International Journal of Geographical Information Science

ISSN: 1365-8816 (Print) 1362-3087 (Online) Journal homepage: https://www.tandfonline.com/loi/tgis20

Coupling urban cellular automata with ant colony optimization for zoning protected natural areas under a changing landscape

Xia Li , Chunhua Lao , Xiaoping Liu & Yimin Chen

International Journal of Geographical Information Science
Vol. 25, No. 4, April 2011, 575–593

Coupling urban cellular automata with ant colony optimization for zoning protected natural areas under a changing landscape

Xia Li*, Chunhua Lao, Xiaoping Liu and Yimin Chen

School of Geography and Planning, Sun Yat-sen University, Guangzhou, PR China

(*Received 6 December 2009; final version received 24 March 2010*)

Optimal zoning of protected natural areas is important for conserving ecosystems. It is an NP-hard problem which is difficult to solve by using common geographic information system (GIS) functions. Another problem is that existing optimization methods ignore potential land-use dynamics in formulating optimal patterns. This article has developed a new method for solving complicated zoning problems by using ant colony optimization (ACO) techniques. Significant modifications have been made, so that traditional ACO can be extended to the solution of area optimization problems. Two strategies, the single-year coupling strategy and the merging-year coupling strategy, have been proposed to couple urban cellular automata with ACO for zoning protected natural areas under a changing landscape. This proposed method has been tested in the metropolitan region of Guangzhou, China, by using Geographical Simulation and Optimization System (GeoSOS) software. The experiments indicate that the modified ACO can effectively solve this optimization problem without getting stuck in local optima. This method has better performances compared to other traditional methods, such as simulated annealing (SA), iterative relaxation (IR), and density slicing (DS). The use of the best coupling strategy can improve the accumulative utility value of the zoning by 4.3%. Moreover, it is also found that the adoption of the best protection pattern could significantly promote the compactness of future urban forms in the study area.

Keywords: cellular automata; ant colony optimization; area optimization; natural protection; GeoSOS

1. Introduction

The establishment of protected areas can serve the purposes of conservation of species and ecosystem diversity, preservation of ecological processes, and promotion of scientific activities and recreation (Snyder *et al.* 2004, Verdiell *et al.* 2005). The zoning can be started by selecting, for each protection level, the units with the highest suitability up to the fulfillment of the land demand (Geneletti and van Duren 2008). However, the zoning based on the ranking of suitability values without the contiguity constraint will result in the fragmentation of land-use patterns. The importance of spatial patterns for ecological protection and management has been widely discussed in the literature of ecological studies (Fahrig 1998). Reducing site fragmentation may aid the likelihood of successful species dispersal, mitigate the effects of human impact, and facilitate reserve management (Pulliam *et al.* 1992, Matisziw and Murray 2006).

*Corresponding author. Email: lixia@mail.sysu.edu.cn; lixia@graduate.hku.hk

ISSN 1365-8816 print/ISSN 1362-3087 online

DOI: 10.1080/13658816.2010.481262
http://www.informaworld.com

Land-use allocation may involve efficient distribution of activities over feasible sites in order to create the maximum benefits and least costs. These optimization problems involve shape or contiguity constraints as well as site attributes (suitability of each cell). Optimal patch design is a hard geometric problem with a huge complex search space and therefore requires efficient optimization methods (Brookes 1998). Heuristics should be adopted since exact enumeration methods cannot be used to solve such hard combinatorial optimization problems. A common heuristic for solving these zoning problems is based on the well-known simulated annealing (SA) paradigm (Bos 1993, Verdiell *et al.* 2005). For example, Verdiell *et al.* (2005) propose an SA method for creating a protected natural area for ecological purposes. However, these methods are usually applied to the spatial data of coarse resolutions with a limited number of cells. For example, a grid of 900 cells is used for the optimization (Verdiell *et al.* 2005).

Most spatial assessments of environmental features, potential habitats, and conservation management areas incorporate implicitly static ecological and geographic relationships (Halpin 1997). Selecting protected areas based on fixed landscape conditions may be problematic without considering land-use changes. Zoning of natural areas for protection is a challenging task in the regions of rapid growing economy and fast urban expansion.

In China, strict mandates for protecting important ecological land areas have to be implemented for specific regions, such as Dongguan and Shenzhen (Environmental Department of Shenzhen 2008). A major problem is to balance the trade-off between natural protection and economic growth. Urban expansion is an inevitable phenomenon because land consumption is crucial for sustaining economic growth in China. Empirical studies have shown that urban land expands by 3% when the economy, measured by gross domestic product, grows by 10% in China (Deng *et al.* 2008). Land consumption is required to accommodate the growing population shifted from rural to urban and the expanding economic activities (Kuznets 1966). It is impractical to restrict urban development by implementing strict protection plans in these fast-growing regions. Under this situation, zoning should reconcile multiple conflicting interests as rationally and transparently as possible (Carsjens and Van der Knaap 2002, Sante-Riveira 2008).

This study aims to develop a dynamic optimization model for zoning the protected areas under a fast-changing landscape. This kind of optimization may arise in many application areas, such as selecting biological reserves (Cova and Church 2000), hazardous waste sites (Van Zee and Lee 1989), and landfill (Minor and Jacobs 1994). First, conventional ant colony optimization (ACO) will be modified so that it can be suited to the solution of area optimization problems. A utility function is incorporated in this area optimization ACO by addressing the factors of ecological suitability and compactness simultaneously. Then this ACO model will be coupled with an urban cellular automaton (CA) for exploring the protection scenarios under rapid land-use dynamics. The coupling is based on two strategies: the single-year coupling strategy and the merging-year coupling strategy. This proposed method will be tested in Guangzhou, the largest city in south China by using Geographical Simulation and Optimization System (GeoSOS) software.

2. Methodology

2.1. Basic ACO algorithm

The ACO algorithm, which was first proposed by Dorigo *et al.* (1991), has the capability of solving various optimization problems by simulating the behavior of ants in seeking foods. ACO was initially used to solve the classical traveling salesman problem (TSP) (Dorigo *et al.* 1996). In the optimization, an artificial ant selects a city to visit with a probability that is related to the amount of pheromone trail $\tau_{uv}(t)$ on the path and the travel distance. This transition probability from city u to city v for the kth ant at time t is defined as follows (Dorigo *et al.* 1996):

$$p_{uv}^{k}(t) = \begin{cases} \dfrac{[\tau_{uv}(t)]^{\alpha}[\eta_{uv}(t)]^{\beta}}{\sum\limits_{x \in \text{allowed}_k} [\tau_{ux}(t)]^{\alpha}[\eta_{ux}(t)]^{\beta}}, & if\ v \in \text{allowed}_k \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

where $\tau_{uv}(t)$ is the amount of pheromone trail on edge (u,v) and $\eta_{uv}(t)$ is a heuristic function related to the visibility (distance). The set $\text{allowed}_k = \{\text{C - tabu}_k\}$ represents the cities that can be visited next time without repetition. The parameters of α and β control the relative importance of trail versus visibility (distance).

The heuristic function $\eta_{uv}(t)$ is calculated as the inverse of the distance between cities u and v (Dorigo *et al.* 1996):

$$\eta_{uv}(t) = \frac{1}{d_{uv}} \tag{2}$$

where d_{uv} is the distance between city u and city v.

At each iteration t, the trail density is updated according to the following formula (Dorigo *et al.* 1996):

$$\tau_{uv}(t+1) = (1-\rho)\,\tau_{uv}(t) + \Delta\tau_{uv}(t) \tag{3}$$

$$\Delta\tau_{uv}(t) = \sum_{k=1}^{m} \Delta\tau_{uv}^{k}(t) \tag{4}$$

where ρ is a coefficient such that $(1 - \rho)$ represents the evaporation of trail between t and $t + n$. $\Delta\tau_{uv}^{k}(t)$ is the quantity per unit of length of trail substance laid on path (u,v) by the kth ant between time t and $t + n$.

$\Delta\tau_{uv}^{k}(t)$ is calculated by using the following equation (Dorigo *et al.* 1996):

$$\Delta\tau_{uv}^{k}(t) = \begin{cases} \frac{Q}{L_k}, & \text{if the } k\text{th ant visits } (u, v) \\ 0, & \text{otherwise} \end{cases} \tag{5}$$

where Q is a constant and L_k is the tour length of a solution for the kth ant.

2.2. *Modified ACO for area optimization*

Recently, Li *et al.* (2009a, 2009b) have proposed the methods of modifying ACO so that the ant algorithm can be used to solve complex optimal facility-sitting and path-finding problems. In this study, ACO is further extended to the solution of area optimization problems. This type of optimization applications (e.g. optimal zoning for natural protection) usually requires maximizing the compactness as well as the total suitability of a formulated pattern (Brookes 2001). The formation of such pattern can be tackled by using the pheromone feedback of ants. In forming a natural protection area, an artificial ant will visit a cell and lay down pheromone on the cell. The amount of deposited pheromone is related to the utility of the formulated protected area. The larger the amount of pheromone, the more the ants will be attracted to select this cell. A more amount of pheromone is in turn deposited on this cell. The communication between ants based on the pheromone feedback plays a key role in forming an optimal zoning pattern which can generate the maximum utility. The detailed modifications of ACO for solving this hard zoning problem are discussed in the following sections.

First, the probability that a cell will be selected for forming the natural protected areas by the kth ant at time t is modified according to Equation (1):

$$p_i^k(t) = \begin{cases} \dfrac{[\tau_i(t)]^\alpha [\eta_i(t)]^\beta}{\sum\limits_{x \in \text{allowed}_k} [\tau_x(t)]^\alpha [\eta_x(t)]^\beta}, & \text{if } i \in \text{allowed}_k \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

where i is a cell to be selected for formulating the protected natural area by the kth ant. The tabu list (allowed$_k$) is defined to mask out the selected cells which should not be visited again by an ant.

Second, the heuristic function $\eta_i(t)$ in Equation (2) should be significantly revised to incorporate the ecological suitability at cell i for guiding the walking of ants. An artificial ant is more likely to move toward (select) the cells of higher suitability value so that plausible protection patterns can be formed. This heuristic function is represented by using the following equation:

$$\eta_i = \frac{S_{e_i}}{\sum\limits_x S_{e_x}} \times 10 \tag{7}$$

where S_{e_i} is the ecological suitability at cell i, and $\sum\limits_x S_{e_x}$ is the sum of the suitability for all the cells in the study area.

Third, a crucial part of the modification is to incorporate a utility (goal) function in ACO so that the protected area can be formulated. The variable, L_k, in Equation (5) can represent the tour length or the total cost of a site visited by the kth ant. The term $1/L_k$ can then be replaced by the utility of the natural protection. Therefore, Equation (5) should be revised by incorporating the utility function:

$$\Delta\tau_i^k(t) = \begin{cases} \dfrac{QU_{\text{goal}}}{d(x)+1}, & \text{if } x \text{ falls within } 5 \times 5 \text{ window of cell } i \text{ at time } t \\ 0, & \text{otherwise} \end{cases} \tag{8}$$

where U_{goal} is the utility of the protected pattern, and $d(x)$ is the distance from the central cell i.

The variable of $d(x)$ is used to address the neighborhood influences in site selection. A site should have a higher probability to be selected if its neighbors have already been included in the protected area. The window size is related to the neighborhood influences of CA. It is rather common to use 3×3 or 5×5 window size to address these neighboring influences in many CA studies (Wu 2002, Li and Yeh 2002). In this study, the window size for a number of indicators will be set to 5×5 in the following sections.

According to the criteria of natural protection, the utility (U_{goal}) of a protected pattern consists of two parts: (1) the average total ecological suitability of all the selected cells and

(2) the compactness of the pattern. It is expected that the optimal protected pattern should yield the highest values for the average total ecological suitability and the compactness. This utility is defined as follows:

$$U_{\text{goal}} = w_e S_e + w_c P_c \tag{9}$$

where U_{goal} is the utility, S_e is the average total ecological suitability, and P_c is the compactness index of a protected pattern. The parameters w_e and w_c are the weights for the total ecological suitability and the compactness, respectively.

Suitability analysis is carried out to estimate the average total ecological suitability in Equation (9). Each site is evaluated according to its significance (ecological suitability) for nature conservation. Ecological suitability can be estimated from a series of spatial variables which are retrieved from remote sensing and GIS data (Eastman *et al.* 1998, Malczewski 1999). These spatial variables include the following:

(1) Normalized difference vegetation index (NDVI)
Many studies indicate that vegetation indices are well correlated with various vegetation properties including green leaf area, biomass, percent of green cover, productivity, and photosynthetic activity (Huete 1988). NDVI can be used as an important indicator for the suitability analysis. This index is calculated according to the following equation (Tucker 1979):

$$\text{NDVI} = \frac{\text{TM}_4 - \text{TM}_3}{\text{TM}_4 + \text{TM}_3} \tag{10}$$

(2) Standard deviation of normalized difference vegetation index (NDVI_{std})The standard deviation of NDVI should be a good indicator for representing the biodiversity of vegetation. This indicator is obtained by using a moving 5 × 5 window:

$$\text{NDVI}_{\text{std}} = \sqrt{\frac{\sum_i (\text{NDVI}_i - \overline{\text{NDVI}})^2}{5 \times 5}} \tag{11}$$

(3) Modified normalized difference water index (MNDWI)
The aquatic natural areas which are critical to water quality and supply should be included in the protection. Identification of aquatic natural areas is the most important step for such protection. MNDWI can be used to identify aquatic natural areas effectively (Xu 2006). MNDWI is calculated as follows:

$$\text{MNDWI} = \frac{\text{TM}_2 - \text{TM}_5}{\text{TM}_2 + \text{TM}_5} \tag{12}$$

(4) Relief amplitude (DEM_{amp})
Topography is a major factor for natural protection. Areas of high topographic heterogeneity should be included in the protection zone for maximizing local variation in climatic, edaphic, and hydrologic habitat features (Halpin 1997). Relief amplitude (DEM_{amp}), which can be used to represent this topographic heterogeneity, is calculated by using a moving 5 × 5 window:

$$\mathrm{DEM}_{\mathrm{amp}} = \max(\mathrm{DEM}_i) - \min(\mathrm{DEM}_i) \tag{13}$$

(5) Human-disturbance factors (H_{disturb})

The factors related to human-disturbance dynamics are considered in the selection of protection areas. Natural habitats could be subject to stress from the activities of adjacent human-dominated landscapes. There are negative effects if a site is close to a series of disturbance centers (e.g. urban centers, town centers, roads, and expressways), or it has a large amount of urban development in the neighborhood. Therefore, two spatial variables, proximity disturbance and development intensity disturbance, are defined to represent these human disturbances for ecological conservation.

The negative factor of proximity disturbance is estimated by using a series of proximity variables:

$$\begin{aligned} P_{\mathrm{disturb}} = {} & b_1 D_{\mathrm{MainCenter}} + b_2 D_{\mathrm{DistrictCenters}} + b_3 D_{\mathrm{LTownCenters}} + b_4 D_{\mathrm{STownCenters}} \\ & + b_5 D_{\mathrm{Railways}} + b_6 D_{\mathrm{Subways}} + b_7 D_{\mathrm{Expressways}} + b_8 D_{\mathrm{Roads}} \end{aligned} \tag{14}$$

where $D_{\mathrm{MainCenter}}$ is the distance to the main center, $D_{\mathrm{DistrictCenters}}$ is the distance to the district centers, $D_{\mathrm{LTownCenters}}$ is the distance to the large town centers, $D_{\mathrm{STownCenters}}$ is the distance to the small town centers, D_{Railways} is the distance to the railways, D_{Subways} is the distance to the subways, $D_{\mathrm{Expressways}}$ is the distance to the expressways, and D_{Roads} is the distance to the roads; b_m (m = 1, 2, . . ., 8) is the weight of each variable.

The negative factor of development intensity disturbance in the neighborhood is estimated according to the following equation:

$$D_{\mathrm{disturb}} = \frac{\sum_x D_x}{5 \times 5} \begin{cases} D_x = 1, & \text{if the cell is developed in the neighborhood} \\ D_x = 0, & \text{otherwise} \end{cases} \tag{15}$$

where D_{disturb} represents the negative factor of development intensity disturbance, and D_x is a binary variable indicating if a cell is developed or not in the 5×5 neighborhood of cell i.

The multicriteria evaluation (MCE) method is used to estimate the ecological suitability according to the above spatial variables (Eastman *et al.* 1998). These spatial variables should be standardized within the range of [0, 1] before the estimation. The suitability surface is created by a liner weighted combination of all these spatial variables:

$$\begin{aligned} S_{\mathrm{e}_i} = {} & w_1 \mathrm{NDVI}'_i + w_2 \mathrm{MNDWI}'_i + w_3 (1 - P'_{\mathrm{disturb}}) \\ & + w_4 (1 - D'_{\mathrm{disturb}}) + w_5 \mathrm{DEM}'_{\mathrm{amp}_i} + w_6 \mathrm{NDVI}'_{\mathrm{std}_i} \end{aligned} \tag{16}$$

where NDVI'_i is the standardized normalized difference vegetation index, MNDWI'_i is the standardized modified normalized difference water index, P'_{disturb} is the standardized proximity disturbance, D'_{disturb} is the standardized development intensity disturbance, $\mathrm{DEM}'_{\mathrm{amp}_i}$ is the standardized relief amplitude, $\mathrm{NDVI}'_{\mathrm{std}_i}$ is the standardized standard deviation of normalized difference vegetation index, and w_m $(m = 1, \ldots, 6)$ is the weight for each variable.

The selection of above spatial variables is subject to data availability. If other ecological data possibly available in the future are included, the suitability surface and optimization results may be different to some extent. However, the applicability of the proposed method remains unchanged.

The average total ecological suitability of a pattern in Equation (9) is then calculated according to the following equation:

$$S_\mathrm{e} = \frac{\sum_{i \in \Omega} S_{\mathrm{e}_i}}{A} \tag{17}$$

where Ω is the cells falling within the protected area, and A is the total area of the protection.

The compactness index in Equation (9) is used to avoid the fragmentation of land-use patterns. It is calculated according to the total area and its perimeter of a protected scenario which is composed of all the selected cells. This index is defined according to the following ratio function:

$$P_\mathrm{c} = \frac{\sqrt{A}}{L} \tag{18}$$

where L is the perimeter of a protected scenario.

2.3. *Dynamic optimization for zoning protected natural areas*

An important feature of this proposed method is to couple CA with the modified ACO, so that the effects of land-use dynamics can be considered in formulating protected natural areas. Economic growth and urbanization has resulted in the change of landscape patterns which can be simulated and predicted by CA models. In the past two decades, a number of CA models have been developed to solve complex spatial simulation problems related to urban dynamics and land-use changes. These commonly used models may include SLEUTH-CA (Clarke *et al.* 1997), MCE-CA (Wu and Webster 1998), logistic-CA (Wu 2002, Li *et al.* 2008), ANN-CA (Li and Yeh 2002), and decision-tree CA (Li and Yeh 2004b). Most of these CA models can be calibrated by using empirical information since they have well-defined structures. More importantly, CA models have the advantages of simulating large-scale regions within a feasible computation time (Li *et al.* 2008). These are the main reasons why CA models, instead of agent-based models, are chosen for the coupling with optimization models.

In this study, the landscape dynamic, which provide the background conditions (suitability surfaces) for generating the optimal zoning patterns, will be simulated according to the logistic CA (Wu 2002, Li *et al.* 2008). The development probability, which is the core of urban CA, is estimated by incorporating a stochastic factor, a logistic component, a local interaction factor, and a series of physical constraints (Li *et al.* 2008):

$$p_i^t = (1 + (-\ln \gamma)^\alpha)\frac{1}{1 + \exp(-z_i^t)} f(\Omega_i^t)\xi_i \tag{19}$$

where γ is a stochastic factor ranging from 0 to 1, α is a parameter to control the stochastic degree, z is the global suitability score for urban development, $f(\Omega_i^t)$ is the local interaction factor (development intensity in the neighborhood of Ω_i), and ξ_i is the total constraint score of cell i.

The global suitability score for urban development in Equation (19) is estimated according to a linear combination of various proximity variables (Wu 2002):

$$z_i^t = b_0 + b_1 D_{\mathrm{MainCenter}} + b_2 D_{\mathrm{DistrictCenters}} + b_3 D_{\mathrm{LTownCenters}} + b_4 D_{\mathrm{STownCenters}} + b_5 D_{\mathrm{Railways}} + b_6 D_{\mathrm{Subways}} + b_7 D_{\mathrm{Expressways}} + b_8 D_{\mathrm{Roads}} \tag{20}$$

A coupling paradigm is adopted by integrating urban CA with spatial optimization for the zoning of protected natural areas. An optimal zoning pattern can be generated based on the existing land-use pattern or the predicted one in the future years. There are various optimal patterns subject to the simulated land-use dynamics at different years. The effects of each optimal pattern can be assessed according to the accumulative utility by considering the land-use dynamics. If a protection zone is implemented, its effects should not be just evaluated by the current situations. The benefits of a protection zone should be calculated under a dynamic environment within a planning period. The accumulative utility value (U_{accum}) is thus defined to address these effects by using the following equation:

$$U_{\mathrm{accum}} = \sum_t U_{\mathrm{goal}} = \sum_t \left(w_{\mathrm{e}} \frac{\sum_{i \in \Omega} S_{\mathrm{e}_i}}{A} + w_{\mathrm{c}} P_{\mathrm{c}} \right) \tag{21}$$

where t is the planning period (e.g. 2008, 2013, 2018, 2023, 2028, 2033, 2038, and 2043).

The coupling of the simulation model with the optimization model is implemented by using the following procedures:

(1) Establishing a calibrated urban CA model to simulate regional development according to logistic regression (Wu 2002, Li *et al.* 2008)
(2) Running the calibrated CA model to predict possible land-use changes which will be incorporated in the optimization
(3) Obtaining the suitability surface in Equation (16) and the utility function in Equation (9) for natural protection according to the updates of urban simulation
(4) Creating the optimal patterns by using two strategies of coupling:

- A single-year coupling strategy: The layers of simulated urban areas in the future years are used to create the suitability surfaces according to Equation (16). The ACO model is then used to generate optimal patterns based on these suitability surfaces.
- A merging-year coupling strategy: All these suitability surfaces are summed up to create a single merged suitability layer. The ACO model is then used to create the optimal pattern by directly searching on this merged surface.

For the single-year coupling strategy, there are various optimal patterns with the inputs of the existing land-use pattern or the predicted one in the future years (e.g. 2008, 2013, 2018, 2023, 2028, 2033, 2038, and 2043). The most optimal zoning pattern is identified according to the maximum accumulative utility value as described by Equation (21).

3. Model implementation and application results

3.1. *Study area and spatial data*

The study area covers the metropolitan region of Guangzhou, which has an area of 7434.4 km^2. Guangzhou is located at the center of the Pearl River Delta in Guangdong.

The landscape has been experiencing significant changes because of rapid urban expansion and population growth in this metropolitan region (Li and Yeh 2004a).

Guangzhou Landsat TM images (Scene No. 122–44 in China Remote Sensing Ground Station reference system) in 2003 and 2008 were classified to obtain the information about land-use changes. The classified urban areas in 2003 and 2008 reveal a fast urban expansion in this period, and provide empirical information for training the CA model. This calibrated CA model was used to simulate future distributions of urban areas. Urban dynamics is related to a number of independent variables, such as the distance to the main center, distance to the district centers, distance to the large town centers, distance to the small town centers, distance to the railways, distance to the subways, distance to the expressways, and distance to the roads (Wu and Webster 1998, Li *et al.* 2008). Figure 1 presents these proximity variables created by using common GIS functions.

The ACO optimization model involves a number of spatial variables which determine the suitability of natural protection. The suitability analysis was accomplished by using remote sensing and GIS data (Eastman *et al.* 1998, Malczewski 1999). Figure 2 shows the spatial variables selected for the suitability analysis.

All these spatial variables were converted into a raster format for implementing the proposed model. The computation is quite intensive because the study region is as large as 7434.4 km^2. The resample function of ArcGIS was used to reduce the data burden according

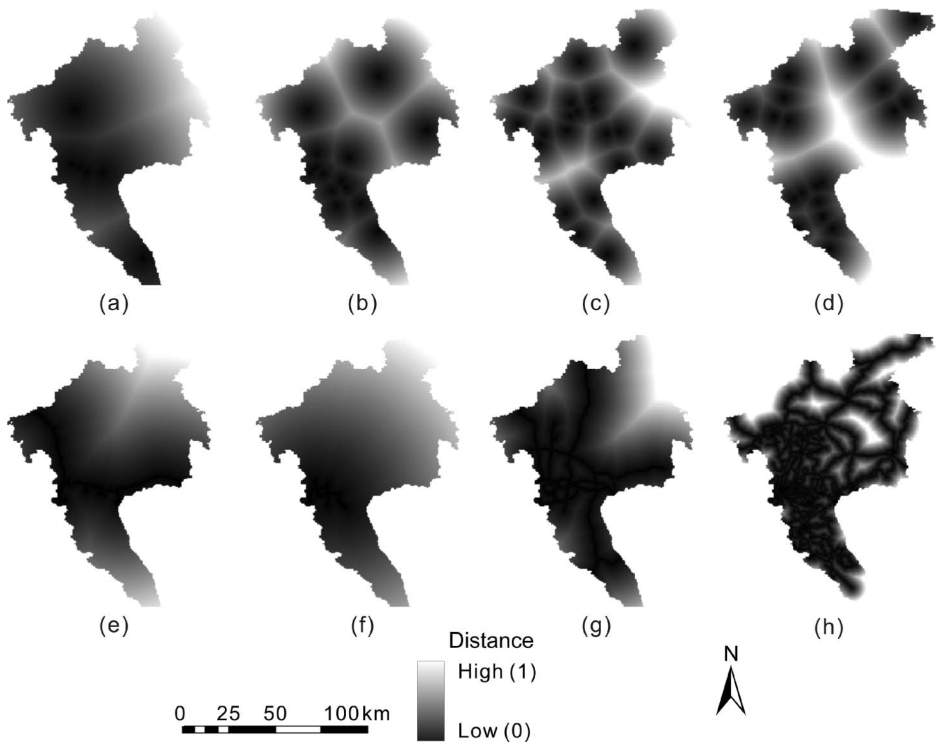

Figure 1. Various proximity variables related to urban dynamics: (a) distance to the main center, (b) distance to the district centers, (c) distance to the large town centers, (d) distance to the small town centers, (e) distance to the railways, (f) distance to the subways, (g) distance to the expressways, and (h) distance to the roads.

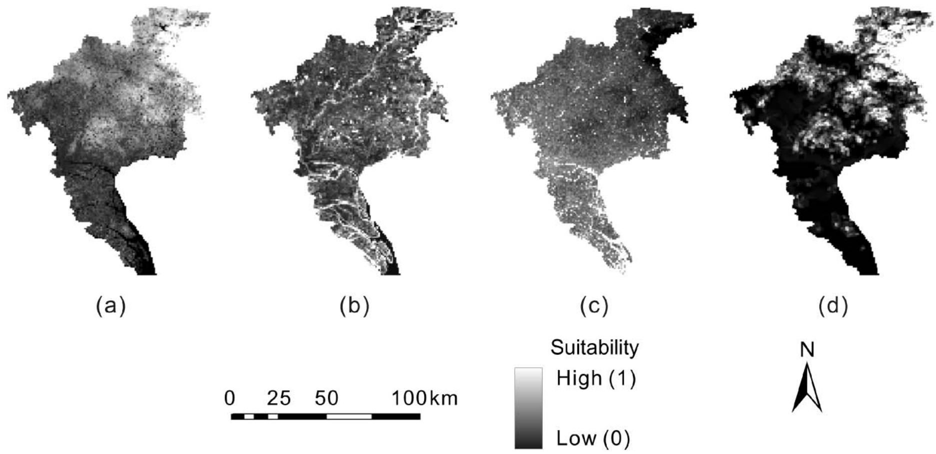

Figure 2. Spatial variables for suitability analysis using remote sensing and GIS data: (a) NDVI, (b) $NDVI_{std}$, (c) MNDWI, and (d) DEM_{amp}.

to the cubic convolution sampling method. All these data were resampled to the layers of 90 m grid-cell size with 1698 × 1244 cells for urban simulation. However, the computation of the area optimization is much more intensive than that of the urban simulation. These original spatial data were further resampled to the layers of 800 m grid-cell size with 191 × 140 cells for spatial optimization.

3.2. Model implementation and results

The optimal zoning of protected natural areas under a changing landscape was accomplished by using a GeoSOS. GeoSOS is equipped with a number of urban simulation and optimization modules which can be used to simulate urban dynamics and optimize natural protection patterns. The software can be downloaded at http://www.geosimulation.cn.

3.2.1. Simulation of landscape changes

First, the module of logistic CA in GeoSOS was used to simulate the urban dynamics in the study area. It is important to calibrate CA models so that future urban development can be predicted in a more realistic way. GeoSOS provides useful tools for calibrating urban CAs. The parameters (weights) of CA will be automatically obtained after the empirical data about urban development and independent spatial variables (e.g. proximity variables) have been defined. Table 1 lists the weights of the calibrated urban CA for Equation (20) by using the calibration procedures of GeoSOS.

Table 1. Parameters of the calibrated urban CA for the study area by using GeoSOS.

Weights	Constant	$D_{MainCenter}$	$D_{DistrictCenters}$	$D_{LTownCenters}$	$D_{STownCenters}$
	1.570	−3.812	−1.314	−0.338	0.184
Weights	$D_{Railways}$	$D_{Subways}$	$D_{Expressways}$	D_{Roads}	
	−3.110	1.152	1.954	−3.397	

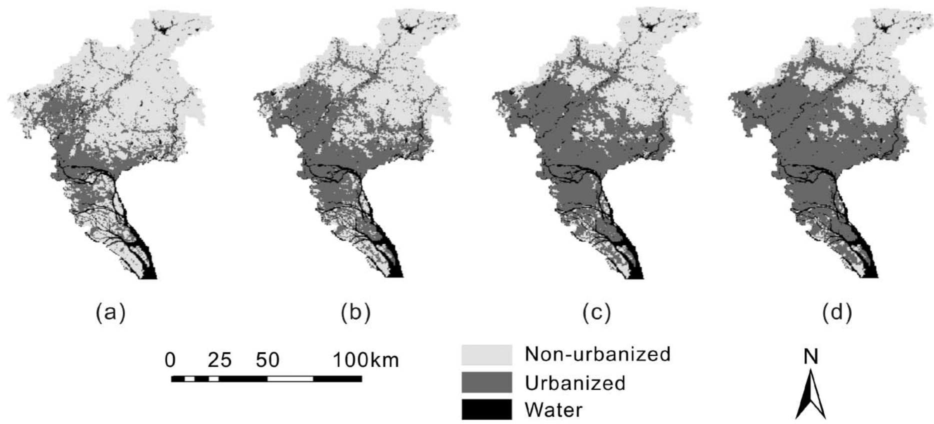

Figure 3. Simulated urban areas in (a) 2008, (b) 2018, (c) 2028, and (d) 2038.

This calibrated CA was then used to simulate the distribution of urban areas in 2008, 2013, 2018, 2023, 2028, 2033, 2038, and 2043 respectively. Figure 3 just shows the simulated patterns of 2008, 2018, 2028, and 2038. The simulated urban areas in 2008 can be compared with the actual urban areas classified from the 2008 TM image for validating this simulation model. A cell-by-cell overlay indicates that the total simulation accuracy is 88.7%.

3.2.2. Optimizing protected areas under a changing landscape

The modified ACO was used to search for the optimal pattern for the protected natural area. The required area for the protection is assumed to be 3840 km^2 with reference to the strategic planning of Guangzhou. This optimization model involves some parameters which could have impacts on the optimization results. These parameters can be determined according to previous studies (Dorigo *et al.* 1996, Li *et al.* 2009a, 2009b). Table 2 lists the detailed parameters for implementing this area optimization ACO.

Suitability analysis involves a number of spatial variables. The weights for each variable should be decided according to expert experiences and domain knowledge. Table 3 provides these weights for Equations (14) and (16) based on Saaty's pairwise comparison method (Eastman 1998).

The weights in the utility function of Equation (9) may have different combinations. The sensitivity of these combinations on optimization effects can be analyzed by varying these weights. The experiments were based on four typical combinations of these weights: (1) $w_e = 1$ and $w_c = 0$; (2) $w_e = 0.7$ and $w_c = 0.3$; (3) $w_e = 0.3$ and $w_c = 0.7$; and (4) $w_e = 0$ and $w_c = 1$. Figure 4a shows the optimization patterns by using these four settings of weights, based on the existing urban areas in 2008. It is clear that the compactness factor has an

Table 2. Parameters used in this area optimization ACO.

Iteration	α	β	ρ	Q
1000	5	1	0.01	0.1

Table 3. Weights for calculating ecological suitability.

(a) Weights for calculating Equation (14)

Weights	$D_{\text{MainCenter}}$	$D_{\text{DistrictCenters}}$	$D_{\text{LTownCenters}}$	$D_{\text{STowCenters}}$
	0.109	0.137	0.153	0.184
Weights	D_{Railways}	D_{Subways}	$D_{\text{Expressways}}$	D_{Roads}
	0.192	0.029	0.171	0.025

(b) Weights for calculating Equation (16)

	NDVI_i'	MNDWI_i'	$1 - P_{\text{disturb}}'$	$1 - D_{\text{disturb}}'$	$\text{DEM}'_{\text{amp}i}$	$\text{NDVI}'_{\text{std}i}$
Weights	0.064	0.214	0.150	0.237	0.058	0.267

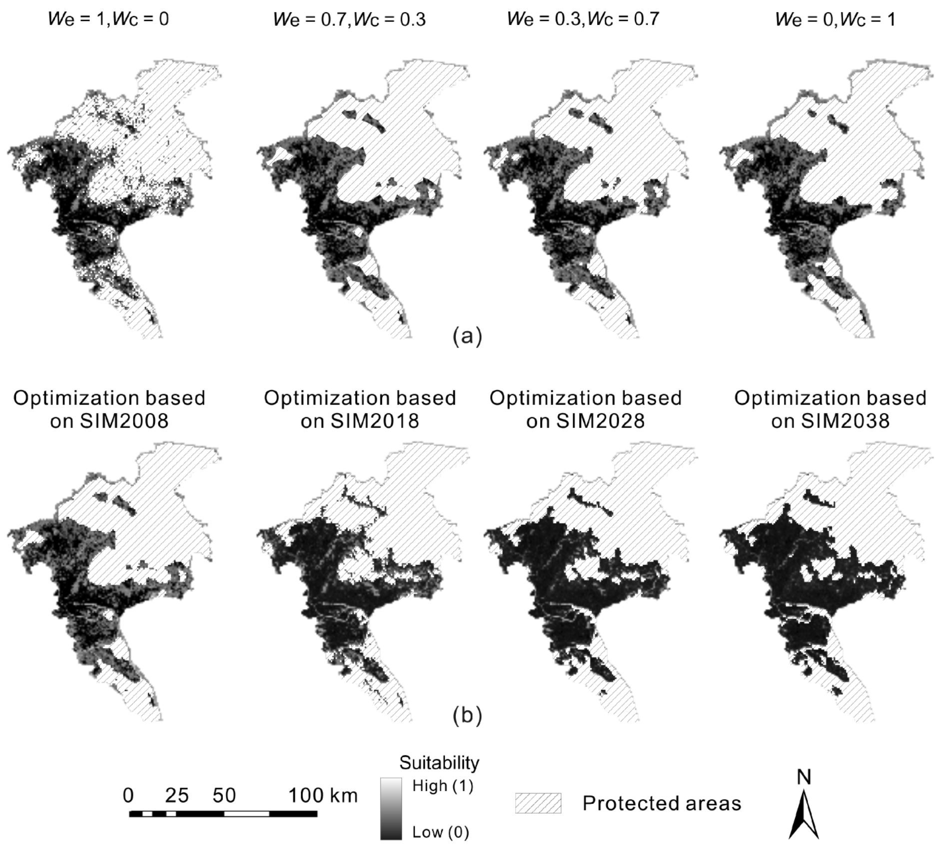

Figure 4. Sensitivity analysis (a) and various optimal protected patterns based on urban simulations (b).

important role for deriving a feasible protection zone. The first combination is an extreme case which does not consider the compactness factor ($w_c = 0$). As a result, the optimal pattern (Figure 4a) is too fragmented to be implemented as a practical protection plan. Actually, the compactness is a high priority factor for ecological conservation (Matisziw and Murray 2006). However, the increase in compactness is at the cost of ecological suitability (Table 4). An appropriate choice of these weights is important for generating plausible zoning solutions. It is found that the second set of weights ($w_e = 0.7$ and $w_c = 0.3$) can produce a very

Table 4. Sensitivity analysis of this area optimization ACO.

Weights	Ecological suitability	Compactness
$w_e = 1$ and $w_c = 0$	0.4700	0.8606
$w_e = 0.7$ and $w_c = 0.3$	0.4625	0.9622
$w_e = 0.3$ and $w_c = 0.7$	0.4615	0.9628
$w_e = 0$ and $w_c = 1$	0.4511	0.9712

satisfactory compact pattern for natural protection according to the visual interpretation and the comparison of the trade-off (Table 4). Therefore, the second set of weights was used to define the utility function of ACO for further analyses.

The optimal pattern for natural area protection can be created according to the existing urban areas in 2008 (Figure 4b). However, such optimization cannot yield the maximum accumulative utility under a dynamic landscape. Instead, the optimization model should use all the land-use patterns in the planning period as inputs. For this purpose, the simulated urban areas in 2013, 2018, 2023, 2028, 2033, 2038, and 2043 were used to calculate the negative factor of development intensity disturbance (D'_{disturb}). Various suitability surfaces were then obtained with regard to these simulated scenarios. Two strategies of coupling were adopted for the implementation of this dynamic optimization. For the single-year coupling strategy, the suitability surfaces were created based on the simulated urban areas in 2013, 2018, 2023, 2028, 2033, 2038, and 2043 (SIM2013, SIM2018, SIM2023, SIM2028, SIM2033, SIM2038, and SIM2043). The ACO model was then used to generate each optimal pattern for these single-year inputs (suitability surfaces) respectively (Figure 4b). The best optimal pattern was identified among these patterns based on the accumulative utility value as described in Equation (21). For the merging-year coupling strategy, all these suitability surfaces related to SIM2013, SIM2018, SIM2023, SIM2028, SIM2033, SIM2038, and SIM2043 were summed up to create a single merged suitability layer. The optimal pattern was generated by directly searching on this merged surface.

For the single-year coupling strategy, the highest accumulative utility value (5.0998) of the optimization is obtained if it is based on the simulated land use in 2028 (Figure 5). The non-coupling optimization based on the static (existing) land use in 2008 only yields the

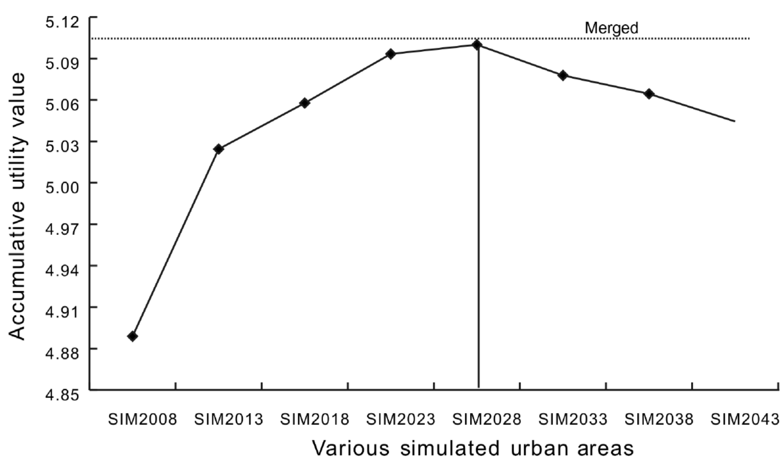

Figure 5. Comparison of the accumulative utility values between the single-year coupling strategy and the merging-year coupling strategy.

accumulative utility value of 4.8889. Therefore, the best pattern from the single-year coupling strategy has a utility improvement of 4.3%, compared to the non-coupling optimization. For the merging-year coupling strategy, the accumulative utility value (5.1015) is slightly higher than that of the best pattern from the single-year coupling strategy (5.0998). Therefore, the performances of the spatial optimization can be significantly improved by using the proposed single-year coupling strategy or merging-year coupling strategy.

The above analysis has demonstrated that urban dynamics have significant impacts on spatial optimization. However, optimization patterns should also affect urban dynamics if these zoning patterns are implemented. The proposed coupling method provides a convenient way for exploring such mutual influences. In this experiment, the zoning patterns are regarded as a constraint factor for the CA model according to Equation (19). By inputting different zoning patterns, the CA model will generate different scenarios of urban development. Figure 6 clearly shows that the compactness of urban forms is related to the zoning of natural protected areas. It is found that the future urban forms could be more fragmented if the optimal zoning is based on the static land use in 2008 (non-coupling). However, the urban development will become more compact if the zoning is based on the single-year coupling strategy or merging-year coupling strategy. Therefore, the adoption of proper zoning for natural protection is also important for creating a more reasonable urban form in this fast-growing region.

3.2.3 *Model validation*

There is a question if the proposed ACO model will be trapped at local optima during the search for optimal zoning patterns. However, it is difficult to know if heuristic models can reach global optima for high-dimensional data because the optima are unknown. A practical way is to use well-structured hypothetical data with known optima to validate this ACO model. Figure 7 shows such a hypothetical suitability surface with multiple peaks, which has a peak of the highest suitability value at the center and four minor peaks at four corners. The peak at the center is designed to be large enough to accommodate all the selected cells. It is obvious that the known optimum is a compact circle at the center. This special design can help to examine if the proposed model will stop at some local optima (the four minor 'hills')

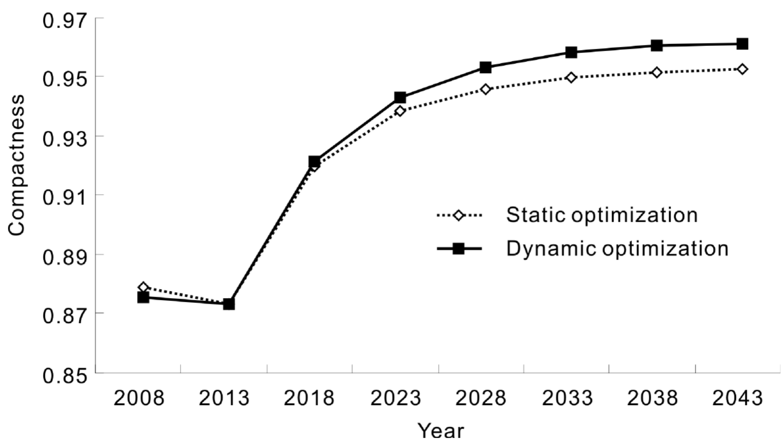

Figure 6. Compactness of future urban forms affected by various optimization schemes.

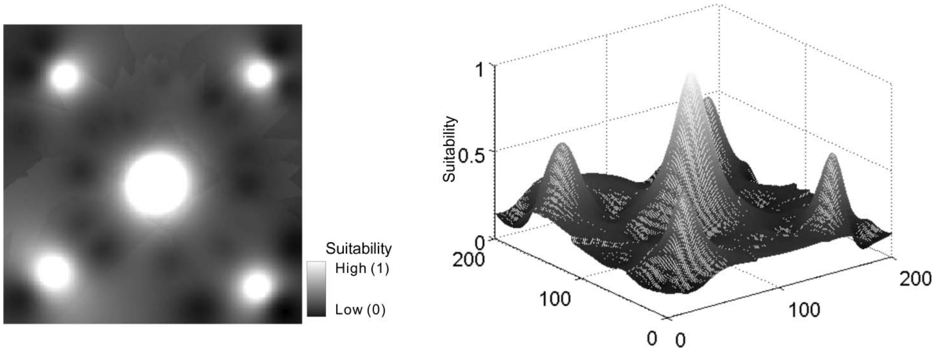

Figure 7. Hypothetical suitability surface with multiple peaks.

(Figure 8). At the beginning, artificial ants are randomly located in the raster suitability space. These ants will explore the space and try to identify the best locations for forming an optimal pattern. At the early stage (e.g. 50 iterations), some of these ants did occupy these local optima (the four minor 'hills'). However, these ants can quickly get out of these traps because of the cooperation between ants. Actually, these ants have almost occupied the best locations for forming the protection zone after 200 iterations. The overlay of the final formulated pattern with the known optimum pattern reveals a very good fit between them (Figure 8). This suggests that the proposed model can find the near optimum for area optimization by using the bottom-up approach of ant intelligence.

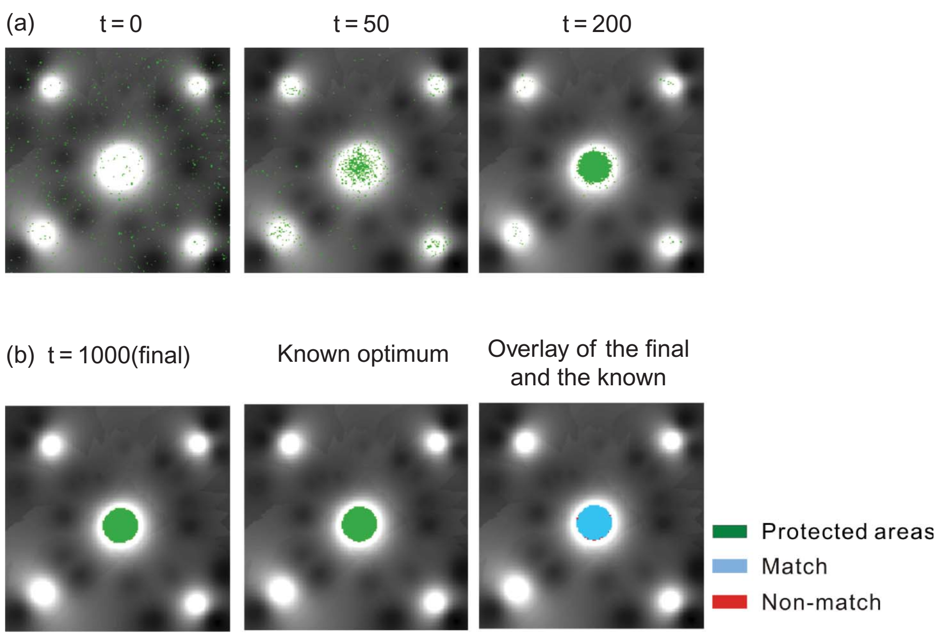

Figure 8. Model validation using the hypothetical data: (a) escaping from local optimums by ants; (b) overlay of the modeling optimum and the known optimum.

As the comparison, the zoning of protected natural areas was also implemented by using three common methods: (1) SA (Aerts and Heuvelink 2002); (2) the iterative relaxation (IR) (Eastman *et al.* 1995); and (3) the density slicing (DS) (Li and Yeh 2001). These methods are applied to the same data set by using the same utility function so that their performances for the optimal zoning can be compared with that of the modified ACO. The optimization was just based on the existing land use in 2008 without considering the coupling effects for simplicity.

In the SA method, initial locations for n sites which are composed of the protected area were randomly generated and the initial utility, $U_{goal}(0)$, was obtained. A small perturbation was used to move these initial locations and a new utility value $U_{goal}(1)$ was calculated. The acceptance of such move at each iteration was subject to the following rules (Aerts and Heuvelink 2002):

If $U_{goal}(t+1) > U_{goal}(t)$

Then, the move is accepted, and

If $U_{goal}(t+1) < U_{goal}(t)$ and $\exp(U_{goal}(t+1) - U_{goal}(t))/T_C(t) >$ Random[0,1] then the move is also accepted.

$T_C(t)$ is the freezing parameter which is gradually deceasing by using a multiple formula,

$$T_C(t+1) = \delta T_C(t) \tag{22}$$

where δ is a cooling constant ($0 < \delta < 1$).

Typical values for this cooling constant are between 0.80 and 0.98 (Laarhoven 1987). The initial value of temperature ($T_C(0)$) and the cooling constant were set to 100 and 0.98, respectively. As a result, $T_C(t+1)$ decreases at each step by multiplying the cooling constant. In this way, the acceptance probability of the worse solution becomes smaller for reaching the global optimal solution.

The IR program was executed according to the following steps (Eastman *et al.* 1995): (1) cells with the utility value greater than a predefined threshold are selected; (2) after removing those spatially discontinuous areas, the rest of them become candidates; (3) a further selection is performed, based on the condition that the area of a candidate region should be greater than the threshold; and (4) the above process is continued until the utility improvement becomes stabilized.

The DS method, which is unable to include the compactness factor, is very simple to implement. The zoning is only based on the suitability value by slicing the density of suitability score. Cells of higher suitability values were selected for formulating the protected area.

Figures 9 and 10 show the effects of these methods for generating the optimal protected natural area. ACO can generate the maximum utility value with a very plausible pattern (Figure 9a). SA seems to be the second best in terms of the total utility value and the pattern (Figure 9b). However, the computation time of SA is much longer than that of ACO according to the experiments. The computation times are 41 and 11 min for SA and ACO respectively, by using a computer with Pentium D 3.4GhZ CPU.

The IR method generally has the capability of deriving the most suitable and continuous area (Figure 9c). However, its compactness is much poorer than that of ACO and SA. Its total utility value is also ranked at the third position. The DS method has the poorest performances although it is the simplest one for computation. This method will result in a fragmented pattern, which can hardly be implemented for practical applications (Figure 9d). In summary, ACO has the improvement of the compactness over SA, IR, and DS by 4.81%, 7.59%,

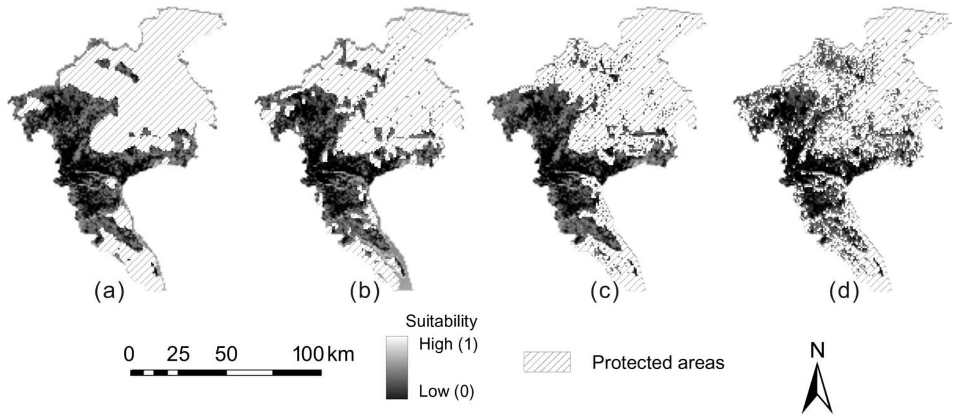

Figure 9. Zoning of protected natural areas using (a) ACO, (b) SA, (c) IR, and (d) DS methods.

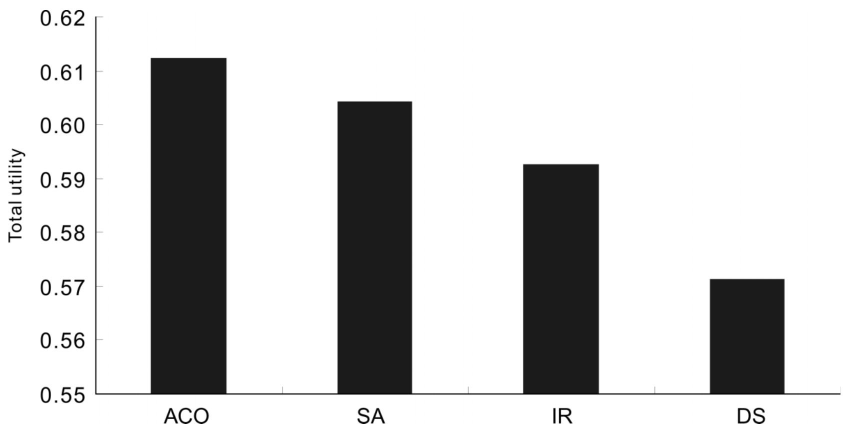

Figure 10. Total utility values for the zoning of protected natural areas using ACO, SA, IR, and DS methods.

and 21.23%, respectively. ACO also has the improvement of the total utility over SA, IR, and DS by 1.33%, 3.21%, and 6.71%, respectively (Figure 10). Therefore, the proposed modified ACO seems to be very attractive for area optimization.

4. Conclusion

This study has demonstrated that ACO can be used to solve the natural-area protection problems effectively. However, three important steps of modifications are required for extending ACO to the solution of area optimization problems. These modifications include: (1) revising the transition probability; (2) defining a heuristic function by encouraging ants to select the cells of higher suitability values; and (3) incorporating a utility (goal) function by addressing the criteria of natural protection.

This proposed model has much better performances than other common conventional models. The comparison between ACO and other common methods, such as SA, IR, and

DS, indicates that ant intelligence has advantages for solving complex spatial optimization problems. ACO has the improvement of the compactness over SA, IR, and DS by 4.81%, 7.59%, and 21.23%, respectively. ACO also has the improvement of the total utility over SA, IR, and DS by 1.33%, 3.21%, and 6.71% respectively. Although SA may have close optimization results with ACO, the computation time of SA is much longer than that of ACO.

Existing static optimization methods have limitations without considering future land-use changes in the landscape. This study suggests that this coupling simulation-optimization model may have contributions for the advances in spatial optimization that can help to formulate the plans for protecting ecosystems. The coupling is based on two strategies: the single-year coupling strategy and the merging-year coupling strategy. There is a utility improvement of 4.3% by using the single-year coupling strategy. The merging-year coupling strategy can also have a slightly higher value of the accumulative utility value than the single-year coupling strategy. It is also found that the optimization patterns can further affect the future development patterns. The compactness of urban forms can be improved if the best optimization pattern for natural conservation is adopted in the study area.

This area optimization ACO is validated by using hypothetical data in which the optimal solutions are known. It is expected that a sound optimization model should identify the best location for site selection after sufficient iterations of simulation. The experiments by using the hypothetical suitability surface with multiple peaks have indicated that this proposed model will not be trapped at the local optima. It can reach the global optimum after several hundreds of iterations.

Acknowledgments

This study was supported by the Key National Natural Science Foundation of China (grant no. 40830532), the National Outstanding Youth Foundation of China (grant no. 40525002), and the National Natural Science Foundation of China (grant no. 40901187).

References

Aerts, J. and Heuvelink, G.B.M., 2002. Using simulated annealing for resource allocation. *International Journal of Geographical Information Science*, 16, 571–587.

Bos, J., 1993. Zoning in forest management: a quadratic assignment problem solved by simulated annealing. *Journal of Environmental Management*, 37, 127–145.

Brookes, C.J., 1998. A genetic algorithm for locating optimal sites on raster suitability maps. *Transactions in GIS*, 2, 201–212.

Brookes, C.J., 2001. A genetic algorithm for designing optimal patch configurations in GIS. *International Journal of Geographical Information Science*, 15, 539–559.

Carsjens, G.J. and van der Knaap, W., 2002. Strategic land-use allocation: dealing with spatial relationships and fragmentation of agriculture. *Landscape and Urban Planning*, 58, 171–179.

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: Planning and Design*, 24, 247–262.

Cova, T.J. and Church, R.L., 2000. Contiguity constraints for single-region site search problems. *Geographical Analysis*, 32, 306–329.

Deng, X.Z., *et al*., 2008. Growth, population and industrialization, and urban land expansion of China. *Journal of Urban Economics*, 63, 96–115.

Dorigo, M., Maniezzo, V., and Colorni, A., 1991. Positive feedback as a search strategy. *Technical Report*. Milan: Politecnico di Milano, 91–106.

Dorigo, M., Maniezzo, V., and Colorni, A., 1996. Ant system: optimization by a colony of cooperating agents. *IEEE Transactions on Systems, Man, and Cybernetics – Part B*, 26, 1–13.

Eastman, J.R., Jiang, H., and Toledano, J., 1998. Multi-criteria and multi-objective decision making for land allocation using GIS. *In*: E. Beinat and P. Nijkamp, eds. *Multicriteria analysis for land-use management*. Dordrecht, The Netherlands: Kluwer Academic Publishers, 227–251.

Eastman, J.R., *et al*., 1995. Raster procedures for multi-criteria/multi-objective decisions. *Photogrammetric Engineering and Remote Sensing*, 61, 539–547.

Environmental Department of Shenzhen, 2008. *Environmental planning outline of Shenzhen in 2007–2020*, Shenzhen: Government Report, pp. 72.

Fahrig, L., 1998. When does fragmentation of breeding habitat affect population survival? *Ecological Modelling*, 105, 273–292.

Geneletti, D. and van Duren, I., 2008. Protected area zoning for conservation and use: A combination of spatial multicriteria and multiobjective evaluation. *Landscape and Urban Planning*, 85, 97–110.

Halpin, P.N., 1997. Global climate change and natural-area protection: management responses and research directions. *Ecological Applications*, 7, 828–843.

Huete, A.R., 1988. A soil-adjusted vegetation index (SAVI). *Remote Sensing of Environment*, 25, 295–309.

Kuznets, S., 1966. *Modern economic growth: rate, structure and spread*. New Haven, CT: Yale University Press.

Laarhoven, P.J.M. van, 1987, *Theoretical and Computational Aspects of Simulated Anneating*, PhD Thesis, Erasmus University, Rotterdam.

Li, X., He, J.Q., and Liu, X.P., 2009a. Ant intelligence for solving optimal path-covering problems with multi-objectives. *International Journal of Geographical Information Science*, 23, 839–857.

Li, X., He, J.Q., and Liu, X.P., 2009b. Intelligent GIS for solving high-dimensional site selection problems using ant colony optimization techniques. *International Journal of Geographical Information Science*, 23, 399–416.

Li, X., Yang, Q.S., and Liu, X.P., 2008. Discovering and evaluating urban signatures for simulating compact development using cellular automata. *Landscape and Urban Planning*, 86, 177–186.

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 and Remote Sensing*, 67, 471–477.

Li, X. and Yeh, A.G.O, 2002. Neural-network-based cellular automata for simulating multiple land use changes using GIS. *International Journal of Geographical Information Science*, 16, 323–343.

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

Li, X. and Yeh, A.G.O., 2004b. Data mining of cellular automata's transition rules. *International Journal of Geographical Information Science*, 18, 723–744.

Malczewski, J., 1999. *GIS and multicriteria decision analysis*. Toronto: John Wiley &Sons.

Matisziw, T.C. and Murray, A.T., 2006. Promoting species persistence through spatial association optimization in nature reserve design. *Journal of Geographical Systems*, 8, 289–305.

Minor, S.D. and Jacobs, T.L., 1994. Optimal land allocation for solid and hazardous waste landfill sitting. *Journal of Environmental Engineering*, 120, 1095–1108.

Pulliam, H.R., Dunning, J.B. Jr., and Liu, J., 1992. Population dynamics in complex landscapes: a case study. *Ecological Applications*, 2, 165–177.

Sante-Riveira, I., *et al*., 2008. Algorithm based on simulated annealing for land-use allocation. *Computers and Geosciences*, 34, 259–268.

Snyder, S.A., Haight, R.G., and ReVelle, C.S., 2004. Scenario optimization model for dynamic reserve site selection. *Environmental Modeling and Assessment*, 9, 179–187.

Tucker, C., 1979. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment*, 8, 127–150.

Van Zee, C. and Lee, J., 1989. Hazardous waste disposal site selection using iterative GIS technology. *AUTO CARTO 9*. Falls Church: ASPRS/ACSM, 391–396.

Verdiell, A., *et al*., 2005. A mathematical model for zoning of protected natural areas. *International Transactions in Operational Research*, 12, 203–213.

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

Wu, F.L. and Webster, C.J., 1998. Simulation of land development through the integration of cellular automata and multicriteria evaluation. *Environment and planning B*, 25, 103–126.

Xu, H.Q., 2006. Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. *International Journal of Remote Sensing*, 27, 3025–3033.