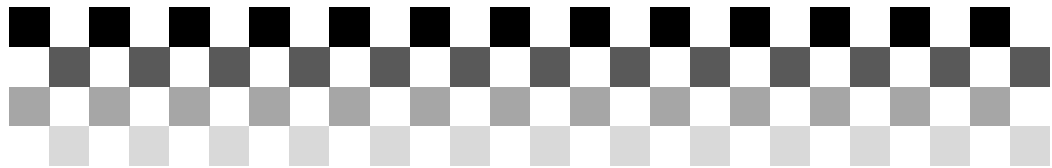# Simulating Urban Growth with the SLEUTH Model:
A Training Manual

Prepared by:
Dr. Claire Jantz
Shippensburg University
Department of Geography-Earth Science
Center for Land Use
June 2009

cajant@ship.edu
717-477-1399

CENTER *for*
Land Use
Shippensburg University

Explore our World With The Department of Geography & Earth Science
Shippensburg

Table of Contents

Introduction

This set of lab exercises is intended to guide you through the use of the SLEUTH model for forecasting alternative future scenarios. To get started, you need:

1. A calibrated model for your study area, including all of the input data sets to go along with that.

2. A good working knowledge of GIS. All of the data preparation and analysis is done in a GIS, and this manual is written for ArcGIS 9.3.

3. The SLEUTH3r model installed on Windows using the Cygwin Linux emulator. (See Appendix A for step-by-step instructions to install Cygwin to run SLEUTH).

This set of labs will walk you through the creation of a scenario map using ArcGIS, running a scenario forecast in SLEUTH, and bringing the output of the scenario forecast back into ArcGIS for analysis. These labs use existing data and a calibrated model for Pike County, PA, but the concepts and methods will apply to any study area.

Every attempt is made in these labs to both present step-by-step instructions and to explain what is being done in each operation. It is impossible, however, to anticipate all questions or provide exhaustive information. Because SLEUTH is such a widely used model, there is a discussion board on the SLEUTH website and a Yahoo e-mail list serv (http://tech.groups.yahoo.com/group/sleuth-users/join).

Some background reading and resources that you may find useful:

- The SLEUTH website: http://www.ncgia.ucsb.edu/projects/gig/

- Jantz, C.A., M. Mrozinski, and E. Coar (2009). Forecasting Land Use Change in Pike and Wayne Counties, PA. Shippensburg University Department of Geography-Earth Science. Final report, April 2009. 42 p.

- Jantz, C.A., S. J. Goetz, D. Donato, P. Claggett (in press). Designing and Implementing a Regional Urban Modeling System Using the SLEUTH Cellular Urban Model. *Computers, Environment and Urban Systems.*
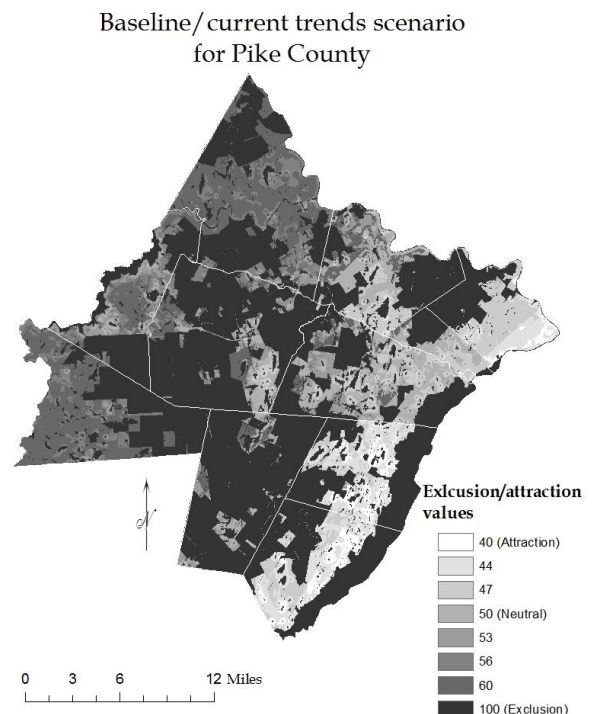
Lab 1

Creating alternative scenarios using exclusion and attraction layers

Introduction

The exclusion/attraction layer is a user-defined map that describes where development is more or less likely to occur in the future. It is the primary means to create alternative future scenarios in SLEUTH. The exclusion/attraction layer is created in a GIS using *raster overlay modeling*. The final layer is exported as a GIF image for input into SLEUTH.

In this lab, you will be taking a variety of raster layers representing different elements that will be included in an excluded/attraction layer. Then, these raster layers will be combined in a series of overlay models to produce the final exclusion/attraction layer. This layer will be exported from ArcGIS as a GIF image file for input into the SLEUTH model.

At the end of this lab, you will have created a layer that looks like this figure, where black represents areas that are excluded from development, gray shows areas that are neutral for development, and white shows areas that will attract development. This is the exclusion/attraction layer that represents the current trends scenario for Pike County and is made up of a combination of layers that include parks, easements, soils, the distance to the New York/New Jersey metropolitan area, population density, proximity to roads, and others.

Baseline/current trends scenario
for Pike County

Exlcusion/attraction values

	40 (Attraction)
	44
	47
	50 (Neutral)
	53
	56
	60
	100 (Exclusion)

0   3   6   12 Miles

Step 1: Copy files

- From the Pike County data disc, copy the entire "Pike_growth_model_data_disc" folder to your local hard drive (~675 MB).

Step 2: Setting up ArcGIS

- Open ArcGIS.

- Turn on the Spatial Analyst extension, Tools → Extensions, check the Spatial Analyst box.

- Add the Spatial Analyst tool bar. Right click on the tool bar and check Spatial Analyst.

- Set your options in Spatial Analyst. *This is an important step to perform each time you start a new map document since it ensures that all of your layers will be the correct extent and cell size.*

  o Spatial Analyst menu → Options

  o Under the General tab, set your workspace where temporary files will be written.

  o Under the Extent tab, set the analysis extent to the 2005 urban layer for Pike County (Pike/Model_Inputs/Urban_Data/pike05).

  o Under the Cell Size tab, set the cell size to the 2005 urban layer.

- SAVE your map document now and frequently!

Step 3: Create a layer that represents all lands that are completely off-limits to development

- Add the following layers to your map document, all of which can be found in the Pike/Model_Inputs/Protected folder or the Pike/Model_Inputs/Exclusions folder. Each of these layers is described in detail in the "Forecasting Land Use Change in Pike and Wayne Counties, PA" final report (Appendix A) that is located on the data disc.

  o Bloom_grove: Blooming Grove hunt club

  o Dela_sf: Delaware State Forest lands

- o Hunt_clubs: All hunt clubs

- o Open_space: A set of open space lands

- o Os_quasi: Privately owned lands that are considered protected (i.e. hunt clubs)

- o Other_clubs: Other club lands that are considered protected

- o Promised_land: Promised Land state forest

- o Water_all: Water bodies

- o Wetlands: Wetlands

- o Rds_excl: Roads must be excluded from (re)development, and are provided to the model as a separate layer.

- Note that most of these layers started out as polygon features, or were derived as distances from point or line features, or were derived from line density (e.g. road density). If you wanted to include new vector features in new scenarios, below are some instructions for doing this.

  - o *Polygon features (i.e. parks):* Polygons that you want to include in the exclusion/attraction layer need to be converted to a raster. I suggest adding a column to your polygon attribute table called "Grid" and use the field calculator to give all features an attribute value of "1." Now, you can use the new Grid attribute as the raster conversion field. To convert a shape file to a raster: Spatial Analyst menu → Convert → Features to Raster.

  - o *Points and Lines:* Point and line features that will be directly incorporated into the exclusion/attraction layer can be handled the same way as polygon features—create a new "Grid" column in the attribute table to use as the raster conversion field. When points and lines are converted into a raster, they are constrained by the designated cell size. For example, when a point file is converted into a raster, each point will be represented as a 28.5 x 28.5 m pixel. Likewise, roads will consist of a line of pixels that are 28.5m x 28.5 m.

  - o *Creating distance buffers:* You may want to create distance buffers around vector data elements (i.e. a 100 m buffer around wetlands that will be off-

limits to development). To calculate distance from point, line or polygon features: Spatial Analyst menu → Distance → Straight Line. The resulting grid will be a continuous raster, where every pixel's value represents the distance of that pixel from the nearest feature. To make things simple to work with, you will probably want to reclassify this grid into discrete classes. First, in the layer's properties, create your desired classification. Then, reclassify the raster: Spatial Analyst → Reclassify.

- o *Handling NoData.* In raster overlay modeling, pixels coded as NoData will persist throughout the analysis. I recommend reclassifying all NoData pixels as "0" before proceeding with the overlay. To reclassify: Spatial Analyst menu → Reclassify.

- o *A few important notes about working with grids.*

    - ▪ Make sure that your shape files are all in the correct projection. This project uses UTM, Zone 18, units meters, datum NAD83.

    - ▪ Make sure that the projection for your data frame is correct.

    - ▪ Your shape file names should not have spaces in them. Replace spaces with "_" or "-".

    - ▪ Grid names cannot exceed 13 characters and cannot have spaces.

- Now you can create the layer that represents lands that are 100% excluded from development.

    - o Use the Raster Calculator to ADD all of your grid layers together. Spatial Analyst menu → Raster Calculator. The expression you create should look something like this: `[bloom_grove] + [dela_sf] + [hunt_clubs] + [open_space] + [os_quasi] + [other_clubs] + [promised_land] + [water_all] + [wetlands]`

    - o Click Evaluate. The resulting grid will appear in your map as a new layer. In this new grid, zero values are lands that are not protected and all values greater than zero are lands that are protected through one or more agencies or programs.

    - o Reclassify this grid so that it only has values of 0 and 100. Spatial Analyst menu → Reclassify. In the reclassification table, 0 should remain 0, but all the other values *(including NoData)* should be classified as 100. You can

also change this output raster from "Temporary" to permanent by giving it a name (i.e. excl100) and folder location.

<u>Step 4: Create a layer that represents lands that are weighted for or against development</u>

- Add the following layers to your map document, all of which can be found in the Pike/Model_Inputs/Exclusions folder. Each of these layers is described in detail in the "Forecasting Land Use Change in Pike and Wayne Counties, PA" final report (Appendix A) that is located on the data disc.

    o Nyc_dist: Distance from New York City classified into 10 categories, where 1 is the closest to New York and 10 is the furthest.

    o Pike_rd_dens: Road density "villages"

    o Pop_dens2: Population density for Pike County's municipalities, classified into four categories, where 1 is the lowest population density and 4 is the highest population density.

    o Rds_dis1500: A 1,500 meter buffer around major roads.

    o Soils: Engineering capabilities of soils, classified into four categories, where 0 = water, 1 = no limitation, 2 = somewhat limited, and 3 = severely limited.

    o Water_dis300: A 300-meter buffer around water bodies.

- Now you can create the layer that represents lands that will either attract or repel development.

    o Use the Raster Calculator to ADD or SUBTRACT all of your grid layers together, using weighting if appropriate. Spatial Analyst menu → Raster Calculator. This is a trial and error process. For Pike County, the expression that was formulated to weight some layers more than others was this: `(([pop_dens2] + [rds_dis1500] + [water_dis300] + [pike_rd_dens]) * 3) - [nyc_dist] - [soils]`. The last two layers are subtracted because of how they are scaled (nyc_dist) or because they represent a resistance to development (soils).

- Click Evaluate. The resulting grid will appear in your map as a new layer. The values in this grid represent the combination of all the layers, but are not meaningful to the model yet.

- Right-click on the layer and go to the Properties menu. Under the Symbology tab, create a new classification with 7 classes according to the natural breaks classification scheme. (Note: for your own purposes, you can experiment with different numbers of classes and classification schemes.)

- Reclassify this grid so that it only has values ranging from 40 to 60. Spatial Analyst menu → Reclassify. In the reclassification table, the classes should be 60, 57, 53, 50, 47, 43, and 40. NoData should be classified as 100. You can also change this output raster from "Temporary" to permanent by giving it a name (i.e. GISModel_1) and folder location. In the reclassified grid, values of 50 are neutral for development, values less than 50 attract development, and values more than 50 will resist development.

Step 5: Include a layer that will be neutral in some areas and resistant in others

- In this project, for the current trends scenario, Act 319 lands were considered to be less effective in areas of high growth pressure. That means that we have to treat Act 319 lands that line up with grid values > 50 (low growth pressure) differently than in places where they line up with grid values < 50 (high growth pressure).

- In ArcGIS we can make this distinction using a conditional statement ("con") in the Raster Calculator (Spatial Analyst menu → Raster Calculator). The "con" command is a powerful and efficient way to combine grids. The general syntax for "con" is: `Con(<condition>, <true_expression>, <false_expression>)`

- To build a layer that includes the Act 319 easements only where growth pressures are low, your conditional statement in Raster Calculator would be: `Con(GISmodel_1>50 & act319==1, 100, GISmodel_1)`. You can "read" this expression as "Where GISmodel_1 is greater than 50 (i.e. low growth pressure) AND where Act319 lands are equal to 1 (i.e. present), recode the output grid to be 100 (i.e. protected from development). Otherwise, recode the output grid so that its values are the same as GISmodel_1."

    - In the resulting output grid, easements that are located in areas where GISmodel1 is greater than 50 will be given the grid value of 100.

Otherwise, the pixels in the new output grid will retain their original value from GISmodel_1.

- o Note: you may have to reclassify this final, but intermediate, grid so that the NoData values are coded as 100. You can then make this grid permanent by giving it a name (i.e. GISmodel_2) and folder location.

<u>Step 6: Add in the lands that are 100% excluded from development and export your final scenario map</u>

- The next step is to add the lands that are 100% excluded from development.

  - o Here we can use another conditional expression in raster calculator. In this case, your conditional expression should be: `Con(excl100==100, 100, GISmodel_2)`.

  - o The resulting grid is your final exclusion/attraction layer for the model. Save it by right-clicking on the layer and choosing the Data → Make permanent option.

- The last step is to export this layer as a GIF image file.

  - o Right click on the layer and chose the Data → Export option. Change the format to GIF. The naming convention for SLEUTH is [location].excluded.[identifier].gif where the terms in [ ] are user-specified. So, for this excluded layer we could name it: pike.excluded.current_trends.gif.

  - o Double-check the file in a Windows file browser. Navigate to the folder where you saved the file. You should see a set of files all with the same name, but that are different types: .aux, .gfw, .gif, .gif.aux.xml, .gif.vat.dbf, and .rrd. The important files for you are the *.gif file* (this is the actual image file) and the *.gfw file* (the GIF world file that will help you bring SLEUTH's outputs back into the GIS in Lab 3).

Lab 2

Running a forecast with SLEUTH

<u>Introduction</u>

Now that you have created a new exclusion/attraction layer, we need to begin working with the SLEUTH model. This lab outlines how to install and use SLEUTH to create a new forecast.

<u>Step 1: Install Cygwin</u>

- If you haven't already done so, follow the instructions in Appendix A to install Cygwin to run SLEUTH. SLEUTH is a Linux program and will not run as a program in Windows, so we need Cygwin, a Linux emulator, to be able to run SLEUTH.

- Some useful Cygwin commands:

    o cd = change directory (cd c:\Geotemp\ SLEUTH3.0d_little)

    o cd .. = move up one directory

    o ls = list files

    o ls | more = list files one screen at a time

    o pwd = print working directory (this tells you where you are in the file structure)

    o rm filename = remove file

    o rmdir directoryname = remove directory

    o cp oldfilename newfilename = copy a file

    o Using TAB autocompletes your file names.

Step 2: Install SLEUTH

- Copy the SLEUTH3r_cygwin directory to your local hard drive. Note SLEUTH's file structure:

  o SLEUTH3r_cygwin\

    ▪ Input\ This is where the input data is located. There are already folders for Pike and Wayne counties.

    ▪ Output\ This is where output from the SLEUTH forecasts is saved. There are already output folders for the mid-range, high and low forecasts for each county.

    ▪ Scenarios\ This is where SLEUTH stores the Scenario files. There are already some scenario files here for you to work with.

    ▪ GD\ This is a folder for a program called GD that handles GIF image files.

    ▪ Whirlgif\ This is a folder for a program called Whirlgif that handles GIF image files.

    ▪ There are lots of other files with .o, .h, .c extensions. These are SLEUTH's program files and are required to run the model.

  o Now you need to (re)compile SLEUTH's code so that it will run on your machine.

    ▪ Open Cygwin and navigate to the directory where you have copied the SLEUTH directory (i.e. cd c:\SLEUTH3r_cygwin).

    ▪ Change directories into the Whirlgif directory (cd Whirlgif). Type the command "make clean" (to clean up old program files), then type the command "make" to compile the Whirlgif program.

    ▪ Change directories into the GD directory (cd ../GD). Type "make clean" and then type "make."

    ▪ Go back up to the main SLEUTH directory (cd ..). Type "make clean" and then type "make." Now SLEUTH is ready to run.

- Most error messages you get while compiling are minor and should not affect the running of the program.

- o Test SLEUTH to make sure it is installed correctly.

  - Change directories into the Scenarios folders (cd Scenarios).

  - List the files (ls). You will see several existing scenario files.

  - Type ../grow.exe predict scenario.pike_mid_range

  - If SLEUTH starts running (you will start to see SLEUTH step through the forecast years), it is installed correctly. Interrupt SLEUTH by typing Ctrl C.

Step 3: Copy your new scenario map to the correct input directory

- In Windows, copy the pike.exclusion.current_trends.gif file to SLEUTH3r_cygwin/Input/pike.

Step 4: Edit the scenario files

- Back in Cygwin, make sure that you are in the Scenarios folder (pwd to see where you are), and navigate to it if necessary. Type ls again to see the existing scenario files.

  - o Note the naming convention for these files, scenario.[location].[identifier]

  - o Note that there is a separate scenario file to produce the mid-, high- and low-range forecasts.

  - o These scenario files are already set to run the calibrated model for both counties. You will only need to change a few things, as outlined below, to run forecasts with your new scenario map.

- The scenario file is a simple text file where you set up the parameters to run the model. Even though it looks complicated, to run predictions you only need to make a couple of changes to the scenario file. DO NOT use Word, Notepad, or Wordpad to edit the scenario file. You need to use a text editor that can save the output in Unix or Linux format. I suggest TextPad (http://www.textpad.com/index.html). A free trial version is available for download, and a single-user license only costs around $30.00.

- You can also use a text editor in Cygwin called Pico. This manual will include instructions for editing in Pico. We will start with the mid-range forecast file. To open the file in Pico, type `pico scenario.pike_mid_range`.

    - Note that Pico does not respond to the mouse; it is all between you and your keyboard.

    - At the bottom of the Pico screen is a series of keyboard commands. The "^" symbol means to hold down the Ctrl key. So, for example, if you want to scroll up a page the command is Ctrl Y and to scroll up the command is Ctrl V. Save is Ctrl O ("WriteOut") and Exit is Ctrl X.

    - In the scenario file, all the lines that begin with "#" are comment lines, meaning they are there for your benefit and are not read in by the model. Any line that does not begin with a # is information that the model reads in.

- Edit or check the following sections in the scenario file:

    - Section I. Make sure the correct input and output directory names are correct. For these purposes they probably do not need to be changed, but if you were to create a new output folder for a new scenario, you would need to 1) create the directory and 2) specify the new directory name in this location in the scenario file. Also make sure that there is a slash (/) at the end of the Input and Output directory lines.

    - Section VII. Check the number of Monte Carlo iterations. Ideally you would want to run 100 iterations, but if you have a slow computer or are pressed for time, you can drop the number as low as 25.

    - Section IX. Prediction date range. The current setting is to create forecasts from 2005 to 2030. You will only need to change this if you want to forecast out to a different year.

    - Section X. Change the name for the excluded input image to match the new layer you created for this scenario.

- Exit Pico (Ctrl X) and key in "Y" when Pico asks if you want to save changes. Note that Pico will give you the opportunity to save to a new file at this point, but for our purposes we can keep the same file name.

- If you also want to create high- and low-range forecasts you will need to edit the corresponding scenario files, making the same changes as you did with the mid-range file. The one difference is to check the self-modification parameters in section XIII.

    o The last two lines of the file are the "BOOM" and "BUST" multipliers. In the high-range scenario file, the model is set to go into boom mode with a multiplier of 1.02. Since the multiplier is greater than one it will increase the growth rate through time. Likewise, the low-range file is set to go into bust mode with a multiplier of 0.95.

    o The only reason you would need to change these numbers is if you run a forecast and find that the amount of growth is more or less than what might be expected to occur.

- Finally, enter the command to begin running the model: `../grow.exe predict scenario.pike_mid_range`

Step 5: Verify the outputs

- In Windows, navigate to SLEUTH3r_cygwin\Output\pike_mid_range and check the output files.

- SLEUTH generates a lot of output files:

    o An image for each year in the forecast. These images are classified according to probabilities of development and are generally only useful for visualization, not for further data analysis.

    o A "cumcolor_urban" image for the final year of the forecast. Again, this image is classified according to probabilities of development and is generally only useful for visualization.

    o A "cumulate_urban" image. This image has values that range from 0-100 according to each pixel's probability of development in the final forecast year. If you want to do further analysis in a GIS, this is the image to use.

    o Log files (see the SLEUTH website for a full description):

        ▪ An avg.log file that contains averaged metrics for your run. The "pop" or "area" column in this file has the total number of urban pixels that exist in each year. You can convert this pixel count to an

area measurement by multiplying it by the area of each pixel (28.5 x 28.5 = 812.25 m2).

- A std_dev.log file that contains the standard deviation of the urban metrics.

- A coeff.log file that keeps track of the coefficient values in each year. This file is only useful if we apply SLEUTH's "self-modification" function, as in the high- and low- range forecasts.

- Each of these log files can be opened in Excel as will be covered in Lab 3.

- If all the output files appear to be okay, exit Cygwin (Ctrl X).

Lab 3

Working with SLEUTH's output files

<u>Introduction</u>

As you found in lab 2, SLEUTH generates a lot of output files. Some are images that can be analyzed in a GIS, and some are log files that can be imported into Excel. This lab will focus on working with SLEUTH's outputs in ArcGIS and Excel.

<u>Step 1: Getting a SLEUTH image back into ArcGIS</u>

- As noted in Lab 2, SLEUTH's best and most complete forecast map is the "cumulate_urban" image, which shows the probability for development in each pixel for the final forecast year. However, SLEUTH only outputs the raw GIF images, not the spatial reference information. So, we have to provide the spatial information so that the image can be brought into ArcGIS.

- First, I suggest copying the cumulate_urban GIF image from SLEUTH's output directory to a directory where you want to work with it. Then, make a copy of the GIF world file (.gfw) that you created at the end of Lab 1 and place it in the same directory with the image. Last, give the .gif image and the .gfw file *the exact same name*. As long as the .gif image and the .gfw file are in the same directory with the same name, ArcGIS will automatically associate them when you load the image into the GIS.

- Start ArcMap. Make sure to set your extensions and the Spatial Analyst options as outlined in Lab 1.

- Before adding the cumulate_urban image to the map document, you will need to define its project (ArcToolbox → Data Management Tools → Projections and Transformation → Define Projection)

    o If it is not added to the document automatically, add the cumulate_urban image after the projection is defined.

- The image will appear in grayscale. To change the appearance, right-click on the layer and go to the properties menu. Under the Symbology tab, choose a "Classified" classification and color scheme. The image should have values that

range from 0 to 100 that represent the probability of each pixel becoming developed by 2030. You can use this map as is, or continue on for further analysis.

Step 2: Summarizing growth information to areas

- It is often useful to summarize the forecast information to areas, like municipalities or watersheds, to show how much development is forecasted for each of those areas.

- As an example, we will calculate the amount of forecasted growth for municipalities in Pike County. Add the pike_muni.shp file to the map document (located in the municipal_boundaries folder).

- Now we will perform a Zonal Statistics operation that will summarize information from the raster layer to the municipalities.

    o Spatial Analyst menu → Zonal Statistics

    o The "Zone dataset" should be the pike_muni.shp.

    o The "Zone field" should be the ID—or any field that has a unique value for each polygon.

    o The "Value raster" is the cumulate_urban image.

    o Check the "Join output table to zone layer" option and uncheck the "Chart statistic" box.

    o Save the table then click okay.

- Open the table for the pike_muni.shp (right click → Open Attribute Table). You will notice several new fields have been joined to this table: Value, County, Area, Min, Max, etc. These are all the statistics that have been summarized for the raster for each municipality. So, for example, the minimum raster value in each municipality is 0 and the maximum value is 100.

- You can use the SUM field to calculate the area of development.

    o First, add a new field to the table. Attribute table Options → Add field. Give the field a name that makes sense (i.e. Dev2030_ct) and make it a floating point field.

    o Right-click on the title of the new field and choose the "Field Calculator" option. Build this expression to calculate the area of development in

square meters: `([SUM]/100)*812.25` where [SUM]/100 is equal to the number of pixels that are forecasted to be developed and 812.25 is the area in square meters of each 28.5 x 28.5 m pixel.

- o You can also calculate a variety of other statistics (i.e. % developed area) by adding new fields and using the field calculator.

- o The same zonal statistics operation can be used for any shape file (i.e. watersheds).

Step 3: Working with SLEUTH's log files in Excel

- In Lab 2 you found out that SLEUTH also produces a lot of tabular files. Each log file is described in detail on the SLEUTH website (http://www.ncgia.ucsb.edu/projects/gig/).

- One of the most useful files is the avg.log file. You can open this file in Excel.

  - o Open Excel and go to File → Open.

  - o Navigate to the directory where the log files are stored. You will have to view all file types to be able to see the log files.

  - o Select the avg.log file. In the Text Import Wizard, select "Delimited" and click "Next." Then, click the "Space" option and click "Next" through the next screen, then click "Finish."

  - o Of the 27 attributes that SLEUTH writes to the avg.log file, one of the most useful is the "pop" or "area" fields. These fields contain the same information: the number of urban pixels present in the whole county in each year. This pixel count can be converted into an area because you know that each pixel is 28.5 x 28.5 m, or 812.25 m². Once this calculation is complete, you can create graphs that compare different scenarios.

Appendix A: Installing Cygwin to run the SLEUTH model