Comparison of Dasymetric Mapping Techniques for Small-Area Population Estimates

Paul A. Zandbergen and Drew A. Ignizio

ABSTRACT: Dasymetric mapping techniques can be employed to estimate population characteristics of small areas that do not correspond to census enumeration areas. Land cover has been the most widely used source of ancillary data in dasymetric mapping. The current research examines the performance of alternative sources of ancillary data, including imperviousness, road networks, and nighttime lights. Nationally available datasets were used in the analysis to allow for replicability. The performance of the techniques used to examine these sources was compared to areal weighting and traditional land cover techniques. Four states were used in the analysis, representing a range of different geographic regions: Connecticut, New Mexico, Oregon, and South Carolina. Ancillary data sources were used to estimate census block group population counts using census tracts as source zones, and the results were compared to the known block group population counts. Results indicate that the performance of dasymetric methods varies substantially among study areas, and no single technique consistently outperforms all others. The three best techniques are imperviousness with values greater than 75 percent removed, imperviousness with values greater than 60 percent removed, and land cover. Total imperviousness and roads perform slightly worse, with nighttime lights performing the worst compared to all other ancillary data types. All techniques performed better than areal weighting.

KEYWORDS: Dasymetric mapping, population, imperviousness, land cover, roads

Introduction

Demographic information available through the Census Bureau is aggregated using census enumeration units, including blocks, block groups, and tracts. The smallest unit is represented by blocks. In urban areas this typically corresponds to about a city block or smaller. The demographic information available at the level of blocks is limited: population, households, race, Hispanic origin, gender, and age. More detailed demographic and socioeconomic information is available at the level of block groups, which are the aggregation of a number of blocks into a larger unit, typically several dozen blocks, with a total population size of several hundred to several thousand. Block groups are further aggregated into tracts—a tract typically consists of between one and five block groups but may contain up to nine block groups. Delineation of census enumeration areas is partly driven by trying to obtain relatively consistent population counts within each unit. For example, population counts for census tracts range from 1500 to 8000, with an average of about 4000 people. As a result, census units in urban areas are much smaller compared to rural areas (U.S. Census Bureau 2000).

The aggregation of census data to these enumeration units represents a challenge when trying to compare census data to other boundaries. Many other commonly used boundaries such as neighborhoods and police beats are not delineated with census boundaries in mind, resulting in a spatial mismatch between boundaries. This spatial mismatch is also referred to as "spatial incongruity" (Voss et al. 1999). Estimating the population (and its associated demographic characteristics) within these boundaries can be problematic due to this spatial incongruity.

This problem is not limited to administrative units or other socio-economic boundaries but also applies to natural boundaries. Natural science data works with geographic units defined by such features as hydrology, land use, soil type, and others which do not correspond to census enumeration units. The lack of a common set of boundaries is one of the obstacles to the integration of social

Paul Zandbergen, Associate Professor, University of New Mexico, Department of Geography, Bandelier West Room 111, MSC01 1110, 1 University of New Mexico, Albuquerque, New Mexico 87131. E-mail: <zandberg@unm.edu>. **Drew A. Ignizio**, Department of Geography, University of New Mexico. E-mail: <drew.ignizio@gmail.com>.

and natural science. The spatial mismatch may also result from temporal differences. For example, census data are typically collected every 10 years. From census to census, enumeration area boundaries change as a result of changes in population densities. Temporal differences in census area boundaries are a persistent problem (Gregory 2002) which limits reliable analysis of long-term changes using fine-scale spatial units. Research in several areas has documented the challenges of working with mismatched boundaries, including environmental justice (Higgs and Langford 2009), flood hazards (Maantay and Maroko 2009; Patterson and Doyle 2009), air pollution (Maantay et al. 2008), and time-series demographic analysis (Gregory and Ell 2006).

Spatially mismatched boundaries between geographic data sets present a persistent problem in geography, planning, regional science, landscape ecology, and other fields. A number of approaches have been developed to spatially allocate attributes between sets of spatially mismatched boundaries. One of these approaches is dasymetric mapping which is the process of disaggregating spatial data into finer units of analysis using ancillary data to help refine locations of population of other phenomena (Wright 1936; Mennis 2003). The purpose of this paper is to compare a number of different types of ancillary data for use in dasymetric mapping, including imperviousness, road density, and nighttime lights. The performance of dasymetric techniques using these types of ancillary data is compared to that of areal weighting and traditional land cover-based dasymetric techniques. National datasets are used to ensure results are replicable across the entire United States.

Background

Several approaches have been developed to address the problem of spatial incongruity between sets of boundaries. These vary in complexity and in how they address the spatial mismatch. The most basic approaches employ rules of inclusion or exclusion based on the boundaries of the geographic units. For example, in the case of "centroid containment," census enumeration areas (source areas) are represented by their centroids. All source area centroids falling inside a particular polygon (target area) are then assigned to that polygon. If the source areas are relatively small compared to the target areas, this can be a sufficiently robust technique, but if the source and target areas are approximately similar in size, this approach will result in very large errors.

More advanced methods which go beyond simple inclusion/exclusion rules are commonly referred to as areal interpolation (Goodchild and Lam 1980). Areal interpolation describes a variety of methods which generally employ weights based on the area of intersection between source and target areas in order to allocate characteristics from the source areas to the target areas. The most basic form of areal interpolation is "areal weighting" in which the population within the source areas is spatially apportioned into the target areas based on how much of each source area falls within each target area (Flowerdrew and Green 1992). This effectively assumes that the population density is uniform within the source areas. In a GIS environment, areal weighting can easily be accomplished with a polygon overlay operation. This approach works for source and target areas of any size, but its performance is very dependent on the assumption of uniform population density. Without any additional information, however, it is the most logical technique.

Alternative methods have been developed to improve upon the areal weighting technique. The first set of methods can be referred to as "surface fitting." In this approach, a surface is fitted to the data in the source areas, and this surface is used to interpolate values for the target areas. Fitting a surface to the data typically employs inferential statistics and a number of different approaches have been developed and tested (Bracken 1991; Bracken and Martin 1989; Martin 1996; Tobler 1979).

A second set of methods is referred to as "zonal methods." Several approaches that fall under this category employ ancillary information from the target areas or from an external set of areas in the areal interpolation process. For example, Goodchild et al. (1993) employed a set of external control zones with an assumed uniform population distribution. Population is allocated from the source areas to the control zones using a variety of regression-based techniques. Then, it is allocated from the control zones to the target areas based on areal weighting. In another example, Langford et al. (1991) used linear regression models to interpolate population based on land-use zones. Flowerdrew and Green (1994) adopted an iterative expectation/maximization (E/M) algorithm originally designed to estimate missing data (Dempster et al. 1997) for use in areal interpolation. Finally, Martin (2003) developed an iterative automatic zoning algorithm based on Openshaw's (1977) automatic zoning procedures. This method employs an intermediate layer of boundary processing and attempts to minimize the mismatch between two sets of boundaries.

A third set of methods is referred to as "dasymetric mapping." In this approach, ancillary data are used to gain information about the distribution of population within the source areas. Dasymetric mapping is the process of disaggregating spatial data into finer units of analysis using ancillary data to help refine locations of population of other phenomena (Mennis 2003). The dasymetric map was conceived as a type of thematic map during the early to mid nineteenth century. While dasymetric mapping has been a well established cartographic technique for many years, in recent years it has gained interest as an approach to estimate populations for small areas, and to improve upon the assumptions made in areal weighting (Eicher and Brewer 2001). A comprehensive review of dasymetric mapping is provided by Mennis (2009).

In its most basic form, dasymetric mapping distributes population (or another variable) within a polygon using ancillary data to provide finer units of analysis. This ancillary data most often consists of land cover data. The actual mapping procedure consists of an overlay between population data in the form of polygons and land cover data. This yields a set of dasymetric zones, nested within both the population polygons and the land cover data. Population is then spatially apportioned from the population polygons to the dasymetric zones based on the relationship between land cover and population density. The resulting dasymetric map can be used for areal interpolation to a set of target areas. To accomplish the areal interpolation the dasymetric zones are overlaid with the target areas and population is spatially apportioned based on areal weighting. This assumes the population density within each dasymetric zone is uniform, but since these zones are typically much smaller than the source areas, the result is a more accurate estimate of the population in the target areas compared to the estimate based on areal weighting without ancillary data.

While dasymetric mapping techniques represent one approach to address the problem of spatial incongruity between sets of boundaries, it can also be employed for other applications. For example, in regions of the world where census data are very coarse, ancillary data can be used to develop more fine-grained estimates of population distribution. A good example of this is the LandScan dataset which provides a worldwide 1-km population grid in which census counts are apportioned to each grid cell based on proximity to roads, slope, land cover, nighttime lights, and other information (Dobson et al. 2000; Bhaduri et al. 2007).

In creating the dasymetric map, the relationship between land cover and population density can take one of several forms. The simplest form is the use of a Boolean style mask which identifies areas where population density is zero; for example, a data layer of water bodies. This is referred to as binary dasymetric mapping (e.g., Langford 2007). While conceptually simple, binary dasymetric mapping can result in substantial improvements in population estimates relative to areal weighting (Langford 2007). When more detailed land cover data are employed, multiple population density categories are used instead of a Boolean mask, although one of these categories is often still exclusionary (i.e., a population density of zero). In a typical dasymetric mapping scenario land cover would be classified into classes of low, medium and high population density in addition to a category of zero population. The population density of each land cover type can be assumed *a priori*, but more commonly these values are estimated using the source areas and the land cover data. Training areas are identified where one of the land cover classes is dominant (e.g., a single land cover type accounts for a minimum of 70 or 80 percent of the total surface of a source area). The population density of a particular land cover class is then derived from these training areas, and the values are used in the redistribution of population from source areas to dasymetric zones. In effect this means that the population density estimates for each land cover type are used as relative weights in the redistribution of population. This approach ensures that the total number of people within each source area remains the same, which is referred to as the *pycnophylactic property* of dasymetric maps (Tobler 1979).

Land cover has been the most widely used type of ancillary data for dasymetric mapping (e.g., Eicher and Brewer 2001; Holt et al. 2004; Langford 2006; Mennis 2003; Mennis and Hultgren 2006; Reibel and Agrawal 2007). In theory, however, any spatial data type that correlates with population density could be used. Street networks have received some attention (Mrozinski and Cromley 1999; Reibel and Buffalino 2005; Xie 1995) as well as parcel-based land use information (Maantay et al. 2007; Maantay et al. 2008). Despite the relatively well established methods for dasymetric mapping, a number of other ancillary data types remain unexplored. The purpose of this paper is to employ alternative ancillary data types and determine how well they perform compared to more established methods. Specifically, the following ancillary data types will be tested: imperviousness, road density, and nighttime lights.

Imperviousness has been firmly established as a robust and meaningful measure of urban development (e.g., Schueler et al. 2009; Theobald et al. 2009). In recent years it has in fact emerged as a promising measure of the global human footprint (Sutton et al. 2009). Imperviousness correlates strongly with population density and can be measured relatively easily. It is typically derived from land cover maps or directly from the original source imagery. One of the potential benefits of imperviousness is that it provides greater detail than traditional land cover categories. Since it is typically recorded as a percentage (of total area) it also does not require the same type of calibration that land cover data requires. Road density is also a well established measure of urban development and has been used as a source of ancillary data in dasymetric mapping. Similar to imperviousness, road density (in km/km^2) provides an easily measurable quantity which does not require extensive calibration. Voss et al. (1999) used both total road segment length and the count of road network nodes in dasymetric mapping. Finally, nighttime lights have received much attention as an alternative way to characterize population distribution (e.g., Sutton et al. 2001; Sutton et al. 2003). The intensity of nighttime lights corresponds closely to the concentration of urban centers, but it is also correlated to economic wellbeing since more affluent areas are likely to produce more nighttime lights (per person) than poorer areas.

Areal interpolation methods that result in the least amount of error are preferred. A number of factors are expected to influence the errors in areal interpolation. First, the relative size of the source and target areas will dictate to some extent the magnitude of errors. For example, it has been well established that interpolating from small source areas to large target areas introduces relatively small amounts of error (Fisher and Langford 1995; Sadahiro 2000). Conversely, interpolating from large to small source areas will typically introduce much larger errors. Second, the spatial organization of source and target areas relative to each other can influence the amount of error. For example, relatively modest changes in census unit boundaries over time may result in small errors in areal interpolation, but comparing census units and watersheds is likely to result in large errors. Simpson (2002) refers to this as the degree of fit; the greater the similarity between source and target areas, the lower the error. Third, the quality of the ancillary data will influence the amount of error. In the case of land cover, for example, the resolution and number of land cover categories are likely determinants of the performance of areal interpolation. Given the influence of these three factors, it is not possible to quantify the "typical" error introduced by areal interpolation methods. Studies that have evaluated the performance of areal interpolation methods, therefore, have compared the results to those for areal weighting, i.e., they should perform better than areal weighting for a given scenario.

Several other observations are in order with respect to the performance of areal interpolation methods. First, methods that preserve the *pycnophylactic* properties of the data are preferred, since this provides population distributions of greater internal consistency. Second, it should be recognized that all methods have errors and that their performance may vary with conditions. No single best method has emerged from the research so far, and the best approach for a particular set of source and target areas is likely to vary with specific circumstances and the intended uses of the estimates for the target zones (Gregory 2002).

The current study compares a number of different types of ancillary data for use in dasymetric mapping. Ancillary data types include imperviousness, road density, and nighttime lights. All methods using ancillary data are expected to outperform areal weighting, but the relative strengths of these various data sources are largely unknown. The performance of each technique is determined by using tract-level populations to estimate block group-level populations which can then be compared to the known block group populations.

In addition to comparing multiple techniques which have not previously been tested, the current study also seeks to contribute to the body of knowledge on dasymetric mapping by employing datasets which are available at the national level. Highly detailed information (e.g., parcels) may result in more accurate population estimates, but the results may not be replicable across the entire United States. By using national level datasets the results should be insightful for any jurisdiction in the nation.

Methods

Four U.S. states were selected for the analysis: Connecticut, New Mexico, Oregon, and South Carolina. These four states represent a range of different eco-regions and population densities.

State	Population	Area (km²)	Population Density (#/km²)	Census Tracts
Connecticut	3,405,565	12,885	264.30	816
New Mexico	1,819,046	314,915	5.78	456
Oregon	3,421,399	251,424	13.61	755
South Carolina	4,012,012	80,168	50.04	867

Table 1. Population characteristics of the four study states.

Total population counts for the year 2000 were obtained from the U.S. Census for these four states at the tract and block group level. Table 1 provides a brief summary of the population characteristics of the four states.

Following sections outline the sources and processing steps for each of the ancillary data sources. All data were projected in a U.S. Contiguous Albers Equal Area projection prior to analysis.

Land cover data were obtained from the 2001 National Land Cover Dataset (NLCD). This data consists of a nationwide 30-m grid. The land cover data includes 30 different categories for the continental U.S., including four different urban categories (developed open space, developed low intensity, developed medium intensity, and developed high intensity). Details on the land cover data are reported in Homer et al. (2007). Figure 1A shows an example of the land cover data for Eugene, Oregon.

Dasymetric mapping using land cover data employed the general methods as described by Mennis (2003). This consisted of reclassifying the land cover grid into fewer population density categories. Initially, four different density categories were employed (based on the four urban categories), and all other categories were assigned zero population. Calibration of the land cover model was accomplished for each state independently, using a 70-percent threshold for land cover dominance. In other words, only census tracts where one population density category covered at least 70 percent of the population areas were selected as calibration areas. In each state this resulted in at least one category with too few calibration areas (i.e., fewer than two). In order to achieve a sufficient number of calibration sites some of the four original density categories had to be merged together. All possible combinations were explored to determine which combination preserved as much as possible of the original detail in the land cover information while at the same time providing sufficient calibration areas for each population density category. In each of the four states the two urban land cover types with the highest densities (i.e., developed medium density and developed high density) had to be combined in order to accomplish this. The final classifications are shown in Table 2. While the optimal classification turned out to be identical between the states, the calibration step was carried out separately for each state. Several alternatives were tried, including assigning agriculture its own non-zero population density, but these scenarios resulted in larger errors in block group level population estimates and the results of these scenarios are therefore not reported. Figure 1B shows an example of the reclassified land cover data using the final categories as shown in Table 2.

Original NLCD Land Cover Class	Population Density Category
Developed open space	Low
Developed low intensity	Medium
Developed medium intensity	High
Developed high intensity	High
All others	Zero

Table 2. Reclassification of National Land Cover dataset (NLCD) land cover classes into population density categories.

Imperviousness data were obtained from the same 2001 National Land Cover Dataset (NLCD). This data consists of a nationwide 30-m grid, with cell values between zero and 100 indicating the percentage of imperviousness. The imperviousness information was derived from Landsat imagery. Details on the methodology are reported in Yang et al. (2003). Figure 2A shows an example of the imperviousness data for Eugene, Oregon.

The imperviousness grid was used to estimate the total amount of impervious area (in km²) for each tract and block group. To redistribute the population from tracts to block groups, these total amounts were used as weight factors. An example of this is shown in Figure 3. Implicitly this assumes a linear relationship between percent imperviousness and population density.

Four different versions of imperviousness were used and are illustrated in Figure 2. First, total imperviousness was used as provided directly by the

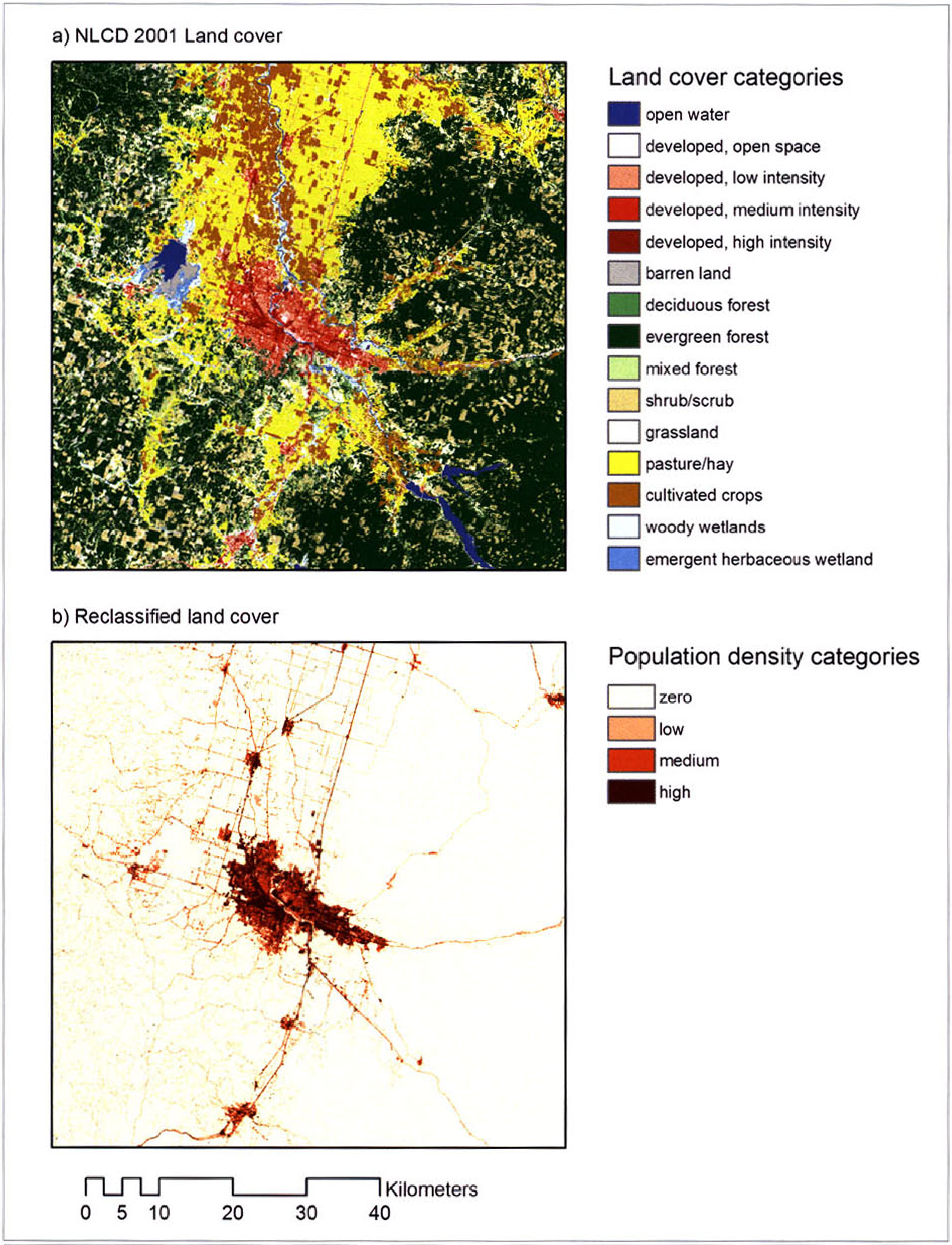

Figure 1. Original and reclassified land cover data for a sample study area in Eugene, Oregon.

NLCD data (Figure 2A. Second, the imperviousness grid was modified by running a boundary clean operation (Figure 2B). Detailed inspection of the imperviousness data revealed that many of the non-zero imperviousness cells consisted of roads and the boundary clean operation removes these linear features. Third, cells with an imperviousness value greater than 75 percent were removed (Figure 2C).

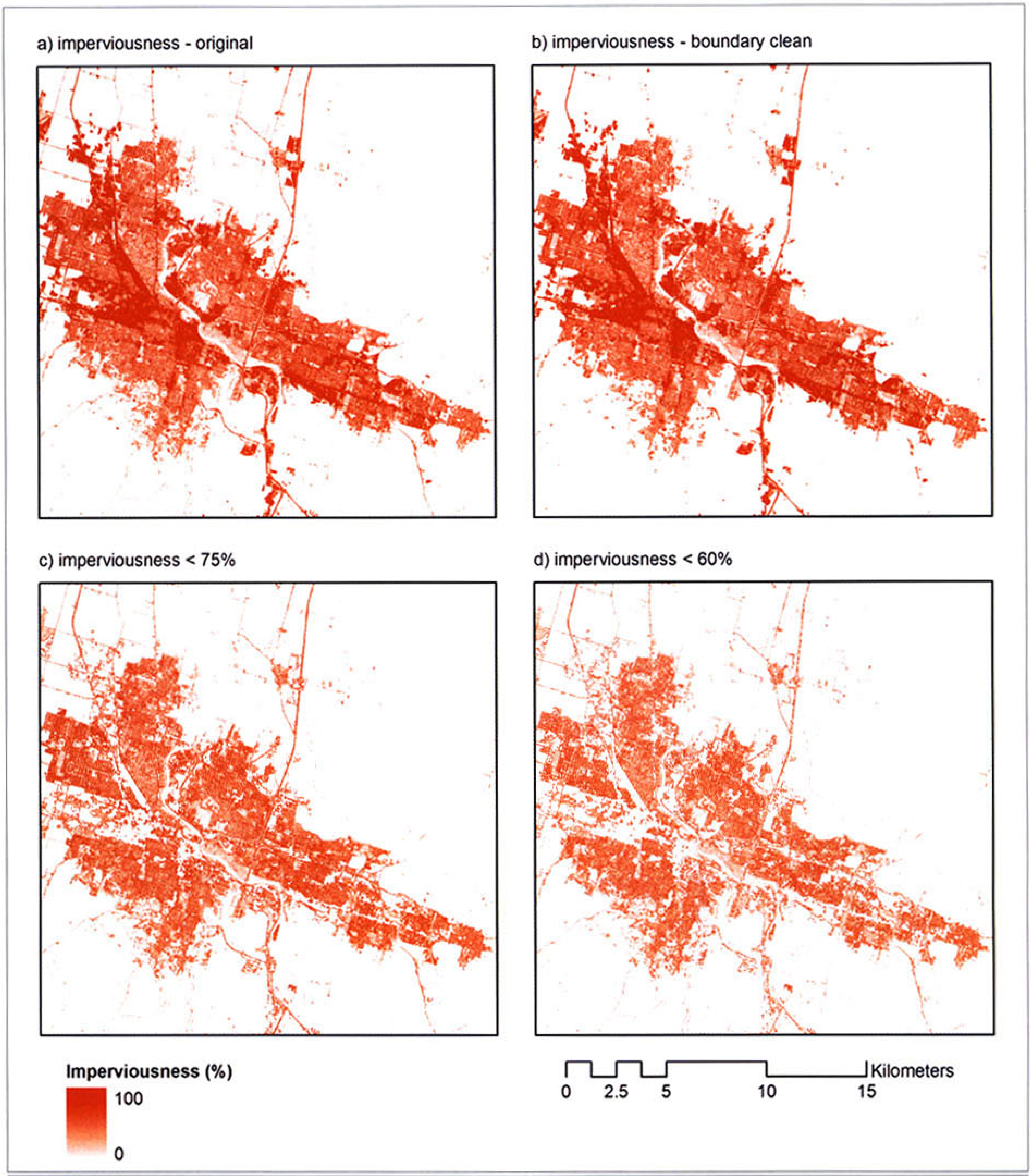

Figure 2. Original and modified versions of imperviousness data for a sample study area in Eugene, Oregon.

Fourth, cells with an imperviousness value greater than 60 percent were removed (Figure 2D). Cells with high imperviousness values are expected to consist of nonresidential areas. This is supported by other studies that have correlated imperviousness and population density. For example, Morton and Yuan (2009) removed cells greater than 75 percent imperviousness to estimate population density. The 60-percent threshold was added based on detailed inspection of the imperviousness data combined with high resolution orthophotos for selected urban areas within the four states.

Road networks were obtained from the TIGER 2000 line files. These files consist of polylines at a scale of 1:100,000 with codes for road types. Figure 4 provides an example of the TIGER road data for Eugene, Oregon. Road length (in km) was determined for each tract and block group.

Similar to imperviousness, to redistribute the population from tracts to block groups, these total lengths were used as weight factors. This assumes a linear relationship between road density (in km/km^2) and population density. Two different versions of road length were used: total road length, which included all possible road types; and local roads, which excluded all interstates, highways, major arterials, and unpaved roads. The logic behind using only the local roads is that many road types may not correspond directly to population but instead serve only as connections between populated places.

Data on nighttime lights was obtained from the Operational Linescan System (OLS) of the Defense Meteorological Satellite Program (DMSP). The specific image employed consisted of the average value of stable nighttime light intensity for the year 2000. The global grid has a resolution of 30 arc seconds, roughly equivalent to about 2.7 km in the continental U.S. Figure 5 gives an example of the DMSP data for Eugene, Oregon. The DMSP nighttime lights data were employed in a manner identical to imperviousness. The only difference is the fact that the nighttime lights data do not have a particular unit, i.e., the grid consists of values ranging from 0 to 63, with 0 indicating no lights and 63 being the maximum on an otherwise unit-less scale. Since the intensity values are used as relative weight factors this has no bearing on the results other than the fact that a linear relationship between intensity values and population is assumed.

Unit	Population Count	Imperviousness	Population Estimate	Error
BG 1	998	21.3%	773	-225
BG 2	948	23.8%	861	-87
BG 3	938	42.5%	1540	602
BG 4	739	12.4%	450	-289

Note: imperviousness in this table represents the amount of impervious surfaces in each block group as a percentage of the tract

Total number of incorrectly placed people = (225+87+602+289)/2 = 602

Figure 3. Example of the redistribution of population from tract to block groups based on total imperviousness.

When comparing the various datasets there is strong agreement between the land cover and imperviousness information because they are derived from the same Landsat imagery. Imperviousness, however, provides slightly more detail within each of the four urban land cover classes. Road density also corresponds fairly closely to land cover and imperviousness, but there are obviously many roads outside of urban areas. The nighttime lights data are of much lower resolution, and the set provides a much more generalized image of the distribution of urban development. Based on the comparison of datasets, the following nine techniques were identified:

- Areal weighting
- Land cover
- Total imperviousness
- Imperviousness < 60 percent
- Imperviousness < 75 percent
- Cleaned imperviousness
- Total roads
- Local roads
- Nighttime lights

After each of the dasymetric mapping techniques was applied to estimate the population count for block groups within each state, the estimated values were compared to the known population counts. An error analysis was carried out to determine measures of agreement. Tracts with a population count of less than 100 and tracts consisting of a single block group were excluded from this analysis to avoid issues of sample size. For each state, the following error metrics were determined:

- The number of people placed incorrectly. This is determined by the sum of the absolute values of the difference between estimated and known population counts for each block

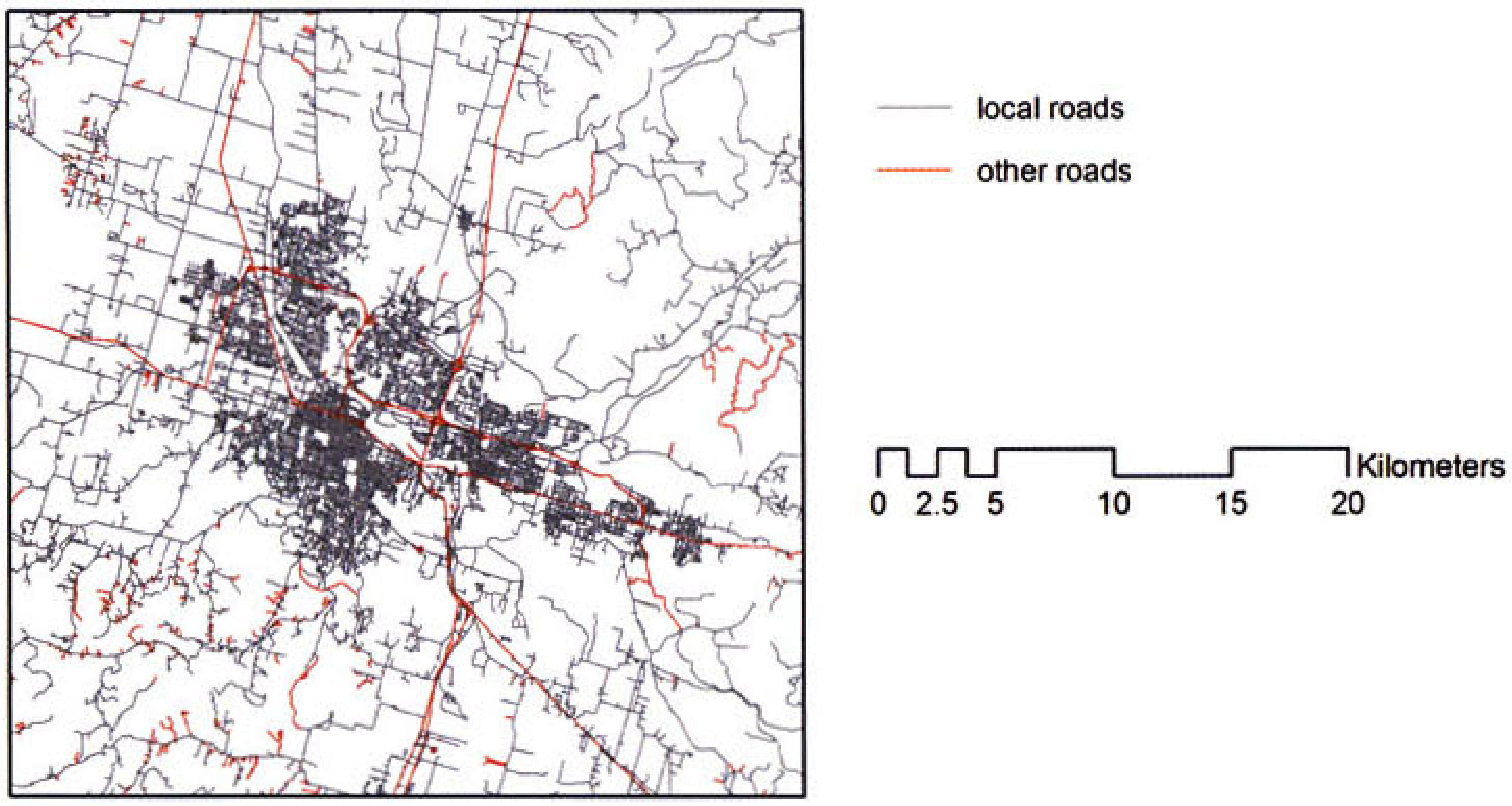

Figure 4. TIGER 2000 local roads and other roads for a sample study area in Eugene, Oregon.

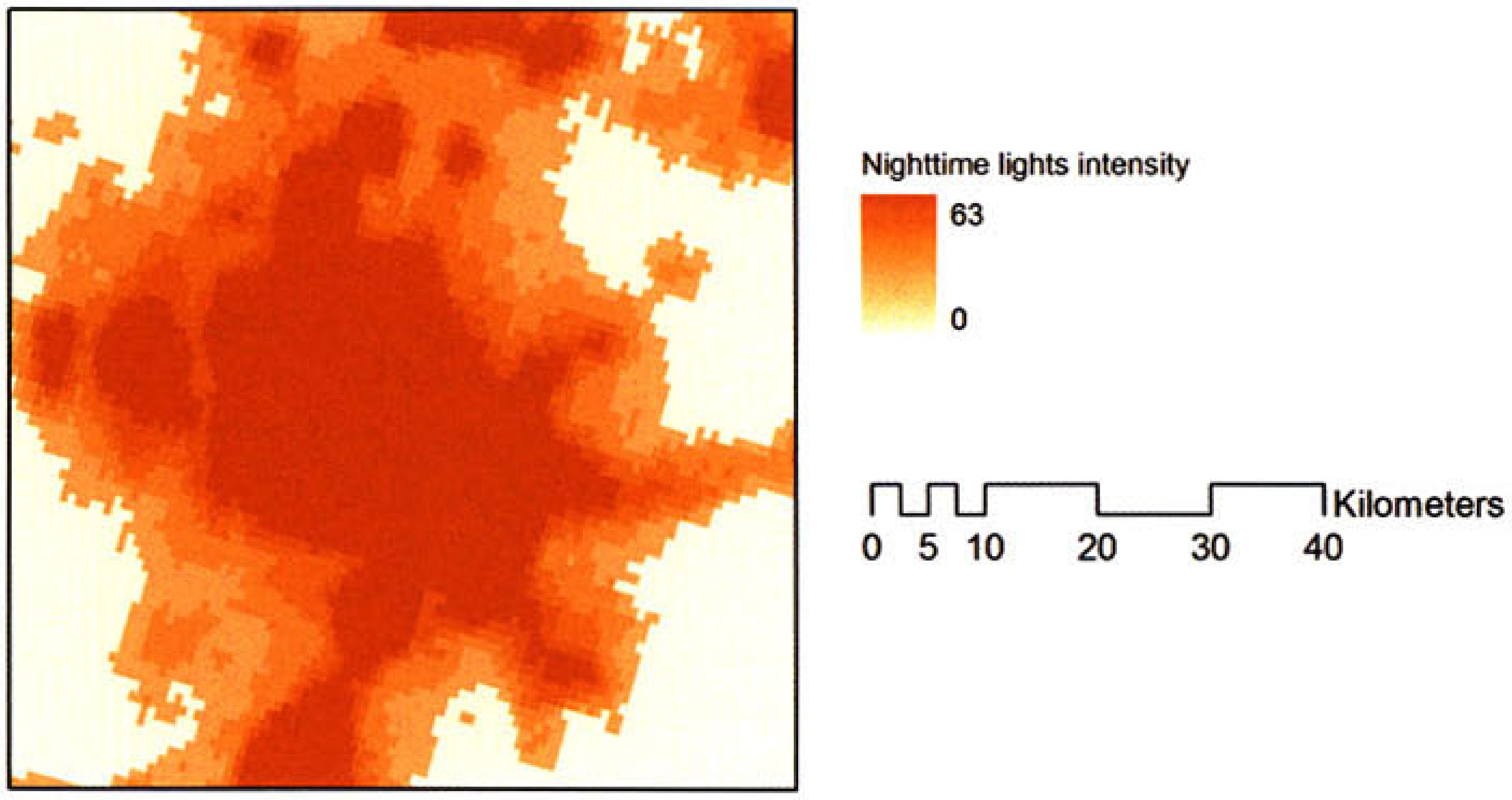

Figure 5. Defense Meteorological Satellite Program 2000 nighttime lights intensity for a sample study area in Eugene, Oregon.

group, divided by two. The logic is that a single incorrectly placed person results in an error of 2—minus 1 in one block group and plus 1 in another. Figure 3 includes an example of this type of error calculation.

- The median of the absolute value of the percent error for each block group, using the known population as the base.
- The R-squared value of the correlation between estimated and known population counts for each block group.

These error metrics were determined for each of the nine techniques and for every state. Subsequently, each state was further divided into three different categories based on the average population density of tracts: less than 250 people per km^2, between 250 and 1000 people per km^2, and more than 1000 people per km^2. The selection of the density classes is somewhat arbitrary but corresponds roughly to density categories employed by other research. For example, the U.S. Census Bureau employs densities of 500 and 1000 people per square mile to identify urban areas. The State of Pennsylvania defines an area as rural if the population density is below 274 persons per square mile. Cayo and Talbot (2003) used a population density of 250 people per km^2 to separate suburban and rural areas. Error metrics were determined for each density category to test the performance of the nine techniques relative to population density.

Results and Discussion

The performance of each of the nine dasymetric mapping techniques for each of the four states is summarized in Table 3. Performance metrics include the number and percentage of the people placed incorrectly, the median value of the absolute percentage error, and the R-squared value of the correlation between estimated and known block group population counts. The three error metrics result in virtually the same ranking of the nine techniques. For the discussion of the results in Table 3, the rank (1 through 9) based on the percentage of people placed incorrectly will be used.

The first observation is that all techniques indeed outperform areal weighting—not always by a great margin, but consistently across all four states. The performance of areal weighting itself does vary substantially, from a low error of 16.9 percent for Connecticut to a high error of 29.3 percent for New Mexico. The performance of areal weighting corresponds to the overall population density, with Connecticut having the highest population density and New Mexico the lowest. This reflects the fact that estimating small area populations is more difficult in areas with lower population densities. The performance of areal weighting represents the baseline against which the performance of dasymetric mapping should be assessed.

The best performing technique, somewhat surprisingly, is different for every state. Minor road density is the best technique for Connecticut, land cover is the best technique for New Mexico, imperviousness under 60 percent is the best method for Oregon, and imperviousness under 75 percent is the best method for South Carolina. The differences in performance between the top-ranked methods are relatively small, but the lack of consistency in ranking between the four states indicates that there is no single technique that consistently outperforms all others.

The overall improvement that can be accomplished using dasymetric techniques is also noteworthy. Compared to areal weighting, the overall error in terms of the number of incorrectly placed people can be reduced by a factor of approximately two, based on the best performing methods. For example, in the case of New Mexico, the error of 511,848 people for areal weighting is reduced to 252,371 using land cover. Error reductions for Oregon and South Caroline are similar in magnitude, but Connecticut falls a bit short of this with a reduction from 559,739 to 351,506. Apparently, since the baseline error for Connecticut is relatively small (16.9 percent) there is less to be gained in relative terms by employing dasymetric mapping techniques.

The use of land cover performed very well across the four states, with a rank of 1 for New Mexico, a rank of 2 for Oregon and South Carolina, and a rank of 4 for Connecticut. This illustrates the robustness of land cover as a source of ancillary data. Imperviousness also performs well, but some variations perform better than others. The imperviousness techniques where cells with the highest values have been removed (60 and 75 percent) perform slightly better than the original total imperviousness. This supports the logic that these highest values indeed correspond to non-residential areas and that removing them improves the performance of dasymetric mapping. The cleaned version of imperviousness, however, performs very poorly. This approach removes linear and other small features from the imperviousness data and is not robust in terms of estimating the distribution of population.

The performance of roads is highly variable. For Connecticut, minor roads and total roads ranked 1 and 2, respectively, but their rank in other states is much lower, in particular New Mexico and Oregon. This suggests that the use of roads performs better in areas with greater population density and that as a measure of population distribution, this method is less robust across study areas with varying characteristics. The use of nighttime lights performed poorly. In both Connecticut and Oregon it was the lowest ranked technique with the exception of areal weighting, and in the other two states the performance was only slightly better. This suggests that the coarse resolution of the nighttime lights data makes it only moderately useful for dasymetric mapping. However, the technique does improve upon areal weighting, presenting a possible source of ancillary data when higher resolution data are not available.

In terms of the best performing methods it is a close three-way tie between land cover, impervious with values greater than 60 percent removed, and imperviousness with values greater than 75 percent removed. The latter has a slight edge over the other two because it consistently ranks in the top three, but the differences in performance are relatively small. This confirms the usefulness of imperviousness as a source of ancillary data for dasymetric mapping, but the improvements that can be made over traditional land cover techniques are fairly modest.

To further compare the performance of the various techniques, bivariate correlations were determined between the various methods using the error in the population estimates at the block group level. Table 4 summarizes the correlation coefficients for South Carolina as an example, with results for the other states revealing similar patterns. As expected, the correlations between the various versions of imperviousness are very strong, as is the correlation between the two versions of road density. The correlation between land cover and imperviousness is also quite strong (with the

Technique	People Placed Incorrectly			Median Abs. % Error	R-squared
	Count	Percentage	Rank		
Connecticut					
Areal weighting	559,739	16.9	9	27.9	0.447
Land cover	394,724	11.9	4	18.8	0.582
Total imperviousness	408,770	12.4	6	19.7	0.561
Imperviousness < 60%	407,589	12.3	5	20.2	0.589
Imperviousness < 75%	374,961	11.3	3	18.8	0.618
Cleaned imperviousness	503,041	15.2	7	22.6	0.409
Total road density	364,320	11.0	2	18.7	0.642
Minor road density	351,506	10.6	1	17.9	0.664
Nighttime lights	525,464	15.9	8	26.7	0.487
New Mexico					
Areal weighting	511,848	29.3	9	46.7	0.143
Land cover	252,371	14.5	1	21.1	0.519
Total imperviousness	297,869	17.1	4	25.2	0.416
Imperviousness < 60%	282,282	16.2	2	21.8	0.411
Imperviousness < 75%	283,816	16.3	3	22.8	0.413
Cleaned imperviousness	306,141	17.5	5	26.3	0.415
Total road density	378,025	21.6	8	31.9	0.254
Minor road density	371,693	21.3	7	29.3	0.255
Nighttime lights	370,561	21.2	6	36.1	0.378
Oregon					
Areal weighting	863,315	26.1	9	41.3	0.234
Land cover	452,171	13.6	2	21.2	0.628
Total imperviousness	496,891	15.0	4	23.4	0.572
Imperviousness < 60%	442,384	13.4	1	20.8	0.631
Imperviousness < 75%	478,797	14.5	3	23.1	0.596
Cleaned imperviousness	505,111	15.2	5	23.9	0.591
Total road density	624,702	18.9	7	25.8	0.376
Minor road density	620,544	18.7	6	25.4	0.375
Nighttime lights	660,438	19.9	8	32.1	0.394
South Carolina					
Areal weighting	859,254	22.3	9	38.1	0.3768
Land cover	512,542	13.3	2	22.5	0.6785
Total imperviousness	596,205	15.4	6	25.7	0.6089
Imperviousness < 60%	517,606	13.4	3	22.2	0.6832
Imperviousness < 75%	476,633	12.3	1	20.3	0.7224
Cleaned imperviousness	829,252	21.5	8	36.1	0.4491
Total road density	568,710	14.7	5	23.8	0.6069
Minor road density	551,732	14.3	4	22.9	0.6243
Nighttime lights	628,932	16.3	7	27.2	0.5775

Table 3. Performance metrics of dasymetric mapping techniques.

exception of cleaned imperviousness), confirming the relatively similar overall performance of these methods. The correlations for road density are on the low side, especially with imperviousness. This indicates that while imperviousness and road density demonstrate a similar overall performance for South Carolina, these methods disagree substantially in terms of the specific block

	Areal weighting	Land cover	Total imperviousness	Imperviousness < 60%	Imperviousness < 75%	Cleaned imperviousness	Total road density	Minor road density	Nighttime lights
Areal weighting	1								
Land cover	0.398	1							
Total imperviousness	0.128	0.736	1						
Imperviousness < 60%	0.167	0.797	0.954	1					
Imperviousness < 75%	0.202	0.812	0.873	0.970	1				
Cleaned imperviousness	0.004	0.408	0.766	0.704	0.616	1			
Total road density	0.820	0.521	0.209	0.265	0.314	0.024	1		
Minor road density	0.791	0.460	0.164	0.216	0.266	0.011	0.963	1	
Nighttime lights	0.679	0.543	0.364	0.402	0.419	0.132	0.574	0.530	1

Table 4. Correlation coefficients between dasymetric mapping techniques for South Carolina based on block group level errors in population counts ($n = 2{,}800$).

groups where they perform well. The strongest correlation for road density is in comparison with areal weighting, which suggests that many of the block groups where road density performs poorly correspond to block groups where areal weighting performs poorly too.

The performance of the nine techniques is also broken down by population density category in Table 5. In this case only the percentage of people placed incorrectly is included as a metric together with the rank. Several patterns emerge from the results in Table 5. As a general rule, the performance improves for areas of higher population densities, although there are some notable exceptions. A good example of this general pattern is New Mexico where the error for areal interpolation drops from 40.1 percent for a population density less than 250 people per km^2, to 29.0 percent for a population density between 250 and 1000 people per km^2, and to 13.3 percent for a population density greater than 1000 people per km^2. Most other methods follow this same general trend, confirming that small area population estimates are less prone to error with increasing population densities. One notable exception to this is the use of imperviousness with values greater than 60 percent removed, for Connecticut, where this method performed worst in the highest population density category. This suggests the assumption that cells with very high population densities represent non-residential areas may be flawed in areas with very high average population densities, for example urban areas with many apartment complexes and very few unpaved areas. The other notable exception is the use of nighttime lights which consistently performs the worst in areas of medium population density.

Despite some of these notable exceptions, the results support the conclusion that dasymetric mapping is more difficult in areas with low population densities. This is related to the fact that population distribution in rural areas is typically more heterogeneous compared to urban areas. In areas with substantial urban and suburban development (e.g., Connecticut), large areas may be relatively homogeneous. In areas characterized by very low population densities (e.g., New Mexico), large areas with very sparse population may be interrupted by small areas with very high population concentrations. Patterns in the clustering of population likely play an important role in determining the accuracy of dasymetric mapping techniques, which is not fully captured in the stratification of study areas based on population density alone.

Another pattern that emerges from Table 5 is that the performance of each method within a state varies with the population density category. For example, minor roads are ranked number 1 in Connecticut for low- and medium-density areas, but ranked number 3 for high-density areas. For South Carolina this trend

Technique	< 250 km^2		250 to 1000 km^2		> 1000 km^2	
	(%)	Rank	(%)	Rank	(%)	Rank
Connecticut	(n= 633)		(n=874)		(n=1,057)	
Areal weighting	18.3	8	18.5	9	17.3	9
Land cover	12.7	4	12.2	5	11.2	2
Total imperviousness	14.1	6	12.3	6	11.5	5
Imperviousness < 60%	12.2	3	10.4	3	14.2	8
Imperviousness < 75%	13.1	5	10.4	3	11.1	1
Cleaned imperviousness	23.4	9	13.5	7	11.3	3
Total road density	11.3	2	10.2	2	11.6	6
Minor road density	11.2	1	9.5	1	11.3	3
Nighttime lights	15.2	7	18.0	8	14.4	7
New Mexico	(n=589)		(n=323)		(n=439)	
Areal weighting	40.1	9	29.0	9	13.3	9
Land cover	19.0	1	12.3	3	9.2	3
Total imperviousness	23.6	4	13.3	4	10.1	4
Imperviousness < 60%	23.7	5	11.1	1	8.8	1
Imperviousness < 75%	23.5	2	11.7	2	8.9	2
Cleaned imperviousness	23.5	2	15.4	6	10.2	5
Total road density	32.0	8	16.0	7	10.4	7
Minor road density	31.9	7	15.1	5	10.2	5
Nighttime lights	24.9	6	25.0	8	12.7	8
Oregon	(n=1,049)		(n=426)		(n=976)	
Areal weighting	40.5	9	29.8	9	11.6	9
Land cover	17.8	1	15.0	4	9.4	4
Total imperviousness	19.7	3	16.2	7	10.3	6
Imperviousness < 60%	19.3	2	13.4	1	8.1	1
Imperviousness < 75%	20.1	4	13.6	2	9.8	5
Cleaned imperviousness	20.6	5	15.6	6	10.4	7
Total road density	32.4	7	15.2	5	8.5	3
Minor road density	32.7	8	14.6	3	8.3	2
Nighttime lights	26.1	6	27.0	8	11.4	8
South Carolina	(n=1,691)		(n=849)		(n=258)	
Areal weighting	25.3	9	18.4	9	13.3	9
Land cover	13.5	2	13.2	5	11.7	5
Total imperviousness	15.4	4	16.0	6	13.0	6
Imperviousness < 60%	13.9	3	13.1	4	10.7	3
Imperviousness < 75%	13.0	1	11.5	2	10.4	2
Cleaned imperviousness	24.7	8	17.1	7	13.2	7
Total road density	16.8	7	11.7	3	10.8	4
Minor road density	16.5	6	10.9	1	10.1	1
Nighttime lights	16.3	5	17.1	7	13.2	7

Note: Sample size is reported as the number of block groups employed in the calculation of the percentage.

Table 5. Percentage of people placed incorrectly by population density category.

is somewhat in reverse, with minor roads ranked number 1 in the medium- and high- density areas and ranked number 6 in the low-density areas. The use of land cover performs best in the high population density in Connecticut but in the other three states it performs best in the low population density. These observations suggest that no single method outperforms all others across different population densities. The differences in performance between methods also cannot be explained by differences in population densities.

While the search for the single best technique appears to be somewhat elusive, the results in Table 5 can be used to identify the most robust technique. A robust technique would perform well for different study areas across a range of different conditions. In this context, land cover has a lowest rank of 5 and this occurs three times in Table 5. Imperviousness with values greater than 60 percent removed has a lowest rank of 8, but this occurs only once, and the next lowest rank is 5, and this occurs only once. Imperviousness with values greater than 75 percent removed has a lowest rank of 5, and this occurs twice. This suggests these three methods are very similar in performance. Another way to look at this comparison is to calculate the sum of ranks from Table 5; the results are presented in Table 6. A method that consistently outperforms all others would have a sum rank of 12. The results indicate that imperviousness with values greater than 75 percent removed has a slight edge, followed by imperviousness with values greater than 60 percent removed, followed by land cover. Next is minor roads followed by a tie between total roads and total imperviousness. Next is cleaned imperviousness, followed by nighttime lights. While this sum of ranks approach has limitations and only takes relative performance into consideration, our final ranking of techniques generally follows the more detailed patterns observed in Tables 3 and 5.

Technique	Sum of Ranks
Areal weighting	107
Land cover	39
Total imperviousness	61
Imperviousness < 60%	35
Imperviousness < 75%	31
Cleaned imperviousness	72
Total road density	61
Minor road density	43
Nighttime lights	85

Table 6. Sum of ranks based on performance metrics by state and population density.

Conclusions

The comparison of multiple ancillary data sources for dasymetric mapping has provided some meaningful insights into their performance. First, the results confirm the robustness of the existing land cover techniques. While land cover does not emerge as the single best technique in each scenario, it consistently ranks high across different study areas. Imperviousness is very useful as a source of ancillary data for dasymetric mapping, and after removal of the cells with the highest imperviousness values this method regularly outperforms the land cover technique. Despite the improvement represented by adjusted imperviousness data over land cover, the reduction in the overall error is relatively modest. The use of total imperviousness, however, does not perform as well, and neither does the cleaned version of imperviousness which removes linear and small features. The use of road density (minor roads in particular) performs quite well for some areas, but its performance overall is inconsistent across different study areas. Nighttime lights consistently perform poorly relative to the other methods, although the use of this data does present some improvement over areal weighting.

The current study has a number of limitations. First, only four states were used in the analysis. While this represents a large number of census tracts across many different geographic regions, the sample is not completely representative of the entire U.S. A related limitation is that the analysis was conducted at the level of states. While convenient for making comparisons with other studies, an analysis of study areas based on physiographic or other characteristics may provide different insights. Second, the land cover and imperviousness data are derived from the same Landsat imagery and are highly correlated as a result. This partially explains the relatively high agreement in the performance of the land cover and imperviousness techniques. The performance of these techniques may differ if different data sources were employed. Third, the performance of the dasymetric techniques was determined using block groups as the target zones. Block groups are relatively large in areas of low population densities, and the performance of the various techniques for population estimates for smaller areas is not known. Fourth, a linear relationship was assumed between the ancillary data (imperviousness, road density, and nighttime lights) and population density, which may not hold true across a wide range of population densities. Despite these limitations, the current study represents one of the more comprehensive comparisons of dasymetric mapping techniques to date.

One of the strengths of the current study is that nationally available datasets were employed, which allows for greater replicability. Only the nighttime lights dataset is available outside the U.S. and in fact, this dataset has global coverage and has been available annually since 1992. So while nighttime lights is the lowest performing technique, it presents opportunities for temporal dasymetric mapping at a global scale.

The results also provide some insights into the nature of the persistent problem of small-area population estimates. As expected, the error in these estimates is greatest in areas of lowest population density. In terms of the magnitude of these errors, in the lowest density areas the error introduced by areal interpolation is approximately 40 percent (based on the number of people placed incorrectly) for New Mexico and Oregon and somewhat less for Connecticut and South Carolina. The best performing methods are able to reduce this error to about half, which still represents a substantial amount of error. This suggests that for rural areas in particular more detailed ancillary is needed to further improve the robustness of dasymetric mapping. This data could consist of parcels, address points, or more detailed land cover data, but such data are not available in a consistent format for the entire U.S. In higher- density urban areas the error of population estimates is reduced to approximately 10 percent, and this is fairly consistent between the techniques considered and across different study areas. A state-wide error of about 10 percent may thus represent the limit of what is achievable for the national level datasets currently available.

Future research on dasymetric mapping should consider using even larger study areas (e.g., all U.S. States) to test for robustness across a wider range of conditions. Additional efforts should also be focused on comparing datasets from different sources and/or different resolutions to examine the degree to which more detailed ancillary data could lead to further reductions in error. The more widespread availability of parcel and address point data are a promising development in this regard. A final area of interest is the development of dasymetric mapping approaches which rely on multiple ancillary data sources which may improve upon the performance of single data source techniques.

REFERENCES

Bhaduri, B., E. Bright, P. Coleman, and M.L. Urban. 2007. LandScan USA: A high-resolution geospatial and temporal modeling approach for population distribution and dynamics. *GeoJournal* 69: 103-17.

Bracken, I. 1991. A surface model approach to small area population estimation. *The Town Planning Review* 62: 225-37.

Bracken, I., and D. Martin. 1989. The generation of spatial population distributions from census centroid data. *Environment & Planning A* 21: 537-43.

Cayo, M. R., and T. O. Talbot. 2003. Positional error in automated geocoding of residential addresses. *International Journal of Health Geographics* 2(10).

Dempster, A.P., N.M. Laird, and D. Rubin. 1977. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society. Series B. Methodological* 39: 1-38.

Dobson, J.E., E.A. Bright, P.R. Coleman, R.C. Durfee, and B.A. Worley. 2000. LandScan: A global population database for estimating populations at risk. *Photogrammetric Engineering and Remote Sensing* 66(7): 849-57.

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

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

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

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

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

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

Gregory, I.N., and P.S. Ell. 2006. Error-sensitive historical GIS: Identifying areal interpolation errors in time-series data. *International Journal of Geographical Information Science* 20(2): 135-52.

Gregory, I.N. 2002. The accuracy of areal interpolation techniques: Standardizing 19th and 20th census data to allow long-term comparisons. *Computers, Environment and Urban Systems* 26: 293-314.

Higgs, G., and M. Langford. 2009. GIScience, environmental justice and estimating populations at risk: The case of landfills in Wales. *Applied Geography (Sevenoaks, England)* 29: 63-76.

Holt, J.B., C.P. Lo, and T.W. Hodler. 2004. Dasymetric estimation of population density and areal interpolation of census data. *Cartography and Geographic Information Science* 31(2): 103-21.

Homer, C., J. Dewitz, J. Fry, M. Coan, N. Hossain, C. Larson, N. Herold, A. McKerrow, J.N. VanDriel, and J. Wickham. 2007. Completion of the 2001 National Land Cover Database for the conterminous United States. *Photogrammetric Engineering and Remote Sensing* 73(4): 337-41.

Langford, M. 2007. Rapid facilitation of dasymetric-based population interpolation by means of raster pixel maps. *Computers, Environment and Urban Systems* 31: 19-32.

Langford, M. 2006. Obtaining population estimates in non-census reporting zones: An evaluation of the 3-class dasymetric method. *Computers, Environment and Urban Systems* 30: 161-80.

Langford, M., D. Maguire, and D. Unwin. 1991. The areal interpolation problem: Estimating populations using remote sensing in a GIS framework. In: I. Masser and M. Blakemore (eds), *Handling geographical information: Methodology and potential applications*. Harlow, UK: Longman. pp. 55-77.

Maantay, J.A., and A.R. Maroko. 2009. Mapping urban risk: Flood hazards, race and environmental justice in New York. *Applied Geography (Sevenoaks, England)* 29: 111-24.

Maantay, J.A., A.R. Maroko, and H. Porter-Morgan. 2008. Research note – A new method for mapping population and understanding the spatial dynamics of disease in urban areas: Asthma in the Bronx, New York. *Urban Geography* 29(7): 724-38.

Maantay, J.A., A.R. Maroko, and C. Hermann. 2007. Mapping population distribution in the urban environment: The Cadastral-based Expert Dasymetric System (CEDS). *Cartography and Geographic Information Science* 34(2): 77-102.

Martin, D. 2003. Extending the automated zoning procedure to reconcile incompatible zoning systems. *International Journal of Geographical Information Science* 17(2): 181-96.

Martin, D. 1996. An assessment of surface and zonal models of population. *International Journal for Geographical Information Systems* 10(8): 973-89.

Mennis, J. 2009. Dasymetric mapping for estimating population in small areas. *Geography Compass* 3(2): 727-45.

Mennis, J., and T. Hultgren. 2006. Intelligent dasymetric mapping and its applications to areal interpolation. *Cartography and Geographic Information Science* 33(3): 179-94.

Mennis, J. 2003. Generating surface models of population using dasymetric mapping. *The Professional Geographer* 55(1): 31-42.

Morton, T.A., and F. Yuan. 2009. Analysis of population dynamics using satellite remote sensing and US census data. *Geocarto International* 24(2): 143-63.

Mrozinski, R., Jr., and R. Cromley. 1999. Singly- and doubly-constrained methods of areal interpolation for vector-based GIS. *Transactions in GIS* 3: 285-301.

Openshaw, S. 1977. A geographical solution to scale and aggregation problems in region-building, partitioning and spatial modeling. *Transactions of the Institute of British Geographers, NS* 2: 459-72.

Patterson, L.A., and M.W. Doyle. 2009. Assessing effectiveness of national flood policy through spatiotemporal monitoring of socioeconomic exposure. *Journal of the American Water Resources Association* 45(1): 1-16.

Reibel, M., and A. Agrawal. 2007. Areal interpolation of population counts using pre-classified land cover data. *Population Research and Policy* 26: 619-33.

Reibel, M., and M.E. Buffalino. 2005. Street-weighted interpolation techniques for demographic count estimation in incompatible zone systems. *Environment & Planning A* 37: 127-39.

Sadahiro, Y. 2000. Accuracy of count data transferred through the areal weighting interpolation method. *International Journal of Geographical Information Science* 14: 25-50.

Schueler, T.R., L. Fraley-McNeal, and K. Cappiella. 2009. Is impervious cover still important? Review of recent research. *Journal of Hydrologic Engineering* 14(4): 309-15.

Simpson, L. 2002. Geography conversion tables: A framework for conversion of data between geographical units. *International Journal of Population Geography* 8: 69-82.

Sutton, P.C., S.J. Anderson, C.D. Elvidge, B.T. Tuttle, and T. Ghosh. 2009. Paving the planet: Impervious surface as a proxy measure of the human ecological footprint. *Progress in Physical Geography* 33(4): 510-27.

Sutton, P.C., C. Elvidge, and T. Obremski. 2003. Building and evaluating models to estimate ambient population density. *Photogrammetric Engineering and Remote Sensing* 69: 545-53.

Sutton, P., D. Roberts, C. Elvidge, and K. Baugh. 2001. Census from heaven: An estimate of the global human population using nighttime satellite imagery. *International Journal of Remote Sensing* 22(16): 3061-76.

Theobald, D.M., S.J. Goetz, J.B. Norman, and P. Jantz. 2009. Watersheds at risk to increased impervious surface cover in the conterminous United States. *Journal of Hydrologic Engineering* 14(4): 362-8.

Tobler, W.R. 1979. Smooth pycnophylactic interpolation for geographical regions. *Journal of the American Statistical Association* 74: 519-30.

US Census Bureau. 2000. *Census 2000 TIGER/line files technical documentation*. U.S. Census Bureau, Washington, DC.

Voss, P.R., D.D. Long, and R.B. Hammer. 1999. When census geography doesn't work: Using ancillary information to improve the spatial interpolation of demographic data. Center for Demography and Ecology, University of Wisconsin – Madison, Working Paper No. 99-26.

Wright, J.K. 1936. A method of mapping densities of population with Cape Cod as an example. *Geographical Review* 26: 103-11.

Xie, Y. 1995. The overlaid network algorithms for areal interpolation problem. *Computers, Environment and Urban Systems* 19: 287-306.

Yang, L., C. Huang, C. Homer, B. Wylie, and M. Coan. 2003. An approach for mapping large-area impervious surfaces: Synergistic use of Landsat 7 ETM+ and high spatial resolution imagery. *Canadian Journal of Remote Sensing* 29(2): 230-40.