Environment and Planning B: Planning and Design 2005, volume 32, pages 231 – 246

DOI:10.1068/b31148

Diffusion and coalescence of the Houston Metropolitan Area: evidence supporting a new urban theory

Charles Dietzel
Department of Geography, University of California Santa Barbara, 3611 Ellison Hall, Santa Barbara, CA 93106, USA; e-mail: dietzel@geog.ucsb.edu
Hakan Oguz
Department of Forest Science, Texas A&M University, Forest Science Building, College Station, TX 77843, USA; e-mail: hakan_oguz@yahoo.com
Jeffery J Hemphill, Keith C Clarke, Nicholas Gazulis
Department of Geography, University of California Santa Barbara, 3611 Ellison Hall, Santa Barbara, CA 93106, USA; e-mail: jeff@geog.ucsb.edu, kclarke@geog.ucsb.edu, ngazu@umail.ucsb.edu
Received 10 May 2004; in revised form 6 December 2004

Abstract. The authors build on a recent development in urban geographic theory, providing evidence of an oscillatory behavior in spatiotemporal patterns of urban growth. With the aid of remotely sensed data, the spatial extent of urban areas in the Houston (USA) metropolitan region from 1974 to 2002 was analyzed by spatial metrics. Regularities in the spatial urban growth pattern were identified with temporal periods as short as thirty years by means of spatial metric values, including mean nearest-neighbor distance, mean patch area, total number of urban patches, and mean patch fractal dimension. Through changes in these values, a distinct oscillation between phases of diffusion and coalescence in urban growth was revealed. The results suggest that the hypothesized process of diffusion and coalescence may occur over shorter time periods than previously thought, and that the patterns are readily observable in real-world systems.

1 Introduction

The process of urban growth, from local sprawl to global urbanization, affects natural and human systems at all scales, and urban geography has attempted to create multiple models conceptualizing this process (for example, Burgess, 1925; Harris and Ullman, 1945; Hoyt, 1939). Although the fundamentals used to address urban systems have been widely recognized, they are based largely on social and economic theories that do not completely represent the spatiotemporal patterns of urban change (Batty, 2002). Borrowing methods from other disciplines, geographers have recently begun to use pattern analysis as a technique for analyzing the processes of urban growth and sprawl (Herold et al, 2003; 2002).

Basic theories of the form of individual cities and urban regions can be traced to Von Thünen's (1826) bid-rent theory, Burgess's (1925) concentric-zone model for Chicago, Christaller's (1933) geometrically driven central place theory, Lösch's (1938) similar work on economic regions, Hoyt's (1939) sector model, based on housing data, and a more general multiple-nucleus model by Harris and Ullman (1945). Zipf's (1949) work provided convincing evidence for a power-law relationship within entire urban systems with respect to population size. Although these theories and their variants have formed the foundation for subsequent work, they are predominantly descriptive models that assume cities grow in a uniform or linear manner. Most are not relevant to questions about interurban relations, or to the spatiotemporal dynamics of urban form; nor do they provide details about urban land-use change. However, most of these models share the concept of a nominal invariant surface over which the idealized form evolves, often called the 'isotropic plane', with a uniformly distributed population and no boundary effects.

Beyond the monocentric city or system of isolated cities, Gottmann (1961) first presented the idea of the coalescence of towns and cities into urban regions in his book *Megalopolis*, in which he empirically described an emerging urban corridor stretching from Washington DC to Boston. Megalopolis was driven by social and economic factors, and was not primarily a spatial model. Contemporary urbanization in the developed world, and indeed in the rapidly growing cities of the developing world, now seems dominated by this agglomeration rather than by the frontier spread of new settlements—even in areas such as Amazonia. Therefore, recent work has been more centered on the dynamics of form, and on the linking of form with processes driven by factors that originate at multiple scales within the urban system. An analysis of urban growth that provides some means of accounting for the 'path dependency of system evolution', stemming from influential factors such as the initial conditions, or distortions caused by random events and nonuniform landscapes, would be a valuable research contribution (Verburg et al, in press).

Batty and Longley (1994) considered urban growth as a cellular fractal stochastic process, which had already been described in physics as diffusion-limited aggregation (DLA). DLA yields spatial forms that are the result of growth by constrained diffusion. In this model, particles (or urban areas) spread randomly outward from a center and may find locales for establishment at the edge of the current form, resulting in a dendritic pattern of spread—with seeds growing outward like tentacles from an initial center. This model was based on the cellular framework outlined by Tobler (1979), and has provided the basis for growth patterns found in several urban models (Maske et al, 1995). Inherent in the DLA approach developed by Batty and Longley (1994) are the two spatial forms/processes that drive growth: diffusion and coalescence. Diffusion is defined as a process in which particles disperse, moving from regions of higher density to regions of lower density. In the spatial growth context, of course, no actual motion takes place but new urban areas are dispersed from the origin point or 'seed' location. Coalescence is the union of individual entities into one body, form, or group; or the growing together of parts. The 'growing together of parts' definition most suits the type of coalescence exhibited by an expanding urban area. In this paper we explore the complex interaction between these processes in the context of urban growth.

Geographers have long examined similar processes, in innovation diffusion (Hägerstrand, 1967), and the propagation of contagious disease (Gould, 1993). Also, a component of the DLA is scale sensitivity. Batty and Longley (1994) made a sound conceptual link between the spatial spreading process and fractal forms: that is, a form that behaves in the same way statistically at a range of spatial scales. In models of dynamic urban processes, especially those involving cellular automata, it has generally been assumed that forms that are equivalent across scale are generated. In these models change is treated linearly, with the resulting pattern being the consequence of an equilibrium that cannot be extended over long time periods (Verburg et al, in press).

Rarely has the scaling of model applications been subjected to empirical testing. Scaling relationships are deeply embedded in the classical theories of Zipf, Christaller, and Lösch, but scaling relationships are essentially only emergent properties in cellular models. Urban modeling and theories of urban dynamics have been used to address a variety of spatial scales, including global and superregional urban networks, metropolitan agglomerations, as well as urban growth and land-use change within individual cities at the local level (Alberti, 1999; Alberti and Waddell, 2000). Given the rapidly increasing array of available data sources that portray urban extent, spatial urban dynamics can now be observed and mapped at almost any scale. For this reason it is important to develop a multiscale model that will support the study of cross-scale urban growth dynamics to address some of the modeling limitations first identified by Lee (1973; 1994). There is also a significant gap between models for processing remotely sensed images to yield data on urban land use, such as work based on Ridd's (1995) vegetation-impervious-surface-soil (VIS) framework, and theory for understanding urban growth. Our proposed model and the approach used in this paper fill the gap, and provide the foundation for future work in this area.

Knowledge about the operational scale(s) of urban form and process, and the interaction and parallelism among different scales, is poor. This line of research was theoretically touched on in the late 1960s and early 1970s, but has not been followed up since the revitalization of urban growth models (Bourne, 1971; Boyce, 1966; Guttenberg, 1964). Guttenberg (1964), for example, described the drivers of urbanization, which contribute to urban form. His early work hints at scaling relationships in urbanization, and he discussed the "gradual adjustment of the regional structure to a larger territorial scale" (pages 205 – 206). Changes in the spatial scale may strongly impact modeled interactions, and may result in an inappropriate representation of specific processes of interest and their impact on model results. For example, at the street scale, neighborhood social and economic factors are far more likely to influence change than they are at the citywide or regional metropolitan scale. Because of the fact that most model-based studies of urban growth follow the 'one scale, one extent' mantra that has guided research into urban growth dynamics for so long, the cross-scale dynamics of the urbanization process have not been directly addressed (Verburg et al, 2005). Most valuable for the future will be the ability to scale up one model from fine-scale data sources to coarser scales, allowing for the study of interactions at the regional, continental, and, eventually, global scale. Should this prove possible, such a model and the theoretical framework supporting it might indeed be called 'universal'. We suggest that the scaling process and the coalescence – diffusion relationship in urban dynamics are cyclical. We are not the first to suggest a harmonic repetition: an analogy to ocean waves was developed for describing the spatiotemporal characteristics of urban growth by Boyce (1966) and by Blumenfeld (1954). Borrowing wave concepts from physics to simulate the urbanization process was originally proposed as a way to compensate for the lack of consideration of dynamics in the urbanization process in prevalent theory (Batty and Longley, 1994).

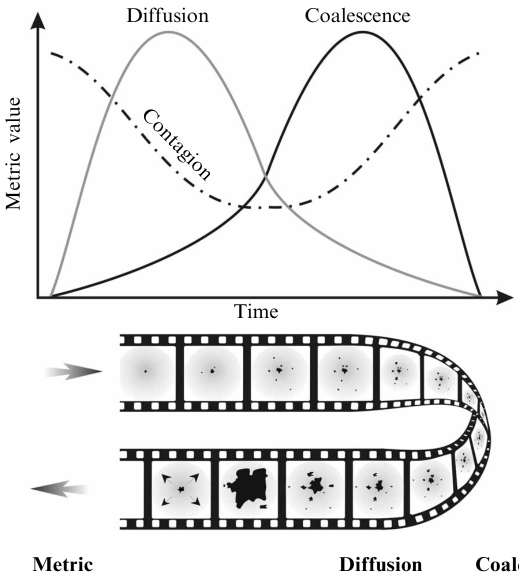
More recently, Dietzel et al (2005) have built on the work of Boyce (1966) and Blumenfeld (1954), proposing a more formal theory of spatiotemporal urban growth dynamics, which suggests that the process of urban growth can be characterized into two phases: diffusion and coalescence. This theory suggests that the processes are continuously observable, even after a landscape becomes completely urbanized, simply by scaling up to cover a greater spatial area. Initial work, in which time-series data from the Central Valley of California were used, showed that the processes of diffusion and coalescence were observable on a 100-year timescale. The lack of a sufficiently long time series of historical data led to the use of modeling results in the extension of the time span of analysis, formalization, and illustration of the theory; but the hypothesized patterns were not definitively observable in the data. These results fostered the current research. Assuming the theoretical basis for this theory of diffusion and coalescence is correct, the hypothesized patterns should be more readily observable in a rapidly urbanizing area over a shorter time span.

With the aid of spatial metric analysis, the spatiotemporal pattern of urban growth in the Houston metropolitan area of the United States was examined to determine if it exhibited any of the quantitative measures characteristic of diffusion and coalescence. The time span of the study was from 1974 to 2002. The study was based on remotely sensed data used in the application of the SLEUTH urban growth model (Clarke et al, 1997) to the Houston (Texas, USA) metropolitan area.

2 Urban diffusion and coalescence

The spatial evolution of cities can be described as a two-phase process of diffusion and coalescence (Dietzel et al, 2005). In the proposed model, the evolution of a city starts with the expansion of an urban seed, or core area. As this seed grows, it diffuses (grows) to new urban centers—or cores. As the process of diffusion continues, it is paralleled by organic growth which leads to expansion away from existing urban areas and the infilling of gaps in between them. This theoretical approach is different from a classical physics model of a diffusion process because, once established, no single zone 'moves' or deurbanizes. The model is more analogous to spilling a viscous liquid onto a surface, without evaporation: the liquid splashes outward and spreads at the same time.

As the urban system evolves, there comes a point at which the urban areas have become so diffuse that they begin to coalesce towards a saturated urban landscape. The full build-out of the urban landscape can also be seen as a seed urban area for the hypothesized model to evolve at a coarser spatial scale. This process of 'scaling up' is similar to the concept found in traditional urban studies, whereby the spatial extent is changed through the use of concentric rings, or increasing distances, around a central urban core or between urban centers (Blumenfeld, 1954; Luck and Wu, 2002). Batty and Longley (1994) made a similar assumption of self-similar scaling in their analysis of fractal cities.



Metric	Diffusion	Coalescence
Contagion	decreasing	increasing
Amount of urban land	increasing	increasing
Number of urban patches	increasing	decreasing
Nearest-neighbor distance	increasing	decreasing
Mean patch area	decreasing	increasing
Patch perimeter-to-area ratio	increasing	decreasing

Figure 1. The hypothetical wave patterns of the harmonic oscillation between the spatiotemporal process of urban diffusion and coalescence. The waves of diffusion are accompanied by declines in metric signatures for nearest-neighbor distance and number of urban patches. Conversely, coalescence is indicated by increases in the values of these metrics. Contagion is expected to decrease until the landscape becomes more urban than rural, at which point it increases again.

Through the use of spatial metrics, the hypothesized process of urban growth (figure 1) can be detected. Spatial metrics are quantitative measurements derived from digital categorical maps that quantify the spatial patterns and structure of a landscape at a specific scale and resolution. Calculation of these measures is based on a definitive, patch-based depiction of the landscape as developed for landscape ecology (Gustafson, 1998). Patches are homogeneous regions comprised of one category, such as 'urban', 'forest', or 'water'. This perspective involves the assumption that there are sudden spatial transitions between individual patches, which result in distinct edges with no gradual change between categories.

The spatiotemporal characteristics for a hypothetical cycle of urbanization and uniform isotropic growth at a fixed scale are shown in figure 1. The graph reflects the influence of diffusion in the early stages of urbanization. The heterogeneity of the landscape, described by the contagion metric, is hypothesized to be highest in the transitional period of development, when the system is switching from being dominated by diffusion to coalescence. As coalescence increases, the heterogeneity of the landscape decreases until it is completely urbanized. In the early stages of diffusion, the nearest-neighbor distances between individual urban patches are highest and decrease until more individual urban areas are distributed and a peak in urban patch density occurs. With the onset of coalescence, the decrease in the nearest-neighbor distances is less significant because nearby patches are the first to aggregate spatially. A high urban patch density is characteristic of the dominance of diffusion, and decreases once coalescence begins. During this time the difference between the total urban area and the amount of urban land in the urban core is highest because urban areas are the most spatially dispersed. The edge density peaks when the process of coalescence results in larger, heterogeneous urban agglomerations, and then decreases as the process moves towards the complete urbanization of the landscape.

This hypothesized process of spatiotemporal urban dynamics stresses that the spatial evolution of urban areas oscillates between diffusion and coalescence of individual urban areas in relation to the urban core. The patterns represent the dynamics at a defined spatial extent, but it could be expected that similar growth characteristics could be observed for varying spatial extents. The growth periodicity is expected to be longer with increasing distance from the central core, as has been suggested by Alonso (1964) and White et al (2001).

Twelve spatial metrics were used to evaluate the presence of the hypothesized process: number of patches; patch density; total number of edges; edge density; landscape-shape index; largest-patch index; mean patch area; perimeter-to-area fractal dimension; perimeter-to-area mean fractal dimension; mean patch fractal dimension; mean Euclidean nearest-neighbor distance; and contagion. It was believed that there are four metrics that identify the presence of harmonic urban dynamics through time: number of urban patches; mean patch size; patch density; and mean Euclidean nearest-neighbor distance—but the additional metrics were tested to see if there were others that could also be useful. Theoretically, the analysis based on the four metrics should result in the identification of two temporal waves that are typical of the hypothesized process (figure 1). These waves will have the same amplitude and wavelength, but will be offset by half of the periodicity. As urbanization spreads from an initial core, the number of patches should increase until a process of coalescence takes over, merging the patches back together. The process should then repeat itself, so that there is an oscillation between higher and lower numbers of urban patches through time—depending on whether the system is in a phase of diffusion or coalescence. Patch density should exhibit similar behavior: decreasing during periods of diffusion, and increasing during coalescence. These two metrics characterize the first wave. The second wave is evident in the mean patch size and Euclidean nearest-neighbor distance. These two metrics should exhibit similar trends, but function as a mirror image of the trends exhibited by the number of patches and patch density. As a system is undergoing diffusion (increase in the number of urban patches), the mean patch size will decrease, and then increase as the system coalesces. The same happens for the Euclidean nearest-neighbor distance: during the initial stages of diffusion, the nearest-neighbor distances between individual urban patches are highest, and decrease as diffusion occurs until they reach their minimum. This point represents the start of coalescence and the merging together of urban patches. The total amount of urbanized area is assumed to be increasing throughout this process.

Compared with previous work developing the theory of urban diffusion and coalescence, in this paper we research a much more rapidly growing region: the Houston (Texas) metropolitan area, over a thirty-year period. With the aid of the FRAGSTATS program (McGarigal et al, 2002), spatial metrics were calculated to derive the spatiotemporal patterns of urban growth at a fixed extent. The results suggest that the hypothesized process of diffusion and coalescence may occur over shorter time periods than was previously thought. Although the findings are just for one city, it is logical to suggest that the same processes may occur in other cities—but possibly at different spatiotemporal scales. The results presented provide strong empirical evidence that the theoretical patterns of urban growth hypothesized in Dietzel et al (2005), are real and are readily observable.

3 Data and methods

The Houston metropolitan region (figure 2) is one of the fastest growing regions in the United States: the fourth-largest city in the United States, the population of Houston grew by 25.8% between 1990 and 2000—well above the national growth rate of 13%

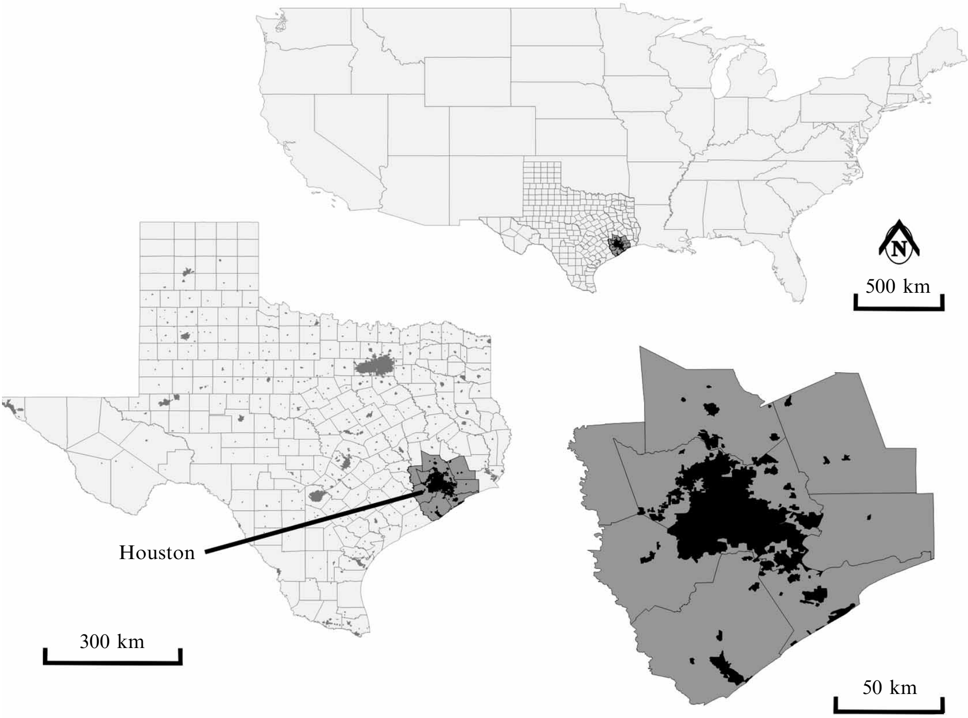


Figure 2. Location of Houston, Texas (USA).

(Clemonds and Liu, 2004). Data on the urban extent of this rapidly growing region were collected for 1974, 1984, 1992, and 2002, and used as input data for the SLEUTH urban growth model (Clarke et al, 1997) as part of a research initiative to forecast urbanization and land-use change for the region. The almost regular interval of observations for Houston made the data ideally suited for investigating the processes of urban diffusion and coalescence; and to test whether they occur in time periods of less than one hundred years in a rapidly growing region. As there was no precedent to determine the time scale necessary for determining if diffusion and coalescence were present, the use of four time periods allowed for the detection of one and a half complete cycles—if the hypothesized processes did occur. Requirement for the detection of one and a half cycles to confirm the hypothesized model of urban growth ensured that the results were not caused by noise around a constant trend in urbanization.

The Houston metropolitan region was defined as comprised of Waller, Montgomery, Fort Bend, Harris, Liberty, Brazoria, Galveston, and Chambers Counties. The urban extents for 1974 and 1984 were derived from four Landsat Multispectral Scanner (MSS) Triplicate scenes (25:39, 40 and 26:39, 30) with 60 m spatial resolution. The Iterative Self-Organizing Data Analysis Technique (ISODATA) clustering algorithm was used to perform an unsupervised classification of the images into urban/nonurban, with a convergence threshold of 0.95, and a maximum of thirty iterations. Once classified, the MSS scenes were mosaiced to create urban extent layers for the entire study area. Urban extent for 1992 was derived from the National Land Cover Dataset (NLCD) (http://landcover.usgs.gov/index.asp), with a spatial resolution of 30 m, and the land classified into twenty-one classes by the two-digit NLCD classification system. These twenty-one classes were then reclassified into six broader land-use categories (table 1).

Table 1. Initial input values from the 1992 National Land Cover Database (NLCD), and their reclassification into six land-use classes, from which urban extent for 1992 was extracted.

NLCD land-use class	NLCD classification values	Reclassified land-use class
Water	11 open water 12 perennial ice/snow	Water
Developed	21 low-intensity residential 22 high-intensity residential 23 commercial/industrial/transportation	urban
Barren	31 bare rock/sand/clay 32 quarries/strip mines/gravel pits 33 transitional	Other
Forest upland	41 deciduous forest 42 evergreen forest 43 mixed forest	Forest
Shrubland	51 shrubland	Agriculture
Nonnatural woody	61 orchards/vineyards/other	Agriculture
Herbaceous upland natural	71 grasslands/herbaceous	Agriculture
Herbaceous planted/cultivated	81 pasture/hay 82 row crops 83 small grains 84 fallow 85 urban/recreational grasses	Agriculture
Wetlands	91 Woody wetlands 92 emergent herbaceous wetlands	Wetlands

These included an urban class that was the aggregate of low-intensity residential, high-intensity residential, and commercial/industrial/transportation. To derive urban extent for 2002, a land-use data layer was compiled from three Landsat Enhanced Thematic Mapper (ETM) scenes (25:39, 40 and 26:39, 30), with a spatial resolution of 30 m. The images were classified into the six land-use classes in table 1 with the aid of the ISODATA unsupervised classification technique; they were then mosaiced together to form one image. Table 2 shows the overall accuracy assessment of the 2002 imagery classification, with the 2002 Houston – Galveston Area Council (http://www.h-gac.com) Land Use Land Cover Maps used as a reference. The overall classification accuracy was 87.33%, with a Kappa (Khat) coefficient of 0.82.

Table 2. Confusion matrix and kappa coefficient for the 2002 land-use/land-cover data from which urban extent for 2002 was derived.

Land-use class 2002 imagery classification	Reference data points (Houston Glaveston Area Council LULC 2002)								
	urban	agriculture	forest	water	wetland	other	row	producers' accuracy (%)	users' accuracy (%)
Urban	28	5	1	1	1	0	36	93.33	77.78
Agriculture	1	117	4	0	4	0	126	89.31	92.86
Forest	1	5	67	0	7	0	80	89.33	83.75
Water	0	0	1	35	0	0	36	97.22	97.22
Wetland	0	2	2	0	15	0	19	53.57	78.95
Other	0	2	0	0	1	0	3		
Total	30	131	75	36	28	0	300		

Overall classification accuracy = 87.33%
Kappa (Khat) coefficient = 0.82

Although the original use of these data was as input for the SLEUTH model, SLEUTH was not used in any capacity for the analysis presented here. Because of the processing time required for the calibration of the SLEUTH model, the spatial resolution of the data were resampled from 30 m to 100 m by means of the nearest-neighbor technique. This resampling method was chosen because it does not alter the cell values by averaging: instead, during resampling the nearest-cell value is assigned to the target cell. The use of the nearest-neighbor method may have introduced artifacts into the data—a modal filter may have been more appropriate. However, the data were initially gathered for another project in which sampling by the nearest-neighbor method was suitable. It was not thought that after processing these techniques would have a significant, if any, impact on the detection of the processes of urban diffusion and coalescence when the data were used to test the hypothesized model, as a 7×7 kernel-size majority filter was used to smooth the image slightly. The final image sizes were 1843×2100 (width $\times$ height). At 100 m resolution, each image of urban extent covered an area of 387 km^2, with a total of 3.87 million pixels. It is not thought that the use of the 7×7 filter had a significant impact on the results. If anything, the slight smoothing may have reduced the amount of diffusion and increased the coalescence observed.

Binary grids of urban/nonurban were derived from these input data for use in the spatial analysis program FRAGSTATS (McGarigal et al, 2002) (figure 3). Measures of the number of patches, patch density, total number of edges, edge density, landscape-shape index, large-patch index, mean patch area, perimeter-to-area fractal dimension, perimeter-to-area mean fractal division, mean patch fractal dimension, mean Euclidean

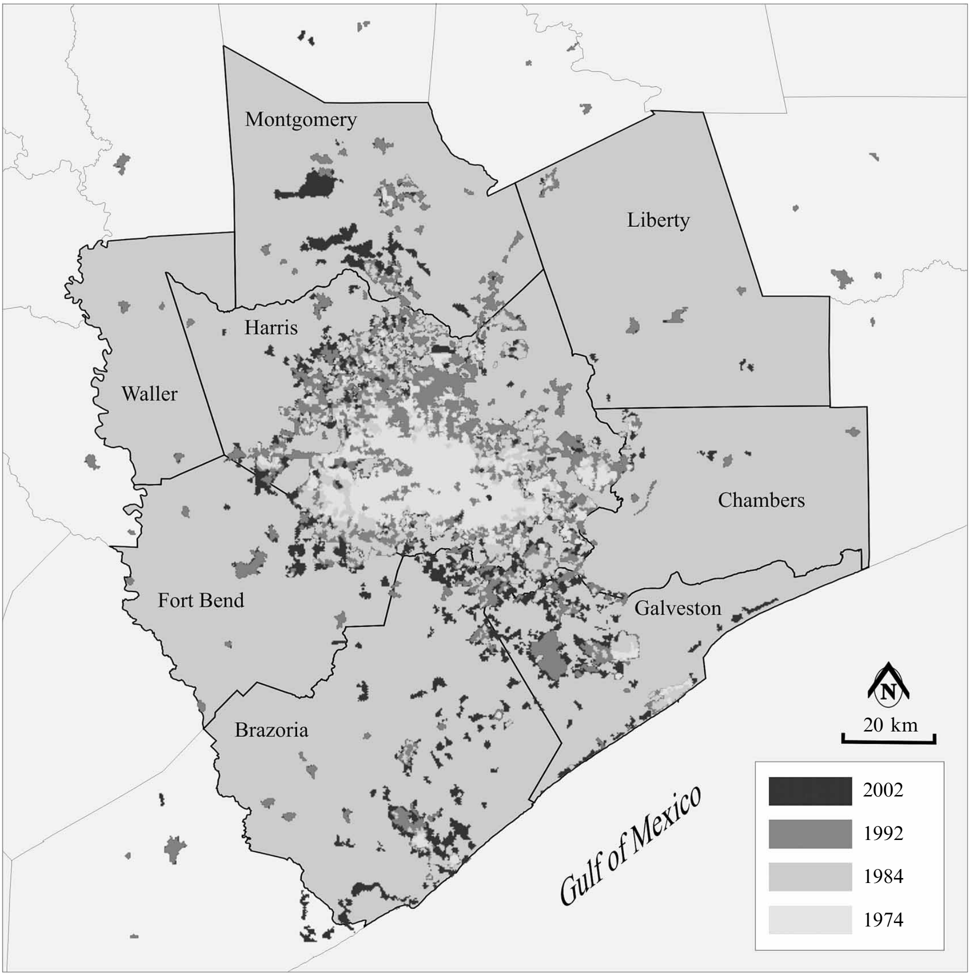


Figure 3. Counties of the Houston Metropolitan region, and the urban extent for 1974, 1984, 1992, and 2002.

nearest-neighbor distance, and contagion, were calculated for all years. The selection of metrics was based on those identified in previous research on spatial metric analysis of urban areas (Alberti and Waddell, 2000; Herold et al, 2003), and from the theoretical assumptions, although admittedly other metrics may be useful. Although some of these spatial metrics are related to or correlated with other metrics (for example, number of patches and patch density), our final conclusions were based on the mean Euclidean nearest-neighbor distance and number of patches—two metrics for which the calculation is not correlated. A more detailed description, including the mathematical equations, for all of the metrics can be found in McGarigal et al, 2002.

3.1 Landscape metrics

Landscape ecologists have played a critical role in the development of spatial metrics for the analysis of spatial patterns, such as deforestation and land-use change (O'Neill et al, 1988). When spatial metrics are used for landscape analysis, landscapes are viewed as a mosaic of patches. Spatial metrics can be used to quantify the spatial heterogeneity of individual patches, all patches in a class, and the landscape as a collection of patches. Some metrics are spatially nonexplict scalar values, but still capture important spatial properties—such as number of patches. Spatially explicit metrics can be computed as patch-based indices (for example, size, shape, edge length, patch density, fractal dimension) or as pixel-based indices (for example, contagion) computed for all pixels in a patch. Spatial metrics have proved an invaluable tool for measuring composition and spatial pattern at one or many points in time, allowing the examination of pattern processes occurring at various geographic and temporal scales.

Although four metrics (figure 1) were used for this research, twelve landscape metrics were calculated to determine if other metrics might aid in identifying stages of diffusion and coalescence. An eight-cell window was used to calculate the metrics, and the landscape was treated with a fixed geographic extent and spatial resolution was held constant. The twelve landscape metrics were

(1) Number of patches (NP)—total number of individual patches in a landscape. As urbanization diffuses throughout the landscape this number is expected to increase, and then decrease as the patches coalesce.

(2) Patch density (PD)—the number of patches in the landscape, divided by total landscape area. As diffusion increases, the number of urban patches within a fixed extent and their density should increase until coalescence starts, after which they should decrease in number.

(3) Total edges (TE)—the sum of the lengths (m) of all edge segments in the landscape.

(4) Edge density (ED)—the sum of all edge segments divided by the landscape area. This value should have a positive correlation with NP. As the landscape becomes more fragmented, the number of edges increases.

(5) Landscape-shape index (LSI)—total length of edge in the landscape, divided by the minimum total length of edge possible. This is a standard measure of total edge, or edge density, that includes adjustment for the extent of the landscape. LSI is an interpreted measure of patch aggregation that is expected to increase as the landscape becomes increasingly disaggregated.

(6) Largest-patch index (LPI)—the percentage of the landscape encompassed by the largest patch. As LPI approaches 100, the landscape is increasingly dominated by one patch.

(7) Mean patch area (AREA_MN)—the average size of the patches within a given landscape, calculated by dividing the sum of all patch areas by the number of patches; inversely correlated with NP. As the urban landscape becomes more fragmented, the number of patches will increase until the point in the cycle where patches begin to coalesce back together—increasing the mean patch area.

(8) Perimeter-to-area fractal dimension (PAFRAC)—reflects shape complexity across a range of spatial scales.

(9) Mean perimeter-to-area (PARA_MN)—the ratio of the patch perimeter (m) to area (m^2). This provides a simple measure of shape complexity, but without standardization to a simple Euclidean shape.

(10) Mean patch fractal dimension (FRAC_MN)—the mean fractal dimension of all individual patches within the landscape. Fractal dimension for each patch is calculated as 2 divided by the slope of regression line obtained by regressing the logarithm of patch area (in m^2) against the logarithm of patch perimeter (in m). The fractal dimension increases as the urban edge increases at a greater rate than the urban area. As edge growth proceeds, this metric will increase initially and there will be periods of decrease as core-area growth reaches the perimeter.

(11) Mean Euclidean nearest-neighbor distance (ENN_MN)—the average distance between two patches within a landscape (m), ENN_MN will decrease as patches grow together, and increase as there is diffusion and the urban areas expand.

(12) Contagion (CONTAG)—the negative value of the sum of the proportional abundance of each patch type multiplied by the proportion of adjacencies between cells of that patch type and another patch type, multiplied by the logarithm of the same quantity, summed over each unique adjacency type and each patch type; divided by 2× the logarithm of the number of patch types; multiplied by 100 (to convert to a percentage) (McGarigal et al, 2002). This metric is computationally complex and confusing; but it increases as urban patches become increasingly aggregated, and decreases as they become dispersed.

4 Observed spatiotemporal patterns

Spatial metric analysis of the urban extent for the Houston metropolitan area from 1974 to 2002 reveals the presence of harmonic spatiotemporal patterns in several of the spatial metrics, providing evidence in support of the hypothesized diffusion and coalescence phases of urban growth. The initial theory (Dietzel et al, 2005) suggested that patterns of diffusion and coalescence could be found by using four metrics (number of patches, patch density, Euclidean nearest-neighbor distance, and mean patch size). The authors (Dietzel et al, 2005) used a combination of cartographic sources, remotely sensed satellite imagery, and highly detailed aerial photography to map the urban extent in California's Central Valley from 1940 to 2000. With the aid of a similar approach, and FRAGSTAT, the spatiotemporal signature of only these spatial metrics were examined. Results presented in the present paper suggest that there may be other metrics that capture the oscillatory properties of urban dynamics.

The images of growth used discrete samples in time, and dates between these samples were not examined. This forced the assumption that there was a linear growth trend between the sample dates. Although this is not certain, it is nearly impossible to validate what the proper interval for investigating urban growth at a metropolitan scale is without capturing an exhaustive set of data. Because of the planning process in the United States, and the length of time that it takes for significant development to occur on a metropolitan scale, it was felt that the roughly ten-year interval of the data was appropriate for the testing of the hypothesized model of urban growth. Based on these assumptions, harmonic properties were found for seven of the twelve metrics (figure 4, over). As was outlined above, in the description of the hypothesized theory, measures of the number of patches, patch density, Euclidean nearest-neighbor distance, and mean patch size were believed to be best suited for capturing harmonic spatiotemporal properties. What was not expected was that metrics relating to the fractal dimension and perimeter-to-area ratio would exhibit similar properties (table 3, over). This was an interesting result in terms of the use of spatial metrics for the study of urban growth.

The number of urban patches in the Houston metropolitan region increased from 136 in 1974 to 447 in 1984. This was a period of diffusion, which was followed by coalescence between 1984 and 1992 at which time the number of patches had declined to 191. During the time period 1974–92, one cycle of diffusion and coalescence was completed. The time period 1992–2002 is the start of the next cycle, indicated by the increasing number of urban patches, from 191 to 323. Urban patch density is derived from the number of urban patches and exhibits the same spatiotemporal pattern. The oscillatory behavior of the number of urban patches and patch density should have a corresponding wave that is a mirror of the Euclidean nearest-neighbor distance and mean patch size. Mean patch size (and its standard deviation) decreased from 1974 to 1984, increased from 1984 to 1992, and then decreased again from 1992 to 2002. Euclidean nearest-neighbor distance (and its standard deviation) behaved in the same way as mean patch size: decreasing, increasing, and then decreasing again. The behavior of these four metrics, and the patterns in figure 5, are very similar to the hypothetical pattern shown in figure 1. Although the two waves are not exact mirrors of one another,

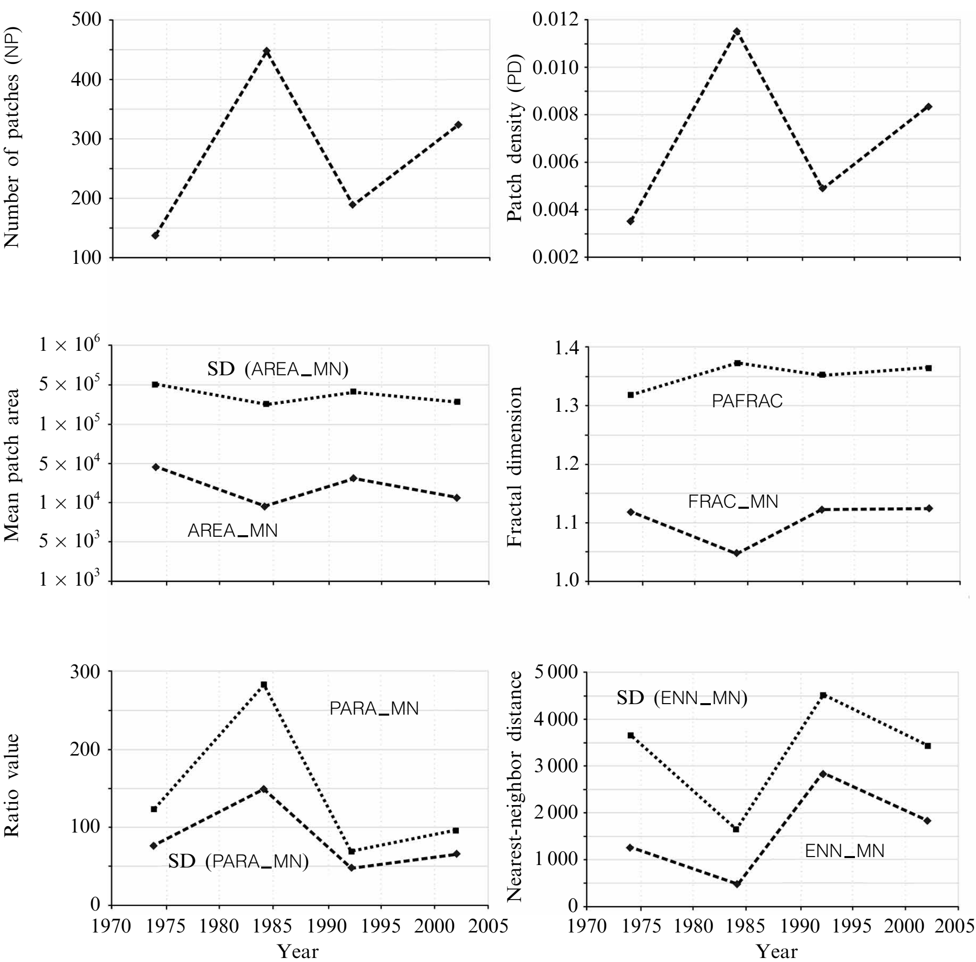


Figure 4. Plots of spatial metrics calculated based on the urban extent of the Houston metropolitan region, 1974 – 2002.

and their values differ in magnitude, they demonstrate the presence of a harmonic oscillation in the urban system between stages of diffusion and coalescence.

Perimeter-to-area fractal dimension (PAFRAC), mean perimeter-to-area ratio (PARA_MN), the standard deviation of PARA_MN and mean patch fractal dimension (FRAC_MN) were another four metrics that also exhibited harmonic spatiotemporal behavior; this was an unexpected result. As the amount of diffusion increased, so did PAFRAC—most likely because of an increase in the number of patches, which led to a more complex landscape. As coalescence 'filled in' the area between patches, the landscape became less complex and PAFRAC decreased. PARA_MN (and its standard deviation) increased during diffusion and decreased during coalescence. This was similar to PAFRAC, and the measures are clearly linked. As diffusion occurs within the system, the perimeter of urban patches grows at a more rapid rate than does patch area; this increase in perimeter leads to a more complex landscape shape—as indicated by the increase in PAFRAC dimension. FRAC_MN followed an opposite trend from PAFRAC and PARA_MN: it decreased with diffusion, and increased with coalescence. This suggests that, whereas diffusion was creating a more complex landscape pattern, the individual patches were simple in structure, and the coalescing of these simple shapes created larger

Table 3. Observed metric values from the application of FRAGSTATS to the urban extent data for Houston for 1974, 1984, 1992, and 2002.

Metrics	Year			
	1974	1984	1992	2002
Number of patches (NP)	136	447	191	323
Patch density (PD)	0.0035	0.0115	0.0049	0.0083
Total edges (TE)	2 651 800	4 970 000	6 052 800	8 831 700
Edge density (ED)	0.6852	1.2841	1.5639	2.2819
Landscape-shape index (LSI)	4.3715	7.3172	8.693	12.224
Largest-patch index (LPI)	97.5962	95.4138	91.7553	88.8785
Mean patch area (AREA_MN)	28 458.0882	8 658.3893	20 263.3508	11 982.3529
Standard deviation of AREA_MN	322 697.7087	174 525.0126	256 627.3459	191 676.83
Perimeter-to-area fractal dimension (PAFRAC)	1.318	1.373	1.3523	1.3637
Mean perimeter-to-area ratio (PARA_MN)	122.3779	281.9571	72.8406	97.28
Standard deviation of PARA_MN	78.5933	150.1409	48.514	67.1846
Mean patch fractal dimension (FRAC_MN)	1.1184	1.0478	1.1245	1.1252
Standard deviation of FRAC_MN	0.0385	0.0666	0.0314	0.0361
Mean Euclidean nearest-neighbor distance (ENN_MN)	1 291.8268	473.0953	2 818.0927	1 823.337
Standard deviation of ENN_MN	3 636.4484	1 653.1462	4 461.0563	3 432.3573

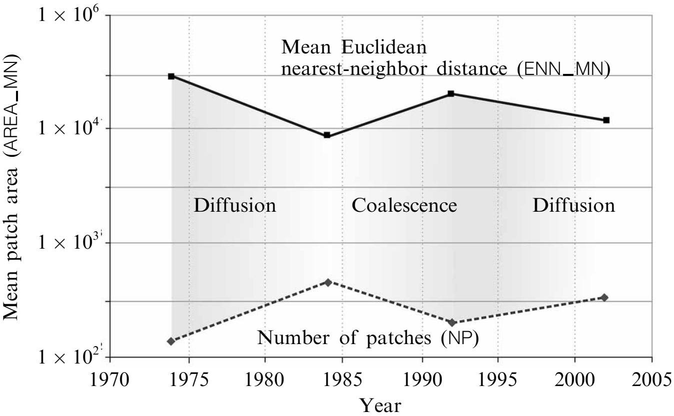


Figure 5. Plot of number of urban patches and mean patch area through time for the Houston metropolitan region, 1974 – 2002. As the number of patches increases, the patch area decreases—indicating that diffusion is taking place. When the number of patches decreases, the patch area increases as the patches merge together in the process of coalescence.

patches that were more complex than their predecessors. The harmonic behavior of these metrics was not expected, and was not thought to be significant when developing the theory of spatiotemporal urban harmonics, yet they seem to provide important information about urban evolution, and need to be refined and incorporated into the theory.

Total edges, edge density, landscape-shape index (LSI), and largest-patch index (LPI) did not exhibit any harmonic behavior. Total edges, edge density, and LSI increased through time; LPI decreased. As well as these four metrics, contagion also only decreased with time. Previous work suggests that contagion is a metric that would decrease with time until coalescence was greater than diffusion, at which point it would increase again. This prior work was done with a 100-year time scale, not the 28-year timeframe of the current work. It may be that the harmonic oscillation of the contagion metric can only be observed over a longer temporal period than the oscillations between the metrics of number of patches, patch density, Euclidean nearest-neighbor distance, and mean patch size. Similar ideas relating to this with regard to urban growth processes oscillating at different temporal scales have been suggested by earlier researchers (Cressy, 1939; Duncan et al, 1962; Hoover and Vernon, 1959; Winsborough, 1962), and help provide an explanation as to why the contagion metric for the Houston metropolitan area did not exhibit the oscillatory behavior that was observed with other metrics. The next step in this area of research should be to find other examples of this process occurring in urban systems, so that the theory can be more formalized with a broader set of examples. Formalization of the theory will lead to the development of a predictive model of urban dynamics, as opposed to the mostly descriptive models of Von Thünen (1826), Burgess (1925), Hoyt (1939), and Harris and Ullman (1945).

5 Furthering urban theory

Metric analysis of spatiotemporal urban growth data from the Houston metropolitan region provides empirical evidence that the processes of urban diffusion and coalescence occur in real-world systems, providing a link between theory and empirical evidence. The harmonic oscillation between stages of diffusion and coalescence is apparent in the temporal behavior of the spatial metrics used: specifically, the number of urban patches, patch density, Euclidean nearest-neighbor distance, and mean patch size (figure 5). The behavior of these metrics confirms the hypotheses presented by Dietzel et al (2005). In this previous work on this topic, the possibility that there were other metrics that might identify the presence of a spatiotemporal oscillation in urban growth was ignored; conclusions were based on nearest-neighbor distance, and patch and edge density. We can now conclude that there are other metrics, including perimeter-to-area fractal dimension, mean patch perimeter-to-area ratio, and mean patch fractal dimension, which can be helpful in detecting harmonic oscillation between phases of urban diffusion and coalescence.

Scaling up refers to a change in extent, whereby after one cycle of diffusion and coalescence it becomes necessary to increase the spatial extent of the study area to detect the next harmonic cycle. The results in this paper suggest that this may not be the case. Over the time period of the study, the data suggest that Houston has diffused, coalesced, and is in the process of diffusing again—yet the contagion metric is still decreasing. This has two implications. First, diffusion and coalescence may occur multiple times within a fixed spatial area until the contagion metric reaches a value of 50 (that is, half of the landscape is urban and half is nonurban). This is the time when scaling up would be more appropriate. At present it is still unclear what degree of scaling up is necessary to observe diffusion and coalescence best at the next spatial extent, and this provides ground for future work in theory development. Second, as was previously suggested, the harmonic oscillation of urban diffusion and coalescence, as reflected by the spatial metrics, occurs at different temporal scales. Metric values for the number of urban patches and Euclidean nearest-neighbor distance may exhibit harmonic trends on a shorter time scale than that of other metrics, such as contagion. This suggests that the processes of diffusion and coalescence are actually comprised of multiple waves, each with different periods.

Although this study has established a clear link between empirical measurements and the hypothesized theory of urban growth, there is no clear link between the processes that lead to the observed spatiotemporal patterns. The suggestion that the spatiotemporal patterns of urban diffusion and coalescence are comprised of multiple waves with different periods may be indicative of the interactions between multiple urban processes. The fact that urban growth is driven by local factors (that is, topography, transportation networks, policy, and initial conditions) may lead to discrepancies between observations and expected theoretical patterns in different cities. Differences are to be expected and may be found in the form of amplifications, lagging, or damping of the metric signatures. In the case of metric analysis, the initial conditions are not the true initial conditions—rather, they are the first observation recorded of an urban area—so it may be difficult to tell at what point in a cycle of diffusion and coalescence a developing urban area lies.

The results presented here provide supporting empirical evidence for a new theoretical framework that addresses the dynamics of urbanization. Evidence of urban diffusion and coalescence suggest that development is not just a diffusive process, in which development merely disperses outward from existing areas, but one followed by a temporal lag during which gaps in open space are filled in. In the development of this theory, one of the goals was to provide a means of improving the modeling of the spatiotemporal dynamics of urban growth. The next step will be to develop an experimental model to replicate these patterns. Analysis of the Houston metropolitan region has provided adequate information to begin the development of a general model of diffusion and coalescence. This type of model might be able to serve as a guide or reference for more accurate representations of dynamic spatial processes—something that is greatly needed if spatial models are to be taken seriously outside of the academic community.

References

Alberti M, 1999, "Modeling the urban ecosystem: a conceptual framework" *Environment and Planning B: Planning and Design* **26** 605 – 630

Alberti M, Waddell P, 2000, "An integrated urban development and ecological simulation model" *Integrated Assessment* **1** 215 – 227

Alonso W, 1964 *Location and Land Use* (Harvard University Press, Cambridge, MA)

Batty M, 2002, "Thinking about cities as spatial events" *Environmental and Planning B: Planning and Design* **29** 1 – 2

Batty M, Longley P, 1994 *Fractal Cities* (Academic Press, San Diego, CA)

Blumenfeld H, 1954, "Tidal wave of metropolitan expansion" *Journal of the American Institute of Planners* Winter, 3 – 14

Bourne L S, 1971 *Internal Structure of the City* (Oxford University Press, New York)

Boyce R R, 1966, "The edge of the metropolis: the wave theory analog approach" *British Columbia Geographic Series* **7** 31 – 40

Burgess E W, 1925, "The growth of the city: an introduction to a research project", in *The City* Eds R E Park, E W Burgess, R D McKenzie (Chicago University Press, Chicago, IL) pp 47 – 62

Christaller W, 1933 *Die zentralen Orte in Süddeutschland* (Jena); republished 1966 as *Central Places in Southern Germany* (Prentice-Hall, Englewood Cliffs, NJ)

Clarke K C, Hoppen S, Gaydos L, 1997, "A self-modifying cellular automaton model of historical urbanization in the San Francisco Bay area" *Environment and Planning B: Planning and Design* **24** 247 – 261

Clemonds M, Liu H, 2004, "Exploring urban heat island effects in the Houston metropolitan area using satellite remote sensing data", paper presented at the 100th Annual Meeting of the American Association of Geographers, Philadelphia, PA; copy available from the authors, Department of Geography, Texas A & M University, College Station, TX

Cressy P F, 1939, "Population succession in Chicago, 1898 – 1930" *American Journal of Sociology* **44** 59 – 69

Dietzel C, Herold M, Hemphill J J, Clarke K C, 2005, "Spatio-temporal dynamics in California's Central Valley: empirical links to urban theory" *International Journal of Geographic Information Science* **19** 175 – 195

Duncan B, Sabagh G, Van Arsdol M D, 1962, "Patterns of city growth" *American Journal of Sociology* **67** 418 – 429

Gottman J, 1961 *Megalopolis: The Urbanized Eastern Seaboard of the United States* (Twentieth Century Fund, New York)
Gould P, 1993 *The Slow Plague: A Geography of the AIDS Pandemic* (Blackwell, Oxford)
Gustafson E J, 1998, "Quantifying landscape spatial pattern: what is the state of the art?" *Ecosystems* **1** 143 – 156
Guttenburg A Z, 1964, "The tactical plan", in *Explorations into Urban Structure* Ed. M M Webber (University of Pennsylvania Press, Philadelphia, PA) pp 197 – 219
Hägerstrand T, 1967 *Innovation Diffusion as a Spatial Process* (University of Chicago Press, Chicago, IL)
Harris C, Ullman E, 1945, "The nature of cities" *Annals of the American Academy of Political and Social Science* **242** 7 – 17
Herold M, Scepan J, Clarke K C, 2002, "The use of remote sensing and landscape metrics to describe structures and changes in urban land uses" *Environment and Planning A* **34** 1443 – 1458
Herold M, Goldstein N C, Clarke K C, 2003, "The spatiotemporal form of urban growth: measurement, analysis and modeling" *Remote Sensing of Environment* **86** 286 – 302
Hoover E M, Vernon R, 1959 *Anatomy of a Metropolis: The Changing Distribution of People and Jobs within the New York Metropolitan Region* (Harvard University Press, Cambridge, MA)
Hoyt H, 1939 *The Structure of Growth of Residential Neighborhoods in American Cities* Federal Housing Authority, Washington, DC
Lee D B, 1973, "Requiem for large-scale models" *Journal of the American Institute of Planners* **39** 163 – 178
Lee D B, 1994, "Retrospective on large-scale urban models" *Journal of the American Planning Association* **60** 35 – 40
Lösch A, 1938, "The nature of economic regions" *Southern Economic Journal* **5** 71 – 78
Luck M, Wu J, 2002, 'A gradient analysis of urban landscape pattern: a case study from the Phoenix metropolitan region, Arizona, USA" *Landscape Ecology* **17** 327 – 339
McGarigal K, Cushman S A, Neel M C, Ene E, 2002 *FRAGSTATS: Spatial Pattern Analysis Program for Categorical Maps* http://www.umass.edu/landeco/research/fragstats/fragstats.html
Maske H A, Havlin S, Stanley H E, 1995, "Modeling urban growth patterns" *Nature* **377** 608 – 612
O'Neill R V, Krummel J R, Gardner R H, Sugihara G, Jackson B, Deangelis D L, Milne B T, Turner M G, Zygmunt B, Christensen S W, Dale V H, Graham R L, 1988, "Indices of landscape pattern" *Landscape Ecology* **1** 153 – 162
Ridd M, 1995, "Exploring a V-I-S (vegetation-impervious-surface-soil) model for urban ecosystem analysis through remote sensing: comparative anatomy for cities" *International Journal of Remote Sensing* **16** 2165 – 2185
Tobler W, 1979, "Cellular geography", in *Philosophy in Geography* Eds S Gale, G Olsson (D Riedel, Dordrecht) pp 379 – 386
Verburg P, Schot P, Dijst M, Veldkamp A, in press, "Land use change modelling: current practice and research priorities" *GeoJournal*
Von Thünen J H, 1826 *Die Isoliste Staat* [The isolated state] (Hamburg)
White R, Luo W, Hatna E, 2001, "Fractal structures in land use patterns of European cities: form and process", paper presented at the 12th European Colloquium on Quantitative and Theoretical Geography, 7 – 11 September, St-Valery-en-Caux, France; copy available from C Dietzel
Winsborough H H, 1962, "City growth and city structure" *Journal of Regional Science* **4** 35 – 50
Zipf G K, 1949 *Human Behavior and the Principle of Least Effort* (Addison-Wesley, New York)