ICLUS v1.2 USER'S MANUAL:

ARCGIS TOOLS AND DATASETS FOR MODELING US

HOUSING DENSITY GROWTH

```
ICLUS v1.2 Tools
    1) Population and Housing Density Tools
        Batch Clip ICLUS Rasters
        Create Population Projection Rasters
        Reclassify Housing Density
        SERGoM v3
    2) Impervious Surface Tools
        Estimate Percent Impervious Surface
        Summarize Impervious Surface by Housing Density Class
```

NOTICE

DISCLAIMER

This document is an internal draft for review purposes only. This information is distributed solely for the purpose of pre-dissemination peer review under applicable information quality guidelines. It has not been formally disseminated by the U.S. Environmental Protection Agency. It does not represent and should not be construed to represent any agency determination or policy. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

Preferred Citation:
U.S. Environmental Protection Agency (EPA). 2009. <u>ICLUS v1.2 User's Manual: ArcGIS Tools and Datasets for Modeling US Housing Density Growth</u>. Global Change Research Program, National Center for Environmental Assessment, Washington, DC; EPA/600/R-09/XXX. Available from the National Technical Information Service, Springfield, VA, and online at http://www.epa.gov/ncea.

CONTENTS

TABLES

FIGURES

AUTHORS AND REVIEWERS

1. PURPOSE

The purpose of this document is to describe the tools and models generated for the Integrated Climate and Land Use Scenarios (ICLUS) project led by the US EPA's Global Change Research Program in the Office of Research and Development's National Center for Environmental Assessment (EPA 2008). The tools and models described in this document provide users with the ability to generate, visualize, and analyze spatial datasets about growth in the coterminous US.

SERGoM, the Spatially Explicit Regional Growth Model, has been under development for nearly a decade, and a number of papers describe it more fully (Theobald 2001, 2003, 2005). It has also been used for a number of applications to assess effects on other land cover classes or environmental variables, including the Forests on the Edge project, which examines effects of housing density changes on forest products and services (Stein et al. 2005, 2007). Unlike the majority of land use change models focusing on urban growth, SERGoM uses and models a full continuum of housing density, from urban to rural. This allows a more comprehensive examination of growth patterns, since exurban/low-density development has become more prevalent in the past decades and is an important aspect of possible future growth scenarios. SERGoM is based on a relatively high resolution (~1 ha) representation of land use patterns, but is best applied at a regional to national extent. This is because there are relatively few assumptions underlying the model – it is not a data-rich approach (e.g., agent-based modeling). Essentially, decadal population projections for each county drive the production of new housing units, which are allocated in response to the spatial pattern of previous growth (e.g., 1990 to 2000), transportation infrastructure, and other basic assumption. An important technical advantage of this model is that it provides a consistent, comprehensive, nationwide model at 1 ha resolution.

1.1 KEY ASSUMPTIONS AND USES

As with any model, there are a number of key assumptions that should be understood when using the SERGoM model and results. In particular, these assumptions are:

- future growth will be similar to those patterns experienced during the previous decade (e.g., 1990-2000);

- this approach is applicable to understanding regional and broader-scale development patterns, it is not suitable for finer-scale analysis;

- projected population level of a county is the major driver of growth, particularly in the long term;

- only residential housing density is estimated – other land uses typically considered to be part of urbanization are industrial, commercial, or transportation land use;

- housing units are added – a loss of housing units due to conversion to commercial or demolition is not modeled;

- the default resolution to run this model is 100 m (1 ha), but with appropriate datasets it could be run at 30 m, or at a more coarse resolution up to 400 m; and

- All input data should be in the same projected coordinate system (e.g., Albers, NAD1983).

Although there are a large possible number of uses, the tools provided in the ICLUS SERGoM toolset are designed to assist with, at a minimum, the following common anticipated user scenarios:

- to replace the ICLUS projected population values to reflect different growth rate assumptions for the nation;

- to upscale maps of housing density to a coarser resolution;

- to summarize patterns by region, watershed, county, or NLCD 2001 classes;

- to reclassify housing density into classes different than those already provided; and

- to generate a map of estimated impervious surface based on a housing density map.

Note that originally we intended to develop a tool to compute population density from housing density. However, in developing such a tool, we discovered that there can be some numerical instability issues because of the sensitivity of the assumption about people per housing unit (household size) that is used as a multiplier to convert housing units to population. As a result, we recommend computing this ratio at a coarser resolution data such as census tracts or 1 km^2 to reduce these artifacts, and then to carefully examine the results for outliers.

2.   SOFTWARE, HARDWARE, AND EXPERIENCE REQUIREMENTS

These tools were written as a series of tools in a Geoprocessing toolbox, using ArcGIS 9.3.1. The ArcToolbox geoprocessing tools require the Spatial Analyst extension and Python 2.5.1, which is installed by default with ArcGIS v9.3.1. If needed, a Python 2.5.1 installer has been included in the "software" folder of the ICLUS v1 zip file. Because this is a complex spatial model, with numerous large (fine-grain, broad-extent) spatial datasets, a high-end computing environment is required. Typically, this means >1GB RAM, 100-300 GB working space of memory, and a RAID drive for fast disk I/O. The experience level of the GIS analyst using these tools should be advanced (at least 3 years experience), knowledge of raster datasets and spatial modeling, and experience with very large (>1 GB) datasets.

2.1.   TECHNICAL DESCRIPTION OF SERGOM V3

SERGoM, unlike the majority of land use change models, allocates a full continuum of housing density, from urban to rural. This allows a more comprehensive examination of growth patterns, since exurban/low-density development generally has a footprint 10 times as large as urban areas and is growing at a faster rate than urban areas (Theobald 2005). In addition, SERGoM forecasts housing development by establishing a relationship between neighboring housing density, population growth rates, and transportation infrastructure (Theobald 2005). The model is dynamic in that as new urban core areas emerge, the model re-calculates travel time from these areas. However, the expected changes in functional connectivity that would result from such emerging urbanization were not fed back into the functional connectivity calculations used to calculate domestic migration. SERGoM also incorporates a detailed layer of developable/un-developable areas that incorporates public protected lands as well as private protected (e.g., through conservation easements) lands. Finally, population forecasts are a principal driver of SERGoM; in the model, population growth is converted to housing units, which are spatially allocated in response to the spatial pattern of previous growth and transportation infrastructure (Figure 2-1). Growth rates and other model parameters can be specified spatially-explicitly, so different regions (even census tracts or neighborhoods) have different parameters (e.g., lower household size in amenity areas, etc.). The benefit of this approach is that there are fewer (internal to coterminous United States) discrete differences across artificial analytical boundaries imposed by "piecing" individual model runs into a nationwide map, although the allocation of new housing units is restricted to counties.
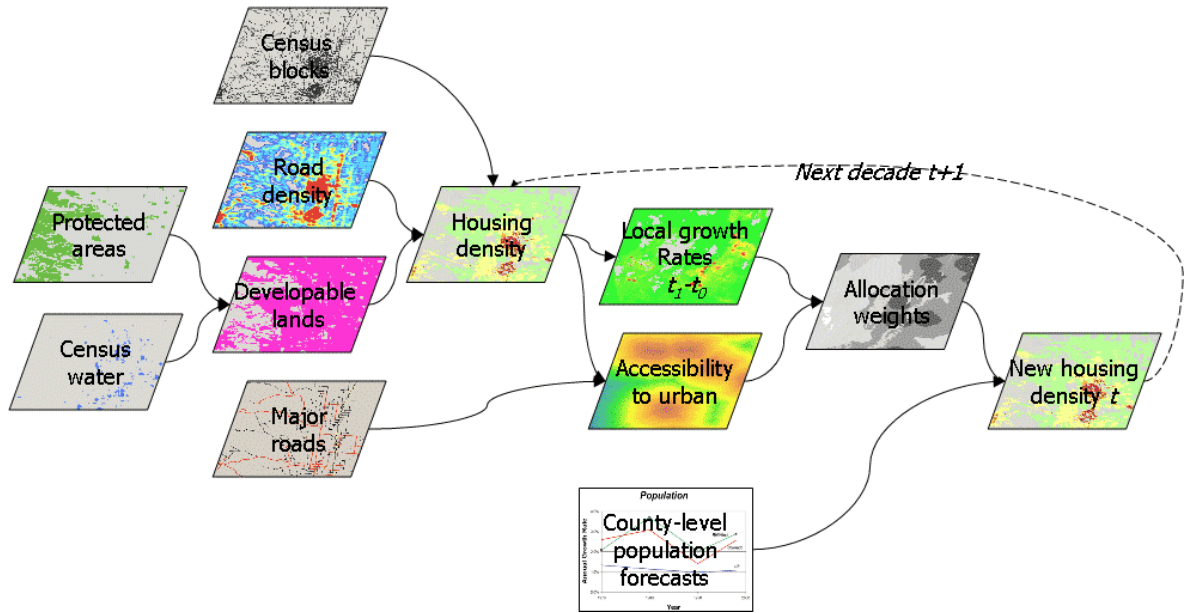
Figure 2-1. Diagram of SERGoM data inputs and modeling process

Housing density (number of housing units per acres) was computed for each 1 ha cell (100 m x 100 m raster; 2.47 acres). There are five main input spatial datasets used to estimate housing density:

1. <u>2000 Census Bureau</u> – Data were compiled from the 2000 census on the number of housing units and population for each block in 2000 and the geography or polygon boundary for each census block (from the SF1 dataset). Block-groups, which are a coarser-level aggregation of block polygons, and attributes of the number of housing units built by decade were used to estimate the historical number of housing units in each block. An operating assumption in estimating historical housing units is that they have not declined over time, so that the number of housing units in any past decade (back to 1940) did not exceed the number of units in any subsequent decade (up to 2000). Reservoirs, lakes, and wide rivers that were identified as "water blocks" were also removed, so that no housing units were placed in these undevelopable areas.

2. <u>Undevelopable lands</u> – Spatial data on land ownership were compiled from a variety of sources to create the most current and comprehensive dataset – called the UPPT (unprotected, private protected, public protected, and tribal/native lands). The UPPT dataset was generated by starting from the Conservation Biology Institute's PAD-US database. We updated the PAD-US dataset with more current data for 21 states. The operating assumption is that housing units do not occur on publicly owned lands (e.g., national parks, forests, state

wildlife areas, etc.) or on privately-owned, protected lands. Some state lands in the western United States (the so-called "school lands" sections, but not "stewardship" lands) were kept in the developable category because they are in practice sold to generate revenue for state school systems. Also, tribal lands are often considered federal (public), but here we included tribal lands as developable (except for known tribal parks). The portions of blocks that overlapped with public (and other non-developable lands) were deleted to create a modified or refined block. All housing units associated with each block are then assumed to be located in the refined (developable) portion of the blocks. Housing units were apportioned within the refined block using a dasymetric mapping approach described below.

3. <u>Road and groundwater well density</u> – The existence of major roads (interstates, state highways, county roads) was used to better allocate the location of housing units within a block. In a previous SERGoM model (v1, v2), housing units were spread evenly throughout the refined blocks. Here, in v3 of SERGoM, housing units were disproportionately weighted to areas of higher likelihood of being developed according to fine-grained land use/cover data from NLCD 2001 (see Table 2-1). Because major road infrastructure is included in the NLCD (actually burned in as values 21, 22, and some 23), road density *per se* was not included. Also, in the western US where the rural blocks are particularly large, groundwater well density was included to refine the allocation of units. Also, the analytical hierarchy process (AHP; Saaty 1980) was used to provide an estimate of logical consistency during the development of the weights (the consistency index was 0.035, which is less that then 0.15 threshold, showing that these estimates were logically consistent). Note also that these weights are applied in a relative, not absolute context. That is, the number of units that will be distributed in a given area is specified by the census block and so units are allocated in proportion to the weights found within a given block. This is robust in the face of potential mis-classification of land cover types, because all the known housing units will be allocated to a given block, regardless of the land cover (but note that the undevelopable – water and public lands – portions of the blocks are excluded.

4. <u>County population projections</u> – Population projections for each county are used to drive the future growth forecasts. Additional housing units needed were computed by determining the number of new housing units needed to meet the needs of the additional population, assuming the same (in 2000) population to housing unit ratio in each tract, using 2000 U.S. Census of Housing data.

5. <u>Commercial/industrial land use</u> – We also mapped locations with land uses that would typically preclude residential development (increased housing density), especially commercial, industrial, as well as transportation land uses. Using urban/built-up categories of NLCD 2001 (not open space developed), we identified locations (1 ha cells) that had >25%

urban/built-up land cover but that had also had lower than suburban levels of housing density (because high-density residential areas would otherwise be included in the urban/built-up land cover categories). Although some re-development of central business districts ("gentrification") is occurring, SERGoM works from the operating assumption that these are relatively smaller portions of the landscape and typically brown-field settings.

Table 2-1. Description of the weights that are applied to NLCD which are used to allocate the block housing units within each block

NLCD Classes	Description	Weight[a]
Open water & perennial ice/snow (11, 12)	No housing units on water or snowfields	0.0
Developed, open space (21)	This is typically either open space (urban parks and greens space) or roads in rural areas), so very low likelihood of housing units	0.085
Developed, low, medium & high intensity (22, 23, 24)	These cover types are where housing units most likely are located	0.549
Transitional (31, 32, 33)	Includes barren, transitional areas such as cleared or recently cleared areas, but also mines, etc.	0.115
Wildland vegetation (41, 42, 43, 51, 52, 71, 72, 73, 74)	Not likely to have high-density (urban) in these areas.	0.150
Agricultural (pasture/hay, cultivated) (61, 81, 82)	Will have some housing density, but most housing infrastructure is clustered and near roads.	0.050
Wetlands (90, 91, 92, 93, 94)	Not likely to develop in wetlands (without filling).	0.050

[a]Note that the weights are applied to each 30 m cell in the NLCD, then the weights are aggregated up to 100 m resolution by averaging the 30 m weight values.

SERGoM is a demand/allocation/supply model, where the number of new housing units needed for the next decade is computed to meet the demands of the projected population, computed here for each county (but could be other analytical unit boundaries). The average growth rate for each state-housing density class is computed from the previous to current time step (e.g., 1990 to 2000). These average growth rates are computed using a moving neighborhood for 12 development classes. These classes are formed by overlaying three housing

density classes – urban/suburban, exurban, and rural – with four accessibility classes measured as travel time (minutes one way) from the nearest urban/suburban core along (existing) major roads: 0-10, 10-30, 30-60, and >60 minutes. The resulting combination creates a "surface" of raster values that reflect historical patterns of growth – called allocation weights -- and are used to allocate the new housing units for a given time step.

Based on the Census definition of urban areas, we defined urban housing densities as less than 0.1 ha per unit and suburban as 0.1 – 0.68 ha per unit. We defined exurban density as 0.68 – 16.18 ha per unit (to 40 acres) to capture residential land use beyond the urban/suburban fringe that is composed of parcels or lots that are generally too small to be considered productive agricultural land use (though some high-value crops such as orchards are a notable exception). Rural is defined as greater than 16.18 ha per unit where the majority of housing units support agricultural production.

2.2.   SRES-SPECIFIC PARAMETERS FOR SERGOM

We used a number of population growth scenarios that are based on the Intergovernmental Panel on Climate Change (IPCC) Special Report on Emissions Scenarios (SRES) social, economic, and demographic storylines. These scenarios are rendered using a combination of models representing demography. The SRES provide storylines along two major axes, economic vs. environmentally-driven development (A-B) and global vs. regional development (1-2), which make up the four combinations of storylines, A1, A2, B1, and B2. In addition to changes in population that resulted from the various demographic assumptions associated with each SRES-compatible scenario developed for the ICLUS project, the spatial location of growth was modified using SERGoM in two ways, through household size and travel time (Table 2-2). With SERGoM, household size is expected to reflect demographic changes due to changes in fertility and socioeconomic changes that affect household formation. Travel time from urban "central city" locations is used to help express how the evolution of the urban form might be affected by changing priorities and increases in the cost of transportation.

First, weighting values can be adjusted as a function of distance away (travel time) from urban cores. Urban area (<5 minutes) weights can be lowered by a given percentage to reflect a carrying capacity or saturation of an area, specified by zoning perhaps; or raised to reflect increased desire for urban living (lofts, gentrification, etc.). Exurban area weights (~30-60 minutes) can be lowered to reflect assumptions of lower rates of development due to increased fuel prices or can be used as a surrogate for lower land availability because of increased conservation purchases (or easements). It can also be raised for exurban areas to reflect increased "urban flight" of baby-boomer retirees and rural amenities. This weighting surface is re-

computed at each time step. We modified the weights of travel times for the B1 and B2 storylines to model a "compact" growth scenario (see Table 2-2). Given the environmental orientation of the B1 and B2 storylines, we assumed that growth patterns in these scenarios would place a greater emphasis on promoting denser growth patterns closer to existing urban centers.

Table 2-2. Summary of Adjustments to SERGoM v3 for SRES scenarios.

Scenario	SERGoM parameters		
	Household size	*Travel time (minutes)* *<5; 10; 20; 30; 45; >45*	*Travel time consequences for urban growth form*
A1	Smaller (-15%)	75; 75; 85; 90; 95; 100	No change
B1	Smaller (-15%)	90; 95; 85; 90; 95; 100	Slight compact
A2	Larger (+15%)	75; 75; 85; 90; 95; 100	No change
B2	No change	90; 95; 85; 90; 95; 100	Slight compact
Baseline	No change	75; 75; 85; 90; 95; 100	No change

A second type of modification is changing assumptions about households, particularly household size (roughly family size), defined as the number of people living in a single housing unit. Currently this ratio (population per unit) is static and is computed at the tract level from the 2000 U.S. Census data. We modified this ratio to reflect assumptions in the SRES scenarios to adjust for assumed changes in demographic characteristics. For example, SRES A1 and B1 assume smaller household sizes (a reduction by 15% throughout the future decades), whereas scenarios B2 and baseline are not changed and A2 assumes a 15% increase in household size (Jiang and O'Neill, 2007). The changes in household size correspond to changes in fertility rates that are assumed under the different storylines. Under A1 and B1, where fertility is lowest, smaller average household sizes are also expected. Conversely, A2 has the highest fertility rates, so an increase in household sizes is expected. In B2, which uses the medium fertility rates, household sizes are not changed.

SERGoM was parameterized to reflect the SRES scenarios in the following ways. First, the A1 and B1 scenarios were modeled to reflect a 15 percent decline in average household size. A2 was modeled to show a 15 percent increase in average household size. B2 was modeled with no change in household size. The household size for each census tract was modified by 0.85 for A1 and B1 scenarios, and 1.15 for A2. Second, to model the "compact" growth scenarios in SRES for runs B1 and B2, SERGoM was run with modifications to the spatial allocation of new housing units as a function of travel time from urban cores. Urban cores were defined as a contiguous grouping of urban/suburban cells at least 10 ha in size.
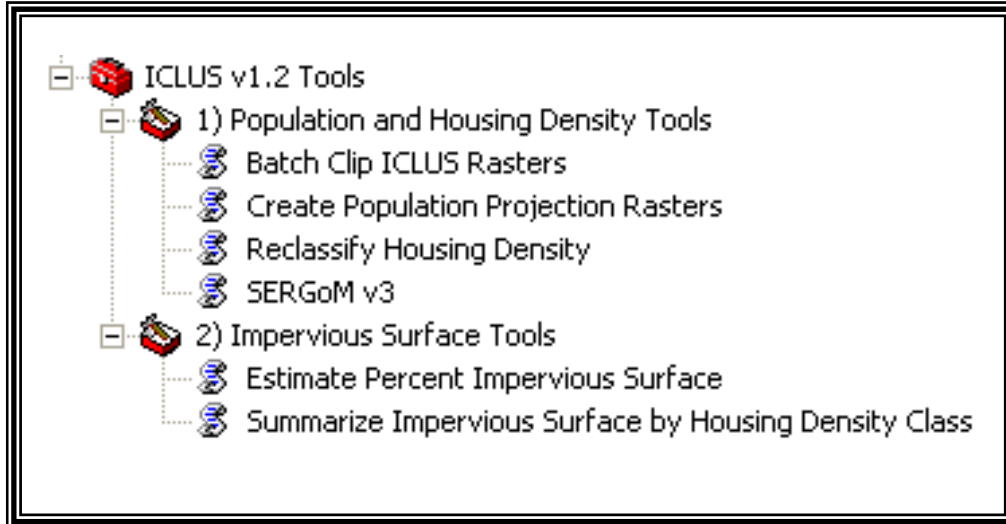
3. TOOLS

3.1. SETTING UP THE ICLUS TOOLS

Extract the ICLUS_v1.2 zip file (ICLUS v12.zip) to whatever location you desire other than your desktop. Inside the folder you will see four folders titled: "base", "population", "output", and "software".

Start either ArcMap or ArcCatalog, right-click in the ArcToolbox panel, and select "Add Toolbox". Navigate to the "software" folder and add the red toolbox named "ICLUS v1.2".
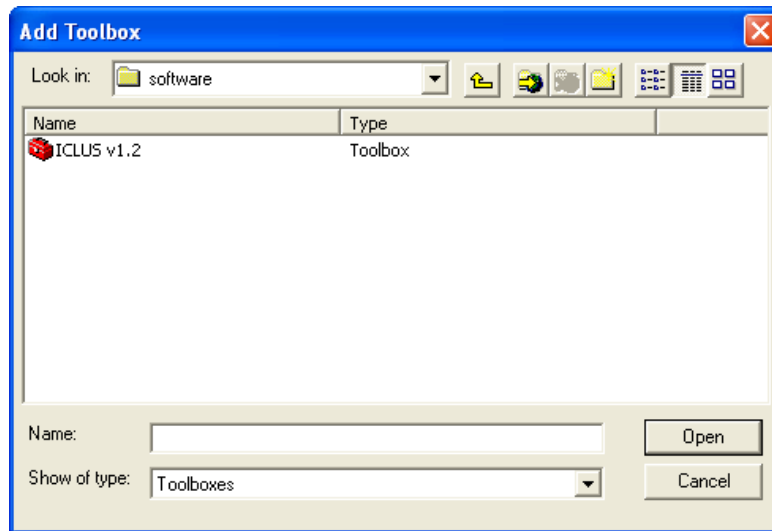
**Figure 3-1: Loading the ICLUS Toolbox into ArcGIS**

Figure 3-2: The Expanded ICLUS Tools

The tools are split into two general groupings. The two sections are **Population and Housing Density Tools** (this includes the SERGoM model itself), and **Impervious Surface Tools**. Select the particular tool you wish to run by double clicking. Some of the tools will already have certain fields filled out. Those fields reference required files, and will automatically populate with a network path relative to where you have unzipped your ICLUS folder. All fields with a green dot next to them are required.

Please see below for a description of each tool screen.

3.2.   BATCH CLIP ICLUS RASTERS
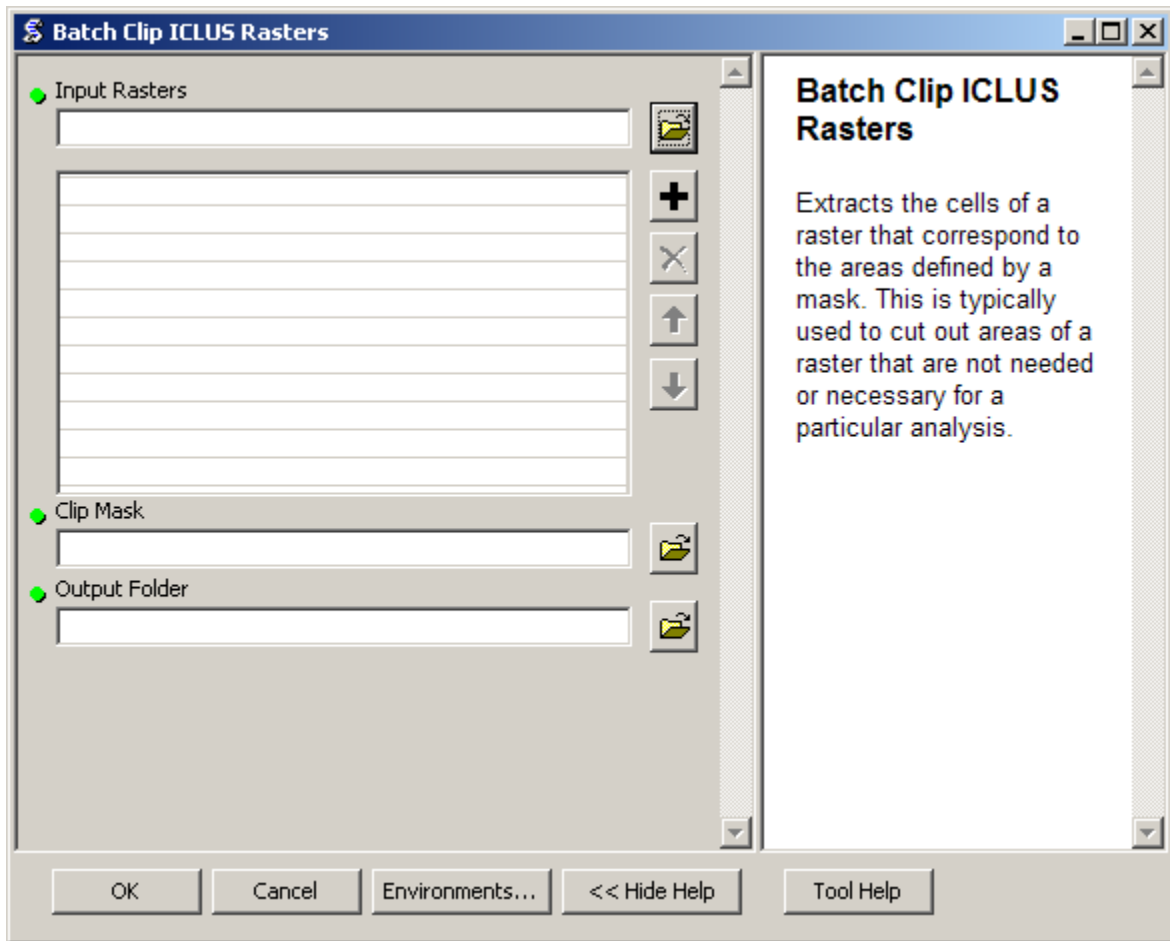
Figure 3-3: Opening Screen for the Batch Clip ICLUS Rasters Tool

You can use this tool to create and work with smaller rasters, this is especially useful if you are only interested in a certain part of the country rather than analyzing the whole. The clip mask can be any polygonal boundary that you want to "cut out" and save from the main input rasters. You select where you want the finished products to be stored in the Output Folder choice.

3.3.   CREATE POPULATION PROJECTION RASTERS

This tool creates population projection rasters from "cofips_upp" shapefile located in the "base" folder. These rasters are used by the SERGoM model, but are also appropriate for analyses requiring county level populations. The various ICLUS scenarios can be selected by clicking on the pull down next to the field. Only valid choices are accepted.

- These rasters can also be created "on the fly" by simply running the SERGoM tool.

- Population rasters will automatically be placed in the "population" folder.
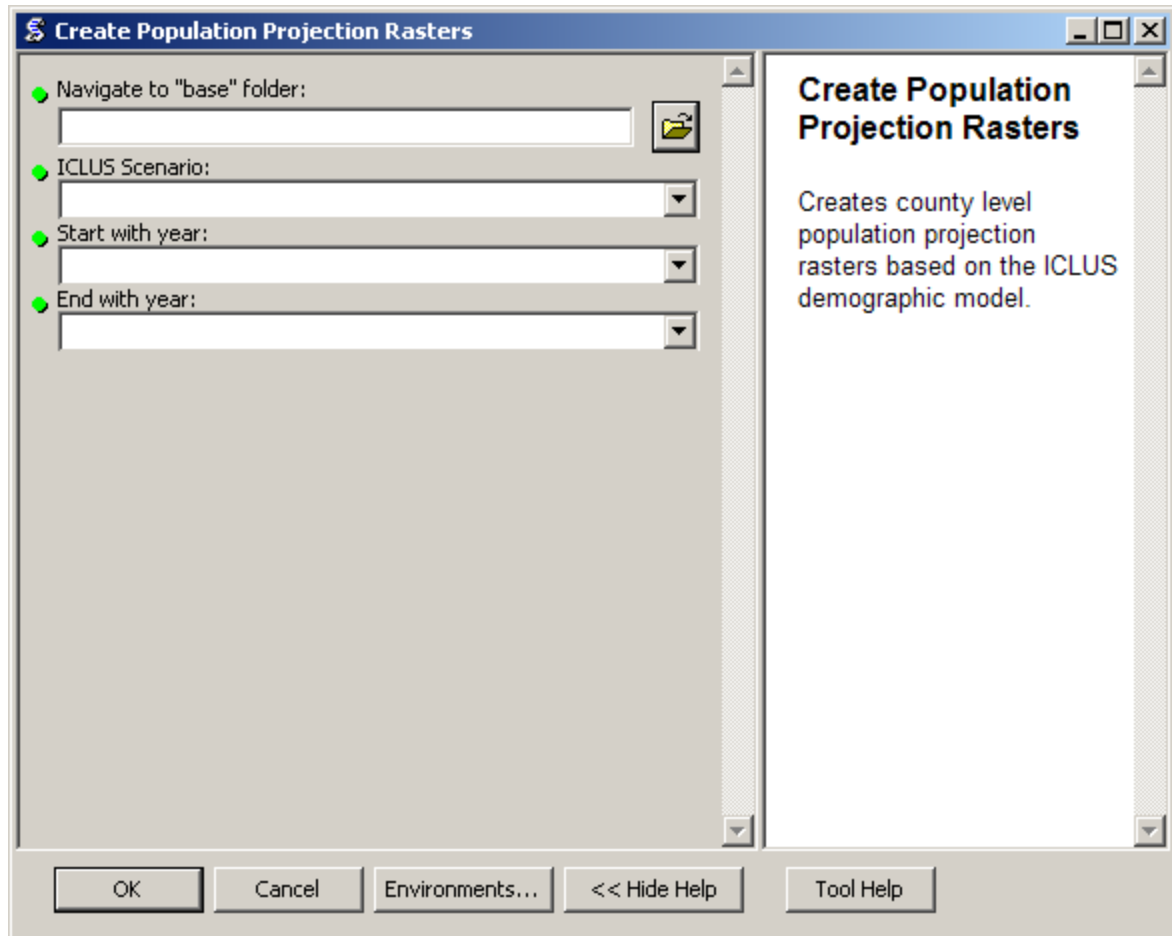
Figure 3-5: Opening Screen for Create Population Projection Rasters Tool

3.4.   RECLASSIFY HOUSING DENSITY TO CLASSES

- This tool reclassifies raw block housing density values into density classes.

- Commercial land pixels, which are static in the ICLUS SERGoM model, may also be incorporated into the output of this tool.
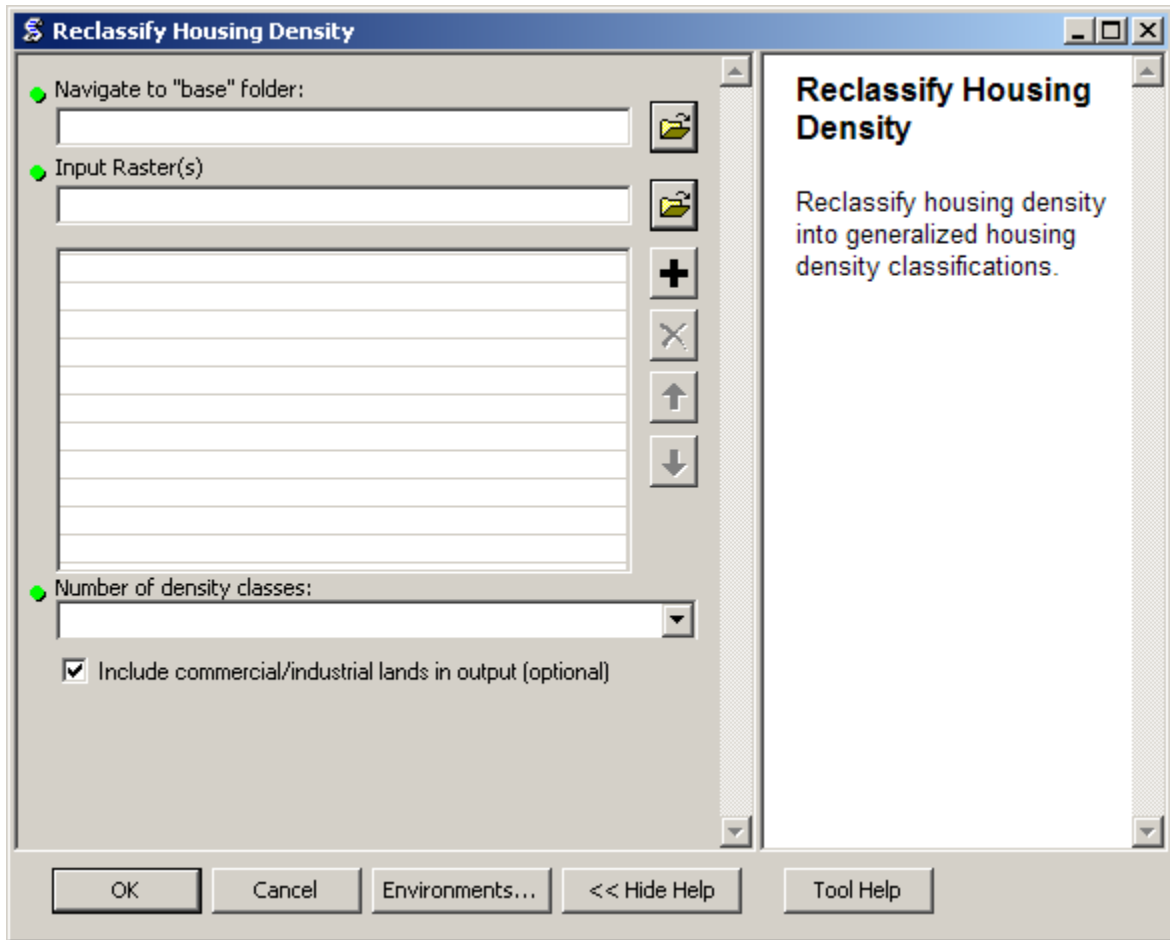
Figure 3-6: Opening Screen for the Reclassify Housing Density Tool

3.5. SERGOM V3 GROWTH MODEL

Populate the required fields. Travel time weights can be adjusted manually to affect the density of growth. The default weights for each scenario are listed in Table 2-2.

Block housing density rasters will automatically be placed in the "output" folder.

Clicking inside any blank field will bring up text in the Tool Help window.
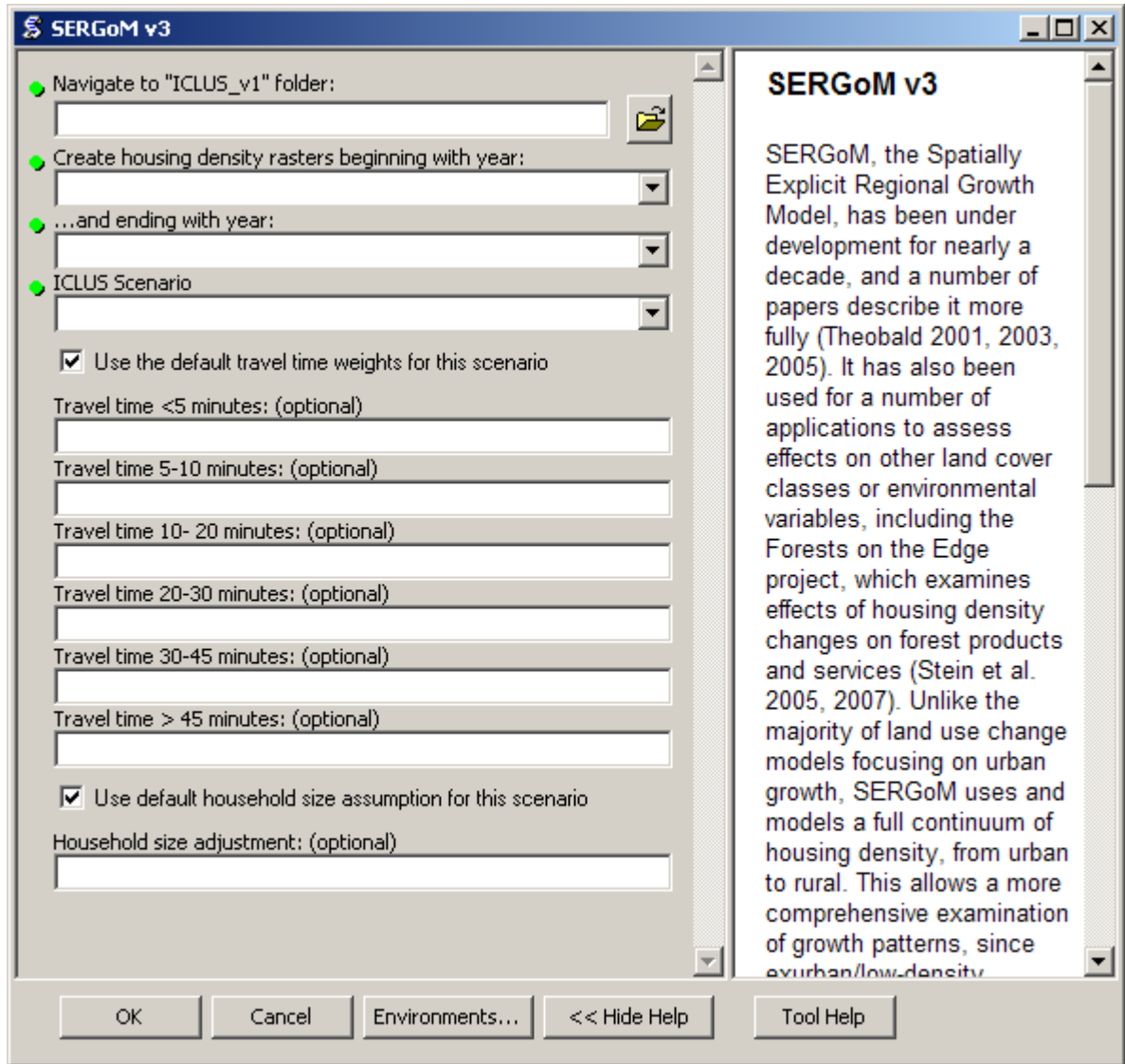
Figure 3-7: Opening Screen for SERGoM v3 Growth Model Tool

Only the fields with green dots are necessary for the tool to run. Selecting the various pulldowns will select the various scenarios one can run through the tool.

3.6.   IMPERVIOUS SURFACE TOOLS

Impervious Surface Tools are described below. These tools are designed to be used with ICLUS generated housing density data at resolutions ≤ 1 kilometer.

3.7.   ESTIMATE PERCENT IMPERVIOUS SURFACE

This tool reclassifies block housing density values into estimated percent impervious categories.
A Categorical Regression Tree model was used to develop a regression equation that describes
the relationship between the Percent Urban Imperviousness (PUI) dataset produced by the
MRLC Consortium and housing density. This relationship was then converted into a set of
conditional statements which form the basis for this tool. For a detailed description of the
methods and sources used to derive the percent impervious surface values, see the full report.
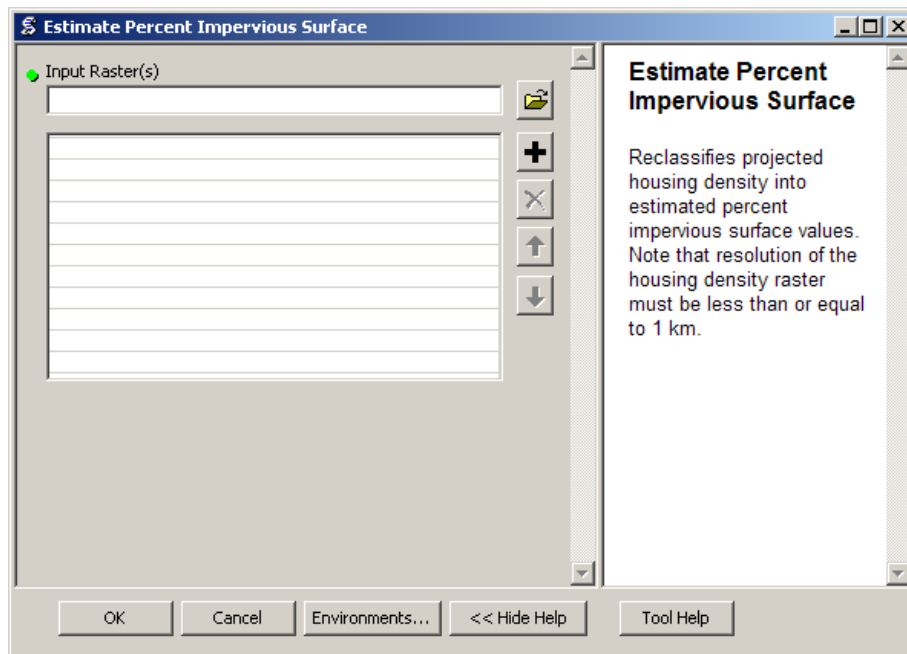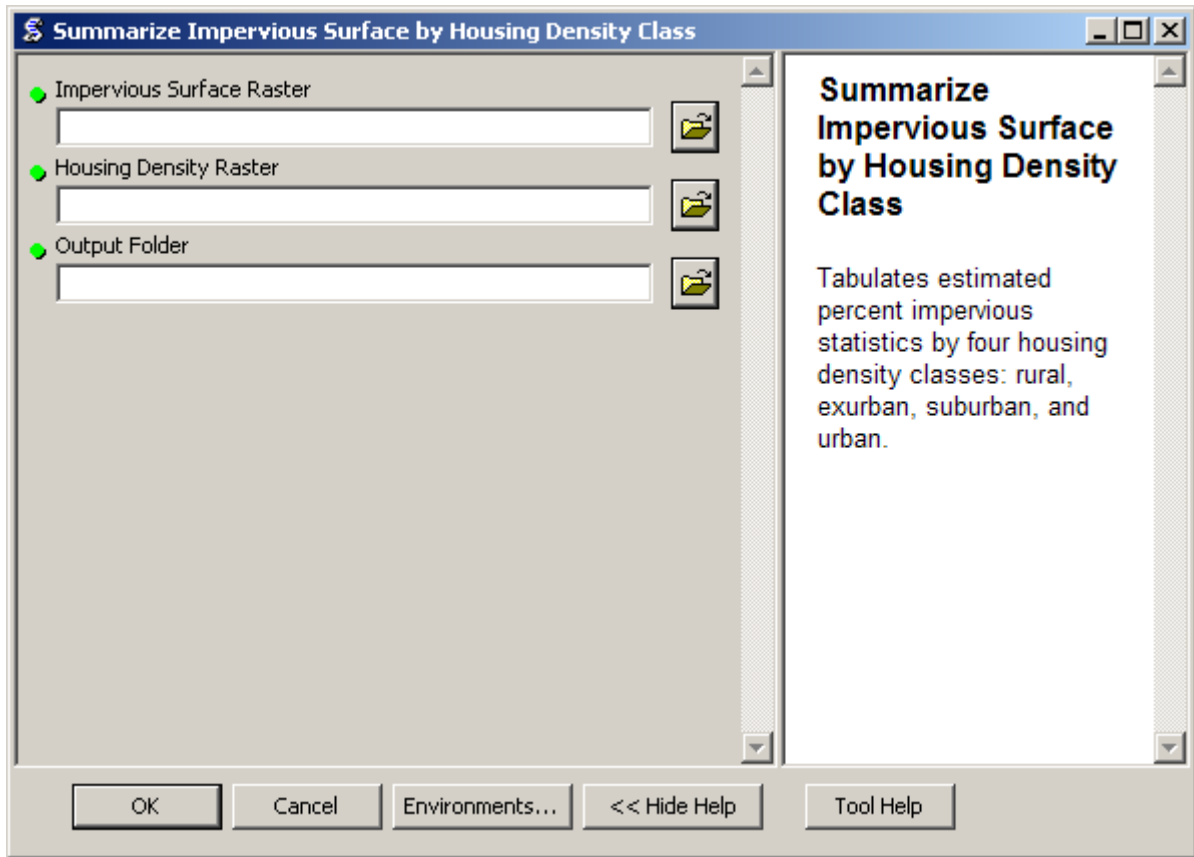

Figure 3-9: Opening Screen for the Estimate Percent Impervious Surface Tool

3.8.   SUMMARIZE IMPERVIOUS SURFACE BY HOUSING DENSITY CLASS

- This tool uses housing density classes to summarize estimated percent impervious surface.

Table 3-1. Housing density class definitions

Class	Zone Value	Density	Housing Density Raster Value
Rural	1	>16.18 ha/unit	< 62
Exurban	2	0.68 - 16.18 ha/unit	62 – 1471
Suburban	3	0.1 – 0.68 ha/unit	1471 - 10000
Urban	4	< 0.1 ha/unit	> 10000

Figure 3-10: Opening Screen for the Summarize Impervious Surface by Housing Density Class Tool

4. REFERENCES

Jiang, L; O'Neill, BC. (2007) Impacts of demographic trends on US household size and structure. Population and Development Review 33(3): 567-591.

Saaty, A. (1980) The Analytic Hierarchy Process: planning, priority setting, resource allocation, McGraw-Hill.

Stein, SM; McRoberts, RE; Alig, R; et al. (2005) Forests on the Edge: housing development on America's private forests. U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station PNW-GTR-636. May 2005. <Online> URL: http://www.fs.fed.us/projects/fote/reports/fote-6-9-05.pdf

Stein, SM; Alig, R; White, EM; et al. (2007) National Forests on the Edge: development pressures on America's National Forest System. U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station, PNW-GTR-728. August 2007.

Theobald, DM. (2001) Land use dynamics beyond the American urban fringe. Geographical Review 91(3):544-564.

Theobald, DM. (2003) Targeting conservation action through assessment of protection and exurban threats. Conservation Biology 17(6):1624-1637.

Theobald, DM. (2005) Landscape patterns of exurban growth in the USA from 1980 to 2020. Ecology and Society 10(1): 32. [online] URL: http://www.ecologyandsociety.org/vol10/iss1/art32/.

U.S. EPA (Environmental Protection Agency). (2009) Land-Use Scenarios: National-Scale Housing-Density Scenarios Consistent with Climate Change Storylines (Final Report). U.S. Environmental Protection Agency, Washington, DC; EPA/600/R-08/076F. Available from the National Technical Information Service, Springfield, VA, and online at http://www.epa.gov/ncea.