Landscape Ecol (2012) 27:447–463
DOI 10.1007/s10980-011-9684-1

RESEARCH ARTICLE

An integrated approach of remote sensing, GIS and swarm intelligence for zoning protected ecological areas

Xiaoping Liu · Chunhua Lao · Xia Li · Yilun Liu · Yimin Chen

Received: 8 November 2010 / Accepted: 8 November 2011 / Published online: 27 November 2011

Abstract Interest in protecting ecological areas is increasing because of land uses conflicts and environmental pressures. The optimal zoning of protected ecological areas belongs to a NP-hard problem because it is subject to both box and spatial constraints. A challenge in solving area optimization problems emerges with the increasing size of a study region. In this article, an integrated approach of remote sensing, GIS and modified ant colony optimization (ACO) is proposed for application in zoning protected ecological areas. Significant modifications have been made in the conventional ACO so that it can be further extended to solve zoning problems in large regions. An improved selection strategy is designed to accelerate the progress of sites selection for artificial ants. Another important modification in ACO is to incorporate the neighborhood diffusion strategy into pheromone updating. The optimal objective is to generate protected areas that maximize both ecological suitability and spatial compactness. The modified ACO model has been successfully applied to a case study involving an area of 25,483 cells in Dongguan, Guangdong, China. The experiments have demonstrated that the proposed model is an efficient and effective optimization technique for generating optimal protection. The modified ACO model only requires approximately 119 s for determining near-optimal solutions. Furthermore, the proposed method performs better than other methods, including simulated annealing, genetic algorithm, iterative relaxation, basic ACO, and density slicing.

Keywords Remote sensing · GIS · ACO · Zoning · Protected ecological areas

X. Liu (✉) · C. Lao · X. Li · Y. Liu · Y. Chen
School of Geography and Planning, and Guangdong Key Laboratory for Urbanization and Geo-simulation, Sun Yat-sen University, Guangzhou 510275, People's Republic of China
e-mail: yiernanh@163.com

C. Lao
e-mail: 107683069@qq.com

X. Li
e-mail: lixia@mail.sysu.edu.cn

Introduction

Land-use changes in China have been dominated by unprecedented urban transformation since the launch of economic reform and the open-door policy in 1978 (Yeh and Li 1999). This transformation is particularly true in numerous coastal regions and cities, such as the Pearl River Delta. Urban areas expanded by more than 300% between 1988 and 1996 (Seto et al. 2002). As urban areas expand, natural ecosystems, farmland, water and vegetation are converted into urban areas. Rampant urban growth has resulted in a series of environmental and ecological problems, including encroachment of agricultural land, local and regional climate change, destruction of sensitive ecosystems, loss of wildlife habit and biodiversity, water and air pollution, soil erosion, and severe flooding because of the lack of appropriate land-use planning (Seto et al. 2002; Li and Liu 2008). These problems have had a significant impact on sustainable land development.

As a method for reducing the problems associated with land conversion and development, the establishment of a system of protected ecological areas can serve the purposes of conservation of species and ecosystem diversity, preservation of ecological processes, protection of agricultural land, limitation of natural resource exploitation, and promotion of scientific activities and recreation (Snyder et al. 2005; Verdiell et al. 2005).

In China, the zoning of protected ecological land areas has been implemented in fast-growing regions, such as Shenzhen (Environmental Department of Shenzhen 2008). Disorderly urban development has been effectively controlled after protected areas were officially demarcated. For example, illegal buildings were put to an end in Yantian District of Shenzhen in the past 2 years. The protected areas play an important role in solving and preventing the land use problems, such as excessive urban development and increasing eco-environmental pressures.

Difficult decisions must be made to evaluate the trade-offs among ecological protection goals, economic development, and spatial constraints. Empirical studies have shown that urban land expands by 3% when the economy, measured by the gross domestic product, grows by 10% in China (Deng et al. 2008). Land consumption is required to accommodate the growing population shifting from rural to urban and the expanding economic activities (Kuznets 1966). A fragment pattern will be produced without incorporating spatial constraints in the protection zoning. An ongoing dilemma in the design of protected areas is whether the spatial pattern of selected sites is compact (Diamond et al. 1976). A compact protection may enhance the long term persistence of species (Önal and Briers 2002). Under this situation, zoning is a complex multi-objective optimization problem because a trade-off is inevitably involved (Carsjens and Van der Knaap 2002). To date, zoning methods applied to protected areas are mostly qualitative and rely heavily on the involvement of subject experts in developing countries (Verdiell et al. 2005). Planners need a quantitative and effective tool to evaluate alternative ecological protection plans in terms of achieving intended objectives measured against costs.

A number of quantitative zoning methods have been developed over the past two decades. An initial approach to the problem is to select the set of sites with the highest suitability up to the fulfillment of the land demand (Geneletti and van Duren 2008). However, zoning based on the ranking of suitability values without considering spatial constraints will result in the fragmentation of protected areas. Mathematical techniques, such as integer programming (IP) and mixed-integer programming (MIP), are the most commonly used methods for zoning protected areas (Church et al. 2003). Cocks and Baird (1989) applied IP to address the multiple reserve selection problems in South Australia. Hof and Joyce (1993) developed a MIP approach for spatially optimizing wildlife and timber in managed forest ecosystems. These techniques can ensure optimal solutions, but noting that mathematical techniques may not yield a solution within a reasonable amount of time. To overcome this issue, researchers resort to heuristic algorithms that are more efficient to solve complex optimization problems (Xiao et al. 2007). A common heuristic for solving these zoning problems is based on the well-known simulated annealing (SA) algorithm (Bos 1993; Verdiell et al. 2005). For example, Possingham et al. (2000) presented a modified SA method that incorporated spatial considerations into a reserve network design. Verdiell et al. (2005) developed an SA method for the zoning of protected natural areas subject to both box and spatial constraints. However, most of these methods are primarily applied to the spatial data with coarse resolutions. For example, Bos (1993) used the SA algorithm to create forest management zones with an area of only 83 cells. Verdiell et al. (2005) applied the SA method to select and design a national park, but the study region has an area of only 900 cells. The challenge lies in solving zone allocation problems with an increase in the size of the study area. Thus, exploring efficient and effective optimization methods for zone allocation in large areas is academically interesting and may also result in useful practical applications.

This research proposes an approach based on the integrated use of remote sensing, GIS and swarm intelligence for the zoning of protected ecological areas. The zoning of protected areas usually involves an analysis of a large amount of spatial data. The integration of remote sensing and GIS can provide the tools and data required for such a purpose. Remote sensing data can be used to obtain information on land use, vegetation indices and water indices. GIS is capable of generating spatial variables, such as slope, elevation, traffic accessibility, and habitat diversity. Furthermore, GIS is used to overlay these spatial variables to create a composite map that acts as an ecological suitability map. Then, swarm intelligence is designed to generate protected ecological areas based on the suitability map. As a bottom-up approach, swarm intelligence (SI) is a complex multi-agents system, consisting of numerous simple individuals (e.g., ants, birds, and so on), which exhibit their swarm intelligence through coordination and competition among the individuals (Liu et al. 2008a). SI mainly involves two algorithms, namely, ant colony optimization (ACO) and particle swarm optimization (PSO). In this article, ACO is attempted to be introduced into the solution of area optimization problems. ACO, a computational method inspired by observations of the behavior of real ants, was first proposed by Dorigo (1992). This algorithm is composed of a set of simple artificial ants that cooperate through self-organization. Studies indicate that ACO has better performance than other nature-inspired algorithms, such as SA and evolutionary computations, in solving complex combinatorial optimization problems, because the mechanism of pheromone updating is effective for finding the optimal solutions using cooperating artificial ants (Dorigo and Gambardella 1997). In recent years, ACO algorithms have been used successfully to solve geographical problems, such as urban simulation (Liu et al. 2008b), remote sensing classification (Liu et al. 2008c), and path-covering optimization (Li et al. 2009a). Previous studies have demonstrated that ACO is a potentially useful algorithm for tackling complex spatial optimization problems.

In this article, conventional ACO will be modified so that it can be adapted to the solution of zoning problems. It is expected that the positive feedback mechanism of ACO can produce better performance in handling complex, heterogeneous spatial data for the search of an optimal solution. The proposed ACO method is used to generate protected ecological areas in Dongguan, a rapidly growing region in the Pearl River Delta, China. Finally, the modified ACO method will be compared with other methods, such as SA, genetic algorithm (GA), iterative relaxation (RI), basic ACO, and density slicing (DS), in terms of performance in area optimization.

ACO algorithm for the zoning of protected ecological areas

ACO for the traveling salesman problem (TSP)

ACO is a novel heuristic approach for the solution of combinatorial optimization problems. The optimization is carried out by simulating the natural behavior of ant colonies in their search for food, including mechanisms of cooperation and adaptation (Dorigo 1992). When searching for food, ants initially explore the area surrounding their nest in a random manner. As soon as an ant finds a food source, it deposits some pheromone to mark the path between the nest and the food source. The quantity of pheromone may depend on the quantity and quality of the food. Other ants can detect the pheromone trail and are attracted to follow it. Pheromone evaporates with time, so the pheromone on a long path will decrease when fewer ants select it. The shorter path is reinforced, and more ants are attracted to follow the trail, resulting in an increase in the quantity of pheromone on the shorter path. In this way, ants can find the shortest route from their nest to food sources through the communication.

ACO was first applied to the well-known traveling salesman problem (TSP), which is to find the shortest tour between N cities, visiting each only once and ending at the starting point (Dorigo and Gambardella 1997). In the algorithm, an artificial ant selects a city to visit with a probability that is determined by the following equation (Dorigo and Gambardella 1997):

$$p_{gh}^{k}(t) = \begin{cases} \dfrac{\left[\tau_{gh}(t)\right]^{\alpha} \cdot \left[\eta_{gh}(t)\right]^{\beta}}{\sum_{s \in \text{allowed}_k} \left[\tau_{gs}(t)\right]^{\alpha} \cdot \left[\eta_{gs}(t)\right]^{\beta}}, \text{if h} \in \text{allowed}_\text{k} \\ 0 \quad \text{otherwise} \end{cases} \tag{1}$$

where $p_{gh}^{k}(t)$ is the transition probability from city g to h for the kth ant at time t, $\tau_{gh}(t)$ is the amount of pheromone on edge (g, h), $\eta_{gh}(t)$ is a heuristic function which is defined as the inverse of the distance between cites g and h, and allowed_k is a set of the cites that can be selected by the kth ant at city g for the next step. The relative weight of the pheromone trail and the heuristic value are controlled by the parameters α and β.

After all ants have completed their tour, the intensity of pheromone is updated based on the following equations (Dorigo et al. 1996):

$$\tau_{gh}(t+1) = (1-\rho) \times \tau_{gh}(t) + \Delta\tau_{gh}(t) \tag{2}$$

$$\Delta\tau_{gh}(t) = \sum_{k=1}^{m} \Delta\tau_{gh}^{k}(t) \tag{3}$$

$$\Delta\tau_{gh}^{k}(t) = \begin{cases} \dfrac{Q}{L_k} & \text{if the } k\text{th ant visits } (g,\ h) \\ 0 & \end{cases} \tag{4}$$

where $\rho \in (0, 1]$ is the pheromone trail decay coefficient, $\Delta\tau_{gh}(t)$ is the increment of $\tau_{gh}(t)$, $\Delta\tau_{gh}^{k}(t)$is the pheromone trail deposited by the kth ant on edge (g, h) at time t, m is the number of ants, Q is a constant, andL_k is the total length or cost of current tour traveled by the kth ant.

Formulation for the zoning of protected ecological areas

According to the criteria of ecological protection, there are two planning objectives for the utility of a protected pattern. First, it is expected that the optimal protected pattern should yield the highest values for the average total ecological suitability. Second, compact configuration is always more desirable than fragmented protection. The conservation biology theory indicates that compact arrangements can protect a larger, contiguous habitat area and may be more effective for some species than unconnected sites (Reid and Murphy 1995). Furthermore, reducing site fragmentation may mitigate the effects of human activity and can facilitate the management of protected areas (Diamond et al. 1976; Morris 1991; Murcia 1995). Thus, the zoning protection problem can be expressed using the following Eqs.:

$$\text{Maximize} \quad \sum_{i} \text{Suit}_i x_i \tag{5}$$

$$\text{Maximize} \quad C_e \tag{6}$$

$$C_e = \frac{L_{\text{MaxSum}} - L_{\text{Sum}}}{L_{\text{MaxSum}} - L_{\text{MinSum}}} \tag{7}$$

$$\sum_{i} x_i = Q \tag{8}$$

$$x_i = \begin{cases} 1\text{if the cell } i \text{ is included in the protection} \\ 0 \quad \text{otherwise} \end{cases} \tag{9}$$

where Suit$_i$ is the ecological suitability of cell i, C_e is the compactness index of a protected pattern, and $L_{\text{Sum}}$is the sum of perimeter of a protected scenario. Once the area is known, the most compact form would be circular and the minimum sum of perimeter (L_{MinSum}) can then be calculated. On the contrary, if the selected sites are separate from each other, the maximum sum of perimeter L_{MaxSum} can then be obtained. Q is the total area of protection. Generally, a simple additive weighting method is employed to create a composite score for solving a multi-objective problem. Accordingly, the utility of protection can be defined as follows:

$$\text{Utility} = w_s \times S_e + w_c \times C_e \quad \forall\ w_s + w_c = 1 \tag{10}$$

$$S_e = \frac{\sum_i \text{Suit}_i x_i}{Q} \tag{11}$$

where S_e is the average total ecological suitability, w_s is the weight of ecological suitability, and w_c is the weight of compactness.

Ecological suitability serves as an important aid for zoning protection, and it can be estimated from a series of spatial variables (factors) that are retrieved from remote sensing and GIS data (Eastman et al. 1998). These factors include:

(1) The normalized difference vegetation index (NDVI)

Vegetation indices are commonly used for monitoring vegetation biomass and forecasting crop production. Many studies indicate that vegetation indices are well correlated with various vegetation properties including green leaf area, biomass, vegetation abundance, gross primary productivity, and photosynthetic activity (Sellers 1985; Huete 1988). NDVI is very sensitive to vegetation. In areas with more vegetation and better growth, the values of NDVI will be higher, and the ecological benefits will be better. Similarly, in areas with less vegetation and poorer growth, the values of NDVI will be smaller. Therefore, NDVI can be used as an important indicator for ecological suitability analysis. NDVI is calculated according to the following equation (Tucker 1979):

$$\mathrm{NDVI} = \frac{\mathrm{TM4} - \mathrm{TM3}}{\mathrm{TM4} + \mathrm{TM3}} \tag{12}$$

(2) Modified normalized difference water index (MNDWI)

An important task in ecological conservation is the protection of water resources in densely populated areas. The aquatic natural areas that are important for water quality and supply should be included in the protection. The identification of aquatic natural areas is the most important step for such protection. The Normalized Difference Water Index (NDWI) was first proposed by McFeeters (1996). This index can help separate water class from other land use classes in thematic mapper (TM) satellite images. The modified normalized difference water index (MNDWI) was further proposed by the substitution of a middle infrared band for the near infrared band so that water features can be efficiently enhanced (Xu 2006). MNDWI can be expressed as follows:

$$\mathrm{MNDWI} = \frac{\mathrm{TM2} - \mathrm{TM5}}{\mathrm{TM2} + \mathrm{TM5}} \tag{13}$$

(3) Urban development potential

The potential of urban development should also be considered because ecological conservation should not completely hinder future economic development. Furthermore, ecological protection could be subject to the stress from the activities of adjacent human-dominated landscapes. Negative effects emerge if the selected site is close to an area that has high potential for urban development. Therefore, development potential is regarded as a negative factor for conservation, and it can be estimated using the following Eq.:

$$\begin{aligned} S_{\mathrm{dev}} = {} & b_1 D_{\mathrm{District}} + b_2 D_{\mathrm{Towns}} + b_3 D_{\mathrm{Railways}} \\ & + b_4 D_{\mathrm{Expressways}} + b_5 D_{\mathrm{Roads}} + b_6 \mathrm{Slope} \end{aligned} \tag{14}$$

where D_{District} is the distance to district centre, $D_{\mathrm{Towns}}$is the distance to towns, D_{Railways} is the distance to railways, $D_{\mathrm{Expressways}}$ is the distance to expressways, D_{Roads} is the distance to roads, and $b_u (u = 1, 2, \ldots 6)$ is the weight of each variable and is subject to $b_1 + b_2 + b_3 + b_4 + b_5 + b_6 = 1$.

(4) Habitat heterogeneity

Habitat heterogeneity can be used to estimate the spatial distribution pattern of the heterogeneous environmental conditions of the study area (Svoray et al. 2005). Local species richness has long been known to be influenced by habitat heterogeneity (Goetz et al. 2007). Hence, an increase in habitat types will result in more species. Moreover, many studies indicate that ecological diversity should increase significantly with habitat heterogeneity on the landscape scale (Fahr and Kalko 2011). The habitats of land units are generally composed of three variables, including wetness index, slope orientation, and soil attributes. The wetness index is calculated using Eq. 15, which represents the ratio between flow accumulation and loss of moisture in a given cell:

$$\text{Wetness index} = \ln\left(\frac{Asi}{\tan\beta}\right) \tag{15}$$

where Asi is the specific catchment area (the upslope contributing area), which can be calculated using the ArcGIS 9.3 flow accumulation, andtan β is the tangent of the slope angle of the surface (Barling et al. 1994). The slope orientation can be obtained using the ArcGIS 9.3 spatial analyst function. The detailed map of soil attributes was prepared based on field surveys and expert knowledge. Habitat heterogeneity was calculated using the Shannon–Weaver Index:

$$H_{\mathrm{h}} = -\sum_j P_j(\ln P_j) \tag{16}$$

where P_j is the relative abundance of each individual habitat, which can be calculated based on 16 cells that occupy the grid.

The multi-criteria evaluation (MCE) (Eastman et al. 1998) method is used to estimate the ecological suitability according to the above mentioned spatial variables, which should be standardized within the range of [0, 1]. The total score of ecological suitability was calculated using a weighted linear combination method (Malczewski 2006):

$$\begin{aligned} S_{\mathrm{e}} = {} & w_1 \times \mathrm{NDVI} + w_2 \times \mathrm{MNDWI} + w_3(1 - S_{\mathrm{dev}}) \\ & + w_4 \times H_{\mathrm{h}} \end{aligned} \tag{17}$$

where w_1, w_2,w_3, and w_4 are the weights for each factor, and the total of all the criterion weights is equal to 1.

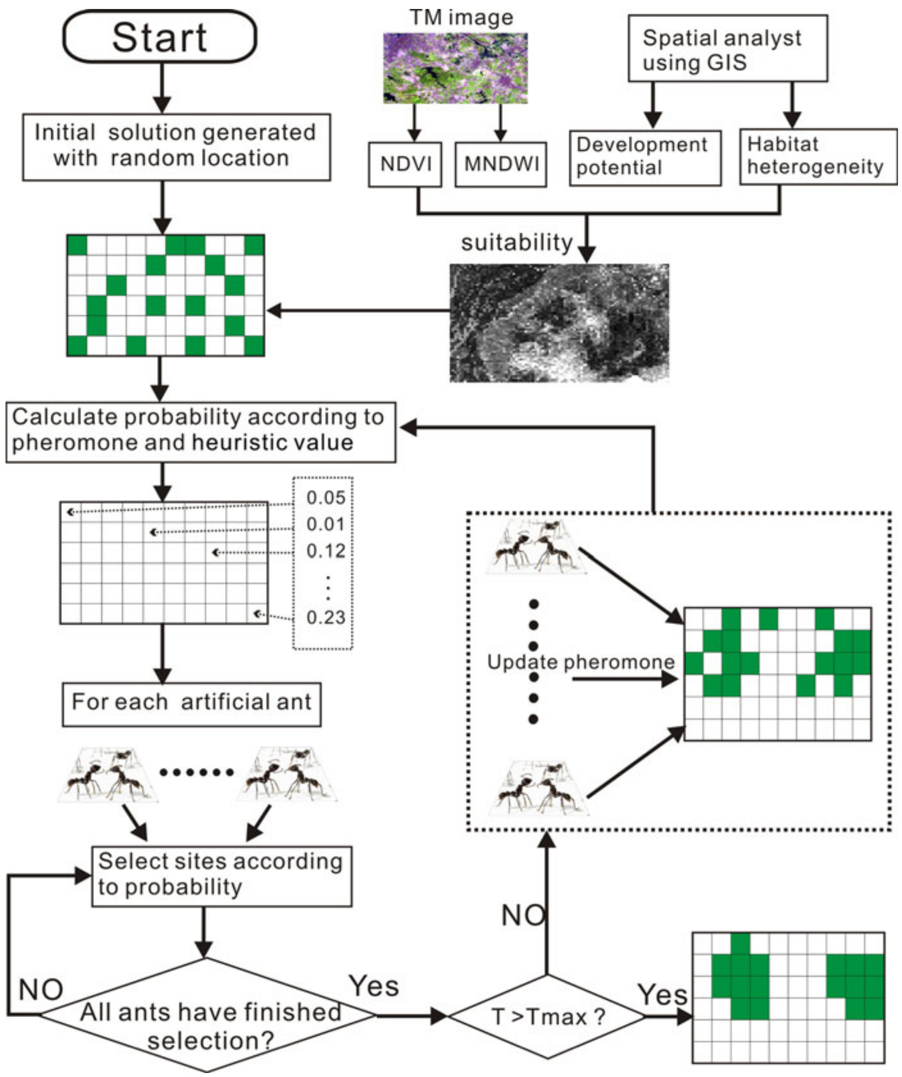


Fig. 1 The procedure of zoning protected ecological areas by integrating remote sensing, GIS and modified ACO

Modified ACO algorithm for zoning protection

ACO is a meta-heuristic technique that uses artificial ants to find solutions to combinatorial optimization problems (Dorigo et al. 1996). Recently, ACO has been modified to solve complex point and path optimization problems by using a rich set of spatial information (Li et al. 2009a, 2009b). This article further modifies and extends the ACO algorithm to solve zoning protection problems in large areas. Figure 1 illustrates the procedure of zoning protected ecological areas by integrating remote sensing, GIS and modified ACO. Remote sensing data can be used to obtain the information of vegetation indices and water indices. GIS can provide the tools for the analysis of a large amount of spatial data. Modified ACO is designed to form protected ecological areas based on suitability map. The detailed modifications of ACO for solving zoning protection problem are provided in the following sections.

Solution construction

In forming protected ecological areas, artificial ants are randomly positioned in the study region at the start of optimization. Artificial ant will visit a cell and lay down pheromone on the cell, the amount of deposited pheromone is related to the total utility of this cell in forming the entire protected ecological areas. The utility of protection is composed of ecological suitability and compactness. The larger the amount of pheromone, the more the ants will be attracted to select this cell. A larger amount of pheromone is in turn deposited on the cell. The solution is constructed after all ants have located their best sites.

Improved selection strategy

Sites selection is an important step for ants in the formation of protection areas. The probability that a cell i will be selected by the kth ant at time t is modified according to Eq. 1:

$$p_i^k(t) = \begin{cases} \frac{[\tau_i(t)]^{\alpha} \cdot [\eta_i(t)]^{\beta}}{\sum_{x \in \text{allowed}_k} [\tau_x(t)]^{\alpha} \cdot [\eta_x(t)]^{\beta}}, & \text{if i } \in \text{allowed}_k \\ 0 & \text{otherwise} \end{cases} \tag{18}$$

where allowed_k represents the tabu list, which is defined to mask out the selected cells which should not be visited again by other ants.

A heuristic function is designed to guide the path searching of artificial ants so that the computation time is significantly reduced. In this paper, the heuristic function $\tau_i(t)$ is incorporated into the ecological suitability at cell i to guide the walking of ants. An artificial ant is more likely to select a cell with a higher suitability value so that a plausible protection area can be formed. Therefore, the heuristic function is defined as follows:

$$\eta_i = \frac{\text{Suit}_i}{\sum_x \text{Suit}_x} \tag{19}$$

where Suit_i is the ecological suitability at cell i, and $\sum_x \text{Suit}_x$ is the sum of the suitability for all cells in the study area.

According to the selection probability, ants select a site through the roulette wheel selection technique, which can be achieved using random pointers. Assuming there are four grids with the probabilities of being chosen as 0.1, 0.2, 0.3 and 0.4, respectively, the sum of the probability is 1 (Fig. 2). The cumulative probabilities are obtained by summing up the selection probability of these four grids, that is, 0.1, 0.3, 0.6 and 1. A random number between 0 and 1 representing the pointer, saying 0.7, is generated and compared with the cumulative probabilities. If the cumulative probability of the first n grids is greater than 0.7, then the n-th grid is selected. In this example, the cumulative probability of the first four grids is greater than 0.7 and hence the fourth grid is selected. If the random number generated is 0.45, then the third grid is selected, as shown in Fig. 2.

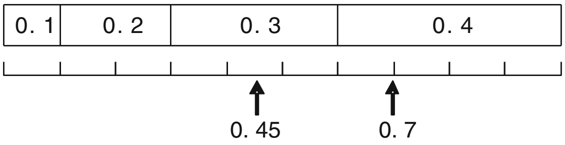


Fig. 2 Sites selection through the roulette wheel selection technique

In the basic TSP algorithm, a random number is generated for one ant to select a city. When an ant is deciding which city to visit at the current step, the selection probability of all cities excluded by the tabu list are accumulated until the sum exceeds the value of a pre-generated random number. This procedure will be executed each time when an ant selects a new city to visit. Therefore, if the number of city is set as m, an ant visiting all cities requires the implementation of such selection procedure for m-1 times. Apparently, this selection strategy requires intensive computation, so it is, thus, inadequately efficient.

In this study, a more efficient strategy is designed to select sites for artificial ants. In this improved selection strategy, a set of m random numbers instead of one are generated in one time for all ants, and the sequencing of the selection probability is executed before accumulation. Taking the previous example in Fig. 2, two random numbers are simultaneously generated, say 0.45 and 0.7. The selection probability of each grid is then sorted in ascending order, and the cumulative probabilities are then calculated. The cumulative probability of the first three grids is 0.6, which is greater than 0.45, so the third grid is selected. On the other hand, the cumulative probability of the first four grids exceeds 0.7 and is also selected. Using this strategy, the selection can be accomplished within only one round of accumulation. Compared with the basic selection strategy, the improved strategy is more efficient, especially when a large number of sites are involved.

Modified pheromone updating mechanism

Pheromone updating is the core of ACO for spatial optimization. Some modifications in pheromone updating mechanism are designed to address zoning protection problem. Firstly, it is important to incorporate the utility function in ACO so that the protected area can be formulated using the ant algorithm. The variable, $\frac{1}{L_k}$, in Eq. 4 can then be replaced by the utility of the ecological protection. Another important modification is to introduce the strategy of neighborhood pheromone diffusion. This technique is conducted by incorporating a distance decay function in pheromone updating. Therefore, Eq. 4 should be modified as follows:

$$\Delta\tau_x^k(t) = \begin{cases} \dfrac{Q \times \text{Utility}}{d(x)+1}, & \text{if } x \text{ falls within } 5 \times 5 \text{ window of cell } i \text{ at time } t \\ 0 & \text{otherwise} \end{cases} \tag{20}$$

where $d(x)$ is the distance from the central cell (i), and is used to address the effects of neighborhood on site selection. A site should have a higher probability of being selected if its neighbors have already been included in the protection.

Model implementation and results

Study area and spatial data

Dongguan City in the Pearl River Delta has been selected for the testing of the proposed model to solve a practical problem, that is, the zoning of protected ecological areas. The city is along the corridor between Guangzhou and Shenzhen (Fig. 3), with a total area of 2,465 square kilometers. A large amount of agricultural land has been converted into urban areas in Dongguan. A series of environmental and ecological problems driven by rapid urban development in this region have been reported in many studies (Yeh and Li 1997; Seto et al. 2002). There is an urgent demand to establish protected ecological areas to sustain the environmental quality and the ecological functions for human welfare and further development. In this article, the previously mentioned ACO-based model described is applied to establish ecological conservation areas in Dongguan City.

A satellite Landsat TM image of Dongguan acquired on March 04, 2008 was classified to obtain land use information. Then, NDVI and MNDWI of the study aera were calculated using the Landsat TM Image. Then, 30-m DEM data is used to produce the slope, wetness index and slope orientation of the region. The soil associations of the study area were classified into 13 categories based on field surveys and expert knowledge. Then, habitat heterogeneity is obtained by integrating wetness index, slope orientation and soil attributes based on Eq. 16. Five proximity variables (distance to district, distance to towns, distance to railways, distance to expressways, and distance to roads) and slope are used to produce the potential of urban development. All these spatial variables (factors) are converted into a raster format using GIS techniques (Fig. 4).

Ecological suitability analysis

Suitability analysis is the process to determine whether the land resource is suitable for some specific use (Steiner et al. 2000). This type of analysis involves a number of spatial variables (factors) which evaluate the suitability score. The weights for each variable should be decided according to expert experiences and domain knowledge. A common method is adopting Saaty's pairwise comparison to obtain these weights (Eastman et al. 1998). Table 1 shows the weights of different factors as suitability evaluation indicators. A consistency ratio was used to examine the consistency of the matrix. The weight matrix should be re-evaluated if the ratio value is greater than 0.10. Figure 5 demonstrates the final ecological suitability map, which was created by integrating the above mentioned spatial variables and weights using the ArcGIS 9.3 spatial overlay analyst function. This suitability map servers as an important aid for optimizing spatial pattern of protected areas.

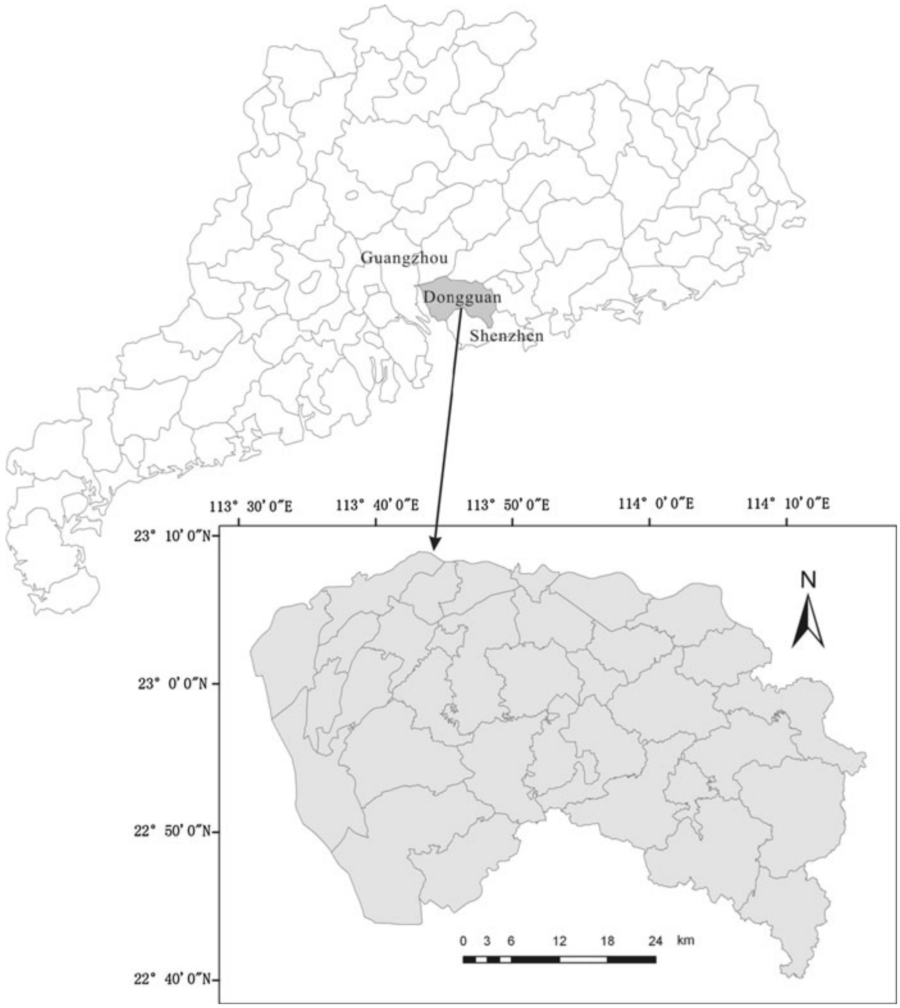


Fig. 3 Location of Dongguan in the Pearl River Delta

Zoning protected areas using modified ACO

The modified ACO model was used to search for the optimal pattern for the protected ecological area. This optimization model involves some parameters, which could affect the optimization results. In Eq. 18, the parameters of α and β control the relative weight of the pheromone trail and the heuristic value. These parameters should be defined before running the model as they directly affect the results of optimization. Therefore, the key lies in how to determine the parameters of α and β in order to improve the optimization utility. A greedy search strategy is designed to determine these parameters. We assigned values to parameter α from 0.5 to 10, with increments of 0.5. Parameter β was varied between 0.25 and 5, with increments of 0.25. Then, the average utility value of the formed protection is calculated with different combinations of (α, β). Once the average utility value reaches its maximum, the values of α and β are considered as the optimal parameter settings. Figure 6 illustrates the utility value obtained using different combinations of α and β values with running the ACO model for 1,000 iterations. As shown in Fig. 6, when $\alpha > 2$, the utility value does not appear to be sensitive to the choice of α; and differences along parameter α are relatively insignificant. When α takes small values (i.e., α varies from 0 to 2), the utility value is very sensitive to parameter α, and it is relatively small. When $\alpha > 2$ and $\beta < 1$, the modified ACO model can obtain a relatively high utility value. In the experiment, the maximum value of average utility was 0.67151 when $\alpha = 8$ and $\beta = 0.75$. These parameter settings ($\alpha = 8$, $\beta = 0.75$) were then used to optimize protected areas in Dongguan.

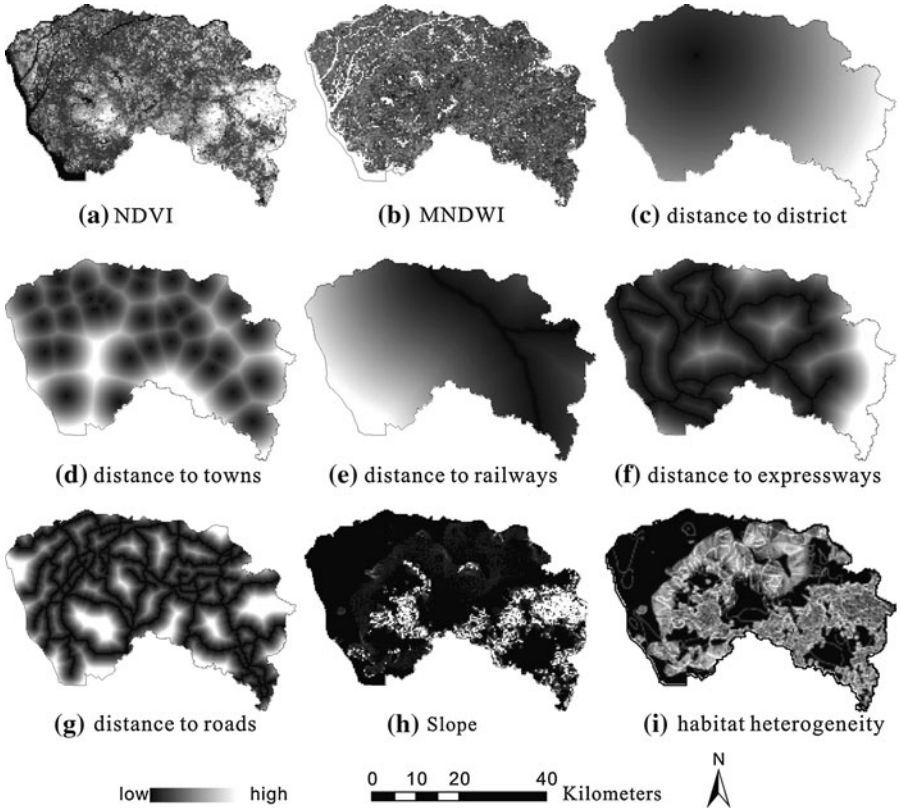


Fig. 4 Various spatial variables for suitability analysis using remote sensing and GIS data

Table 1 Weights for calculating ecological suitability

Factors	NDVI	MNDWI	Development potential	Habitat heterogeneity
Weights	0.345	0.155	0.286	0.214

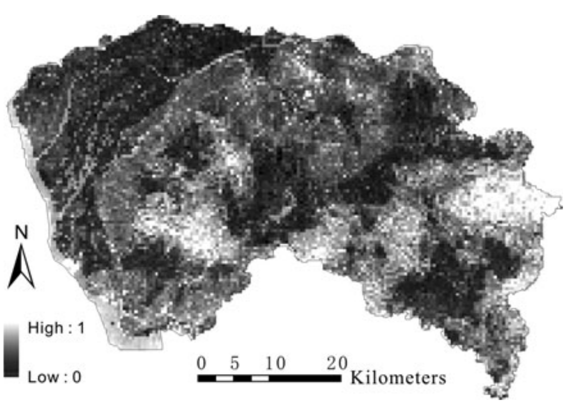


Fig. 5 Ecological suitability map of Dongguan

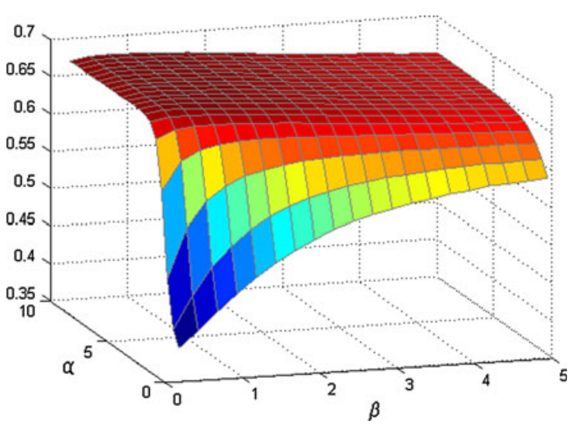


Fig. 6 Utility values obtained using different combinations of parameters (α, β) settings

The required area for protection is set to be 880 square kilometers with reference to the strategic planning of Dongguan. The weights for the average total ecological suitability and the compactness in Eq. 10 are both assigned as 0.5. Figure 7 illustrates the optimization process for the spatial patterns of protection by using modified ACO model. At the initial stage, artificial ants are randomly located in the study area. As the iteration continues, these ants will explore the space and try to find the best locations through cooperation and interaction. The formed patterns become more and more compact because of the feedback effects of ants. It is found that artificial ants have almost occupied the best locations only after 500 iterations. The spatial pattern becomes stabilized when the iteration reaches approximately 1,000. As illustrated in Fig. 7, the final optimized pattern clearly shows that each artificial ant is allocated at the sites with high ecological suitability and the spatial form is very compact.

As shown in Fig. 8, the value of the utility function will increase significantly during the initial stage. The utility value will become stabilized when the iteration is greater than 400. The optimization spends about only 119 s when using a computer with a Pentium IV 3.2 GHz CPU.

Similar to other heuristic methods, the ACO algorithm may be affected by some randomness. It is necessary to carry out optimizations for several times to determine whether ACO can produce stable results. To test the robustness of the modified ACO model, the optimization is repeated ten times. Then, the variance of utility value under different optimization processes is calculated. The variance, which can provide a way to measure the robustness of the ACO model, is defined as follows:

$$V_t = (1/T)\sum_{i=1}^{T}(U_t^i - \widehat{U}_t)^2 \tag{21}$$

where V_t is the variance of the utility value for the tth iteration with different optimization times, T is the number of optimization times (here $T = 10$), $\widehat{U}_t$ is the mean utility value of the tth iteration under optimization times, and U_t^i represents the utility value of the tth iteration for the ith optimization.

Then, the mean variance of all iteration is calculated to measure the robustness of the ACO model:

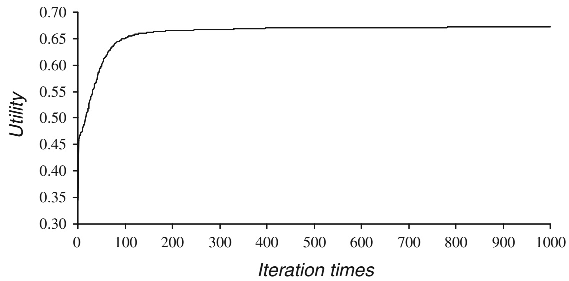
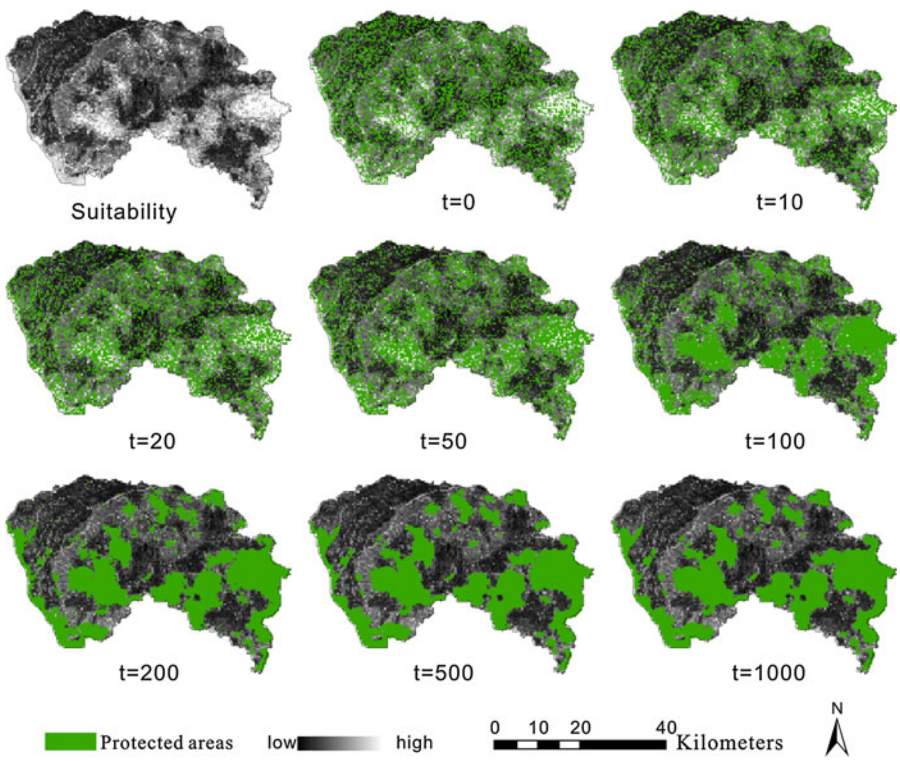


Fig. 8 Utility improvements with iterations by the modified ACO model

Fig. 7 The optimization process of protection spatial patterns by using modified ACO model



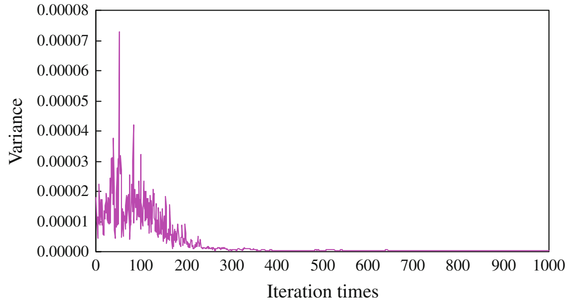


Fig. 9 The variance of utility value with iterations by the modified ACO model

$$MV = \sum_{t=1}^{I} \frac{V_t}{I} \qquad (22)$$

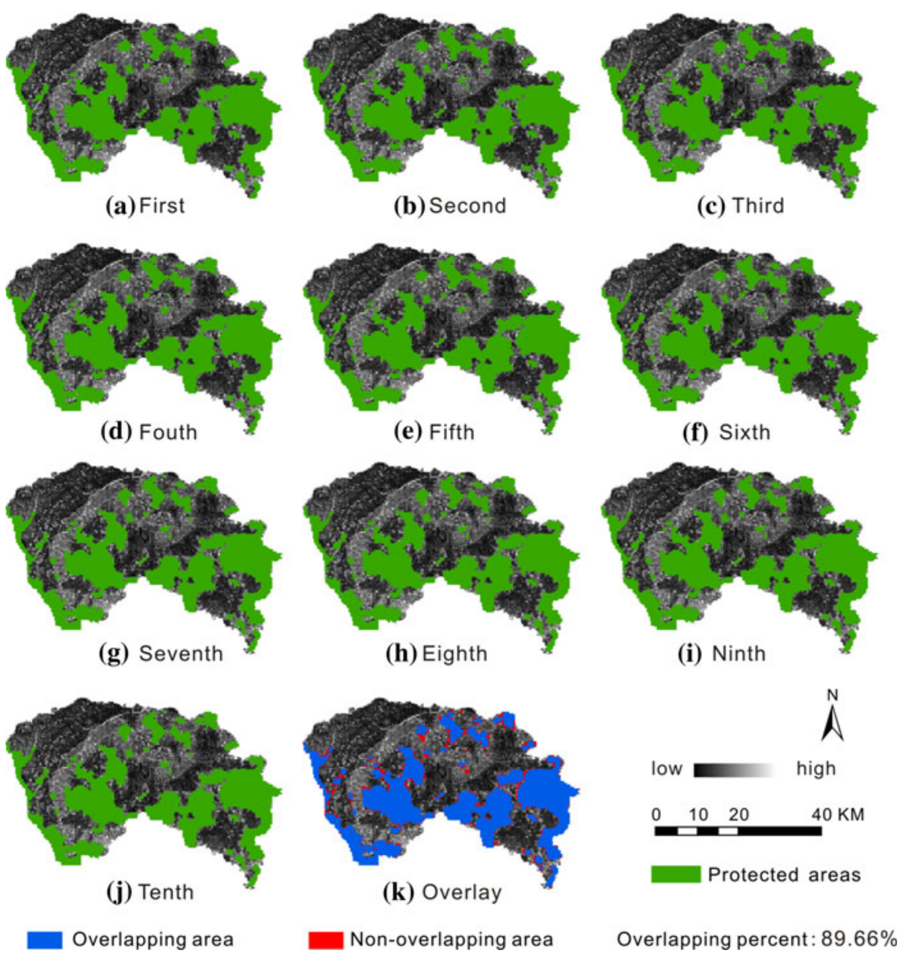
where I is the number of iteration times.

Figure 9 shows the variance of the utility value with iterations using the modified ACO model. The variance value is not stabilized during the initial stage. However, the variance value is almost equal to 0 after the iteration is greater than 400. The MV value of modified ACO model is 0.000013. A smaller MV value indicates that the optimization model is more stable. The above analysis shows that the robustness of the ACO model is good.

Finally, the overlapping of the zoning results is examined. As shown in Fig. 10, there are only small non-overlapping areas in the fringe of the formed protection. The overlapping areas account for a considerable proportion at 89.66%. The good overlapping indicates that the modified ACO model can produce quite stable results for forming protected areas.

A further experiment was conducted to compare the performances of the modified ACO model with those of the five algorithms: (a) the simulated annealing (SA) (Verdiell et al. 2005); (b) the genetic algorithm

Fig. 10 Overlay of the optimization results by repeatedly running modified ACO ten times



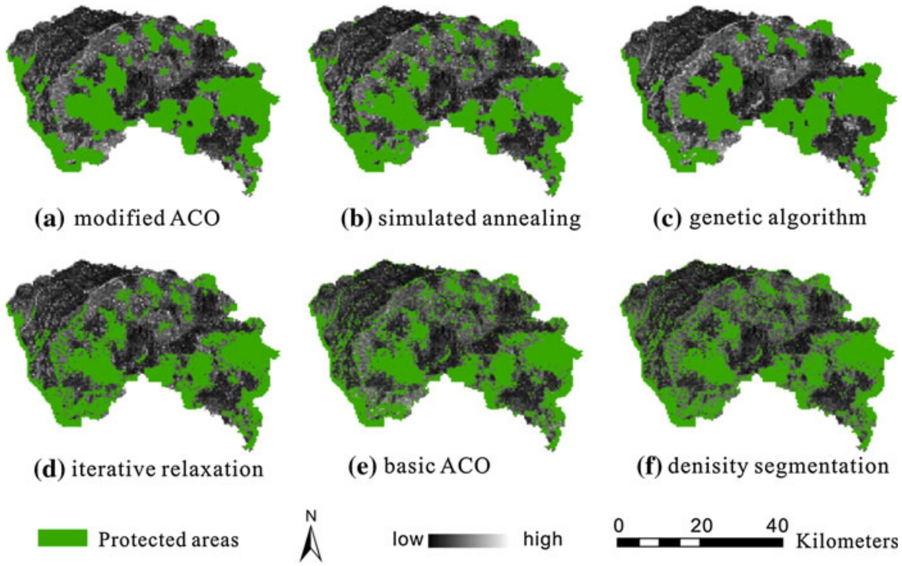


Fig. 11 Zoning of protected ecological areas using modified ACO, SA, GA, IR, basic ACO and DS methods

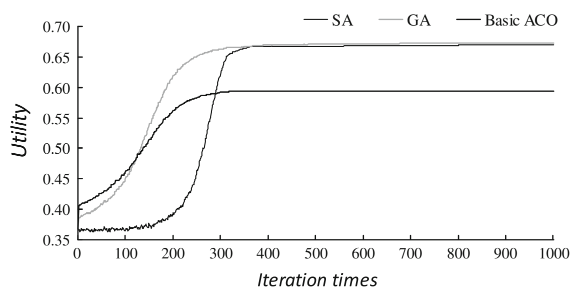


Fig. 12 Utility values with iterations by using SA, GA and basic ACO

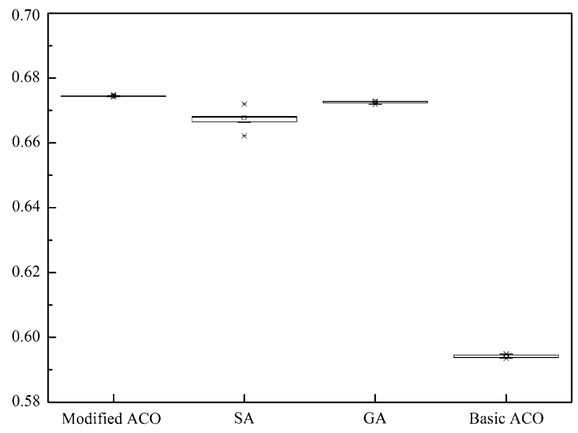


Fig. 13 Boxplot of the utility value for modified ACO, SA, GA and basic ACO

(GA) (c) the iterative relaxation (RI) (Eastman et al. 1995); (d) the basic ACO model; and (e) the density slicing (DS) (Li and Yeh 2001). The performance of these four methods is compared based on the average total utility using the same dataset so that their performances can be compared with that of the modified ACO model.

The SA method is a well-known heuristic algorithm that has been widely used to solve combination optimization problems. At the beginning of the procedure, initial locations of sites, which are composed of the protected areas (solution), were randomly generated, and the initial value of the utility function was calculated. Then, a small change was made to the current solution and the utility value of new solution was obtained. If the new utility value is greater than the previous, then the new solution is accepted; otherwise, the new solution is accepted on a random basis as specified by the Metropolis procedure (Kirkpatrick et al. 1983).

Table 2 Comparison of the *MV* value and the computation time using the modified ACO, SA, GA and basic ACO

	Modified ACO	SA	GA	Basic ACO
MV value	0.000013	0.000380	0.000017	0.000021
Time(s)	119	3383	864	1558

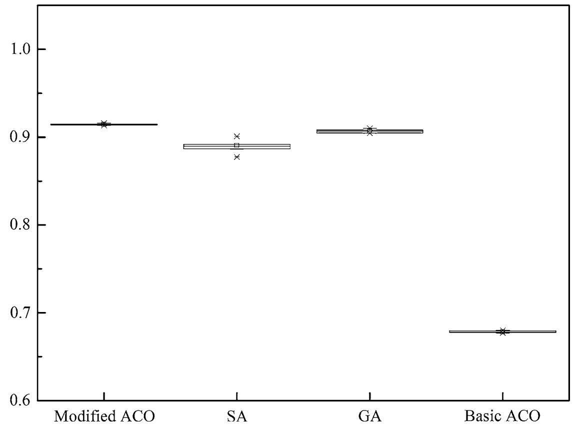


Fig. 14 Boxplot of compactness for modified ACO, SA, GA and basic ACO

The concept of GA was developed by Holland (1975). The genetic algorithm adopts an evolutionary process to solve optimization problems based on the natural selection theory. GA and ACO both work with a number of initial solutions, the so-called population, which is usually randomly generated. Then new solutions are created by applying a series of genetic operators (selection, crossover, mutation and replacement). GA has been proven excellent in quickly finding solutions for complex optimization problems (Mitchell 1996).

The RI method can be implemented using the following steps (Eastman et al. 1995): First, all cells with a suitability value greater than an initial threshold are selected. Then, spatially discontinuous cells are removed, and contiguous groups of feasible cells are considered to be candidate regions. Third, candidate regions having an area less than that of the minimum size are also removed. The remaining candidates meet both the suitability and size requirements. Finally, the above processes are repeated until the utility improvement becomes stabilized. This method can be used to generate feasible alternatives, but it does not generate regions with a given size.

The basic ACO model also incorporates the utility function by addressing the criteria of natural protection, which involves the average total ecological suitability and the compactness of the pattern. The values of parameters settings (α, β) in the basic ACO model are similar to that of the modified ACO model so that their performances can be compared. However, the neighborhood pheromone diffusion strategy and the improved selection strategy are not adopted in this basic ACO algorithm.

The DS method is based only on the ranking of suitability values by slicing the density of suitability score (Li and Yeh 2001). Cells with higher suitability values are selected to establish the protected areas. The DS method is very simple to implement, but the spatial constraint cannot be considered in this method.

The zoning results of the above mentioned approaches are shown in Fig. 11a–f. Figure 12 shows the utility values with iterations by using SA, GA and basic ACO. It indicates that GA can achieve convergence more quickly than SA and basic ACO. As shown in Table 2, the values of *MV* for modified ACO, SA, GA and basic ACO are 0.000013, 0.000380, 0.000017 and 0.000021, respectively. The values of *MV* for modified ACO, GA and basic ACO are very small, which indicates that these three methods are robust. Then, utility and compactness values for modified ACO, SA, GA and basic ACO were computed in ten runs. Matlab7.0 was used to create boxplots of utility and compactness (Figs. 13, 14), which show no outliers for modified ACO, GA and basic ACO. However, some outliers are identified in the boxplots for the SA method. This indicates that the optimization results of SA are unstable.

Table 3 depicts the obtained utility values and compactness of these six methods. As illustrated in Fig. 11 and Table 3, the modified ACO model can generate protected areas with a maximum utility value and compact pattern. GA and SA also can obtain good performances in terms of the utility value and spatial pattern. However, GA and SA require a significantly

Table 3 Comparison of the utility value and compactness using modified ACO, SA, GA, IR, basic ACO and DS methods

	Modified ACO	SA	GA	IR	Basic ACO	DS
Utility value	0.67509	0.65969	0.67214	0.63928	0.60803	0.59832
Compactness	0.91553	0.87010	0.90628	0.80005	0.70220	0.68288

longer computation than the modified ACO model (Table 2). The computation time for the modified ACO method only amounts to 3.52 and 13.77% of those of the SA and GA methods, respectively. The RI method generally has the capability to derive suitable and continuous area (Fig. 11c). However, the spatial pattern of this method is less compact than that of modified ACO and GA (Table 3). Neither the basic ACO method nor the DS method achieves good performances as both methods generate fragmented protected areas. The basic ACO method only has a slight improvement of in utility value and compactness compared with the DS method. It is noted that the modified ACO model can generate more compact patterns and much more efficient results than the basic ACO model. This is because the neighborhood pheromone diffusion strategy and the improved selection strategy are adopted in the modified ACO algorithm. The experiment indicates that the modified ACO method is an efficient and effective optimization technique for generating alternative protection.

Conclusion

Rapid changes in land use patterns, especially urban expansion, will impose significant threats to natural ecosystems. The establishment of protected areas appears to be a good strategy for conserving our ecosystems. Existing zoning methods applied to protected areas are mostly qualitative and are highly dependent on the knowledge of experts (Verdiell et al. 2005). Zoning a protected ecological area under spatial constraint is belonging to the NP-hard problem because of its huge combinatorial solution space. It is impossible to solve such difficult problems within a reasonable amount of time using a precise enumeration method. Heuristic algorithms have the advantage of speed and simplicity in solving combination optimization problems.

ACO, a recently developed heuristic method, has already been proven useful for providing a good solution to complex spatial optimization problems (Li et al. 2009a). This study demonstrates how ACO can be integrated with remote sensing and GIS for the zoning of protected ecological areas involving large amounts of spatial data. This paper makes a number of contributions in four aspects by adopting and modifying the ACO-based algorithm:

1. The utility function by addressing the criteria of ecological protection is incorporated into ACO algorithm.
2. The neighborhood pheromone diffusion strategy is designed to improve the compactness of protection patterns. This is implemented by incorporating a distance decay function into pheromone updating.
3. An improved selection strategy is adopted to accelerate the progress of sites selection for artificial ants. This strategy makes the modified ACO much more efficient than SA, GA and basic ACO in forming protected areas.
4. To improve the optimization utility, a greedy search strategy is proposed to determine the optimal parameters of the ACO algorithm.

The modified ACO model was then applied to the zoning of protected ecological areas in Dongguan, a rapidly urbanized region, which is involved an area of 25,483 cells. The proposed model can find the near-optimal solutions with a good convergence rate. Furthermore, this model can generate protected areas with the maximum ecological suitability value and compact pattern. This indicates that the modified ACO method is an efficient and effective optimization technique for generating optimal protections.

The proposed method is compared with other zoning methods, such as SA, GA, IR, basic ACO, and DS. The comparison indicates that the modified ACO method can yield much better performance than other zoning methods. It is found that modified ACO can improve the compactness over SA, GA, IR, basic ACO and DS by 5.22, 1.02, 14.43, 30.37 and 34.07%, respectively. Modified ACO also has the improvement of the total utility over SA, GA, IR, basic ACO and DS by 2.33, 0.44, 5.60, 11.03 and 12.83%, respectively. In addition, Modified ACO is much more efficient than SA, GA and the basic ACO method, the computation time of modified ACO, SA and basic ACO are 119, 3383, 864 and 1558 s, respectively. Although GA may have close optimization results with modified ACO, the computation time of modified ACO is only 13.77% of that of GA.

There are some limitations in using the modified ACO model in zoning protection. First, the movement of the ant is a random walk in the proposed model, without constrains or barriers. In practical protection design, some barriers, like major highways or large rivers, should be excluded from protection areas. Second, the optimization model is static. However, urban dynamics have significant effects on spatial optimization. Hence, the optimization model should use the dynamic urban patterns as inputs during the planning period. In future work, this ACO model will be integrated with an urban cellular automaton for the exploration of protection scenarios under land-use dynamics.

Although the present study only demonstrates how the ACO-based zoning model is applied to generate protected ecological areas, by altering the objective functions, the model can be applied to other application, such as environmental planning, reserve selection, resource management and farmland protection.

Acknowledgments This study was supported by the National Basic Research Program of China (973 Program) (GrantNo. 2011 CB707 103), the National Natural Science Foundation of China (Grant No. 40901187), the National Natural Science Foundation of China (Grant No. 41171308) and the Key National Natural Science Foundation of China (Grant No. 40830532).

References

Barling R, Moore I, Grayson R (1994) A quasi-dynamic wetness index for characterizing the spatial distribution of zones of surface saturation and soil water content. Water Resour Res 30:1029–1044

Bos J (1993) Zoning in forest management: a quadratic assignment problem solved by simulated annealing. J Environ Manag 37:127–145

Carsjens G, van der Knaap W (2002) Strategic land-use allocation: dealing with spatial relationships and fragmentation of agriculture. Landsc Urban Plan 58:171–179

Church R, Gerrard R, Gilpin M, Stine P (2003) Constructing cell-based habitat patches useful in conservation planning. Ann Assoc Am Geogr 93:814–827

Cocks K, Baird I (1989) Using mathematical programming to address the multiple reserve selection problem: an example from the Eyre Peninsula, South Australia. Biol Conserv 49:113–130

Deng X, Huang J, Rozelle S, Uchida E (2008) Growth, population and industrialization, and urban land expansion of China. J Urban Econ 63:96–115

Diamond JM, Terborgh J, Whitcomb RF, Lynch JF, Opler PA, Robbins CS, Simberloff DS, Abele LG (1976) Island biogeography and conservation: strategy and limitations. Science 193(4257):1027–1032

Dorigo M (1992) Optimization, learning and natural algorithms. PhD Thesis, Politecnico di Milano, Dipartimento di Elettronica, Italy

Dorigo M, Gambardella L (1997) Ant colony system: a cooperative learning approach to the travelingsalesman problem. IEEE Trans Evol Comput 1:53–66

Dorigo M, Maniezzo V, Colorni A (1996) Ant system: optimization by a colony of cooperating agents. IEEE Trans Syst Man Cybern 26:29–41

Eastman J, Jin W, Kyem P, Toledano J (1995) Raster procedures for multi-criteria/multi-objective decisions. Photogr Eng Remote Sens 61:539–547

Eastman J, Jiang H, Toledano J (1998) Multi-criteria and multi-objective decision making for land allocation using GIS. Multicriteria analysis for land-use management. Dordrecht, Kluwer Academic Publishers, pp 227–251

Fahr J, Kalko EKV (2011) Biome transitions as centres of diversity: habitat heterogeneity and diversity patterns of West African bat assemblages across spatial scales. Ecography 34:177–195

Geneletti D, van Duren I (2008) Protected area zoning for conservation and use: a combination of spatial multicriteria and multiobjective evaluation. Landsc Urban Plan 85:97–110

Goetz SJ, Steinberg D, Dubayah R, Blair B (2007) Laser remote sensing of canopy habitat heterogeneity as a predictor of bird species richness in an eastern temperate forest, USA. Remote Sens Environ 108:254–263

Hof J, Joyce L (1993) A mixed integer linear programming approach for spatially optimizing wildlife and timber in managed forest ecosystems. For Sci 39:816–834

Holland JH (1975) Adaptation in natural and artificial systems: an introductory analysis with application to biology. In Control and artificial intelligence. University of Michigan Press, Ann Arbor

Huete A (1988) A soil-adjusted vegetation index (SAVI). Remote Sens Environ 25:295–309

Kirkpatrick S, Gelatt CD, Vecchi MP (1983) Optimization by simulated annealing. Science 220(4598):671–680

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

Li X, Liu XP (2008) Embedding sustainable development strategies in agent-based models for use as a planning tool. Int J Geogr Inf Sci 22:21–45

Li X, Yeh AGO (2001) Zoning land for agricultural protection by the integration of remote sensing, GIS, and cellular automata. Photogr Eng Remote Sens 67:471–477

Li X, He JQ, Liu XP (2009a) Ant intelligence for solving optimal path-covering problems with multi-objectives. Int J Geogr Inf Sci 23(7):839–857

Li X, He JQ, Liu XP (2009b) Intelligent GIS for solving high-dimensional site selection problems using ant colony optimization techniques. Int J Geogr Inf Sci 23:399–416

Liu XP, Li X, Peng XJ, Li H, He JQ (2008a) Swarm intelligence for classification of remote sensing data. Sc China Ser D 51:79–87

Liu XP, Li X, Liu L, He JQ, Ai B (2008b) A bottom-up approach to discover transition rules of cellular automata using ant intelligence. Int J Geogr Inf Sci 22:1247–1269

Liu XP, Li X, Liu L, He JQ, Ai B (2008c) An innovative method to classify remote-sensing images using ant colony optimization. IEEE Trans Geosci Remote Sens 46:4198–4208

Malczewski J (2006) GIS-based multicriteria decision analysis: a survey of the literature. Int J Geogr Inf Sci 20:703–726

McFeeters S (1996) The use of the normalized difference water index (NDWI) in the delineation of open water features. Int J Remote Sens 17:1425–1432

Mitchell M (1996) An introduction to genetic algorithms. MIT Press, Cambridge
Morris MG (1991) The management of reserves and protected areas. In: Spellerberg IF, Goldsmith FB, Morris MG (eds) The scientific management of temperate communities for conservation: the 31st symposium of the British Ecological Society Southampton 1989. Blackwell, Boston, pp 232–247
Murcia C (1995) Edge effects in fragmented forests: implications for conservation. Trends Ecol Evol 10:58–62
Önal H, Briers RA (2002) Incorporating spatial criteria in optimum reserve network selection. Proc R Soc Lond B 269:2437–2441
Possingham H, Ball I, Andelman S (2000) Mathematical methods for identifying representative reserve networks. In: Ferson S, Burgman M (eds) Quantitative methods for conservation biology. Springer-Verlag, New York, pp 291–306
Reid T, Murphy D (1995) Providing a regional context for local conservation action. Bioscience 45:84–90
Sellers P (1985) Canopy reflectance, photosynthesis and transpiration. Int J Remote Sens 6:1335–1372
Seto K, Woodcock C, Song C, Huang X, Lu J, Kaufmann R (2002) Monitoring land-use change in the Pearl River Delta using Landsat TM. Int J Remote Sens 23:1985–2004
Snyder S, Haight R, ReVelle C (2005) A scenario optimization model for dynamic reserve site selection. Environ Model Assess 9:179–187
Steiner F, McSherry L, Cohen J (2000) Land suitability analysis for the upper Gila River watershed. Landsc Urban Plan 50(4):199–214
Svoray T, Bar P, Bannet T (2005) Urban land-use allocation in a Mediterranean ecotone: habitat heterogeneity model incorporated in a GIS using a multi-criteria mechanism. Landsc Urban Plan 72(4):337–351
Tucker C (1979) Red and photographic infrared linear combinations for monitoring vegetation. Remote Sens Environ 8:127–150
Verdiell A, Sabatini M, Maciel M, Iglesias R (2005) A mathematical model for zoning of protected natural areas. Int Trans Oper Res 12:203–213
Xiao N, Bennett D, Armstrong M (2007) Interactive evolutionary approaches to multiobjective spatial decision making: a synthetic review. Comput Environ Urban Syst 31:232–252
Xu H (2006) Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. Int J Remote Sens 27:3025–3033
Yeh AGO, Li X (1997) An integrated remote sensing and GIS approach in the monitoring and evaluation of rapid urban growth for sustainable development in the Pearl River Delta, China. Int Plan Stud 2:21–193
Yeh AGO, Li X (1999) Economic development and agricultural land loss in the Pearl River Delta. China Habitat Int 23:373–390