. income and family size). These factors are reflected in a combined utility function, which is defined to assess the value of residency of each site for a resident agent. The main objective of resident agents is to maximize the following utility function as much as possible in site selection. This combined utility function of location (i) for agent k can be represented as follows:

$$\begin{aligned} U(k,i) = {} & w_{\text{price}} \cdot B_{\text{price}}(i) + w_{\text{env}} \cdot B_{\text{env}}(i) \\ & + w_{\text{access}} \cdot B_{\text{access}}(i) + w_{\text{facil}} \cdot B_{\text{facil}}(i) + w_{\text{edu}} \cdot B_{\text{edu}}(i) + \varepsilon_{\text{tij}} \end{aligned} \tag{12}$$

where $w_{\text{price}} + w_{\text{env}} + w_{\text{access}} + w_{\text{facil}} + w_{\text{edu}} = 1$; the variables of $B_{\text{price}}(i)$, $B_{\text{env}}(i)$, $B_{\text{access}}(i)$, $B_{\text{facil}}(i)$, and $B_{\text{edu}}(i)$ are the utilities (benefits) related to land price, surrounding environment, accessibility, general facilities, and education for the development of location (i); the parameters of w_{price}, w_{env}, w_{access}, and w_{edu} are the preferences (weights) for these variables, respectively; and the term of ε_{tij} is a stochastic variable which accounts for unexplained factors in site selection.

These weights are dependent on agents' status, such as income and family size. In this study, resident agents are classified into a number of categories according to their attributes, such as income and family size. The weights are then determined for each group of agents according to Saaty's pairwise comparison procedure (Eastman 1999). The heterogeneity of resident agents is reflected by the weights in this combined utility function.

The probability of selecting a site is estimated according to the utility function. For resident k, the probability of location (i) to be selected is equal to the utility probability that the utility value at that location is greater than or equal to those at other locations (McFadden 1978):

$$P^t_{\text{resident}}(k,i) = P(U(k,i) \geq U(k,i')) = \frac{\exp(U(k,i))}{\sum_k \exp(U(k,i))} \tag{13}$$

3.2.4 Interactions between government agents, developer agents and resident agents. Although the initial approval probability is determined by governments, it is subject to changes with the influences from residents and property developers. The following equation can be used to represent this type of interaction between government agents, developer agents, and resident agents in affecting the development probability of a cell (*i*):

$$P^t_{\mathrm{gov}}(i) = P^{t-1}_{\mathrm{gov}}(i) + g \cdot \Delta P_1 + + h \cdot \Delta P_2 \ \left(\text{if } P^t_{\mathrm{gov}}(i) > 1, \text{ then } P^t_{\mathrm{gov}}(i) = 1\right) \tag{14}$$

where the initial value of $P^t_{\mathrm{gov}}(i)$ is $P^0_{\mathrm{gov}}(i)$, which is related to land-use types; the coefficients g and h are the total numbers applied for development at cell (*i*) by resident agents and developer agents, respectively; and ΔP_1 and ΔP_2 are the incremental probability for each application by developer agents and resident agents, respectively.

3.3 *Integrating CA with ABM*

CA are an important component in this integrated model. In this study, the development probability related to the local interactions of physical factors is estimated using a logistic-CA model (Wu 2002):

$$P^t_{\mathrm{ca}}(i) = \frac{1}{1 + \exp\left[-\left(D_0 + \sum_h D_h \cdot x_h(i)\right)\right]} \cdot \mathrm{con}^t(i) \cdot \Omega^t(i) \tag{15}$$

where $P^t_{\mathrm{ca}}(i)$ is the development probability of location (*i*), determined by the neighbourhood function, x_h is the hth spatial variable, D_0 is a constant, and D_h is the weight of the hth variable. The function of *con(i)* is a combined physical constraint, and $\Omega(i)$ is the percentage of developed cells in the neighbourhood.

The final decision is made according to a joint development probability, which reflects the combined effects of human factors (government agents, resident agents, and developer agents) and environmental factors. The joint probability is represented as follows:

$$P^t_i = A \cdot P^t_{\mathrm{resident}}(k,i) \cdot P^t_{\mathrm{developer}}(k,i) \cdot P^t_{\mathrm{gov}}(i) \cdot P^t_{\mathrm{ca}}(i) \tag{16}$$

where A is an adjusted coefficient.

The Monte Carlo method is used to determine the final selection of a location for development (Wu and Webster 1998). The final land-use conversion is determined by comparing the development probability with a random variable:

$$S_{t+1}(i) = \begin{cases} \text{Development}, & P^t_i > \mathrm{Rand}() \\ \text{Non} - \text{development}, & \text{Others} \end{cases} \tag{17}$$

where Rand() is a random variable ranging from 0 to 1.

This simulation is to determine which sites will be developed based on the combined assessment from various individuals. The final decision is based on the joint probability calculated by equation (16). This equation consists of four components of interactions for determining land-use conversion. The first three components are obtained by the ABM method, and the last component is obtained by the CA method.

4. Model implementation and results

4.1 *Programming*

The prototype of this integrated model is developed using the Visual Basic and ArcObjects component of ARCGIS. The use of ArcObjects can allow this model to access the spatial data in a GIS database directly. The computation will be too intensive if the model is implemented using common pure agent-modelling shells (e.g. the swarm package). They have difficulties in coupling with GIS and CA directly. Figure 4 shows the interface of this proposed prototype. Agents are only implied through model results for simplicity.

4.2 *Preparing spatial variables and determining model coefficients*

The original layers of land-use types, land price, living environment, accessibility, general public facilities, and education were transformed into a raster format with the resolution of 100×100 m for the programming. The utility (benefit) of each spatial variable for land development was estimated using GIS.

Some coefficients should be estimated before calculating the utilities in equations (2)–(4). The values of b_1, b_2, and b_3 in equation (2), which are related to transport conditions, can be estimated according to empirical traffic data. It is assumed that a transport tool with a larger traffic density will have a larger area of influence (a smaller value for the decay coefficient). For example, expressways which have larger traffic densities will be assigned smaller values for the coefficient. The following equation can be used to represent this relationship:

$$b_1/b_2 = f_{\text{express}}/f_{\text{road}} \tag{18}$$

where f_{road} and f_{express} are the average traffic densities for roads and expressways, respectively.

The same method can be applied to the estimation of b_3. If there are z_1 number of roads and z_2 number of expressways connected to urban centres, the equation becomes:

$$b_1/b_3 = \left(z_1 \cdot f_{\text{road}} + z_2 \cdot f_{\text{express}}\right)/f_{\text{road}} \tag{19}$$

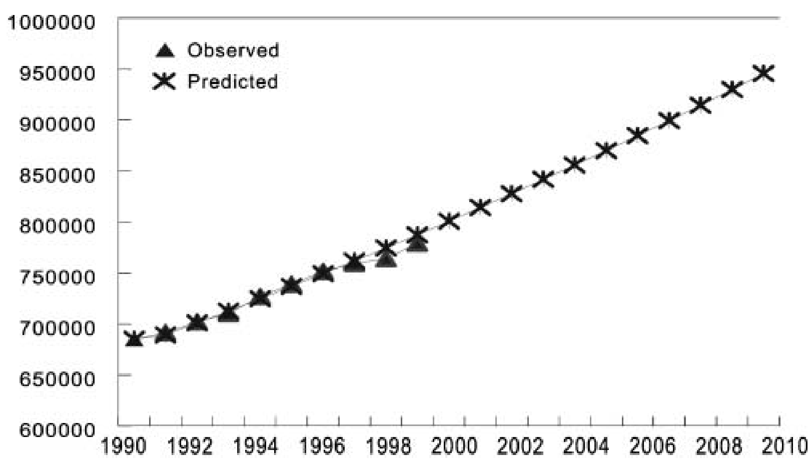

Figure 4. Estimating population growth of Guangzhou in 1990–2010.

Table 1 provides estimates of f_{road} and f_{express} according to statistical data. When b_1 is set to 0.00100, b_2 and b_3 become 0.00023 and 0.000125, respectively, according to the above equations. The same value of b_1 is also used in equations (3) and (4).

The original spatial variables were normalized into the range of [0, 1] before they were used for the calculation. The values of the incremental probabilities, ΔP_1 and ΔP_2, were decided by experiments. In this study, ΔP_1 was set to 0.005 and ΔP_2 was set to 0.1. The coefficients of the CA component were calibrated according to logistic regression (Wu 2002). Landsat TM images dated on 30 December 1995 and 13 June 2004 were used to obtain training data about actual land-use conversion. Table 2 lists the coefficients of the logistic-CA model in equation (15) based on the regression analysis.

CA and ABM are based on discrete time steps in simulating urban dynamics. Too few time steps will neglect local interactions and cannot allow spatial details to emerge (Yeh and Li 2006). An increase in the number of time steps can help to generate more accurate simulation results. In many applications, 200–300 time steps are required to guarantee sufficient temporal accuracy for simulation (Yeh and Li 2006). In this study, 1500 time steps are adopted to simulate land development in the period of 1995–2010. Therefore, 100 time steps correspond to 1 year in the simulation.

4.3 *Implementing sustainable use of land resources by government agents*

The initiatives for sustainable use of land resources should be implemented by government agents, who determine the proper distribution of land consumption across different planning periods. The first step is to estimate the population growth before the appropriate land consumption can be obtained for each planning period. A regression model was established for estimating the population growth using empirical data (table 3):

$$Y(T+1) = 673300.46e^{0.0166T} \tag{20}$$

where $Y(T+1)$ is the predicted population in year $T+1$ based on the initial population in 1990.

Table 1. Traffic densities of roads and expressways according to statistical data[a].

Transport types	Length (km)	Traffic (1000 persons)	Average Traffic Density (1000 persons km^{-1})
Roads	4637.2	147 330	31.77
Expressways	382.8	52 310	136.7

[a]Sources: Guangdong Statistical Yearbook (2000).

Table 2. Coefficients of the logistic-CA model.

	D_1	D_2	D_3	D_4	D_5
D0	Distance to main centres	Distance to sub-centres	Distance to main roads	Distance to roads	Distance to expressways
0.625	−0.002	0.005	−0.009	−0.006	0.002

Table 3. Empirical data about the population growth in the Haizhu district of Guangzhou in 1990–1999.

Year	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999
Population	684 887	691 557	702 004	710 153	727 045	738 910	751 486	759 256	763 959	778 984

Figure 5 indicates that the regression model can predict the population growth satisfactorily. The predicted population was used to calculate the optimized land consumption for each period according to the modified Tietenberg model. The available land for development was estimated according to land-use information. The whole study area was 101.40 km^2, of which the urban area made up 34.09 km^2 in 1995. There was only 66.38% left for future land supply since 33.62% of the area had been urbanized. The detailed provision of land consumption at each period in 1995–2010 was obtained by using equation (9), assuming that 50% of the area could be

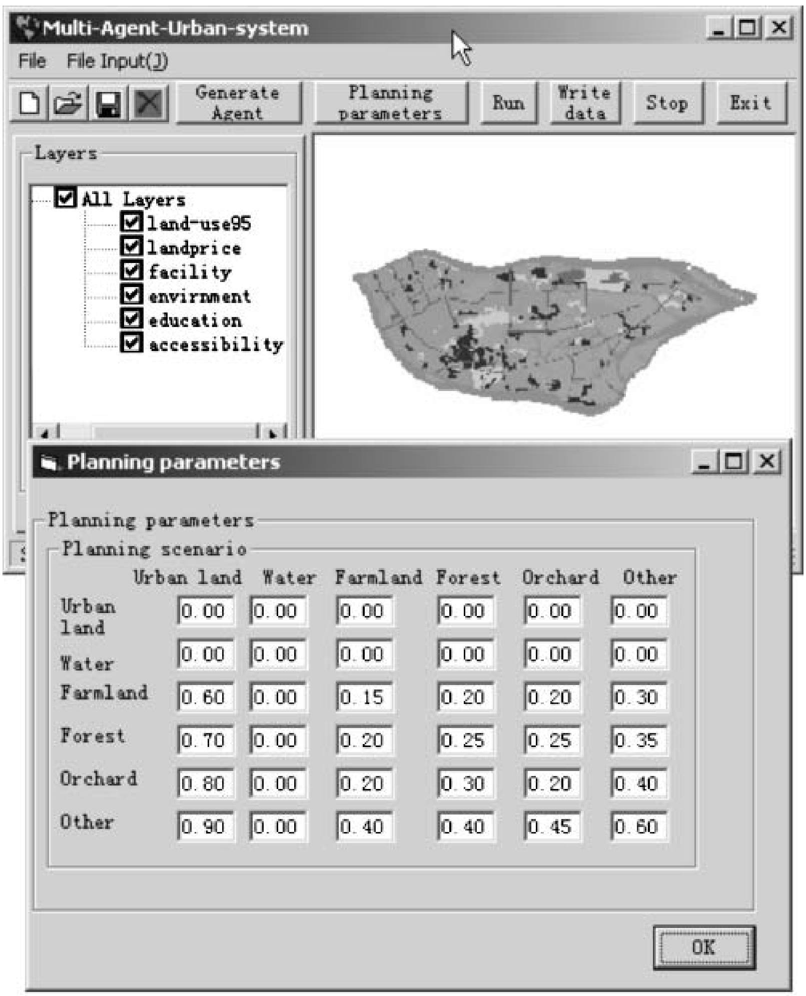

Figure 5. Interface of the proposed model.

Table 4. Optimal land consumption for different periods with various discount rates according to the Tietenberg's model.

Year	Population growth	Land consumption (km^2)		
		r=0	r=0.02	r=0.1
1995–2000	63 880	5.08	6.04	7.66
2000–2005	69 420	5.52	5.39	5.08
2005–2010	75 439	6.01	5.18	3.87

urbanized in 2010 (table 4). The number of new urban cells was then determined for the simulation.

4.4 *Defining resident agents' properties using empirical data*

The decision behaviours of resident agents are defined using aggregated census data because of the lack of detailed information. Some simplification procedures have to be carried out for obtaining the attributes of resident agents. First, resident agents should be classified into a few categories so that their properties can be heuristically defined. The attributes for the aggregated agents are obtained using social and economic data. This study considers two major attributes, income and household size, which are obtained from the statistical yearbook of Guangzhou in 2004, and the *Fifth National Census*, respectively.

Residents can be classified into three groups by their income—low-income class (income<9600 RMB $year^{-1}$), middle-income class (9600 RMB $year^{-1}$<income< 60 000 RMB $year^{-1}$), and high-income class (income>60 000 RMB $year^{-1}$). (1 US$ is roughly equivalent to 7.6951 RMB as of 8 May 2007.) They can also be classified into two groups by household size: without children and with children. Six classes of residents were obtained using these two attributes. The actual percentages for these six groups were calculated according to the statistical yearbook of Guangzhou in 2004, and the *Fifth National Census* (table 5). These percentages were used to create the actual numbers for various groups of resident agents in the simulation.

Each group of resident agents has distinct behaviours or preferences in the location choice of residency. In this model, their preferences are reflected by the weights in the utility function as described in equation (12). The weights were obtained by using Saaty's pairwise comparison procedure (Eastman 1999). The comparison was mainly based on experts' knowledge and preferences. A higher value of the weight means that the variable will be treated more importantly. A matrix can be constructed to indicate the relative importance based on the comparison. Saaty (1990) proposes a consistency ratio (CR) to examine the consistency of the matrix. He suggests that the matrix should be re-evaluated if

Table 5. Proportion of each group of resident agents.

	Types of resident agent					
Household size	Without children			With children		
Income	Low income	Middle income	High income	Low income	Middle income	High income
Proportion (%)	9	39	9	6	31	6

Table 6. Weights for different groups of resident agents obtained using Saaty's method.

Types of residents	Weights					Total	CR
	Land price	Surrounding environment	Accessibility	Public facilities	Education		
Low income without children	0.443	0.093	0.206	0.155	0.103	1	0.042
Low income with children	0.401	0.081	0.154	0.081	0.283	1	0.087
Middle income without children	0.175	0.379	0.165	0.194	0.087	1	0.057
Middle income with children	0.220	0.276	0.142	0.140	0.222	1	0.094
High income without children	0.048	0.526	0.194	0.141	0.091	1	0.072
High income with children	0.084	0.434	0.171	0.076	0.235	1	0.064

the ratio value is greater than 0.10. Table 6 shows the results of the weights derived from Saaty's method.

4.5 *Generating planning scenarios*

This simulation assumes that each new urbanized cell can accommodate one resident agent. The total number of resident agents was determined according to the allowed amount of land consumption. The discount rate r was set to 0.1 for calculating the optimal distribution of land consumption. The detailed procedures for generating development alternatives are as follows:

1. Determining the total number of resident agents according to the allowed amount of land consumption for each planning period.
2. Using the Monto Carlo method to create a resident agent according to the proportion of various types of residents based on census data (table 3).
3. The development probability related to the local interactions of physical factors is estimated by the logistic-CA model, which is calibrated using the classified satellite images.
4. Using equation (12) and table 6 to compute the utility function for this resident agent.
5. Selecting the locations with the highest utility values and estimating development probability for these places according to the interactions described in equation (16).
6. Determining whether the locations of the highest combined probability values will be developed using the Monto Carlo method. If yes, the location will be marked and go to step 2 to create a new resident agent. If no, the next site of the second highest utility value will be evaluated until this existing agent has been accommodated.
7. This procedure continues until all the required resident agents have been accommodated.

This model was used to simulate both baseline development scenarios and planning development scenarios. The baseline scenarios were generated according to the development trend. Planning scenarios were produced by incorporating the initiatives of sustainable development in the modelling. The intervention from government agents is crucial for producing planning scenarios instead of baseline scenarios. In this study, the intervention was first represented by using the initial pre-defined approval probability (P^0_{gov}) for government agents. There are five regimes of land development for the simulation, as follows.

4.5.1 Baseline scenario. This simulation is based on the trajectory of past development. This regime assumes that no sustainable development strategies are adopted to regulate existing development trends. No spatial and temporal efficiencies are implemented in this simulation. Table 7 lists the initial pre-defined approval probability (P^0_{gov}) from government agents for this regime. This simulation can generate the scenario provided that the city continues to develop without any constraints. Urban planners can compare this baseline scenario with the following planning scenarios.

4.5.2 Planning scenario 1: compact development. Some government intervention is implemented by controlling land consumption in the spatio-temporal dimension. The equity of using land resources is emphasized by properly arranging land-use conversion at each planning stage according to equation (9). This planning scenario also adopts a high priority on implementing the spatial efficiency in terms of compact development by using the initial pre-defined approval probability (P^0_{gov}) (table 8). Land-use types will not impose restrictions on land development so that compact patterns can be formulated in the simulation.

Table 7. Initial pre-defined approval probability (P^0_{gov}) for simulating baseline development.

	Planning					
Existing	Urban land	Water	Farmland	Forest	Orchard	Other
Urban land	1.00	0.00	0.00	0.00	0.00	0.00
Water	0.02	0.00	0.00	0.00	0.00	0.00
Farmland	0.60	0.00	0.15	0.20	0.20	0.30
Forest	0.65	0.00	0.20	0.25	0.25	0.35
Orchard	0.68	0.00	0.20	0.30	0.20	0.40
Other	0.90	0.00	0.40	0.40	0.45	0.60

Table 8. Initial pre-defined approval probability (P^0_{gov}) for simulating planning scenario 1: compact development.

	Planning					
Existing	Urban land	Water	Farmland	Forest	Orchard	Other
Urban land	1.00	0.00	0.00	0.00	0.00	0.00
Water	1.00	0.00	0.00	0.00	0.00	0.00
Farmland	1.00	0.00	0.00	0.00	0.00	0.00
Forest	1.00	0.00	0.00	0.00	0.00	0.00
Orchard	1.00	0.00	0.00	0.00	0.00	0.00
Other	1.00	0.00	0.00	0.00	0.00	0.00

4.5.3 Planning scenario 2: farmland protecting development. This regime is to ensure the equity of using land resources across generations and avoid the encroachment on agricultural land as well. The former is realized by arranging the proper land consumption at different planning stages, and the latter is implemented by using the initial pre-defined approval probability (P^0_{gov}) for government agents (table 9). A very small probability will be given to the development of agricultural land. This planning scenario can allow land resources to be used more efficiently than the baseline pattern.

4.5.4 Planning scenario 3: green-land protecting development. This regime pays special attention to implementing the concept of 'garden cities' while land consumption is also constrained by the equity criterion. It addresses the growing concern for a better living environment after residents have secured their basic housing demand. This is a further development stage compared with planning scenario 2. Table 10 shows the initial pre-defined approval probability (P^0_{gov}) which imposes extreme restrictions on converting green land and orchard land into residential use.

4.5.5 Planning scenario 4: housing-demand development. This scenario just completely satisfies housing demand from resident agents at each location without government controls. Land-use types do not impose any restrictions on land development. Therefore, all the initial probability values are set to 1 for this regime.

Figure 6 is the outcome from the simulation of baseline patterns in the study area in 2000–2010 according to historical growth. Planning scenarios can be simulated by

Table 9. Initial pre-defined approval probability (P^0_{gov}) for simulating planning scenario 2: farmland protecting development.

	Planning					
Existing	Urban land	Water	Farmland	Forest	Orchard	Other
Urban land	0.00	0.00	0.00	0.00	0.00	0.00
Water	0.00	0.00	0.00	0.00	0.00	0.00
Farmland	0.10	0.00	0.00	0.05	0.05	0.10
Forest	0.75	0.00	0.25	0.30	0.30	0.35
Orchard	0.80	0.00	0.05	0.35	0.25	0.40
Other	0.95	0.00	0.15	0.40	0.45	0.70

Table 10. Initial pre-defined approval probability (P^0_{gov}) for simulating planning scenario 3: green-land protecting development.

	Planning					
Existing	Urban land	Water	Farmland	Forest	Orchard	Other
Urban land	0.00	0.00	0.00	0.00	0.00	0.00
Water	0.00	0.00	0.00	0.00	0.00	0.00
Farmland	0.65	0.00	0.35	0.20	0.20	0.50
Forest	0.10	0.00	0.05	0.00	0.00	0.10
Orchard	0.15	0.00	0.05	0.00	0.00	0.15
Other	0.95	0.00	0.50	0.30	0.35	0.70

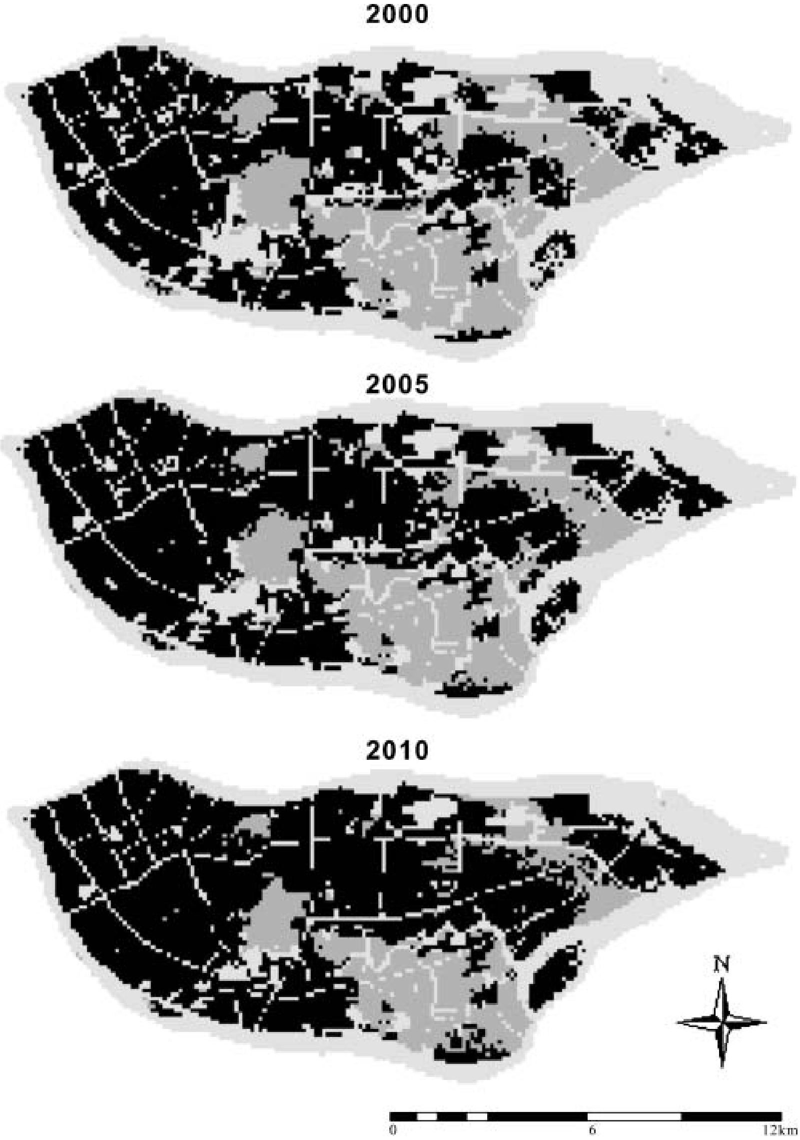

Figure 6. Simulation of baseline development patterns of Guangzhou based on historical trends.

incorporating the criteria of sustainable development and properly modifying the parameters of this agent-based model. Figure 7(*a*) is to simulate compact development, which can reduce the energy consumption in transportation. It is also able to reduce the encroachment on agricultural land by introducing government intervention (figure 7(*b*)). However, this scenario may result in some fragmented patterns. Planning scenario 3 (green-land protecting development) emphasizes the preservation of green land and orchard land (figure 7(*c*)). Planning scenario 4 (housing-demand development) is associated with significant dispersed development patterns, since it just satisfies housing demand from resident agents (figure 7(*d*)).

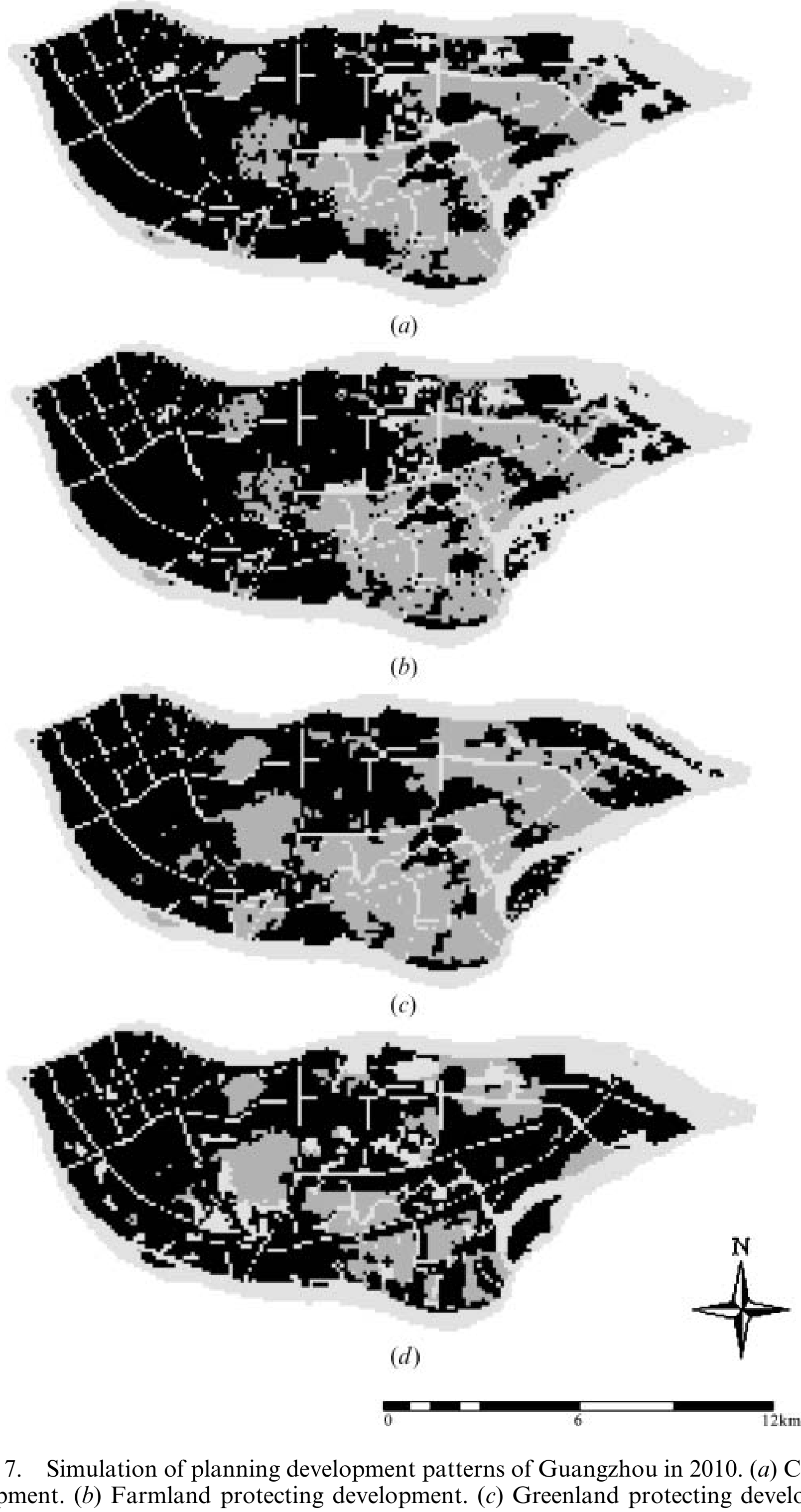

Figure 7. Simulation of planning development patterns of Guangzhou in 2010. (*a*) Compact development. (*b*) Farmland protecting development. (*c*) Greenland protecting development. (*d*) Housing-demand development.

4.6 *Metrics for the comparisons among the simulated scenarios*

Statistical comparisons among the simulated scenarios were carried out for providing planning implications according to a number of metrics. These metrics, which will indicate the gain and loss of land development, include the indicators of compactness, development suitability gain, agricultural suitability loss, green-land loss, and farmland loss. The first two indicators are related to the gain, and the last three are related to the loss of land development.

The compactness of land development can be calculated according to the average comparison between the perimeter of a developed cluster and the standard perimeter of the circle which has the same area (Li and Yeh 2004). The index can be represented using the following equation:

$$\mathrm{CI} = \sqrt{\sum_j S_j} \Big/ \sum_j P_j \tag{21}$$

where CI is the value of the compactness index, and S_j and P_i are the area and perimeter of the developed cluster (polygon) j_j. It is obvious that land development with average narrow shapes or dispersed development patterns will have low values for the index.

The developed sites should have higher values of development suitability and lower values of agricultural suitability for spatial efficiency (Li 2005). Therefore, the development suitability gain can be obtained by summing up the urban development suitability for all the developed cells:

$$D_{\mathrm{gain}} = \sum_i S_{\mathrm{ur}}(i) \tag{22}$$

where D_{gain} is the development suitability gain, and $S_{\mathrm{ur}}(i)$ is the urban development suitability at cell i where land development takes place.

The agricultural suitability loss can be calculated using this similar method:

$$A_{\mathrm{loss}} = \sum_i S_{\mathrm{ag}}(i) \tag{23}$$

where A_{loss} is the agricultural suitability loss, and $S_{\mathrm{ag}}(i)$ is the agricultural suitability at cell i where land development takes place.

The last two indicators are to sum up the total amounts of green-land loss and farmland loss for each scenario. Table 11 is the analysis results from these five metrics. Figure 8 shows the gain of land development for these simulated scenarios. The baseline scenario and planning scenario 4 (housing-demand development) have lower values of the compactness, although they have higher values of development suitability gain.

Figure 9 further displays the loss related to land development for these scenarios. The baseline scenario and planning scenario 4 (housing-demand development) have larger values for the indicators of agricultural suitability loss, green-land loss, and farmland loss. Planning scenario 1 (compact development) and planning scenario 3 (green-land protecting development) have lower values for the green-land loss. Planning scenario 2 (farmland protecting development) has the lowest value for farmland loss.

The final assessment of these scenarios is based on a linear combination of these five indicators. The values from these five indicators should be normalized into the

Table 11. Comparisons among the simulated scenarios using various metrics.

Development patterns	Year	Compactness ($\times 10^{-3}$)	Development suitability gain ($\times 10^3$)	Agricultural suitability loss ($\times 10^3$)	Green-land loss ($\times 10^6 m^2$)	Farmland loss ($\times 10^6 m^2$)
Baseline scenario	2000	19.1	46.3	72.5	1.4	1.9
	2005	19.5	73.7	109.2	4.9	3.1
	2010	19.7	91.2	135.8	6.5	4.9
Planning scenario 1: compact development	2000	20.1	44.3	60.1	1.0	1.2
	2005	22.3	71.9	95.7	2.2	2.5
	2010	23.4	89.6	106.5	3.0	3.6
Planning scenario 2: farmland protecting development	2000	19.9	41.4	57.1	2.9	0.5
	2005	21.2	68.4	89.8	4.7	0.8
	2010	21.0	85.2	99.6	6.2	1.1
Planning scenario 3: green-land protecting development	2000	20.2	39.1	64.8	1.2	0.9
	2005	22.4	66.4	103.5	2.5	1.5
	2010	23.8	82.4	115.8	3.4	2.2
Planning scenario 4: housing-demand development	2000	18.7	54.9	79.0	3.0	3.1
	2005	18.9	81.4	112.0	7.2	4.3
	2010	19.3	102.4	148.6	10.5	5.2

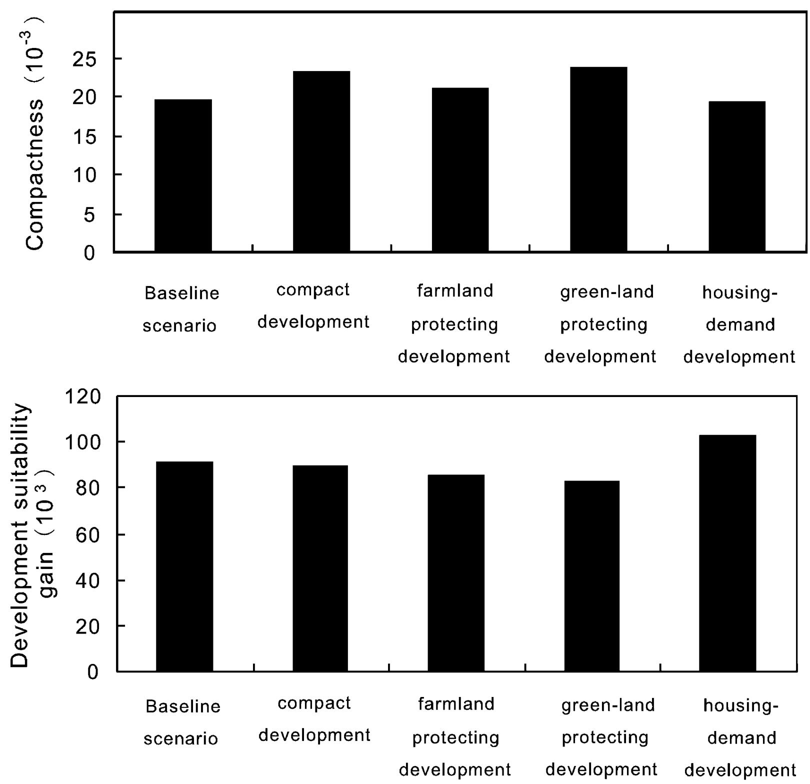

Figure 8. Gain of the simulation scenarios.

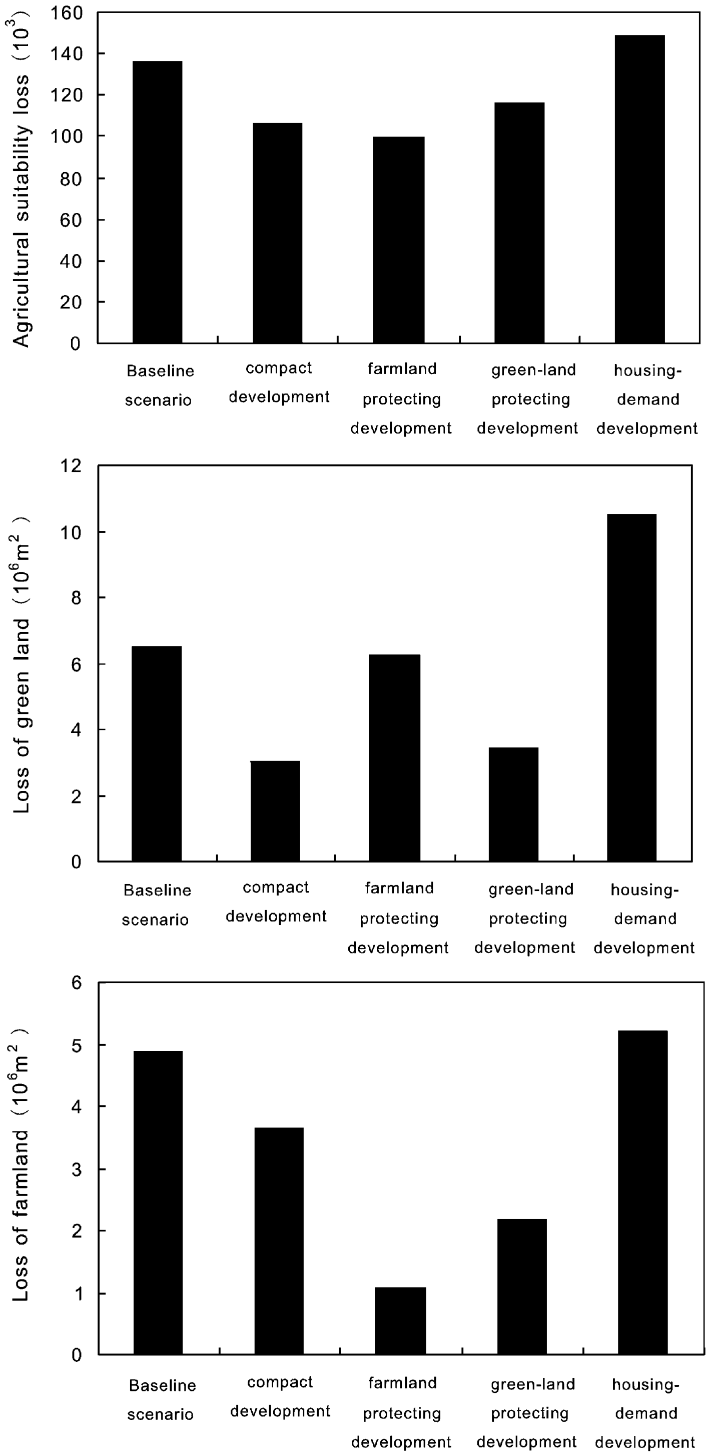

Figure 9. Loss of the simulation scenarios.

range of 0–1 before the use of the linear combination. The normalization is different for these two types of factors: gain (higher scores are better; e.g. compactness and development suitability gain) and loss (lower scores are better, e.g. agricultural suitability loss, green-land loss, farmland loss).

The gain factors are as follows:

$$x' = \frac{x - \mathrm{Min}}{\mathrm{Max} - \mathrm{Min}} \tag{24}$$

The loss factors are as follows:

$$x' = \frac{\mathrm{Max} - x}{\mathrm{Max} - \mathrm{Min}} \tag{25}$$

where x is the original data, Max and Min are the maximum and minimum values, x' is the normalized value.

Figure 10 shows the final result from the linear combination of these five normalized metrics. The same weight is applied to each indicator, since there is no prior knowledge. It is also clear that the baseline scenario and planning scenario 4 (housing-demand development) have a poorer performance from the assessment. Planning scenario 1 (compact development) and planning scenario 3 (green-land protecting development) have a better performance according to the combined indicator.

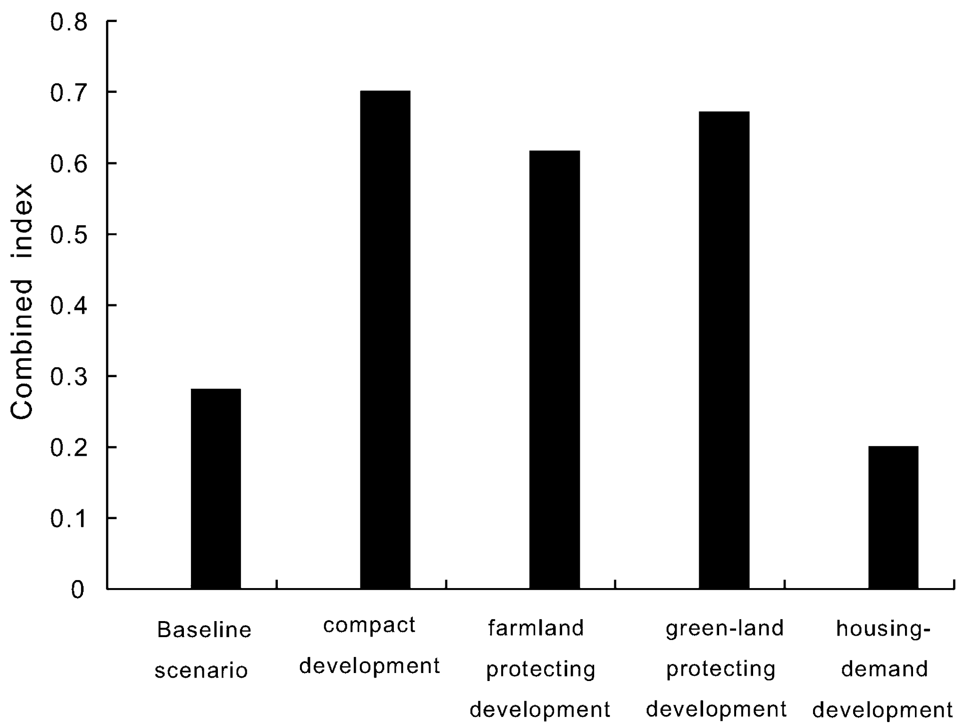

Figure 10. Final assessment of the simulated scenarios using a linear combination of five d metrics.

5. Conclusion

This paper has demonstrated that agent-based modelling techniques can be further extended to the simulation of development alternatives. The strategies for sustainable development are incorporated in the modelling by properly defining agents' behaviours. Spatial efficiency of using land resources is implemented by selecting suitable sites for development according to planning objectives. The efficient allocation of land resources over the temporal dimension is realized by defining decision behaviours of government agents.

Sustainable land development is a complex issue which involves negotiations and compromises of various stakeholders. Local interactions from this integrated model are essential for dealing with these complex situations. In this study, the heterogeneity of agents is reflected by using different sets of weights according to GIS data. Development plans can be generated to implement sustainable development initiatives through the interactions between government agents, developer agents, and resident agents. Since several compromises have been adopted in the simulation, the simulated alternatives should be more realistic and practical for planning practice.

Five scenarios of land development have been simulated by using this proposed model. The effects of these scenarios are compared according to a number of metrics, such as compactness, development suitability gain, agricultural suitability loss, green-land loss, and farmland loss. These metrics are used to quantify the gain and loss of land development by providing planning implications. The comparison can identify the best scenario for a certain planning objective. For example, the baseline scenario and planning scenario 4 (housing-demand development) have lower values for gain, and larger values for loss. However, planning scenario 2 (farmland protecting development) has the lowest values of farmland loss.

A combined index can be devised to take account of all these five metrics. This index also indicates that the baseline scenario and planning scenario 4 (housing-demand development) have a poorer performance, whereas planning scenario 1 (compact development) and planning scenario 3 (green-land protecting development) have a better performance for land development.

Like other agent-based models, this model also involves many parameters that have critical effects on the simulation results. Although some calibration procedures have been carried out, a finer tuning of this model still requires considerable effort. Further studies should be carried out in developing the methods of calibrating agents' properties in a more consistent way. There is still a general lack of detailed spatial information that can be used to define the decision behaviours of agents for Chinese cities. More detailed resident agents can be defined to improve simulation performance when such spatial information is available.

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 PhD development program from the Ministry of Education of China (Project No. 20040558023).

References

BASU, N. and PRYOR, R.J., 1997, *Growing a Market Economy*, Technical Report SAND-97-2093 (Albuquerque, NM: Sandia National Laboratories).

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

COURDIER, R., GUERRIN, F., ANDRIAMASINORO, F.H. and PAILLAT, J.M., 2002, Agent-based simulation of complex systems: application to collective management of animal wastes. *Journal of Artificial Societies and Social Simulation*, **5**. Available online at: http://jasss.soc.surrey.ac.uk/5/3/courdier.html (accessed 12 January 2006).

EASTMAN, J.R., 1999, Multi-criteria evaluation and GIS. In *Geographical Information Systems*, P.A. Longley, M.F. Goodchild, D.J. Maguire and D.W. Rhind (Eds), pp. 493–502 (New York: Wiley).

GILBERT, N. and CONTE, R., 1995, *Artificial Societies: the computer simulation of social life*. UCL Press.

HENTON, D. and STUDWELL, K., 2000, *Informed Regional Choices: How California's Regional Organizations are Applying Planning and Decision Tools* (Oakland, CA: California Center for Regional Leadership).

JANTZ, P., GOETZ, S. and JANTZ, C., 2005, Urbanization and the loss of resource lands in the Chesapeake bay watershed. *Environmental Management*, **36**, pp. 808–825.

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**, pp. 131–152.

LI, X. and YEH, A.G.O., 2001, Zoning land for agricultural protection by the integration of remote sensing, GIS and cellular automata. *Photogrammetric Engineering & Remote Sensing*, **67**, pp. 471–477.

LI, X. and YEH, A.G.O., 2004, Analyzing spatial restructuring of land use patterns in a fast growing region using remote sensing and GIS. *Landscape and Urban Planning*, **69**, pp. 335–354.

LI, X., 2005, A four-component efficiency index for assessing land development using remote sensing and GIS. *Photogrammetric Engineering & Remote Sensing*, **71**, pp. 47–57.

LOIBL, W. and TOETZER, T., 2003, Modelling growth and densification processes in suburban regions—simulation of landscape transition with spatial agents. *Environmental Modelling & Software*, **18**, pp. 553–563.

MARKANDYA, A. and RICHARDSON, J., 1992, The economics of the environment: an introduction. In *Environmental Economics*, A. Markandya and J. Richardson (Eds), pp. 7–25 (London: Earthscan).

MCFADDEN, D., 1978, Modelling the choice of residential location. In *Spatial Interaction Theory Planning Models*, A. Karlqvist, L. Lundqvist, F. Snickars and J. Weibull (Eds), pp. 75–96 (Amsterdam: North Holland).

NIJKAMP, P., OUWERSLOOT, H. and RIENSTRA, S.A., 1997, Sustainable urban transport systems: an expert-based strategic scenario approach. *Urban Studies*, **34**, pp. 693–712.

SAATY, T.L., 1990, *The Analytic Hierarchy Process: Planning, Priority Setting, Resource Allocation* (Pittsburgh, PA: University of Pittsburgh).

STATISTICAL BUREAU OF GUANGZHOU 2000, Guangzhou Statistical Yearbook (China Statistics Press), pp. 187–190.

TIETENBERG, T., 1992, *Environmental and Natural Resource Economics* (New York: HarperCollins).

TORRENS, P.M. and BENENSON, I., 2005, Geographic automata systems. *International Journal of Geographical Information Science*, **19**, pp. 385–412.

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**, pp. 1175–1199.

WU, F., 2002, Calibration of stochastic cellular automata: the application to rural–urban land conversions. *International Journal of Geographical Information Science*, **16**, pp. 795–818.

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

YEH, A.G.O. and LI, X., 1998, Sustainable land development model for rapid growth areas using GIS. *International Journal of Geographical Information Science*, **12**, pp. 169–189.

YEH, A.G.O. and LI, X., 1999, Economic development and agricultural land loss in the Pearl River Delta, China. *Habitat International*, **23**, pp. 373–390.

YEH, A.G.O. and LI, X., 2006, Errors and uncertainties in urban cellular automata. *Computers, Environment and Urban Systems*, **30**, pp. 10–28.

ZANDERA, P. and KÄCHELE, H., 1999, Modelling multiple objectives of land use for sustainable development *Agricultural Systems*, **59**, pp. 311–325.