Accepted for publication in ASCE Journal of Urban Planning and Development

Sustainable Land Use Planning for a Downtown Lake Area in Central China: a Multi-objective Optimization Approach Aided by Urban Growth Modeling

Bo Huang *
Department of Geography and Resource Management, The Chinese University of Hong Kong, Hong Kong
Institute of Space and Earth Information Science, The Chinese University of Hong Kong, Hong Kong
E-mail: bohuang@cuhk.edu.hk

Wenting Zhang
Department of Geography and Resource Management, The Chinese University of Hong Kong, Hong Kong
E-mail: wentingjj@cuhk.edu.hk

* Corresponding author. Department of Geography and Resource Management, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong
Abstract: Urban land use planning is becoming increasingly complex in China because it requires estimating the future urban extent and balancing its various aspects, e.g., provision of enough housing and employment opportunities and preservation of the environment. This paper presents an integrative approach to the problem of sustainable land use planning. Specifically, an urban growth simulation tool, a cellular automata model, was used to determine the extent of the urban area, and a neural network approach was used to quantitatively predict land use structure in the projected year. A multi-objective optimization genetic algorithm was then developed to search for Pareto-optimal urban land use plans within the urban extent determined by the urban growth model. To search for a sustainable urban land use plan, social, economic, and environmental objectives were defined that reflect the multiple objectives of the urban system. In addition to these non-spatial goals, objectives concerning the spatial distribution of land usages were proposed. These are classified as local and global spatial objectives. A case study of the Donghu Lake watershed was conducted. Donghu Lake is one of the largest downtown lakes in central China, and the watershed area is undergoing rapid urbanization and suffering from non-point source water pollution. The study was carried out to validate the proposed method, and it demonstrated that the models are appropriate for areas undergoing urban expansion.

Key words: multi-objective optimization; watershed; cellular automata; genetic algorithm; non-point source pollution

Introduction

Land use planning can be defined as the process of allocating different activities or uses to specific units of area within a region (Stewart et al., 2004). Such allocation provides rules and guidance for urban management and future development. The activities or uses always involve residential land, industry, commercial activities, green space, and public service (Cao et al., 2011). Because it involves multiple and usually conflicting demands from different land use groups such as government, merchants, residents, and drivers, urban planning has become a multi-objective problem. Increased inclusion of objectives leads to different demands on the expected results (Steward et al., 2004). In addition, the increased complexity of the urban land use planning problem follows not only from the involvement of multiple objectives but also from the definition of involved objectives. These objectives may be unstructured and nonlinear and thus difficult to handle. Within this context, computer-based techniques have been developed to assist the planners with decision making.

The linear programming (LP) model was first articulated in the 1960s to solve the linear or quadratic equations that address problems in urban planning systems (Guldmann, 1979; Aerts, 2003). However, because the LP model cannot handle nonlinear and unstructured requirements like spatial interactions between land use types, it is not suitable for complex urban problems. In addition, because multiple objectives are involved, e.g., providing enough housing and
employment for an increasing population, reducing traffic congestion, preserving the
environment, and so on, and because LP is efficient only when a single objective is clearly
identified (Stewart et al., 2004), the model is not suitable. Within this context, a heuristic
algorithm, the genetic algorithm (GA), which is capable of handling the unstructured urban
issues, was proposed in this field in 1970s (Hopkins, 1977; Los, 1978). The GA is a type of
general global optimization algorithm, and it has been shown to be robust and efficient for
searching large, complex, and little-understood search spaces such as those of multi-objective
land use planning problems (Zhang et al., 2010). Furthermore, because the GA works with a
population of plans and a number of Pareto-optimal solutions can be generated, it becomes
possible to provide a set of alternative solutions from which planners can choose (Srinivas &
Deb, 1994), rather than offering one “best” solution. It is well suited for practical applications.
For example, if implementation of an optimal plan is difficult or impossible, another land use
plan can be selected from the pool of Pareto-optimal solutions (Srivastava et al., 2002). In fact,
many researchers have applied a GA to solve multi-objective land use planning problems, and
some meaningful outcomes have been achieved (Balling et al., 2004; Cao et al., 2012;
Chandramouli et al., 2009; Ligmann-Zielinska et al., 2008; Matthews, 2001; Matthews et al.,
2006; Stewart et al., 2004). For example, Stewart et al. (2004) used a GA to incorporate multiple
objectives in land use planning and found that the heuristic algorithm tended to invalidate the use
of traditional LP models as an optimization approach for land use planning systems. Brookes
(2001) used a GA combined with a region-growing program that generated alternative patch
configurations, and discovered that a GA can effectively resolve multiple patch problems and is
better than a LP method.

Problems remain in current studies of land use optimization. First, existing research has
separated urban growth from land use planning. Scholars have focused on the objectives and on
optimization techniques. For example, numerous techniques such as GA (Balling et al., 2004;
Cao et al., 2011), GIS (geographical information system) (Batty et al., 1999; Wang et al., 2004),
and DSS (computer-based decision support system) techniques (Matthews et al., 2006) have
been proposed as approaches to land use optimization. However, attention to the matter of urban
extent for which optimization could be conducted is lacking. In recent studies, urban extent was
given or fixed subjectively in advance, and such determination is not appropriate for urban land
use planning. For example, Balling et al. (2004) and Cao et al. (2012) identified urban extent on
the basis of the administrative boundary in which sustainable land use planning was conducted.
Chang et al. (1995) searched for optimal land use programs in the Tweng-Wen reservoir
watershed on the basis of a natural boundary. Because the urban extent determines the difficulty
of achieving some objectives, ignorance of the extent will result in errors in optimization. For
instance, the extent determines the total urban land area and thus the objectives associated with
the land use area, such as housing capacity and employment opportunities provided by the
available residential and commercial lands, respectively (Balling et al., 1999), are affected by the
urban extent. Thus, it is important to account for the extent of the urban area in the projected year.
In the study described herein, a widely used urban growth model, the cellular automata (CA) model (Al-kheder et al., 2008; Barredo et al., 2003; Batty et al., 1999; Feng et al., 2011), is employed to identify the urban extent in the projected year before land use optimization. The CA approach is a bottom-up approach modeling a complex system by a set of simple rules (He, 2006). It is used by professionals for urban growth simulation and for predicting the extent of an urban area (Clarke & Gaydos, 1998). Given the advantages of CA in spatial and temporal simulations of urban land use change, a lot of work has been done by using CA to simulate the urban growth process. For example, Rabbani et al. (2012) used CA, together with particle swarm optimization to calculate the urbanization probability of cells in simulating the urban growth of the city of Tehran, Iran, from 1988 to 2010. The results showed the lower scale sensitivity of the proposed model. Feng et al. (2011) applied CA in the Fengxian District of Shanghai Municipality, eastern China, to simulate the spatio-temporal process of urban growth from 1992 to 2008, with the results demonstrating that the model outperformed other spatial statistical models. Han et al. (2009) applied an integrated system dynamics and CA model to an analysis of the socio-economic driving forces and evaluation of urban spatial patterns to simulate the urban growth of Shanghai. Mahiny & Gholamalifard (2007) used CA to model and forecast the likely change in the extent of Gorgan, and illustrated the utility of modeling in explaining the spatial pattern of urban growth. Despite their superiority in simulating dynamic systems, conventional CA models have problems defining the parameter values of transition rules (Batty et al., 1999). Thus, scholars have made efforts to refine transition rules. For example, Mantelas et al. (2010) used a fuzzy system to assist in determining transition rules and Guan et al. (2011) defined transition rules on the basis of a Markov model. Among the various approaches, the neural networks (NNs) approach seems to be most suitable for the simulation of complex relationships (Pijanowski et al., 2002), of which urban areas stand as major examples (Almeida et al., 2008). It is generally accepted that the NNs approach can achieve greater accuracy in modeling (Wang, 1994). Hence, in this study, the NN approach was integrated with a CA model to determine the transition rules.

In addition to inadequate consideration of the extent of the urban area in the projected year, the global spatial trend in land use distribution is ignored. The global spatial trend can be used to represent land use in the urban system on a large scale. Local spatial trends, such as compatibility and contiguity, have been widely proposed as objectives (Aerts et al., 2003; Cova & Church, 2000; Ligmann-Zielinska et al., 2008; Shirabe, 2005). These factors are measured as the consistency between a plot and its neighborhood (Wu, 2005). Focus on the global spatial trend in land use optimization is somewhat lacking. To date, only limited attention has been paid to global spatial trends in land use planning optimization systems. For example, Ligmann-Zielinska et al. (2008) tried to minimize the distance between land set aside for new development and land already developed so as to ensure easy transition between the two. However, to reflect use, Ligmann-Zielinska et al. (2008) classified plots simply as developed or undeveloped, despite the fact that the urban system involves various land usages, and it is not possible to put their research results into practical application. Thus, in our study, in addition to objectives that
reflect the local spatial distribution trend, the global spatial distribution trends of commercial and
industrial land are proposed according to their distances from the city center and the density of
the road lines.

To validate our method by which an urban growth model assists in determining the urban
extent for optimization and the global spatial trend is added, a downtown lake area, the Donghu
Lake watershed area, which is undergoing rapid urban sprawl, was selected as a case study area.
The proposed method is used to search for the optimal land use plan for Donghu watershed in the
projected year 2020. The area is undergoing rapid urbanization according to the land cover
change data (Li et al., 2006). Within this context, persistent expansion of the urban area is clearly
expected for Donghu watershed. Therefore, it is important to identify the urban extent in the
projected year before carrying out the optimization. As a downtown lake, Donghu Lake is also
greatly affected by rapid development of the downtown area. Despite the furtherance of
downtown economic development, some adverse effects such as non-point source (NPS)
pollution of the water body are brought about and have been measured by some scholars (Xu et
al., 2000; Zhou et al., 2002; Zhu et al., 1993). NPS pollution in particular has become the major
cause of water quality deterioration (Leone et al., 2009; Ouyang et al., 2010). Hence, in addition
to economic and spatial distribution objectives, minimizing NPS is proposed as an objective for
the Donghu watershed area.

The remainder of this paper is organized as follows. In the methods section, an urban growth
model, specifically an NN-improved CA, is proposed to identify the urban extent in the projected
year. A multi-objective optimization approach based on a GA is then proposed with the
assistance of the global spatial trend. In the case study section, an actual case study is described:
(1) the land use characteristic of Donghu Lake is introduced; (2) the extent of the urban area in
the projected year is forecasted by a CA model; (3) the maximization of housing capacity, the
maximization of employment capacity, the minimization of NPS pollution, and the maximization
of compatibility between land uses are set as objectives in association with the global spatial
trend; (4) the optimized land use plan is represented, and the results are analyzed.

Methods

Urban growth model: CA

CA modeling has become a preferred technique (Feng et al., 2011) for urban sprawl
modeling because it can simulate the complex dynamic (spatial and temporal) process through
relatively simple rules (Guan et al., 2011). The process of urban growth modeling via CA can be
presented as equation (1) shows (Almeida et al., 2008; Batty et al., 1999):

\[ S_{ij}^{t+1} = f(S_{ij}^t, \Omega_{ij}, Con) \]  

where \( S_{ij}^{t+1} \) means the land use state of one cell (i, j) at step t+1 with 0 indicating a non-urban
grid and 1 indicating an urban grid, and it is determined according to the state at step t (named
In practice, the state of cell \((i,j)\) at time \(t + 1\) can be determined by equation (2) where, if the conversion probability is greater than a random constant (ranging from 0 to 1), the land use cell \((i,j)\) will convert to urban land use; otherwise, the land use cell will not convert. The conversion probability of a certain land use cell \((i,j)\) can be determined using equation (3):

\[
S_{ij}^{t+1} = \begin{cases} 
\text{convert to urban} & \text{if } P_{ij}^{t+1} > P_{\text{random}} \\
\text{unchanged} & \text{if not}
\end{cases}
\]

\(P_{ij}^{t+1} = (P_\Omega)_{ij} \times C_{nk} \times P_r \)  

where \((P_\Omega)_{ij}\) is the state conversion probability of the cell, which depends on the number of urban cells in its neighborhood, and it equals the ratio of urban cells in its neighborhood to the total number of neighbors; \(C_{nk}\) is the prior change of the \(k\)-th land usage; and \(P_r\) represents the stochastic disturbance of any unknown errors.

\(C_{nk}\) in equation (3) depends on the area that the \(k\)-th land usage should be in the projected year, and the actual area of \(k\)-th land usage at current step \(t\) as represented by equation (4)

\[
C_{nk} = \frac{P_k - C_k}{P_u - C_u}
\]

where \(P_k\) is the predicted area of \(k\)-th land use in the projected year, \(C_k\) is the area at current step \(t\) of \(k\)-th land use, and \(P_u\) and \(C_u\) are the predicted area and actual area of urban land use at step \(t\), respectively.

In particular, NNs were employed to predict \(P_k\) under the CA transition rules. An NN can be defined as a mathematical model or calculation model that imitates the structure and function of a biological neural network. It is a non-parametric approach (Almeida et al., 2008). An NN consists of three layers: the input, hidden, and output layers. These can identify non-linear relationships in nature (Pijanowski et al., 2002). The basic unit in an NN is the neuron or node, which is organized in a couple of layers. All neurons except those in the input layer execute two functions, receiving a signal (activation) from a previous layer and transmitting a new signal (activation) to the next layer (Almeida et al., 2008). The function for addressing the signals between neurons is given as

\[
\text{net}_n = \sum_m W_{m,n} I_m
\]

where \(I_m\) is the generated signal from neuron \(i\) in the previous layer, \(\text{net}_i\) is the collected signal for receiver neuron \(n\) in the next layer, and \(W_{m,n}\) is the weight of the sum of the activations from different input nodes. The receiver neuron creates activation, which will become the input for the next layer, in response to the signal \(\text{net}_n\). The activation is usually generated in
the form of a sigmoid function. According to equation (3) and the sigmoid function, the activation will be processed again and again and passed to the next layer as input signal. With regard to $W_{m,n}$, the values of the weight are not set by the users but rather are determined by the network during training (Foody, 1996; Rumelhart et al., 2002). Once the optimized weights have been retrieved from the training data set, the network is ready for classification or prediction. In this study, the NN algorithm was carried out by MATLAB 7.0, where the quantities of historical land uses and the corresponding year act as the input layer, and the areas of all land uses in the projected year function as the output layer which will act as $P_k$ in equation (2).

Optimization model: GA

The GA process

A GA was selected as the multi-objective optimization technique to search for optimal urban land use allocation. Three hundred sixty-four plots in the study area separated by road lines are the smallest units of the GA and referred to as genes. A straightforward chromosome representation can be a list of genes in which the land use type of the gene is determined by a value (Cao et al., 2012). In particular, each gene is set as an integer, ranging from 1 to the number of possible land uses; and the 364 genes constitute one chromosome, which, in this study, is also referred to as a land use plan. One generation consists of numerous chromosomes (plans) that will undergo selection, crossover, mutation, and elitism.

In practice, the GA first generated one plan with 364 random values. A generation size of 100 is selected, indicating that generation will not stop until 100 plans exist. The plans in the first generation satisfying constraints are then ordered by their fitness and termed feasible plans. In particular, fitness can be denoted as the goodness of the plan that satisfies objectives. The plans with greater fitness are better than the plans with relatively low fitness. The second generation is created from the previous generation according to the processes of selection, crossover, mutation, and elitism. Two plans with high fitness are selected, as father and mother, from the generated feasible plans. This is the selection process. Parent genes are then exchanged via crossover to generate children, i.e. new next-generation plans. Because the children inherit the goodness of both father and mother, the next generation is generally expected to be better than the previous generation. In addition, to avoid local optima, the process of mutation is carried out with a 0.05 probability of mutation for each children plan. This means that for each gene in each children plan, there is a probability of 0.05 for random change to any other land use type. All of these processes are executed until 100 plans are generated for the next generation. The algorithm will not terminate until 1000 iterations have been carried out. The execution of the model on 365 plots with four objectives and fitness computation requires approximately 2.9 hours for 1000 iterations for the population of 100 on a computer with Iterl(R) Intel Core 2 Duo CPU E8400 at 3.00 Ghz with 3.48 GB RAM. The entire GA process is diagrammed in Figure 1.
Fitness Evaluation

Fitness is calculated according to the degree to which a plan dominates the other plans of its generation. Therefore, the greater the fitness value, the better the plan (Lowry & Balling, 2009). There are numerous methods for computing fitness. In this study, the maximin fitness function, proposed by Balling (2002), was used, as expressed by equation (3).

\[ \text{Fitness}_i = \max_j (\min(\text{Obj}_{1,j} - \text{Obj}_{1,i}, \ldots, \text{Obj}_{k,j} - \text{Obj}_{k,i})) \]  

The objectives need to be normalized before fitness evaluation. In this study, a simple method was employed, expressed by equation (4).

\[ \text{NormObj}_{k,i} = \frac{\text{Obj}_{k,i} - \text{Obj}_{\text{min},k,i}}{\text{Obj}_{\text{max},k,i} - \text{Obj}_{\text{min},k,i}} \]

where \( \text{obj}_{\text{max},k} \) and \( \text{obj}_{\text{min},k} \) are the maximum and minimum values of the k-th objective (\( \text{obj}_k \)) in one generation and \( \text{NormObj}_{k,i} \) is the normalized value of the k-th objective for the \( i \)-th plan in this generation.

Each plan must be compared with all other plans within the same generation. Then \( \text{plan}_i \) in a particular generation is deemed to be dominated by another \( \text{plan}_i \) within the same generation if

\[ \max_j (\min(\text{Obj}_{1,j} - \text{Obj}_{1,i}, \ldots, \text{Obj}_{k,j} - \text{Obj}_{k,i})) > 0. \]

Case study

Study area and land cover change

Donghu watershed (114° 18′ -114° 30′ E, 30° 30′ -30° 38′ N; area, 18075 ha) is located in the eastern portion of Wuhan in central China (Figure 2). Donghu Lake is one of the largest downtown lakes in China. In addition to its general function in regulating climate, degrading pollutants, providing living space for aquatic life, and preventing flooding, Donghu Lake is part of an urban area and plays an important role in urban development. The effect of intense anthropogenic activities and urbanization on the watershed water quality is profound.

Figure 3 shows the land use patterns of our site in 1980, 1990, 2000, and 2005 and the quantitative land use changes from 1980 to 2005 are listed in Table 1. The land use data were retrieved from a Landsat Thematic Mapper (TM) image acquired by remote sensing, downloaded
from the Internet, and classified by means of an ERDAS software package. Additionally, land
use data from 1980 and 2005 were obtained through the Earth System Scientific Data Sharing
Network, which can be used to confirm accuracy of the classification results.

In 1980, the watershed was mainly rural, with urban land accounting for 45.31% of the whole
land area. By 2000, against the background of rapid economic development in Wuhan, the
western basin exhibited features of a city, as the western development rate was significantly
higher than that of the eastern area. With urban expansion, the built-up land increased to 71.68
km\(^2\) (53.61% of the land area) by 2005. Generally, it has been concluded that Donghu watershed
has been undergoing rapid urbanization over the last 20 years and that this process will continue
into the future.

Objectives

The urban extent predicted by the CA model is divided into 364 plots by road lines. The
optimal plan will incorporate the following 11 types of land use into each plot: low density
residential, medium density residential, high density residential, central business district,
commercial center, commercial, light industrial, heavy industrial, mixed residential-commercial,
green space, and public service uses. Possible land use types and their abbreviations are listed in
Table 2. In addition, it is wise to identify the plots that will remain as they are by asking planners
and experts before the optimization. For example, land used by the university in the watershed
would not be allocated to any other use. Furthermore, in so doing, the search space and
programming time can be reduced (Balling et al., 2004).

Four objectives are formulated to achieve a sustainable urban land use plan for the watershed.
The proposed social and economic development objectives are described in objectives 1 and 2:
the maximization of housing capacity and the maximization of employment capacity for the
whole watershed, respectively. Specific goals for the watershed are proposed as objective 3 and 4.
The first is the minimization of NPS pollution in the urban area, with an export coefficient model
(ECM) being used to evaluate NPS pollution for a candidate plan. Second, the local spatial
objective is carried out as represented by the minimization of incompatibility.

The first objective

The first objective is to maximize the housing capacity (\(NumRP\)) of the watershed area. The
residential area can be categorized by housing density into low density residential, medium
density residential and high density residential zones, as listed in Table 2. The total housing capacity of candidate plan \( i \) can be computed by multiplying the residential land area by unit housing capacity in persons per hectare per land use type (Balling et al., 2004; Chandramouli et al., 2009). Unit housing capacities of individual land use types are given in Table 3 according to existing studies and characteristics of our watershed (Balling et al., 2004). Therefore, the first objective is represented by the following equation:

\[
\text{NumRP}_i = Ar \times LDR_i \times 50 + \text{AreaMDR}_i \times 80 + \text{AreaHDR}_i \times 100 + \text{AreaRC}_i \times 50
\]  

(6)

where \( \text{AreaLDR}_i \), \( \text{AreaMDR}_i \) and \( \text{AreaHDR}_i \) refer to the total area of LDR plots, MDR plots and HDR plots of candidate plan \( i \), respectively, and \( \text{NumRP}_i \) is the housing capacity of plan \( i \).

The second objective

The second objective is to maximize the employment capacity for residents (\( \text{NumEmploy} \)). Unit employment capacities in persons per hectare vary per land use type (Chandramouli et al., 2009) and are listed in Table 3. The second objective for Donghu watershed is represented by the following equation:

\[
\text{NumEmploy}_i = \text{AreaLI}_i \times 25 + \text{AreaHI}_i \times 25 + \text{AreaCDB}_i \times 150 + \text{AreaCC}_i \times 50 + \text{AreaC}_i \times 30 + \text{AreaRC}_i \times 30
\]  

(7)

where \( \text{AreaLI}_i \), \( \text{AreaHI}_i \), \( \text{AreaCDB}_i \), \( \text{AreaCC}_i \), \( \text{AreaC}_i \), and \( \text{AreaRC}_i \) represent the total LI, HI, CDB, CC, C, and RC areas, respectively, and \( \text{NumEmploy}_i \) is the employment capacity of plan \( i \).

The third objective

The third objective is to minimize NPS pollution. To meet this objective, an empirical codon model (ECM) is employed. The ECM is considered reliable for simulating NPS pollution (Worrall & Burt, 1999). The ECM approach was originally developed in North America to predict nutrient inputs into lakes and streams (Beaulac & Reckhow, 1982; Dillon & Kirchner, 1975), and the ECM process is represented by the following equation:

\[
L = \sum_{k=1}^{n} A_k E_k
\]  

(8)

where \( L \) is the annual pollutant load (kg), \( E_k \) is the export coefficient from land use \( i \); \( A_k \) is the catchment area covered by land use \( i \) (ha), and \( n \) is the number of land uses (Khadam & Kaluarachchi, 2006). The export coefficient (\( E_k \)) expresses the rate at which nitrogen or
phosphorus is exported from each land use type in the watershed. It is conventionally derived
from the literature and by field experiments (Johnes, 1996). In particular, a literature review is
conducted to determine the export coefficient of Donghu watershed (Cao et al., 2007; Kim et al.,
1993; Liu et al. 2006; Shields et al. 2008; Winter & Duthie, 2000). To simplify the model,
normalized export coefficient values without units are used that range from 0 to 100, with high
values suggesting high NPS pollution export coefficients (see Table 3). The ECM makes it
possible to estimate total annual loads of phosphorus and nitrogen for a water body from NPSs
(Winter & Duthie, 2000). However, the typical ECM does not account for the uneven spatial
distributions of precipitation and slope, which are believed to be the primary factors that affect
NPS pollution. Rainfall is the main driving force of NPS contamination, and terrain plays an
important role in NPS pollutant transport. Therefore, it is important to modify the ECM
according to rainfall and topography. Ding (2010) improved the ECM by introducing
precipitation and terrain impact factors, as represented by equation (9). In this study, the factors
of precipitation and terrain were identified by equations (10) and (11).

\[
L = \sum_{i \in U} Area_i E_i \alpha_i \beta_i \quad (9)
\]

where \(Area_i\) is the area of the \(l\)-th plot, \(E_i\) is the export coefficient of the \(l\)-th plot, \(\alpha_i\), and \(\beta_i\) are
the precipitation factor and terrain factor, respectively, at the plot \(l\) location. \(U\) denotes all plots
in the optimization.

\[
\alpha_i = \frac{Pre_i}{AverPre_{year}} \quad (10)
\]

\[
\beta_i = \frac{Slope_i}{AverSlope} \quad (11)
\]

where \(Pre_i\) is the precipitation of \(l\)-th plot; \(AverPre_{year}\) is the annual precipitation of the study
area; \(Slope_i\) is the slope at the plot \(l\) location; and \(AverSlope\) is the average slope of the study
area.

The slope is calculated by digital elevation model (DEM) data obtained from the TM remote
sensing image. The study area is generally flat (Figure 4), with the exception of some steeply
sloped plots in the south. The precipitation of the study area is extracted through interpolation of
national rainfall station data (Figure 5), and the result shows that rainfall decreases from the
south-east to the north-west.

[Figure 4 Slope of Donghu watershed]

[Figure 5 Precipitation of Donghu watershed]
The fourth objective

The fourth objective is compatibility between land uses. In reality, a given land use, such as industrial use, tends to cluster into zones that can be described as compact or contiguous. These zones minimize conflict between neighboring land uses (Ligmann-Zielinska et al., 2008). In actuality, there are land use-related preferences that are reflected in the choice of neighborhoods. Therefore, we can define the degree of conflict between any two land use types as ranging from 0 to 100, with high values indicating high conflict. The compatibility of each plan can be calculated by adding the degrees of conflict between neighboring plots. The lower the sum, the more compatible the scenario will be. In the process of determining the degree of conflict, the opinions of experts, decision makers, and stakeholders involved in land use allocation must be taken into account. The final conflicting degrees are presented in Table 4. A plan that is not compatible can be formulated as follows:

\[
InCom = \sum_{i \in U} \sum_{j \in \Omega_i} ConflictDegree(plot_i, plot_j)
\]  

where \( InCom \) is incompatibility of plan \( i \), \( plot_i \) is the \( l \)-th plot; \( \Omega_i \) is the neighborhood of plot \( l \).

The above-defined objectives do not account for the global spatial trend of land usage in the urban system, even though the fourth objective reflects the local land-use related preferences. Because Donghu Lake watershed is mixed urban-rural area and the characteristic varies from place to place, the large scale spatial trend is essential to land use allocation. For example, a land use plan that puts a CBD in an area without a convenient traffic system is unreasonable even if it can provide as much employment as a land use plan that locates the CBD in the city center. Furthermore, HI and LI should be located far away from the city center to reduce disturbance. To reflect the global spatial distribution trend, the distance to the city center and the density of road lines are determined. The global spatial distribution tends to be associated with the previous objectives rather than proposed as a new objective. If proposed as a new objective, several objectives are needed rather than one, and this clearly will increase the cost of computation. In addition, errors may result when there is no HI or LI and the objective value of the distance of the HI and LI to the urban center is minimized. For the reasons stated above, the global spatial trend is added by refining existing objectives rather than proposing new objectives.

Because the first objective does not deal with LI, HI, or CBD, it is not necessary to refine it. The other three objectives are refined as equations (13), (14), and (15), respectively.
\[\text{NumEmploy} = 50 \sum_{p \in LI} \text{Area}_p \text{DisC}_p + 40 \sum_{p \in HI} \text{Area}_p \text{DisC}_p + 250 \sum_{p \in CBD} \text{Area}_p \text{DenR}_p + \text{AreaCC}_i \times 100 + \text{AreaC}_i \times 50 + \text{AreaRC}_i \times 30\]  

(13)

\[L = A \sum_{i \in HI \cup LI} E_i \alpha_i \beta_i \text{DenR}_i + A \sum_{i \in CBD} E_i \alpha_i \beta_i \text{DisC}_i + A \sum_{i \in others} E_i \alpha_i \beta_i\]  

(14)

\[\text{InCom} = \sum_{i \in HI \cup LI} \sum_{j \in \Omega_i} \text{ConflictDegree}(plot_i, plot_j)(1 - \text{DisC}_i)\]

\[+ \sum_{i \in CBD} \sum_{j \in \Omega_i} \text{ConflictDegree}(plot_i, plot_j)(1 - \text{DenR}_i)\]  

(15)

\[+ \sum_{i \in others} \sum_{j \in \Omega_i} \text{ConflictDegree}(plot_i, plot_j)\]

where \(\text{DenR}_i\) is the road density of the \(l\)-th plot and \(\text{DisC}_i\) is the distance between the \(l\)-th plot and the city center. Both \(\text{DenR}\) and \(\text{DisC}\) are normalized (ranging from 0 to 1) to make formulation simple. Distances to the center and road density are shown in Figure 6 as achieved by ArcGIS. By the equations (13), (14), and (15), the global spatial trend is attached to each objective that is associated with CBD, HI, and LI, respectively. Then, with the optimization process, a CBD with high road density and HI and LI with appropriate distance from the city center are expected.

With the maximin fitness as defined above, a better plan with higher fitness is generated (Balling, 2003). The maximin fitness is available to uncover minimum objectives, whereas some objectives such as housing capacity and employment capacity, \(\text{NumRP}\) and \(\text{NumEmploy}\), have tried to maximize the goals. To simplify the problem, in the process of optimization, \(\text{VNumRP} = \max \text{NumRP} - \text{NumRP}\) and \(\text{VNumEmploy} = \max \text{NumEmploy} - \text{NumEmploy}\), which are the inverse of \(\text{NumRP}\) and \(\text{NumEmploy}\), are defined. In particular, \(\max \text{NumRP}\) is the housing capacity when all the urban plots are assigned to HDR, and \(\max \text{NumEmploy}\) is the employment capacity when all the urban plots are assigned to CDB.

**Results and discussion**
To validate the benefit of the global spatial trend, optimization without a global spatial objective was conducted for comparison. Two optimized plans selected randomly from a set of Pareto-optimal plans are shown, one with (Figure 7) and one without (Figure 8) a global spatial objective. In the conventional model, some HI and LI zones are allocated close to the center city (see Figure 7, plot A) where residential and commercial lands are concentrated and the detrimental effects are profound; meanwhile, some CBDs are allocated to the east (see Figure 7, plot B), too far from the center city and where the traffic system is inconvenient. Clearly, even if the optimized plan can provide as much employment capacity as is desired, the global spatial trends of HI, LI, and CBD are ignored. With the refined objectives, the global spatial trend is added, and the model provides a solution, with HI and LI zones far away from the city center and the CBD near to convenient transportation system (see Figure 8). In particular, according to Figure 8 only a few commercial land (C) plots are denoted in the southeast, and the CBD is assigned to the city center, the northwest of our study area.

In addition to reflecting the global spatial tendency of some land usages, the proposed optimization model achieved a trade-off land map where urban extent is 7805.88 ha and is covered by LDR, MDR, HDR, CBD, CC, C, LI, HI, RC, GS, and PS with an appropriate proportion for each land use. According to the map, the urban area in Donghu watershed should be assigned mainly as residential land accompanied by some commercial land. In terms of industry, only LI is allocated in our study area. The land use structure satisfies the environmental requirements of a watershed and fulfills the economic objective because some commercial land is assigned and provides much employment capacity. In terms of spatial distribution, LDR, MDR, and HDR are clustered and surrounded by SP and GP. Similarly, CBDs are concentrated and close to residential lands, and CC and C are scattered among residential lands, conferring an environmental benefit to the whole watershed and convenience for residents. Generally, the optimal spatial allocation is available.

In addition to the spatial distribution, analysis of variation in average objective values for one generation is carried out. Figure 9 represents the optimizing process. As the plot shows, there are two breaks in optimization processing, one around the 700th step and the other around the 850th step. Clearly, at the 700th step, employment capacity increased at the cost of growth of NPS pollution. Around the 820th step, the incompatibility decreased but the employment capacity decreased simultaneously. We can say that compromise between these four objectives occurs in the process of optimization. Finally, a relatively balanced condition emerges in which all objectives are satisfied to the greatest extent.
The Pareto-optimal plans searched out the trade-offs for all objectives; thus, the best performances aimed at single objectives are presented and analyzed.

Concerning housing capacity, when the urban area is completely covered by HDR the maximum value of $NumRP$ can be achieved. Within this context, the value of $NumRP$ equals the product of the area of urban place and unit housing capacity in people per hectare of HDR which is $7805.88 \times 100 \text{ per/ha.} = 780588 \text{ person}$. In contrast, maximum NumRP plan affords no employment and a relatively high NPS pollution and the InCom value will be zero because the whole area is covered by a single land use (HDR).

With regard to employment capacity, the urban area completely covered by CBD provides the largest employment capacity. Similarly, the maximal $NumEploy$ value equals to the product of area of urban site and unit employment capacity of CBD in people per hectare as $7805.88 \times 150 \text{ per/ha.} = 1170882 \text{ person}$. In this context, maximal $NumEploy$ plan will yield no housing capacity and high NPS pollution; similar to maximal NumRP plan, the InCom value of maximal NumEploy is zero because the whole area is covered by a single land use (CBD).

Regarding NPS pollution, according to the minimal $L$ plan, the whole urban area is covered by GS, and the NPS pollution will be minimal. Under this situation, the minimal $L$ plan does not afford any housing or employment capacity at all; similarly, the InCom value will be zero because the whole area is covered by a single land use (GS).

As for incompatibility, the minimal $InCom$ plan will emerge when the whole urban area undergo just one certain kind of land usage. When the one certain land usage is HDR, the minimal $InCom$ plan is also maximal $NumRP$ plan and the minimal $InCom$ plan is also capable of being maximal $NumEploy$ plan if the one certain land usage is CBD. Because the state in the minimal $InCom$ plan will differ with respect to the particular land use, there are no assured values for the other three objectives (see minimal $InCom$ plan in Table 5).

Moreover, objective values of the optimized plan that is represented by Figure 8 are listed in Table 5. Corresponding to its spatial distribution, the plan yields a large housing capacity (353510 persons) and the employment capacity is relatively low (106040 persons). Even if the NPS pollution is 312920, which is much larger than 39029 in the minimal $L$ plan, the pollution load is much lower than that of the maximal $NumRP$ plan or maximal $NumEploy$ plan. Clearly, regarding the single objective, the optimized land use plan is not as good as the maximal $NumRP$ plan, maximal $NumEploy$ plan, minimal $L$ plan, or minimal $InCom$ plan. On the whole, however, the optimized plan accounted for all conflicting objectives.

**Conclusion**
In this study, using a GA we established a multi-objective optimization model to successfully search for Pareto-optimal land use plans for Donghu watershed in Wuhan, central China. To determine the urban extent for the case study area in the projected year, the CA model is proposed to simulate the urban growth process. In addition, an NN approach is introduced to assist in the generation of transition rules in urban growth according to the historical land use data pattern of Donghu watershed from 1980 to 2005. Following the determination of the extent, four objectives are proposed as multiple objectives for the urban planning: housing capacity, employment capacity, reduced NPS water pollution, and compatibility between land uses. However, even if the objective of compatibility reflects the local spatial distribution of land use preferences, the global spatial distribution trend is ignored. Hence, with the measurement of the distance from the city center and the density of road lines, we introduce the trend in global spatial distribution of commercial land and industrial land to the process of optimization.

It is concluded that CA and GA are successfully used to search for an optimized urban land use plan for Donghu watershed. First, all conflicting objectives are incorporated into the process of optimization. Even if the optimal plans do not yield the best solution for single objectives, the optimal plans on the whole provide solutions that satisfy all objectives to the maximum extent. Second, by consideration of the global spatial trend, the spatial distribution of optimal plans is more reasonable than that of the plans that do not consider global distribution. According to the results, our plans allocate HI and LI far from the city center and the CBD close to a convenient transportation system.

What makes our approach innovative is the determination of urban extent before land use optimization. According to the historical land use data, our studied watershed area is undergoing rapid urbanization, and this expansion is expected to continue well into the future. Unlike other studies where the urban extent is fixed or unchanged, ignorance of the projected extent of the urban area will yield errors in optimization. Because the urbanization process is occurring worldwide (Geyer, 1996; Tilly, 1967), it is meaningful to apply our proposed method in other areas.

Despite the remarkable outcomes, there are still limitations to this study. For example, the GA is stochastic (Matthews et al., 2006), and the results are not exactly the same from application to application. In addition, like other artificial intelligence techniques, the GA cannot assure a constant optimization response time. Meanwhile, as the land use map is provided for planners’ information, a three-dimensional representation would be more informative and more easily comprehended than a two-dimensional representation (Chandramouli et al., 2009). These problems will be addressed in our future work.

References


Balling, R.J., (2002). The maximin fitness function for multiobjective evolutionary optimization. OPTIMAIZATION IN INDUSTRY, 135-147


### Table 1 Areas and percentages of land uses, 1980-2005

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Area (km²)</td>
<td>Percent (%)</td>
<td>Area (km²)</td>
<td>Percent (%)</td>
</tr>
<tr>
<td>Forest</td>
<td>15.18</td>
<td>8.40</td>
<td>14.52</td>
<td>8.03</td>
</tr>
<tr>
<td>Grass</td>
<td>0.93</td>
<td>0.51</td>
<td>1.09</td>
<td>0.60</td>
</tr>
<tr>
<td>Agricultural</td>
<td>53.03</td>
<td>29.34</td>
<td>52.05</td>
<td>28.80</td>
</tr>
<tr>
<td>Urban</td>
<td>58.56</td>
<td>32.40</td>
<td>63.59</td>
<td>35.18</td>
</tr>
<tr>
<td>Rural settlement</td>
<td>0.92</td>
<td>0.51</td>
<td>1.07</td>
<td>0.59</td>
</tr>
<tr>
<td>Water</td>
<td>51.50</td>
<td>28.49</td>
<td>47.78</td>
<td>26.43</td>
</tr>
<tr>
<td>Unused</td>
<td>0.63</td>
<td>0.35</td>
<td>0.66</td>
<td>0.36</td>
</tr>
<tr>
<td>Total</td>
<td>180.75</td>
<td>100.00</td>
<td>180.75</td>
<td>100.00</td>
</tr>
<tr>
<td>Index</td>
<td>Abbrev.</td>
<td>Land use</td>
<td>Index</td>
<td>Abbrev.</td>
</tr>
<tr>
<td>-------</td>
<td>---------</td>
<td>---------------------------</td>
<td>-------</td>
<td>---------</td>
</tr>
<tr>
<td>1</td>
<td>LDR</td>
<td>Low density residential</td>
<td>7</td>
<td>LI</td>
</tr>
<tr>
<td>2</td>
<td>MDR</td>
<td>Medium density residential</td>
<td>8</td>
<td>HI</td>
</tr>
<tr>
<td>3</td>
<td>HDR</td>
<td>High density residential</td>
<td>9</td>
<td>RC</td>
</tr>
<tr>
<td>4</td>
<td>CBD</td>
<td>Central business district</td>
<td>10</td>
<td>GS</td>
</tr>
<tr>
<td>5</td>
<td>CC</td>
<td>Commercial center</td>
<td>11</td>
<td>PS</td>
</tr>
<tr>
<td>6</td>
<td>C</td>
<td>Commercial</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 3 Coefficients for multi-objective optimization

<table>
<thead>
<tr>
<th>Index</th>
<th>Land usage</th>
<th>Export coefficient</th>
<th>Unit housing capacity (persons per hectare)</th>
<th>Unit employment capacity (person per hectare)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>LDR</td>
<td>35</td>
<td>50</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>MDR</td>
<td>45</td>
<td>80</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>HDR</td>
<td>60</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>CBD</td>
<td>50</td>
<td>0</td>
<td>150</td>
</tr>
<tr>
<td>5</td>
<td>CC</td>
<td>30</td>
<td>0</td>
<td>50</td>
</tr>
<tr>
<td>6</td>
<td>C</td>
<td>20</td>
<td>0</td>
<td>30</td>
</tr>
<tr>
<td>7</td>
<td>LI</td>
<td>70</td>
<td>0</td>
<td>25</td>
</tr>
<tr>
<td>8</td>
<td>HI</td>
<td>80</td>
<td>0</td>
<td>25</td>
</tr>
<tr>
<td>9</td>
<td>RC</td>
<td>45</td>
<td>50</td>
<td>30</td>
</tr>
<tr>
<td>10</td>
<td>GS</td>
<td>5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>11</td>
<td>PS</td>
<td>10</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>LDR</td>
<td>MDR</td>
<td>HDR</td>
<td>CBD</td>
</tr>
<tr>
<td>-------</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>LDR</td>
<td>0</td>
<td>10</td>
<td>20</td>
<td>60</td>
</tr>
<tr>
<td>MDR</td>
<td>0</td>
<td>10</td>
<td>20</td>
<td>60</td>
</tr>
<tr>
<td>HDR</td>
<td>0</td>
<td>10</td>
<td>20</td>
<td>40</td>
</tr>
<tr>
<td>CBD</td>
<td>0</td>
<td>20</td>
<td>30</td>
<td>80</td>
</tr>
<tr>
<td>CC</td>
<td>0</td>
<td>20</td>
<td>70</td>
<td>80</td>
</tr>
<tr>
<td>C</td>
<td>0</td>
<td>65</td>
<td>75</td>
<td>75</td>
</tr>
<tr>
<td>LI</td>
<td>0</td>
<td>20</td>
<td>80</td>
<td>80</td>
</tr>
<tr>
<td>HI</td>
<td>0</td>
<td>90</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>RC</td>
<td>0</td>
<td>20</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>GS</td>
<td>20</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>PS</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 5 Comparison of objective values

<table>
<thead>
<tr>
<th></th>
<th>Max NumRP</th>
<th>Max NumEploy</th>
<th>Min L</th>
<th>Min InCom</th>
<th>Optimized plan</th>
</tr>
</thead>
<tbody>
<tr>
<td>NumRP (in persons)</td>
<td>780588</td>
<td>0</td>
<td>0</td>
<td>-</td>
<td>353510</td>
</tr>
<tr>
<td>NumEploy (in persons)</td>
<td>0</td>
<td>1170882</td>
<td>0</td>
<td>-</td>
<td>106040</td>
</tr>
<tr>
<td>$L$</td>
<td>468352.8</td>
<td>390294</td>
<td>39029</td>
<td>-</td>
<td>312920</td>
</tr>
<tr>
<td>InCom</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>27940</td>
</tr>
</tbody>
</table>
Figure 1

1. Start
2. Initialize the first generation randomly
3. \( \text{gen} = 1 \)
4. Fitness Evaluation
5. Terminate
   - Yes: Stop
   - No: Selection → Crossover → Mutation → Elitism
6. \( \text{gen} = \text{gen} + 1 \)
Figure 2
Click here to download high resolution image
Figure list

Figure 1 Diagram of the GA process

Figure 2 Location of the study area

Figure 3 Temporal land use patterns
   (a) 2005; (b) 2000; (c) 1990; (d) 1980

Figure 4 Slope of Donghu watershed

Figure 5 Precipitation of Donghu watershed

Figure 6 Spatial distributions of distance to the city center and road density
   (a) Distance to center   (b) Road density

Figure 7 Optimized plan without spatial tendency

Figure 8 Optimized plan with spatial tendency

Figure 9 Trends in the variation of objective values