Generating Surface Models of Population Using Dasymetric Mapping*

Jeremy Mennis
University of Colorado

Aggregated demographic datasets are associated with analytical and cartographic problems due to the arbitrary nature of areal unit partitioning. This article describes a methodology for generating a surface-based representation of population that mitigates these problems. This methodology uses dasymetric mapping and incorporates areal weighting and empirical sampling techniques to assess the relationship between categorical ancillary data and population distribution. As a demonstration, a 100-meter-resolution population surface is generated from U.S. Census block group data for the southeast Pennsylvania region. Remote-sensing-derived urban land-cover data serve as ancillary data in the dasymetric mapping. **Key Words: areal interpolation, dasymetric mapping, population data, surface modeling.**

Introduction

Most publicly available demographic datasets, such as those generated by the U.S. Census, are aggregated to areal units, such as counties and census tracts. A variety of analytical issues associated with these aggregated datasets concern the typically arbitrary nature of the areal unit partitioning. Perhaps the most prominent of these issues is the modifiable areal unit problem (MAUP), defined as a situation in which modifying the boundaries and/or scale of data aggregation significantly affects the results of spatial data analysis (Openshaw 1983). Consequently, it is often unclear whether the results of census-data analysis indicate some reality about the individuals living in that region or are strictly a function of the particular areal unit used in the analysis (Openshaw 1984). A related problem concerns the display of demographic data. Choropleth maps of population by administrative areal unit give the impression that population is distributed homogeneously throughout each areal unit, even when portions of the region are, in actuality, uninhabited (Dorling 1993).

One potential solution to these problems is a surface-based demographic data representation, in which data are modeled as a continuous field that is not dependent on an irregular partitioning into arbitrary areal units. Areal unit versus surface models of population can be understood via the object versus field representations of geographic reality considered in geographic information science (Goodchild 1992). The object view treats population as a set of individual geographic entities (i.e., administrative areal units) to which population attributes may be attached. The field view, on the other hand, treats population as a continuously varying surface whose value (i.e., population density) may be measured at any given location. Of course, population is in reality composed of individual people; both object and field representations of population are thus abstractions of that reality. In geographic information systems (GIS), the object view is typically represented using points, lines, and polygons in the vector data model, and the field view is typically represented by an exhaustive tessellation of square grid cells in the raster data model.

Surface-based population representation offers certain advantages over areal unit representation. A surface-based representation allows for population data aggregation to nearly any desired areal unit and hence is not subject to the MAUP and other areal unit-derived problems (Bracken 1993). In addition, because surface representations can present a graphic unit of display (a grid cell) that is uniform in size across a region, surfaces of population may offer a more accurate cartographic

*The author would like to thank Cory Eicher and Alan MacEacren for many helpful conversations regarding dasymetric mapping and Barbara Buttenfield for her constructive comments on a previous draft of this article.

The Professional Geographer, 55(1) 2003, pages 31–42 © Copyright 2003 by Association of American Geographers.
Initial submission, November 2001; revised submission, June 2002; final acceptance, August 2002.
Published by Blackwell Publishing, 350 Main Street, Malden, MA 02148, and 9600 Garsington Road, Oxford OX4 2DQ, UK.

representation of population distribution than do conventional choropleth maps (Langford and Unwin 1994). While most demographic data are not available in a surface-based format, raster GIS provides an environment in which to develop reliable and useful surface-based representations of population and population character from aggregated census data (Langford, Maguire, and Unwin 1991; Martin and Bracken 1991). This article describes a methodology for generating surface-based representations of demographic data using a dasymetric mapping technique that incorporates satellite-derived urban land-cover data.

Areal Interpolation and Dasymetric Mapping

The problem of creating a population surface from areal unit data is essentially one of *areal interpolation*, the transformation of geographic data from one set of boundaries to another. Areal interpolation is typically used to compare two or more spatial datasets that are stored in incompatible areal units, such as congressional districts and census tracts. In a sense, the generation of a raster population surface is a special case of areal interpolation, because the desired (target) areal unit (a raster grid cell) is intended to approximate a continuous surface; hence, it is necessarily much smaller than the size of the original areal unit of data aggregation.

Flowerdew and Green (1992) and Goodchild, Anselin, and Deichmann (1993) offer descriptions of various areal interpolation techniques that can be applied to the transformation of population data. Perhaps the most straightforward technique for transformation to raster format is *areal weighting*, whereby each grid cell is assigned a population value based on its percentage area of the host areal unit. This method meets the requirements for preserving the *pycnophylactic* property (Tobler 1979), defined as when the summation of population data to the original set of areal units is preserved in the transformation to a new set of areal units—that is, "[P]eople are not destroyed or manufactured during the redistribution process" (Langford and Unwin 1994, 24).

Bracken and Martin (1989) and Martin (1989) describe a more sophisticated approach to developing surface representations of demographic data for census enumeration districts in the United Kingdom. Their method is a variant of inverse distance weighted (IDW) interpolation. Population counts are assigned to a set of summary points generated from the centroids of the original areal units. A moving window operation over an "empty" raster grid then assigns to the window kernel a value according to the population values of those centroids contained within the window, with closer centroids having more "weight" than those centroids farther away. The relative density of centroids around the kernel determines the size of the window. This method assumes that population density decreases away from the centroid according to some distance-decay function and allows for some areas of the raster surface to contain zero population.

Other approaches to population surface generation have used dasymetric mapping, first popularized in the U.S. by Wright (1936). Dasymetric mapping may be defined as a kind of areal interpolation that uses ancillary (additional and related) data to aid in the areal interpolation process. Dasymetric mapping differs from choropleth mapping in that the boundaries of cartographic representation are not arbitrary but reflect the spatial distribution of the variable being mapped (Eicher and Brewer 2001). Wright demonstrated dasymetric mapping by first redistributing population from a set of areal units into inhabited and uninhabited regions as indicated on USGS topographic maps. He then subdivided the inhabited regions into smaller portions, using settlement pattern data also gathered from USGS topographic maps. Population density values are derived subjectively for the different types of settlement patterns, and this information is used to estimate population density for the portions of the inhabited regions according to the fraction of inhabited region area each portion occupies.

Langford and colleagues (1991) describe a dasymetric mapping procedure for generating raster population surfaces. These authors used land-use data derived from Landsat Thematic Mapper (TM) multispectral imagery to build a series of predictive models regressing population density on land use. These models were then used to redistribute U.K. census-ward

population data for Leicestershire to a 1-km-resolution raster surface. While the results of this approach proved promising, the models typically overestimated population for urban areas and underestimated population for suburban and rural areas (Langford et al. 1991). In later work, Langford and Unwin (1994) used satellite imagery to determine residential and nonresidential pixels for the same region. A raster surface was then generated in which the population of each census ward was divided evenly among the residential pixels within that census ward.

Eicher and Brewer (2001) offer a review and evaluation of a number of dasymetric mapping techniques, including the use of raster-based approaches to areal interpolation. Of particular import to this discussion is what these authors call the "grid three-class" method, which they demonstrate by using raster land-use data (with three classifications) to redistribute county-level population data to subcounty units. In their study, they assigned a predetermined percentage of a county's population to a given land-use area in that county. For instance, they assigned 70% of the population of a county to urban, 20% to agricultural/woodland, and 10% to forested land uses. They note that while improving the accuracy of population distribution, this method suffers from two weaknesses: first, like Wright's (1936) approach, the percentages are subjectively determined, and second, the method does not account for differences in area among the three land-use classes within a county.

Demonstration of the Dasymetric Mapping

The raster population surface generation methodology described in the present article is related to Langford and Unwin's (1994) use of remotely sensed residency data to redistribute population, Eicher and Brewer's (2001) grid three-class method, and Wright's (1936) initial dasymetric mapping technique. Like Langford and Unwin, I use remote sensing-derived ancillary data to redistribute population to a raster grid. However, instead of a binary assignment of population to a raster surface based on residential/nonresidential pixel classification, I use a three-tier classification of urban land-cover as ancillary data. I use these ancillary data within a dasymetric mapping framework similar to the grid three-class method described by Eicher and Brewer (2001).

I propose two techniques for addressing the weaknesses of the grid three-class method. The first technique uses empirical sampling to determine appropriate percentage assignment values. This technique mitigates the subjectivity of the assignment of a percentage of population to a given ancillary data class (i.e., land use or urban land-cover). The second technique draws from Wright's (1936) approach to area-based weighting to address the differences in area among ancillary data classes within a given areal unit. These techniques are demonstrated by generating a raster population surface from 1990 U.S. Census block group data (Geolytics 1998) using ArcView GIS and ArcView's Spatial Analyst extension for raster data handling.

This demonstration maps population in the five-county southeast Pennsylvania region, including Bucks, Chester, Delaware, Montgomery, and Philadelphia counties (Figure 1). This region provides an example in which block group area and population density vary significantly from urban to rural parts of the region (Figure 2). Block groups in urban areas may be relatively small and of homogeneous population density, while block groups in rural areas are typically much larger, and have a

Figure 1 *The five-county region of southeast Pennsylvania. Note that the county and city boundaries for Philadelphia are identical.*

much more heterogeneous population distribution. Dasymetric mapping may therefore be used to generate a surface model that provides a more accurate representation of population within rural block groups, as well as in urban block groups that contain parks, cemeteries, and other features that may control the within-block group distribution of population.

Urban Land-Cover Data

The urban land-cover dataset used here was acquired from the Pennsylvania Explorer CD released by the Pennsylvania State University (Digital Equipment Corporation et al. 1998). This categorical polygon dataset exhaustively partitions Pennsylvania into high-density urban, low-density urban, and nonurban areas (Figure 3). Note that urban "density," in this case, refers not to population density but to "urbanization," or degree of urban development. For the sake of clarity, and to distinguish urban density from population density, I refer to the urban land-cover data as "urbanization" data with high, low, and nonurban classes. The urbanization data were generated from 1993 Landsat TM imagery that was classified to yield a raster grid of land cover for Pennsylvania (see Digital Equipment Corporation et al. 1998 for details of the classification process). The land-cover data were graphically overlaid with 1996 Pennsylvania Department of Transportation road-network data and photo-interpreted according to land cover and road density to generate the urbanization dataset.

While the use of urbanization data as a proxy for population distribution is not perfect, it is justified by its utility. Satellite remote sensing cannot indicate population density directly but can describe the urban morphology of built-up and nondeveloped areas.

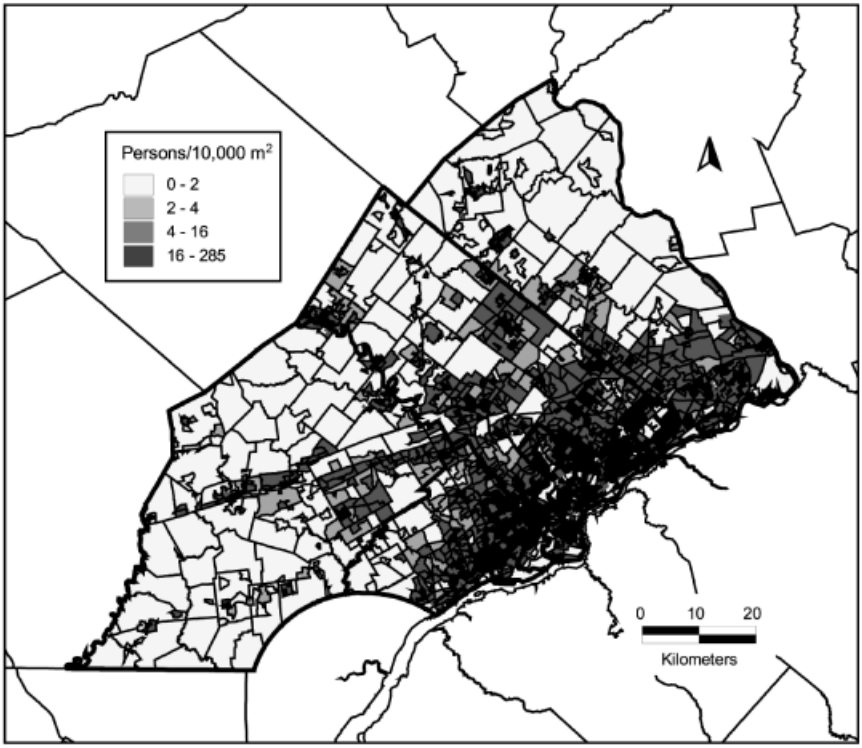

Figure 2 *Population density by block group for southeast Pennsylvania. Note that for graphical clarity, only those block-group boundaries for block groups with an area greater than 1,000,000 m² are shown. The use of an exponential class-break scheme is necessitated by the extreme skewness of the data distribution.*

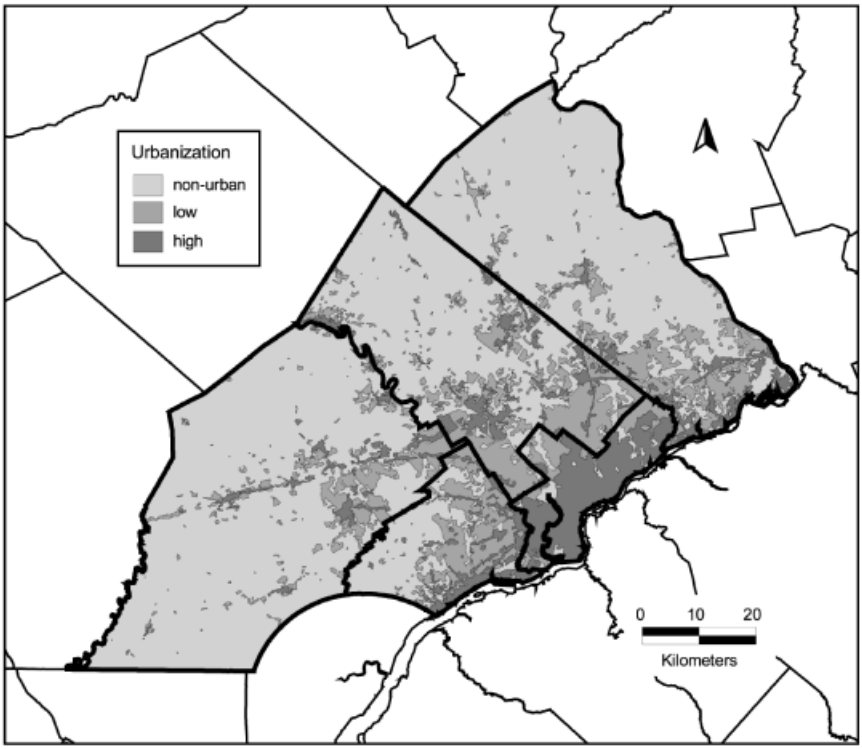

Figure 3 *Urbanization classes for southeast Pennsylvania.*

A generally predictable positive relationship occurs between population density and the degree of urban development as indicated by satellite imagery (Langford, Maguire, and Unwin 1991; Mesev 1999), although there are a number of problematic issues associated with the use of satellite remote sensing for monitoring urban areas. Forster (1985) notes some of these issues, such as pixel resolution and the intrapixel heterogeneity of urban regions, attenuation of the electromagnetic signal due to the earth's atmosphere, and error due to variations in sensor calibration and platform-target geometry over time.

Perhaps the most troubling issue for this particular project is that certain regions, such as industrial complexes, that are sparsely populated (in terms of residence) may be classified as high urbanization due to their dense road networks and large areas of impervious surface, the spectral signature of which in TM imagery may resemble residential and commercial areas. While the use of land-use data instead of urbanization data may solve some of these problems (Monmonier and Schnell 1984), timely land-use data for the five-county southeast Pennsylvania region that distinguished between industrial, commercial, and residential land uses were not available. The anomaly of sparsely inhabited, high urbanization areas in the dataset must be acknowledged as a source of error when using this dataset for areal interpolation of population. However, because the method described here preserves the pycnophylactic property (as outlined below), all error is inherently limited to variation *within* each individual original areal unit; the population of the area described by each block group is preserved in the transformation to raster data surface.

Methods

The urbanization data were initially converted to a 100-m-resolution raster grid. This grid cell

resolution serves as the resolution for the final raster population surface. The choice of grid cell size is important, as the resolution must be fine enough to capture the desired spatial variation of population within the area of interest. Note also that if the grid cell resolution exceeds the size of the smallest areal unit in the original population or ancillary datasets (i.e., block group), data will be lost in the vector to raster transformation. A 100-m-resolution was chosen to capture the population variation within urban areas of Philadelphia and to keep computations at acceptable processing times.

The population of each block group is distributed to each grid cell in the population surface based on two factors: (1) the relative difference in population densities among the three urbanization classes; and (2) the percentage of total area of each block group occupied by each of the three urbanization classes. Factor one concerns the fact that a grid cell with a high urbanization class has a higher population density than a grid cell with a low or nonurban urbanization class (as derived from empirical measurement, described below). Thus, the high-urbanization grid cell should receive a greater share of the total population assigned to a block group than a low or nonurban urbanization grid cell in the same block group.

In order to determine the relative difference in population density among urbanization classes, population density values for each urbanization class were sampled. The sampling process selected all block groups that are *entirely contained within* each urbanization class, found their total population and area, and calculated their aggregated population density. This calculation was performed independently for each county, because the relative difference in population density among urbanization classes varies from county to county—population density in high-urbanization areas in Philadelphia County (the urban core) is much higher than that in high-urbanization areas in Bucks County (the suburban-rural fringe) (Table 1). Note that it is this difference in population density for the same urban morphologic classification (or land-use classification) that disrupts the creation of a predictive model, as Langford and colleagues (1991) testify. In other words, one cannot accurately predict the population density of a block group in the region based on, say, the percentage of that block group occupied by a particular urbanization class. However, the actual population density values themselves are not of concern here; what is important is the within-county *relative* difference in population density among the urbanization classes.

Each urbanization class within each county was assigned a "population density fraction" number that indicates, based on the relative differences in population density, the percentage of a block group's total population that should be assigned to a particular urbanization class

Table 1 *Aggregated Population Densities for Block Groups Completely Contained within Each Urbanization Class for Each County in Southeast Pennsylvania*

County	Urban Class	Population	Area (m^2)	Population Density (persons/10,000 m^2)	Sum Density	Population Density Fraction
Bucks	High	1,562	43,000	36.33	56.38	64.44
Bucks	Low	35,637	1,841,000	19.36	56.38	34.34
Bucks	Nonurban	8,345	12,041,000	0.69	56.38	1.22
Chester	High	12,945	244,000	53.05	61.07	86.87
Chester	Low	6,514	893,000	7.29	61.07	11.94
Chester	Nonurban	15,849	21,767,000	0.73	61.07	1.2
Delaware	High	66,815	1,460,000	45.76	73.42	62.33
Delaware	Low	53,124	2,218,000	23.95	73.42	32.62
Delaware	Nonurban	4,702	1,267,000	3.71	73.42	5.05
Montgomery	High	33,952	1,079,000	31.47	47.11	66.80
Montgomery	Low	77,011	5,238,000	14.70	47.11	31.2
Montgomery	Nonurban	2,807	2,979,000	0.94	47.11	2.00
Philadelphia	High	1,195,621	15,263,000	78.33	116.60	67.18
Philadelphia	Low	79,485	2,195,000	36.21	116.60	31.05
Philadelphia	Nonurban	3,816	1,848,000	2.06	116.60	1.77

within the block group. The population density fraction is calculated by dividing an urbanization class's population density by the sum of the population density values for all three urbanization classes, which may be expressed as:

$$d_{uc} = p_{uc}/(p_{hc} + p_{lc} + p_{nc}) \quad (1)$$

where d_{uc} = population density fraction of urbanization class u in county c, p_{uc} = population density (persons/10,000 m^2) of urbanization class u in county c, p_{hc} = population density (persons/10,000 m^2) of urbanization class h (high) in county c, p_{lc} = population density (persons/10,000 m^2) of urbanization class l (low) in county c, and p_{nc} = population density (persons/10,000 m^2) of urbanization class n (nonurban) in county c. This operation was performed for each individual county. For example, if, hypothetically, in a particular county, the high, low, and nonurban urbanization classes had sampled population densities of 6 persons/10,000 m^2, 3 persons/10,000 m^2, and 1 person/10,000 m^2, respectively, their population density fractions would be 0.6, 0.3, and 0.1, respectively. Thus, if a block group in that county had a population of 100 people, 60 (100 people $\times$ 0.6) of those people would be assigned to the portion of that block group that is classified as high urbanization. Table 1 reports the population density fractions for each urbanization class for each county.

This approach assumes that the block group is evenly spatially partitioned among the three urbanization classes—that each urbanization class occupies 33.3% of the total block group area. This occurrence, of course, is an extremely rare event. Factor two thus addresses the difference in block group area occupied by each urbanization class. In order to accurately assign a portion of a given block group's total population to an urbanization class, the population density fraction must be adjusted by the percentage of that block group's total area that that urbanization class occupies. In other words, the county-level urbanization class's population density fraction must be adjusted for each individual block group according to the difference in area occupied by each urbanization class within that block group.

This adjustment is made by calculating the "area ratio" number for each urbanization class for each block group. This number represents the ratio of the percentage of area that an urbanization class actually occupies within a block group to the "expected" percentage of 33.3%. The area ratio for a given urbanization class within a given block group may be calculated by dividing the number of grid cells (i.e., area) of the urbanization class by that block group's total number of grid cells (i.e., area) and dividing that sum by 33.3, which may be expressed as:

$$a_{ub} = (n_{ub}/n_b)/0.33 \quad (2)$$

where a_{ub} = area ratio of urbanization class u in block group b, n_{ub} = number of grid cells of urbanization class u in block group b, and n_b = number of grid cells in block group b.

This operation was performed for each urbanization class in each individual block group. For example, if, hypothetically, in a particular county, the high, low, and nonurban urbanization classes had areas of six grid cells, three grid cells, and one grid cell, respectively, their area ratios would be 1.8, 0.9, and 0.3, respectively.

The population density fraction and area ratio may be integrated into one term, referred to as the "total fraction," which represents the fraction of a given block group's total population that should be assigned to a given urbanization class within that block group, accounting for variation in both population density and area of the different urbanization classes. The total fraction may be calculated by multiplying the population density fraction and area ratio of a given urbanization class in a given block group and dividing that result by the result of that same expression for all three urbanization classes in that block group. This calculation may be expressed as:

$$f_{ubc} = (d_{uc} \times a_{ub}) / [(d_{hc} \times a_{hb}) + (d_{lc} \times a_{lb}) + (d_{nc} \times a_{nb})] \quad (3)$$

where f_{ubc} = total fraction of urbanization class u in block group b and in county c, d_{uc} = population density fraction of urbanization class u in county c, a_{ub} = area ratio of urbanization class u in block group b, d_{hc} = population density fraction of urbanization class h (high) in county c, d_{lc} = population density fraction of urbanization class l (low) in county c, d_{nc} = population density fraction of urbanization class n (non-urban) in county c,

a_{hb} = area ratio of urbanization class h (high) in block group b, a_{lb} = area ratio of urbanization class l (low) in block group b, and a_{nb} = area ratio of urbanization class n (nonurban) in block group b.

Once the total fraction for each urbanization class for each block group is determined, a portion of a block group's total population may be assigned to each urbanization class's grid cells within that block group. To assign a population value to a given grid cell with that urbanization class, one merely divides the population assigned to that urbanization class evenly among the grid cells in that block group that have that urbanization class. The calculation to assign a population value to a given grid cell within a given block group and within a given county may therefore be expressed as:

$$\mathrm{pop}_{ubc} = (f_{ubc} \times \mathrm{pop}_b)/\mathrm{n}_{ub} \qquad (4)$$

where pop_{ubc} = population assigned to one grid cell of urbanization class u in block group b and in county c, f_{ubc} = total fraction for urbanization class u in block group b and in county c, pop_b = population of block group b, and n_{ub} = number of grid cells of urbanization class u in block group b.

This calculation was implemented in ArcView GIS by adding and calculating values for fields added to the polygon attribute table of the block group shapefile. One of the original fields in this table stored the 1990 total population for each block group. New fields were then created to store the population density fraction (three new fields) and area ratio (three new fields) for each urbanization class for each block group. Based on the values of the population density fraction fields and the area-ratio fields, three new fields representing the total fraction for each urbanization class for each block group were added and their values calculated. Another three new fields were created that stored the number of grid cells of each urbanization class for each block group. The values for these three fields were calculated by summing the number of grid cells per block group by urbanization class, a raster/vector zonal operation provided by ArcView's Spatial Analyst extension. After all of these table operations, there were fields in the block group polygon attribute table that represented each of the terms in Equation (4). The values in these fields were then used to calculate the values for another three fields that stored the population value assigned to a grid cell of each urbanization class within each block group.

Three raster grids were then created by converting the vector block group coverage to a raster grid. In the creation of the first grid, each high-urbanization grid cell was assigned its appropriate population value from the block group polygon attribute table by "masking" all those cells that were not classified as high urbanization. The same procedure was performed for the low urbanization class and the nonurban urbanization class to generate the two other raster grids. The three grids were then merged to produce one complete raster grid of population in southeast Pennsylvania. Note that this population surface simultaneously describes population-count data (number of persons stored within each grid cell) and population density (persons/10,000 m^2, the area of each grid cell).

Results

Figure 4 shows a map of the raster population surface for southeast Pennsylvania. Note that in the urban core areas, it does not appear to differ significantly from the vector block group map (Figure 2). In urban areas where there are parks or cemeteries, however, the raster grid is significantly more detailed. Figure 5 shows an area on the boundary of Philadelphia County and Montgomery County, denoted as rectangle A in Figure 4. Because this area lies along the urban/suburban boundary, and because of the presence of a relatively large park that sits adjacent to densely populated urban neighborhoods, considerable intra–block group variation exists in urbanization class and, thus, population density, in this area. Figure 5 compares the representation of population in the raster surface with the vector block group shapefile. The dasymetric mapping procedure redistributed the within-block group population to certain raster grid cells according to those grid cells' urbanization class. This is especially evident in the southwest quadrant of the area, in which the vector representation indicates a relatively homogeneous population density, whereas the raster surface has concentrated that population in certain sub-block group regions.

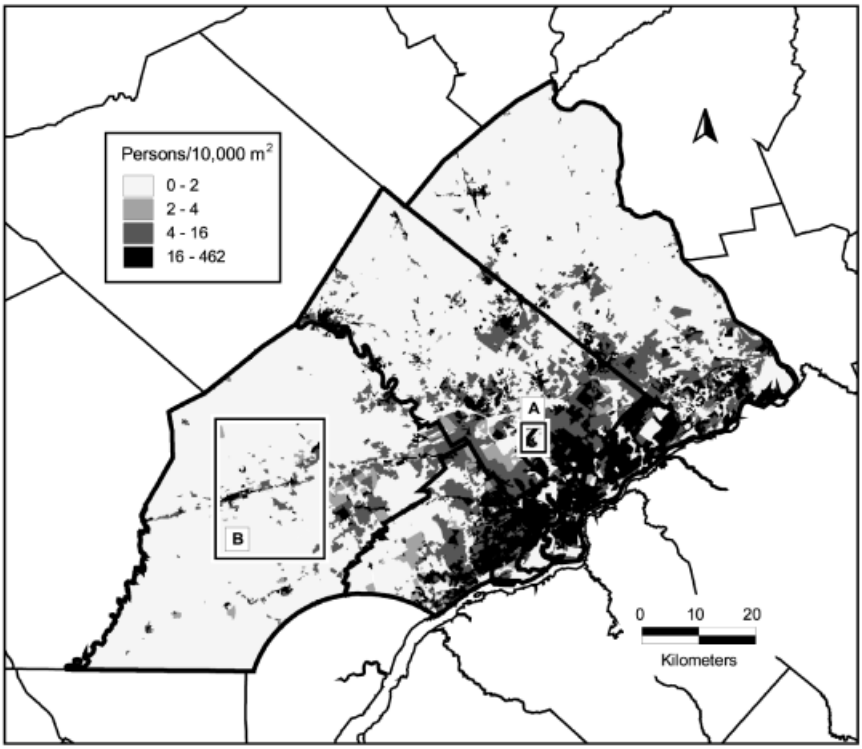

Figure 4 *Raster surface of population density for southeast Pennsylvania. Rectangles A and B show areas of detail presented in Figures 5 and 6, respectively.*

In rural areas, block groups are typically much larger than in urban areas, and these block groups may contain urbanized areas within which a majority of that block group's population resides. Figure 6 shows a rural area in Chester County, delineated as rectangle B in Figure 4, that demonstrates how the dasymetric mapping technique redistributes population within such block groups. This redistribution of population is particularly evident in the center of the area, where population in the raster surface is clearly concentrated in a number of sub-block group-sized high urbanization areas.

Conclusion

This research demonstrates the use of dasymetric mapping to create raster surface representations of population and population density. Two new techniques are presented that may improve previous methods in dasymetric mapping. The first technique suggests an empirical sampling approach to assessing the relationship between categorical ancillary data and population density. While this approach does not provide a predictive model of this relationship, Langford and colleagues (1991) have shown that the derivation of such models is made difficult due to the spatial variation in the nature of land use, urbanization, or other land-cover-based classifications as they relate to population density. The sampling approach described here does offer an improvement over the relatively simplistic approach whereby, say, 70% of an areal unit's population is assigned to a particular land-cover class based solely on the subjective decision of the analyst. The obvious drawback to the approach described here is that it assumes that at least some of the original areal units of population are small enough to be

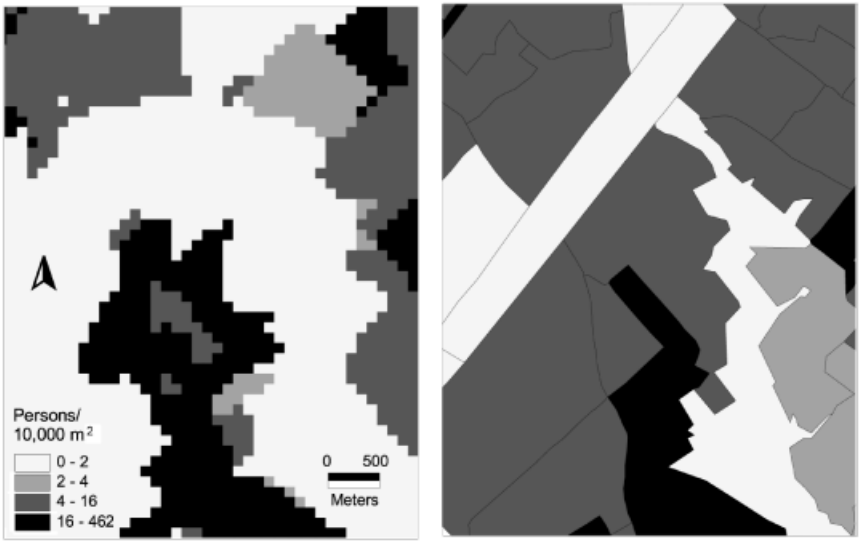

Figure 5 *Detail of an urban area, specified in Figure 4 as rectangle A, showing the difference between raster population surface (left) and vector block group (right) representations of population density.*

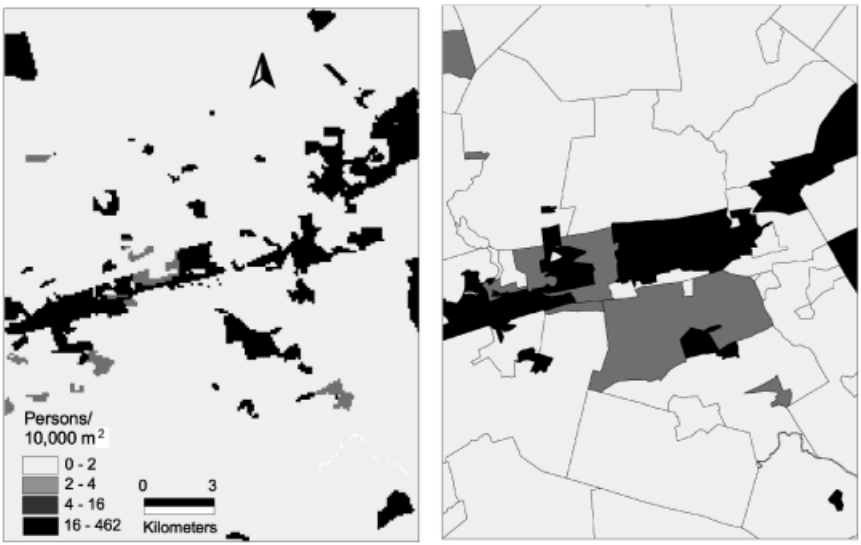

Figure 6 *Detail of a rural area, specified in Figure 4 as rectangle B, showing the difference between raster population surface (left) and vector block group (right) representations of population density. Note that the scale of this figure is significantly smaller than that of Figure 5.*

contained entirely within the area of each ancillary data class, so that the population density of the ancillary data classes may be sampled. This drawback may be countered by loosening the rules by which the sampling occurs. For instance, while I used the criteria

"entirely within" to sample individual block groups associated with a given urbanization class, one could instead select those block groups the centroids of which are contained within a given urbanization class. While I acknowledge that this would provide a less accurate sampling of population density per urbanization class, it is still an improvement over a non-empirically-derived subjective assignment.

The second technique proposed here is the use of areal weighting to improve the accuracy of the redistribution of population according to ancillary data. Whereas previous approaches to dasymetric mapping have assigned population to sub-areal unit ancillary data classes (e.g., land cover) based solely on the nature of the class (e.g., urban versus forested), the technique described here demonstrates how this assignment of population may be improved by incorporating the relative difference in area among different ancillary classes within the redistribution calculation.

These techniques may be easily incorporated within a dasymetric mapping analysis using raster GIS packages. While I demonstrated these techniques here using ArcView GIS, most GIS packages that support vector and raster data handling also support the calculations necessary for these techniques using simple drop-down menus or push-button controls. In other words, programming is not required to implement these techniques in ArcView GIS and most other GIS packages. In addition, the techniques described here are generalizable to a variety of settings. While I used urbanization data for the dasymetric mapping, one could easily use these techniques with other categorical ancillary data that have a demonstrable spatial relationship with the distribution of population. ■

Literature Cited

ArcView GIS. Version 3.2. Environmental Systems Research Institute, Inc., Redlands, CA.

Bracken, Ian. 1993. An extensive surface model database for population-related information: Concept and application. *Environment and Planning B: Planning and Design* 29:13–27.

Bracken, Ian, and David Martin. 1989. The generation of spatial population distributions from census centroid data. *Environment and Planning A* 21:537–43.

Digital Equipment Corporation, GAP Analysis Program, The Pennsylvania State University, Pennsylvania Mapping and Geographic Information Consortium, and Environmental Systems Research Institute. 1998. *Pennsylvania Explorer CD: Urban Areas of Pennsylvania*, CD-Rom. University Park: The Pennsylvania State University.

Dorling, Daniel. 1993. Map design for census mapping. *The Cartographic Journal* 30:167–83.

Eicher, Cory L., and Cynthia A. Brewer. 2001. Dasymetric mapping and areal interpolation: Implementation and evaluation. *Cartography and Geographic Information Science* 28 (2):125–38.

Flowerdew, Robert, and M. Green. 1992. Developments in areal interpolation methods and GIS. *Annals of Regional Science* 26:67–78.

Forster, B. C. 1985. An examination of some problems and solutions in monitoring urban areas from satellite platforms. *International Journal of Remote Sensing* 6 (1):139–51.

Goodchild, Michael F. 1992. Geographical data modeling. *Computers and Geosciences* 18 (4):401–8.

Goodchild, Michael F., Luc Anselin, and U. Deichmann. 1993. A framework for the areal interpolation of socioeconomic data. *Environment and Planning A* 25:393–97.

Langford, Mitchel, David J. Maguire, and David J. Unwin. 1991. The areal interpolation problem: Estimating population using remote sensing in a GIS framework. In *Handling Geographical Information: Methodology and Potential Applications*, ed. I. Masser and M. Blakemore, 55–77. London: Longman.

Langford, Mitchel, and David J. Unwin. 1994. Generating and mapping population density surfaces within a geographical information system. *The Cartographic Journal* 31:21–26.

Martin, David. 1989. Mapping population data from zone centroid locations. *Transactions of the Institute of British Geographers* 14:90–97.

Martin, David, and Ian Bracken. 1991. Techniques for modelling population-related raster databases. *Environment and Planning A* 23:1069–75.

Mesev, Victor. 1999. From measurement to analysis: A GIS/RS approach to monitoring changes in urban density. In *Geographic Information Research: Trans-Atlantic Perspectives*, ed. M. Craglia and H. Onsrud, 303–17. London: Taylor and Francis.

Monmonier, Mark S., and G. A. Schnell. 1984. Land-use and land-cover data and the mapping of population density. *The International Yearbook of Cartography* 24:115–21.

Openshaw, Stan. 1983. *The modifiable areal unit problem*. Concepts and Techniques in Modern Geography, vol. 38. Norwich: Geobooks.

———. 1984. Ecological fallacies and the analysis of areal census data. *Environment and Planning A* 16:17–31.

Tobler, Waldo. 1979. Smooth pycnophylactic interpolation for geographic regions. *Journal of the American Statistical Association* 74:519–30.

Wright, John K. 1936. A method of mapping densities of population with Cape Cod as an example. *Geographical Review* 26:103–10.

JEREMY MENNIS is an assistant professor in the Department of Geography at the University of Colorado, Boulder, CO 80309. E-mail: jeremy@colorado.edu. His research interests are in the design and analysis of geographic databases.