USING SPACE/TIME TRANSFORMATIONS TO MAP URBANIZATION IN BALTIMORE/WASHINGTON

Lee De Cola, U.S. Geological Survey*
521 National Center, Reston VA 20192, ldecola@usgs.gov

ABSTRACT

During the past year researchers at the U.S. Geological Survey have been using historical maps and digital data for a 168-km × 220-km area of the Baltimore-/Washington region to produce a dynamic database that shows growth of the transportation system and built-up area for 270-meter grid cells for several years between 1792 and 1992. This paper presents results from the development of a *Mathematica* package that spatially generalizes and temporally interpolates these data to produce a smoothly varying urban intensity surface that shows important features of the 200-year urban process. The boxcount fractal dimension of a power-2 grid pyramid was used to determine the most appropriate level of spatial generalization. Temporal interpolation was then used to predict urban intensity for 4320-m cells for 10-year periods from 1800 to 1990. These estimations were spatially interpolated to produce a 1080-m grid field that is animated as a surface and as an isopleth (contour) map (see USGS 1997 for the Internet address of the animation). This technique can be used to experiment with future growth scenarios for the region, to map other kinds of land cover change, and even to visualize quite different spatial processes, such as habitat fragmentation due to climate change.

In 1994 a team of U.S. Geological Survey (USGS) and academic researchers produced an animation of the growth of the San Francisco/Sacramento region using a temporal database extracted from historical maps, USGS topographic maps, digital data, and Landsat imagery (Gaydos and Acevedo 1995). Publicly televised videotapes of this work received sufficient attention to support a larger team that had planned to work on the development the Boston/Washington megalopolis (Gottmann 1990) (The

---

*Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

current research involves staff from USGS, National Air and Space Administration, the Smithsonian Institution, and University of Maryland Baltimore County.) Resource and time constraints, however, limited efforts to the southern part of the region shown in figure 1 (Crawford-Tilley et al 1996, Clark et al. 1996). The animation of urbanization in this region is based on a 512^2-cell grid data structure that represents whether or not a given 270-meter cell is built-up in each of 8 base years (figure 2). This raster was interpolated for intervening years, but still represents a binary condition for each of the grid cells. Throughout this work there was interest in how we might analyze the intensity of development, perhaps by sacrificing spatial resolution for temporal and measurement resolution (table 1). Because the urban phenomenon (cartographic feature) is self-organized, complex, and probably also critical (Bak 1996), it is reasonable to suppose that scaling properties would assist in this transformation (Quattrochi and Goodchild 1997).

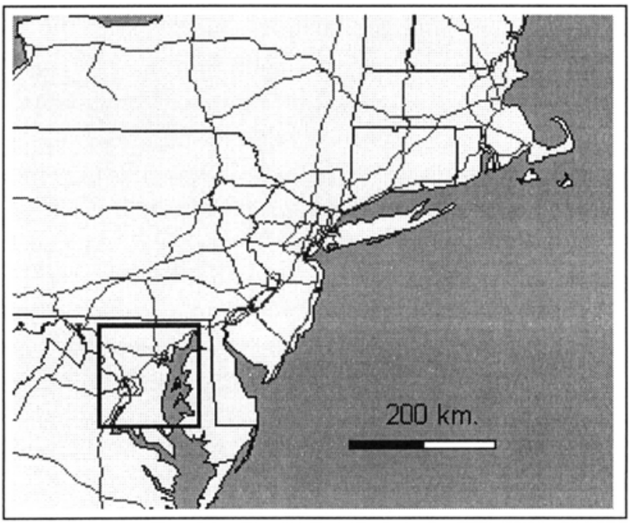

Figure 1. The study area in the Boston-Washington megalopolis.

Table 1. Dimensions of the data

DIMENSION		EXTENT	RESOLUTION l DATA	ANIMATION
SPACE x	Balt/Wash	138 km	270-meter grid	1080-meter
TIME t	1772-1992	200 years	~25 years	10 years
FEATURE f	Land cover	Built-up	[0, 1] binary	[0, 256]

Consider therefore a location in space x at a given time t and spatial resolution level l for which a measurement f is made; call this measurement $f_l(x_t)$. For example, in the present case we are interested in whether or not a given grid cell of a certain size is built-up (covered by buildings, has a dense road network, etc.). In this simplest case we have a binary function $f_l(x_t)$ = {1 if x_t is built-up, 0 otherwise}. Assume at the finest scale level $l = 0$ that this measurement is reliable—but what can be said of the phenomenon at other spatial scales? Table 2 shows how a 10-level power-2 image pyramid can be built upon the 0-level data in the present case. One (not necessarily obvious) way to examine data at coarser scales is simply $fbox_{l+1}(x_t)$ = {0 if all $f_l(x_t) = 0$, 1 otherwise}, i.e. the value of a higher-level $l + 1$ cell will be "on" if any of the lower-level l cells is on. This is called a box-covering algorithm because a high-level box is needed to cover 1 or more lower-level boxes (De Cola 1997). Consider the 0-level image of figure 2, which contains 41,183 built-up cells, as reported in the last row of table 3, which presents the box counts for each level and each of the raw data years The table shows at the next highest level $l = 1$ that 14,892 cells are necessary to cover these cells. This number is 45% larger than the (41,183 / 4 =) 10,296 level-1 cells that would be necessary if all the level-0 cells were spatially compact. The excess number is due to spatial complexity of the urban phenomenon, which has fractal dimension $D < 2$, where $D = 2$ would be the dimension of say a perfect disk (for a comprehensive discussion of the fractal nature of cities see Batty 1995).

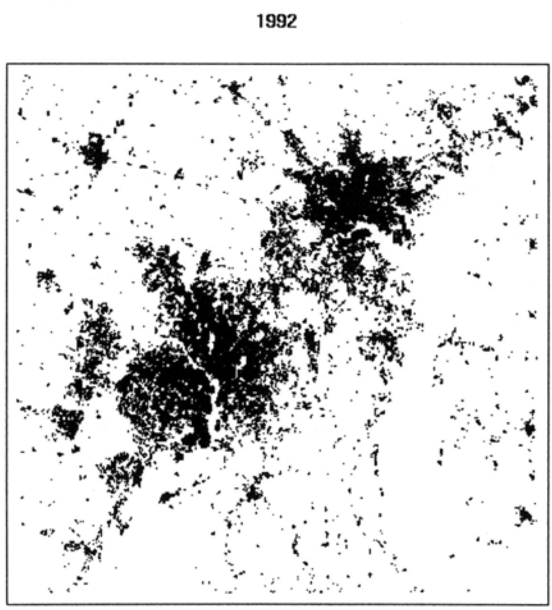

Figure 2. Level 0 grid for 1992.

Table 2. Characteristics of the image pyramid

LEVEL	CELLS PER ROW	CELL SIZE (meters)	MAXIMUM VALUE	EXAMPLE
9	1	138,240	262,144	Size of the study area
8	2	69,120	65,536	
7	4	34,560	16,384	
6	8	17,280	4,096	
5	16	8,640	1,024	
4	32	4,320	256	Interpolation level
3	64	2,160	64	
2	128	1,080	16	Animation
1	256	540	4	
0	512	270	1	BaltWash Pixel

The 0-level row of the table 3 illustrates that for at least 200 years there has been some urbanization in the region (A fit of a linear model to the 0-level data yields $\ln[f_0(x_t)] = -40 + 0.026\ t$ which predicts a y-intercept at about the year 1575). The table cells that are shaded represent completely covered pyramid levels, showing how in later years the windows rapidly become saturated. This happens at $l = 8$ in 1792 and by level 6 in 1972 and later. One way to avoid this saturation is to expand the extent of the study area, and this indeed is underway. But another problem with this analysis is that traditional maps (1772-1850) produced to widely varying cartographic styles, are being analyzed along with carefully standardized USGS maps (1900-1953) and satellite imagery (1972-1992). Nevertheless—and this is another advantage of multiscale analysis—at coarser scales the difference among these disparate data sources diminishes.

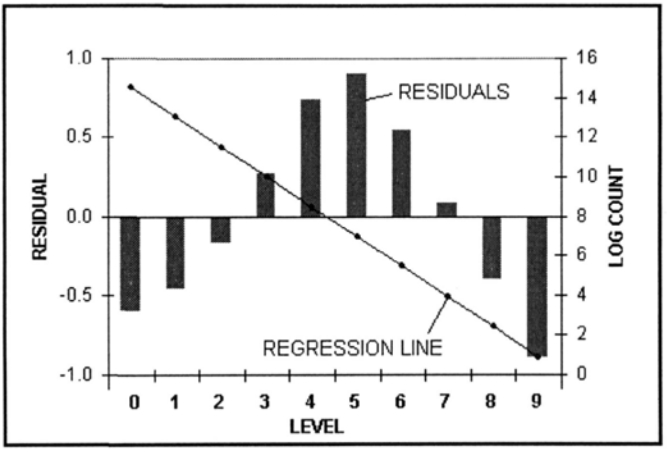

Figure 3. Fractal dimension estimation 1953.

Table 3. Box Counts for each year and level

LEVEL	1792	1850	1900	1925	1953	1972	1982	1992
9	1	1	1	1	1	1	1	1
8	4	4	4	4	4	4	4	4
7	15	16	16	16	16	16	16	16
6	31	42	62	63	63	64	64	64
5	40	75	190	218	230	242	243	245
4	52	99	360	457	588	750	770	784
3	59	126	539	763	1216	1894	1985	2076
2	83	197	909	1402	2560	4741	5118	5448
1	142	412	1790	2897	5956	12296	13564	14892
0	286	1069	4431	7089	15463	33092	36742	41183

The box counts in table 3 can be used to compute the fractal dimension of the built-up area for each year. For example, figure 3 shows the regression line estimating $\log_2[f_l(x_{1953})] = 0.89 - 1.51\ l$ for 1953, which yields a fractal dimension of $D_{1953} = 1.51$ and an $R^2 = .99$ (Falconer 1990). The box counts for each level and each year are used to compute the 8 values of D_t, the fractal dimensions for each of the data years, shown in figure 4. There is a continuing debate in urban studies about how regions develop. One school argues that so-called "primate" metropolitan regions continue to grow from a point to a centralized but spreading metropolitan pole. But another school envisions a dispersed metropolis that may eventually completely disperse, returning to a collection of isolated points (Alonso 1980, De Cola 1985). Figure 3 certainly shows the early stages of this process; we can only speculate about whether D_t will eventually decline, although its rate of increase seems to be leveling off. This scenario suggests the possibility of future dispersion in which the urban complex not only breaks up into dispersed centers but even perhaps returns to the low-dimension post-industrial "village" system similar to that of the 18th century.

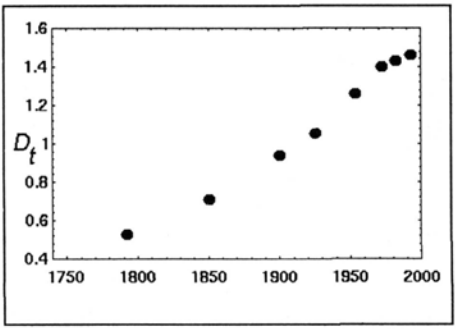

Figure 4. Boxcount fractal dimensions 1792-1992.

Each of the fractal dimensions D_t for the data years is a linear estimate of the behavior of the box counts over the scale levels. Yet the fit is not perfect, as figure 3 shows for 1953; there is a similar pattern of parabolic residuals among all the years. In general the middle scale levels $l = 4$ and 5 have higher residuals, suggesting that at about the 6-km scale the urban area has its most compact representation. But the box count aggregation algorithm, which yields 0/1 values, cannot be used to generalize the data. Another way to aggregate grid data is to sum lower-level values using $fsum_{l+1}(x_t) = \Sigma f_l(x_t)$ where the aggregation is over subwindows of 4 cells each. The algorithm *fsum* is like a mean filter that aggregates subregions into a higher-level region whose value is the average of lower-level elements. The generalized animation is therefore based on the level-4 generalization, which gives for each of $32^2 = 1024$ cells of size 4320-m an 8-bit dynamic range of [0, 256] (see table 2). Figure 5 shows what happens to the 1992 data for 5 successive levels of aggregation. The lower-level images allow us to focus on the individual features of the region, while the higher-level images highlight the unified nature of the BaltWash metropolis.

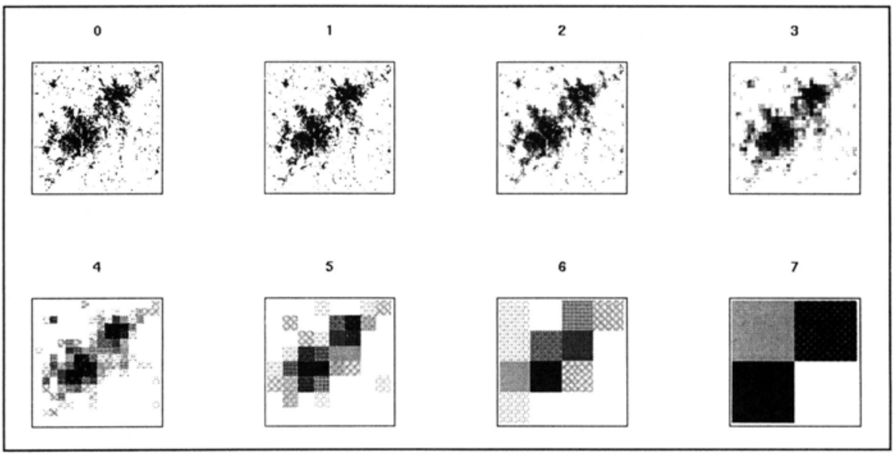

Figure 5. Sum pyramid for 1992.

Let $l = 4$ and consider the central-cell x = (col, row) = (16, 16) for each of the $t = 1,\dots,8$ data years. The values of $fsum_4(x_t)$ for this cell are shown in figure 6 and (as did D_t in figure 4) these points suggest a logistic curve, which can be estimated with an interpolation (prediction) function $fsum^p$ that predicts *fsum* for any year and not just the 8 data years. Figure 6 shows $\{fsum^p_4(x_t)$: $x = (16,16)$, $t \in [1750 \text{ to } 2000]\}$. When this function is used at level-4 we only get 32^2 predictions. This is how we obtain a gain in feature resolution (from [0,1] data to [0, 256] values), and a gain in temporal resolution (from 8 irregularly spaced measurements to 20 decadal interpolations), by sacrificing a loss in spatial resolution (from 270-m to 4321-m cells).

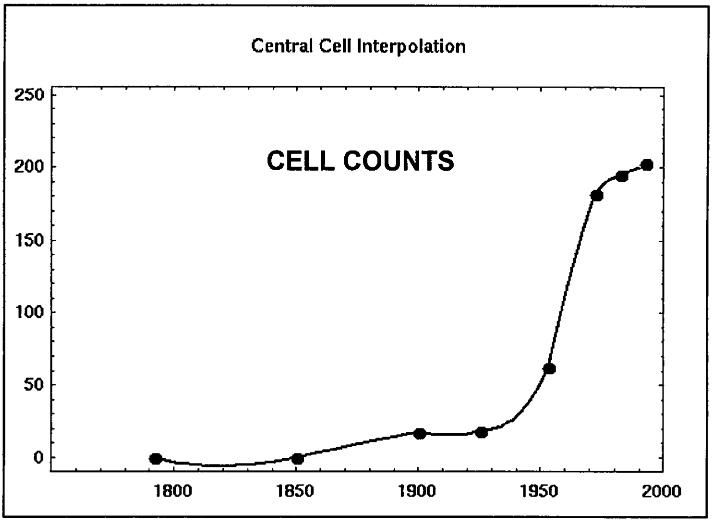

Figure 6. Actual and interpolated values for cell (16,16).

The unique temporal interpolation functions for each of the (32^2 =) 1024 level-4 cells can be arrayed into a *Mathematica* table that provides a grid of predictions for any year in the study period. A sample for 1990 is shown in figure 7, taken from the animation (USGS 1997). The data have been spatially linearly interpolated to level 1 (540 meters) to provide a smooth surface for visualization (for a alternative approaches to the interpolation problem see Tobler 1979 and Bracken and Martin 1989). The image, which is one frame of a 20-period animation, illustrates the polycentric nature of the Baltimore/Washington urban process. The animation shows reveals a self-organizing system that has been growing along the Northeast U.S. transportation corridor. During the past 200 years urban leadership has shifted between the two centers at least three times, and since World War II there has arisen a polycentric post-industrial system whose fractal dimension has been growing logistically and may be leveling off.

Another way to visualize the growth process is isopleths or contours, which emphasize the geographic location of urbanization. figure 8 shows not only the 2 urban centers in 1992, but such other features as the edge cities of Frederick, Annapolis, and La Plata, MD as well as Potomac Mills, VA. The picture also highlights the linear nature of the whole system, oriented along Interstate 95, which continues from Boston to Miami.

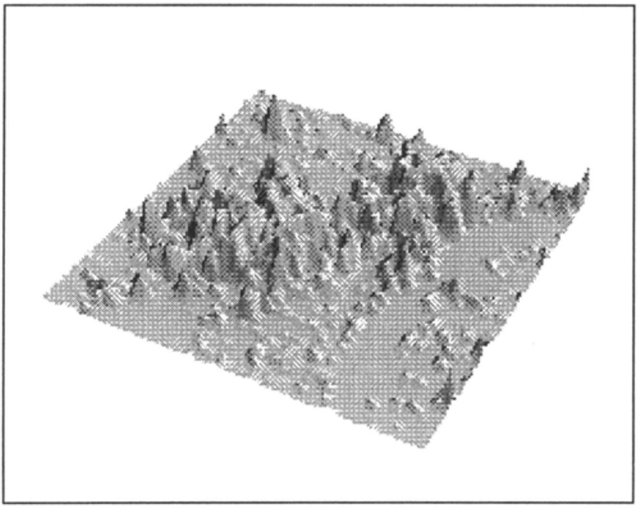

Figure 7. 3-d plot of estimated built-up areas for 1990.

Naturally we are interested in the future of the region, and the analysis suggests approaches. (A logistic curve fitted to the 0-level data in table 2 yields $f_0(x_t) = 55800\ [1 + \mathrm{Exp}(2.09 - 0.0469(t - 1923))]^{-1}$, which has a maximum growth rate of 2.1% in 1923 (Haggett, Cliff and Frey 1977:238)). This expression has an asymptotic value of 55,800 pixels, which is only about 20% of the window at level-0.

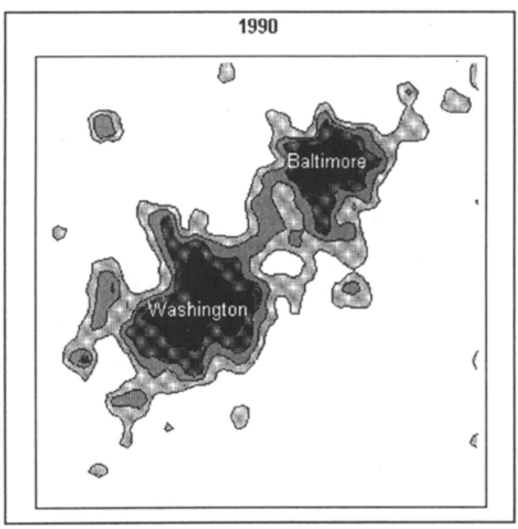

Figure 8. Contour plot of estimated built-up areas 1990.

The analysis of the last 3 data years (1972, 1982, 1992) was based on Landsat imagery, and the growth both of the fractal dimension D_t (figure 4) as well as of one of the generalized cells $fsum_4(x_t)$ (figure 6) show a linear growth trend. The growth rate for 1972-1992 is mapped in figure 9; darker shades show faster growth—up to 2% per year. Recent metropolitan development displays the doughnut patterns typical of U.S. cities (Whyte 1968). The Baltimore growth ring is broken by Patapsco Bay and the Washington ring by a Potomac River "greenbelt" that would clearly be the fastest growing edge city were the river bridged from Sugarland Run VA to Seneca Creek MD. It is interesting how strongly topography still influences the development of this region.

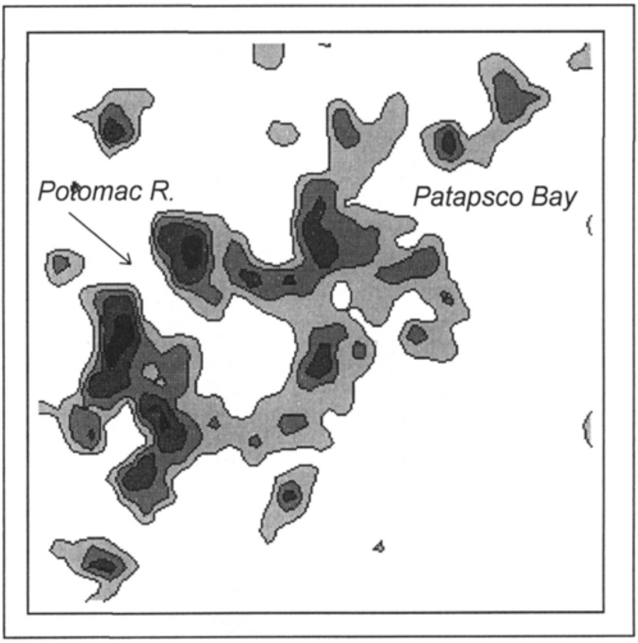

Figure 9. Contour plot of growth rates, 1972-1992.

The research presented here is part of a 118-year history of the use of USGS core skills in the physical, and—more recently—human and biological sciences to understand human-induced land transformations. These efforts exhibit not only institutional expertise but also rich historical databases that can be used to understand spatial processes, to forecast change, and help to shape future policy. The dimensions highlighted in table 1 suggest new directions for this research. First, the analysis can profit from a broader spatial view, expanding to Megalopolitan and even world urbanization. Second temporal extrapolation and deeper "data mining" will help planners envision the future of the region—as well as its distant past. Third, more features (shoreline, land cover, climate) need to be studied and animated. A central theoretical and policy problem highlighted by this work therefore is the development of rigorous, informative, and visually effective transformations of data along and among spatial, temporal, and phenomenological scales.

REFERENCES

Acevedo, William, Timothy Foresman, and Janis Buchanan 1996 Origins and philosophy of building a temporal database to examine human transformation processes, ASPRS/ACSM Technical Papers I:148-161.

Alonso, William 1980 Five bell shapes in development, *Papers and proceedings of the Regional Science Association*, 45:5-16.

Bak, Per 1996 *How nature works*, NY: Cambridge University Press.

Batty, Michael 1994 *Fractal cities*, NY: Wiley.

Bracken, I and D. Martin 1989 The generation of spatial population distributions from census centroid data, *Environment and Planning A*, 21:537-543.

Clark, Susan C., John Starr, William Acevedo, and Carol Solomon 1996 Development of the temporal transportation database for the analysis of urban development in the Baltimore-Washington region, ASPRS/ACSM Technical Papers, 3:77-88.

Crawford-Tilley, Janet S., William Acevedo, Timothy Foresman, and Walter Prince 1996 Developing a temporal database of urban development for the Baltimore/Washington region, ASPRS/ACSM Technical Papers, 3:101-110.

De Cola, Lee 1985 Lognormal estimates of macro-regional city size distributions, 1950-1970 *Environment and Planning A* 17:1637-1652.

De Cola, Lee 1997 Multiresolution covariation among Landsat and AVHRR vegetation indices, in Quattrochi and Goodchild 1997.

Falconer, K.J. 1990 *Fractal geometry: mathematical foundations and applications*, New York: Wiley.

Gaydos, Leonard J and William Acevedo 1995 Using animated cartography to illustrate global change, International Cartographic Association, Barcelona.

Gottmann, Jean 1990 *Since Megalopolis: the urban writings of Jean Gottmann* Baltimore: Johns Hopkins University Press.

Haggett, Peter, Andrew Cliff and Allan Frey 1977 *Locational analysis in human geography* London: Edward Arnold.

Quattrochi, Dale and Michael Goodchild 1997 *Scale in remote sensing and GIS* Boca Raton FL: CRC Press.

USGS1997 The animation is available on the WorldWideWeb at: http://geog.gmu.edu/gess/classes/geog590/gis_internet/ldecola/baltwash/

Tobler, Waldo R. 1979 Smooth pyncnophylactic interpolation for geographical regions *Journal of the American Statistical Association* 74(367):519-530.

Whyte, William H. 1968 *The last landscape*, Ch. 8, NY: Doubleday.