

Appendix I:

Detailed Methodology:

1) Gather all necessary datasets.

- a) National Landcover Dataset 1992. Available for download at <http://landcover.usgs.gov/>. Downloaded California North half, and subset the data by county.
- b) U.S. Census Bureau, 1990 Census Block Groups, Population; Sub-County Cartographic Boundary Files. Available for download at <http://www.census.gov/>.
- c) Vector Layer showing open space, public land ownership (Federal, State, and City). Received from Green Info. Network. <http://www.greeninfo.org>
- d) Landsat ETM+ 1990 Path44, Row34

2) Re-project all datasets into common projection and datum.

- a) UTM, Zone 10 N, WGS_84, Meters

3) Create Choropleth Map of Population Density, by census block group.

4) Processing NLCD

- a) *Neighborhood* command :
 - i) Using Erdas Imagine, use the *Neighborhood* command. Neighborhood functions are specialized filtering functions that are designed for use on thematic layers. Each pixel is analyzed with the pixels in its neighborhood. The number and location of the pixels in the neighborhood are determined by the size and shape of the filter, which you define. In this case, the 3x3 pixel filter was used. Each filtering function results in the center pixel value being replaced by the result of the filtering function. This function was performed to filter the NLCD for pixels that were classified incorrectly, due to the fact that NLCD is a national dataset, and this study is at the county scale.
- b) Raster Recode

- i) The NLCD needed to be recoded into the classes chosen for this study. Out of the twenty-one NLCD classes, only three classes are identified, and the remainders are classified as non-urban. Below the three classes are highlighted, Open Water, Low Intensity Residential, and High Intensity Residential. The residential classes are the urban classes, and the open water signifies an area that can later be eliminated, or given a value of zero population for the interpolation.
- c) Recode NLCD into classes 0, 1,2,3,4.
 - (a) 0 = No Data
 - (b) 1 = High Intensity
 - (c) 2 = Low Intensity
 - (d) 3 = Non Urban (18 classes)
 - (e) 4 = Water (this is separate from non-urban because later it will be part of the elimination class).

5) Process Open Space Layer

- a) Create raster layer from vector open space layer.
 - i) ArcMap-Spatial Analyst - Convert - Features to Raster
- b) Raster Recode
 - i) Erdas Imagine - Recode into two classes 0, 10.
 - (a) 0 = No Data
 - (b) 10 = Open Space
- c) Combine Two Rasters to produce the sum of all values.
 - i) Erdas Imagine- Utilities - Two Input Operators
 - ii) Output :
 - (1) 0 = No Data
 - (2) 1 = High Intensity
 - (3) 2 = Low Intensity
 - (4) 3 = Non Urban
 - (5) 4 = Water

(6) 10 = Open Space

(7) 11 = Open Space + High Intensity

(8) 12 = Open Space + Low Intensity ↔ Recode all Values to

(9) 13 = Open Space + Non Urban 4 = Eliminator

(10) 14 = Open Space + Water

iii) Water becomes part of the eliminator class as well as all of the areas that are excluded due to the fact that we know that there is no population in these areas.

6) Creating Point Feature Class Representing Land-cover ID

a) New LandCover Table

i) 0 = No Data

ii) 1 = High Intensity

iii) 2 = Low Intensity

iv) 3 = Non Urban

v) 4 = Eliminator

b) Convert Raster to Feature using Spatial Analyst, outputting POINTS.

i) For every point in the new vector layer, there is a `grid_code` attribute which represents the thematic land-cover code.

ii) Each Point Represents 30 meters on the ground, relative to the input pixel size.

iii) The output point file for Alameda County created approximately 4.7 million points. In order to process the statistics with efficiency, the data needs to be subset into a smaller, more manageable area. To emulate Mennis' study, the population density of Alameda should be calculated by sub-county, because the counties on the west coast are much larger than the east coast, where Mennis conducted his research. In addition, the large point file is subset into the POINTS the *have their center within*, each sub-county.

7) Join Land-cover data (points) with Census Block Group data (polygons)

- a) Each Point needs to have the Block Group Unique ID associated to it.
- b) To do this, conduct a *Spatial Join* (Polygons to Points)
- c) Now each point has all block group attributes associated to it creating a layer that will be the final statistical table.
- d) Using the Livermore-Pleasanton Sub-County Division as an example, the dataset should now contain around 1.2 million points.

8) Areal Interpolation

- a) Before the field calculations can take place, population density values for the sub-county needs to be established. This is a sampling method which calculates the relative difference in population density among urbanization classes (Mennis, 2003). To sample this, three block groups were selected for each urbanization class (block groups that had a majority of high, low, non-urban points). If possible, a block group that was entirely high, low or non-urban points would be the best sample. The total population and area were calculated for each sample and an aggregated population density was the outcome. (See Table 3.1) The sampled population density was calculated for each sub-county, independently because the relative difference in population density varies within the county due to the fact that population density is going to be much higher in the urban core compared to the urban fringe of the county where there lies a lot of open space.
- b) Now that there is a POINT feature class that represents the foundation for the interpolation, new fields can be added to the table. All of the calculations were performed with the *Field Calculator* in ArcMap.
- c) The first field added was the Population Density Fraction which will be calculated by equation (1) of Mennis' study.

$$d_{uc} = P_{uc} / (P_{bc} + P_{lc} + P_{nc}) \quad (1)$$

where d_{uc} = population density fraction of urbanization class u in county c , P_{uc} = population density (persons/10,000 m²) of urbanization class u in county c , P_{bc} =

population density (persons/10,000 m²) of urbanization class *h* (high) in county *c*, P_{lc} = population density (persons/10,000 m²) of urbanization class *l* (low) in county *c*, and P_{nc} = population density (persons/10,000 m²) of urbanization class *n* (nonurban) in county *c*. (Mennis, 2003) This fraction will be used throughout the interpolation.

- d) The next calculation is to determine the Area Ratio, which focuses on the difference in area between the urbanization classes, within each block group.

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

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

- e) After the Area Ratio is calculated the Total Fraction is calculated which determines the percentage of the block group population goes to each urbanization class. This is the most crucial value of the interpolation.

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

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

- f) Finally the block group population can be distributed to the points within each block group.

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

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