Methodology for Projecting Land Cover Change in the CARA Region

Richard Ready
Department of Agricultural Economics and Rural Sociology
Pennsylvania State University

and

Dennis Guignet
Department of Agricultural and Resource Economics
University of Maryland

The purpose of this study was to estimate a county-level statistical model that explains historical rates of land cover change caused by changes in land use, and use that model to predict future rates of land cover change. The specific land use change studied is the conversion of land from less-developed, open space uses (forests, pastures, cropland, wetland) to more developed uses (residential, commerical, industrial, roads). The statistical model was estimated based on data from 79 nonurban counties in Pennsylvania, New Jersey, Delaware, Maryland, Virginia and North Carolina. Land cover change projections were calculated for all counties in the CARA region.

Data Sources for the Statistical Model

Data on Land cover

The statistical model was estimated based on land cover change over the period 1992 to 2001. The USGS, working with other agencies, has produced maps of land cover for the United States based on satellite images. These maps divide the United States into square pixels, each about 30 meters by 30 meters. Using a combination of satellite images, the USGS determines the land cover for each pixel. The resulting national map is called the National Land Cover Database (NLCD). More information about the NLCD is available at http://www.mrlc.gov. This process was first done using satellite images taken during or around 1992. It is currently being done again using satellite images taken during or around 2001. Maps showing 2001 land cover are currently available only for some regions of the U.S.

NLCD maps for 1992 and 2001 were obtained for part of the mid-Atlantic region. The counties included in the analysis are shown in Figure 1. The NLCD land cover determinations were aggregated into two broad categories: land in open space and developed land[1]. Land in agricultural uses, forests, and wetlands were considered open space. Barren land, which includes quarries and land that has been cleared for development, was considered a developed use.

[1] Details on which land cover categories were considered open space and which were considered developed are provided in Appendix Table 2. Note that land in urban parks, lawns and golf courses were considered developed uses.

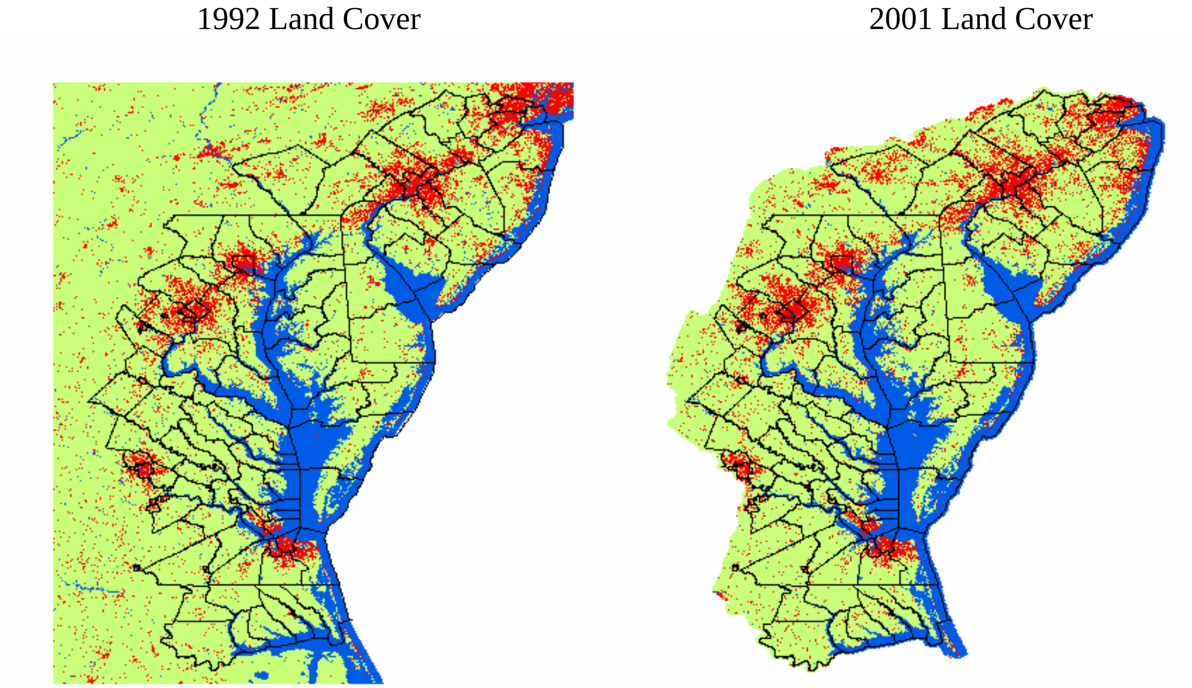

Figure 1. Study Area showing NLCD land cover in 1992 and 2001

Protected Areas

The purpose of the statistical model was to explain the rate at which open space was converted to developed uses during the period 1992 to 2001. Some open space land is protected from development. This includes national and state forests, game lands, wilderness areas, etc. Maps showing protected land were developed as part of the national GAP analysis program (http://gapanalysis.nbii.gov). These maps were obtained for the states included in the study area. Land that was protected at GAP level 3 or higher was excluded from the analysis. The protected lands in the study area are shown in Figure 2. Note that many lands that are in fact protected from development (such as land protected under many agricultural conservation easements) are not included in the GAP analysis.

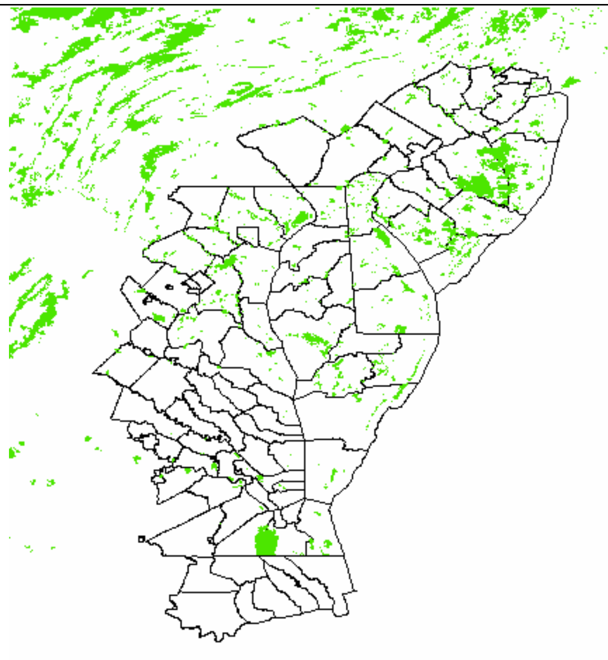

Figure 2. Protected Lands

Land Cover Change

For each nonurban county in the study area, the total land area in developed and open space uses (excluding protected lands) was calculated for 1992 and for 2001. Two land cover measures were calculated based on this information. First, the percent of total land area in open space was calculated as

$$\%\ \text{Open Space in 1992} = \frac{\text{Open Space Area in 1992}}{\text{Open Space Area in 1992} + \text{Developed Area in 1992}}$$

Second, the percent loss in open space was calculated as

$$\%\ \text{Open Space Loss} = \frac{\text{Open Space Area in 1992} - \text{Open Space Area in 2001}}{\text{Open Space Area in 1992}}$$

To convert this to a rate of open space loss per decade, this variable was multiplied by 10/9.   For some counties that experienced very little land cover change, the calculated loss in open space was negative due to differences in interpretation of the satellite images in the two years. Negative values of open space loss were replaced by a value of 0. The average decadal rate of open space loss during the study period was 3.8%.

It was expected that land cover change in heavily urbanized counties would follow different rules than in nonurban counties. Consequently, four counties with less than 60% open space in 1992 were excluded from the analysis (Washington D.C.; Baltimore City, MD; Union County, NJ; and Philadelphia, PA).

In Virginia, several small cities constitute independent counties, surrounded by larger rural counties. To make these observations more similar in character to other counties, small city/counties in Virginia were combined with surrounding counties. The details of which cities were combined with which counties are provided in Appendix Table 2.

Data on Population

The most important driver of land cover change is population. Data on population in each county in 1990 and 2000 was obtained from the U.S. Census. This was used to calculate population density (persons per square kilometer) in 1990 and 2000, and the change in population density over the decade[2].

Estimation of the Statistical Model

The purpose of the statistical model is to explain why some counties lose open space quickly, and others lose it more slowly. Based on our understanding of the processes that drive land cover change, we had the following expectations about the factors that will determine the percent loss of open space in a county:

- Counties that experience a greater increase in population density will lose more open space. New residents need housing, shopping opportunities, and workplaces. As more people move into a county, more land is needed for these purposes.

- Counties that start out with higher population densities will lose more open space. Even if a county does not grow in population, it will tend to lose open space. The existing residents will use up more and more land over time. For example, now shopping centers may be built while old shopping centers go vacant. As another example, average household sizes are decreasing, so that the same population requires more housing units over time. We refer to this process as “spreading out.”

- Counties that start out with more open space may lose that open space at a different rate than counties that start out with little open space. One the one hand, if a county has very little open space at the beginning of a decade, it may take actions to protect its remaining open space. To the extent that occurs, counties that have little open space remaining will tend to lose it at a slower rate. On the other hand, a single development project will have a proportionately higher impact on open space in a county that has little remaining. To the extent that is important, counties with little open space remaining may lose it at a higher rate. The combined effect could be in either direction, and may be nonlinear.

[2] For these calculations, the total land area for each county was determined using the 1992 NLCD maps. Small differences in total land area exist between the 1992 NLCD and the 2001 NLCD, due to differences in interpretation of the satellite images.

Based on these expectations, a mathematical model was developed to explain the percent loss of open space in the study area counties:

$$\begin{aligned} \textit{\% Open Space Loss} \quad = \quad \beta_0 &+ \beta_1 \textit{*(Population Density in 1990)} \\ &+ \beta_2 \textit{*(Change in Population Density 1990 to 2000)} \\ &+ \beta_3 \textit{*(\% Open Space in 1992)} \\ &+ \beta_4 \textit{*(\% Open Space in 1992)}^2 \end{aligned}$$

Two terms are included to model the impact of initial quantity of open space, a linear and a non-linear term.

The parameters (β's) of this model were estimated using data for the 79 counties included in the study area[3]. The estimated values of the parameters are provided in Table 1.

Table 1. Estimated open space loss model

Parameter	Explanatory Variable	Estimated Parameter Value	t-statistic
β_0	Constant	-0.68428	-3.548
β_1	Population density in 1990	0.000191	3.809
β_2	Growth in population density 1990 to 2000	0.000566	2.819
β_3	% Open Space in 1992	1.450658	3.362
β_4	% Open Space in 1992 squared	-0.76148	-3.025
σ	Standard deviation of regression error term	0.02783	11.486

A positive parameter estimate means that a higher level of that factor results in higher rates of open space loss. Each of the estimated parameters is statistically significant at the 1% level, meaning that we have high confidence that the explanatory variables identified do in fact influence open space loss.

The results of the statistical estimation confirm our expectations. Population growth leads to higher rates of open space loss. However, even if population does not change, a higher initial population density leads to higher rates of open space loss, due to the spreading out effect. By comparing the two estimated parameters, we see that each existing resident is responsible for about 1/3 as much open space loss as each new resident. This means that it does not make sense,

[3] Because we assume that open space can be converted to developed use, but not vice versa, % Open Space Loss can only take values between 0 and 1, inclusive. A Tobit model that includes that restriction was used to estimate the parameter values.

as is sometimes done, to calculate a “land used per new resident” statistic equal to the total increase in developed land by the increase in population. Such a statistic would be misleading, because it implies that new residents are the only driver of new development.

When the two terms that measure the initial quantity of open space are combined, the net effect is that counties that have more initial open space lose proportionately more than counties that start with little open space, and the relationship is nonlinear.

Additional models were estimated at a smaller scale of resulution (the township or minor civil division). These models were consistent with the county model regarding the impact of population, population change, and initial quantity of open space. However, these models showed that other factors are also important in determining the rate at which land cover changes. Specifically, the rate of open space loss for a township was higher if 1) the township is located closer to an interstate highway, 2) the township has a lower proportion of land with steep slopes (slopes greater than 10%), 3) the township has a higher proportion of elderly residents, 4) the township has a lower proportion of African American residents, 5) the township has residents with lower median income. While the smaller resolution model allows investigation of more factors than could be included in the county level model, it was not used to project land cover change because projections for those other drivers are not available.

Projecting Land Cover Change

The estimated land cover change model was used to project land cover change for every county in the CARA region. Projections were made by decade from 2000 to 2050. This required knowledge of the initial conditions in 2000 for each county (% land area in open space and population density) and projections of population change for each county.

Data Sources

The population projections used here and shown on the CARA website are provided by NPA Data Services, Inc. (www.npadata.com), a company that develops and maintains databases of historical and projected U.S. population, employment, earnings, and personal income. The projections were developed by NPA Data Services in 1999. More recent projections can be purchased from them.

In deriving county-level population projections of the type used here and shown on the website, NPA Data Services begins with national-level projections for population, employment by industry, earnings by industry, and personal income. These population projections are based on various assumptions about birth and death rates and rates of immigration. The employment, earnings and income projections are based on economic growth models.

NPA Data Services then projects employment at the industry level for each geographic area (county, state, MSA) based on historical differences between that area's employment growth rates at the industry level and the corresponding national-level projected growth rates. Growth rate differences are assumed to narrow over time: each area's employment gradually converges to the national trend in each industry. Area projections are constrained so that they add up to the national projections.

Employment projections, along with area-specific employment/population ratio projections, are used to derive population projections for each area, on the assumption that people move to wherever the jobs are. Area projections are constrained so that they add up to the national projections.

Because the population projections were developed prior to the 2000 Census, they are based on estimates of population, rather than actual counts. This means that some projections were inaccurate already in 2000. To account for this problem, the NPA projections were calibrated according to the difference between the projected population in 2000 and the actual measured population. This difference was added or subtracted to each of the subsequent population projections.

The second piece of information needed for projecting land cover change is the initial quantity of open space in 2000. For each county, the total amount of land area was calculated based on the 1992 NLCD. The amount of land in open space in 2000 was based on one of two sources. For counties where the 2001 NLCD was available, that source was used to measure the percent open space in 2000. For counties where the 2001 NLCD was not available, the open space loss model (described below) was used to calculate the predicted quantity of open space in 2000, which was then used as the basis for future projections. In all cases, protected lands were excluded from the calculations.

Applying the Open Space Loss Model

The land cover change projections were calculated using the estimated parameters shown in Table 1. The specific procedure was as follows. For the period 2000 to 2010, an intermediate statistic called X was calculated using the estimated values for the β's:

$$\begin{aligned} X \quad = \quad [\beta_0 \quad & + \beta_1 \text{*(Population Density in 2000)} \\ & + \beta_2 \text{*(Change in Population Density 2000 to 2010)} \\ & + \beta_3 \text{*(\% Open Space in 2000)} \\ & + \beta_4 \text{*(\% Open Space in 2000)}^2 \;] / \sigma \end{aligned}$$

While this statistic could take on values greater than 1 or less than 0, the actual rate of open space loss cannot. An adjustment was made to restrict the projected open space loss to lie between 0 and 1.

$$\begin{aligned} \text{Projected \% Open Space Loss} = [1-\Phi(1-X)] + [\Phi(1-X) - \Phi(-X)]^* X \\ + \sigma^*[\phi(-X) - \phi(1-X)] \end{aligned}$$

where $\Phi(\bullet)$ is the cumulative density function and $\phi(\bullet)$ is the probability density function for the standard normal distribution. Using this formula, the projected % open space loss will always lie between 0 and 1.

Maps and Tools on the Web Site

This formula was used to project open space loss between 2000 and 2010. That projection was then used as the starting point for another projection of open space loss between 2010 and 2020, based on initial conditions in 2010 and projection population change between 2010 and 2020. This process was repeated for each decade until 2050.

The maps shown on this web site show the projected percent open space in each county by decade and the projected decrease in open space relative to the quantity that existed in 2000. The model is only valid for counties that have at least 60% of their land in open space. When the projected open space in a county falls below 60% of the county, projections are no longer made for that county, and the county is shaded gray.

The land cover projection tool allows the user to explore the impact that population change has on land cover. High and low projections of population are provided for each county, or the user can enter other numbers. Remember that population change is not the only land cover change driver. Initial population can be as important as or more important than population change.

The projections included on this web site are our best predictions of what will happen, based on what we know now and our understanding of historical land cover change. They are not, however, predictions. The actual rate at which land cover will change in a given county will depend on myriad factors including local zoning and planning decisions, local economic conditions, and drivers not included in the model estimated here. Still, the projections are useful for identifying which counties are at highest risk of open space loss, and for evaluating the role that population change has on land cover.

Appendix

Table A1. NLCD categories considered Open Space and Developed Land.

Aggregate Categories	1992 NLCD Categories Class #	Name	2001 NLCD Categories Class #	Name
Open Water	11	Open Water	11	Open Water
Open Space	41	Deciduous Forest	41	Deciduous Forest
	42	Evergreen Forest	42	Evergreen Forest
	43	Mixed Forest	43	Mixed Forest
	51	Shrubland	52	Shrub/Shrub
	61	Orchards/Vineyards/Other	71	Grasslands/ Herbaceous
	71	Grasslands/Herbaceous	81	Pasture/Hay
	81	Pasture/Hay	82	Cultivated Crops
	82	Row Crops	90	Woody Wetlands
	83	Small Grains	95	Emergent Herbaceous Wetlands
	91	Woody Wetlands		
	92	Emergent Herbaceous Wetlands		
Developed	21	Low Intensity Residential	21	Developed, Open Space
	22	High Intensity Residential	22	Developed, Low Intensity
	23	Commercial/Industrial/ Transportation	23	Developed, Medium Intensity
	31	Bare Rock/Sand/Clay	24	Developed, High Intensity
	32	Quarries/Strip Mines/Gravel Pits	31	Barren Land (Rock/Sand/Clay)
	33	Transitional		
	85	Urban/Recreational Grasses		

*For more detail of the NLCD classification criteria visit the following websites:

1992 NLCD – http://landcover.usgs.gov/classes.asp

2001 NLCD – http://www.mrlc.gov/nlcd_definitions.asp

Table A2. Virginia County Summations.

Aggregate County FIPS (used in analysis)	Actual Census FIPS	Name
51901	51003	Albemarle County
	51540	Charlottesville City
51903	51005	Alleghany County
	51560	Clifton Force City
	51580	Covington City
51905	51013	Arlington County
	51510	Alexandria City
51907	51015	Augusta County
	51790	Staunton City
	51820	Waynesboro City
51909	51019	Bedford County
	51515	Bedford City
51911	51031	Cambell County
	51680	Lynchburg City
51913	51035	Carroll County
	51640	Galax City
51937	51550	Chesapeake City
	51710	Norfolk City
	51740	Portsmouth City
51918	51053	Dinwiddie County
	51570	Colonial Heights City
	51730	Petersburg City
51919	51059	Fairfax County
	51600	Fairfax City
	51610	Falls Church City
51921	51069	Frederick County
	51840	Winchester City
51923	51081	Greensville County
	51595	Emporia City
51083	51083	Halifax County
	51083	South Boston City
51927	51087	Henrico County
	51760	Richmond City
51929	51089	Henry County
	51690	Martinsville City
51931	51095	James City County
	51830	Williamsburg City
51933	51121	Montgomery County
	51750	Radford City
51939	51143	Pittsylvania County
	51590	Danville City
51941	51149	Prince George County
	51670	Hopewell City

Table 1. Virginia County Summations (continued)

Aggregate County **Actual Census**

FIPS (used in analysis)	FIPS	Name
51153	51153	Prince William County
	51683	Manassas City
	51685	Manassas Park City
51943	51161	Roanoke County
	51770	Roanoke City
	51775	Salem City
51945	51163	Rockbridge County
	51530	Buena Vista City
	51678	Lexington City
51947	51165	Rockingham County
	51660	Harrisonburg City
51949	51175	Southampton County
	51620	Franklin City
51951	51177	Spotsylvania County
	51630	Fredericksburg City
51953	51191	Washington County
	51520	Bristol City
51955	51195	Wise County
	51720	Norton City
51957	51199	York County
	51650	Hampton City
	51700	Newport News City
	51735	Poquoson City