Replication of Spatio-temporal Land Use Patterns at Three Levels of Aggregation by an Urban Cellular Automata

Charles Dietzel[1] and Keith C. Clarke[2]

[1] University of California, Santa Barbara, Department of Geography, 3611 Ellison Hall, Santa Barbara, CA 93106, USA
dietzel@geog.ucsb.edu

[2] National Center for Geographic Information Analysis, University of California, Santa Barbara, Department of Geography, 3611 Ellison Hall, Santa Barbara, CA 93106, USA
kclarke@geog.ucsb.edu

Abstract. The SLEUTH urban growth model [1] is a cellular automata model that has been widely applied throughout the geographic literature to examine the historic settlement patterns of cities and to forecast their future growth. In this research, the ability of the model to replicate historical patterns of land use is examined by calibrating the model to fit historical data with 5, 10, and 15 different land use classes. The model demonstrates it robustness in being able to correctly replicate 72-93% of the land use transitions over an eight-year time period, in both space and time.

1 Introduction

The logic and mechanisms of cellular automata (CA) allow linking the local to the global, not just through simulation and model forecasting, but also in the sense that global patterns and forms can be illustrated through local processes [2]. While there is a difficulty in simulating large systems at the micro-scale, understanding the processes and evolution of form at the smallest level allows for a better understanding and modeling of process taking place on the next hierarchical scale. The idea of cellular automata-like geographic models, simulating change at the local scale can be traced to Tobler [3], but a more formal outline of CA models and their possible use in simulating urban systems was made by Couclelis [4]. The use of CA for urban systems was slow to catch on, taking nearly a decade before there was a broad body of literature. Adaptation, experimentation, and application of these models to urban systems has been quite prolific in more recent years. One of the lures of these models is as a metaphor for urban growth and change [4], [5], but the models have the ability to at least attempt to simulate real-world systems, if not accurately mirror them [2], [7], [8], due to several advantageous properties of CA.

There are certain properties of two-dimensional CA that make them especially advantageous to use in the modeling of geographic systems. The most obvious is that

P.M.A. Sloot, B. Chopard, and A.G. Hoekstra (Eds.): ACRI 2004, LNCS 3305, pp. 523–532, 2004.

CA models are spatial in the same manner that an urban or any other geographic system is. This treatment of space in an absolute manner is an advantage over other urban and regional models (spatial interaction, gravity, econometric and location-allocation) where the treatment of space is relative. The spatial aspect of CA is a natural link to geographic and remotely sensed data; much of which is used as input for these models. The raster structure of GIS and remotely sensed data sources is the same lattice structure as that present in CA models, making them ideal sources for regular grid data.

The process of simultaneous computation in CA allows for modelers to view urban systems growing over time in increments instead of just the beginning and end points. This is not to say that CA model a system at a continual temporal scale, just that the increments between time periods can be set to such a small temporal period that the result is a model that appears to be continuous in time. The flexibility in temporal dynamics that a CA model provides allows the simulation of events that occur at various timeframes; from pedestrian behavior to the growth of cities. CA act within a localized neighborhood, creating micro-scale dynamics; but when the overall micro-scale behavior of the system is taken collectively, there emerges macro-scale pattern. These dynamics are typical of complex systems where the local elements are allowed to interact, creating the macro-scale perspective.

The lattice structure and link to geographic and remotely sensed data makes CA models highly visual – giving modelers and model users the ability to visualize the results of model forecasts. This is especially helpful when models are being simulated for multiple scenarios and analysis is done within the results. Incorporating the temporal dynamics of CA with visualization, CA models allow users to view the dynamic growth process as it takes place within the model. With such advantages, the capabilities of CA for modeling urban systems are powerful. CA models easily incorporate common forms of geographic data, enables processes to take place at multiple scales, and produce outputs that are highly visual, increasing the understanding and appeal.

The SLEUTH urban growth model [1] has capitalized on these advantages to be successfully applied to a wide variety of geographic areas [9], [10], [11]. Yet the accuracy of the model in evaluating different quantities of land use classes has not been rigorously evaluated. The goal of this research is to examine the ability of the SLEUTH model to replicate the spatio-temporal patterns of land use change over an eight year period with the same dataset classified into 5, 10, and 15 land use classes. The purpose is to test the sensitivity of SLEUTH as a typical CA model to land cover class aggregation.

2 The SLEUTH Model

The SLEUTH model has the ability to model urban/non-urban dynamics as well as urban-land use dynamics, although the latter has not been widely used; presumably due to the limitations of gathering consistently classified land use data. The dual ability has led to the development of two subcomponents within the framework of the model, one that models urban/non-urban growth, the urban growth model (UGM) [1] and the other that models land use change dynamics (Deltatron). Regardless of whether each of these components is used, the model has the same calibration routine. The input of land use data during calibration activates the Deltatron part of SLEUTH.

SLEUTH is a moniker for the data required to calibrate and forecast this urban growth model. The model requires topographic data in the form of Slope and Hillshade maps, although the hillshade is used only for visualization purposes, and does not play a role in determining model outputs. Land use with consistent classification for two time periods are needed to implement the Deltatron submodel, they are not necessary to simulate urban growth, but are recommended. An Exclusion layer is used to place constraints on urban growth. Through the Exclusion layer, a user can specify where urban growth is allowed to occur, or where it is prohibited. This layer can also be a weighted layer so that 'resistances' against growth can be put in place in an attempt to slow or alter the rate of urbanization. Urban extent data is critical and necessary for this model. Four different temporal layers are needed, showing the extent of urban areas at different points in time. These maps serve as the control points, against which the model is calibrated, and a goodness of fit is determined. The last layer required for using SLEUTH is Transportation. The creation of these input maps is typically done within a geographic information system (GIS), and then they are converted to GIF format files which are the actual data used in the model.

For this model, the transition rules between time periods are uniform across space, and are applied in a nested set of loops. The outermost of the loops executes each growth period, while an inner loop executes growth rules for a single year. Transition rules and initial conditions of urban areas and land use at the start time are integral to the model because of how the calibration process adapts the model to the local environment. Clarke et al. [1] describe the initial condition set as the 'seed' layer, from which growth and change occur one cell at a time, each cell acting independently of the others, until patterns emerge during growth and the 'organism' learns more about its environment. The transition rules that are implemented involve taking a cell at random and investigating the spatial properties of that cell's neighborhood, and then urbanizing the cell, depending on probabilities influenced by other local characteristics [1]. Five coefficients (with values 0 to 100) control the behavior of the system, and are predetermined by the user at the onset of every model run. These parameters are:

1. *Diffusion* – Determines the overall dispersiveness nature of the outward distribution.
2. *Breed Coefficient* – The likelihood that a newly generated detached settlement will start on its own growth cycle.
3. *Spread Coefficient* – Controls how much contagion diffusion radiates from existing settlements.
4. *Slope Resistance Factor* – Influences the likelihood of development on steep slopes.
5. *Road Gravity Factor* – An attraction factor that draws new settlements towards and along roads.

These parameters drive the four transition rules which simulate spontaneous (of suitable slope and distance from existing centers), diffusive (new growth centers), organic (infill and edge growth), and road influenced (a function of road gravity and density) growth. By running the model in calibration mode, a set of control parameters is refined in the sequential 'brute-force' calibration phases: coarse, fine and final calibrations [9].

2.1 Deltatron Dynamics Within SLEUTH

Three types of land use transitions are assumed to take place in the Deltatron model. First are state changes where land makes a transition from one use to another (e.g. Forest to Agriculture). The Deltatron assumes that urbanization is the driver of change within the system, and that once a cell has become urbanized, it is not possible for it to transform to another state. Neighborhood transitions are also assumed where if one cell is transformed into another land use class, that similar surrounding cells have a higher probability of changing to that land use class as well. The last transition that is assumed to take place is the discrete location change, where a particular cell, while influenced by neighborhood change, changes land use on an individual level. This may be from water to wetlands or vice versa. Thus land use changes in the model are assumed to have spatio-temporal autocorrelation. This violates the "classical" CA assumptions because changes have a "memory" beyond one time cycle, albeit localized.

Three types of change influence transitions. The first is a simple Markov transition matrix that calculates the annual probability of change between all pairs of land uses. The matrix is calculated by differencing the two input land use datasets and is normalized by the number of years between the control years. Topography is also assumed to influence land use change, with slope being a factor in determining where land use classes may occur within the landscape. Urbanization will occupy the flattest land available, while areas with steep slopes are better fit for other classes such as forests and natural vegetation. The driver of change is urbanization. As the amount of urbanization increases, so too does the model's attempts at land use changes, with urban land cover consuming land that is currently in use by other land uses.

Clarke's [12] paper provides the initial outline of the Deltatron model and its primary assumptions:

1. That land transitions be considered to take place on a uniform spacing grid.
2. That transition is between and among a finite set of states, where the number of states is small
3. That the transition matrix accurately estimates land use state transition probabilities from observed counts.
4. That an external model be used to change the state of the dominant or driving class.
5. That there should exist considerable spatial autocorrelation in land transitions.
6. That there exists temporal correlation between land transitions.
7. That specific land transitions are influenced by context.

8 That land transitions happen to some degree at random, i.e. independent from the driving force.

Deltatrons themselves are defined as bringers of change within land use change space; they are the successful culmination of forces of change within the system and have their own life cycles. Deltatrons are born out of land use change between time periods. In proceeding years the likelihood of similar conversion is still high due to the assumption of spatial autocorrelation between land use transitions, and the Deltatron has the ability to affect and promote similar transitions in its neighborhood. Yet when a new land use class has been established, the probability of an immediate transition to another land use class is low. When this is the case, a Deltatron acts as a placeholder or memory and prevents subsequent transition of the land class for the duration of its lifetime. How many cycles, or years, a Deltatron lives may be able to be used as a modifiable parameter to fine-tune the model per application. Due to their sensitivity to local influences, and ability to modify behavior over time, Deltatrons are critical to the spatio-temporal autocorrelation of the land transitions. This process of Deltatron land use change is initiated within the model through two distinct phases. The first is the creation of change within the landscape, and the second is spreading the change throughout it. The result has been described as the dropping of a pebble in pool of water and then the diffusion of ripples throughout the remainder of the pool.

2.2 Creation of Landscape Change

The creation of landscape change is driven by the number of cells that were newly urbanized in the UGM. The change cycles are initiated by choosing a cell at random and then testing if it is suitable for change. There are four conditions that are enforced which prevent some cells from changing: (1). the cell contains no data. (2). the cell is already urban. (3) the cell is in some land use class that has been determined by the user to be incapable of change, water for example. (4). a Deltatron is already present at that location. When a suitable (not meeting the four conditions) cell is found, then two land use classes are randomly chosen, and the class that has the average slope value closest to the selected cell is selected, allowing topography and land cover to both play a role in the transition process. The probability of transitioning between the initial class and the randomly chosen class is calculated, and if a randomly drawn number is greater than the probability of change, then the cell does not change states, and the next random cell is selected. On the other hand, if the randomly drawn number is less that the probability of transition between the two classes, the transition takes place and is encouraged to randomly spread to its neighbors, creating a cluster. At the end of this first phase of growth, several clusters of land use transitions have been made, and a deltaspace is created, tracking the locations of changes in space and time, from which Deltatrons are 'born,' and their ages monitored. This process is summarized in Figure 1.

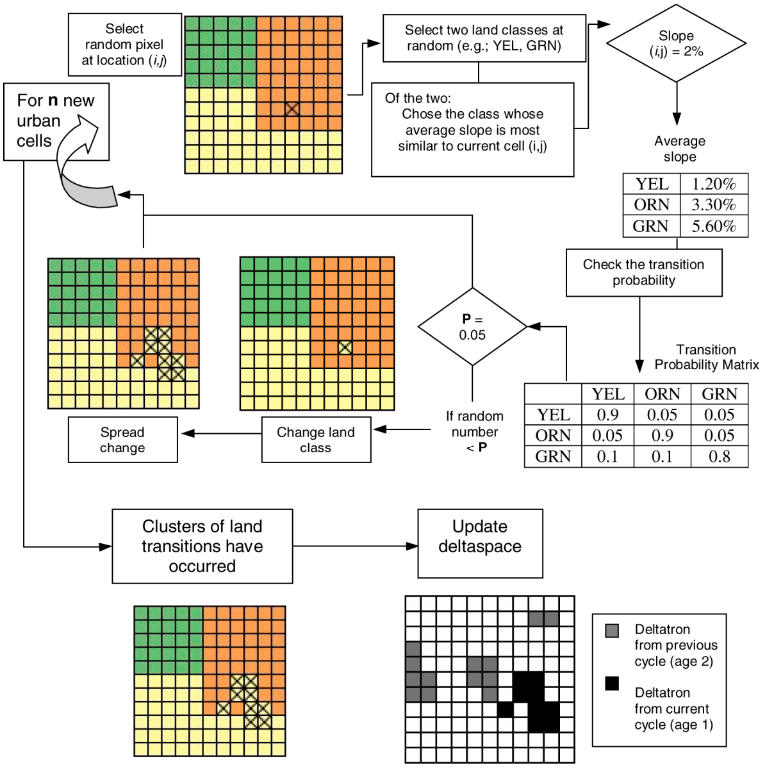

Fig. 1. Summarization of the first phase of the Deltatron model, creation of land use change.

2.3 Proliferation of Landscape Change

In the second phase of the Deltatron model, existing Deltatrons attempt to initiate change on available land within their local neighborhood. The rules for initiating this change are quite simple (Figure 2). If a suitable cell is neighbored by two or three Deltatrons that were created in a previous time step, then an attempt is made to create a transition within that cell. The requirement of two or three neighbors is randomly enforced. If a cell is selected for transition, then the neighboring Deltatrons are queried to determine the land use change that is being proliferated. As is done in the creation of land use change, the probability of land use transition is tested, and a random draw is made to determine if the cell maintains its current state or it is transformed, creating a new Deltatron. Once the process of proliferating change is completed, the deltaspace is updated, and the Deltatrons are either aged or killed.

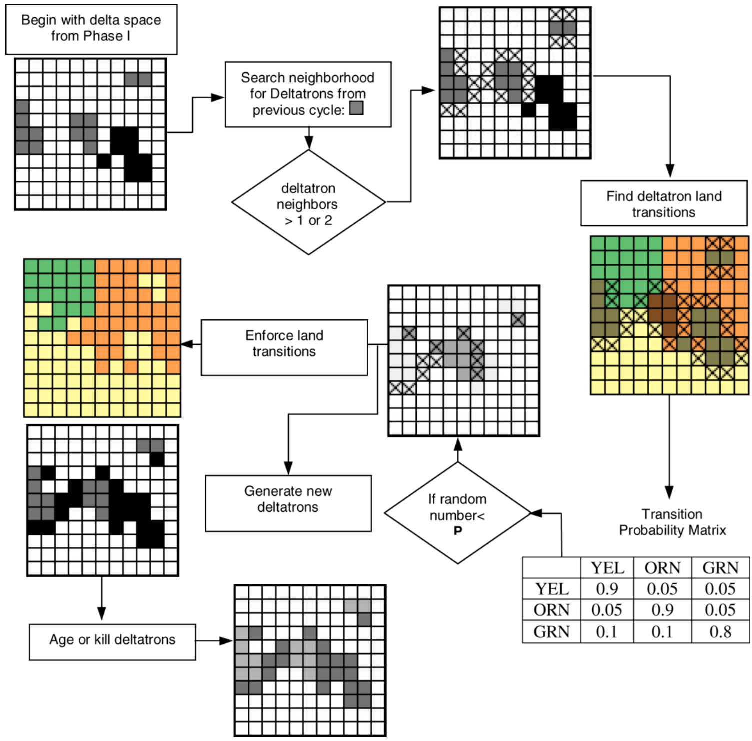

Fig. 2. Summarization of the second phase of the Deltatron model, proliferation of land use change.

The calibration process is done in three stages: coarse, fine, and final. The coarse calibration begins with parsing the parameter space into five areas and using the values of 1, 25, 50, 75, and 100 for each of the five parameters. This gives 3,125 different parameter sets that are tested to determine which range of parameters the one parameter set that best describes the data is located within. Results from the coarse calibration are examined to determine the goodness of fit for each of the parameter sets based on a set of spatial metrics. Narrowing of the parameter set can be based on a variety of different goodness of fit measures and there is not one sole metric that has been shown to be the most effective. For this research a metric that was the product of four individual metrics was used. The four individual metrics looked at the models ability to model the urban area in each of the input dataset, replicate the spatial shape of the input datasets, and replicate the patterns of land use in space and time.

3 Land Use Data

Land use data for San Joaquin County (California, USA), from 1988 and 1996 was downloaded from the California Department of Water Resources webpage. Data were converted from polygons to raster (100m resolution) in a GIS. The data originally consisted of 20 classes that were reclassified to 5, 10, and 15 land use classes as shown in Table 1. Urban extent data were obtained for 1988, 1992, 1994, and 1996 as described in [13].

Table 1. Reclassification scheme of the land use data from San Joaquin County (California, USA) into data sets aggregated into 5, 10, and 15 land use classes

Land Use Class	Reclassified	Code 5	Code 10	Code 15
Citrus and Subtropical	Fruit, Nut, & Vegetables	1	1	1
Deciduous Fruits & Nuts	Fruit, Nut, & Vegetables	1	1	2
Field Crops	Field Crops	1	2	3
Grain and Hay	Field Crops	1	2	4
Idle	Field Crops	1	2	5
Pasture	Pasture	1	3	6
Rice	Field Crops	1	2	7
Truck, Nursery, and Berry	Fruit, Nut, & Vegetables	1	1	8
Vineyards	Vineyards	1	4	9
Barren and Wasteland	Barren	2	5	10
Riparian Vegetation	Riparian	3	6	11
Native Vegetation	Native Vegetation	3	7	12
Water Surfaces	Water	4	8	13
Semi Agricultural	Feedlots	1	9	14
Urban	Urban	5	10	15
Commercial	Urban	5	10	15
Industrial	Urban	5	10	15
Landscape	Urban	5	10	15
Residential	Urban	5	10	15
Vacant	Urban	5	10	15

4 Calibration Results

Results from the calibration show that SLEUTH was able to accurately replicate 93% of the land use changes for the dataset with five land use classes, and 77% and 72% for the dataset with ten and fifteen land use classes (Table 1).

SLEUTH was exceptionally good at modeling the total number of urban pixels in the last year of input data as indicated by the *compare* statistic. It also performed well in replicating the overall urbanization of the input time-series as shown in the *population* statistic, which is a least squares regression score (r^2) for modeled urbanization compared to actual urbanization for the time series. The Lee-Sallee metric [14] was used to evaluate the ability of the model to spatially match historical data; that is, how good was the model at replicating the number of urban pixels and their location in space. The model was able to accurately replicate 74% of these patterns. Not surprisingly, the model's performance with regard to these three statistics was very similar since they all used the same urban extent data over the same time period. The most stringent measure of ability of the model to accurately reproduce land use patterns was measured by the F-match statistic.

The F-match statistic measured the proportion of goodness of fit across land use classes[14]. SLEUTH was able to accurately replicate 93% of the land use changes for the dataset with five land use classes, and 77% and 72% for the dataset with ten and fifteen land use classes.

Table 2. Results from calibrating the SLEUTH urban growth model to replicate land use change patterns of datasets with 5, 10, and 15 land use classes between 1988 and 1996. Four spatial statistics (all scaled 0 to 1, with 1 being perfect replication of data) were used as measure of the models ability to replicate the input data. Details on these statistics are described in [14]

	Goodness of Fit Measures				
Land Use Classes	Compare	Population	Lee-Sallee	F-Match	Composite
5	1	0.83331	0.74278	0.92973	0.575471
10	0.99152	0.83802	0.74427	0.76968	0.475989
15	0.99429	0.83768	0.74267	0.72439	0.448084

5 Conclusions

This research has shown the robust ability of the SLEUTH urban growth and land use change model to replicate land use patterns in both space and time. While the model was better able to replicate the evolution of land use patterns with less detailed data, it was to be expected since the Deltatron submodel has a stochastic component that decreases the probability of picking the correct land use transition as more classes are added. In its current format the model is only capable of handling land use transitions between two time periods. With the increasing availability of multi-temporal spatial data, it may be possible to alter the Deltatron model to handle land use data for multiple time periods. If possible, it will be interesting to reevaluate the robustness of the model in handling multiple classes of land use data.

While previous use of this model has focused on the implementation and application of this model to specific case studies [10], [11], this research is the first to evaluate the ability of the land use model to replicate numerous land use patterns. This has been largely due to two reasons, one that is strikingly concerning. First, the public availability of multi-temporal land use data is not widespread, and there is a significant time lag between when imagery (the basis of most classifications) is taken, and when it is classified by land use and released. The other reason that this has not been done is the general attitude in the geographic modeling community that if a model works and gives credible results, then it should be accepted without further poking and prodding. While the results of this research show the robustness of the model for simulating a variety of different land use classes, it is somewhat disturbing that other researchers [10], [11] have not tested the sensitivity of the model's performance to the number of land use classes. Hopefully this research will encourage other modelers to make strides in their work to further test the mechanics and sensitivity of their models.

References

1. Clarke, K. C., Hoppen, S., Gaydos, L.: A self-modifying cellular automaton model of historical urbanization in the San Francisco Bay area. Environment and Planning B 24 (1997) 247-261
2. Batty, M., Xie, Y.: Modelling inside GIS: Part 1. Model structures, exploratory spatial data analysis and aggregation. International Journal of Geographical Information Systems 8 (1994) 291-307
3. Tobler, W.: Cellular Geography. In Gale, S., Olsson, G. (eds): Philosophy in Geography. D. Reidel Publishing Company, Dordrecht Boston London (1979) 379-386
4. Couclelis, H.: Cellular worlds: a framework for modeling micro-macro dynamics. International Journal of Urban and Regional Research 17 (1985) 585-596
5. Torrens, P., O'Sullivan, D.: Cellular automata and urban simulation: where do we go from here? Environment and Planning B 28 (2001) 163-168
6. Couclelis, H.: Of mice and men: What rodent populations can teach us about complex spatial dynamics. Environment and Planning A 29 (1988) 99-109
7. Batty, M., Xie, Y.: Possible urban automata. Environment and Planning B 24 (1997) 175-192
8. White, R., Engelen, G.: Cellular automata and fractal urban form: a cellular modelling approach to the evolution of urban land-use patterns. Environment and Planning A 25 (1993) 1175-1199
9. Silva, E.A., Clarke, K.C.: Calibration of the SLEUTH urban growth model for Lisbon and Porto, Portugal. Computers, Environment and Urban Systems 26 (2002) 525-552
10. Yang, X., Lo, C.P.: Modelling urban growth and landscape change in the Atlanta metropolitan area. International Journal of Geographical Information Science 17 (2003) 463-488
11. Jantz, C.A., Goetz, S.J., Shelley, M.K.: Using the SLEUTH urban growth model to simulate the impacts of future policy scenarios on urban land use in the Baltimore-Washington metropolitan area. Environment and Planning B 31 (2004) 251-271
12. Clarke, K.C.: Land use modeling with Deltatrons. The Land Use Modeling Conference, Sioux Falls, South Dakota. June 5-6 (1997) http://www.ncgia.ucsb.edu/conf/landuse97/
13. Dietzel, C.: Spatio-temporal difference in model outputs and parameter space as determined by calibration extent. In: Atkinson, P., Foody, G., Darby, S., Wu, F. (eds): Geodynamics. Taylor and Francis, London (2004)
14. Project Gigalopolis Webpage. www.ncgia.ucsb.edu/projects/gig