GEOSPATIAL MODELING OF POPULATION GROWTH SCENARIOS FOR THE HUMBOLDT BAY, CALIFORNIA REGION: ADAPTING SLEUTH TO A RURAL ENVIRONMENT

by

Jess A. Morrison

A Thesis
Presented to
The Faculty of Humboldt State University

In Partial Fulfillment
Of the Requirements for the Degree
Masters of Science
In Natural Resources, Planning and Interpretation

July, 2012
GEOSPATIAL MODELING OF POPULATION GROWTH SCENARIOS FOR
THE HUMBOLDT BAY, CALIFORNIA REGION: ADAPTING
SLEUTH TO A RURAL ENVIRONMENT

by

Jess A. Morrison

Approved by the Master’s Thesis Committee:

______________________________
Steven J. Steinberg, Major Professor

______________________________
Sharon Tuttle, Committee Member

______________________________
Dennis Fitzsimons, Committee Member

Coordinator, Natural Resources Graduate Program

Date

12-ES-786-07/27

Natural Resources Graduate Program Number

______________________________
Jena' Burges, Vice Provost

Date
ABSTRACT

Geospatial Modeling of Population Growth Scenarios for the Humboldt Bay, California Region: Adapting SLEUTH to a Rural Environment

Jess A. Morrison

Humboldt County is a modestly populated region on the northern coast of California. Much of the county is held in State and National parks and unavailable for development. While growth rates remained steady for several decades, the region is not immune to population growth and poorly-planned expansion. Geospatial modeling provides a means to explore various development scenarios examining the likelihood of urban sprawl encroaching upon agricultural lands and other protected areas. Communities surrounding Humboldt Bay are particularly susceptible to rapid growth, representing approximately 60 percent of the population. Understanding impacts of growth around existing population centers is necessary to assess potential benefits of regional smart growth strategies. The SLEUTH urban growth model uses cellular automata, terrain mapping, and land cover modeling to generate potential population growth scenarios. It has been successfully applied to numerous metropolitan areas; however, its application to rural environments has not been fully explored. Using SLEUTH, I generated multi-scenario population growth estimates projecting 100 years into the future. These will inform city and county planning departments in developing smart and sustainable growth strategies and provide resource managers information useful in selecting an optimal expansion plan for the region.
ACKNOWLEDGMENTS

I extend a thank you first and foremost to Dr. Steve Steinberg for his support and guidance in the development and completion of this project, even after his departure from Arcata, and also for the wealth of geospatial knowledge he imparted to me over the course of my study at Humboldt State University. I would also like to thank Dr. Dennis Fitzsimons for providing encouragement and broadening my horizons by introducing me to new ways of looking at maps. In addition, Dr. Sharon Tuttle's computing knowledge and tireless enthusiasm for teaching were invaluable in giving me the necessary skills to complete this project.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>iii</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>iv</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>vi</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>vii</td>
</tr>
<tr>
<td>LIST OF APPENDICES</td>
<td>viii</td>
</tr>
<tr>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>MATERIALS AND METHODS</td>
<td>11</td>
</tr>
<tr>
<td>RESULTS</td>
<td>21</td>
</tr>
<tr>
<td>DISCUSSION AND CONCLUSIONS</td>
<td>31</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>34</td>
</tr>
</tbody>
</table>
### LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Humboldt County Population from 1960-2010 with Department of Finance Predictions for 2020-2050 and percent population change by decade.</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>Computer processing times for each model process and sub-process in days, hours, minutes, and seconds.</td>
<td>20</td>
</tr>
<tr>
<td>3</td>
<td>Accuracy assessment showing producer's accuracy for classifications of the 1984 Landsat image, including the unsupervised classifications of TM and no TM composites, and the supervised classification of the raw image.</td>
<td>22</td>
</tr>
<tr>
<td>4</td>
<td>Accuracy assessment showing producer's accuracy for classifications of the 2011 Landsat image, including the unsupervised classifications of TM and no TM composites, and the supervised classification of the raw image.</td>
<td>23</td>
</tr>
<tr>
<td>5</td>
<td>Selected statistics from Scenario 1 process including changes in coefficient values, total number of urbanized pixels, and characteristics of predicted expansion areas, including mean size of clusters and number of edges.</td>
<td>28</td>
</tr>
<tr>
<td>6</td>
<td>Selected statistics from Scenario 2 process including changes in coefficient values, total number of urbanized pixels, and characteristics of predicted expansion areas, including mean size of clusters and number of edges.</td>
<td>29</td>
</tr>
<tr>
<td>7</td>
<td>Selected statistics from Scenario 3 process including changes in coefficient values, total number of urbanized pixels, and characteristics of predicted expansion areas, including mean size of clusters and number of edges.</td>
<td>30</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

Figure                                                                                                                             Page

1 Location of the Humboldt Bay, California region and the communities included in this SLEUTH analysis with a shaded relief background................. 2

2 Diagram of SLEUTH input images, needed starting conditions, and model process flow (Clarke, 2004)............................................................... 7

3 SLEUTH model workflow including all processes, coefficients, and steps needed to complete each growth scenario..................................................... 18

4 Initial urban conditions in the Humboldt Bay Region, 2011. Urbanized areas are shown in pale yellow................................................................. 24

5 Scenario 1 urbanized predictions, 2110. Original 2011 urban extent is shown in pale yellow. Areas in red have 90-100 percent probability of urbanization, non-ocean blue areas are 30-40 percent and green are 1-10 percent...................... 25

6 Scenario 2 urbanized predictions, 2110. Original 2011 urban extent is shown in pale yellow. Areas in red have 90-100 percent probability of urbanization, non-ocean blue areas are 30-40 percent and green are 1-10 percent...................... 26

7 Scenario 3 urbanized predictions, 2110. Original 2011 urban extent is shown in pale yellow. Areas in red have 90-100 percent probability of urbanization, non-ocean blue areas are 30-40 percent and green are 1-10 percent...................... 27
## LIST OF APPENDICES

<table>
<thead>
<tr>
<th>Appendix</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Selected log data from Scenario 1 SLEUTH run, edited for clarity. This log shows all the necessary components to execute an identical SLEUTH run, including input details, coefficient values, random number seed, and number of Monte Carlo iterations.</td>
<td>38</td>
</tr>
<tr>
<td>B</td>
<td>Selected log data from Scenario 2 SLEUTH run, edited for clarity. This log shows all the necessary components to execute an identical SLEUTH run, including input details, coefficient values, random number seed, and number of Monte Carlo iterations.</td>
<td>42</td>
</tr>
<tr>
<td>C</td>
<td>Selected log data from Scenario 3 SLEUTH run, edited for clarity. This log shows all the necessary components to execute an identical SLEUTH run, including input details, coefficient values, random number seed, and number of Monte Carlo iterations.</td>
<td>46</td>
</tr>
</tbody>
</table>
INTRODUCTION

Humboldt County, California is situated between Del Norte and Mendocino counties on California’s North Coast (Figure 1). It covers approximately 930,000 hectares and is characterized as rural with a density of 0.15 persons per hectare. Recent estimates place the total population at 134,623 (U.S. Census Bureau, 2011). With just over 600 hectares of timberland area, Humboldt is second in the state of California in forestry/timberland only to Siskiyou County. Although its influence has declined, the county’s economy and population still depend heavily on the timber industry. In 1992, 22 percent of jobs in Northern California were in the forest products industry. As of 2000, Humboldt County still accounted for 19.7 percent of California’s timber harvest volume (Laaksonen-Craig, 2003). The region is also home to Humboldt Bay, one of California’s largest coastal estuaries and the only major commercial port on the Pacific coast between San Francisco and Coos Bay, Oregon (Barnhart et al, 1992).

The area around Humboldt Bay is the population center of Humboldt County, with approximately 60 percent of the county’s residents living in this region. The majority of recent growth is concentrated in the cities of McKinleyville and Arcata, and in the Eureka suburbs of Humboldt Hill and Cutten (Figure 1). Although Humboldt is the largest county on California’s north coast, 80 percent of its lands are undevelopable, protected forestlands (California Economic Development Department, 2010). While most of the growth areas have an adequate supply of land and public facilities needed to support additional expansion, upturns in growth could prove problematic if not
Figure 1. Location of the Humboldt Bay, California region and the communities included in this SLEUTH analysis with a shaded relief background.
adequately planned. My study focuses on the primary growth areas of the county: the incorporated cities of Eureka, Arcata, and Fortuna and the unincorporated city of McKinleyville. The region is characterized as the Eureka-Arcata-Fortuna Micropolitan Statistical Area by the U.S. Census Bureau (2003).

Since 1960, Humboldt County has experienced just over 45 percent total population growth. The growth rate varied by decade, from a low of about -4.96 percent 1960 through 1970, to a high of approximately 9.77 percent 1980 through 1990 (Table 1) (U.S. Census Bureau, 1995). An economic downturn in 2007 reversed this trend, and updated projection figures from the California Department of Finance predict a sharp drop in population growth over the next 40 years (Table 1) (2007). Nevertheless, these are only predictions and it is impossible to determine how economic and other factors may influence future growth. Using models to simulate the continuation of current growth trends is an excellent way to plan for changes that could be otherwise difficult for city and county planning departments to conceptualize. Often, the most challenging aspect of conveying the relevance of Geographic Information System (GIS) modeling is communicating the results to policy makers. My study aims to bridge that gap using animated dynamic mapping. Powerful analysis and visualization tools such as these improve the interpretability and communicability of results and provide additional troubleshooting data during the stages of model calibration (Clarke et al., 1996). The resulting visualizations are particularly helpful when conveying a model’s findings to the general public and non-model users (Hoppen et al., 1996).
Table 1. Humboldt County Population from 1960-2010 with Department of Finance Predictions for 2020-2050 and percent population change by decade.

<table>
<thead>
<tr>
<th>Humboldt County</th>
<th>Year</th>
<th>Population</th>
<th>Percent change</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1960</td>
<td>104,892</td>
<td></td>
<td>Actual</td>
</tr>
<tr>
<td></td>
<td>1970</td>
<td>99,692</td>
<td>-4.96</td>
<td>Actual</td>
</tr>
<tr>
<td></td>
<td>1980</td>
<td>108,514</td>
<td>8.85</td>
<td>Actual</td>
</tr>
<tr>
<td></td>
<td>1990</td>
<td>119,118</td>
<td>9.77</td>
<td>Actual</td>
</tr>
<tr>
<td></td>
<td>2000</td>
<td>126,839</td>
<td>6.48</td>
<td>Actual</td>
</tr>
<tr>
<td></td>
<td>2010</td>
<td>134,785</td>
<td>6.26</td>
<td>Actual</td>
</tr>
<tr>
<td></td>
<td>2020</td>
<td>142,167</td>
<td>5.48</td>
<td>Predicted</td>
</tr>
<tr>
<td></td>
<td>2030</td>
<td>147,217</td>
<td>3.55</td>
<td>Predicted</td>
</tr>
<tr>
<td></td>
<td>2040</td>
<td>150,121</td>
<td>1.97</td>
<td>Predicted</td>
</tr>
<tr>
<td></td>
<td>2050</td>
<td>152,333</td>
<td>1.47</td>
<td>Predicted</td>
</tr>
</tbody>
</table>

Total Change (%) 1960-2000: 45.23
A variety of urban modeling approaches have been developed and explored over the last several decades. Traditional urban growth prediction methods have been criticized by numerous authors for their lack of spatial reference, static nature, limited temporal capabilities, and failure to address the physical characteristics of developing cities (e.g. Batty et al., 1989; White and Engelen, 1993; Clarke et al., 1997; Chen et al., 2002). These shortcomings prompted a marked increase in the development of complex, dynamic models to better represent the nature of urban expansion (Couclelis, 1985; Batty et al., 1989). The broader implications of these models was more fully realized in the 1990’s, sparking significant research development and a number of further applications (White and Engelen, 1993; Batty and Longley, 1994; Clarke et al., 1995, 1996, 1997; Batty et al., 1999). Some of the most commonly utilized tools in the modern urban planning field include artificial intelligence, agent-based models (ABM), multi-criteria evaluation (MCE), genetic algorithms, and cellular automata (CA) (Riveira and Maseda, 2006; Torrens and O’Sullivan, 2000). Modeling complex spatio-temporal processes such as the development of cities necessitates an organic, dynamic model, strongly consistent with the properties of a CA-based approach (Batty et al., 1999). This connection explains the rise in popularity such models have seen over the past two decades. The use of cellular automata in urban modeling originated in part from the research of Batty and Longley, who identified the ability of diffusion limited aggregation (DLA), a technique closely related to cellular automata used for simulating dendritic growth, to more accurately replicate urban form (Batty et al., 1989; Batty and Longley, 1994). According to their research, DLA provided an excellent starting point for the development of urban
growth models. Couclelis (1985) showed that complex patterns could arise from simple growth rules using CA, and brought attention to the possible applications for urban simulation. In the early 1990’s, White and Engelen (1992, 1993) pioneered one of the first urban growth and land use models implementing CA, paving the way for two decades of CA-based models.

Although they were originally created for mathematics and biological sciences, cellular automata have since been used in many fields to model and predict natural phenomena. Cellular automata were originally developed by Ulam in the 1940’s, and implemented soon after by John von Neumann in the study of self-replicating systems (Madden, 2009). In their most basic form, cellular automata are characterized by a discrete grid of cells, each cell having a value within a finite set of parameters. A cell’s value depends only on its surrounding cells (its neighborhood) and the entire grid evolves in discrete time steps defined by its transition rules (Clarke and Gaydos, 1998; Madden, 2009). In urban growth models, the cells correspond to pixels and a cell’s value can be a simple urbanized/non-urbanized binary system or a qualitative land use system (e.g. 1 – Urban, 2 – Agriculture, 3 – Forest). The starting values of the cells are then dictated by the starting date of the calibration data, and the transition rules governed by the specific model’s parameters (Clarke and Gaydos, 1998).

For this study I selected an urban growth model known as SLEUTH. The SLEUTH model derives its name from the input data layers used: Slope, Land cover, Exclusion, Urban extent, Transportation, and Hillshade (Clarke and Gaydos, 1998) (Figure 2). The
Figure 2. Diagram of SLEUTH input images, needed starting conditions, and model process flow (Clarke, 2004).
model itself is a C-language computer program using pseudo-random numbers generated by the standard C math library function `rand()`. These pseudo-random numbers are multiplied by a counter for each seed, or starting point. Every iteration of the model is unique, but can be replicated using a known seed. The simulation begins with initial seed values and from this point the cellular automata rules are applied iteratively in each growth cycle. Operationally, the model has an outer loop that repeatedly executes each growth history, saving statistical and cumulative data for Monte Carlo operations. Monte Carlo is one variety of stochastic simulation which produces probability distributions using random number generators (Chan, 2001). This allows confidence margins to be calculated for the model’s predicted growth trends, a critical component given the potentially controversial nature of such predictions (Clarke and Gaydos, 1998). An inner loop runs the cellular automata with each sequence processing the data layer through one time cycle, typically representing a year (Clarke et al, 1997). The model was designed for, and typically used in, urban growth scenarios, and was selected because of its successful implementation in numerous metropolitan areas worldwide. In fact there have been over 100 different SLEUTH applications to date, providing an informative body of research (Clarke, 2008). In my study, the model is applied to a much smaller geographic region and population center than previous research. However, this does not exclude it as a candidate for representation by this model. Given the necessary input data and calibration, SLEUTH should model the spread of inhabited land for any populated area, regardless of density or size. In fact, the original research that prompted creation of the model was based on wildfire spread, but the initial growth rules were general enough to
model any type of organic development given the proper modifications. For SLEUTH, the parameters were specifically adapted to simulate urban expansion (Clarke et al., 1996).

The SLEUTH model takes into account four types of urban growth behavior: spontaneous growth, diffusive growth, organic growth, and road-influenced growth. Five coefficients control these growth types, each applied sequentially during the specified growth cycle. The growth coefficients are: diffusion, breed, spread, slope resistance and road gravity (Clarke and Gaydos, 1998). Executing the model is a three step process, including processing and creation of input data followed by model calibration and ending with scenario design and utilization of derived growth coefficients in prediction.

The first, and often most time-consuming, phase of the modeling process is data layer preparation. This step is time-intensive due to SLEUTH’s precise input data specifications. The slope layer is usually derived from a Digital Elevation Model (DEM) and must be expressed as percent slope. The land cover layer is only needed if the project scope necessitates modeling of urban growth and land use change. The exclusion layer represents water bodies and/or any protected or undevelopable land areas to remove from consideration. Classification can be in binary form (1 for developable, 0 for not developable) or can be a scaled value from zero to 100, representing relative barriers to development. An example would be marshlands that are developable, but would present regulatory or economic difficulties that should act as an obstruction to development (Clarke, 2004). The urban extent layer can also be set up with a binary classification or graded categories. Consequently, the definition of urban extent is determined by the
individual creating the input data set, and necessarily varies given the purpose of the study. The transportation layer can also follow the binary or weighted classification method. Pixels can be classified as road/non-road, or an accessibility weighting value between zero and 100 can be assigned to reflect the urban growth influence of highways over unimproved roads. The final input layer, the hillshade, is included for visualization purposes only and may aid analysis by providing spatial and topographic context to the urban extent data (Clarke, 2004).

The second, and most rigorous, phase of the modeling process is calibration. Historic urban classification data is analyzed by SLEUTH to generate optimal coefficients for modeling known growth. The number of historical files required depends on the scope of the project and data availability, with a minimum of four urban extent and two road layers. Generally speaking, the more calibration data used, the more robust the results of the calibration process. Calibration sources may include satellite imagery, aerial photography, topographic maps, or other verified geospatial references. Because of the importance of quality and consistent calibration data, availability of calibration sources dictates the scope of a SLEUTH modeling project.

Road networks and excluded development areas help establish constraints and trends in the model. The calibration process calculates ideal values for the growth coefficients based on historical trends, with different variables representing different types of growth. Once the calibration for historical data is complete and accuracy estimates are known, the model can be applied to any desired future scenario.
MATERIALS AND METHODS

Data Preparation

The SLEUTH model requires input data formatted to specific standards allowing for little flexibility in final data processing. All inputs must be in raster format as grayscale GIF (Graphics Interchange Format) images. All images must also be in the same map datum and projection, covering exactly the same geographical extent. Image resolution also must be consistent, and strict naming conventions are required. The benefit of these standards is that the geospatial processing steps for each data layer within a specific category are identical. Each category of input data requires its own geoprocessing steps, which I implemented as follows.

The slope layer was derived from a United States Geological Survey (USGS) 30m resolution Digital Elevation Model (DEM) (USGS, 2010a). The DEM was clipped to the project extent. Slope values were calculated in percent using the ArcGIS Spatial Analyst extension. The resulting image was then converted to an eight-bit GIF image.

Land cover data are optional and only required if land use changes are included in the model. For this study, I examined six classes of land cover change: forest, water, urban, agriculture, barren, and rangeland. These classes represent the range of landscape variation in the study region without being overly complex. The current USGS National Land Cover Dataset (NLCD), for example, uses a 15 class system. Given that the goal of this analysis was to provide actionable information to city planners and policy makers the
additional complexity of a higher level classification system was determined to outweigh its potential benefits.

The exclusion layer represents areas of land prohibited or discouraged from urban growth, depending on the detail of classification. The exclusion layer was assembled from water features and protected lands. The water features were based upon the USGS Land Use Land Cover (LULC) dataset and the USGS National Hydrography Dataset (NHD) (USGS, 2010b). These were combined with a statewide exclusion layer of protected lands including Department of Defense and Bureau of Indian Affairs lands (Cal-Atlas, 2010). Williamson Act farmland was added as an additional constraint in the exclusion layer because of the prevalence of agricultural lands in Humboldt County. The Williamson Act is a program in the state of California that reduces property taxes for farm or related open-space land owners, in return for restricting their parcel’s use. Although the program’s initial agreement term is only 10 years, the contract automatically renews unless either party files a notice of non-renewal (California Department of Conservation, 2007). According to the Humboldt County 2010 Williamson Act GIS dataset (Humboldt County, 2010), over half of the region’s developable lands are parcels protected under the act, making it a critical component of the exclusion layer. Once combined, these data were merged to develop a single exclusion layer for the analysis. This layer was converted to eight-bit pixel depth, clipped to the study area, and exported to a GIF file.

Urban extent was derived from USGS LULC maps (USGS, 2012). This decision was based on the need for consistency of the input layers in the model. The model
calibration process is dependent upon the accuracy of historical data, so maintaining as much consistency as possible across the range of input years is crucial to the analysis. The USGS LULC data are derived from Landsat Thematic Mapper (TM) satellite images, and the USGS maintains a database of LULC maps dating back to 1980. Given the date range of existing Landsat imagery and USGS land cover data, my two primary sources for modeling historical inhabited land extent, the range of urban growth years available was limited to that time frame. Given these constraints, I selected five dates: 1980, 1992, 2000, 2006 and 2011 for calibration of the model. USGS LULC data was available for four of the five calibration years (1980, 1992, 2000 and 2006), but I ultimately deemed 1980 data unusable due to the coarser resolution of the input data from which it was derived. For years where USGS classifications were unavailable (1980 and 2011), I identified inhabited areas from National Association of Space and Aeronautics (NASA) Landsat TM imagery using multiple classification approaches in ERDAS IMAGINE 2010 (ERDAS, 2010). The earliest available Landsat TM scene satisfying my minimum cloud cover requirement of less than 10 percent was a Landsat 4 Multi-Spectral Scanner (MSS) image collected on April 4, 1984, necessitating adjustment to the initial calibration year. The most recent available data at the time of this analysis was a Landsat 5 TM image collected on January 3, 2011.

Many aspects of digital image processing can be improved by combining multiple data layers of the same region (Lillesand et al., 2008). Generating composites with a variety of derived layers is a common technique for better differentiating land cover types (Chen et al., 2002). Landsat scenes obtained to calculate urban extent and land use were
preprocessed to enhance classification accuracy. I assessed multiple classification methods, parameters, and inputs to determine the most appropriate for the model. After downloading the raw Landsat bands, I combined bands one through five and band seven (eliminating the thermal band) using the Layer Stack function in ERDAS IMAGINE 2010. The thermal band is not included in either Landsat composite because it has a coarser spatial resolution (120 m) than the other Landsat 5 TM bands and is not present in Landsat 4 MSS data (NASA, 2005). A subset of the composite was developed to limit the image to the study area. Derived layers were then calculated from the resulting Landsat composite images for each date. These include a Principal Components Analysis (PCA), Normalized Difference Vegetation Index (NDVI), and Tasseled Cap Transformation (TC). These derivations are used to enhance the image’s spectral variability and reduce interband correlation, improving the classification results (Lillesand, et al., 2008). Two composites were created for each needed urban extent year (1984, 2011), again using the Layer Stack function. The first composite included NDVI, Principal Components (PCs) 1-3, and the TC 1 band. Composite two consisted of NDVI, PC 1, and the six TM bands.

Using these composites, I generated an unsupervised classification with 20 clusters and a convergence threshold of 0.990 in ERDAS IMAGINE 2010 (ERDAS 2010). An unsupervised classification identifies spectrally similar pixels and groups them into a pre-specified number of categories based on commonalities inherent in the image. This classification method does not utilize any field data, and is based only on the spectral characteristics of pixels (Lillesand et al., 2008). For comparison purposes, I also generated a supervised classification using the original composited Landsat data. The
supervised classification process involves identifying training areas representing each land cover type to be classified. I identified 20 representative areas for each land cover type. The supervised classification identifies spectrally similar zones and assigns them to the most likely category based on a parametric rule. I elected to use the Maximum Likelihood classifier, a preferred method for normally distributed data (Repaka et al., 2004). The data fit well into my six-class system. As previously mentioned, the original USGS NLCD images use a 15 class system. To create the required consistency between SLEUTH land cover inputs, I reclassified the USGS LULC images to the same six class system as the Landsat scenes. After classification the land cover data for each year are limited to eight-bit radiometric resolution and converted to a GIF image.

Results from both classification methods and composites were compared with an accuracy assessment in ERDAS IMAGINE 2010. The accuracy assessment is performed using the 2010 National Agricultural Imagery Program (NAIP) orthoimagery as a reference layer. The NAIP has a one meter resolution, compared with Landsat’s 30m pixels, making it an appropriate accuracy-checking tool (Guindon et al., 2009). I chose a stratified random sampling method including 50 points from each land cover type, an appropriate sample size for land cover data with less than 12 categories (Lillesand et al., 2008). These points were used to perform an accuracy assessment of my classifications in ERDAS IMAGINE 2010. The accuracy assessment compares the land cover type at each reference point to the type in the classified layer. I chose the highest accuracy classifications for each year in model calibration. In both cases, this was the supervised classification.
Transportation layers in this study were limited to simple road or non-road classification. While some studies include shopping malls in their road network (Yang and Lo, 2003), this is not a significant factor in an area such as Humboldt county, where shopping malls are rare. Road classification data were obtained from US Census TIGER/line files. Due to limited availability of historical transportation network data, only two years (2000 and 2009) were used in this study, satisfying the minimum SLEUTH requirements. Once converted to raster format, they were exported to eight-bit GIF images.

The hillshade layer, only used for visualization purposes, was derived from the same USGS DEM as the slope layer. It can aid in the interpretation of results by providing spatial and topographic context to the urban extent data (Clarke, 2004). The water pixels on the hillshade were reclassified to a value of zero after zero elevation land pixels are reclassified to one. In the SLEUTH specifications, a particular color can be assigned to these zero value water pixels to improve the appearance of the visualization. This was especially important for my study, where Humboldt Bay is a predominant feature in all image scenes.

In this case, three unique scenarios for possible growth were analyzed based on successes of researchers including Yang and Lo (2003), Jantz et al. (2003) and Liu and Phinn (2004). The first scenario is an extension of the growth coefficients established in the calibration phase. In other words, this is what would happen if growth trends were to continue following historic rates. The second scenario maintains the same input data but modifies the growth coefficients to discourage organic growth, the natural tendency of a
populated area to expand, and root cause of urban sprawl. The third and final scenario is the “smart growth” strategy. This adjusts the growth coefficients as in the previous scenario to promote other forms of expansion and discourage organic growth, but also institutes additional environmental constraints. This strategy has demonstrated effectiveness in reducing total urban expansion and promoting more sustainable growth trends (Yang and Lo, 2003). Environmental constraints can also be added through buffers around water bodies and protected “green space” areas such as forests and grasslands or critical habitat areas. In this third scenario, I specified additional 30 meter buffers around water bodies and forests and added an exclusionary provision to protected Williamson Act agricultural lands.

Calibration

The calibration process for this SLEUTH application is shown as a flow chart in Figure 3. The model calibration and prediction time runs are listed in Table 2. The model was run on a laptop computer with an Intel Core i7 Q720M 1.6 GHz quad-core processor, 8 gigabytes of memory, and a 64-bit Windows 7 Ultimate operating system.
Initial Set Up

Install Cygwin, a Linux emulator, to run SLEUTH model in a Microsoft Windows environment
Download, compile, and verify model functions running demo data in test mode
Compare demo test results with sample results to ensure proper functioning

Humboldt Bay SLEUTH Calibration (Diff: Diffusion, Brd: Breed, Sprd: Spread, Slp: Slope, RG: Road Gravity)

Coarse Calibration
Edit scenario file for coarse calibration: run 5 brute force Monte Carlo iterations, testing coefficients from 0-100 with step of 25
Examine Monte Carlo runs, sort based on Lee Sallee metric and select ranges for fine calibration from top ten Lee Sallee values
Edit scenario file for fine calibration. Coefficient values: Diff 0-10, step 2; Brd 0-75, step 15; Sprd 20-30, step 2; Slp 0-25, step 5; RG 0-100, step 20

Fine Calibration
Run fine calibration with 8 Monte Carlo runs
Evaluate control stats log file, again sorting by highest Lee Sallee metric
Edit scenario file for final calibration. Coefficient values: Diff 0-2, 1; Brd 15-60, 9; Sprd 28-32, 1; Slp 1-5, 1; RG 60-100, 10

Figure 3. SLEUTH model workflow including all processes, coefficients, and steps needed to complete each growth scenario.
Run final calibration with 10 Monte Carlo iterations

Examine control stats log file, sorting on Lee Sallee metric, take top score. Three scores tied, take lowest value of coefficients

Edit scenario file for forecast calibration with one value. Coefficients: Diff 1, Brd 24, Sprd 31, Slp 1, RG 80

Run forecast calibration with 100 Monte Carlo iterations

Examine logs for coefficient values taking last year and rounding values to nearest integer, then initialize a prediction run of SLEUTH

Final best fit value. Coefficient values: Diff 1, Brd 24, Sprd 32, Slp 1, RG 84

Run forecast calibration with 100 Monte Carlo iterations

Examine logs for coefficient values taking last year and rounding values to nearest integer, then initialize a prediction run of SLEUTH

Final best fit value. Coefficient values: Diff 1, Brd 24, Sprd 32, Slp 1, RG 84

Figure 3. SLEUTH model workflow including all processes, coefficients, and steps needed to complete each growth scenario (continued).
Table 2. Computer processing times for each model process and sub-process in days, hours, minutes, and seconds.

<table>
<thead>
<tr>
<th>Operation</th>
<th>Elapsed Time (dd:hh:mm:ss)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test</td>
<td>00:00:00:28</td>
</tr>
<tr>
<td>Calibration Steps</td>
<td></td>
</tr>
<tr>
<td>Coarse</td>
<td>00:02:34:29</td>
</tr>
<tr>
<td>Fine 1</td>
<td>00:04:11:25</td>
</tr>
<tr>
<td>Fine 2</td>
<td>01:08:02:27</td>
</tr>
<tr>
<td>Final</td>
<td>04:05:00:30</td>
</tr>
<tr>
<td>Forecast</td>
<td>00:00:18:19</td>
</tr>
<tr>
<td>Predict</td>
<td></td>
</tr>
<tr>
<td>Scenario 1</td>
<td>00:02:46:33</td>
</tr>
<tr>
<td>Scenario 2</td>
<td>00:06:01:49</td>
</tr>
<tr>
<td>Scenario 3</td>
<td>00:04:19:29</td>
</tr>
</tbody>
</table>
RESULTS

My accuracy assessment results from Landsat 4 MSS classifications are shown in Table 3 and the Landsat 5 TM results in Table 4. Initial urbanized extent for 2011 is illustrated in Figure 4. Urbanized predictions are then shown in Figures 5, 6, and 7. Logging results from each SLEUTH run, including elapsed time and coefficient values, edited for clarity, are shown in the Appendices. The Scenario 1 log is found in Appendix A, Scenario 2 in Appendix B, and Scenario 3 in Appendix C. Selected statistics from each scenario are given in Tables 5, 6, and 7. The statistical tables include detailed data on how the coefficient values change as the model progresses, the total number of urbanized pixels, and comprehensive characteristics of predicted expansion areas, including mean size of clusters and number of edges.
Table 3. Accuracy assessment showing producer's accuracy for classifications of the 1984 Landsat image, including the unsupervised classifications of TM and no TM composites, and the supervised classification of the raw image.

<table>
<thead>
<tr>
<th>Accuracy Assessment TM 4 Unsupervised</th>
<th>Classes</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Water</td>
<td>92%</td>
<td></td>
</tr>
<tr>
<td>Urban</td>
<td>53%</td>
<td></td>
</tr>
<tr>
<td>Barren</td>
<td>48%</td>
<td></td>
</tr>
<tr>
<td>Forest</td>
<td>79%</td>
<td></td>
</tr>
<tr>
<td>Grassland</td>
<td>52%</td>
<td></td>
</tr>
<tr>
<td>Agriculture</td>
<td>70%</td>
<td></td>
</tr>
<tr>
<td>Overall Accuracy</td>
<td></td>
<td>66%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Accuracy Assessment No TM 4 Unsupervised</th>
<th>Classes</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Water</td>
<td>95%</td>
<td></td>
</tr>
<tr>
<td>Urban</td>
<td>39%</td>
<td></td>
</tr>
<tr>
<td>Barren</td>
<td>43%</td>
<td></td>
</tr>
<tr>
<td>Forest</td>
<td>76%</td>
<td></td>
</tr>
<tr>
<td>Grassland</td>
<td>67%</td>
<td></td>
</tr>
<tr>
<td>Agriculture</td>
<td>64%</td>
<td></td>
</tr>
<tr>
<td>Overall Accuracy</td>
<td></td>
<td>64%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Accuracy Assessment TM 4 Supervised</th>
<th>Classes</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Water</td>
<td>97%</td>
<td></td>
</tr>
<tr>
<td>Urban</td>
<td>67%</td>
<td></td>
</tr>
<tr>
<td>Barren</td>
<td>51%</td>
<td></td>
</tr>
<tr>
<td>Forest</td>
<td>69%</td>
<td></td>
</tr>
<tr>
<td>Grassland</td>
<td>61%</td>
<td></td>
</tr>
<tr>
<td>Agriculture</td>
<td>64%</td>
<td></td>
</tr>
<tr>
<td>Overall Accuracy</td>
<td></td>
<td>68%</td>
</tr>
</tbody>
</table>
Table 4. Accuracy assessment showing producer's accuracy for classifications of the 2011 Landsat image, including the unsupervised classifications of TM and no TM composites, and the supervised classification of the raw image.

<table>
<thead>
<tr>
<th>Accuracy Assessment TM 5 Unsupervised</th>
<th>Classes</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Water</td>
<td>95%</td>
<td></td>
</tr>
<tr>
<td>Urban</td>
<td>57%</td>
<td></td>
</tr>
<tr>
<td>Barren</td>
<td>54%</td>
<td></td>
</tr>
<tr>
<td>Forest</td>
<td>83%</td>
<td></td>
</tr>
<tr>
<td>Grassland</td>
<td>66%</td>
<td></td>
</tr>
<tr>
<td>Agriculture</td>
<td>65%</td>
<td></td>
</tr>
<tr>
<td>Overall Accuracy</td>
<td>70%</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Accuracy Assessment No TM 5 Unsupervised</th>
<th>Classes</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Water</td>
<td>96%</td>
<td></td>
</tr>
<tr>
<td>Urban</td>
<td>45%</td>
<td></td>
</tr>
<tr>
<td>Barren</td>
<td>47%</td>
<td></td>
</tr>
<tr>
<td>Forest</td>
<td>78%</td>
<td></td>
</tr>
<tr>
<td>Grassland</td>
<td>56%</td>
<td></td>
</tr>
<tr>
<td>Agriculture</td>
<td>70%</td>
<td></td>
</tr>
<tr>
<td>Overall Accuracy</td>
<td>65%</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Accuracy Assessment TM 5 Supervised</th>
<th>Classes</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Water</td>
<td>98%</td>
<td></td>
</tr>
<tr>
<td>Urban</td>
<td>61%</td>
<td></td>
</tr>
<tr>
<td>Barren</td>
<td>55%</td>
<td></td>
</tr>
<tr>
<td>Forest</td>
<td>79%</td>
<td></td>
</tr>
<tr>
<td>Grassland</td>
<td>67%</td>
<td></td>
</tr>
<tr>
<td>Agriculture</td>
<td>69%</td>
<td></td>
</tr>
<tr>
<td>Overall Accuracy</td>
<td>72%</td>
<td></td>
</tr>
</tbody>
</table>
Figure 4. Initial urban conditions in the Humboldt Bay Region, 2011. Urbanized areas are shown in pale yellow.
Figure 5. Scenario 1 urbanized predictions, 2110. Original 2011 urban extent is shown in pale yellow. Areas in red have 90-100 percent probability of urbanization, non-ocean blue areas are 30-40 percent and green are 1-10 percent.
Figure 6. Scenario 2 urbanized predictions, 2110. Original 2011 urban extent is shown in pale yellow. Areas in red have 90-100 percent probability of urbanization, non-ocean blue areas are 30-40 percent and green are 1-10 percent.
Figure 7. Scenario 3 urbanized predictions, 2110. Original 2011 urban extent is shown in pale yellow. Areas in red have 90-100 percent probability of urbanization, non-ocean blue areas are 30-40 percent and green are 1-10 percent.
Table 5. Selected statistics from Scenario 1 process including changes in coefficient values, total number of urbanized pixels, and characteristics of predicted expansion areas, including mean size of clusters and number of edges.

<table>
<thead>
<tr>
<th>Year</th>
<th>Area (sq m)</th>
<th>Edges</th>
<th>Clusters</th>
<th>Size (sq m)</th>
<th>Diff</th>
<th>Spri</th>
<th>Brd</th>
<th>Slp</th>
<th>RG</th>
<th>%Urban</th>
<th>Growth rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>2011</td>
<td>206447.57</td>
<td>90974.45</td>
<td>7869.52</td>
<td>25.25</td>
<td>1.00</td>
<td>41.00</td>
<td>32.00</td>
<td>1.00</td>
<td>84.00</td>
<td>21.06</td>
<td>4.82</td>
</tr>
<tr>
<td>2020</td>
<td>298715.15</td>
<td>101594.28</td>
<td>6024.58</td>
<td>48.75</td>
<td>1.09</td>
<td>44.84</td>
<td>35.00</td>
<td>1.00</td>
<td>86.14</td>
<td>27.07</td>
<td>3.23</td>
</tr>
<tr>
<td>2030</td>
<td>382329.08</td>
<td>94157.66</td>
<td>4665.84</td>
<td>80.91</td>
<td>1.21</td>
<td>49.53</td>
<td>38.66</td>
<td>1.00</td>
<td>89.11</td>
<td>32.53</td>
<td>1.88</td>
</tr>
<tr>
<td>2040</td>
<td>441199.61</td>
<td>87843.15</td>
<td>3833.96</td>
<td>113.92</td>
<td>1.30</td>
<td>53.43</td>
<td>41.71</td>
<td>1.00</td>
<td>91.69</td>
<td>36.37</td>
<td>1.08</td>
</tr>
<tr>
<td>2050</td>
<td>451680.56</td>
<td>86778.66</td>
<td>3699.00</td>
<td>120.93</td>
<td>0.24</td>
<td>0.85</td>
<td>0.85</td>
<td>29.99</td>
<td>88.79</td>
<td>37.05</td>
<td>0.02</td>
</tr>
<tr>
<td>2060</td>
<td>452365.53</td>
<td>86696.12</td>
<td>3690.48</td>
<td>121.45</td>
<td>0.24</td>
<td>0.85</td>
<td>0.85</td>
<td>67.06</td>
<td>85.09</td>
<td>37.10</td>
<td>0.01</td>
</tr>
<tr>
<td>2070</td>
<td>452988.86</td>
<td>86608.39</td>
<td>3682.63</td>
<td>121.85</td>
<td>0.24</td>
<td>0.85</td>
<td>0.85</td>
<td>100.00</td>
<td>81.37</td>
<td>37.14</td>
<td>0.01</td>
</tr>
<tr>
<td>2080</td>
<td>453601.65</td>
<td>86522.95</td>
<td>3674.80</td>
<td>122.27</td>
<td>0.24</td>
<td>0.85</td>
<td>0.85</td>
<td>100.00</td>
<td>77.66</td>
<td>37.18</td>
<td>0.01</td>
</tr>
<tr>
<td>2090</td>
<td>454201.85</td>
<td>86434.99</td>
<td>3668.92</td>
<td>122.62</td>
<td>0.24</td>
<td>0.85</td>
<td>0.85</td>
<td>100.00</td>
<td>73.94</td>
<td>37.22</td>
<td>0.01</td>
</tr>
<tr>
<td>2100</td>
<td>454801.78</td>
<td>86349.16</td>
<td>3661.98</td>
<td>123.05</td>
<td>0.24</td>
<td>0.85</td>
<td>0.85</td>
<td>100.00</td>
<td>70.22</td>
<td>37.26</td>
<td>0.01</td>
</tr>
<tr>
<td>2110</td>
<td>455392.20</td>
<td>86256.90</td>
<td>3655.56</td>
<td>123.39</td>
<td>0.24</td>
<td>0.85</td>
<td>0.85</td>
<td>100.00</td>
<td>66.49</td>
<td>37.29</td>
<td>0.01</td>
</tr>
</tbody>
</table>
Table 6. Selected statistics from Scenario 2 process including changes in coefficient values, total number of urbanized pixels, and characteristics of predicted expansion areas, including mean size of clusters and number of edges.

<table>
<thead>
<tr>
<th>Year</th>
<th>Area (sq m)</th>
<th>Edges (sq m)</th>
<th>Clusters</th>
<th>Size (sq m)</th>
<th>Diff</th>
<th>Spd</th>
<th>Brd</th>
<th>Slp</th>
<th>RG</th>
<th>%Urban</th>
<th>Growth rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>2011</td>
<td>200400.19</td>
<td>88486.93</td>
<td>8043.80</td>
<td>24.00</td>
<td>25.00</td>
<td>15.00</td>
<td>50.00</td>
<td>1.00</td>
<td>100.00</td>
<td>20.66</td>
<td>1.94</td>
</tr>
<tr>
<td>2020</td>
<td>259295.20</td>
<td>101137.08</td>
<td>7747.25</td>
<td>30.05</td>
<td>27.34</td>
<td>16.41</td>
<td>54.68</td>
<td>1.00</td>
<td>100.00</td>
<td>23.20</td>
<td>1.88</td>
</tr>
<tr>
<td>2030</td>
<td>285264.70</td>
<td>105594.54</td>
<td>7248.38</td>
<td>38.56</td>
<td>30.20</td>
<td>18.12</td>
<td>60.41</td>
<td>1.00</td>
<td>100.00</td>
<td>26.20</td>
<td>1.66</td>
</tr>
<tr>
<td>2040</td>
<td>331314.59</td>
<td>104971.16</td>
<td>6792.63</td>
<td>47.94</td>
<td>33.36</td>
<td>20.02</td>
<td>66.73</td>
<td>1.00</td>
<td>100.00</td>
<td>29.20</td>
<td>1.40</td>
</tr>
<tr>
<td>2050</td>
<td>374313.79</td>
<td>103464.48</td>
<td>6366.78</td>
<td>57.92</td>
<td>34.47</td>
<td>20.68</td>
<td>68.93</td>
<td>1.00</td>
<td>100.00</td>
<td>32.01</td>
<td>1.08</td>
</tr>
<tr>
<td>2060</td>
<td>393630.86</td>
<td>102673.51</td>
<td>6163.97</td>
<td>62.97</td>
<td>0.46</td>
<td>0.46</td>
<td>18.57</td>
<td>98.24</td>
<td>33.27</td>
<td>0.04</td>
<td></td>
</tr>
<tr>
<td>2070</td>
<td>395166.16</td>
<td>102550.09</td>
<td>6136.36</td>
<td>63.49</td>
<td>0.46</td>
<td>0.46</td>
<td>0.46</td>
<td>51.88</td>
<td>94.91</td>
<td>33.37</td>
<td>0.04</td>
</tr>
<tr>
<td>2080</td>
<td>396559.59</td>
<td>102413.49</td>
<td>6106.76</td>
<td>64.07</td>
<td>0.46</td>
<td>0.46</td>
<td>0.46</td>
<td>85.29</td>
<td>91.57</td>
<td>33.46</td>
<td>0.03</td>
</tr>
<tr>
<td>2090</td>
<td>397871.22</td>
<td>102263.57</td>
<td>6078.33</td>
<td>64.55</td>
<td>0.46</td>
<td>0.46</td>
<td>0.46</td>
<td>100.00</td>
<td>88.22</td>
<td>33.54</td>
<td>0.03</td>
</tr>
<tr>
<td>2100</td>
<td>399174.16</td>
<td>102118.27</td>
<td>6052.40</td>
<td>65.00</td>
<td>0.46</td>
<td>0.46</td>
<td>0.46</td>
<td>100.00</td>
<td>84.86</td>
<td>33.63</td>
<td>0.03</td>
</tr>
<tr>
<td>2110</td>
<td>400463.59</td>
<td>101973.38</td>
<td>6027.87</td>
<td>65.54</td>
<td>0.46</td>
<td>0.46</td>
<td>0.46</td>
<td>100.00</td>
<td>81.50</td>
<td>33.71</td>
<td>0.03</td>
</tr>
</tbody>
</table>
Table 7. Selected statistics from Scenario 3 process including changes in coefficient values, total number of urbanized pixels, and characteristics of predicted expansion areas, including mean size of clusters and number of edges.

<table>
<thead>
<tr>
<th>Year</th>
<th>Area (sq m)</th>
<th>Edges</th>
<th>Clusters</th>
<th>Size (sq m)</th>
<th>Diff</th>
<th>Sprd</th>
<th>Bnd</th>
<th>Slp</th>
<th>RG</th>
<th>%Urban</th>
<th>Growth rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>2011</td>
<td>200142.16</td>
<td>88393.25</td>
<td>8061.19</td>
<td>24</td>
<td>25</td>
<td>15</td>
<td>50</td>
<td>1</td>
<td>100</td>
<td>21.32</td>
<td>1.82</td>
</tr>
<tr>
<td>2020</td>
<td>236335.28</td>
<td>100463.94</td>
<td>7873.02</td>
<td>29.09</td>
<td>27.34</td>
<td>16.41</td>
<td>54.68</td>
<td>1</td>
<td>100</td>
<td>23.75</td>
<td>1.77</td>
</tr>
<tr>
<td>2030</td>
<td>278953.57</td>
<td>105168.84</td>
<td>7426.08</td>
<td>36.77</td>
<td>30.2</td>
<td>18.12</td>
<td>60.41</td>
<td>1</td>
<td>100</td>
<td>26.62</td>
<td>1.57</td>
</tr>
<tr>
<td>2040</td>
<td>321688.12</td>
<td>105349.03</td>
<td>7016.11</td>
<td>45.02</td>
<td>33.27</td>
<td>19.96</td>
<td>66.55</td>
<td>1</td>
<td>100</td>
<td>29.5</td>
<td>1.33</td>
</tr>
<tr>
<td>2050</td>
<td>361319.57</td>
<td>104546.91</td>
<td>6642.06</td>
<td>53.52</td>
<td>33.71</td>
<td>20.22</td>
<td>67.41</td>
<td>1</td>
<td>100</td>
<td>32.17</td>
<td>1.03</td>
</tr>
<tr>
<td>2060</td>
<td>373811.69</td>
<td>104263.44</td>
<td>6516.82</td>
<td>56.45</td>
<td>0.63</td>
<td>0.63</td>
<td>0.63</td>
<td>23.76</td>
<td>97.72</td>
<td>33.01</td>
<td>0.04</td>
</tr>
<tr>
<td>2070</td>
<td>375223.1</td>
<td>104162.77</td>
<td>6489.38</td>
<td>56.94</td>
<td>0.63</td>
<td>0.63</td>
<td>0.63</td>
<td>56.81</td>
<td>94.42</td>
<td>33.11</td>
<td>0.04</td>
</tr>
<tr>
<td>2080</td>
<td>376495.13</td>
<td>104050.2</td>
<td>6462.68</td>
<td>57.4</td>
<td>0.63</td>
<td>0.63</td>
<td>0.63</td>
<td>89.96</td>
<td>91.1</td>
<td>33.19</td>
<td>0.03</td>
</tr>
<tr>
<td>2090</td>
<td>377703.14</td>
<td>103924.48</td>
<td>6436.44</td>
<td>57.85</td>
<td>0.63</td>
<td>0.63</td>
<td>0.63</td>
<td>100</td>
<td>87.78</td>
<td>33.28</td>
<td>0.03</td>
</tr>
<tr>
<td>2100</td>
<td>378891.92</td>
<td>103799.35</td>
<td>6411.97</td>
<td>58.2</td>
<td>0.63</td>
<td>0.63</td>
<td>0.63</td>
<td>100</td>
<td>84.45</td>
<td>33.36</td>
<td>0.03</td>
</tr>
<tr>
<td>2110</td>
<td>380072.11</td>
<td>103675.23</td>
<td>6386.82</td>
<td>58.63</td>
<td>0.63</td>
<td>0.63</td>
<td>0.63</td>
<td>100</td>
<td>81.11</td>
<td>33.44</td>
<td>0.03</td>
</tr>
</tbody>
</table>
DISCUSSION AND CONCLUSIONS

My SLEUTH modeling outputs reinforce established predictions from results identified in research by Chen et al., 2002; Jantz et al., 2003; Li and Yeh, 2002; Liu and Phinn, 2004, and several others. Scenario 1 had the highest predicted urbanization at 37.29 percent of total area, and while scenario 2 (33.71 percent) and 3 (33.44 percent) were similar, scenario 3 demonstrated a lower probability of urban expansion into protected agricultural lands as visible when comparing Figures 5 and 6. As expected, the location of roads has an impact on where urbanization spreads. The marked increase in expansion onto the lower slope lands near Ferndale, and lack of spread in steeper regions, demonstrates the effect of the slope resistance coefficient. Development will tend to spread to these areas first when permitted, potentially putting many agricultural lands at risk.

The commonalities between scenarios 2 and 3 illustrate the importance of the growth coefficients, which these scenarios shared. The effect of adding environmental constraints did much less to reduce urban spread than modifying the coefficients, as the transition from scenario 1 to 2 showed a notable decrease in quantity and probability of urbanized pixels. These findings highlight the need for policy makers to implement smart growth strategies rather than simply rely on environmental or zoning restrictions. Ideally, these approaches could be implemented in tandem to best control urban sprawl and prevent loss of critical habitat and agricultural lands.
This SLEUTH application opens many potential areas for future research and exploration in the Humboldt Bay region. Alternate scenarios can be quickly designed and implemented with the calibration complete. Some possibilities include crafting scenarios to take into account cultural and political constraints or additional land use restrictions. Encouraging development in certain areas is even possible using a method described in Jantz et al., 2010. By giving all unavailable areas a base value of 50 instead of 100 in the exclusion layer, regions where development should be encouraged can be set to a value below 50. In addition, public input could be incorporated into the SLEUTH scenarios using a Public Participation GIS (PPGIS) approach. Some ideas for this include the use of open public forums or even a web-based application that would allow community members to design custom SLEUTH scenarios incorporating their own constraints and see how these affect the model’s outputs. Designing a graphical user interface for the SLEUTH model is another possibility for improving its usability and opening up its capabilities to a broader audience. However, due to the potentially controversial nature of urbanization predictions, care needs to be taken when implementing these participatory kinds of methods. Political and cultural issues can arise due to the content of SLEUTH outputs, and these possibilities should be considered and discussed in the design of a public participation SLEUTH model.

Modeling land cover change in addition to urban extent provides some other avenues for future research in the Humboldt Bay region. The SLEUTH deltatron model can generate both animated and cumulative land use image outputs. These can be used to
build additional GIS layers, such as change detection and land use probability maps.
These facilitate the study of urbanization’s impact on land use change and offer policy
makers another potentially valuable source of information for strategic urban growth
planning.
REFERENCES


 Humboldt County, 2005. Humboldt County California, URL: http://co.humboldt.ca.us/portal/about.asp (last date accessed: 4 March 2012).


Appendix A. Selected log data from Scenario 1 SLEUTH run, edited for clarity. This log shows all the necessary components to execute an identical SLEUTH run, including input details, coefficient values, random number seed, and number of Monte Carlo iterations.

DATE OF RUN: Wed May 23 21:01:52 2012

USER: J. Morrison
HOST: (null)
HOSTTYPE: (null)
OSTYPE: (null)
Type of architecture: 32 bit

Number of CPUs 1

PWD: /cygdrive/c/sleuth1/scenarios
Scenario File: scenario.humboldt_predict1
Type of Processing: PREDICTING

scenario.filename = scenario.humboldt_predict1
scenario.input_dir = ../Input/humboldt_full/
scenario.output_dir = ../Output/humboldt_predict1/
scenario.whirlgif_binary = ../Whirlgif/whirlgif
scenario.urban_data_file[0] = humboldt.urban.1980.gif
scenario.road_data_file[0] = humboldt.roads.2000.gif
scenario.excluded_data_file = humboldt.excluded.gif
scenario.slope_data_file = humboldt.slope.gif
scenario.background_data_file = humboldt.hillshade.gif
scenario.echo = 1
scenario.logging = 1
scenario.log_processing_status = 1
scenario.random_seed = 12
scenario.num_working_grids = 8
scenario.monte_carlo_iterations = 100
scenario.start.diffusion = 1
scenario.stop.diffusion = 1
scenario.step.diffusion = 1
scenario.best_fit.diffusion = 1
Appendix A. Selected log data from Scenario 1 SLEUTH run, edited for clarity, showing the necessary components to execute an identical SLEUTH run, including input details, coefficient values, random number seed, and number of Monte Carlo iterations (continued).

scenario.start.breed = 24
scenario.stop.breed = 24
scenario.step.breed = 1
scenario.best_fit.breed = 32
scenario.start.spread = 31
scenario.stop.spread = 31
scenario.step.spread = 1
scenario.best_fit.spread = 41
scenario.start.slope_resistance = 1
scenario.stop.slope_resistance = 1
scenario.step.slope_resistance = 1
scenario.best_fit.slope_resistance = 1
scenario.start.road_gravity = 80
scenario.stop.road_gravity = 80
scenario.step.road_gravity = 1
scenario.best_fit.road_gravity = 84
scenario.prediction_start_date = 2010
scenario.prediction_stop_date = 2110
scenario.date_color = ffffff
scenario.seed_color = f9d16e
scenario.water_color = 1434d6
scenario.probability_color[0].lower_bound = 0
scenario.probability_color[0].upper_bound = 1
scenario.probability_color[0].color = 0
scenario.rd_grav_sensitivity = 0.010000
scenario.slope_sensitivity = 0.100000
scenario.critical_low = 0.970000
scenario.critical_high = 1.300000
scenario.critical_slope = 15.000000
scenario.boom = 1.010000
scenario.bust = 0.090000
scenario.log_base_stats = 1
scenario.log_debug = 0
scenario.log_urbanization_attempts = 0
Appendix A. Selected log data from Scenario 1 SLEUTH run, edited for clarity, showing the necessary components to execute an identical SLEUTH run, including input details, coefficient values, random number seed, and number of Monte Carlo iterations (continued).

```plaintext
scenario.log_coeff = 0
scenario.log_timings = 1
scenario.write_avg_file = 1
scenario.write_std_dev_file = 0
scenario.log_memory_map = 1
scenario.log_landclass_summary = 1
scenario.log_slope_weights = 0
scenario.log_reads = 0
scenario.log_writes = 0
scenario.log_colortables = 0
scenario.log_processing_status = 1
scenario.log_trans_matrix = 1
scenario.view_growth_types = 0
scenario.growth_type_window.run1 = 0
scenario.growth_type_window.run2 = 0
scenario.growth_type_window.monte_carlo1 = 0
scenario.growth_type_window.monte_carlo2 = 0
scenario.growth_type_window.year1 = 1995
scenario.growth_type_window.year2 = 2020
scenario.phase0g_growth_color = ff0000
scenario.phase1g_growth_color = ff00
scenario.phase2g_growth_color = ff
scenario.phase3g_growth_color = ffff00
scenario.phase4g_growth_color = ffffff
scenario.phase5g_growth_color = ffff
scenario.view_deltatron_aging = 0
scenario.deltatron_aging_window.run1 = 0
scenario.deltatron_aging_window.run2 = 0
scenario.deltatron_aging_window.monte_carlo1 = 0
scenario.deltatron_aging_window.monte_carlo2 = 0
scenario.deltatron_aging_window.year1 = 1930
scenario.deltatron_aging_window.year2 = 2020
scenario.deltatron_color[0] = 0
scenario.deltatron_color[1] = 65280
```
Appendix A. Selected log data from Scenario 1 SLEUTH run, edited for clarity, showing the necessary components to execute an identical SLEUTH run, including input details, coefficient values, random number seed, and number of Monte Carlo iterations (continued).

scenario.deltatron_color[2] = 53760
scenario.deltatron_color[3] = 43520
scenario.deltatron_color[4] = 33280
scenario.deltatron_color[5] = 23040
coeff_obj.c 1237 start_coeff.diffusion = 1
coeff_obj.c 1238 start_coeff.spread = 31
coeff_obj.c 1239 start_coeff.breed = 24
coeff_obj.c 1240 start_coeff.slope_resistance = 1
coeff_obj.c 1241 start_coeff.road_gravity = 80
coeff_obj.c 1258 stop_coeff.diffusion = 1
coeff_obj.c 1259 stop_coeff.spread = 31
coeff_obj.c 1260 stop_coeff.breed = 24
coeff_obj.c 1261 stop_coeff.slope_resistance = 1
coeff_obj.c 1262 stop_coeff.road_gravity = 80
coeff_obj.c 1216 step_coeff.diffusion = 1
coeff_obj.c 1217 step_coeff.spread = 1
coeff_obj.c 1218 step_coeff.breed = 1
coeff_obj.c 1219 step_coeff.slope_resistance = 1
coeff_obj.c 1220 step_coeff.road_gravity = 1
coeff_obj.c 1279 best_fit_coeff.diffusion = 1
coeff_obj.c 1280 best_fit_coeff.spread = 41
coeff_obj.c 1281 best_fit_coeff.breed = 32
coeff_obj.c 1282 best_fit_coeff.slope_resistance = 1
coeff_obj.c 1283 best_fit_coeff.road_gravity = 84
memory_obj.c 716 Allocated 219653760 bytes of memory
Appendix B. Selected log data from Scenario 2 SLEUTH run, edited for clarity. This log shows all the necessary components to execute an identical SLEUTH run, including input details, coefficient values, random number seed, and number of Monte Carlo iterations.

DATE OF RUN: Sun Apr 10 19:42:51 2011

USER: J. Morrison
HOST: (null)
HOSTTYPE: (null)
OSTYPE: (null)
Type of architecture: 32 bit

Number of CPUs 1

PWD: /cygdrive/c/SLEUTH/Scenarios
Scenario File: scenario.humboldt_predict2
Type of Processing: PREDICTING

scenario.filename = scenario.humboldt_predict2
scenario.input_dir = ../Input/humboldt_full/
scenario.output_dir = ../Output/humboldt_predict2/
scenario.whirlgif_binary = ../Whirlgif/whirlgif
scenario.urban_data_file[0] = humboldt.urban.1980.gif
scenario.road_data_file[0] = humboldt.roads.2000.gif
scenario.excluded_data_file = humboldt.excluded.gif
scenario.slope_data_file = humboldt.slope.gif
scenario.background_data_file = humboldt.hillshade.gif
scenario.echo = 1
scenario.logging = 1
scenario.log_processing_status = 1
scenario.random_seed = 34
scenario.num_working_grids = 8
scenario.monte_carlo_iterations = 100
scenario.start.diffusion = 25
scenario.stop.diffusion = 25
scenario.step.diffusion = 1
Appendix B. Selected log data from Scenario 2 SLEUTH run, edited for clarity, showing the necessary components to execute an identical SLEUTH run, including input details, coefficient values, random number seed, and number of Monte Carlo iterations (continued).

```python
scenario.best_fit.diffusion = 25
scenario.start.breed = 50
scenario.stop.breed = 50
scenario.step.breed = 1
scenario.best_fit.breed = 50
scenario.start.spread = 15
scenario.stop.spread = 15
scenario.step.spread = 1
scenario.best_fit.spread = 15
scenario.start.slope_resistance = 1
scenario.stop.slope_resistance = 1
scenario.step.slope_resistance = 1
scenario.best_fit.slope_resistance = 1
scenario.start.road_gravity = 100
scenario.stop.road_gravity = 100
scenario.step.road_gravity = 1
scenario.best_fit.road_gravity = 100
scenario.prediction_start_date = 2010
scenario.prediction_stop_date = 2110
scenario.date_color = ffffff
scenario.seed_color = f9d16e
scenario.water_color = 1434d6
scenario.probability_color[0].lower_bound = 0
scenario.probability_color[0].upper_bound = 1
scenario.probability_color[0].color = 0
scenario.rd_grav_sensitivity = 0.010000
scenario.slope_sensitivity = 0.100000
scenario.critical_low = 0.970000
scenario.critical_high = 1.300000
scenario.critical_slope = 15.000000
scenario.boom = 1.010000
scenario.bust = 0.090000
scenario.log_base_stats = 1
scenario.log_debug = 0
```
Appendix B. Selected log data from Scenario 2 SLEUTH run, edited for clarity, showing the necessary components to execute an identical SLEUTH run, including input details, coefficient values, random number seed, and number of Monte Carlo iterations (continued).

```plaintext
scenario.log.urbanization.attempts = 0
scenario.log.coeff = 0
scenario.log.timings = 1
scenario.write.avg.file = 1
scenario.write.std.dev.file = 0
scenario.log.memory.map = 1
scenario.log.landclass.summary = 1
scenario.log.slope.weights = 0
scenario.log.reads = 0
scenario.log.writes = 0
scenario.log.colortables = 0
scenario.log.processing.status = 1
scenario.log.trans.matrix = 1
scenario.view.growth.types = 0
scenario.growth.type.window.run1 = 0
scenario.growth.type.window.run2 = 0
scenario.growth.type.window.monte.carlo1 = 0
scenario.growth.type.window.monte.carlo2 = 0
scenario.growth.type.window.year1 = 1995
scenario.growth.type.window.year2 = 2020
scenario.phase0g.growth.color = ff0000
scenario.phase1g.growth.color = ff00
scenario.phase2g.growth.color = ff
scenario.phase3g.growth.color = ffff00
scenario.phase4g.growth.color = ffffff
scenario.phase5g.growth.color = ffff
scenario.view.deltatron.aging = 0
scenario.deltatron.aging.window.run1 = 0
scenario.deltatron.aging.window.run2 = 0
scenario.deltatron.aging.window.monte.carlo1 = 0
scenario.deltatron.aging.window.monte.carlo2 = 0
scenario.deltatron.aging.window.year1 = 1930
scenario.deltatron.aging.window.year2 = 2020
scenario.deltatron.color[0] = 0
```
Appendix B. Selected log data from Scenario 2 SLEUTH run, edited for clarity, showing the necessary components to execute an identical SLEUTH run, including input details, coefficient values, random number seed, and number of Monte Carlo iterations (continued).

scenario.deltatron_color[1] = 65280
scenario.deltatron_color[2] = 53760
scenario.deltatron_color[3] = 43520
scenario.deltatron_color[4] = 33280
scenario.deltatron_color[5] = 23040
coeff_obj.c 1237 start_coeff.diffusion = 25
coeff_obj.c 1238 start_coeff.spread = 15
coeff_obj.c 1239 start_coeff.breed = 50
coeff_obj.c 1240 start_coeff.slope_resistance = 1
coeff_obj.c 1241 start_coeff.road_gravity = 100
coeff_obj.c 1258 stop_coeff.diffusion = 25
coeff_obj.c 1259 stop_coeff.spread = 15
coeff_obj.c 1260 stop_coeff.breed = 50
coeff_obj.c 1261 stop_coeff.slope_resistance = 1
coeff_obj.c 1262 stop_coeff.road_gravity = 100
coeff_obj.c 1216 step_coeff.diffusion = 1
coeff_obj.c 1217 step_coeff.spread = 1
coeff_obj.c 1218 step_coeff.breed = 1
coeff_obj.c 1219 step_coeff.slope_resistance = 1
coeff_obj.c 1220 step_coeff.road_gravity = 1
coeff_obj.c 1279 best_fit_coeff.diffusion = 25
coeff_obj.c 1280 best_fit_coeff.spread = 15
coeff_obj.c 1281 best_fit_coeff.breed = 50
coeff_obj.c 1282 best_fit_coeff.slope_resistance = 1
coeff_obj.c 1283 best_fit_coeff.road_gravity = 100
memory_obj.c 716 Allocated 219653760 bytes of memory
Appendix C. Selected log data from Scenario 3 SLEUTH run, edited for clarity. This log shows all the necessary components to execute an identical SLEUTH run, including input details, coefficient values, random number seed, and number of Monte Carlo iterations.

DATE OF RUN: Sat May 5 20:58:01 2012

USER: J. Morrison
HOST: (null)
HOSTTYPE: (null)
OSTYPE: (null)
Type of architecture: 32 bit

Number of CPUs 1

PWD: /cygdrive/c/SLEUTH1/scenarios
Scenario File: scenario.humboldt_predict3
Type of Processing: PREDICTING

scenario.filename = scenario.humboldt_predict3
scenario.input_dir = ../Input/humboldt_full/
scenario.output_dir = ../Output/humboldt_predict3/
scenario.whirlgif_binary = ../Whirlgif/whirlgif
scenario.urban_data_file[0] = humboldt.urban.1980.gif
scenario.road_data_file[0] = humboldt.roads.2000.gif
scenario.excluded_data_file = humboldt.excluded1.gif
scenario.slope_data_file = humboldt.slope1.gif
scenario.background_data_file = humboldt.hillshade.gif
scenario.echo = 1
scenario.logging = 1
scenario.log_processing_status = 1
scenario.random_seed = 21
scenario.num_working_grids = 8
scenario.monte_carlo_iterations = 100
scenario.start.diffusion = 25
scenario.stop.diffusion = 25
scenario.step.diffusion = 1
Appendix C. Selected log data from Scenario 3 SLEUTH run, edited for clarity. This log shows all the necessary components to execute an identical SLEUTH run, including input details, coefficient values, random number seed, and number of Monte Carlo iterations (continued).

```plaintext
scenario.best_fit.diffusion = 25
scenario.start.breed = 50
scenario.stop.breed = 50
scenario.step.breed = 1
scenario.best_fit.breed = 50
scenario.start.spread = 15
scenario.stop.spread = 15
scenario.step.spread = 1
scenario.best_fit.spread = 15
scenario.start.slope_resistance = 1
scenario.stop.slope_resistance = 1
scenario.step.slope_resistance = 1
scenario.best_fit.slope_resistance = 1
scenario.start.road_gravity = 100
scenario.stop.road_gravity = 100
scenario.step.road_gravity = 1
scenario.best_fit.road_gravity = 100
scenario.prediction_start_date = 2010
scenario.prediction_stop_date = 2110
scenario.date_color = ffffff
scenario.seed_color = f9d16e
scenario.water_color = 1434d6
scenario.probability_color[0].lower_bound = 0
scenario.probability_color[0].upper_bound = 1
scenario.probability_color[0].color = 0
scenario.rd_grav_sensitivity = 0.010000
scenario.slope_sensitivity = 0.100000
scenario.critical_low = 0.970000
scenario.critical_high = 1.300000
scenario.critical_slope = 15.000000
scenario.boom = 1.010000
scenario.bust = 0.090000
scenario.log_base_stats = 1
scenario.log_debug = 0
```
Appendix C. Selected log data from Scenario 3 SLEUTH run, edited for clarity. This log shows all the necessary components to execute an identical SLEUTH run, including input details, coefficient values, random number seed, and number of Monte Carlo iterations (continued).

```plaintext
scenario.log_urbanization_attempts = 0
scenario.log_coeff = 0
scenario.log_timings = 1
scenario.write_avg_file = 1
scenario.write_std_dev_file = 0
scenario.log_memory_map = 1
scenario.log_landclass_summary = 1
scenario.log_slope_weights = 0
scenario.log_reads = 0
scenario.log_writes = 0
scenario.log_colortables = 0
scenario.log_processing_status = 1
scenario.log_trans_matrix = 1
scenario.view_growth_types = 0
scenario.growth_type_window.run1 = 0
scenario.growth_type_window.run2 = 0
scenario.growth_type_window.monte_carlo1 = 0
scenario.growth_type_window.monte_carlo2 = 0
scenario.growth_type_window.year1 = 1995
scenario.growth_type_window.year2 = 2020
scenario.phase0g_growth_color = ff0000
scenario.phase1g_growth_color = ff00
scenario.phase2g_growth_color = ff
scenario.phase3g_growth_color = ffff00
scenario.phase4g_growth_color = fffff
scenario.phase5g_growth_color = ffff
scenario.view_deltatron_aging = 0
scenario.deltatron_aging_window.run1 = 0
scenario.deltatron_aging_window.run2 = 0
scenario.deltatron_aging_window.monte_carlo1 = 0
scenario.deltatron_aging_window.monte_carlo2 = 0
scenario.deltatron_aging_window.year1 = 1930
scenario.deltatron_aging_window.year2 = 2020
scenario.deltatron_color[0] = 0
```
Appendix C. Selected log data from Scenario 3 SLEUTH run, edited for clarity. This log shows all the necessary components to execute an identical SLEUTH run, including input details, coefficient values, random number seed, and number of Monte Carlo iterations (continued).

scenario.deltatron_color[1] = 65280
scenario.deltatron_color[2] = 53760
scenario.deltatron_color[3] = 43520
scenario.deltatron_color[4] = 33280
scenario.deltatron_color[5] = 23040
coeff_obj.c 1237 start_coeff.diffusion = 25
coeff_obj.c 1238 start_coeff.spread = 15
coeff_obj.c 1239 start_coeff.breed = 50
coeff_obj.c 1240 start_coeff.slope_resistance = 1
coeff_obj.c 1241 start_coeff.road_gravity = 100
coeff_obj.c 1258 stop_coeff.diffusion = 25
coeff_obj.c 1259 stop_coeff.spread = 15
coeff_obj.c 1260 stop_coeff.breed = 50
coeff_obj.c 1261 stop_coeff.slope_resistance = 1
coeff_obj.c 1262 stop_coeff.road_gravity = 100
coeff_obj.c 1216 step_coeff.diffusion = 1
coeff_obj.c 1217 step_coeff.spread = 1
coeff_obj.c 1218 step_coeff.breed = 1
coeff_obj.c 1219 step_coeff.slope_resistance = 1
coeff_obj.c 1220 step_coeff.road_gravity = 1
coeff_obj.c 1279 best_fit_coeff.diffusion = 25
coeff_obj.c 1280 best_fit_coeff.spread = 15
coeff_obj.c 1281 best_fit_coeff.breed = 50
coeff_obj.c 1282 best_fit_coeff.slope_resistance = 1
coeff_obj.c 1283 best_fit_coeff.road_gravity = 1