# PREDICTING URBAN GROWTH WITH REMOTE SENSING AND DYNAMIC SPATIAL MODELLING

**Xiaojun Yang**

Department of Geography, Florida State University, Tallahassee, FL 32306, U.S.A.
xyang@fsu.edu

Commission VII, WG II/4

**KEY WORDS:** Remote Sensing, Spatial Information Sciences, Urban, Modelling, Monitoring, and Visualization

**ABSTRACT:**

This paper presents a research that integrates remote sensing, GIS, and dynamic spatial modeling for predicting urban spatial growth with different development conditions considered. The study area has been a fast growing American metropolis. The prediction is based on a cellular automate urban growth model governed by a set of complex transition rules combining both socio-economic and biophysical conditions. Historical urban extent data derived with remotely sensed imagery are used to calibrate the model. Two possible future growth scenarios are assessed. The first scenario assumes that the current development conditions do not change and therefore, can be termed as 'continuation'. The second is a hybrid growth strategy in which both conventional urban development and alternative growth efforts are addressed. It is found that many small-size urban patches would emerge and smaller ones would merge to form larger urban clusters. If current conditions do not alter, the process of urbanization would deplete vegetation and open space. A restrictive growth plan should be adopted in order to promote the livability and sustainable development in the study area. Overall, this study has demonstrated the usefulness of remote sensing, GIS, and dynamic modeling in urban and landscape planning and management. The methodology developed in this research can be easily adopted to other urban areas with similar growth patterns.

## 1. INTRODUCTION

Restless urban growth throughout the world has called for improved methods and techniques towards a better understanding of the dynamics of complex urban systems. Over the past decades, a great deal of research efforts has been directed to develop dynamic models in connection with urban and landscape applications (*e.g.* Meaille and Ward, 1990; Batty and Xie, 1994; Veldkamp and Fresco, 1996; White and Engelen, 1997; Clarke and Gaydos, 1998; Wu and Webster, 2000; and Wang and Zhang, 2001). Among the documented models, those based on cellular automata (CA) are probably the most impressive because they have grown out of an earlier game-like simulator and evolved into a promising tool for urban growth prediction and forecasting.

This research has been focused on the exploration of cellular automata-based dynamic modeling approach for applied urban studies with Atlanta as a case study area (Figure 1). For the past three decades, Atlanta has been one of the American's fastest growing metropolises as it emerged to become the premier commercial, industrial, and transportation urban center of the southeastern United States. Starting from 1996, the author has been involved in various research projects focusing on the understanding of the dynamics of change in Atlanta through the use of geographic information technologies. This paper reports part of the result of urban growth simulation carried out with a cellular automate model, a key element of the above research effort.

This research was built upon the SLEUTH Urban Growth Model (Clarke, 2000). SLEUTH derives its name from the six types of data inputs: Slope, Land cover, Exclusion, Urban extent, Transportation, and Hillshade. This is a cellular automaton urban growth model and its behavior is controlled by the coefficients of diffusion, breed, spread, slope resistance, and road gravity. The model considers four types of growth behavior: spontaneous neighborhood growth, diffusive growth and creation of new spreading centers, organic growth, and road influenced growth. Detailed description of the model can be found elsewhere (e.g. Clarke, 2000; Silva and Clarke, 2002). Using Atlanta as the study area, this project investigates the effectiveness of the SLEUTH

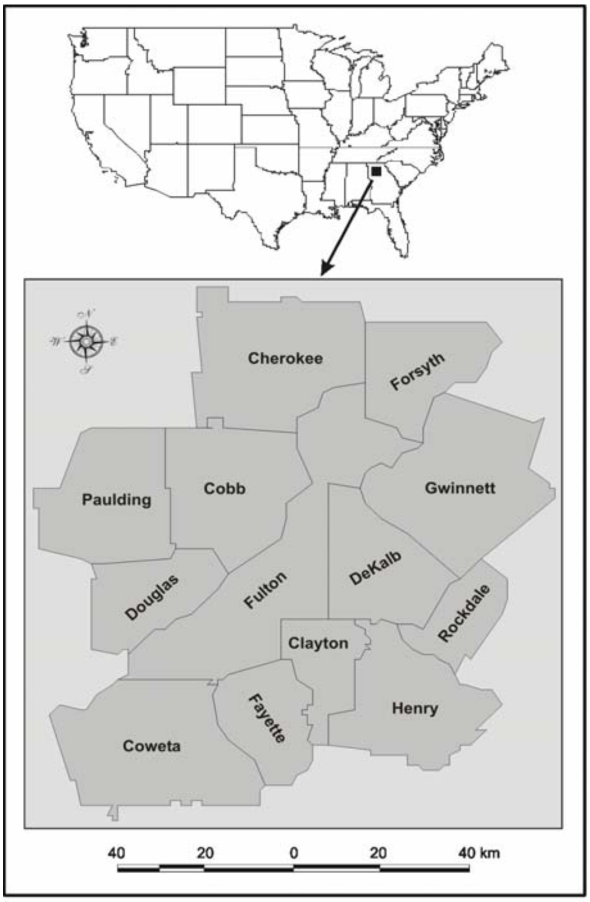


Figure 1 Location of the study area.

model as a tool to imagine, test and choose between possible future urban growth scenarios in relation to different environmental and development conditions.

## 2. RESEARCH METHODOLOGY

### 2.1 Data Assemblages

Five types of input data are needed to run the SLEUTH urban growth model. They are: urban extent, road, excluded area for development, slope, and shaded relief. Most of these input data were assembled from the databases constructed by the author through different research efforts over the past years. Each layer was resampled into three levels of spatial resolutions, namely 60 m, 120 m, and 240 m.

**2.1.1 Urban Extent:** Urban extent is actually urban built-up land, thus including all types of urban uses. Five layers of urban extent data were extracted from a time series of land use/cover maps that were produced with a hybrid approach combining unsupervised classification and knowledge-based spatial reclassification. Detailed description of these procedures can be found from Yang (2002). These layers represent five different dates, namely, 1973, 1979, 1987, 1993, and 1999.

**2.1.2 Road:** It contains not only major road networks but also node points and large shopping malls. For convenience, this layer is still named as 'road'. The major highways were extracted from the AND global highway database (http://www.and.com), and then updated with satellite images to form 1973, 1987, and 1999 highway layers for three years. Major node points are either (major) highway exits, junctions, or towns where major highway(s) runs across. They may be of strategic significance for commercial or industrial development. Three layers of large mall polygons were extracted from the 1973, 1987, and 1999 Landsat images. A weighting system was established for highways, nodes, and malls, respectively.

The layers of highways, nodes, and shopping malls in the same year were combined to form a single 'road' layer. In this way, three 'road' layers were produced for 1973, 1987, and 1999, respectively. The 'road' layer for the year of 2025 was produced by overlaying the 1999 roads with the improved roadways and new roadways according to the 2025 Regional Transportation Plan (Atlanta Regional Commission, 2000).

**2.1.3 Excluded Area for Development:** Two layers of excluded areas were assembled. The first layer is a binary image, consisting of the water extracted from 1973 Landsat MSS image and the public lands. The latter includes national parks/refuge and wilderness areas, archaeological sites/areas, historic sites, off-road vehicle sites/areas, wild and scenic areas, state parks, USDA land, wildlife management areas, and county Parks. These areas were not allowed for urban development. This layer was mainly used for the model calibration.

For the future growth prediction, another layer was built, with probabilities of exclusion included. All excluded areas in the first layer were still preserved and assigned a value of 100. Additionally, this layer contains three levels of buffer zones around major streams in the study area.

**2.1.4 Slope and Shaded Relief Image:** In order to produce terrain slope and shaded relief images, a seamless DEM image was constructed by mosaicking 159 USGS 7.5' DEMs covering the entire modeled area. Then, a terrain slope image was computed and represented in percentage. Furthermore, a layer of the hillshaded image was computed from the DEM. This image shows the topographic relief in the study area. It was used as a background image for visualization purpose only.

### 2.2 Model Calibration

The purpose of model calibration was to determine the best values for five control coefficients, namely, diffusion (overall dispersiveness of growth), breed (likelihood of new settlements being generated), spread (growth outward from existing spreading centers), slope resistance (likelihood of settlements extending up steeper terrain), and road gravity (attraction of urbanization near road networks).

The calibration was built upon on a statistical approach. Thirteen statistical measures were computed to quantify the historical fit between the modeled results and historical urban extent data extracted from remotely sensed imagery. The list of these statistical measures and their detailed description are given elsewhere (Yang and Lo, 2003). They were used to narrow down the range and determine the best value for each control coefficient. The possible range is between 0 and 100 and the possible combinations for the five control coefficients are approximately $5^{100}$ or $7.89 \times 10^{69}$! Ideally, each combination should be assessed. Given the computational resources available (a Sun Ultra Model 1, with 143 Mhz CPU and 64 Mb RAM), however, this would take years to complete according to an earlier test. For the time and computational resources constraints, the calibration was broken down into three phases (Table 1). The coarse calibration was to block out the widest range for each control coefficient. The fine calibration was to narrow down the ranges to approximately 10 or less The final calibration was to determine the best combination, which had the following starting values: diffusion(55), breed (8), spread(25), slope resistance (53), and road gravity (100).

Table 1 Calibration runs: input data, calibration files, number of Monte Carlo iteration, computation time, and outputs.

| Items | | Future Simulations | |
|---|---|---|---|
| | | Scenario 1 | Scenario 3 |
| Time Span | | 2000-2050 | |
| Input Data | Resolution (m) | 240 | |
| | urban extent (year) | 1999 | |
| | 'roads' | 1999 | 1999, 2025 |
| | excluded areas | stream buffered zones not considered | stream buffered zones considered |
| | slope | same (only one layer can be chosen) | |
| | hillshaded relief | same (only one layer can be chosen) | |
| Self Modification Constraints* | critical_high | 1.500 | |
| | critical_low | 0.050 | |
| | boom | 1.010 | |
| | bust | 0.090 | |
| | critical_slope | 21 | 10 |
| Control Coefficients** | diffusion | 71/88 | 100/100 |
| | breed | 10/12 | 100/100 |
| | spread | 32/40 | 15/15 |
| | slope resistance | 73/100 | 10/40 |
| | road gravity | 100/100 | 200/2003*** |
| Number of Monte Carlo Computations | | 100 | |
| Random Samples | | 2840 | |

* *These are about 1 percent of the total pixel numbers.*
** *This number is for the Monte Carlo iterations.*

One more step was conducted to determine the starting values used for future growth simulation. The best values identified above were actually the starting values. Because of self modification incorporated in the model, these starting values tend to be altered when a run is completed. Thus, a coefficient may have different starting and finishing values for each run. At the end, the final values of the control coefficients are: diffusion(71), breed(10), spread(32), slope resistance(73), and road gravity(100).

It should be pointed out that the model calibration was carried out with the use of 240 m resolution dataset only. An earlier test estimates that the time for completing the first stage of calibration using the 120 m resolution data set would be about 32,500 hours or 135 days given the computer resource available. For practical reason, the other two higher resolution data sets were not used in model calibration.

### 2.3 Scenario Design and Simulation

Two possible planning scenarios for future urban development in Atlanta were considered here, which are tied with different policies and environmental conditions.

**2.3.1 Scenario One:** This scenario assumes the factors for the growth remain unchanged and thus, it may be termed as ‘continuation'. It provides therefore a benchmark for comparison with the other alternative growth strategy. To implement this scenario in model simulation, the values of growth control coefficients obtained from the model calibration were used as the starting values. The 1999 urban extent data was actually used in the simulation and other conditions and input data set can be found from Table 2.

Table 2 The conditions applied for each simulation

| | Item | Calibration Runs | | | Parameter Averaging |
|---|---|---|---|---|---|
| | | Coarse | Fine | Final | |
| Input Data | resolution (m) | 240 | | | |
| | urban extent (year) | 1973, 1979, 1987, 1993, 1999 | | | |
| | ‘roads' (year) | 1973, 1987, 1999 | | | |
| | excluded area | stream buffered zones not considered | | | |
| | slope | same (only one layer is available) | | | |
| | hillshaded relief | same (only one layer is available) | | | |
| Calibration File | seed* | 2840 | 2840 | 2840 | 2840 |
| | number_of_times** | 4 | 6 | 10 | 100 |
| | diffusion_coeff_start | 0 | 40 | 52 | 55 |
| | diffusion_coeff_step | 25 | 5 | 3 | 1 |
| | diffusion_coeff_stop | 100 | 60 | 58 | 55 |
| | breed_coeff_start | 0 | 0 | 2 | 8 |
| | breed_coeff_step | 25 | 5 | 3 | 1 |
| | breed_coeff_stop | 100 | 10 | 8 | 8 |
| | spread_coeff_start | 0 | 15 | 22 | 25 |
| | spread_coeff_step | 25 | 5 | 3 | 1 |
| | spread_coeff_stop | 100 | 35 | 28 | 25 |
| | slope_resistance_start | 0 | 40 | 47 | 53 |
| | slope_resistance_step | 25 | 5 | 3 | 1 |
| | slope_resistance_stop | 100 | 55 | 53 | 53 |
| | road_gravity_start | 0 | 80 | 90 | 100 |
| | road_gravity_step | 25 | 10 | 5 | 1 |
| | road_gravity_stop | 100 | 100 | 100 | 100 |
| Number of iteration(s) | | 3125 | 900 | 243 | 1 |

* *The definitions are: critical_high : when the growth rate exceeds this value, self-modification increases the control parameter values; critical_low: when the growth rate falls below this value, self-modification increases the control parameter values); boom: value of the multiplier (greater than one) by which parameter values are increased when the growth rate exceeds critical_high; bust: value of the multiplier (less than one) by which parameter values are decreased when the growth rate falls below critical_low), and Critical_slope: average slope at which system increases spread.*

** *Both starting and ending values are given. It should be noted that the ending values were the averaged values after100 times of Monte Carlo computations.*

*** *Program code was changed to allow up to 200 for road gravity.*

**2.3.2. Scenario Two:** The second scenario considers a hybrid growth strategy in which both conventional suburban development and alternative growth efforts, such as smart growth and new urbanism, are addressed. This scenario also considers environmental conservation by limiting development around several predefined buffer zones.

To implement this idea in model simulation, the starting values for five growth control coefficients used in the first scenario need to be changed in order to slow down the growth rate and to alter the growth pattern. The conditions used in this scenario can be seen from Table 2. Please note that the proposed transportation improvements and new additions as well as environmental conservation introduced in the second scenario are still valid here.

Although the two scenarios are different in policies and environmental conditions, there are several commonalities. The time span is the same, which is from 2000 to 2050. Because of the limitation in computation resources, only the data set with 240 m spatial resolution is used. The two input data layers, namely, slope and hillshaded relief, are used without change for all the runs. The number of times of Monte Carlo computations is 100 and the random samples are 2,840, or about 1 percent of the total pixels available.

## 3. RESULT

The progressive urban development as projected into the future 51 years under two different scenarios can be perceived quite well from Figure 2. The graphical outputs of the two scenarios are quite similar. By evaluating these graphical outputs carefully, it is found that a Los Angeles-like metropolis characterized by huge urban agglomerations would emerge by 2030, if current development conditions are still valid. The vegetation area and open space in the 13 metro counties (excluding the northwestern mountainous area) would be very limited. In contrast, the simulated urbanization under the second scenario appears to be relatively restrictive, indicating that the effort of slowing down urbanization through model parameterization has been quite efficient.

Statistical measures reveal much more information. Under the first scenario, the total urban area for 2050 would be 1,286,692 ha. The total net increment in urban area with at least 50% probability would be 793,561 ha., or 43.6 ha. per day on the average, representing an increase of 160 percent between 1999 and 2050. As a result of such a dramatic growth, urban land would occupy approximately 78.67 percent of the total modeled land by 2050. The averaged slope steepness for urban land would increase from 4.87 percent in 1999 to 8.32 percent in 2050 (Table 3), indicating many woody area would be converted into urban use.

Under the second scenario, by 2050, the total urban area would be 906,134 ha., or approximately 55.40 percent of the entire modeled area. The total net urban increment would be 413,003 ha., or 22.2 ha. per day, indicating an increase of 84 percent between 1999 and 2050. Apparently, the magnitude of urban growth as projected under this scenario has been substantially suppressed. The mean

Table 3. Statistics of simulation results for different development scenarios.

| | | | 2010 | | 2030 | | 2050 | |
|---|---|---|---|---|---|---|---|---|
| | | | Area (hectare) | Percent | Area (hectare) | Percent | Area (hectare) | Percent |
| Scenario One | Probability | 50-59** | 41985 | 2.57 | 21767 | 1.33 | 12390 | 0.76 |
| | | 60-69 | 51817 | 3.17 | 29716 | 1.82 | 17447 | 1.07 |
| | | 70-79 | 60653 | 3.71 | 43419 | 2.65 | 26173 | 1.60 |
| | | 80-89 | 61269 | 3.75 | 76326 | 4.67 | 46276 | 2.83 |
| | | 90-100 | 27740 | 1.70 | 469428 | 28.70 | 691281 | 42.26 |
| | Total urban area*** | | 736595 | 45.03 | 1133787 | 69.32 | 1286692 | 78.67 |
| Scenario Two | Probability | 50-59 | 29843 | 1.82 | 31444 | 1.92 | 22239 | 1.36 |
| | | 60-69 | 22038 | 1.35 | 41507 | 2.54 | 31110 | 1.90 |
| | | 70-79 | 7188 | 0.44 | 54950 | 3.36 | 45942 | 2.81 |
| | | 80-89 | 559 | 0.03 | 75635 | 4.62 | 74920 | 4.58 |
| | | 90-100 | 0 | 0.00 | 75848 | 4.64 | 238798 | 14.60 |
| | Total urban area | | 552758 | 33.79 | 772514 | 47.23 | 906134 | 55.40 |

* It is computed by using urban area divided by the total modeled area (1,635,656 hectares).

** This is the probability of predicted urbanization.

*** It contains 1999 urban area (493,131 hectares)

slope steepness for urban land would decrease from 4.87 percent in 1999 to 4.46 percent in 2050, implying that many crop land would be converted into urban uses.

The spatial distribution of simulated urbanization under different scenarios can be discerned from Figure 2. For the first scenario, the projected urban additions for the period of 1999-2010 are largely adjacent to the 1999 urban pixels, which can be viewed as 'continuation' of urbanization. This is in line with the statistics given in Table 3, which show that more than 99 percent of the net urban growth under this scenario are accounted for by the organic growth. This type of urban growth actually represents the expansion of existing urban pixels into their surroundings. The projected urban additions during 2010-2030 are largely distributed over places far away from the 1999 urban land. Many projected additions are also found in western, northwestern, and southeastern parts. Some large urban clusters can be clearly recognized. The projected urban additions after 2030 are predominately scattered over the western and southeastern parts. Under the second scenario, the projected urbanization for the period of 1999-2010 has been very limited. Most of the new additions are for the period of 2010-2030, represented by blue and green pixels in Figure 2 (Note: The original figure is in color. Readers can contact the author for obtaining a copy of this color plate). Numerous large urban clusters can be clearly recognized, particularly in southern and western parts.

## 4. CONCLUSION

This study has demonstrated the usefulness of remote sensing, dynamic modeling and geographic information technologies for urban planning. The model used here has been tested by its developers for long-term urban growth prediction in two study sites. From a user's perspective, this study has moved forward to investigate the effectiveness of the model as a tool to imagine, test and choose between different scenarios. These scenarios represent different growth strategies that can be adopted by planners. This is an area on which substantial research efforts need to be made in order to adopt dynamic modeling technology for problem solving in applied urban studies.

At the application level, this study has established a well-documented regional case study focusing on Atlanta, a metropolis without any major natural barriers. The two scenarios designed with different environmental and policy conditions have largely represented the major possible planning strategies. These result from the first scenario indicates Atlanta would be the next 'Los Angeles' by approximate 2030 if the current rate and pattern of urban growth do not alter. These will serve as a good warning to planners in Atlanta. In contrast, the result from the second scenario shows that much more greenness and open space, including buffer zones of large streams and lakes, could be preserved. Accordingly, the second scenario should be the most desirable for the future urban

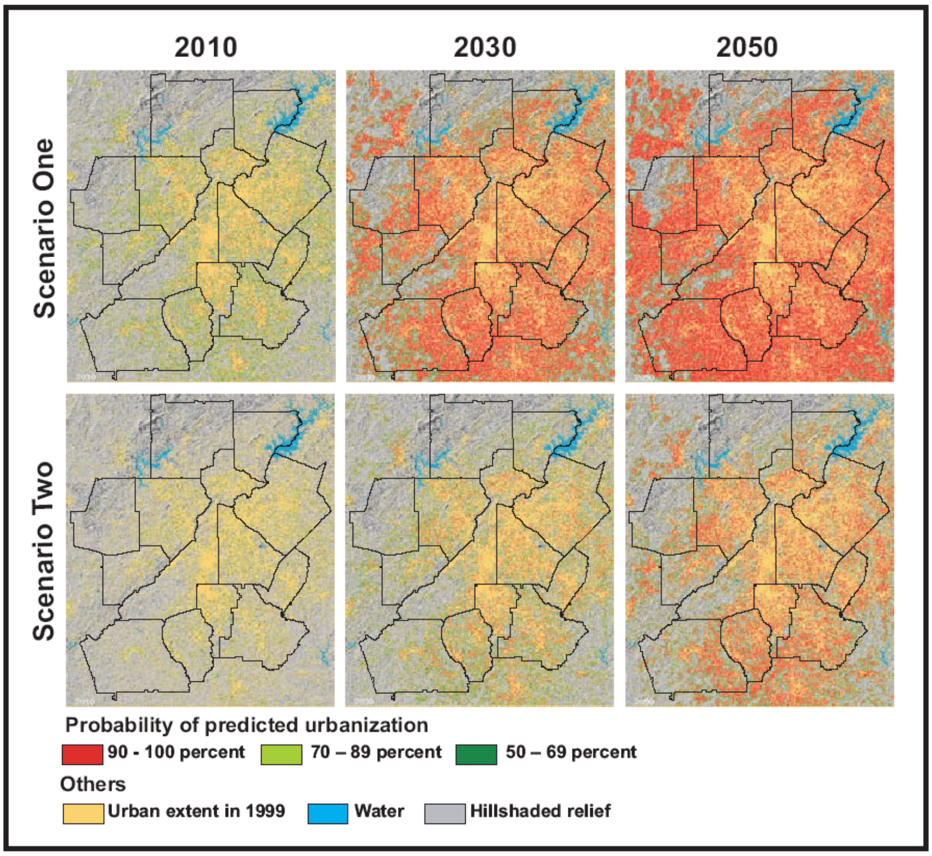


Figure 2 Simulation of the spatial consequences of future urban growth under different scenarios. Note that the county boundary is overlaid.

growth of Atlanta. Given that many major metropolises in the United States face the growing problems caused by restless suburban development, the technical frameworks developed in the current study focusing on Atlanta should be applicable to those metropolises with similar growth styles.

## 5. REFERENCES


Atlanta Regional Commission (ARC), 2000. The 2025 Regional Transportation Plan, URL: http://www.atlantaregional.com (Last date accessed: 10 November 2002).

Batty, M. and Y. Xie, 1994. From cells to cities, Environment and Planning B, 21: 531-548.

Clarke, K., 2000. SLEUTH: Land Cover transition Model, version 3.0, URL: http://www.ncgia.ucsb.edu/projects/gig/project_gig.htm.

Clarke, K. C. and J. Gaydos, 1998. Loose-coupling a cellular automaton model and GIS: long-term urban growth prediction for San Francisco and Washington/Baltimore, International Journal of Geographic Information Science, 12(7):699-714.

Meaille, R. and L. Wald, 1990. Using geographical information system and satellite imagery within a numerical simulation of regional urban growth, International Journal of Geographic Information Systems, 4 (4): 445-456.

Silva, E. A. and K. C. Clarke, 2002. Calibration of the SLEUTH urban growth model for Lisbon and Porto, Portugal, Computers, Environment and Urban Systems, 26(6): 525-552.

Verburg, P.H., A. Veldkamp, and L. O. Fresco, 1999. Simulation of changes in the spatial pattern of land use in China, Applied Geography, 19: 211-233.

Wang, Y. and X. Zhang, 2001. A dynamic modeling approach to simulating socioeconomic effects on landscape changes, Ecological Modelling, 140(2001): 141-162.

White, R. and G. Engelen, 1997. Cellular automata as the basis of integrated dynamic regional modelling, Environment and Planning B, 24: 235-246.

Wu, F. and C. J. Webster, 1998. Simulation of land development through the integration of cellular automata and multicriteria evaluation, Environment and Planning B, 25: 103-126.

Yang, X., 2002. Satellite monitoring of urban spatial growth in the Atlanta metropolitan region, Photogrammetric Engineering & Remote Sensing, 68 (7): 725-734.

Yang, X. and C. P. Lo, 2003. Modelling urban growth and landscape changes in the Atlanta metropolitan area. International Journal of Geographical Information Science, 17(5):463-488.