# A self-modifying cellular automaton model of historical urbanization in the San Francisco Bay area

K C Clarke, S Hoppen
Department of Geology and Geography, Hunter College, The City University of New York
Graduate School and University Center, 695 Park Avenue, New York, NY 10021;
e-mail: kclarke@geog.UCSB.EDU

L Gaydos
US Geological Survey, EROS Data Center, NASA–Ames Research Center, Moffett Field,
CA 94035; e-mail: lgaydos@mail.arc.nasa.gov
Received 23 January 1995; in revised form 24 July 1995

**Abstract.** In this paper we describe a cellular automaton (CA) simulation model developed to predict urban growth as part of a project for estimating the regional and broader impact of urbanization on the San Francisco Bay area's climate. The rules of the model are more complex than those of a typical CA and involve the use of multiple data sources, including topography, road networks, and existing settlement distributions, and their modification over time. In addition, the control parameters of the model are allowed to self-modify: that is, the CA adapts itself to the circumstances it generates, in particular, during periods of rapid growth or stagnation. In addition, the model was written to allow the accumulation of probabilistic estimates based on Monte Carlo methods. Calibration of the model has been accomplished by the use of historical maps to compare model predictions of urbanization, based solely upon the distribution in year 1900, with observed data for years 1940, 1954, 1962, 1974, and 1990. The complexity of this model has made calibration a particularly demanding step. Lessons learned about the methods, measures, and strategies developed to calibrate the model may be of use in other environmental modeling contexts. With the calibration complete, the model is being used to generate a set of future scenarios for the San Francisco Bay area along with their probabilities based on the Monte Carlo version of the model. Animated dynamic mapping of the simulations will be used to allow visualization of the impact of future urban growth.

## Introduction

Although a host of natural and human environmental activities, including fire, agriculture, and deforestation, have profound impacts upon global systems, the most striking human-induced land transformation of the current era is that of urbanization. Urbanization from a global environmental context is the conversion of natural to artificial land cover characterized by human settlements and workplaces. This single transformation involves a wholesale modification of natural processes such as runoff and evapotranspiration, and the short-term and long-term impacts touch every member of the human race every day.

In the short term, several decades, few of us in Europe or North America can fail to be startled when a visit back to the former fields and woodlands of our childhood playgrounds reveals only newly built-up urban land. In a longer timescale, 200 years, total global population has increased six times and the earth's urban population has increased over 100 times (Hauser et al, 1982). Driven by the Industrial Revolution, cities have gone from being a minor feature on our planet to a major one. The impact of urban land on economic and environmental systems is immense compared with its spatial extent although it can be difficult to grasp the notion that incremental growth at a regional level, the summation of myriads of individual human decisions, can amount to a significant global process.

Substantial growth in cities first occurred in Western Europe, America, and Japan but has spread in the latter part of this century throughout Asia, South America, and Africa.

Urban growth at the global scale shows no sign of slowing and is a phenomenon even in nations where population growth has stabilized. Eight of America's twenty largest metropolitan areas grew by at least 20% between 1980 and 1990 (Knox, 1993). The urban transition in less developed countries is also proceeding rapidly, so that cities now account for 36% of the world population (Hauser et al, 1982).

Even when we measure urban extent, we tend to underestimate the full impact. Cities require water, building materials, food, goods, and services from the surrounding region, converting natural land to agriculture, and agricultural land to urban land uses. Pond and Yeates (1994) estimated for a growing county in Canada that, in addition to the actual urban area, 20% of the land was in the process of the urban transition and 2% was in ex-urban uses, fully dependent on the urban areas. In addition, the 20th century has seen this impact on a cross-continental scale. Thus the consumption of a hamburger in California has already resulted in the conversion of land use in South and Central America.

**The human-induced land transformations project**
The United States Geological Survey (USGS) has a long tradition of studying land use and land cover, both current and potential. As a contribution to the US Global Change Research Program, the USGS initiated a human-induced land transformations project (HILT) to understand the urban transition from an historical and a multiscale perspective sufficient to model and predict regional patterns of urbanization 100 years into the future (Kirtland, 1993; Kirtland et al, 1994). The CA urban growth model reported here was developed as part of this study.

The model, now calibrated, will permit regional predictions of urban extent, providing a basis for assessment of the ecological and climatic impacts of urban change and the estimation of the sustainable level of urbanization in a region. The transferability of the model will be tested by calibrating it for other regions, beginning with the Washington, DC – Baltimore area. In this paper existing urban models are reviewed, the application of the HILT model to the San Francisco Bay area is outlined, and the rules governing growth and the tools and processes involved in the calibration of this self-modifying CA are described.

**Modeling urban transformations**
Traditional models of urbanization have sought to model and predict either the economic and size relationships between cities or the internal social and economic patterns within the limits of the city. Of the first type, the central place theory of Christaller, Zipf's rank-size rule, and the land-use transition model of Alonso and Muth in landscape economics have been most carefully examined (Wilson, 1978). A recent reinterpretation has introduced the fractal model as a mechanism behind at least Zipf's rule (Wong and Fotheringham, 1990). The Alonso and Muth model is aspatial, modeling primarily the demand curve relationship for land as a function of linear distance from a central marketplace. Each model allows for urban expansion and for distortions in the assumptions of the uniform isotropic plane beneath the 'economic' city simply by asserting that the evolving patterns will be drawn out along transportation routes because lower transportation costs mean lower overall costs.

Other models have relied less on geometry and economics and more on social and ethnic patterns as determinants of city structure (Jacobs, 1961). A model broadening this tradition predicts the structure and form of a city based on the difference between individuals' intentions and their behavior (Portugali et al, 1997). There are also many urban models which have been developed for a particular region, mainly for use by urban planners. BASS II is a model which forecasts urbanization at the regional

scale for the San Francisco Bay area (Landis, 1992). Although both BASS II and the model developed for HILT predict regional urban growth, they differ vastly in their level of detail, data requirements, and applications: BASS II is tailored to the specifics of the bay area; the growth rules in the HILT model are designed to be general enough to allow it to be applied to other regions.

Few theories have examined specifically the rural to urban transition as a physical process, except for the realization of the critical role that zoning and transportation have played in the ragged outer edge expansion of the rural–urban boundary. Even traditional geographical accounts of regional urban structure emphasize the crucial nature of the structure of transportation from trolley cars to airports, in determining the form of cities (Vance, 1964). In Christaller's model, the urban edge is largely ignored because the model seeks to predict the point provision of goods and services on a spatial tessellation of hexagonal market areas. The physical characteristics of urban expansion remained ill defined until the work of Batty who, together with his colleagues, used a dynamic systems model from physics called 'diffusion-limited aggregation' (DLA) to model urban expansion. The explicit link between the DLA process and the 'stringiness' effect of transportation routes on growth was an attractive one and some productive and interesting work resulted (Batty and Longley, 1994).

The DLA model also lends itself to the techniques of cellular automata, a simple and easily automated method for generating simulations (Couclelis, 1985). White was a pioneer in the application of computational CA models to urban areas and land use (White and Engelen 1992b). White's models used a classical CA approach. The modeling technique involves the following: (1) reduction of space to a grid or tessellation of cells, usually square grids; (2) establishment of an initial set of conditions, which does not have to be the origin of the entire system but can be any spatial arrangement of the phenomenon; (3) establishment of a set of transition rules between iterations; and (4) recursive application of the rules in a sequence of iterations of the spatial pattern.

Development of such a model involves determining the rules from an existing system, calibrating the CA to give results consistent with historical data (that is, predicting the present from the past), and then predicting the future by allowing the model to continue to iterate with the same rules. In White's models he used simple urban growth for some world cities and a more complex model of an island with self-modification and multiple land uses linked by rules. In self-modifying cellular automata, the rules are allowed to change as the system grows or changes (that is, by a feedback mechanism). For example, if all flat urban land is used by existing settlements, the rules penalizing building up slopes can be eased to reflect land pressure.

Urban modeling with cellular automata has become widespread. White and Engelen (1992a) have extended their land-use model to a whole island; Batty and Xie (1994) have modeled the historical growth of Cardiff, Wales, and Savannah, Georgia. In recent work cellular urban modeling is described as a new school of urban modeling, although one with roots in the work of von Neumann (1966), Wolfram (1994), Hägerstrand (1967), and Tobler (1979). The model presented here is a modified version of a CA which features the ability to modify parameter settings when the growth rate of the system exceeds a critical high or drops below a critical low value.

**The San Francisco Bay area as a test case**

The supermetropolitan San Francisco Bay area, which today stretches from the Golden Gate to the Sierra Foothills, was chosen because data were available; growth has been extensive; stresses on the natural systems, especially water supply, have been intense; and major policy questions abound that the research might help address. This region is

an ideal test site for the model because of its diversity: elevations range from sea level to 2500 m and land use ranges from wilderness to metropolitan areas.

Before 1850 settlement in this area consisted of many small enclaves whose distribution followed the inland waterway network. Only after the 1849 Gold Rush and completion of the transcontinental railroad did San Francisco emerge as the dominant city in the region because of its position as a transportation hub. The current distribution of urban population reflects the improvement of the highway system and suburban expansion into the surrounding valleys after World War 2. Corridors of urbanization line the major expressways which link the high-density urbanized areas in the region: Sacramento, San Francisco, San Jose, and the area around San Francisco Bay.

A digital database and geographic information system were assembled to support the animation, descriptive analysis, and modeling of urban land transformations for the bay region. The multiple source database included land cover, topography, climate, population, maps of settlement, historical transportation, aerial photography, and landsat multispectral scanner and thematic mapper data. The data sets have been assembled and made accessible as part of the HILT project at USGS. A full description of the data sets can be found in Kirtland et al (1994) and on the World Wide Web (USGS, 1994).

Four major types of data were compiled for this project: land cover, slope, transportation, and protected lands. Digital elevation model (DEM) data were obtained to represent the topography of the area. Different scale digital elevation data were also acquired from the National Digital Cartographic Data Base (NDCDB) distributed by the USGS. One-degree DEM coverage obtained from the NDCDB was originally produced by the Defense Mapping Agency. These data are generated by digitizing contour lines, spot elevations, and stream and ridge line data from the series of maps of scale 1:250 000. The data are converted to a regular array of elevations referenced horizontally on the UTM coordinate system.

Historical maps provided excellent source material for mapping change in the urban landscape. Paper maps were scanned and converted into digital images which were registered to the UTM grid for the region. Seven raster image maps of urban extent for the years 1850, 1900, 1940, 1954, 1962, 1974, and 1990 were generated (figure 1).
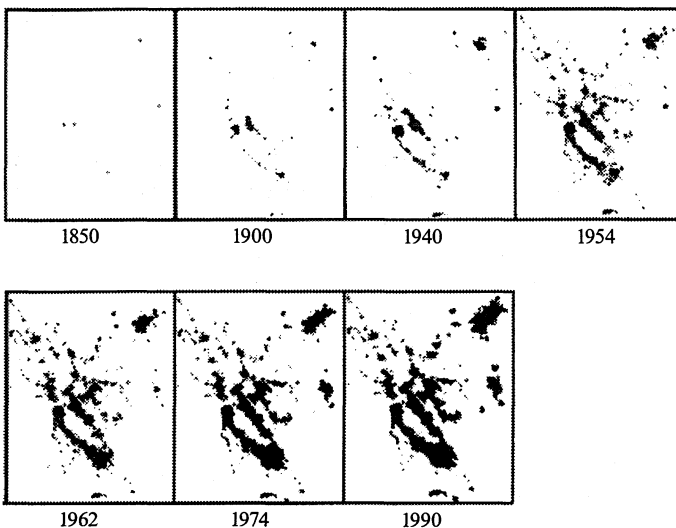


**Figure 1.** Urban growth in the San Francisco Bay area deduced from the human-induced land transformations database.

By temporal interpolation between these years, in a process termed geodynamic mapping, the data were used to build a highly popularized animation of urban growth (Petit, 1994). The time series of the evolving urban landscape provides a useful medium for visualizing change and investigating related spatial consequences.

Transportation and land-cover data sources included maps from the 1979 Atlas of California and maps from the Association of Bay Area Governments (ABAG). Historical map sources were primarily USGS topographic maps of scale 1:62500 and Army Map Service maps of scale 1:24000. Urban tints on the topographic maps were used to delineate historical urban areas. A digital database of transportation routes over time was derived from National Atlas digital line graphs (DLG) of scale 1:2000000. The DLG data were integrated with historical maps of highway development from 1920 to 1978. This information was found in the Atlas of California based on source maps of the Department of Transportation and the Highway Department of California (Donley et al, 1979). The DLG data were converted to a raster grid and edited manually to create time-series images of the major roads in the San Francisco metropolitan area, which generally reflect the physiography of the region (figure 2). A system of north–south interstate highways follows the central valley, rings the San Francisco Bay, and parallels the coast. Two major east–west corridors connect Sacramento and the central valley with the communities around the bay.
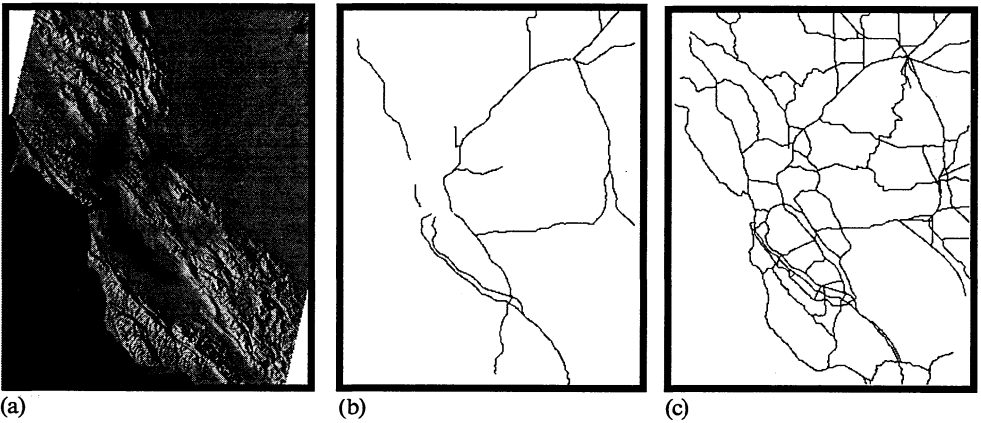


**Figure 2.** The San Francisco Bay area from the human-induced land transformations database: (a) topography; (b) roads in 1920; (c) roads in 1978.

Landsat remotely sensed data provided the most current, spatially continuous, and consistent coverage. Landsat-derived urban boundaries were made by manual photo-interpretation on graphics workstations. Digital landsat satellite data were used to create maps of urban extent for the years 1974 and 1990. Two landsat multispectral scanner scenes acquired in 1974 and two thematic mapper scenes acquired in 1990 were digitally mosaicked to provide coverage of the region. These data have provided a powerful reference base for describing and understanding urban growth in the study area, especially when compared with previous geographical work (Vance Jr, 1964).

**The cellular automaton model**

A cellular automaton model was developed for this study to investigate its utility in constructing scenarios of future urban land transformations. The model uses a basic grid of 300-m cells for the San Francisco Bay area. A set of initial conditions is defined by 'seed' cells which were determined by locating and dating the founding of various settlements identified from historical maps, atlases, and other sources. A summary of
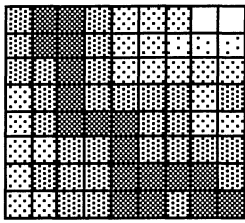
*Slope layer.* The data in this layer was interpolated from a digital elevation model and converted to a slope for every cell. It is used to determine the slope-resistance weighting.
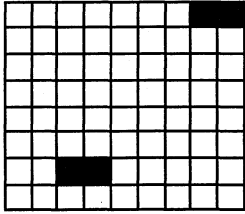


*Excluded areas layer.* Cells entirely exempt from the growth process, including oceans, lakes, and protected areas such as national parks and wetlands.
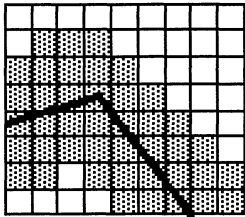


*Roads layer.* A binary array describing roads for a given era (read in when the time is reached) and a buffer whose width is determined by the road gravity, control factor, and defining the road attractiveness for development.
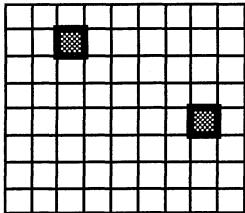


*Seed layer.* The initial distribution of 'urban' areas that act as growth centers. The seed layer can be any distribution, taken either from an actual time period or from a hypothetical starting distribution.

**Figure 3.** Input data layers for the urban growth model.

the input data layers, all derived from the HILT database, used in the model is shown in figure 3. A set of complex behavior rules was developed that involves selecting a location at random, investigating the spatial properties of the neighboring cells (for example, whether or not they are already urban, what their slope is, how close they are to a road, etc) and urbanizing the cell or not, depending on a set of probabilities (weighted by other locational characteristics). This was tested against a pseudo random number generated by the program. The behavior rules are summarized graphically in figure 4.

Five factors control the behavior of the system. These are: a DIFFUSION factor which determines the overall dispersiveness of the distribution both of single grid cells and in the movement of new settlements outward through the road system; a BREED coefficient which determines how likely a newly generated detached settlement is to begin its own growth cycle; a SPREAD coefficient which controls how much normal outward 'organic' expansion takes place within the system; a SLOPE_RESISTANCE factor which influences the likelihood of settlement extending up steeper slopes; and a ROAD_GRAVITY factor which has the effect of attracting new settlements onto the existing road system if they fall within a given distance of a road. These factor values, which affect the acceptance level of randomly drawn numbers, were set by the user for every model run and were varied as part of the calibration process. The values for DIFFUSION, BREED, SPREAD, and SLOPE_RESISTANCE range from 0–100, and ROAD_GRAVITY ranges from 0–20.
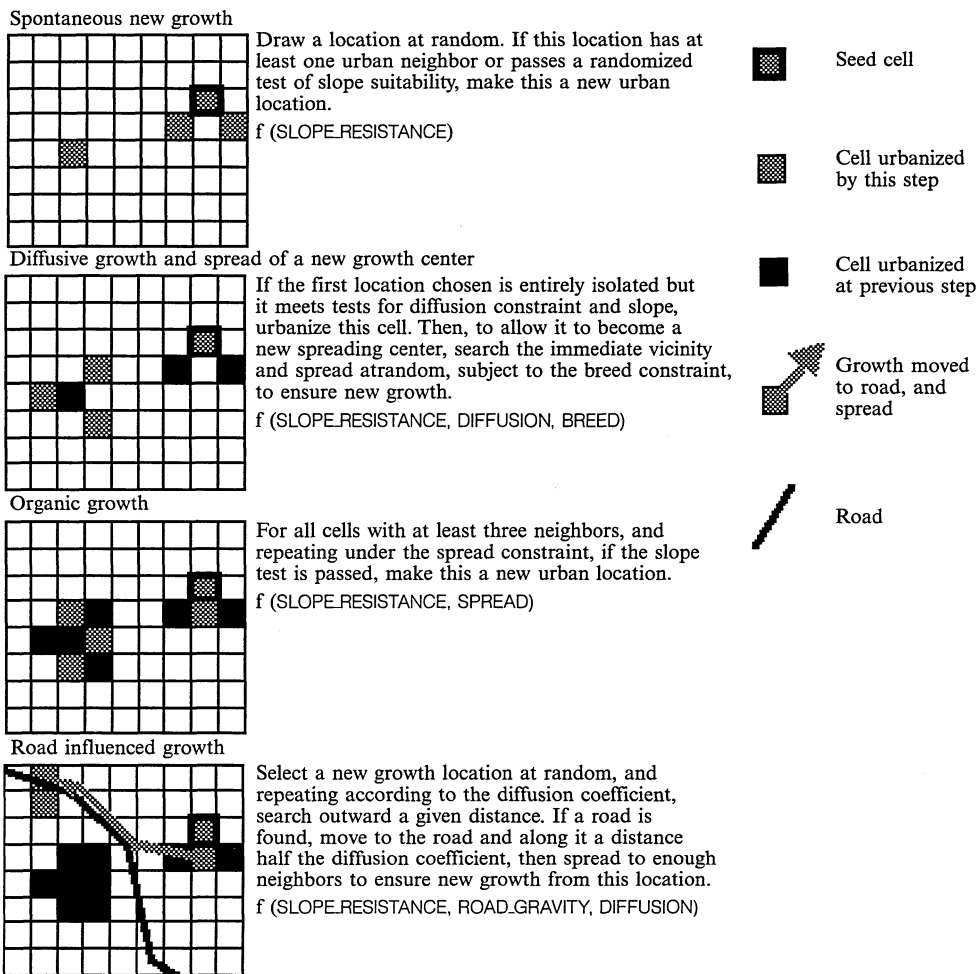
**Spontaneous new growth**

Draw a location at random. If this location has at least one urban neighbor or passes a randomized test of slope suitability, make this a new urban location.

f (SLOPE_RESISTANCE)

**Diffusive growth and spread of a new growth center**

If the first location chosen is entirely isolated but it meets tests for diffusion constraint and slope, urbanize this cell. Then, to allow it to become a new spreading center, search the immediate vicinity and spread atrandom, subject to the breed constraint, to ensure new growth.

f (SLOPE_RESISTANCE, DIFFUSION, BREED)

**Organic growth**

For all cells with at least three neighbors, and repeating under the spread constraint, if the slope test is passed, make this a new urban location.

f (SLOPE_RESISTANCE, SPREAD)

**Road influenced growth**

Select a new growth location at random, and repeating according to the diffusion coefficient, search outward a given distance. If a road is found, move to the road and along it a distance half the diffusion coefficient, then spread to enough neighbors to ensure new growth from this location.

f (SLOPE_RESISTANCE, ROAD_GRAVITY, DIFFUSION)

Legend:
- Seed cell
- Cell urbanized by this step
- Cell urbanized at previous step
- Growth moved to road, and spread
- Road

**Figure 4.** Model operation for a single cycle (year).

Other factors treated as constants after an initial calibration are the upper limit of what is considered high growth and the lower limit that determines low growth. The upper limit is passed when total urban growth for a year is greater than a preset number of pixels (hectares), causing an increase of 1.125 or 12.5% in the parameter values. An absolute value rather than a rate was found most suitable for the upper limit. The lower limit of growth is defined by the overall urban growth rate, causing the parameter values to decrease by 0.9 or 90%.

The growth rate is the sum of the four different types of urban growth defined in the model: spontaneous, diffusive, organic, and road influenced. Spontaneous urban growth occurs when a randomly chosen cell falls close enough to an urbanized cell, simulating the influence of urban areas on their surroundings. Diffusive growth urbanizes cells which are flat enough to be desirable locations for development, even if they do not lie near an already established urban area. Organic growth spreads outward from existing urban centers, representing the tendency of cities to expand. Road-influenced growth encourages urbanized cells to develop along the road network; the accessibility of these locations attracts development. The most prevalent type of urban growth during a model run is organic, followed by spontaneous growth. As road layers from several historical periods are read in at the correct time, road-influenced growth increases.

The model itself consists of a C-language computer program written by the first author. Writing, testing, and calibration of the model took place on Silicon Graphics and Sun computer equipment. Random number calls were to the standard C math library function rand(), seeded by means of the process identification number for the UNIX process invoking the program, multiplied by a counter for every iteration. Thus every iteration of the model was unique, although forced repetition of identical simulations could be generated when required by selecting a known random number seed.

Operation of the program is illustrated in figure 5. An outer control loop repeatedly executes each growth 'history', retaining statistical and cumulative data for the Monte Carlo application. An inner loop executes the CA, with each application cycle processing the whole layer once and considered equivalent to one year or one time cycle. Cycles begin with the 'seed' distribution of the actual settlement pattern in 1900, and the cellular rules are applied forward. The urban extent for 1850 has so few pixels that calibration based on this seed would be arbitrary. When a control year is reached for which actual data are available, the program computes and saves descriptive statistics. Status images can be created and saved for display at any time.

```
Read data layers
Initialize random numbers & control parameters
For n iterations {
        For t time periods {
                Apply change rules
                Apply self modification rules
                Compute and save descriptive statistics
                }
}
Write images
```

**Figure 5.** Sequence of operations for the simulation program.

## Model extension: self-modification of the cellular automaton

The final stage of model development was to define an additional set of rules by coupling the variables, thus allowing the model to modify itself. First, when the absolute amount of growth in any year exceeds a critical value, the DIFFUSION, SPREAD, and BREED factors are increased by a multiplier greater than one. This encourages diffusive, organic, and road-influenced growth, reproducing the tendency of an expanding system to grow ever more rapidly. However, to prevent uncontrolled exponential growth as the system increases in overall size, the multiplier applied to the factors is decreased slightly in every growth year. Second, when the system growth rate falls below another critical value, the DIFFUSION, SPREAD, and BREED factors are decreased by a multiplier less than one. This causes growth to taper off, just as it does in a depressed or saturated system. Third, the ROAD_GRAVITY factor is increased as the road network enlarges, prompting a wider band of urbanization around the roads. Fourth, as the percentage of land available for development decreases, the SLOPE_RESISTANCE factor is increased, allowing expansion onto steeper slopes. Additionally, when new growth in a time cycle takes place on steeper slopes, the spread factor is increased which accelerates urban expansion on flat land.

The self-modification rules, summarized in figure 6, allow much control of the system from only two factors: a 'critical high' growth rate and a 'critical low' growth rate. The system was retested for a complete range of these values and the outcomes examined. A long series of calibration trials of the control factors and the self-modification rules resulted in a stable and operational model. Some of the factors are more system sensitive than others and, in a system this complicated, a set of
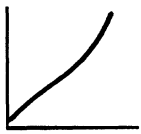
Rapid growth: greater than 'critical' number of hectares per year
DIFFUSION is multiplied by a constant >1.0
SPREAD is multiplied by a constant >1.0
BREED is multiplied by a constant >1.0
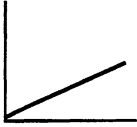
Normal growth: between rapid and little or no growth
If average slope > 10%, increase SPREAD
ROAD_GRAVITY increases by percent of road network
SLOPE_RESISTANCE increases by 0.2 × percent urban land available
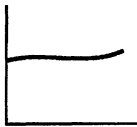
Little or no growth: annual growth rate is less than a critical value
DIFFUSION is multiplied by a constant <1.0
SPREAD is multiplied by a constant <1.0
BREED is multiplied by a constant <1.0

**Figure 6.** Self-modification adjustments to the control parameters.

interfactor dependencies exists. At the outset, a full set of outcomes can be generated by varying the parameters to extremes, for example, outcomes which result in zero and extensive growth can be simulated. Extensive growth patterns (that is, those that completely fill an area) which are linear, exponential, and S-curve type (that is, reaching and stabilizing at an 'optimum' population) can be simulated over time. Interactive and batch versions of the model were written, which allow for calibration, scenario construction, model replication, sensitivity analysis, and browsing of outcomes.

**Calibration of the model**
Statistical and graphical tests were used to calibrate the model. The visual tests were most useful in the initial phases to establish parameter ranges and to make rough estimates of the parameter settings. The visual tests were a necessary step for verifying that the model was in fact replicating the spatial pattern and extent of historical growth, something that could not be determined by statistical tests alone. Once several sets of initial parameter settings had passed the visual tests, a graphics-free version of the model was used to make goodness-of-fit comparisons.

Visual comparison played a key role in the first phase of calibration and involved area, edge, and cluster analysis of urban areas, including a continuously updated set of circles drawn at the center of gravity of the urban distribution and with the same area as the current urban extent. Statistical tests consisted of computing Pearson's $r^2$ for three values: the urban area; the number of edge pixels; and the number of pixel clusters for the modeled and real distributions in the key years.

Calibration consisted of four steps. The first of these was validation. In this step, the model was allowed to run to completion for a single iteration with unit increments in the control parameters and with self-modification disabled. Most parameters varied from 0 to 100, necessitating 101 separate runs per variable for the control variables. In each case, all other control variables were held constant at intermediate levels. Each of the final images was then converted into a single frame in an animation on a Silicon Graphics workstation. This allowed verification of the fact that each control parameter had a unique and controllable impact on the outcomes. In every case this was so, although a few program bugs were detected and resolved at this stage also.

The animation was an excellent tool to verify the outcome and it was noticed that some of the variables had clear saturation levels beyond which increments had relatively little effect.

The second phase of calibration involved writing two versions of the program with a full set of graphical user-interface tools. The first was a prototype written with the Silicon Graphics graphics tools, which allowed easy animation and display of the resulting images. The second version was an X Windows system version suitable for any standard workstation environment. This version used the XView toolkit, so that the critical control parameters could be changed by moving a slider, and execution started and stopped as necessary. A very large number of model runs in this interactive environment provided a means of testing the interaction of basic control parameters with each other and for debugging the self-modification rules. In addition, a set of measures was explored to allow visual comparison between the actual and predicted distributions. Placing symbols on the animated maps showing the actual and predicted centers of gravity for the urban cells proved useful. The time-sequenced display of a circle with the same area as the predicted distribution was also used. This allowed rapid visualization of structural changes in the distribution.

Building upon this second phase, we completed a third-phase batch version of the model, without graphics. This version continued to compute the suite of statistical measures of the distribution but, instead of displaying them, wrote them into a set of files for analysis. The real data were also processed to extract the same set of statistics. These files were then read with statistics and spreadsheet programs and by an additional computer program that calculated correlations between the predicted and observed data. The strategy used was to make minor changes in the control variables and to record the improvements that resulted in the correlation between three critical measurements. These were (1) the total area converted to urban use, (2) the number of pixels defined as edges, that is, with nonurban cell neighbors, which was thought to be a good measure of the rural–urban fringe effect of dispersed distributions, and (3) the number of separate spreading centers or clusters (figure 7).

The calibrated version of the model successfully predicted the total area of urban extent for the San Francisco Bay area from 1900 to 1990, although the historical distribution of growth along the road network is actually less dense than is predicted by the model. The model was also successful at replicating the raggedness of urban edges and the number of independent urban areas from 1900 to 1974. After 1974 a change in the data source used to determine historical extent affected the calculation of these measures. Before 1974 digitized paper maps were used to establish urban extent, whereas in 1974 and 1990 remotely sensed images were used. The shapes on maps tend to have been generalized by the cartographer whereas satellite images have far more salt-and-pepper edges. The cluster measure was particularly sensitive to the changed data source because of the method of calculation. An algorithm was written which systematically eroded the urban array by removing edge cells that did not connect to other clusters. All clusters in the distribution were eroded onto single pixels and simply counting these remaining cells at the end of the process gave the number of clusters in the image.

The influence of self-modification on the parameter values is seen in figure 8. The parameter values, initialized by the user, increase most rapidly at the beginning of the growth cycle when there are still many cells available to become urbanized and while the growth rate exceeds the critical high. The parameters are decreased as urban density increases in the region and expansion levels off and while the growth rate drops consistently below the critical low.
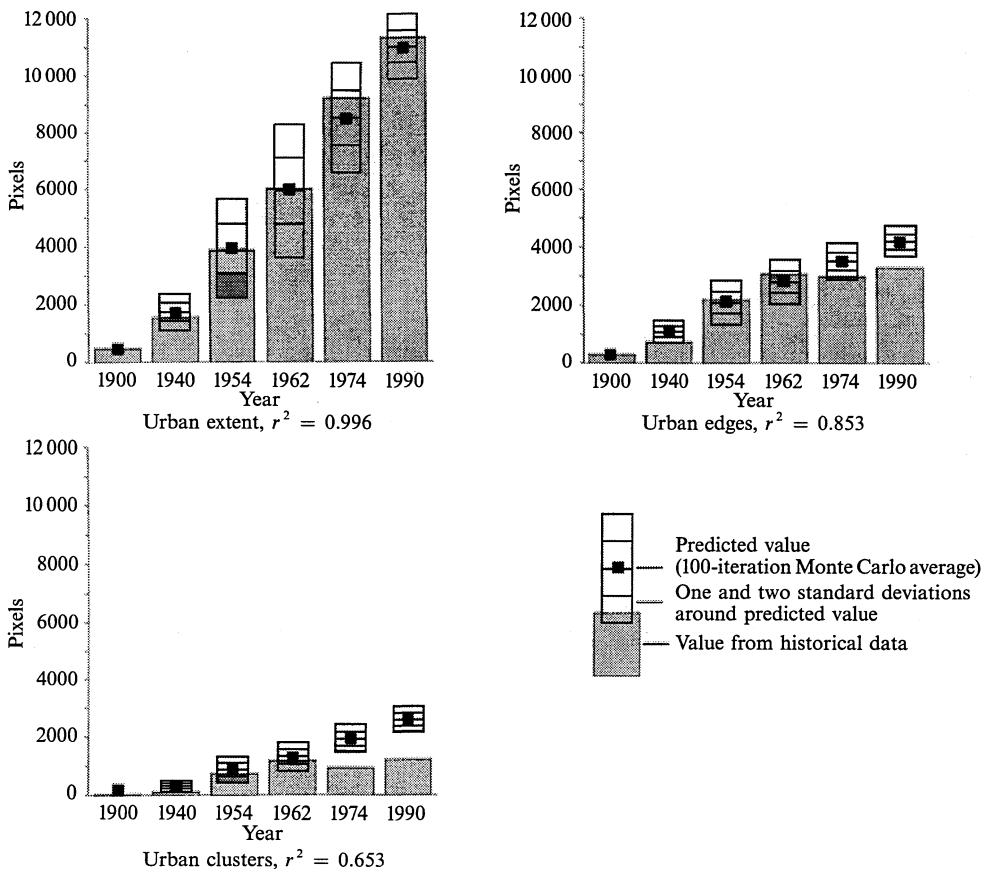
**Figure 7.** Comparison of historical growth with growth predicted by the calibrated model. The 90-year prediction of urbanization of the San Francisco Bay area is a Monte Carlo average of 100 iterations, based on year 1900 as seed. The discrepancy between actual and predicted number of edges and clusters for 1974 and 1990 reflects the switch to a different data source. (Note: $r^2$ is based on 5 base years.)
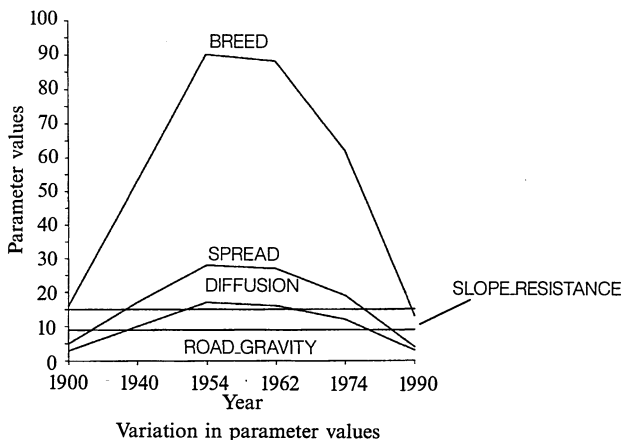


**Figure 8.** Diagram of average parameter values during the same run as for figure 7. The curves are produced by self-modification.

In the fourth and final phase of calibration, Monte Carlo averages of 100 iterations were used as the test statistic and the standard deviations of the predicted outcomes were computed. These variance measures allowed comparison of averages over many runs against observed values of the calibration statistics (figure 9). A final version of the program performed all permutations of the control parameters around the best settings, in each case maximizing the product of the $r^2$ values from the regression of average modeled versus observed. This allowed convergence on the final calibrated model.



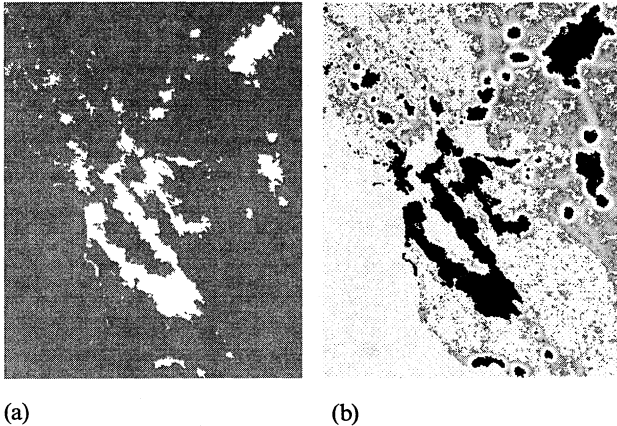(a)                                                            (b)

Figure 9. (a) Actual pattern of urbanization in the bay area, 1990 (white). (b) 100-iteration Monte Carlo image of predicted urban extent for year 1990 with the 1900 'seed' as starting point. White areas have a 50% to 80% probability of being urbanized; for black areas this probability is over 80%. Influence of the road distribution on the prediction is visible.

**Properties and features of the model**
The advantages of the model are many. The step rules are relatively simple to explain and understand. The model is not dependent on generalized probability distributions derived from observed or hypothetical data but allows each cell to act independently according to the rules (that is, every single part acts as part of an ensemble). This is similar to the way in which a city expands, as the result of hundreds of individual personal decisions, made one at a time but susceptible to the physical, social, economic, cultural, and political landscape (for example, the overall trends of the market, mortgage rates, economic climate, transportation technology, etc). One important feature of this model is its conduciveness to interactive and animated computer graphics, allowing point-and-click access to the parameters and immediate visualizations of the outcomes. Furthermore, multiple applications of the model from a variety of starting conditions allow the computation of Monte Carlo–style average aggregate output probabilities of any given cell being urbanized. The resultant maps or anticipated probability of future urbanization, although susceptible to the rules and properties of the model, are extremely useful tools for investigation of urban land transformations in a regional context as part of global change research. Different scenarios for urban outcomes can be linked to simple environmental models and the resultant environmental effects (for example, urban heat islands, loss of other land uses, increased particulate and gas emissions, etc) can be explored more effectively.

Wolfram (1984) has argued that cellular models are predictable in that they eventually converge on a finite set of outcomes, regardless of the initial start conditions. This is termed 'universality' in cellular modeling. Wolfram's outcomes were of three

major types. First, some outcomes are determined. These can take the form of either complete independence of the initial distribution or complete dependence locally. In the urban growth context, one final outcome of the first type is that every nonexcluded cell in the area becomes urban. The second type might imply that growth within a valley enclosed by mountains will take place only if an initial urban area falls within this valley.

The second type of outcome is when the value at any site depends upon the value at an increasing number of other sites. At first, with only a few iterations of the rules, the spatial impact of the rules is local. Later it is universal and very complex at a single site. Such a system, when it involves randomization, always results in chaotic behavior. The strong links between fractal theory and cellular models are then evident, and the cities should show extreme variation between multiple iterations in the Monte Carlo sense, and there seems good evidence that a chaotic model fits the behavioral patterns of rapid urban growth very closely. The statistics derived in this calibration indeed bear this out. Wolfram stated that in such a system the value of a single cell under this type of behavior can be determined by an algorithm. Nevertheless the massive numbers of computations necessary for such a solution seem impossible. A possibility would be to treat cells as individuals interacting with other objects and learning from the outcomes over time. Neural network methods could be used productively in this type of modeling.

The third (Wolfram's fourth) type of behavior for a CA is that the system is beyond prediction by algorithmic means. In such a system, simulation is the only way to predict outcomes. We have chosen to determine the value probabilistically and, as a result, all certainty in the model is de facto eliminated. Consequently, the model in use here can be regarded as a planning tool. It allows the formulation of probabilities of outcomes given starting scenarios based on large numbers of trials. In the real world, of course, there is only one sequence of time. An individual outcome therefore has the potential of being radically different from the expected outcome. The variance measures presented above should therefore be viewed as warnings not to take the probabilities too much to heart. Nevertheless, though self-modification probably increases the range of possible outcomes, its aggregate effect should be damping, that is, moving toward a finite set of outcomes. The Monte Carlo version of the model allows at least a statistical estimate of these possible outcomes to be delineated and mapped. The applications of these maps extend beyond the purposes for which they were compiled here.

Wolfram (1984) nevertheless showed that such simulation modeling is useful. He stated that:

> "This universality implies that many details of the construction of a cellular automaton are irrelevant in determining its quantitative behavior. Thus complex physical and biological systems may lie in the same universality classes as the idealized models provided by cellular automata. Knowledge of cellular automaton behavior may then yield rather general results on the behavior of complex natural systems" (page 1).

Given the failure of prior approaches to understand the urban transition at anything other than a coarse aggregate level, CA modeling does indeed offer the promise of a new approach.

## Conclusion

In this paper we have reported on our initial efforts to build and calibrate a predictive model of urban expansion, part of the HILT study. The search for an effective model has led to the use of a CA as a tool and the extension of the more traditional CA model into a self-modifying CA. Some of the problems of calibration which result

from a self-modifying CA have been researched and discussed. The potential for future work with this type of model and the ability to link external parameters to the self-modification, such as an economic growth rate or global average temperatures, offer some new avenues for future research.

Three different strategies have been employed: animation, description, and prediction. Animation has focused attention on significant changes in urban extent within a region; description has helped to identify and to understand natural and human-induced factors which influence the landscape at a regional scale; and prediction of future regional landscapes by modeling urban land transformations has been accomplished through cellular modeling. Animation has proved to be an effective means of visualizing and communicating the extent of change in an urban area over time as well as a useful format for representing a time series of urban land transformations derived from model output (Gaydos et al, 1996). It has also proved to be an invaluable tool in model calibration.

On the basis of the behavior of the model, urbanization is most likely to occur around the edges or in the vicinity of already established urban centers. After cities, roads have the next most important influence on the location of newly urbanized areas. Hilly terrain that does not lie near a city or a road has a very small chance of becoming urbanized. Overall, the model was successful at replicating urban expansion from 1900 to 1990 in the San Francisco Bay area. Although this region contains urban development on steep terrain, most notably San Francisco, whereas the model favors urbanization of flat areas before steep slopes, the discrepancy is probably not significant because of the small areal extent of this urbanization.

The ability of the model to adjust to its conditions was essential for modeling urban expansion. The growth rate in a traditional CA is limited to a linear or exponential curve, whereas a self-modifying CA permits the modeler to shape and manipulate complex curves. Urban growth takes the form of an S-curve in this region, rapid growth followed by a leveling off, and would not have been adequately represented by a traditional CA.

In the next phase of the project, the model will be used to produce three predictions of growth in the San Francisco Bay area. The simulations will be animated, along with the historical data, as a tool for visualizing these three scenarios: uncontrolled rapid growth, sustained slow growth, and a growth which stabilizes at a desirable, perhaps 'sustainable', level for the bay area. In addition, data collection is now under way to allow the model to be applied to the Washington, DC–Baltimore metropolitan area as a test of the robustness of the entire HILT methodology and approach in another region. Should this prove successful, it is hoped eventually to operate the predictive model at the national scale with the advanced very high resolution radiometer (AVHRR) land-cover data sets, so that America's possible urban futures and their environmental consequences can be visualized, anticipated, and perhaps even improved upon.

**References**
Anderson J R, Hardy E E, Roach J T, Witmer R E, 1976, "A land use and land cover classification system for use with remote sensor data", professional paper 964, United States Geological Survey, Eros Data Center, NASA–Ames Research Center, Moffett Field, CA
Batty M, Longley P, 1994 *Fractal Cities* (Academic Press, London)
Batty M, Xie Y, 1994, "From cells to cities" *Environment and Planning B: Planning and Design* **21** S31–S48

Couclelis H, 1985, "Cellular worlds: a framework for modeling micro–macro dynamics" *Environment and Planning A* **17** 585–596

Donley M, Allan S, Caro P, Patton C, 1979 *Atlas of California* (Academic Book Center, Culver City, CA

Eidenshink J C, 1992, "The 1990 conterminous US AVHRR data set" *Photogrammetric Engineering and Remote Sensing* **58** 809–813

Gaydos L J, Acevedo W, Bell C, 1995, "Using animated cartography to illustrate global change" *Proceedings of the 17th International Cartographic Conference* Barcelona, Institut Cartogràfic de Catalunya, Parc de Montjuic, 08038 Barcelona, Spain

Hägerstrand T, 1967 *Innovation Diffusion as a Spatial Process* (University of Chicago Press, Chicago, IL)

Hauser P N, Gardner R W, Laquian A A, El-Shakhs S, 1982 *Population and the Urban Future* (State University of New York Press, Albany, NY)

Jacobs J, 1961 *The Death and Life of Great American Cities* (Random House, New York)

Kirtland D, 1993, "Progress in global land data sets and land characterization: USGS research program" *World Resource Review* **5** 286–302

Kirtland D, DeCola L, Gaydos L, Acevedo W, Clarke K, Bell C, 1994, "An analysis of human-induced land transformations in the San Francisco Bay/Sacramento area" *World Resource Review* **6** 206–217

Knox P L (Ed.), 1993 *The Restless Urban Landscape* (Prentice-Hall, Englewood Cliffs, NJ) pp 1–13

Landis J, 1992, "BASS II: A new generation of metropolitan simulation models", WP-573, Institute of Urban and Regional Development, University of California at Berkeley, Berkeley, CA

Loveland T R, Merchant J, Ohlen D O, Brown J, 1991, "Development of a land cover characteristics data base for the conterminous US" *Photogrammetric Engineering and Remote Sensing* **57** 1453–1463

Petit C, 1994, "Thumbs up for video on Bay area growth" *San Francisco Chronicle* **65** A23

Pond B, Yeates M, 1994, "Rural/urban land conversion II: identifying land in transition to urban use" *Urban Geography* **15** 25–44

Portugali J, Benenson I, Omer I, 1997, "Spatial cognitive dissonance and sociospatial emergence in a self-organizing city" *Environment and Planning B: Planning and Design* **24** 263–285

Stegner W, 1954 *Beyond the Hundredth Meridian* (Houghton Mifflin, Boston, MA) page 438

Tobler W, 1979, "Cellular geography", in *Philosophy in Geography* Eds S Gale, G Olsson (D Reidel, Dordrecht) pp 379–386

USGS, 1994, http://geo.arc.nasa.gov/usgs/erosames.html, United States Geological Survey, Eros Data Center, NASA–Ames Research Center, Moffett Field, CA

Vance J E Jr, 1964 *Geography and Urban Evolution in the San Francisco Bay Area* Institute of Governmental Studies, University of California, Berkeley, CA

von Neumann J, 1966 *Theory of Self-Reproducing Automata* Ed. A Burks (University of Illinois Press, Champaign, IL)

White R, Engelen G, 1992a, "Cellular dynamics and GIS: modelling spatial complexity", WP-9263, Research Institute for Knowledge Systems (RIKS), Maastricht, The Netherlands

White R, Engelen G, 1992b, "Cellular automata and fractal urban form: a cellular modelling approach to the evolution of urban land use patterns", WP-9264, RIKS, Maastricht, The Netherlands

Wilson A, 1978, "Spatial interaction and settlement structure: toward an explicit Central Place Theory", in *Spatial Interaction, Theory and Planning Models* Eds A Karlqvist et al (North-Holland, Amsterdam) pp 137–156

Wolfram S, 1984, "Universality and complexity in cellular automata" *Physica D* **10** 1–35

Wolfram S, 1994, "Cellular automata", in *Cellular Automata and Complexity: Collected Papers* (Addison Wesley, Steven Wolfram, Reading, MA)

Wong D, Fotheringham A S, 1990, "Urban systems as examples of bounded chaos: exploring the relationship between fractal dimension, rank-size, and rural to urban migration" *Geografiska Annaler* **72B** 89–99