Using the SLEUTH Urban Growth Model to Identify "Drivers" of Land Use Change in the Baltimore Metropolitan Region

Michael Maret

Geoenvironmental research paper

1. Introduction

The objectives of this research are to test how different "human drivers of development" can be incorporated into the SLEUTH urban model and whether they can be utilized to improve the models predictive capabilities. SLEUTH is a cellular automata model that applies transition rules to the states of a gridded series of cells, and in this case the transition is that from undeveloped to developed land cover, otherwise known as urbanization (Clarke, 1995). The model was chosen due to its successful implementation in numerous land use change studies of similar nature and scale (Clarke, et al., 1997; Jantz, et al., 2009; Silva and Clarke, 2002; Oguz, 2007; Onsted and Clarke, 2011; Xian, 2005). Implementation of the study focuses on the changes in land cover experienced within the Baltimore Metropolitan Region (BMR).

To understand this study one must first be provided with a basic overview of SLEUTH. First utilized to model urban land cover in the San Francisco Bay area (Clarke, et al., 1997), SLEUTH is an acronym for the spatial datasets required as inputs from the user to run the model. These datasets are, in order, Slope gradient, Land use, Exclusion, Urban extent, Transportation, and Hillshade. In order to utilize these datasets, they must be rasterized within a GIS to be compatible with the grid format required. Additionally, the urban extent must be a reclassification of land cover into only two classes, urban, and nonurban. A benefit from having data in this format is that SLEUTH and other CA models have results that can be visualized, as well as quantified (Oguz, 2007), both of which benefited this analysis.

A multi-step process is required to accurately prepare the model for predicting geospatial distributions of urban and nonurban lands throughout the study area. The first, and widely acknowledged step, is calibration. Calibration allows SLEUTH to analyze the historic time series of land cover and determine which of the programs growth rules are being applied over time, and the value of their influence. The output data from SLEUTH at this point is a series of goodness of fit statistics used to evaluate how well the model is interpreting historic growth. The fit statistics used in this research include fractional differences between area, edge, and cluster growth between mapped and modeled urbanization. This process can be time consuming due to the labor and computational time needed by SLEUTH. In this research we add to the calibration process by also validating our model with an independent time series, allowed for by a string of four time steps of urban coverage data. Validation allows us to test the predictive capabilities of SLEUTH against a known time step which permits some evaluation of how the model is performing before 'blindly' predicting growth rates into the future.

The last step in modeling is the predictive state. Utilizing the historic trends and values from growth rules a surface representing the future distributions of urban and non-urban coverage can be generated across the study area. It is important to note that the user can influence this process to create different scenarios or outcomes as the future is unknown and trends area anything but linear. In past studies multiple scenarios have been utilized to evaluate land use policy decisions, provide a range of possible future growth across a region, and more recently, to even evaluate portions of the SLEUTH model, like the differential assessment of lands.

As touched upon earlier, during SLEUTH processes the model applies transition rules to a grid of cells, determined to be urban or nonurban, and simulates their change over time using historic trends and input from the user. Factors influencing this transition or change include a series of growth rules, and the 'Exclusion' layer. Both the growth rules and the exclusion layer will be covered in depth later in the research, but it is important to point out another addition. While testing drivers of human development this study utilized three different datasets to create what is considered an exclusion/attraction layer. While in the past the excluded layer has only been used to repel growth from certain areas, a modification of this layer allows for both the repellant and attraction of growth to different areas. The hope is that by testing different drivers the performance of the model will be improved by incorporating factors believed to be influencing development within the BMR today.

Testing of drivers will be conducted upon a base excluded layer, which provides minimal information on protected lands; the sewer/water infrastructure, assumed to be a determinant of urban growth; and population/employment changes represented at a broad scale in Regional Planning Districts (RPDs). Recent studies have shown that the base excluded layer does indeed improve the model (Onsted and Clarke, In Press), but this research plans to find even better datasets for model improvement and advancement. By using a more intensive calibration, validation, and prediction approach in addition to the exclusion/attraction layers portraying drivers of development, our objective can be accomplished.

2. Study Area

The Baltimore Metropolitan Region (Figure 1) provides an ideal location to test metrics of land use change. Consisting of Baltimore City and five surrounding counties (Baltimore, Howard, Harford, Carroll, and Anne Arundel) the area contains both rural and urban expanses. Over the previous years the study site has also experienced increasing amounts of suburbanization and low density development (Jantz and Goetz, 2005) making it a prime location to test SLEUTH's performance with human drivers of development. Figure 2 displays the percentage of urban coverage within RPDs as well as the difference in urban coverage calculated between 1984 and 2006.

Urbanization of the Baltimore Metropolitan Region seems to be occurring for two reasons. First, the City of Baltimore has been experiencing great decline in the number of residents over the past decades. A decrease from nearly one million residents in 1950 to only 600 thousand in 2004 portrays mass exodus from the urban core. During that same period the neighboring, suburban, Baltimore County experienced a population increases of more than 179%, as the number of residents grew from a quarter to three quarters of a million (Short, 2007). Not only were people relocating from the city to its surrounding suburbs, but the population of the metropolitan region as a whole was on the rise. The Baltimore PMSA grew by 23% during the fifty-year span, adding nearly half a million residents (Hanlon, et al., 2010).

New human settlement patterns, in its sprawling/haphazard nature, exact an enormous toll on the nature of land cover across the metropolitan region. Transitions to developed/urban lands affect agricultural land uses particularly hard. Many counties within the study area had high compositions of agricultural lands, usually greater than 50% of total land coverage at the middle of the last century. By the turn of the 21st century however, most had lost more than half of these lands to urban development (Short, 2007). It was not until the late 90's that the state of Maryland realized the threat that exurban sprawl posed to valuable resource lands, causing it to pass legislation limiting low density development. The most notable measures of the statewide smart growth policy act aimed at curbing sprawl were Priority Funding Areas and Rural Legacy Areas (Jantz and Goetz, 2005). These attempted to guide development to existing communities while protecting farmland and open space by purchasing development rights and targeting funding towards already developed areas (Hanlon, et al., 2010). These targeted areas have a high correlation the sewer/water driver tested during this study, and test the assumption that existing infrastructure attracts future expansion.

The population/employment driver of development should also be mentioned in this section as they are linked to Regional Planning Districts. Regional Planning Districts are geographic areas comprised of one or more census tracts, and used mainly for transportation planning (BMC, 2008). Data was gathered by the Baltimore Metropolitan Council and the Cooperative Forecasting Group made up of economists, demographers, and planners within the State of Maryland's planning institutions. Published by this group are population, household, and employment forecasts for the study area that are derived from Census reports, surveys, economic analyses, and finally policies that affect land cover change. Within our study area are 94 RPDs from which population/employment statistics were obtained for this research. It is important to note that several regional planning districts within the central Baltimore City were removed because simulating development trends was difficult as the area already has greater than 90% impervious surface cover, and population/employment trends do not correlate to those in the suburbs where much of the study focuses.

3 Data and Methods

3.1 SLEUTH model

The SLEUTH cellular automata model relies on growth rules to simulate historic growth and forecast future growth scenarios. The four growth rules are: spontaneous new growth; new spreading center; edge growth; and road influenced growth (Clarke, et al., 1997). Spontaneous new growth acts much the same as exurban development today, as it is located outside of the existing urban centers in rural hinterlands. An example of new spreading center growth is the establishment and expansion of a new housing development or shopping mall. Edge growth is simply the continued expansion of existing centers of urban growth, and road influenced growth is that which occurs in proximity to existing transportation networks.

The aforementioned growth rules or types correspond to a set of coefficients that range in value from 0 to 100, and indicate how much of an influence the different growth types have across the study area. Controlling coefficients of growth are diffusion, breed, spread, road gravity, and slope resistance. For the most part these coefficients match up with one of the growth types: spontaneous growth is managed by the diffusion coefficient; edge by the spread coefficient; new spreading center by the breed coefficient. Road influenced growth is controlled by multiple coefficients consisting of road-gravity, dispersion, and breed. The final growth type and coefficient is slope resistance. This affects how likely growth is to extend onto steeper slopes. The exclusion/attraction layer can play a role in this decision as well. Utilized to increase or decrease the probability of development, the layer can prevent, or make unlikely, the occurrence of growth in areas such as protected lands, agricultural easements, and rivers/water bodies. In most previous SLEUTH studies, 0 correlated to neutral growth opportunity while 100 meant no growth could occur. This study however moves the neutral value to 50, making areas assigned a value of 0 an attractor, while 100 remains constant as a growth repellant.

Besides growth rules and coefficients there is another factor influencing how SLEUTH creates and predicts where development will occur. The model has a self-modification feature, also known as “boom and bust” which changes the rate at which growth is occurring. This function is initialized when growth rates shift above (boom) or below (bust) critical thresholds established by the user. If rapid growth is occurring the model will activate a boom phase where growth parameters are multiplied by a value great than one, and if little growth is occurring, or busting, the parameters will be multiplied by a value less than one. This allows SLEUTH to simulate dynamic growth rates over time, so that growth rates replicate the typical S-curve of urbanization and population growth (Silva and Clarke, 2002).

3.2 Changes to the SLEUTH model

While the underlying growth processes modeled by SLEUTH have remained unchanged, there have been some advances to the system and its programming over the years. The more recent SLEUTH-3r model attempts to improve the system by addressing several limitations of the original model (Jantz, et al., 2009). These limitations include an over prediction of edge growth when dealing with fine resolution imagery, like that from the Chesapeake Bay Watershed Land Cover Dataset (CBWLCD) (USGS, 2010) used in this study. The original SLEUTH model also required four historic datasets for calibration. These datasets were required because SLEUTH’s performance measures were based on regression analyses, and the calculation of the coefficient of determination requires a minimum of four data points. Reliance on the r^2 values sometimes lead to unintended under/over prediction in the model. Also when running simulations, especially during the calibration processes, SLEUTH had extensive memory and processing requirements.

The SLEUTH-3r model addressed some of these problems through the modification of the code and modeling process. It allowed for users to set a multiplier for the diffusion coefficient that controls edge growth (Jantz and Goetz, 2005), preventing a bias for that type of development occurring in the original version, and allowing diffusion to take a larger role in some situations. Rather than being a constant value in every type of scenario, it can be tailored towards individual patterns of development. The changes also allowed for decreased numbers of historic control points. While one formerly needed four time steps to run the model, it can now successfully run with only two since SLEUTH-3r calculates fit statistics using ratios of difference between the modeled and actual urban values, rather than using the coefficient of determination. However, if one can still provide four control points they receive additional calculations presenting statistics of modeled versus actual urban footprints. Further information provided by the model were a series of ratios for the fit statistics, in addition to the usual number of urban pixels, number of edge pixels, number of clusters, cluster size, and various other calculations. Lastly, some streamlining of the model was done to improve SLEUTH's memory demands and the speed at which it processed its computations (Jantz, et al., 2009).

3.3 Data requirements

As mentioned earlier, the main datasets required for SLEUTH to perform its operations include a time series of land use and urban data, a DEM, a transportation layer, and the exclusion/attraction layer. For this study the land cover and urban datasets were acquired from the CBWLCD. Collected in 1984, 1992, 2001, and 2006, these Landsat-derived datasets provided the most accurate and up-to-date land use classifications for the study area. The accuracy comes from this data set's ability to capture low density development in its "developed open space" category. Low density development is sometimes overlooked and/or misclassified in other instances of land cover data sets derived from remotely sensed satellite imagery. Capturing the low density development in the exurban counties throughout the study area which experienced large quantities of dispersed growth (Jantz and Goetz, 2007) over the past decade, SLEUTH is able to more realistically simulate urban growth patterns utilizing the urban land cover footprints provided. The datasets were divided into 16 different land use classifications (Anderson, et al., 1976). For this study we considered developed land classes to be urban areas thus contributing to later forecasting of growth, all other natural/agricultural areas were open to development when not taking other layers like slope gradient and the exclusion/attraction layers into account.

The transportation layer used in the SLEUTH analysis contains U.S. and Maryland State highways, interstate highways, and important hand-selected county roads. After compiling the line features a conversion to raster format was needed to make them recognizable to the program. The slope layer was established for the study area using the USGS National Elevation Dataset (NED). By applying the slope tool to the dataset, calculation of slope for the area was achieved.

The final input into the SLEUTH model is the exclusion/attraction layer, which is where the user has the most input into the system. In this dataset an understanding of policy and human growth patterns can be used to coerce and hopefully improve modeling performance. For this study three differing exclusion/attraction layers were created to test drivers of development. These include a base exclude, sewer/water service area, and population/employment and will be explained separately in the next section. Figure 3 provides a visual representation of how these layers were incorporated into the model and the data utilized to create them.

3.4 Exclusion & attraction layers

The first version to the exclusion/attraction layer was the standard base layer either excluding growth completely or allowing it in a neutral state. Particular features which were excluded included public lands, parks, easements, wetlands, and water. Transportation systems were also off limits to building, so roads, clover leaf intersections, and railroads were assigned values of 100. Finally, we added other areas which were off limits to building and had significant urban cover such as airports and military bases. These areas may have possessed lots of open space which SLEUTH believed to be compatible with development, but had to be removed for obvious reasons. All areas not falling within areas assigned a value of 100 for exclusion were given a neutral weight.

Our second exclusion/attraction layer created for this study consisted of the sewer service area. The rationale behind this test was based on the assumption that the location and density of development could be determined by whether or not the area was able to access the sewer/water network. Areas within that area of network connection would presumably be developed more intensely as the infrastructure was already in place. Data for this layer were compiled by collecting service area shapefiles from the various counties within the study area. In terms of the layer itself, the previously mentioned standard exclusion/attraction layer remained the same, with values of 100. However, if a neutral area now fell within the sewer service area, it was to be used as an "attractor" of growth, so the value which was once 50, was now to be tested at several increments (0, 10, 20, 30, and 40) to see which weight best simulated what was actually occurring.

The third and final exclusion/attraction layer put together for this research was meant to capture the all-day human intensity of the present population. All-day human intensity refers to the stresses of development placed upon RPDs by growth in population and employment. The layer was created using population and employment data from the Baltimore Metropolitan Council (BMC, 2010). While explained in detail within a National Science Foundation (NSF) Report (Jantz and Drzyzga, 2011), a brief overview of how these data were incorporated into the model will be given here.

Using past BMC forecasts along with the most recent that run through 2030 we were able to compile a data series that fully integrated into the time series for which we had obtained land use coverages. This required extrapolation of population and employment values to create annual datasets to match up exactly with our CBWLCD series. Growth rates within RPD are not homogenous across the region, with some capturing more of the region's growth than others. To identify the significance of RPDs a location quotient was used. This measured the proportion of population and employment change of a particular regional planning district compared to the regional average. With the location quotient values the next step was to scale them into the 0 to 100 excluded/attraction scores. To accomplish this final step a logarithmic transformation was used to scale the values that could range from 0 - infinity to 0 – 100 and also account for 50 being the neutral middle value.

This series of calculations produced an exclusion/attraction score for each RPD, which was calculated for three different time intervals during the modeling process. For calibration of the model scores were computed for the period of 1984 – 2001, for validation 2001 – 2006, and for forecasting 2006 – 2030. This layer once again overlays the 100 valued areas of exclusion within the original exclusion/attraction layer. Where it is more complicated however, is the value assigned to each RPD. Values may range from 0 to 100 in terms of how much weight the area had in terms of population and employment. Values also ranged differently for each of the steps calibration, validation, and prediction. This layer attempts push development into RPDs where the growth in all-day human intensity was higher than the regional average, and repel growth from RPDs where it was lower.

3.5 Calibration

Once the exclusion/attraction layers and datasets were assembled we established the process through which calibration, validation, and forecasting would occur. The first step, calibration, is where the model tries to replicate growth patterns from the past. Mentioned earlier, this will occur between 1984 and 2001. We first used a "brute force" calibration (Clarke, et al., 1997) that tested parameters for their goodness of fit. In total 3125 different parameter combinations were tested, with values for each parameter ranging from 1 – 100 in increments of 25. To ensure that variability within SLEUTH's random processes is accounted for, several Monte Carlo trials are performed for each combination of parameters. During this research 10 Monte Carlo trials were run for the 3125 parameter combinations. Once results are returned several of the "best" parameter sets are picked based upon their fit statistics. Fractional differences between modeled and actual urbanization trends were measured in terms of area, edge, and cluster. Area measures the difference in the total area of impervious surface cover. Edge measures the difference in the total perimeter of urban clusters. Clusters measure the difference in the number of unique development areas, a measure of fragmentation. These metrics were chosen because of their relevance to the study of spatial patterns of development.

They also measure different aspects of said development, reducing redundancy in fit statistic results. And lastly, metrics attempting to average all metrics into a single "optimal metric" (Dietzel and Clarke, 2007) were avoided. Optimal metrics can be influenced by less relevant metrics to development patterns which decrease the correlation between the metric and development patterns (Jantz and Goetz, 2005). To be considered a good match, a scenario should match all three of these fit statistics within +/- 5% of actual development patterns. If a fit statistic is outside of this range a fine calibration can be done to try to improve the match and scenario's capabilities. In this case it meant breaking down the increments of 25 to 5. Once chosen, the best scenarios were re-tested over 100 Monte Carlo trials. To ensure that the best scenario is chosen the model results are then compared spatially. The best model results are imported into a GIS and evaluated against the actual land cover. Mapped versus modeled comparisons were completed at the 480m x 480m resolution to assess the accuracy of the model and fit statistics at a finer scale using a differences map. This identified the effectiveness of the regional scale fit statistics to capture development patterns correctly. To create the differences map the percentage of urban coverage would be calculated for both mapped and modeled coverages with the following equation; from which their differences could then be compiled:

$$(((\text{Sum of Impervious Surface for 30m Cells} / 100) * 900) / \text{480m cell size})$$

This allows for visible portrayal of over and underestimation by the SLEUTH model at the local scale. Combined with the fit statistics that measure performance at the regional scale, we are able to observe how well each scenario/run matches with the actual development patterns by identifying where and by how much the model is over or underestimating at both local and regional scales.

3.6 Validation

The validation process is divided into two steps, unconstrained and constrained. This will determine which of the base, sewer, RPD calibrations are best and then also begins preparations for forecasting scenarios up until 2030. The process utilizes the urban footprint from 2001 and predicts out to 2006 based upon the coefficient values from scenarios. This will allow us to validate the calibrations based on the model's ability to match the actual mapped 2006 data, which was withheld from SLEUTH during calibration, thus acting as an independent variable. During the validation phase we can also create varied growth projections. Growth projections offer different future outcomes based upon rates of increase in impervious surface coverage.

The unconstrained portion of validation is where we compare the best base, sewer, and RPD calibration results by forecasting them to 2006. During this time SLEUTH is programmed with the parameter sets from the best calibration runs, for 100 Monte Carlo trials, and does not activate any self modification features. No additional information will be provided during these forecasts until 2006. SLEUTH develops land using an empirical constant growth rate which means that all scenarios will linearly predict (based on the 1984 – 2001 growth rate) the amount of growth that occurred during the 2001 – 2006 timeframe, but the scenario that is closest to the actual, or mapped, amount of impervious surface coverage will be deemed the best. To address changing rates of growth a second stage of validation is done called the constrained phase. During actual prediction phases of modeling, extra information is provided to the model, such as forecasts for land development. Self-modification is also allowed, giving SLEUTH the ability to simulate changing trends in growth rather than constant ones. For example the initial time series of this study shows rapid development which tapers off towards later time series and has slowed drastically in present years.

During the constrained portion of validation the best performing unconstrained scenarios moves on and is given a development “target” to hit. This target during validation will be the BMC forecasted estimates for urban land cover within the BMR for the year 2006. Thus while both preparing SLEUTH and later utilizing it for 2030 predictions; the data inputs are of the same caliber and from the same source. To hit these forecasted estimates from the BMC we allow SLEUTH the ability to use its self-modification feature. By setting it to boom or bust it can be coerced quite accurately to hit the amounts of development which were forecasted. From here we will be able to compare the performance of the base, sewer, and population/employment exclusion/attraction layers against one another rather than evaluating them against various runs of the same version, say base against base. Once again these differences were looked at as differences in the percentage of urban coverage within 480 x 480m cells, as explained earlier.

After comparing the base, sewer, and population/employment scenarios, what we consider the best all around performer will begin another step. For this portion of the study a series of three growth projections were created, correlating to rates of urbanization throughout the study area. The scenarios make use of compound annual growth rates and an ‘adjustment exponent’ that simulates future growth throughout the validation period (2001 – 2006) along continuing trend lines. By increasing and decreasing the exponent, the growth rate can be either increased or decreased to meet potential future outcomes. The three projections were named min, max, and status quo. The status quo has an adjustment exponent of 1and is a continuation of compound annual rates of growth where future urban transitions continue along a steady path created by past trends. Min correlates to smart growth practices that conserve lands, therefore minimizing growth in impervious surface cover. By dialing down the adjustment exponent to 0.96 growth within the study area with be around 1 km^2 per year. Max is a worst case scenario or development boom, where of impervious surface increases continuously, but not so much that all-day human density would decrease during later prediction phases. An adjustment exponent of 1.032 met the standards listed for the max scenario (Jantz and Drzyzga, 2011).

To summarize, the best calibration scenarios for the base, sewer, and population/employment move on the validation. From there they are left unconstrained and run in predictive mode with no additional information provided to SLEUTH. The best performers from each of the different scenarios is picked and then constrained by the 2006 population. After comparing modeled results to mapped the best scenario (base, sewer, or RPD) goes on to a projection phase. A variety of projections are created based upon differing rates of growth. After choosing the best performing scenario, the process of predicting out until the year 2030 can begin. SLEUTH will once again be running in predict mode, and it will be given the latest 2006 urban land cover data as a starting point. In predicting out until 2030, it is also possible to continue to create several different scenarios based upon current development trends and other forecast values.

3.7 Forecasting 2030

Forecasting to 2030 occurred in similar fashion to other recently published pieces (Jantz, et al., 2009; Onstead and Clarke, 2011). The initialization of the SLEUTH 3-r model occurred with an urban extent map from 2006. Then a series of growth projections representing rates of urban growth at accelerated, constant, and decelerated levels was created for the 2006 through

2030 time period similar to those created late in the validation phase by compounding annual growth rates from BMC forecasts. These projections will once again be named minimum, status quo, and maximum.

4. Results

4.1 Exclusion/Attraction Layers

Images of exclusion/attraction layers for all three scenarios are shown in Figure 4. The base shows the most areas weighted as neutral with undevelopable lands and waterways incorporated. The sewer service scenario expands upon the base excluded attraction layer by overlaying on top of it the 'attracting' areas with sewer/water connections. Finally the RPD layer with its derived weights is overlaid with the undevelopable lands from the base layer. Figure 5 portrays all of the RPD weights for each of the modeling phases.

4.2 Calibration

Calibration results for each scenarios (base, sewer, and RPD) best run display both the best-fit parameters and the fit statistics used to guide the model (Tables 1 – 3). All of the scenarios matched the impervious surface area within 5% and edge/cluster development for all three were well within 10% of actual trends. While the base and sewer service area scenarios only required a coarse calibration, a fine calibration was used on the RPD scenario following the coarse calibration to try to improve the fractional differences seen in the fit statistics. The spread coefficient was the only one to undergo this change and it was found that a value of 20 brought the fractional differences within the 5% and 10% thresholds set for the varying fit statistics.

4.3 Validation

The same runs were used during calibration were once again used for the validation process, so the best-fit parameters remained the same as in Tables 1 - 3. The fit statistics changed slightly and are shown in the validation portion of previously mentioned tables. In addition to these results boom and bust modifiers are available in Table 4. During the unconstrained validation of all three scenarios the over prediction mentioned earlier was clear, caused by SLEUTHs assumption of constant growth. The RPD scenario however did show the best prediction of urban area at 1,510.9 km^2 compared to the actual 1,456.63 km^2 mapped in 2006. Figure 6 shows a comparison of unconstrained scenario results which portray mapped versus modeled differences in the percentage of urban coverage. The graph in Figure 7 also demonstrates differences in the different scenarios and provides the actual baseline to which they are being compared. It is important to emphasize once again that although all the results are higher than actual area values for the time period, this is consistent with SLEUTH's unconstrained nature to empirically develop lands at a constant rate.

Maps of the differences in the percentage of urban coverage between the three modeled scenarios (Figure 8) also support the claim that RPD best captured urban development trends and should therefore move on in the validation process. The sewer scenario draws much greater proportions of growth into the service area than that which actually occurred. This in turn increased the amount of underprediction in the rural areas which received less than their fair share. The RPD scenario did the best job at minimizing differences in the percentage of urban coverage throughout the entire study area save a few spots that experienced drastically increased growth rates that SLEUTH could not account for.

After constraining the RPD scenarios with population and employment data from past BMC forecasts, three different growth projections were completed. A graph of growth projection values can be seen in Figure 9. This shows that the RPD Min growth projection best matched actual growth trends until 2006 resulting in a near perfect match to the amount of urban coverage. Maps in Figure 10 show all three growth projections once again indicating differences in the percentage of urban coverage. The Min map supports the graph presented earlier that it is indeed the most accurate of the growth projections for 2006.

4.3 2030 Forecasts

The final output from SLEUTH is a series of prediction maps for the year 2030. In this case three forecasts (Figure 11) were made to encompass the same three growth scenarios presented late during the constrained validation. Utilizing the forecast data from the BMR, a minimum growth rate scenario was created along with the status quo and maximum growth counterparts. As stated earlier, these were meant to replicate a range in the amount of urban land coverage experienced across the study area, potentially exposed to more stringent smart growth policy, continuing along recent paths, or enduring even more accelerated unmonitored growth than today. Named accelerated, constant, and decelerated, these growth rate scenarios correlate to the maximum, status quo, and minimum growth projections during validation. Boom and bust modifiers were required to hit these forecasts as they were in the validation phase. Table 4 contains the modifier values for 2030 projections.

Mapped results for the 2030 forecasts can be viewed in Figure 12. This map portrays in the status quo map the percentage of urban impervious surface coverage across the entire study area should growth rates continue along the most current predicted trends. Due to the fact that all three forecast maps would appear almost identical without it, the "what if" decelerated/accelerated forecast maps were made to be difference maps between their percent urban coverage's and the status quo. As shown by the difference values in the map legends and maps, the differences were generally experienced within the urban areas and how intensely development occurred. Minimal differences were common in the rural areas where growth amounts were calculated by SLEUTH was less variable.

5. Discussion and Conclusion

The results presented in this paper support the need to look at drivers of human development within modeling and also begin to question the SLEUTH process, also questioned and examined by its own creator (Clarke, 2008). The basic excluded layer that includes only lands off limits to development has been the default 'go to' in past SLEUTH modeling research due mostly because of its simplistic nature. However, this research has shown that it can be outperformed and for a more accurate modeling result, one can utilize more extensive datasets to build an exclusion/attraction layer that explicitly incorporates "drivers" of change. While adding potentially more relevant information to the SLEUTH model, the exclusion attraction layer can also add functionality that went unutilized in the past. Being able to coerce growth into certain areas can be as useful as keeping it out of others.

In addition to changes in the excluded or excluded/attraction layer and methodology, alteration to the methodology of SLEUTH's typical processes could also be in store. While small datasets may have constrained these methodological additions in the past, the validation phase now seems a necessary requirement within the modeling procedure. Many have chosen to ignore this process, but with multiple time steps becoming more commonly available it seems only logical. Predicting to a point in time where data is available for comparison allows the user to check the accuracy of their calibration and potentially adjust it for improved predictive capabilities. It can also greatly enhance the understanding of urban change processes.

When looking back at the steps taken to improve the modeling process and SLEUTH's predictive capabilities by integrating and supplying more functional data into the system through the exclusion/attraction layer, it can not be denied that this has allowed the model to improve its predictive capabilities. The population and employment data which was converted into exclusion/attraction scores for each RPD within the study area improved SLEUTH's efficiency in allocating urban coverage to areas experiencing development, and prevented it from those that were not. These points are reinforced by the difference maps earlier, but at the same time, present a need to test different scenarios as well. Our sewer/water service area exclusion/attraction layer, most simply put, did not perform the way it was expected to. It drew a much larger portion of development into previously urban areas that were actually headed to the suburbs. For this reason it may be important to test multiple drivers, as was done in this research. While researchers may believe a driver can improve the model by influencing it in a certain way, this demonstrates that this is not always the case. Testing multiple drivers with multiple exclusion/attraction layers should produce a variety of results from which the best can be chosen to run SLEUTH for more accurate predictions.

The more common practice of providing a range of forecasting targets (Jantz, et al., 2009; Onsted and Clarke, 2011) was also reinforced in this research. While the future is unknown, it can typically be predicted to within a range or limits of a certain value. Presented here were scenarios of minimal, constant, and maximum growth. While these results offered a +/- 5%

range from the status quo of a status quo growth rate, it shows how much of an impact unhindered or constrained growth could have upon the landscape of the study area.

To summarize, the goals of the research were met. A variety of human drivers were tested and in the end it was found that the population/employment data acquired from the Baltimore Metropolitan Council could be utilized at the RPD scale to provide the most influential information for SLEUTH to utilize in the form of a exclusion/attraction layer. The RPD data outperformed both the base and sewer exclusion/attraction layers, of which the sewer actually performed the worst. By outperforming the base exclusion layer the research has shown a need to question the information being applied to a model, is it effective, can it be improved? While this is just one instance, others have begun asking questions of SLEUTH (Clarke, 2005; Onsted and Clarke, In Press) and there is a need to continue asking these types of questions in regards to SLEUTH. It is also important to continue to try to advance the model both in the data provided to it and in the methodology by which it operates. SLEUTH is a model that has been around for some time now, and by seeking out answers to questions like these, it can be utilized with future datasets to come.

References

Anderson, J.R., E.E. Hardy, J.T. Roach, and R.E. Witmer. 1976. A Land Use And Land Cover Classification System For Use With Remote Sensor Data. Geological Survey Professional Paper 964. A revision of the land use classification system as presented in U.S. Geological Survey Circular 671.

Clarke, K.C. 2008. A Decade of Cellular Modeling with SLEUTH: Unresolved Issues and Problems, Ch. 3 in *Planning Support Systems for Cities and Regions* (Ed. Brail, R.K., Lincoln Institute of Land Policy, Cambridge, MA, pp 47 – 60

Clarke, K. C. 2005. The limits of simplicity: toward geocomputational honesty in urban modeling. Atkinson, G. Foody, S. Darby, & F. Wu (Eds.), GeoDynamics (pp. 215–232). Boca Raton, USA: CRC Press.

Clarke, K. C., S. Hoppen, and L. Gaydos. 1997. A Self-modifying Cellular Automaton Model of Historical Urbanization in the San Francisco Bay Area. *Environment and Planning B: Planning and Design* 24:247-261.

Dietzel, C. and Clarke, K.C. 2007. Toward Optimal Calibration of the SLEUTH Land Use Change Model. *Transactions in GIS*, 11.1, 29-45

Hanlon, B., S. R. Short, and T. J. Vicino. 2010. Cities and Suburbs: New Metropolitan Realities in the US. New York, NY: Routledge

Jantz, CA and SA Drzyzga. 2011. Annual report for period 09/2010 – 08/2011. CNH: Collaborative Research: Dynamic Coupling of the Water Cycle and Patterns of Urban

Growth. Award ID: 0709537. Submitted to National Science Foundation September 26, 2011.

Jantz, C. A., and S. J. Goetz. 2007. Can smart growth save the Chesapeake Bay? *International Journal of Green Building* 2:41-51.

Jantz, C. A., S. J. Goetz, D. Donato, and P. Claggett. 2009. Designing and implementing a regional urban modeling system using the SLEUTH cellular urban model. *Computers, Environment & Urban Systems* 34:1-16.

Oguz, Hakan, A.G. Klein and R. Srinivasan. 2007. Using the Sleuth Urban Growth Model to Simulate the Impacts of Future Policy Scenarios on Urban Land Use in the Houston-Galveston-Brazoria CMSA. Research Journal of Social Sciences, 1: 72-82

Onstead, J., and K. Clarke. 2011. "Forecasting Enrollment in Differential Assessment Programs Using Cellular Automata." *Environment and Planning B*: *Planning and Design* 38(5) 829 - 849

Onstead, J., and K. Clarke. In Press "The inclusion of differentially assessed lands in urban growth model calibration: A comparison of two approaches using SLEUTH".
IJGIS's Volume 26 issue 5

Regional Data and Forecasting. Baltimore Metropolitan Council, 15 Nov. 2010. Web. 22 Feb. 2011. <http://www.baltometro.org/transportation-planning/regional-data-forecasting>

xShort, J. R. 2007. *Liquid City: Megalopolis and the Contemporary Northeast.* Washington, DC: Resources for the Future.

xSmall Geographic Areas. Baltimore Metropolitan Council, 29 Sept. 2008. Web. 22 Feb. 2011. <http://www.baltometro.org/content/view/468/362/#RPDs>

xSilva, E. A., and K. C. Clarke. 2002. Calibration of the SLEUTH urban growth model for Lisbon and Porto, Portugal. *Computers, Environment & Urban Systems* 26:525-552.

xUSGS Chesapeake Bay Activities. United States Geologic Survey, 25 Feb. 2010. Web. 7 Aug. 2010. <http://chesapeake.usgs.gov/data.html>

xXian, George. 2005. Dynamic modeling of Tampa Bay urban development using parallel computing. Computers & Geosciences 31:920-928

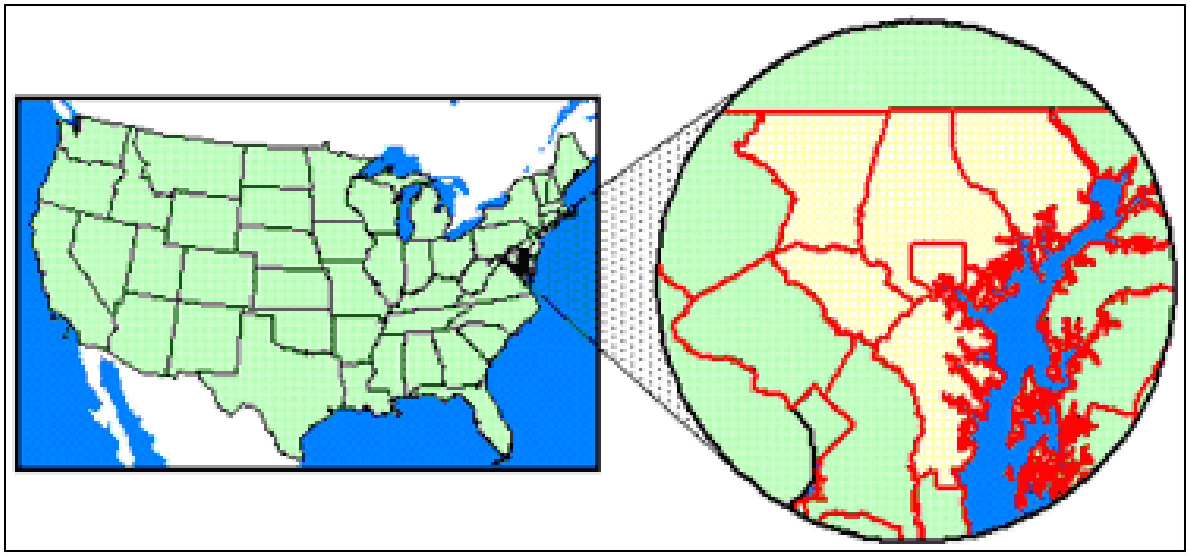


Figure 1: The Baltimore Metropolitan Region (BMR) located in central Maryland. The BMR is comprised of five counties (Anne Arundel, Baltimore, Carroll, Harford, Howard) and Baltimore City (BMC, 2008)

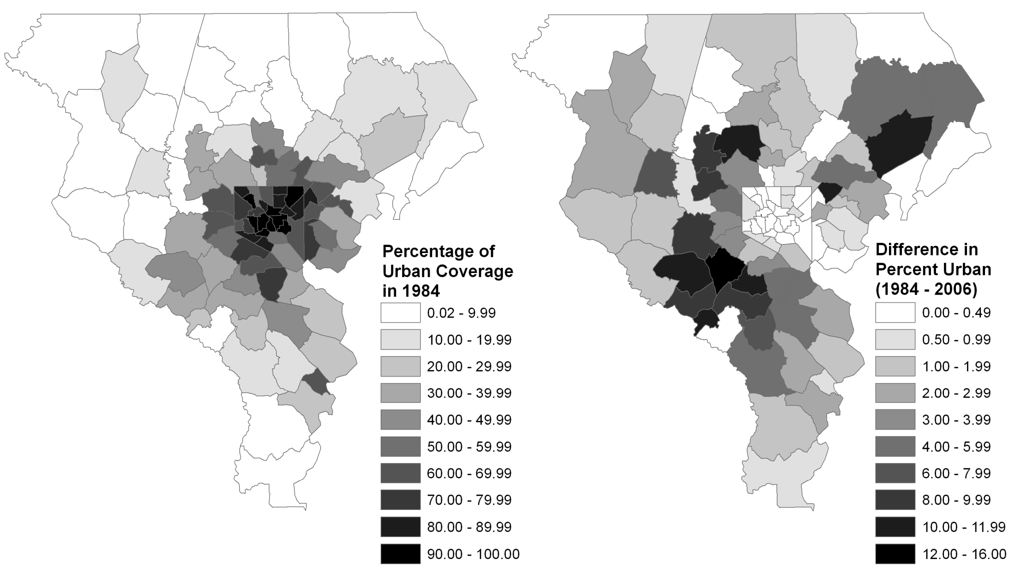


Figure 2. Percentage of urban coverage and difference in percentage of urban coverage (1984 – 2006) for RPDs within the BMR.

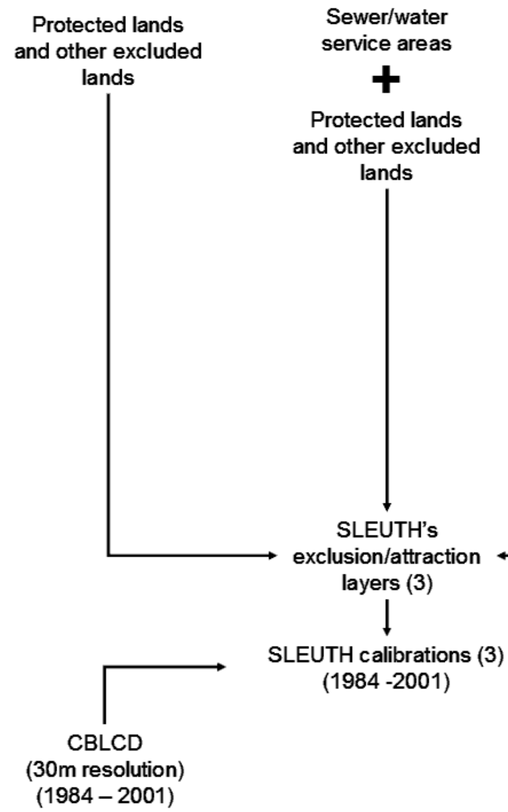


Figure 3: Generalized process flow for testing three different exclusion/attraction layers.

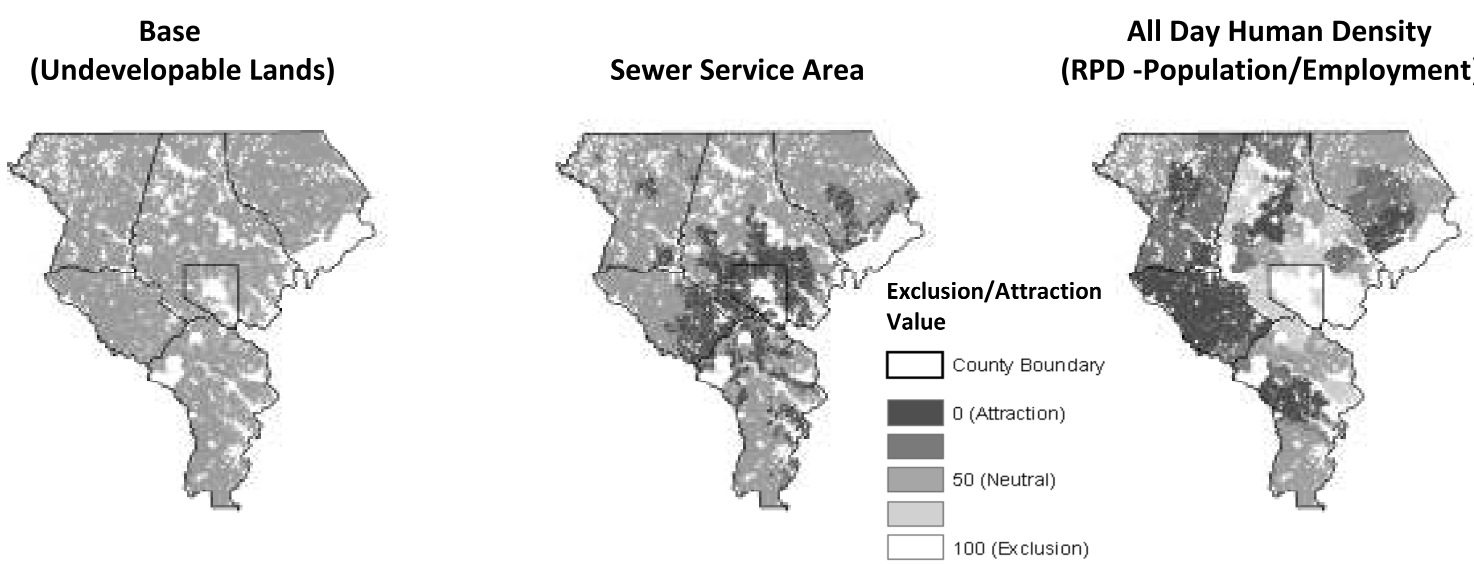


Figure 4: Exclusion/Attraction layers used for SLEUTH analyses.

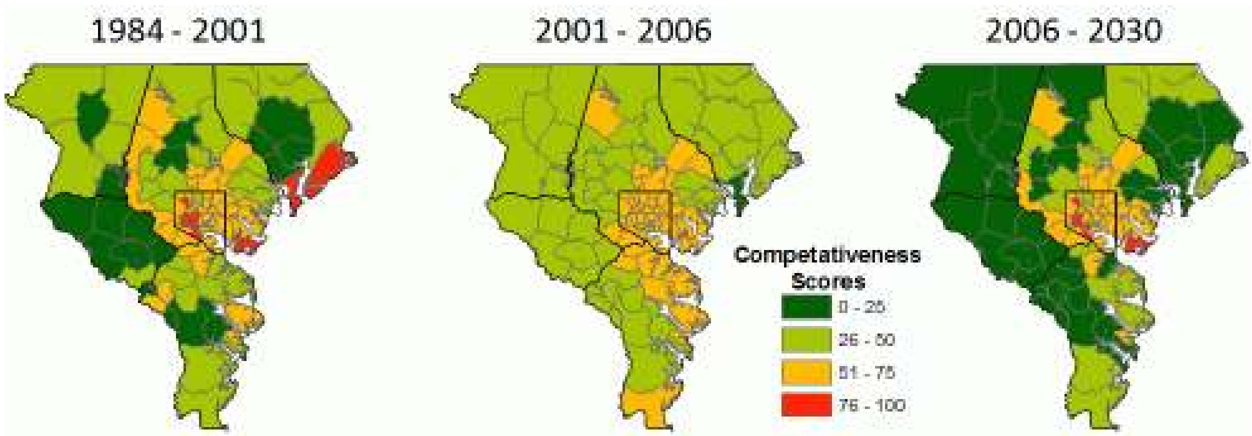


Figure 5: RPD weights for calibration (1984 – 2001), validation (2001 – 2006), and prediction (2006 – 2030).

Table 1:

Base	
Coefficients	**Values**
Diffusion	1
Breed	1
Spread	25
Slope Resistance	75
Road Gravity	50
Calibration Fit Statistics	**Fractional Differences**
Area	0.036264
Edge	0.039932
Cluster	0.079703
Validation Fit Statistics	**Fractional Differences**
Area	0.035516
Edge	0.003493
Cluster	-0.018479

Table 2:

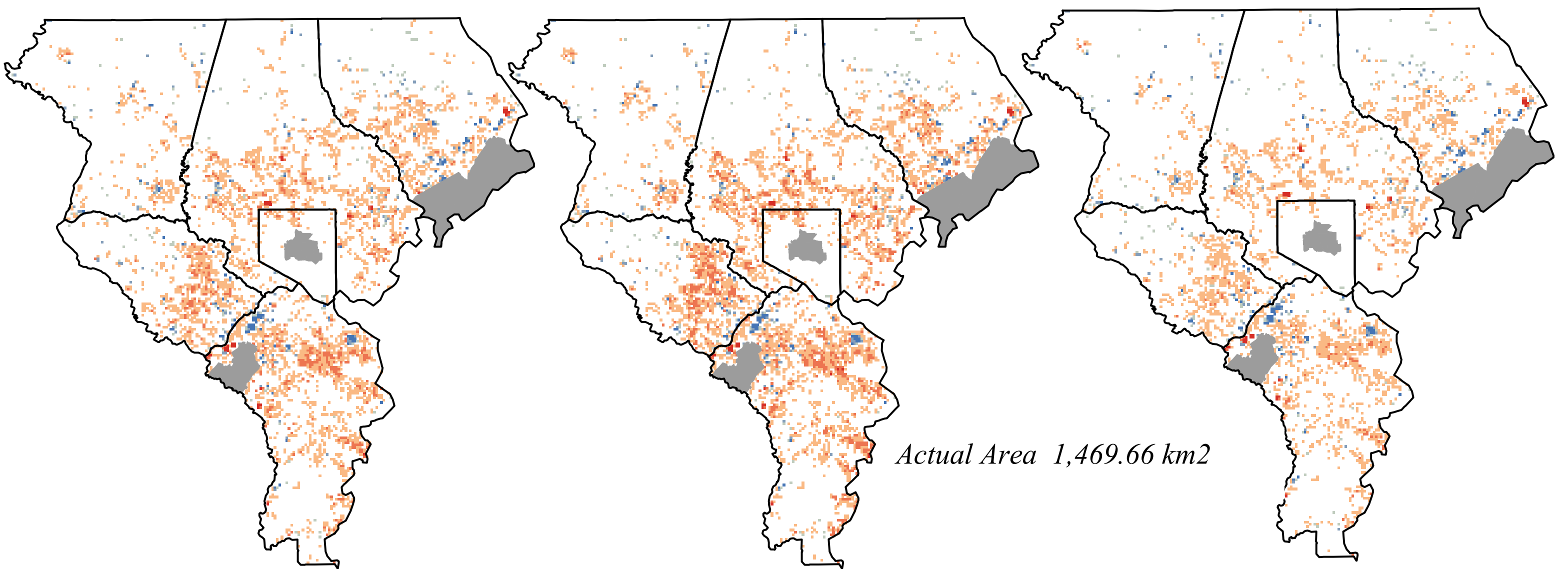
Sewer Service Area	
Coefficients	**Values**
Diffusion	1
Breed	1
Spread	25
Slope Resistance	100
Road Gravity	50
Calibration Fit Statistics	**Fractional Differences**
Area	0.044317
Edge	0.018441
Cluster	0.060596
Validation Fit Statistics	**Fractional Differences**
Area	0.0383612
Edge	-0.0018585
Cluster	-0.023065

Table 3:

RPD	
Coefficients	**Values**
Diffusion	1
Breed	25
Spread	20
Slope Resistance	75
Road Gravity	50
Calibration Fit Statistics	**Fractional Differences**
Area	0.046029
Edge	0.047072
Cluster	0.067445
Validation Fit Statistics	**Fractional Differences**
Area	0.027561
Edge	0.001152
Cluster	-0.017218

Table 4:

Boom/Bust Modifiers		
Validation 2006		
Projection	Boom/Bust	Value
Min	Bust	0.010
Status Quo	Bust	0.820
Max	Boom	1.205
Prediction 2030		
Projection	Boom/Bust	Value
Min	Bust	0.300
Status Quo	Bust	0.710
Max	Bust	0.818



Base Urban Area 1,521.88 km2 Sewer Urban Area 1,526.061 km2 RPD Urban Area 1,510.19 km2 (Best)

Figure 6: Difference Map of unconstrained validation results.

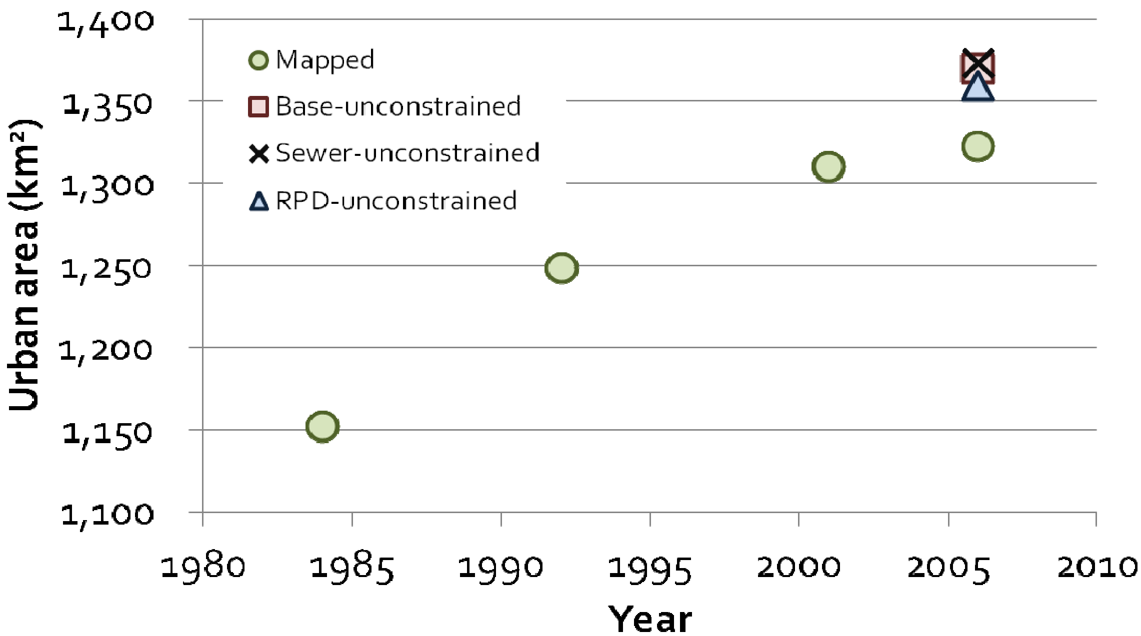


Figure 7: Best performing scenarios in unconstrained validation phase.

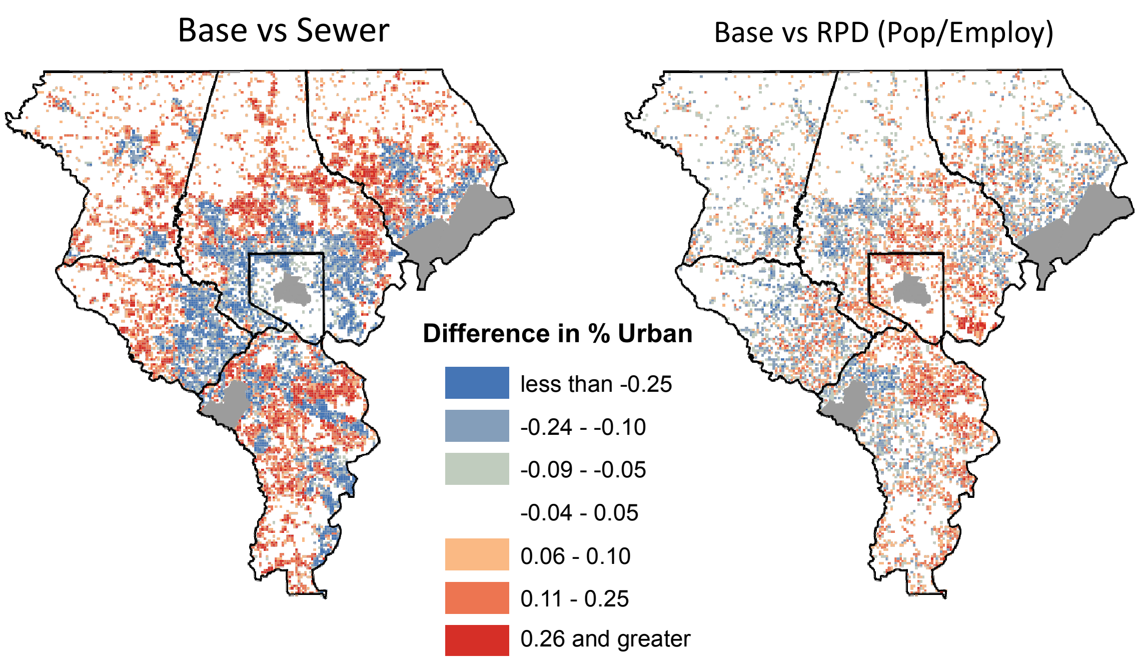


Figure 8: Difference Maps comparing scenarios constrained by 2006 population forecasts.

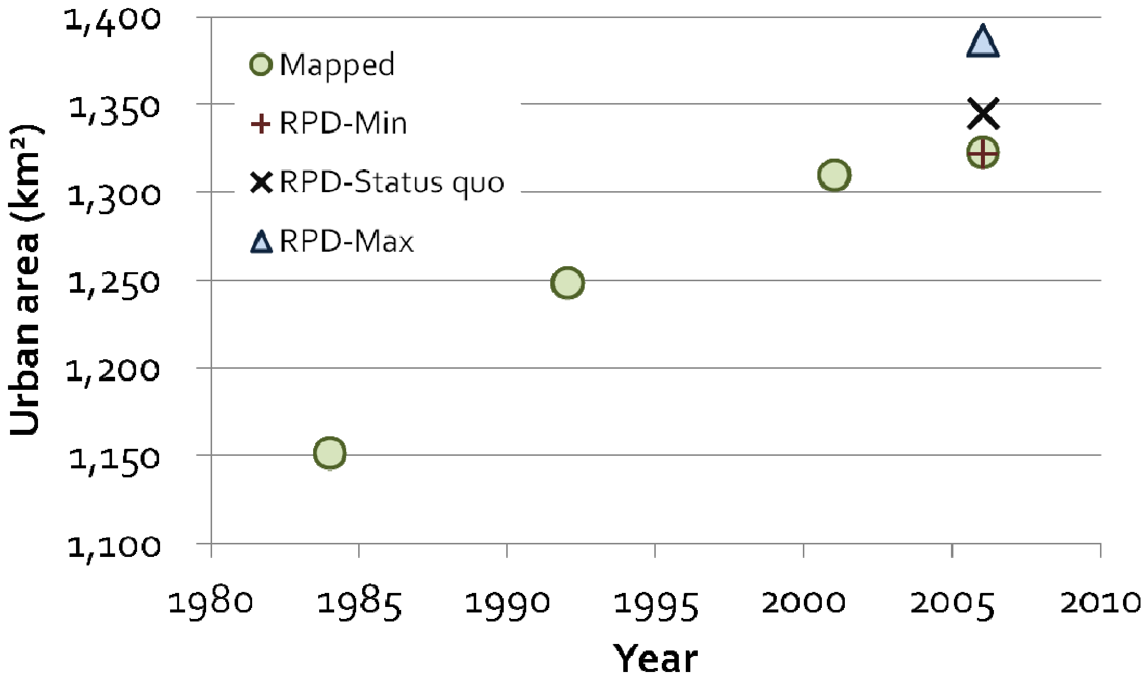


Figure 9: Growth scenarios created during constrained validation phase.

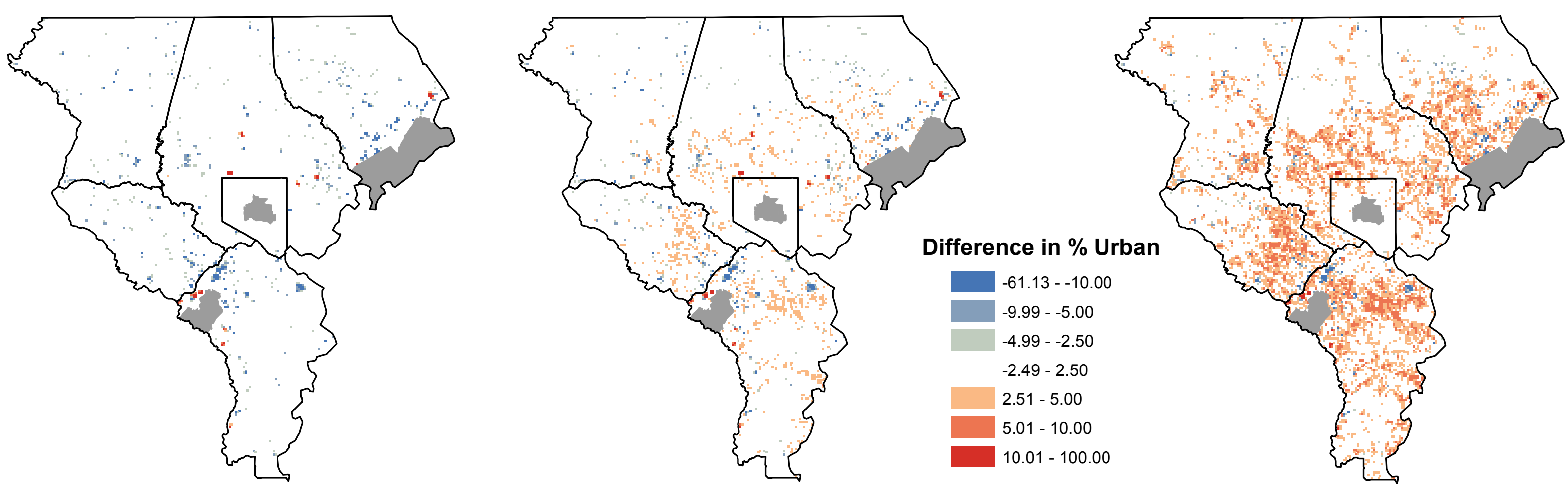


Figure 10: Differences in percent urban coverage for each of the growth scenarios during the constrained validation phase.

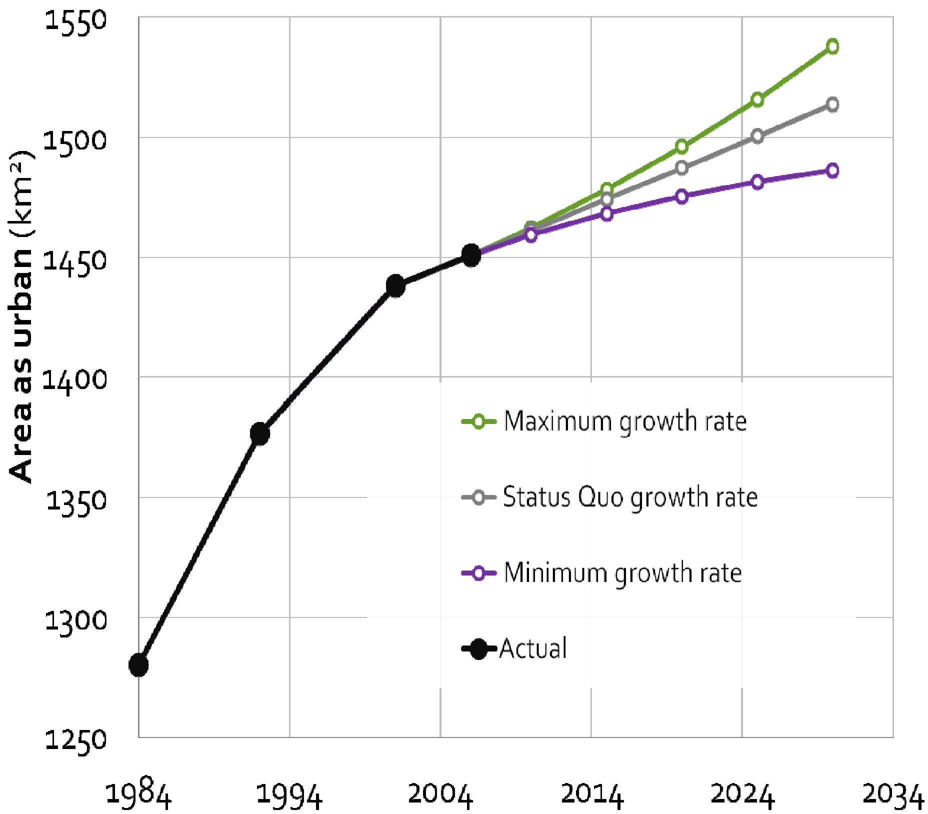


Figure 11: Growth rate scenarios through 2030.

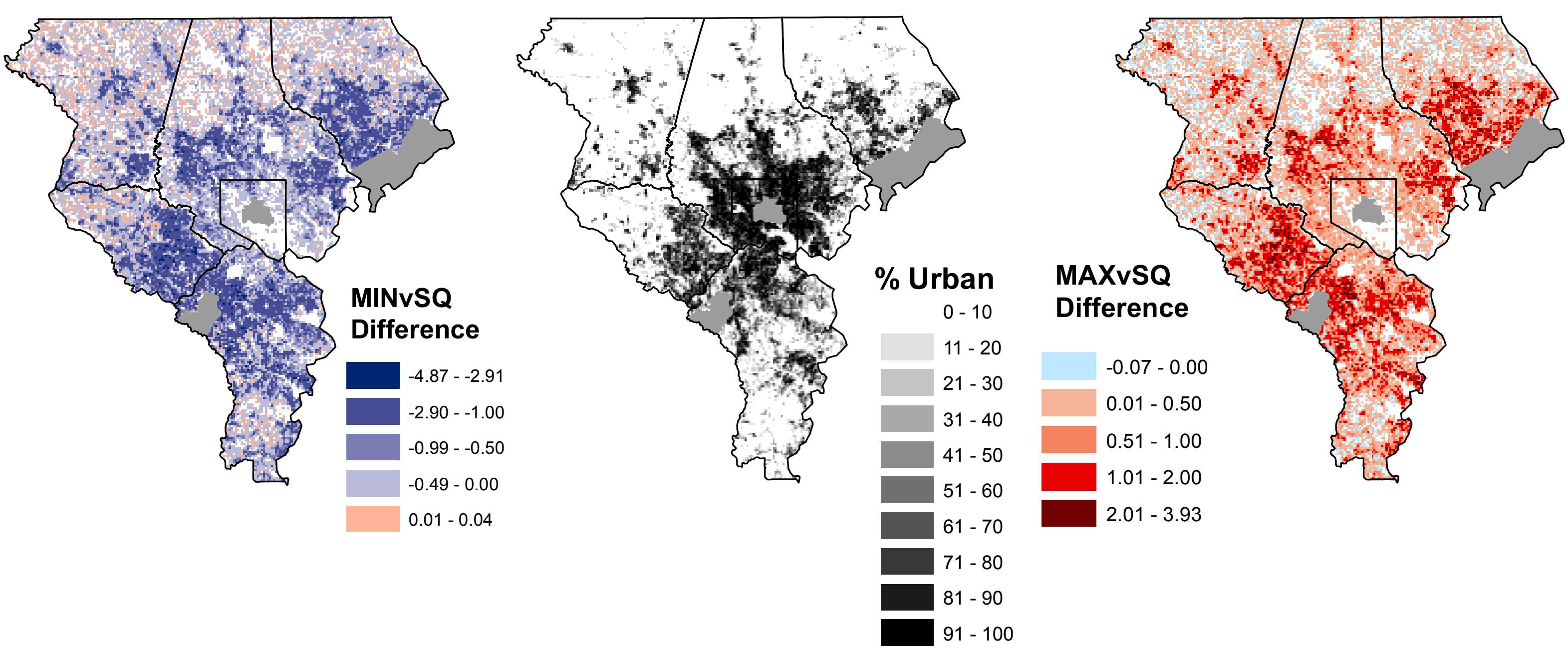


Figure 12: 2030 growth projections.