

# Mapping Population Density: Integrating census data and land cover

## Introduction

POPULATION DENSITY MAPS are typically derived from census data and aggregated to arbitrary areal units such as census tracts, block groups, or at best census blocks. The major assumption with this representation is that within the boundary of the areal unit the population density is homogeneous. While in dense urban areas comprised of small areal units this assumption may be relatively accurate, population densities in larger census block groups outside of the urban center often differ significantly within the block group boundary. Historically cartographers have attempted to improve upon this technique by using ancillary data sources such as topographic maps, satellite imagery, and land cover datasets to estimate the true population distribution. In 1936 Wright popularized a technique in the U.S. called dasymetric mapping. Using USGS topographic maps he redistributed the population density of Cape Cod, MA into inhabitable and uninhabitable regions based on predefined density values for different settlement patterns. Since then **dasymetric mapping**, the technique of displaying the density of a continuous statistical surface, has made some small achievements, but remains mostly unrecognized and underutilized.

**In this study I demonstrate the dasymetric technique developed by Mennis (2003) using the newly released 2001 National Land Cover Database (NLCD) as the ancillary data set.**

## Dasymetric Mapping Method

Initial preparation of the 2001 National Land Cover Database involved reclassifying land cover into three urbanization classes and two exclusion classes based upon previous research by Mennis (2003) and Sleeter (2004). The three reclassified urbanization classes include high intensity development, low intensity development, and nonurban residential development. The exclusion classes designate areas of zero population and include water bodies and wetlands and developed open spaces (Figure 1).

The first step to interpolate population densities of census block groups into a continuous statistical surface is to define the **population density fraction** of each urbanization class. The method here is applied to one county. When a larger geographical area encompassing more than one county is used, population density fractions should be created separately for each county.

**Population density fraction:** the percentage of a block group's total population that should be assigned to an urbanization class within the block group.

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

$d_{uc}$  = population density fraction of urbanization class u in county c,  
 $p_{uc}$  = population density (persons/900 m<sup>2</sup>) of urbanization class u in county c,  
 $p_{hc}$  = population density (persons/900 m<sup>2</sup>) of urbanization class h (high) in county c,  
 $p_{lc}$  = population density (persons/900 m<sup>2</sup>) of urbanization class l (low) in county c,  
 $p_{nc}$  = population density (persons/900 m<sup>2</sup>) of urbanization class n (nonurban) in county c.

In the second step an **area ratio** is calculated for each urbanization class within each block group to adjust for differences in the block group's total area covered by each of the three urbanization classes.

**Area ratio:** the percentage of a block group's total area that is occupied by a particular urbanization class, divided by the assumed percentage of 33.3%.

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

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

The population density fraction and the area ratio are then used to calculate the **total fraction** of a block group's population that should be assigned to one of the three urbanization classes within that block group.

**Total fraction:** the fraction of a block group's total population that should be assigned to one of the three urbanization class within that block group.

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

$f_{ubc}$  = total fraction of urbanization class u in block group b in county c,  
 $d_{uc}$  = population density fraction of urbanization class h (high) in county c,  
 $a_{hb}$  = area ratio of urbanization class h (high) in block group b,  
 $d_{lc}$  = population density fraction of urbanization class l (low) in county c,  
 $a_{lb}$  = area ratio of urbanization class l (low) in block group b,  
 $d_{nc}$  = population density fraction of urbanization class n (nonurban) in county c,  
 $a_{nb}$  = area ratio of urbanization class n (nonurban) in block group b.

The final calculation, and the step in which the actual areal interpolation occurs, is to assign a population value to each grid cell. The result will be the actual population assigned to a single grid cell of a particular urbanization class for each block group in the county.

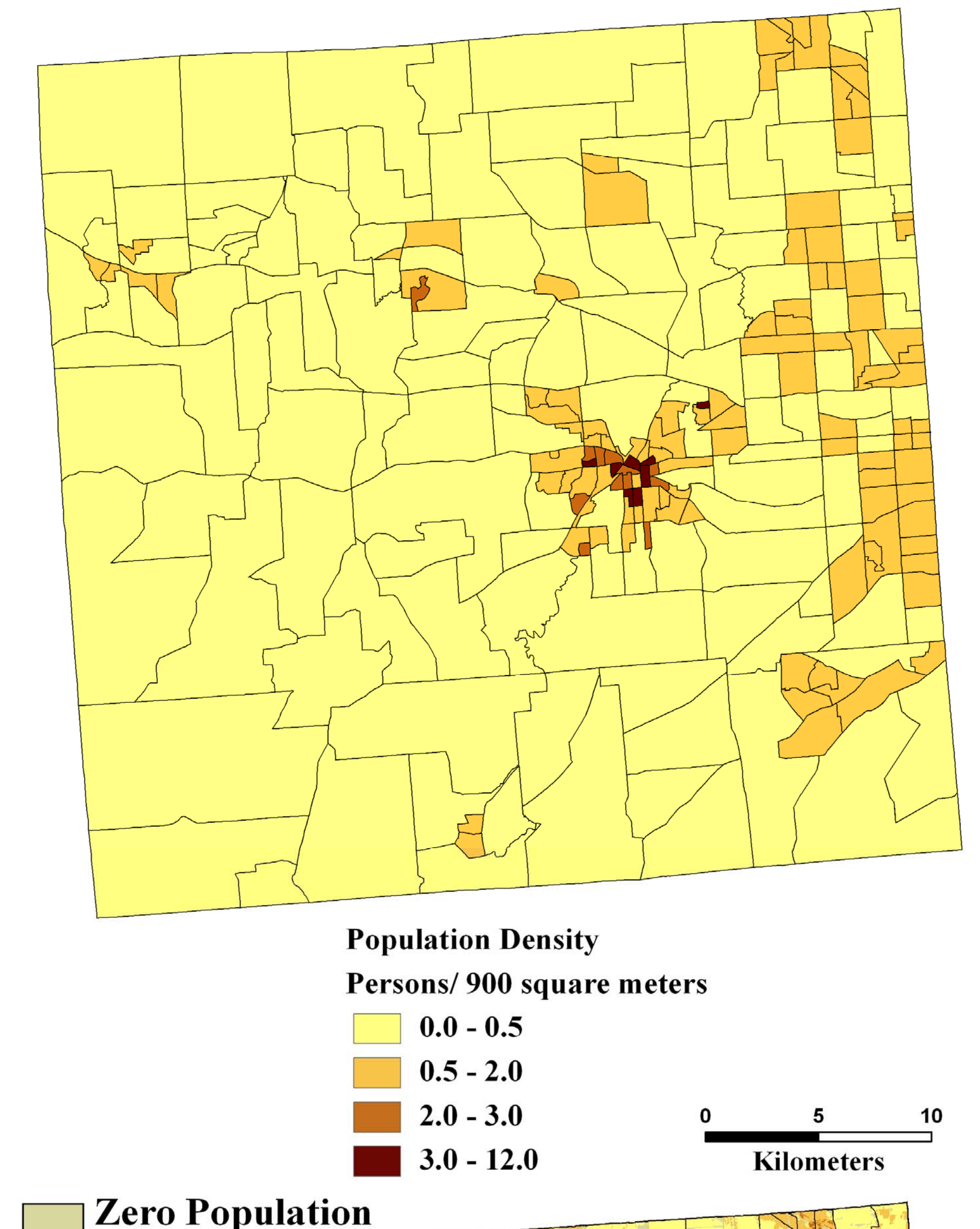
**Grid cell population value:** the number of persons assigned to one grid

cell.

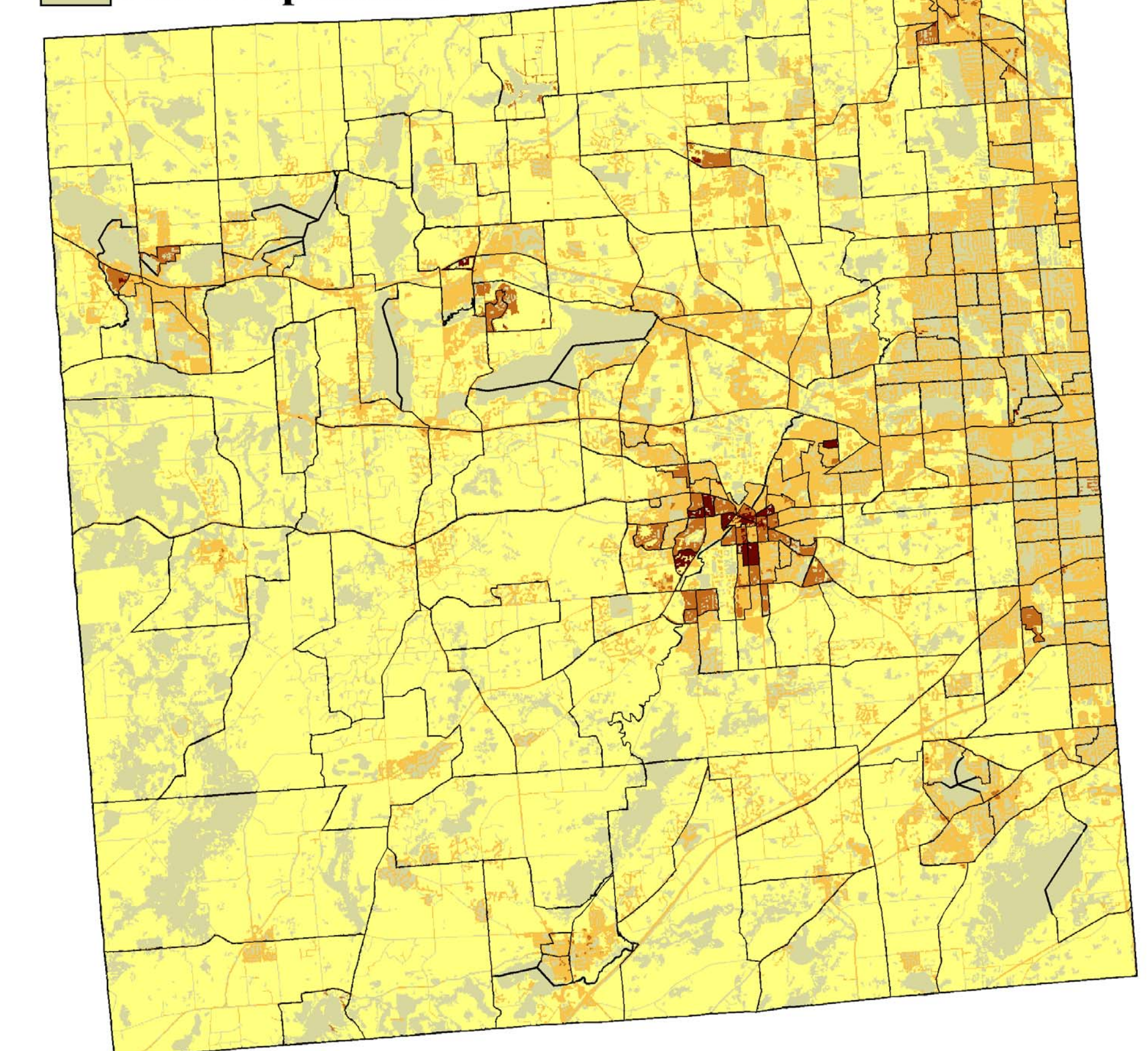
$$pop_{ubc} = (f_{ubc} * pop_b) / n_{ub} \quad (4)$$

$pop_{ubc}$  = population assigned to one grid cell of urbanization class u in block group b in county c,  
 $pop_b$  = total population of block group b.

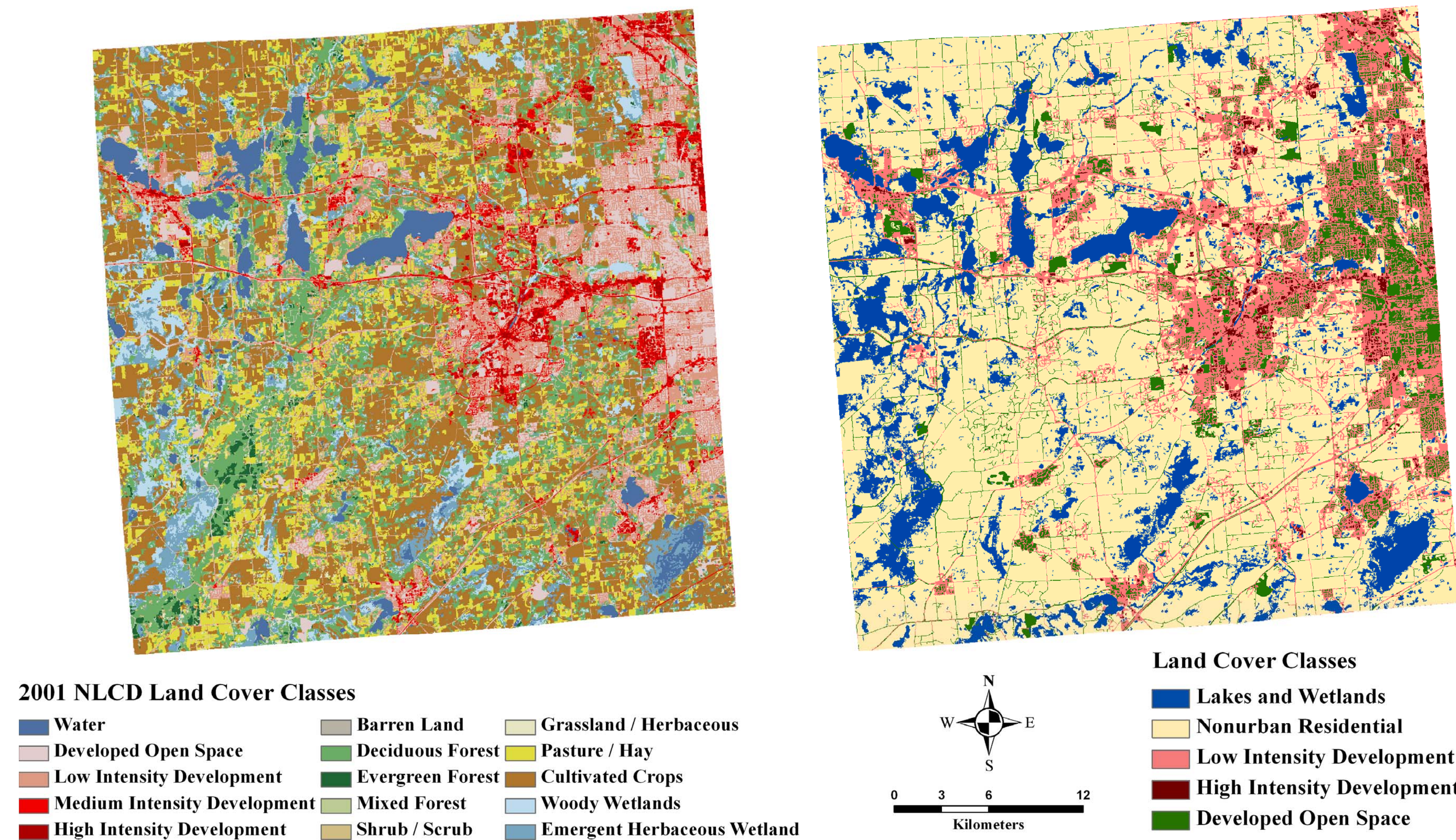
The final step is to create the raster layer that holds the population values from the interpolation as a continuous surface. The resulting surface will display the population value for each grid cell within the county as well as the population density of persons per 900 square meters (the actual size of a grid cell). This surface is shown in Figure 2.



Zero Population



**Figure 2:** Visual comparison between population densities from census block groups (above) and the dasymetric map (below) showing exclusion areas of zero population.



**Figure 1:** Land cover for Waukesha County, Wisconsin. The map on the left shows the original land cover classes as defined by the 2001 National Land Cover Database. The map on the right displays the reclassified land cover with three urbanization classes and two exclusion classes. Low and medium intensity from the 2001 NLCD has been reclassified as low in the new map.

### Map Projection

Albers equal area projection, NAD 1983, meters

### Data Sources

2001 National Land Cover Database, available from: <http://www.mrlc.gov>

2000 U.S. Census, available from: [http://arcdata.esri.com/data/tiger2000/tiger\\_download.cfm](http://arcdata.esri.com/data/tiger2000/tiger_download.cfm)

### References

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

Sleeter R. 2004. Dasymetric mapping techniques for the San Francisco Bay Region, California. Urban and Regional Information Systems Association, Annual Conference, Proceedings. Reno, Nevada: November 2004.

Wright JK. 1936. A method of mapping densities of population. *The Geographical Review*. 103-110.

**Method Description** and supplemental information available at: <https://wikis.uit.tufts.edu/confluence/display/GIS/GIS+for+Urban+and+Environmental+Analysis>

**Cartographer:** Jason Nelson, [jason.nelson@tufts.edu](mailto:jason.nelson@tufts.edu)

May 2, 2007