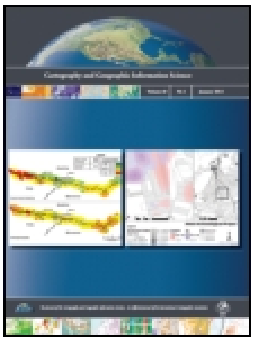

Cartography and Geographic Information Science

ISSN: 1523-0406 (Print) 1545-0465 (Online) Journal homepage: https://www.tandfonline.com/loi/tcag20

Dasymetric Mapping and Areal Interpolation: Implementation and Evaluation

Cory L. Eicher & Cynthia A. Brewer

Dasymetric Mapping and Areal Interpolation: Implementation and Evaluation

Cory L. Eicher and Cynthia A. Brewer

ABSTRACT: Dasymetric maps display statistical data in meaningful spatial zones. Such maps can be preferable to choropleth maps that show data by enumeration zones, because dasymetric zones more accurately represent underlying data distributions. Though dasymetric mapping has existed for well over a century, the methods for producing these maps have not been thoroughly examined. In contrast, research on areal interpolation has been more thorough and has examined methods of transferring data from one set of map zones to another, an issue that is applicable to dasymetric mapping. Inspired by this work, we tested five dasymetric mapping methods, including methods derived from work on areal interpolation. Dasymetric maps of six socio-economic variables were produced for a study area of 159 counties in the eastern U.S. using county choropleth data and ancillary land-use data. Both polygonal (vector) and grid (raster) dasymetric methods were tested. We evaluated map accuracy using both statistical analyses and visual presentations of error. A repeated-measures analysis of variance showed that the traditional limiting variable method had significantly lower error than the other four methods. In addition, polygon methods had lower error than their grid-based counterparts, though the difference was not statistically significant. Error maps largely supported the conclusions from the statistical analysis, while also presenting patterns of error that were not obvious from the statistics.

KEYWORDS: Dasymetric mapping, areal interpolation, mapping census data, map error

Introduction

A dasymetric map depicts quantitative areal data using boundaries that divide the mapped area into zones of relative homogeneity with the purpose of best portraying the underlying statistical surface. The dasymetric map was conceived as a type of thematic map during the early to mid nineteenth century, the formative years of modern thematic cartography. During their early development, the demand for both dasymetric and choropleth maps was driven by interest in population mapping (McCleary 1969; 1984). By 1900, dasymetric and choropleth mapping methods became more clearly differentiated, with the latter becoming overwhelmingly popular in modern cartography and for general use outside the discipline. In contrast, dasymetric mapping has remained relatively unknown even to most geographers. Consistent with their original purpose, dasymetric maps of population are still the most common type found today.

Although dasymetric maps are closely related to choropleth maps, they differ in several ways. First, zonal boundaries on dasymetric maps are based on sharp changes in the statistical surface being mapped, while zonal boundaries on choropleth maps demarcate enumeration units established for more general purposes (e.g., states within the U.S.). The cartographer generates dasymetric zones by using ancillary information. This information can be both objective and subjective, depending on other available data and the cartographer's knowledge of the area. Second, individual dasymetric zones are developed to be internally homogeneous. In contrast, choropleth zones are not defined based on the data and, thus, have varying levels of internal homogeneity. Third, choropleth mapping methods have become standardized (including the development of common classification schemes; Slocum 1999), but the wide range of dasymetric procedures have been under researched.

Surprisingly little literature exists on dasymetric mapping. At a theoretical level, MacEachren (1994) placed dasymetric maps in the continuum between isopleth and choropleth maps, suggesting that dasymetric maps represent data half way between smooth and stepped statistical surfaces. In preparation for this research, we studied the two most frequently cited early pieces by Wright (1936) and McCleary (1969). More recent implementations of dasymetric mapping methods within GIS include Gerth (1993), Holloway et al. (1996), and Charpentier (1997). Their work relies heavily on

Cory Eicher is a Sotware Development Cartography Specialist at ESRI Inc., 380 New York Street, Redlands, CA 92373, USA. **Cindy Brewer** is an Associate Professor of Geography at The Pennsylvania State University, 302 Walker Building, University Park, PA 16802, USA. E-mail: ceicher@esri.com; cab38@psu.edu.

Wright and McCleary and offers insight into the practical challenges of producing dasymetric maps using digital data.

The use of geographic information systems (GIS) for modern mapping has renewed interest in dasymetric mapping (e.g., DeMers 1997; Chrisman 1997). The lack of standardization in production methods is, however, an obstacle that prevents widespread use of dasymetric maps in GIS. Additionally, most production methods have not been evaluated through systematic testing (Robinson et al. 1995). Recent research on areal interpolation (sometimes referred to as cross-area estimation) offers a possible solution to these problems. Though areal interpolation research has not focused on map production, it can advance the study of dasymetric mapping because it has involved structured evaluation of methods that are related in their calculation approaches.

Areal interpolation is the process of transferring spatial data from one set of units to another (Bloom et al. 1996; Fisher and Langford 1996), and is often used to compare multiple datasets, each collected using different enumeration units. For example, U.S. socio-economic and population data are reported by census tract, whereas marketing data may be available by zip code, and pollution data may be collected using watersheds as enumeration units. Areal interpolation may be used to transfer all data to a common set of enumeration units (e.g., watersheds) to permit efficient comparison and analysis.

Many areal interpolation methods can be incorporated into dasymetric mapping methods to improve the detail of a choropleth map below the level of the enumeration unit (Fisher and Langford 1996). The primary difference between the two approaches is that dasymetric mapping omits the final step in areal interpolation of re-aggregating to a preferred enumeration unit type (see Eicher (1999) for a review of areal interpolation methods). A common goal of most work on areal interpolation has been evaluation of accuracy of interpolation methods. This goal contrasts with the largely demonstrative goals of cartographic work on dasymetric mapping.

One of our goals was to build connections between areal interpolation and dasymetric research. The technique we selected from the areal interpolation literature, the binary method, was one of the most accurate in previous testing. The binary method was also chosen because only minor changes would be needed to implement it as the core of a dasymetric mapping procedure. Simple areal weighting is another common areal interpolation method (Goodchild and Lam 1980; Flowerdew and Green 1989; Goodchild et al. 1993; Fisher and Langford 1996; Cockings et al. 1997). Simple areal weighting could not, however, be applied to produce dasymetric maps (at least within the framework of this experiment), because it does not use ancillary variables. We also do not report on the application of regression methods of areal interpolation. These methods are discussed in many of the papers cited, as well as in Moxey and Allanson (1994), Flowerdew et al. (1991), and other work by Flowerdew and Green (1991; 1992; 1994). Regression methods use global information for the entire study area to estimate zone values and are relatively complex when compared to techniques that have traditionally been used to create dasymetric maps in production cartography.

Dasymetric mapping and the dasymetric method have become ambiguous terms with the inception of areal interpolation research. In that literature, the dasymetric method describes a specific form of areal interpolation that we will refer to as the "grid binary" method (Langford et al. 1991; Fisher and Langford 1996; Cockings et al. 1997). In contrast, cartographers use the term "dasymetric" to refer to a general type of thematic map produced using a variety of methods. Adding further confusion to the issue, the term "dasymetric" could be used to describe most forms of areal interpolation, especially intelligent areal interpolation methods. These techniques build from Wright's (1936) idea of reassigning data between units, and they incorporate his approach of using additional information to perform this operation. Chrisman (1998) suggests that Wright's technique was actually a form of areal interpolation.

Methods

Five criteria guided dasymetric map production for the experimental evaluation we undertook. The first two criteria are basic requirements of dasymetric mapping that distinguish it from choropleth and isopleth mapping.

- *Zone homogeneity:* Final dasymetric map zones were internally homogeneous.
- *Abrupt boundaries:* Zone boundaries represented escarpments or marked changes in the mapped variable.
- *Disaggregation of source zones*: Some areal disaggregation of the original or source zones was involved (counties served as the source zones). Source zones were not small enough to allow for the creation of a dasymetric map merely through their re-aggregation.

Figure 1. Dasymetric map zones that result from overlay of county and land-use boundaries in the study area. State boundaries and example major cities are included for reference.

- *Intelligent interpolation:* An areal interpolation was performed, transforming socio-economic data from one set of zones (source counties) to new map zones, defined by the intersection of counties with land-use polygons. This interpolation was "intelligent" because it was guided by use of ancillary land-use data, in contrast to the use of simple areal weighting.
- *Nested data:* The experiment was performed at a scale that allowed checking of mapped values (derived from county data) against enumerated values (summed from block group data). Use of these nested county and block group data allowed statistical and visual evaluations of the resulting dasymetric maps and errors.

The experiment was structured to answer the following three questions:

A. Are there significant differences in the accuracy of dasymetric maps produced using the five methods tested in the experiment?
B. Are there significant differences in the accuracy of polygonal dasymetric maps versus those based on gridded data?
C. On a more qualitative level, what is the visual character of the dasymetric maps produced for the experiment, and what are the patterns of error on these maps?

In the following sections, we first describe the study area and explain our use of the land-use dataset in the mapping process. Secondly, we describe six census variables from which we made the test set of dasymetric maps. Then we outline the general mapping procedure and explain the five dasymetric mapping methods tested. We tested methods derived from traditional cartographic work as well as an areal interpolation method not specifically designed to be applied in a cartographic context. Finally, we describe our methods for calculating map errors that were used in the analysis of the test maps.

Study Area

The study area encompassed 159 counties in parts of four states—Pennsylvania, West Virginia, Maryland, Virginia—and the District of Columbia. Figure 1 shows a map of the study area with a selection of major cities and states labeled. A large number of counties were included to allow for parametric analysis of the results using counties as observations. In Virginia, cities are treated as distinct municipal units, and so these smaller polygons were included as individual observations in the analysis. The varied character of the 159 counties provided a rigorous demonstration of the methods. The chosen study area was also appropriate because it contained a variety of land-use classes and spanned a number of major landform regions, including the coastal plain, ridge-and-valley province, and Allegheny plateau. We expected different population patterns among these regions. Finally, for each mapped variable, a wide range of values were present within the study area. For example, population density varied widely within the study area, from high-density urban areas such as Washington and Baltimore to low-density rural areas in West Virginia and north-central Pennsylvania. Similar variations can be found for black population and housing-value variables (Figure 2).

Though it would have been convenient to map a combination of full states, the study area's rectangular shape lends itself well to compact display. The study area was also designed to avoid the extremely large metropolitan areas of New York (which might overly skew error summaries), while still including major metropolitan areas such as Washington, Baltimore, Pittsburgh, and Richmond (Figure 1). Also, the Chesapeake Bay shoreline was mostly removed to avoid difficulties in overlaying islands and narrow peninsular polygons. Lastly, personal familiarity with the study area allowed us to visually check results for "realistic" values, which was important in implementing several of the mapping methods.

Land-use Dataset

A necessary element for creating dasymetric maps from choropleth maps is an ancillary dataset. This information is used to assist interpolation of data from the original zones (counties) to new map zones. For this experiment, land-use data were used in both polygon and grid form. Monmonier and Schnell (1984) also used land-use data to produce a dasymetric map. Our land-use dataset was obtained from the U.S. Geological Survey in the form of a digital polygon covering the entire country at 1:7,500,000 scale. These data by Hitt (1991) were created from a map of major land uses (Anderson 1967) included in the *National Atlas* (U.S. Geological Survey 1970) and were compared to more recent state-level land-use maps (U.S. Geological Survey 1993). We clipped the data for the study area using ArcView's Geoprocessing Extension to create approximately 330 polygons. The grid dataset was created from the polygon dataset using ArcView's "Convert to Grid" operation. A grid cell of 1km^2 was chosen because we felt it was suitably sized, larger than most census block groups (used for checking the maps), and smaller than the dasymetric zones (shown in Figure 1).

Mapped Variables

Dasymetric mapping is most commonly applied to population density. For this experiment, six socio-economic variables reflecting different aspects of population distribution were mapped. All data were obtained from the 1990 U.S. Census. Two population variables were chosen: total population and black population. Four housing variables were also mapped: counts of housing units valued at less than \$25,000, \$40-49,000, \$50-59,000, and \$60-75,000. These six-count variables were mapped as densities of people or housing units. Altogether, 30 dasymetric maps were tested, six maps for each of five methods. Our goal was to make conclusions about the accuracy of the methods, using a sample of mapped variables. We felt that testing with a sample would produce stronger results than relying on differences with a single variable, which was usual in previous research.

The six variables were selected to obtain a series of original maps that had a variety of spatial patterns. Figure 2 shows pattern differences on county choropleth maps. Socio-economic variables are generally highly correlated, and our variables are no exception. Thus, there are shared character-

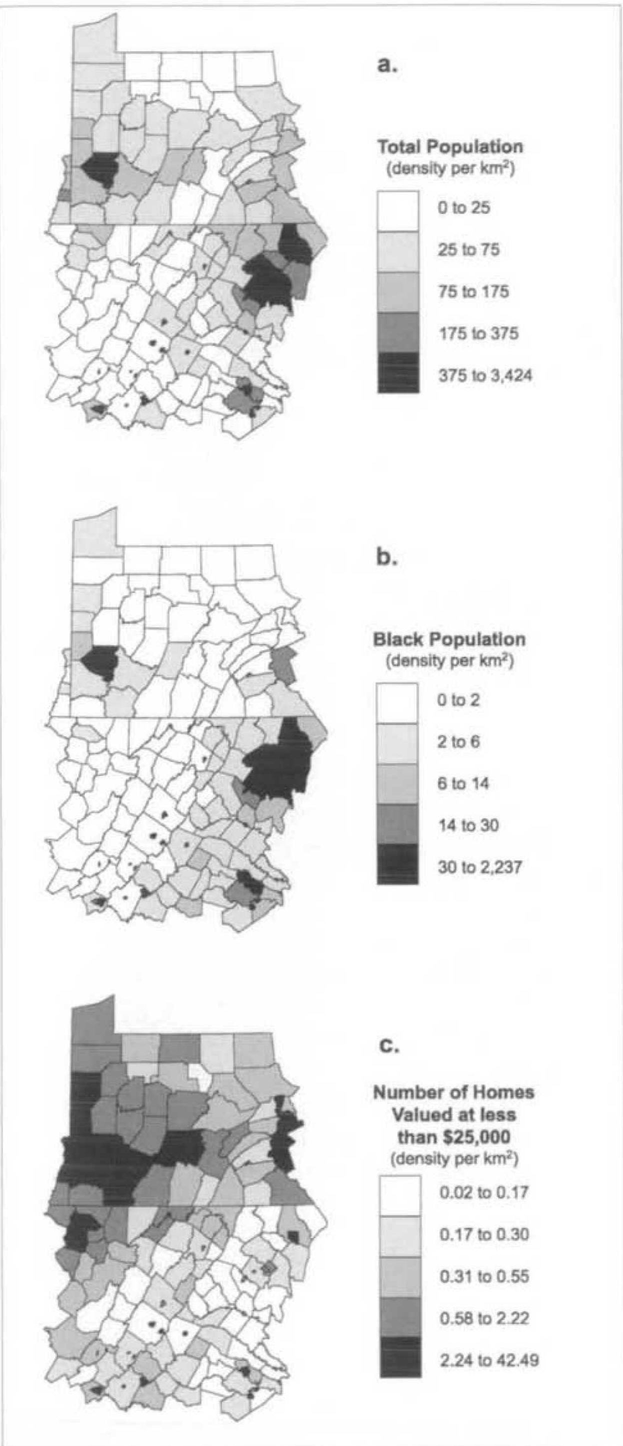

Figure 2. Example choropleth county maps of three variables used in dasymetric map evaluation.

istics in these map patterns. We attempted to select patterns that were distinct from the overall population pattern in at least some areas, such as different clusterings of high values in the north, south, and through the ridge-and-valley region (running northeast to southwest across the study area). The correlation between the density of total population and density of black population was 0.70. The correlations between total population and the four housing variables ranged from 0.60 to 0.79, while the correlations between black population and the housing variables ranged from 0.27 to 0.47. Among the four housing variables, correlations ranged from 0.81 to 0.99.

We selected data that were available for counties and block groups—two nested scales of census enumeration. In the experiment, dasymetric maps were created using the county-level data for the mapped variables. Data available at the block-group level were then used to check the results of the dasymetric mapping methods. Over 13,000 block groups were contained in the study area and were used to calculate errors for 580 dasymetric map zones.

Dasymetric Mapping Methods

Three core methods from the dasymetric and areal interpolation literature were tested. Two methods were based on dasymetric mapping research: the polygon three-class method and the traditional limiting variable method. The polygon three-class method was used to produce a series of maps in an academic production cartography lab (Holloway et al. 1996), while the traditional limiting method has roots in the 65-year-old, heavily cited paper by Wright (1936). The grid binary method, which uses raster land-use data, was selected from the areal interpolation literature.

The dasymetric methods were vector based and the areal interpolation method was grid based. To provide a balanced experiment, we tested two other methods that were adaptations of these previously existing methods: a grid three-class method (complementing the polygon three-class version) and a polygon binary method (complementing the grid binary version). They were developed for this research and have not been tested in the published literature. We felt it was necessary to test both grid and polygon versions of the methods to be able to attribute differences in performance to differences in the methods without worrying that the differences might be attributable to data structures. The limiting variable method is a polygon dasymetric method which we did not adapt for grid processing.

The following sub-sections describe the five methods we tested. We pair related grid and polygon methods in the descriptions, listing the method that originates from the literature before the derivative method (e.g., G1 before P1 and P2 before G2). Methods were implemented in ArcView using Avenue scripting language (see the Appendix in Eicher (1999) for scripts).

Grid and Polygon Binary Methods (G1 and P1)

For the grid binary method (G1), 100 percent of the data in each county was assigned to only the urban and agricultural/woodland cells derived from the raster land-use ancillary dataset. No data were assigned to the other classes (forested, water), hence the label binary. The primary advantage of this method was its simplicity. It was only necessary to reclassify the land-use data into two classes, one for inhabitable areas and another for uninhabitable areas. This subjective decision was dependent on the land-use classification set as well as information known about the mapped area.

The binary method was originally developed for use as a mapping technique (Langford et al. 1990; 1991; Langford and Unwin 1994) but eventually was applied in areal interpolation problems (Fisher and Langford 1995; Cockings et al. 1997). It represents a specialized form of the limiting variable method that McCleary (1969) described. All land uses except urban and agricultural/woodland in our implementation are "limited" to zero for polygons deemed uninhabitable.

The fundamental operations of the G1 method were adapted to work with polygon (versus gridded) land-use data to create method P1. The polygon binary method also assigns 100 percent of the data in each county to only the urban and agricultural/woodland areas within the county, with population or housing densities recalculated for these new areas.

Polygon and Grid Three-class Methods (P2 and G2)

For the polygon three-class method (P2), we used a weighting scheme to assign population or housing data to three different land-use classes within each county: urban, agricultural/woodland, and forested. For counties with all three land-use classes present, 70 percent of the data for each county was assigned to urban land-use polygons, 20 percent to agricultural/woodland polygons, and 10 percent to forested land. Water received no data. The percentages we selected were a subjective decision partly based on a previous implementation by Holloway et al. (1996) who mapped population density for Missoula County, Montana. Working with census tracts as their choropleth units, they assigned 80 percent of each tract population to urban polygons, 10 percent to open polygons, and 5 percent each to agricultural and forested areas. Our percentage breakdown (70-20-10) was chosen because we estimated that a larger percentage of the mapped variable would be found in agricultural/woodland and forested areas in our study region than for the region mapped by the Montana researchers. For this experiment, adjustments in the weighting percentages were not made for different mapped variables; the same scheme was used for all variables.

A major weakness of the three-class method is that it does not account for the area of each particular land use within each county. A worst-case scenario would result for a county that had only one or two small urban polygons. These polygons would still receive 70 percent of the data for the entire county, causing the urban areas in that county to have unusually high densities and the other land-use areas to have lower densities.

Method G2 represents a grid-based implementation of method P2. As with the polygon version, the method used a weighting scheme to assign data to different land-use classes within each county. For counties with all three inhabitable land uses, 70 percent of county data were assigned to urban grid cells, 20 percent to agricultural/woodland cells, and 10 percent to forested cells. The method suffers from the same weaknesses as the P2 method: subjectivity of the weightings and potential problems with very small areas of a particular land use in a county. As with G1, a major strength of the grid three-class method was its easier implementation in Avenue.

Limiting Variable Method (P3)

The final polygon method is the pure form of the limiting variable method described by McCleary (1969). Wright (1936) also made use of this concept, as did Gerth (1993) and Charpentier (1997) in their GIS implementations. The first step in this method was to assign data by simple areal weighting to all inhabitable polygons in each county. For the total population variable, data were assigned so that the three inhabitable land-use types within a county (urban, agricultural/woodland, and forested polygons) had equal population densities. At this step, water was "limited" to zero density. Next, we set thresholds of maximum density for particular land uses and applied these throughout the study area. For example, agricultural/woodland areas were limited to 50 people per square kilometer and forested areas were assigned a lower threshold of 15 people per square kilometer for the total population variable we mapped. The final step in the mapping process was use of these threshold values to make adjustments to the data distribution within each county. If a polygon density exceeded its threshold, it was assigned the threshold density and the remaining data were removed from the area. These data were distributed evenly to the remaining zones in the county (Gerth 1993).

To decide the upper limits on the densities of the mapped variable for agricultural/woodland and forested land uses, we examined data available at the county level. Eighteen counties in the data-set were classed entirely as agricultural/woodland and we used these to set thresholds. For each of the six variables, the agricultural/woodland threshold was based on the value for the sixth highest county (70th percentile). The forested threshold was chosen using the same list of counties and based on the value of the fourth lowest (20th percentile) county. This approach resulted in different thresholds for each variable—a systematic customization necessitated by differing magnitude ranges among variables. Only one county in the study area was fully forested, yielding insufficient information for the selection of the maximum density threshold for forested polygons.

Calculation and Analysis of Error Measures

A variety of ways of measuring error have been used in areal interpolation research. We chose to follow Fisher and Langford (1995) in their use of root mean squared (RMS) error and a coefficient of variation to describe errors in dasymetric zones. Other researchers have used mean percent error and mean absolute percent error to summarize error (Goodchild et al. 1993). We chose RMS error because it could be applied to count data (e.g., total number of people) and was easily interpreted as a value with the same units as the mapped variable.

We began by calculating the RMS error for each county using errors for dasymetric zones within the county. This calculation had three basic steps:

1. Calculate the square of the difference between the correct and interpolated counts for each dasymetric zone;
2. Calculate the mean of these squared errors within a county; and
3. Take the square root of the mean to produce a summary county error measure, the RMS error.

The correct data values for dasymetric zones (needed for step 1) were calculated by summing population, or housing units, for block groups within each dasymetric zone. The coefficient of variation for each county was calculated by dividing the RMS error by the correct county data value (number of people or housing units in 1990). This standardized value allowed for direct comparison across the six mapped variables.

A repeated-measures analysis of variance (ANOVA) was used to test for significant differences between mean coefficients of variation among the five mapping methods. Coefficients of variation

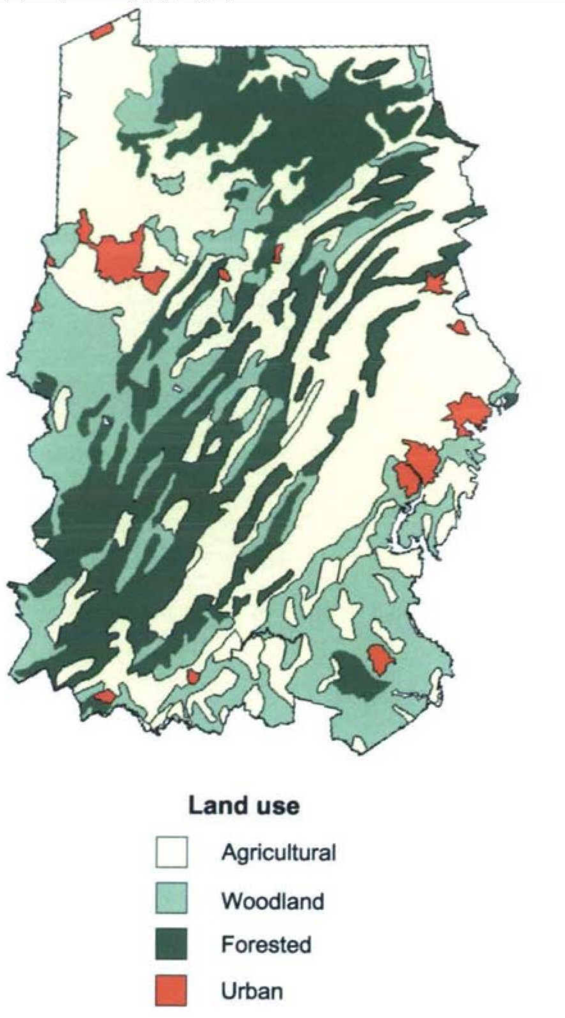

Figure 3. Aggregated land-use data for study area.

were the repeated measures compared within counties (159 counties were the "subjects" in the "within-subjects" parlance of repeated-measures ANOVA).

Results and Discussion

Statistical Analysis of Errors

Table 1 lists means of coefficients of variation for 159 counties for each of 30 maps examined in the analysis. Dasymetric methods are ranked by the right-most column that lists overall means for the rows. P3 had the smallest mean coefficient of variation followed by P2, P1, G2, and G1. This rank order is consistent between variables with a few tied coefficients and only a single pair of methods changing places within any one column. For example, G2 is ranked before P1 in the first two housing variables. P3 is ranked first for all of the variables. This sample of variables contains more housing variables than population variables, but the consistency in rankings of the dasymetric methods within variables indicates that this imbalance in the sample does not bias the overall results.

The ANOVA showed that the dasymetric method had a significant effect overall ($p = 0.000$, $F = 11.164$, $df = 2.451$) based on the Greenhouse-Geiser statistic (SPSS Inc. 1999). Thus, the null hypothesis that there is no difference in the level of error among the methods was rejected. Table 2 lists the mean coefficients of variation for each method (from the right column of Table 1) and shows confidence limits calculated around those means. Significant differences between pairs of mean coefficients of variation were tested using the Bonferroni method, because it accounts for multiple comparisons and has been shown to be useful for large sample sizes (SPSS Inc.). The pairwise results (Table 3) show that P3, the limiting variable method (which had the lowest error for all variables) was significantly different from all other methods. Thus, the limiting variable method (P3) performed well in comparison to the other methods.

Although no other pairings of methods showed significant differences, the three-class methods performed slightly better than the corresponding binary methods (P2 was better than P1 and G2 was better than G1). Though this difference was not significant in our work (Table 3), the result was not consistent with previous research that showed simpler methods performing better than more complex versions (see Eicher 1999, p. 102).

For a straightforward comparison of errors, we adapted G1 and P2 to the alternative data structure (G1 to polygon-based P1 and P2 to grid-based G2). This redundancy was important because the grid methods produced generally higher errors that might have confounded comparison (e.g., G2 had lower errors than G1 but higher error than P1). This approach also allowed us to examine whether there was a significant difference between methods that use polygon land-use data and those methods that use grid land-use data. Methods P1 and G1 and methods P2 and G2 were directly comparable. P1 had lower error than G1 and P2 had lower error than G2, though in both cases the differences were not significant (Table 3).

In sum, the results of the repeated-measures ANOVA supported three conclusions. The traditional limiting variable method (P3) produced dasymetric maps with significantly lower mean error than maps produced using all other methods. Based on pairwise comparison of method P1 to P2

Dasymetric Method	Total Population	Black Population	House Values				Overall mean of c.v. for method
			<$25,000	$40,000 - 49,000	$50,000 - 59,000	$60,000 - 75,000	
P3 (limiting)	0.47	0.64	0.54	0.49	0.45	0.49	0.51
P2 (3-class)	0.62	0.70	0.65	0.64	0.63	0.64	0.65
P1 (binary)	0.62	0.75	0.71	0.67	0.64	0.64	0.67
G2 (3-class)	0.64	0.80	0.67	0.66	0.65	0.66	0.68
G1 (binary)	0.63	0.90	0.72	0.68	0.65	0.80	0.73
Overall mean of c.v. for variable	0.60	0.76	0.66	0.63	0.61	0.65	0.65

Notes: Entries in the center portion of the table are means of 159 county coefficients of variation (c.v.) for 30 individual maps. Values in the right-most column are overall means of c.v. of each dasymetric method, and values in the bottom row are overall means of c.v. of each variable.

Table 1. Summary of county-level coefficients of variation of RMS error measures.

and method G1 to G2, the three-class methods were found to perform slightly better than the binary methods, though the overall difference was not significant. Based on pairwise comparison of method P1 to G1 and method P2 to G2, polygon methods were found to produce slightly more accurate maps than grid methods, though the difference was not significant.

Method	Mean c.v.	95% Confidence Interval
P3 (limiting)	0.51	0.43 - 0.59
P2 (3-class)	0.65	0.55 - 0.74
P1 (binary)	0.67	0.58 - 0.77
G2 (3-class)	0.68	0.58 - 0.79
G1 (binary)	0.73	0.64 - 0.82

c.v. = coefficient of variation of RMS county errors.

Table 2. Estimated mean errors and confidence intervals for dasymetric methods.

	P3	P2	P1	G2
P3				
P2	0.001*			
P1	0.000*	1.000		
G2	0.000*	0.313	0.059	
G1	0.000*	0.609	1.000	1.000

* significant differences.

Table 3. Probabilities for differences in errors for pairs of dasymetric methods.

Visual Analysis of Dasymetric Maps and Errors

Many of the papers we reviewed for this work, especially research on areal interpolation, lacked a visual presentation of results. Papers with maps (e.g., Langford et al. 1990; Fisher and Langford 1996; Charpentier 1997) typically focused on small study areas. We now present dasymetric maps resulting from the three polygon methods tested in this experiment (see also Eicher (1999) for corresponding grid-based maps).

Some of the work in areal interpolation sought to describe the types of error produced (Cockings et al. 1997; Langford et al. 1991). For example, these studies formally tested the connections of geometry and attribute characteristics to map error. While this section will not formally test these relationships, error maps will be presented, allowing visual analysis of spatial error patterns. Presentation of this type of data in tandem with the dasymetric maps may generate hypotheses and ideas for future research.

Example Map Series

Figures 4, 5, and 6 present (a) dasymetric maps, (b) error maps based on percent error, and (c) error maps based on count error for each of the three polygon mapping methods. The dasymetric maps are density representations of the total population variable. Boundaries for all dasymetric zones are shown on these maps. Percent error maps show error relative to the total population of a polygon. One possible weakness of percent error maps is that rural areas generally will have higher percent errors because of lower totals within these zones. Furthermore, when presented in map form, these higher errors dominate, because rural map zones tend to be larger than urban map zones.

Count errors are presented to provide a view of the absolute error present across the study area (map c in Figures 4, 5, and 6). Maps of count error will not signal high errors in rural areas with low populations as seen on the percent error maps. When count errors are mapped, the reader gets more of an impression of the overall quantity of the variable that has been mapped inaccurately. A major weakness of presenting count data with areal fills, however, is that larger polygons will tend to have higher values than smaller polygons with similar characteristics. Despite this recognized problem with mapping counts, consistent symbolization was maintained to aid comparison with the other two map types (a and b) within the figures.

Data ranges for classifications are shared among corresponding maps in Figures 4, 5, and 6 to aid comparison. The extreme classes have varied maximums because the calculated maximum population density varied widely depending on the mapping method. Note that adjacent classes in legends share data breaks, namely, the upper value for the first class (25 persons/km^2) is the same as the lower value for the second class; the upper value for the second class (75 persons/km^2) is the same as the lower value for the third class, etc. Each class should be interpreted as encompassing all values up to but not including the highest value listed in the legend for the class. This method of describing class breaks allows precise definition of breaks without the use of excessive precision in significant digits that clutter a legend and make map use difficult (Monmonier 1982).

Binary Method (P1)

The dasymetric map in Figure 4a clearly shows the binary distinction made in this method between inhabitable and uninhabitable areas. All areas designated as forested (a large portion of the map) have been assigned zero population. The distinction is most obvious in the ridge-and-valley physiographic province running diagonally across the study region, where ridges appear as white (forested areas that are assigned no population) and valleys are yellow-greens (populated urban and agricultural/woodland areas).

The percent error map (Figure 4b), as expected, shows under-prediction (grays) for the forested areas, although urban areas have also been under-predicted. The latter occurs because urban areas are not distinguished from agricultural/woodland areas by the method, receiving the same population density when they are in the same county. Many agricultural/woodland areas are over-predicted for the same reason, especially in areas bounding major metropolitan areas (Washington, Baltimore, Pittsburgh). The count error map (Figure 4c) shows slightly different patterns of error. Most interesting are the relatively low errors in rural, forested areas. Though all these areas have –100 percent error (because they were assigned "zero" population), counts are comparatively low because populations are low.

Three-Class Method (P2)

The dasymetric map in Figure 5a shows that nearly all areas received population, with the exception of water. The map also distinguishes between urban and agricultural/woodland, and thus urban places like York, Erie, and Harrisburg (see Figure 1 for locations) have visibly higher populations than the surrounding rural areas.

The percent error map (Figure 5b) shows that urban areas are either predicted correctly or under-predicted (e.g., Pittsburgh, Washington, Baltimore). This map also shows that forested areas are under-predicted, especially in the ridge-and-valley province. Conversely, the valleys, typically classified as agricultural/woodland, are over-predicted. The under-prediction of forested areas in most counties suggests that not enough population was assigned to the forested land use in the 70-20-10 percent weighting scheme. Again, the map of count error (Figure 5c) shows lower errors throughout the ridge-and-valley province, because populations are lower in the mostly rural area.

Limiting Variable Method (P3)

The dasymetric map produced using the P3 method (Figure 6a), which had the lowest overall error, closely resembles the result from method P2, which had the next best error measures. Slight differences in dasymetric map patterns for the two methods are noticeable in northwest Pennsylvania, along the top edge of the study area, as well as in central Virginia, along the southern edge of the map. The limiting variable method is also similar to method P2 in that few areas of the map have zero population.

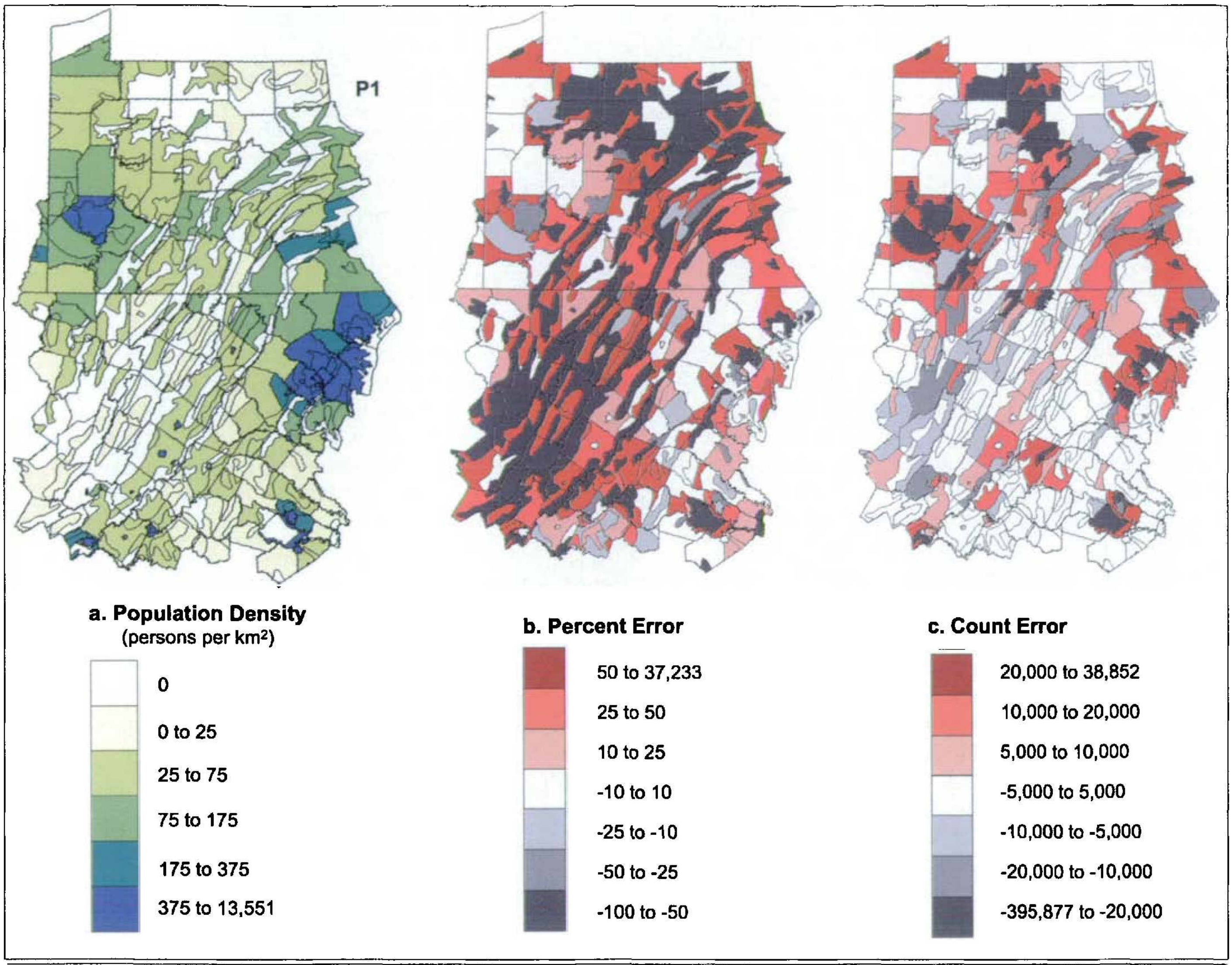

Figure 4. Maps for binary method (P1): a. dasymetric population density map; b. percent error; and c. count error in number of persons.

The pattern of error for the P3 method (Figures 6b and 6c) is different, however, from the first two methods. An interesting radial pattern is evident around both Washington and Baltimore. These areas are unique because they have large metropolitan populations and the municipal boundaries of each designate them as independent cities, essentially small counties in this analysis, so these central cities are well predicted. Immediately outward, however, are over-predicted areas, followed by a ring of under-predicted areas at the fringes of the metropolitan areas. The boundary between over-prediction and under-prediction is the functional border between urban and agricultural/woodland land use. The Pittsburgh area has a similar pattern, with a central urban area that is well predicted, surrounded by areas of agricultural/woodland land use that are under-predicted.

For major metropolitan areas, we conclude that the threshold value for agricultural/woodland areas was set too low. As a result, in the case of Washington, the outermost suburbs (classified as agricultural/woodland) received the threshold density of 15 persons per km^2, while too much data was then carried over and assigned to more truly suburban areas (classified as urban) closer to the city center. In the case of Pittsburgh, a smaller city, suburbs immediately proximate to the city were classified as agricultural/woodland, assigned the corresponding agricultural/woodland density threshold, and therefore under-predicted. The age of the land-use data could also be a factor in both cases, because areas classified as agricultural/woodland may have been developed between 1970 and 1990.

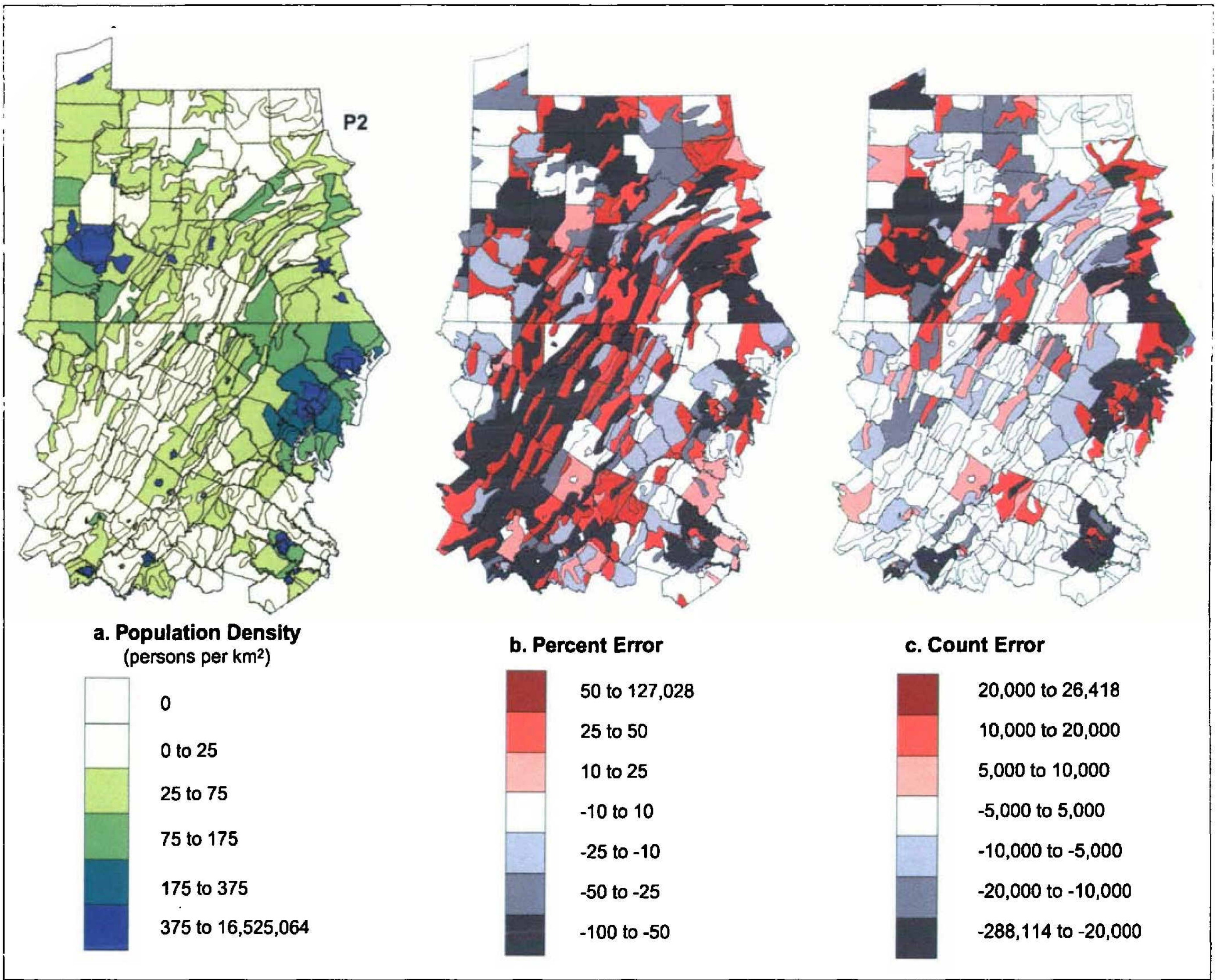

Figure 5. Maps for three-class method (P2): a. dasymetric population density map; b. percent error; and c. count error in number of persons.

Conclusions and Future Research

Research Question A: Mapping Methods

There were significant differences in accuracy for the five dasymetric mapping methods tested in this experiment. The limiting variable method (P3) produced maps with significantly lower error than maps made using the other methods. Three-class methods (P2 and G2) produced maps with errors that were not significantly different than the errors of maps produced with binary methods (P1 and G1).

Though the traditional limiting variable method (P3) is over a half-century old (Wright 1936), we have shown that it can be implemented in a GIS to produce accurate dasymetric maps. The success of this method may, in part, be due to its customized approach. Threshold values used to shift data between zones were based on the data distribution for each mapped variable. This approach contrasts with three-class methods P2 and G2, in which the same 70-20-10 percentage weightings were applied to all variables. Future work could examine the effect of using different threshold values for the land-use classes. Also important are methods of determining the P3 threshold values. Perhaps the reassignment schemes used for P3 could also be made less arbitrary and more specific to the geography of each source zone by making the thresholds dependent on the land-use areas within each source zone (Mennis 1999). For example, a threshold of 20 people per square kilometer used for agricultural/woodland areas could be increased to 30 in counties with more than 50 percent urban land uses. In the future, a standardized and generalizable decision process with a statistical basis might be developed for this purpose.

Research Question B: Polygon Methods versus Grid Methods

Based on a pairwise comparison of polygon and grid versions of the binary method (P1 and G1) and the three-class method (P2 and G2), errors for polygon dasymetric maps were not significantly different than for grid maps. In our implementation, the use of grid methods was more convenient. Within the ArcView environment, grid (raster) versions of our datasets took up less storage space, and the grid methods were easier to implement using Avenue scripting. Processing was also faster in ArcView because shorter scripts were typical of the grid methods (Eicher 1999). Grid methods were developed in the areal interpolation literature but have not been popular for dasymetric mapping. Their effectiveness in producing dasymetric maps in this experiment, as well as their ease of implementation, suggests that they should be used more frequently. In contrast, advantages of the equally accurate polygon methods include greater flexibility for publication-quality mapping, such as ease of adjusting line weights to suit different presentation sizes.

Future work could study the effect of grid cell size on the accuracy of dasymetric maps. The relationship between cell size and map scale is especially interesting. A research project focusing on grid dasymetric maps could determine optimal grid cell sizes for varied map scales. For example, we expect that use of excessively small grid cells for a given scale would lead to higher map error. Such research would also provide an analytical method of balancing increases in accuracy afforded by smaller grid cells with the increased storage requirements and processing times associated with larger data files. See Eicher (1999) for further discussion of this topic.

Research Question C: Dasymetric Maps and Error Maps

Visual inspection of dasymetric maps and error maps supported the results from the statistical analysis portion of the experiment and offered additional insight into the usefulness of dasymetric maps. It is our opinion that the resulting dasymetric maps are visually effective, showing more accurate distributions than county aggregations provide.

Classification (the data ranges represented by map colors) for dasymetric maps was not closely examined in this work. Future work could investigate whether classifications that have been shown to work well for choropleth maps also work well for dasymetric maps. An interesting possibility is the use of contiguity-based classifications (Cromley 1996) which might make dasymetric maps visually "more dasymetric." Contiguity-based classifications consider differences across boundaries between adjacent polygons, and this strategy is particularly well suited for application to boundaries on dasymetric maps. Dasymetric boundaries are more closely related to the underlying data distribution than boundaries on choropleth maps, because dasymetric zones are more internally homogeneous.

Many possibilities exist for extending calculation and comparison of error in dasymetric maps. Future work could include a number of between-county factors or covariates in the repeated-measures ANOVA. These could test the relationship between map error and land-use areas, population density, geographic location, and other factors. Also, different approaches could be taken to evaluating the quality of dasymetric maps in an experiment. A defining property of dasymetric maps is homogeneity within map zones; however, this property was not formally tested in this research. It may be possible to compute the variance for each map zone using values from the block groups comprising that zone. Variances within each map zone could then be compared to variances between map zones to establish a measure of "dasymetricity" (homogeneity within zones combined with distinctness of differences between adjacent zones). To further examine dasymetric mapping methods and the properties of dasymetric maps, a number of different map zone configurations could be used. For example, maps resulting from different ancillary variables such as land use, soil types, or slope zones could be compared. Likewise, maps resulting from different source zones such as counties, voting districts, and economic planning areas could be compared.

Summary

With a structured experiment, we determined that the traditional limiting variable method was the most accurate of those tested and that grid dasymetric mapping methods produced maps of similar accuracy to their polygonal counterparts. Wright's limiting variable concept, developed 65 years ago, has maintained its utility. This traditional method has been updated and adapted successfully to produce dasymetric maps in a GIS mapping environment. Additionally, the availability of GIS technology has made data processing for dasymetric map production a more pragmatic process. Work (such as this paper) on dasymetric mapping, research on areal interpolation, and increased use of GIS by cartographers are first steps toward producing a formal set of dasymetric mapping methods that updates George McCleary's 32-year-old typology. GIS software packages may eventually offer full dasymetric mapping functionality or at least provide the means to make these types of maps more easily.

ACKNOWLEDGEMENTS

Special thanks must be given to Jeremy Mennis and Donna Peuquet for their comments and constructive criticism on this research. We appreciate Terry Slocum's careful edits of the manuscript.

REFERENCES

Anderson, J. R. 1967. Major land uses in the United States. In: U.S. Geological Survey. *National atlas of the United States of America*. U. S. Geological Survey, Washington, D.C., pp. 158-9.

Bloom, L. M., P. J. Pedler, and G. E. Wragg. 1996. Implementation of enhanced areal interpolation using Mapinfo. *Computers & Geosciences* 22(5): 459-66.

Charpentier, M. 1997. The dasymetric method with comments on its use in GIS. In: *1997 UCGIS Annual Assembly & Summer Retreat*. The University Consortium for Geographic Information Science. June 15-20, Bar Harbor, Maine.

Chrisman, N. R. 1997. *Exploring geographic information systems*. New York, New York: John Wiley & Sons.

Chrisman, N. R. 1998. Rethinking levels of measurement for cartography. *Cartography and Geographic Information Systems* 25(4): 231-42.

Cockings, S., P. Fisher, and M. Langford. 1997. Parameterization and visualization of the errors in areal interpolation. *Geographical Analysis* 29(4): 314-328.

Cromley, R. G. 1996. A comparison of optimal classification strategies for choroplethic displays of spatially aggregated data. *International Journal of Geographical Information Systems* 10(4): 405-24.

DeMers, M. N. 1997. *Fundamentals of geographic information systems*. New York, New York: John Wiley & Sons.

Eicher, C. L. 1999. Implementation and evaluation of dasymetric mapping methods. M.S. Thesis. Pennsylvania State University.

Fisher, P., and M. Langford. 1995. Modeling the errors in areal interpolation between zonal systems by Monte Carlo simulation. *Environment and Planning A* 27: 211-24.

Fisher, P., and M. Langford. 1996. Modeling sensitivity to accuracy in classified imagery: A study of areal interpolation by dasymetric mapping. *Professional Geographer* 48(3): 299-309.

Flowerdew, R., and M. Green. 1989. Statistical methods for inference between incompatible zonal systems. In: M. Goodchild and S. Gopal (eds), *Accuracy of spatial databases*. London, U.K.: Taylor and Francis. pp. 239-47.

Flowerdew, R., and M. Green. 1991. Data integration: Statistical methods for transferring data between zonal systems. In: I. Masser and M. B. Blakemore (eds), *Handling geographical information*. Essex, U.K.: Longman Scientific & Technical. pp. 38-53.

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

Flowerdew, R., and M. Green. 1994. Areal interpolation and types of data. In: S. Fotheringham and P. Rogerson (eds), *Spatial analysis and GIS*. London, U.K.: Taylor and Francis. pp. 121-45.

Flowerdew, R., M. Green, and E. Kehris. 1991. Using areal interpolation methods in GIS. *Papers in Regional Science: The Journal of the RSAI* 70(3): 303-15.

Gerth, J. 1993. Towards improved spatial analysis with areal units: The use of GIS to facilitate the creation of dasymetric maps. Masters Paper. Ohio State University.

Goodchild, M. F., and N. S. Lam. 1980. Areal interpolation: A variant of the traditional spatial problem. *Geoprocessing* 1: 297-312.

Goodchild, M. F., L. Anselin, and U. Deichmann. 1993. A framework for the areal interpolation of socioeconomic data. *Environment and Planning A* 25: 383-97.

Holloway, S. R., J. Schumacher, and R. Redmond. 1996. People and place: Dasymetric mapping using Arc/Info. Missoula: Wildlife Spatial Analysis Lab, University of Montana.

Hitt, K. J. 1991. *Digital map file of major land uses in the United States*. U.S. Geological Survey. Accessed October 1998. ArcInfo coverage available from http://water.usgs.gov/GIS/metadata/usgswrd/landuse.html.

Langford, M., D. J. Maguire, and D. J. Unwin. 1990. Mapping the density of population: Continuous surface representations as an alternative to choropleth and dasymetric maps. Leicester, U.K.: Midlands Regional Research Laboratory, University of Leicester and Loughborough, University of Technology, U.K.

Langford, M., D. J. Maguire, and D. J. Unwin. 1991. The Areal interpolation problem: Estimating population using remote sensing in a GIS framework. In: I. Masser and M. B. Blakemore (eds), *Handling geographic information*. Essex, U.K.: Longman Scientific & Technical. pp. 55-77.

Langford, M., and D. J. Unwin. 1994. Generating and mapping population density surfaces within a geographical information system. *The Cartographic Journal* 31: 21-6.

MacEachren, A. M. 1994. *Some truth with maps: A primer on symbolization and design*. Washington, D.C.: Association of American Geographers.

McCleary, G. F., Jr. 1969. The dasymetric method in thematic cartography. Ph.D. Dissertation. University of Wisconsin at Madison.

McCleary, G. F., Jr. 1984. Cartography, geography, and the dasymetric method. In: *Proceeding, 12th Conference of International Cartographic Association*, August 6-13, Perth, Australia. 1:599-610.

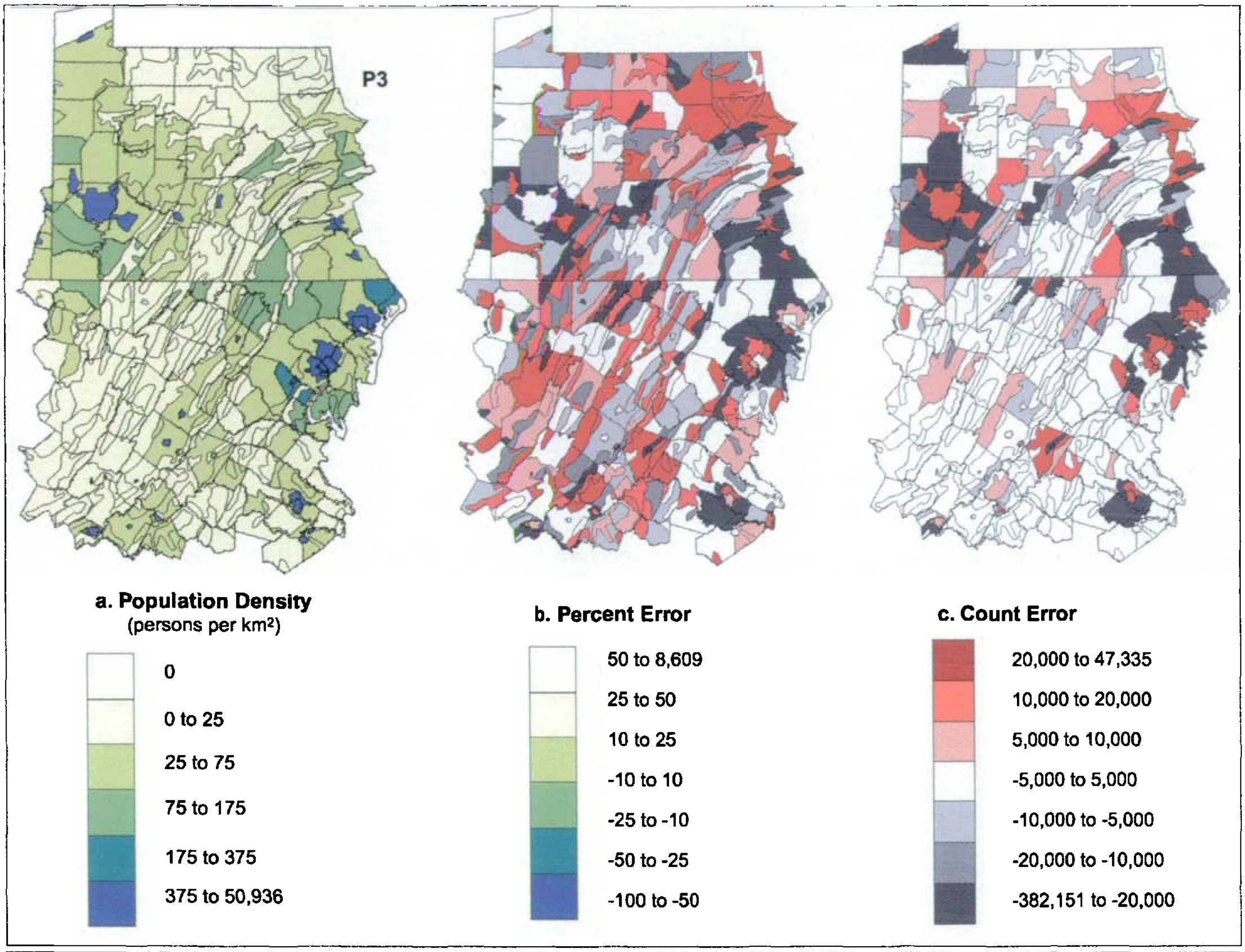

Figure 6. Maps for traditional limiting variable method (P3): a. dasymetric population density map; b. percent error; and c. count error in number of persons.

Mennis, J., 1999. Using GIS, spatial statistics, and visualization to investigate environmental justice. Paper presented at the Joint Statistical Meetings, American Statistical Association, Baltimore, Md, August 8–12, 1999.

Monmonier, M. 1982. Flat laxity, optimization, and rounding in the selection of class intervals. *Cartographica* 19(1): 16-27.

Monmonier, M., and G. A. Schnell. 1984. Land use and land cover data and the mapping of population density. *International Yearbook of Cartography* 24: 24-9.

Moxey, A., and P. Allanson. 1994. Areal interpolation of spatially extensive variables: A comparison of alternative techniques. *International Journal of Geographical Information Systems* 8(5): 479-87.

Robinson, A., J. Morrison, P. Muehrcke, A. Kimerling, and S. Guptill. 1995. *Elements of cartography*, 6th ed. New York, New York: John Wiley & Sons.

Slocum, T. A. 1999. *Thematic cartography and visualization.* Upper Saddle River, New Jersey: Prentice Hall.

SPSS Inc. 1999. *SPSS help, version 9.0.* SPSS Incorporated, Chicago, Illinois.

U.S. Geological Survey. 1970. *National atlas of the United States of America.* U.S. Geological Survey, Washington, D.C.

U.S. Geological Survey. 1993. National water summary 1990-91: Hydrologic events and stream water quality. U.S. Geological Survey water supply paper 2400. U.S. Geological Survey, Washington, D.C.

Wright, J. K. 1936. A method of mapping densities of population. *The Geographical Review,* pp. 103-10.