Ecosystems (1998) 1: 575-594

ECOSYSTEMS

ORGINAL ARTICLE

Land-Use Changes in Southern Appalachian Landscapes: Spatial Analysis and Forecast Evaluation

David N. Wear[1]* and Paul Bolstad[2]

[1]USDA Forest Service, PO Box 12254, Research Triangle Park, North Carolina 27709; and [2]Department of Forest Resources, University of Minnesota, St. Paul, Minnesota 55108, USA.

ABSTRACT

Understanding human disturbance regimes is crucial for developing effective conservation and ecosystem management plans and for targeting ecological research to areas that define scarce ecosystem services. We evaluate and develop a forecasting model for land-use change in the Southern Appalachians. We extend previous efforts by (a) addressing the spatial diffusion of human populations, approximated by building density, (b) examining a long time period (40 years, which is epochal in economic terms), and (c) explicitly testing the forecasting power of the models. The resulting model, defined by linking a negative binomial regression model of building density with a logit model of land cover, was fit using spatially referenced data from four study sites in the Southern Appalachians. All fitted equations were significant, and coefficient estimates indicated that topographic features as well as location significantly shape population diffusion and land use across these landscapes. This is especially evident in the study sites that have experienced development pressure over the last 40 years. Model estimates also indicate significant spatial autocorrelation in land-use observations. Forecast performance of the models was evaluated by using a separate validation data set for each study area. Depending on the land-use classification scheme, the models correctly predicted between 68% and 89% of observed land uses. Tests based on information theory reject the hypothesis that the models have no explanatory power, and measures of entropy and information gain indicate that the estimated models explain between 47% and 66% of uncertainty regarding land-use classification. Overall, these results indicate that modeling land-cover change alone may not be useful over the long run, because changing land cover reflects the outcomes of more than one human process (for example, agricultural decline and population growth). Here, additional information was gained by addressing the spatial spread of human populations. Furthermore, coarse-scale measures of the human drivers of landscape change (for example, population growth measured at the county level) appear to be poor predictors of changes realized at finer scales. Simulations demonstrate how this type of approach might be used to target scarce resources for conservation and research efforts into ecosystem effects.

Key words: landscape modeling; land use; information theory; forecasting; spatial analysis; Southern Appalachians.

INTRODUCTION

Among the most challenging problems in the study of ecosystems is understanding how people, acting

Received 13 March 1998; accepted 30 September 1998
*Corresponding author; r-mail: dwear/srs_rtp@fs.fed.us

in response to various social and economic factors, define patterns of land and resource use (Lee and others 1992). Insights into human drivers of landscape change are needed to understand better how and where human pressures are most likely to lead to detrimental effects on the structure and function of ecosystems. Understanding these interactions between human activities and ecological consequences is especially important in areas of the world that are experiencing rapid change, where the cumulative impacts of development may be realized too late to trigger mitigation measures. In such places, models that forecast land-use change could provide a way to anticipate ecological problems before they are actually observed on the landscape. Projecting land and resource uses is also a necessary first step in developing an effective ecosystem management strategy because it allows managers to define critical elements of landscapes where ecological values and probability of change are both high (Lubchenco and others 1991; Wear and others 1996). This report examines how people have shaped landscapes in the southern Appalachian Highlands with an emphasis on developing and evaluating tools for predicting land-use changes.

Changing land uses can also affect values defined by aesthetic and other environmental qualities of a region. In the Southern Appalachian Highlands, aesthetics, climate, and access to recreation, have fueled substantial population growth over the last 20 years (SAMAB 1996). Resulting changes in land uses reflect shifts in relative land values away from agricultural and toward residential uses, but may also represent long-run costs in the form of reduced water quality and scenic values (Gottfried and others 1996). Development of lands to supply the market for residences and associated commercial uses could therefore reduce the very qualities that have drawn people to this formerly remote region. Spatial models that forecast land use are needed to help planners evaluate the long-run effects of development patterns on the structure of landscapes and the values derived from them.

Economic models of land use link individual choices in an economy to the evolution of landscapes. This body of work is built on the premise that land generally is employed in its highest-value use, that these values are influenced by various features of the land, and, therefore, that patterns of land use should be organized by these land features. The Von Thiinen model of rural land use [for example, see Samuelson (1983)] and various models of urban development around central places [see Katzman (1974), for example] are built upon the premise that the location of land defines its value for various uses. One track of research has focused on the formation of land prices and rents. Theoretical studies [for example, see Capozza and Helsley (1989)] and empirical studies [for instance, see Palmquist and Danielson (1989) and Geoghehan and Bockstael 1997] have shown how land prices are affected by location and other factors. Another track of research has focused on estimating land-use shares as functions of land prices or rents [see Alig (1986) and Hardie and Parks (1997), for example]. This body of work corroborates and extends the Von Thiinen approach and highlights how economic factors and physical landscape features hold important influence over land-use allocations. However, because these models have been applied to spatially broad units (for example, counties or county groupings), they may not provide direct insights into the fine-scale ecological consequences of land-use changes.

Another recent track of land-use research has applied this same general paradigm to spatially explicit analysis of land-use choices. Spatially explicit land-use models that relate location and other features to the probability of land uses have been applied in Belize (Chomitz and Gray 1995), southern Mexico (Nelson and Hellerstein 1997), and Rondônia, Brazil (Dale and others 1993). All three studies, based on single "snapshots" of land use at fine scales (1- to 100-ha units), demonstrate how road construction and access influence patterns of land use in developing areas. Similarly, Turner and colleagues (1996) estimated models of land-cover change for the Southern Appalachian Highlands and Olympic Peninsula regions of the United States for three 5-year periods, finding that location and topography as well as land ownership significantly influence land-cover dynamics. In the Cascade Range of Oregon, Spies and colleagues (1994) identified distinct land-cover changes on lands of different landowner groups. These studies show that land use at various stages of development is significantly influenced by site-specific and institutional factors and, accordingly, that land-use patterns are largely determined by the physical structure of landscapes and socioeconomic conditions of a region. Furthermore, because these spatially explicit models can be applied to specific places and development scenarios at fairly fine resolution (for example, a hectare or less), they provide a means to link human drivers of land-use change to ecological impacts at landscape scales.

Spatially explicit land-use models have, however, been limited in their ability to reflect certain economic processes. One limitation has been the resolution of land-use definitions. For example, previous studies of landscape change in the Southern Appalachians (Wear and Flamm 1993; Turner and others

1996) have focused on changes in land cover interpreted from satellite images. Land cover alone, however, provides only a first approximation of how land is actually utilized and may provide only limited insights into the ecological consequences of human use patterns. For example, forest cover would be observed both for remote areas with wilderness attributes where a human presence is largely absent, and for a low-density suburban area where human impacts are substantial. Additionally, previous research has been limited to either a single point in time (Chomitz and Gray 1995; Nelson and Hellerstein 1997) or to a fairly short period [for example, 5 years in Turner and colleagues (1996)]. A longer time frame may be needed to fully represent complete adjustments to changes in land markets and other social forces.

Another area that has not been fully explored is the validation of spatially explicit land-use models. Although these models can be used to forecast where land uses may change in the future, evaluations of their forecasting power have been limited (Costanza and others 1990). Most validation efforts [for example, see Turner and others (1989), Nelson and Hellerstein (1997), and Chomitz and Gray (1995)] judge predictive power based on the ability of models to classify historical observations correctly based on predicted probabilities. But, as Chomitz and Gray suggest, these measures may fail to reflect important information regarding low-probability events. Furthermore, these approaches to model validation have not been used to develop statistical tests of forecasting power.

We develop a spatially explicit analysis to examine land-use change over a 40-year period in the Southern Appalachians in pursuit of two objectives: (a) to test hypotheses regarding the effects of various physical and human factors in determining where land uses occur, and (b) to construct and evaluate a model for forecasting land uses. We attempt to address the limitations of previous research in this area by (a) defining land use based on human occupancy, (b) examining changes over a long time period, and (c) specifically addressing the ability of these models to forecast landscape structure. We defined land uses based on the intersection of raw land-cover categories and a measure of the intensity of human presence for each of four study sites in the Southern Appalachians. Land cover was derived from satellite images and aerial photos for two points in time: 1950 and 1990. While an ideal measure of human presence would be the local population density defined by the decennial census in these years, fine-scale spatial population data were available only for 1990. In lieu of population data, we assumed that human influence could be approximated by the number of buildings within the neighborhood of a site. We defined aggregate land-use classes by overlaying land cover and building density. To test and forecast the effects of various site and locational factors on these land-use classes, we applied a two-stage regression approach. The first stage estimated the future building density of individual sites as a function of existing building density and other factors that should influence the valuation of land for different uses. These factors include site attributes that influence operability for agriculture and forestry (for example, slope and distance to the closest road) as well as factors that influence the desirability of a site for residential uses (for instance, distance to the local market center and elevation). This stage therefore describes the spread of human populations, approximated by building density, across a landscape. The second stage of the analysis estimated the probabilities of observing land-cover classes as functions of local building density, predicted from the first stage, and other site features. For both stages, model estimates allow testing hypotheses regarding the effects of various site and locational features on the spatial diffusion of human populations and on the location of land-cover types.

The forecasting power of the estimated models was then evaluated by using validation data sets (that is, observations that were not used for estimating the models). The two-stage model was applied to each observation to calculate predicted probabilities for land-use classes in 1990 based on conditions observed in 1950. Comparing predicted probabilities with those derived from a null model, defined by average probabilities, allows construction of information-theoretic statistics for testing the significance and accuracy of the historical forecasts. In addition, we calculated the information gain over the null model.

This type of forecasting model can focus planning and research efforts within a region. To demonstrate these applications and focus our future endeavors, we present 40-year forecasts of land use for our study watersheds and highlight portions that are the most likely to change. We conclude by discussing issues regarding scenario design and the limitations of the model as well as its appropriate application to the assessment of changing ecological and economic conditions within the region.

STUDY AREA

Our study area is in the Blue Ridge province of the Southern Appalachian Highlands and includes all of the mountainous portions of western North Carolina, northern Georgia, and southeastern Virginia

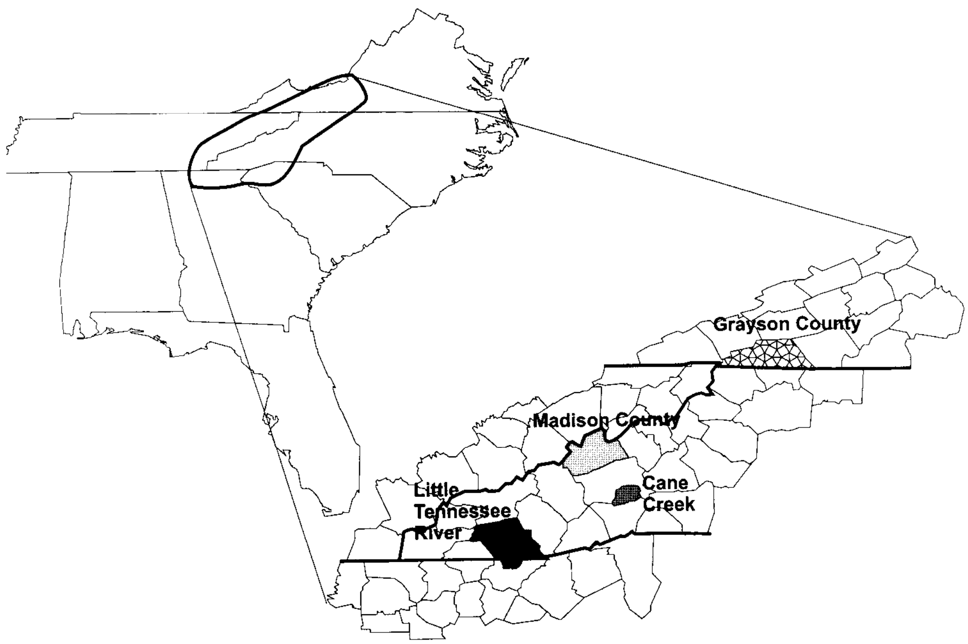

Figure 1. Location of four study sites in the Southern Appalachian Highlands.

(Figure 1) Within this region, we chose to examine land use in four separate study sites that exhibit a broad range of land-use pressures. One is the Little Tennessee River Basin in southwestern North Carolina and northern Georgia, centered at Franklin, North Carolina, and including the Coweeta Hydrologic Laboratory. Further north, we examine two areas within the French Broad River Basin, the Cane Creek drainage in northern Henderson County, and all of Madison County, North Carolina. We also examine Grayson County, Virginia, which borders North Carolina. All four sites were essentially rural in 1950 but have since experienced different levels of development.

The Southern Appalachian region as a whole experienced population growth between 1950 and 1990. However, growth patterns have not been constant across all counties in the region and, as shown in Figure 2, the counties in the vicinity of our study areas show different patterns of growth. Henderson County, which contains the Cane Creek study area and the city of Hendersonville, experienced a 124% increase in population between 1950 and 1990 (US Census Bureau). In contrast, the population of Madison County was 17.4% less than it was in 1950. The population count in Grayson County was relatively stable over this period (roughly 8% more people in 1990 than in 1950), and growth was moderately strong in Macon County-which represents a majority of the land in the Little Tennessee River Basin-with an increase of 45.3%.

Another important force behind landscape change in the Southern Appalachians has been a decline in agriculture. Not surprisingly, as this formerly isolated area has become better integrated into broad regional and national markets for agricultural products, local demand for farmland has declined. This change is clearly reflected in the counties containing our study sites (Figure 3). All of the referenced counties show declines in land in farms between 1978 and 1992 (respective US Censuses of Agriculture). Figure 3 also demonstrates a broad range of agricultural presence in the study sites, from only 7% of land area in farms in Macon County to 47% in Grayson County.

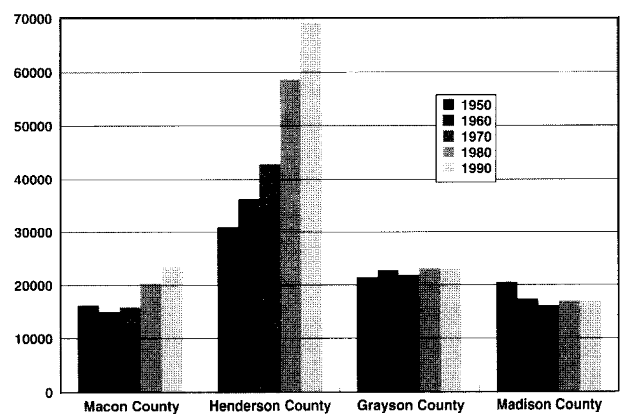

Figure 2. Total population for counties coincident with the four study areas 1950–90. (Macon County, NC, constitutes a majority of the Little Tennessee River Basin study area and a majority of the Cane Creek study area is in Henderson County.) Values are from the decennial censuses of population (US Census Bureau).

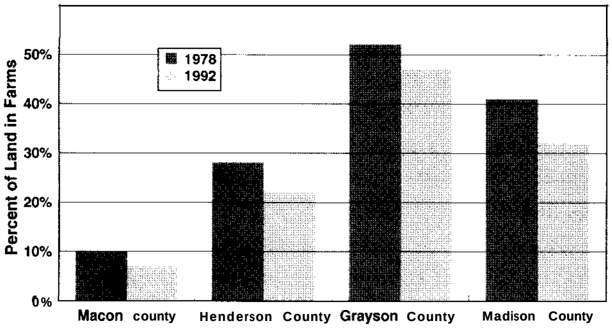

Figure 3. Share of land in farms in 1978 and 1992 for counties coincident with the four study areas.

METHODS

Data-Base Development

To evaluate overall changes and estimate land-use models, spatial data were developed for two points in time-1950 and 1990—and compiled in a Geographic Information System (GIS). All data were developed into the Universal Transverse Mercator (UTM) coordinate system, zone 17. Original data sources (for example, aerial photographs, maps, satellite and imagery) were obtained as close as possible to the target date. In most cases, however, simultaneous coverage was not available for all required data layers, although most data were collected within 4 years of the target years.

Early 1950s' land-cover data were derived from 1:20,000 panchromatic aerial photographs in a 9-inch format taken by the US Soil Conservation Service. Most photographs were leaf-on, early springtime collections although, for some areas, photos were taken during leaf-off condition. Land use was manually interpreted into forest, nonforest, abandoned old-field, and shrub-early successional classes within the neat area of each image, and polygon boundaries were digitized using a high-precision (0.001-inch) coordinate digitizer. Ground control points were photo-identified and marked, and ground coordinates were determined either from field global positioning system readings or from serial transformations from US Geological Survey (USGS) 1:24,000-scale quadrangle maps. Data were terrain and tilt corrected using a single photo resection (Wolf 1983), based on 30-m digital elevation models (DEMs). Single photo images were then combined in a mosaic to create land-cover maps for each study area.

Land-cover data for the 1990s were derived from maximum-likelihood classifications of Landsat thematic mapper (TM) data. To provide complete coverage of the study sites, parts of seven TM scenes were required. Midsummer data collected in the early 1990s were geocorrected and georeferenced. Training data were collected for known land-cover types, based both on field visits and on air-photo interpretation. Classification was aggregated into forest and nonforest classes, and accuracy of both land-cover classifications was verified to be over 95 % (Lillesand and Kiefer 1994).

Road and building location data were determined from 1:24,000-scale USGS maps, and updates were produced in both time periods. Clean, unfolded paper maps were affixed to a coordinate digitizer and georeferenced using the eight corner and central graticule marks. Road centerlines were digitized for all paved and unpaved roads at the time of map compilation. Road type and capacity were attributed to each digitized segment. Building locations and epoch were recorded based on mapped information. Market centers were identified, and point locations were digitized. All data were converted to 1-ha-resolution raster format. Slope and aspect were derived from the DEM data, using a third-order finite difference algorithm (Bernhardsen 1992). Minimum road distance and travel times to nearest market centers were estimated for each point in each study area, based on an off-road average 3-km/h travel speed over steep terrain, S-km/h on off-road flat terrain, and design-estimated speeds on roadways. Neighborhood building densities were determined from a 9-ha moving sum operator (Bernhardsen 1992).

Land-use classes were defined by overlaying land cover with neighborhood building densities in the two periods. We used two classification schemes. One identified three broad land-use classes: *nonforest, forest with no buildings,* and *forest with buildings.* A fine-resolution scheme was also applied that split the forest-with-buildings class into four subcategories: l-2, 3-5, 6-8, and more than 8 buildings per hectare area for each cell, to identify a gradation of intensities of land use. Both schemes were used to evaluate the forecast performance of estimated models.

All referenced variables were compiled as data layers in a grid-based GIS. To define observations for model estimation and validation, a sample of 5000 landscape cells in private ownership was defined by random draw without replacement for each study site. A total of 4000 cells were used to estimate the models. The remaining 1000 cells for each study site were reserved as validation data sets to test the forecasting performance of the estimated models.

Models of Land-Use Changes

Our modeling approach viewed land-use choices as being influenced by features of each site and the spatial contagion of development. The influence of location has been the emphasis of spatial land-use models descended from Von Thünen's analysis of rural land-use specialization [for example, see Samuelson (1983)] and *central business district* models that describe the spatial development of urban areas (Katzman 1983). These models focus on the influence of distance from the site to markets for agricultural products or to the location of essential services on relative land values and land-use choices. We adopt the same general approach but also allow land values and uses to be similarly influenced by several other variables. We also allow for spatial contagion of land-use choices that would define spatial autocorrelation in an uncorrected model.

Historical land-use changes were examined using a two-stage regression approach. The first stage was to estimate the building count within the 9-ha neighborhood of a landscape cell in 1990 (B_{90}) as a function of conditions in 1950. We posited that the value of land for development (and therefore the building count) in 1990 would be positively influenced by the building count in 1950 (B_{50}) and negatively influenced by the cost of access to the site. Access was represented by two variables: the distance from the site to the closest road (DtR_{50}) and shortest travel time to and along the road network to the local market center (DtM_{50}). In addition, because of its influence on costs of access and construction, we anticipated that building count would be negatively influenced by the slope. We also included the elevation (elev) of a site, positing that views afforded by higher elevations would increase the value for development. The general form of the regression equation is:

$$B_{90} = f\,(B_{50}, DtM_{50}, DtR_{50}, \text{slope}, \text{elev}) \qquad (1)$$

The second stage of the analysis estimated the probability that land cover in a landscape cell (LC) would be forest (F) or nonforest (NF) in 1990 as a function of the building count as well as other explanatory variables. We posited that the probability of nonforest cover was positively related to the building density examined in the first stages of the analysis. We also posited that the probability of nonforest cover would be negatively related to the elevation of a site due to increased exposure. Slope should also be negatively related to nonforest cover due to its influence on both the cost of access and on the operating costs for farming. The general equation form is

$$\begin{aligned} pr(LC_{90} = NF) &= P_{NF} \\ &= g\,(B_{90}, NF_{50}, \text{slope}, \text{elev}) \qquad (2) \\ pr(LC_{90} = F) &= 1 - P_{NF} \end{aligned}$$

where NF_{50} is the measure of nonforest cover in adjacent areas and other variables are as previously defined. Several statistical issues needed to be addressed before estimating these equations:

1. *Count data.* Because the dependent variable in Eq. (1), B_{90}, is not a continuous variable but is measured as a count, we estimated Eq. (1) using a negative binomial regression model, a general form of the Poisson regression model. (In particular, it relaxes the Poisson's assumption that the mean of the dependent variable is equal to its variance.) In addition, because few observations have building counts greater than 9, we censored B_{90} by lumping all values greater than 8 into a single category. This censoring was accounted for in the estimation of the regression model and the calculation of predicted values [see Greene (1995: 548–9)].

Assuming that B_{90} is distributed as Poisson defines the following negative binomial model:

$$\begin{aligned} pr(B_{90} = y_i|\gamma) &= \frac{e^{-\lambda_i}\lambda_i^{y_i}}{y_i!} \\ y_i &= 0,1,2,\ldots; \quad i = 1,2,\ldots,n \qquad (3) \\ \text{where} \quad \ln(\lambda_i) &= \ln(\hat{\lambda}_i) + \gamma = \beta'\mathbf{x}_i + \gamma \end{aligned}$$

where γ is a random variable such that $\exp(\gamma)$ has a gamma distribution with mean 1 and variance α, x_i is the vector of independent variables, and β is a vector of coefficients to be estimated [see, generally, Cameron and Trivedi (1986) for a complete development of negative binomial models]. The Poisson regression model would be defined if the random variable γ were not included. Including γ allows for overdispersion in the model [that is, for $\mathrm{var}(B_{90}) > E(B_{90})$] but also requires estimation of the additional parameter α. Equation (3) was estimated using maximum-likelihood estimation in the software package LIMDEP (Greene 1995).

The overall significance of the model was tested using a likelihood ratio test where the null hypothesis holds all slope coefficients equal to zero. The significance of individual variables in predicting building count was tested using a z statistic for each marginal effects coefficient derived from the estimated negative binomial model. The marginal effects coefficients were calculated as:

$$\partial\, E[B_{90}|\mathbf{x}_i]/\partial x_i = \hat{\lambda}_i\beta \qquad (4)$$

with the values of x set at their mean values.

2. *Simultaneity.* Another potential statistical issue defined by Eq. (1) and (2) is the inclusion of the endogenous variable in Eq. (1) (B_{90}) on the right-hand side of Eq. (2). This issue was addressed by first estimating Eq. (3) and then using predicted values of the building count as the explanatory variable in Eq. (2)-that is, predicted values define an instru-

mental variable for B_{90}. The expected value of B_{90} was defined as:

$$EB_{90} = \sum_{i=0}^{c-1} P_i i + \left(1 - \sum_{i=0}^{c-1} P_i\right) c \qquad (5)$$

where the probabilities are those predicted by the fitted negative binomial model and c is the upper limit on building counts (in this case, c = 9). This sequential approach to estimation corrects for simultaneous equation bias and is essential for the forecasting model. By using the expected value of $B_{90}(EB_{90})$ in Eq. (2), land use in 1990 (t) is predicted strictly as a function of conditions in 1950 ($t-1$).

3. *Nominal land cover.* Because land cover is defined by two discrete classes (forest and nonforest), we applied the binomial logit regression model to Eq. (2):

$$pr(LC_{90} = NF) = P_{NF} = \frac{1}{1 + e^{-\delta' z_i}} \qquad (6)$$

where Z, is the vector of variables on the right-hand side of Eq. (2) and δ is the vector of corresponding estimated coefficients. These equations were estimated using maximum-likelihood methods in LIMDEP, and overall significance was tested using a likelihood ratio test. The test statistic for the null hypothesis that all coefficients are equal to zero is distributed as a chi square with degrees of freedom equal to the number of explanatory variables. The significance of individual variables was tested using *z* statistics for the marginal effects coefficients [see Greene (1995: 432)].

4. *Spatial autocorrelation* was also a statistical concern in modeling the probability of land-cover classes. For example, the probability that an area will be in nonforest cover is likely to be, ceteris paribus, higher in the vicinity of other areas that have nonforest cover; that is, we hypothesize that there could be a clumping of land cover, not fully explained by the location or condition of individual landscape cells. Dubin (1995) has shown that excluding terms that represent spatial interactions can lead to inconsistency in estimates of the logit model. To account for spatial autocorrelation here, we defined a spatial weighting matrix of nonforest cover for each cell in the landscape:

$$NF_{50_{ij}} = \frac{1}{9} \sum_{k=i-1}^{i+1} \sum_{l=j-1}^{l=j+1} L_{l,k}$$

$$\text{where } L_{l,k} = 1 \text{ if } LC_{50,l,k} = \text{nonforest} \qquad (7)$$

$$L_{l,k} = 0 \text{ otherwise}$$

where *i* and j are used to identify the row and column location of the referenced landscape cell. NF_{50} is therefore the portion of landscape cells that were in nonforest cover in 1950 within the nine-cell neighborhood (inclusive) of the referenced cell. The neighborhood is made up of horizontally, vertically, and diagonally adjacent cells. This variable was included in the Z vector in Eq. (6), and a significance test of its estimated coefficient was used as a test for spatial autocorrelation.

Forecasting Performance

In addition to testing for the overall significance of estimated equations and the significance of marginal effects in both the building count and the land-cover models, we also tested the ability of the estimated equations to predict land uses in 1990 based on conditions observed for 1950. These tests were conducted on validation data sets made up of 1000 observations for each of the study sites. Validation data were used to test whether the estimated models improved on “naive” or null models of land-use change defined by average probabilities. Because a primary objective was to construct robust projections of land use, tests of forecasting power provided crucial evaluations of model performance.

Evaluating the performance of discrete choice models such as the logit and negative binomial is challenging because the models generate conditional probabilities while observed choices are discrete events. Because the dependent variable is not continuous, standard measures of forecast or simulation performance that are based on direct comparisons of predicted versus observed outcomes cannot be used. Instead, evaluation of model predictions must focus on the ability of the estimated model to provide predictions that are better than the best available naive or null model. Statistical tests address whether the estimated models contribute additional information about the evaluated system.

We evaluated the ability of the two estimated Eqs. (3) and (6) to discriminate among the six aggregate land-use classes: (a) nonforest and forest with (b) no buildings, (c) 1-2 buildings, (d) 3-5 buildings, (e) 6-8 buildings, and (f) more than 8 buildings. For comparison, we also evaluated predictive ability for two coarse land-classification schemes: (a) broad building classes: forest with buildings, forest without buildings, and nonforest; and (b) simple land cover: forest versus nonforest. To construct predictions for each observation in the validation data set, we defined the $\mathbf{x}_i$ and $\mathbf{z}_i$ variables and calculated the conditional probabilities of all possible classes by using estimates of Eqs. (3) and (6). Three statistics based on information theory (Hauser 1978; Judge and others 1985: 777) were then used to evaluate model performance [see Colwell (1974) and Turner and others (1989) for applications of similar information-theory concepts to the evaluation of the aggregate performance of ecological models]. These statistics are derived from the following three basic information concepts [see Hauser (1978)]:

1. The prior entropy of the system, defined as

$$H(A) = -\sum_{j=1} p(a_j) \ln [p(a_j)] \tag{8}$$

where $p(a_j)$ is the probability of land-use class j defined by a prior or null model and A is the set of J potential land-use classes. H(A) measures the total uncertainty inherent in the null model or the maximum uncertainty that could be explained by the estimated model. We define the null model by setting the $p(a_j)$ terms to the proportion of observations in land-use class a_j. Entropy is at a maximum when the probability of each land-use class is 1/J (that is, equal probabilities of all land-use class) and approaches zero as the probability of one land-use class approaches unity (that is, complete certainty about the outcome).

2. The additional information contained in the estimated model is defined as

$$I(A;X,Z) = \frac{1}{m} \sum_{i=1}^{m} \sum_{j=1}^{J} \delta_{ij} \ln \left[\frac{p(a_j|\mathbf{x_i};\mathbf{z_i})}{p(a_j)}\right] \tag{9}$$

where δ_{ij} = 1 if state j is observed at cell i (δ_{ij} = 1 otherwise), x_i and z_i are the independent variables describing cell I, and $p(a_j\, \mathbf{x}_i; \mathbf{z}_i)$ is the conditional probability of land-use class j defined by the estimated models [Eqs. (3) and (6)], and m is the number of cells in the validation sample. The estimated models provide additional information about the land-use classification of a cell [that is, adds to I(o)] if the conditional probability of the observed land-use class is greater than the probability defined by the null model.

3. The expected information provided by the estimated model is defined as:

$$EI(A;X,Z) = \frac{1}{m} \sum_{i=1}^{m} \sum_{j=1}^{J} p(a_j|\mathbf{x_i};\mathbf{z_i}) \ln \left[\frac{p(a_j|\mathbf{x_i};\mathbf{z_i})}{p(a_j)}\right] \tag{10}$$

As with I(.), expected additional information is low if the conditional probabilities are always close to the prior probabilities. Based on these concepts, Hauser (1978) defines three tests for a probabilistic system:

1. Usefulness *test:* Define $U^2 = I(A;X,Z)/H(A)$. U^2 is the proportion of entropy (uncertainty) explained by the model defining a pseudo-r^2 that ranges from 0 (no additional information) to 1 (complete explanation).

2. *Accuracy test:* Hauser shows I(A;X,Z) to be normally distributed with mean EI(A;X,Z) and a variance V(A;X,Z) under the null hypothesis that the estimated model is true. If I(A;X,Z) is outside the confidence interval for EI(A;X,Z), then we reject that the estimated model fully explains the land uses observed in 1990.

3. *Significance test:* The overall significance of the estimated system can be tested by comparison to the null model. The null hypothesis is defined by the prior probabilities. The log-likelihood ratio is defined as L = 2nI(A;X,Z) and is distributed as a chi square with degrees of freedom equal to 11, the number of estimated parameters in Eqs. (3) and (6) [see Kullback (1959: 98)].

In calculating all of these statistics, the prior model was defined as the frequencies of land-use classes calculated from estimation data sets for the study areas.

Results

Descriptive Statistics

Little Tennessee River Basin (LTRB). Between 1950 and 1990, the share of the LTRB in forest cover increased by 12.02% in spite of substantial population growth (Table 2); 26% of the LTRB experienced change in major land-use categories [forest-with buildings (FWB), forest-no buildings (FNB), and nonforest]. The major shift was out of nonforest and FNB to FWB (+ 10.12% and +9.02%, respectively; Table 2); 4.42% of the area shifted from nonforest to FNB and a total of 2.48% shifted from forest to nonforest. As a result and consistent with a developing landscape, the area of forested land without buildings declined by 6% and the area of forests with buildings increased by 18.2 %.

Changes in the means of the independent variables also reflect a developing landscape (Table 3). There was a near tripling of the average building density between 1950 and 1990 (+183.64%). The LTRB lost land in the 0 and 1 building (bldg) per 9-ha land classes and gained area in all other building-density classes (Figure 4). Gains were substantial in the 2, 3, 4, and 5 bldg/9-ha classes, suggesting low-density residential development. There was also a gain in the ≥9 bldg/9-ha class, indicating some expansion in high-density areas. Relative location of cells also changed over this period. The average travel time to the closest market center (MC) declined slightly (-0.66%) while the average distance to roads declined by 15.21%.

Table 1. Land-Use and Land-Cover Characteristics, Both Public and Private Land for the Four Study Areas

	Cane Creek		Grayson County		LTRB		Madison County	
	1950	1990	1950	1990	1950	1990	1950	1990
% Forest	59.7	65.3	41.6	65.6	81.4	89.7	62.5	82.1
% Agriculture	36.1	28.6	50.2	30.3	15.3	7.5	28.1	12.3
% Urban	2.9	5.7	1.4	1.6	1.0	2.0	0.8	1.0
Mean distance to road (m)	253	196	337	319	557	425	433	381
Buildings/ha	0.14	0.25	0.16	0.17	0.11	0.22	0.04	0.09
Mean patch size (ha)	39.8	41.6	35.2	60.0	54.8	79.1	31.8	37.7
Area-weighted median patch size (km^2)	0.32	0.46	0.74	0.56	11.56	12.91	1.27	3.61
% Core area	52.8	59.1	49.2	63.3	73.0	82.1	49.8	70.0
Contagion	55.3	57.3	47.8	62.9	69.3	78.7	49.9	67.0

LTRB, Little Tennessee River Basin.

Table 2. Land Uses for 1950 and 1990 and Changes in Private Land Uses Between 1950 and 1990 Based on a Random Sample of 1 -ha Landscape Cells (n = 5000 for Each Study Area)

	LTRB			Cane Creek			Grayson County			Madison County		
	1950	1990	% Change[a]	1950	1990	% Change	1950	1990	% Change	1950	1990	% Change
Land-use class (% area)												
Forest												
0 Buildings	58.42	52.36	-6.06	46.92	41.62	-5.30	33.62	49.42	15.8	51.90	55.26	3.36
1-2 Buildings	11.88	20.22	8.34	8.68	12.80	4.12	5.76	11.52	5.76	11.42	20.76	9.34
3-5 Buildings	1.20	8.60	7.40	3.70	6.88	3.18	0.90	2.66	1.76	0.42	5.74	5.32
6-8 Buildings	0.14	1.86	1.72	0.70	1.90	1.20	0.04	0.38	0.34	0.02	0.76	0.74
>8 Buildings	0.02	0.64	0.62	0.18	1.42	1.24	0.02	0.06	0.04	0.02	0.24	0.22
Nonforest	28.18	16.06	-12.12	39.72	35.26	-4.46	58.92	35.18	-23.74	35.56	16.36	-19.20
Major land-use changes (% area)												
Forest (w/no bldg)-nonforest	—		1.40	—		3.02	—		2.14	—		1.20
Nonforest-forest (w/o bldg)			4.42	—		4.56	—		18.32	—		9.60
Nonforest-forest (w/bldg)	—		10.12	—		7.42	—	—	8.62	—	—	11.52
Forest (w/o bldg)-forest (w/bldg)			9.02	—		6.82	—		0.44	—		4.88
Forest (w/bldg)-nonforest	—		1.08	—		4.50	—	—	1.14	—	—	0.78
Forest (w/bldg)-forest (w/o bldg)			0.00	—		0.00	—		0.00	—	—	0.00
Total changes			26.04	—	—	26.32	—		30.66	—		27.98

[a]% Change indicates change in the percentage of the referenced landscape in these categories

Taken together, these changes indicate a filling in of the road network rather than an expansion of the network at the remote margin. Among study areas, the LTRB had the highest average slope (14.44”) and the greatest variability in slope (standard deviation, 8.96°), suggesting a fairly rugged landscape and explaining the small share of land dedicated to agriculture (Figure 3).

Cane Creek. The Cane Creek area experienced patterns of change similar to those observed for the LTRB: 26.32% of land shifted among the major land-use classes, and change was dominated by shifts to the FWB class (+14.42%); 3.02% and 4.5% of the landscape shifted from FNB and FWB to nonforest, respectively. Total forested land increased by 4.46%, but there was a 9.76% increase in the

Table 3. Descriptive Statistics for Private Land in the Four Study Sites Based on a Random Sample of l-ha Landscape Cells (n = 5000 for Each Study Site)

	LTRB			Cane Creek			Grayson County			Madison County		
	1950	1990	% Change"	1950	1990	% Change	1950	1990	% Change	1950	1990	% Change
Building density (bldg/9 ha)												
Mean	0.483	1.370	183.64	1.266	2.666	110.58	0.717	0.791	10.32	0.445	1.030	131.46
Range	0–19	0–33		O-36	O-52		O-32	0–33		O-23	O-30	
SD	1.191	2.608		2.719	4.833		1.657	1.839		1.139	2.099	
Travel time to market center (min)												
Mean	26.090	25.918	−0.66	31.401	30.426	-3.10	3x.445	35.894	-6.64	49.298	48.736	−1.14
Range	I-127	I-136		1 o-97	IO-71		3–95	3-99		18–115	IS-105	
SD	14.868	14.729		12.507	10.600		14.693	15414		15.806	15.3X6	
Distance to road (m)												
Mean	245.890	208.484	-15.21	241.818	184.972	−23.51	269.723	266.309	-1.27	323.123	293.629	−9.13
Range	O-2329	I-1831		O-l 603	0–1108		O-1930	O-1930		0–2264	O-2264	
SD	254.007	212.291		256.879	191.055		266.892	266.596		331.913	307.4X6	
Elevation (m)												
Mean	819.581	—		788.732			878.819	—		801.700	—	
Range	565-1493	—		627-l 326			646-1495	—		378-1462	—	
SD	184.738	—		141.240			133.632	—		197.586	—	
Slope (degrees)												
Mean	14.442	—		8.530			7.719	—		12.542	—	
Range	O-87	—		O-68			0–31	—		O-66	—	
SD	8.961	—		6.X32			5.108	—		6.991	—	

% Change indicates change in the percentage of the referenced landscape in the categories. SD, standard deviation.

area of forests with buildings. As a result, the area of forests without buildings declined by 5.3%.

The average building density in the Cane Creek area increased by 110.58% between 1950 and 1990. The shifts in building-density classes (Figure 4) indicate more of an expansion in high-density areas than observed for other study sites (that is, the category with more than 8 buildings saw the greatest gains between 1950 and 1990). As in the LTRB, average travel time to MC and average distance to roads declined (-3.10% and −23.51%, respectively), suggesting a substantial filling in of the road network and some expansion in its extent.

Grayson ***County.*** Unlike the LTRB and Cane Creek study areas, Grayson County had an essentially stable population between 1950 and 1990. A different pattern of change among major land-use classes resulted. Among the study areas, Grayson County experienced the greatest total land-use change, with 30.66% of the area changing major classes: 18.32% of the area shifted from nonforest to FNB, reflecting a substantial decline in farmland in the county; in addition, 8.62% of the area shifted from nonforest to forest with buildings. As a result, the amount of forest area increased by 23.74%. In contrast to all other study areas, the area of forest without buildings also increased between 1950 and 1990 (+ **15.8%).** There was also a 7.94% increase in the area of forest with buildings.

Grayson County was also unique in experiencing only a small change in the average building density (+10.32% while all other areas at least doubled). Figure 4 shows only about 1% of the area changing building classes between 1950 and 1990. Change in location variables were also distinct. Travel time to the MC decreased the most among study areas (-6.64%) while distance to road decreased the least (-1.27%). This suggests that the road network expanded outward at the remote margin but did not fill in substantially. The average slope was only 7.72%, and the range of slopes observed for our observations was less than half that observed in the other study areas. According to these measures, Grayson County had the most gentle terrain of the four study sites.

Madison County. Madison County had 17% fewer people in 1990 than it did in 1950, and land use shifted toward forest cover (+ 19.2 %). Like Grayson

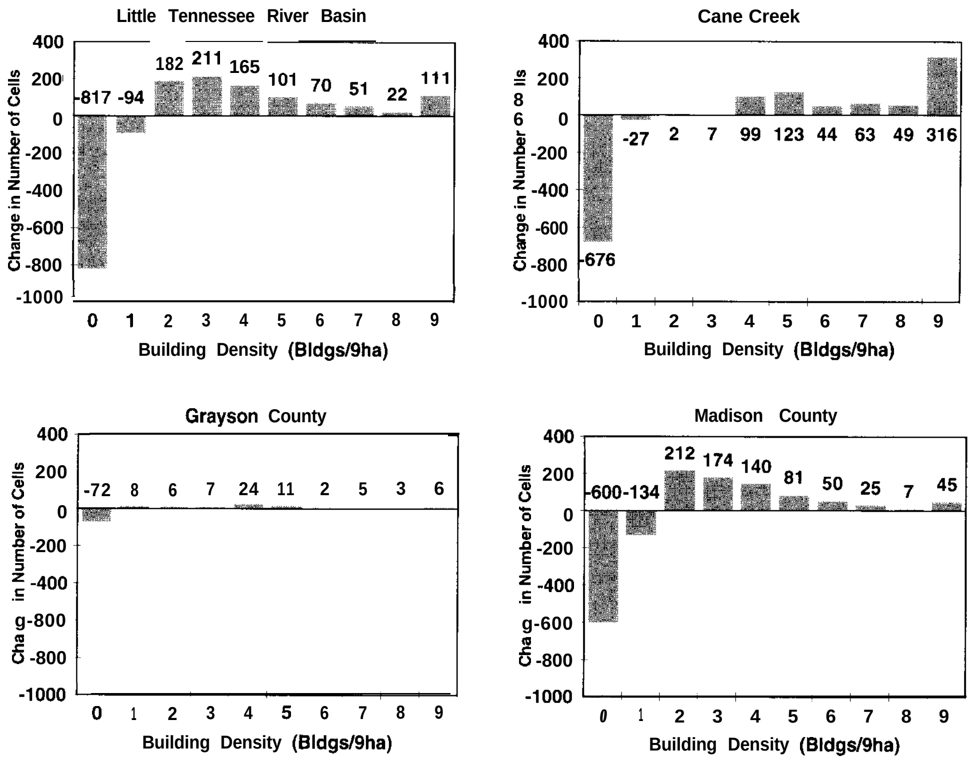

Figure 4. Net changes in building-density classes for land in the four study sites based on sample of landscape grid cells ($n = 5000$ for each site).

County, forest without buildings increased (+3.36%). However, forest with buildings also expanded, similar to Cane Creek and the LTRB.

Average building density more than doubled over this period (+ 131.46%). The amount of change in building-density classes was similar to those observed for the LTRB and Cane Creek, with most gains in the 2, 3, 4 and 5 bldg/9-ha classes, indicating expansion in the area of low-density residential developments for this site. A 1.14% decrease in travel time to MC indicates only a slight expansion in the road network, whereas a 9.13% decrease in distance to roads indicates some filling in of the network, though to a lesser degree than observed for the LTRB and Cane Creek study areas.

All study areas. All study areas showed large increases in forest cover, largely at the expense of agricultural land (Table 1). Agricultural land use was prevalent in all four study areas in the 1950s, claiming from 15% to 50% of the land surface. Agriculture dominated the flatter lands in all study areas; the lower proportion of agricultural land in the Little Tennessee River in part reflects steeper, more variable terrain. Forest cover increased for study areas that had been heavily agricultural (Cane Creek and Grayson County) and for the two less-farmed study areas (the Little Tennessee River and Madison County). Urban land use was a small portion of each study area in the early 1950s and, although increases on a percentage basis were quite large, the absolute increases were generally small.

These land-use changes are reflected in a number of structural indices of land cover and land use. The mean patch size increased for all four study areas, and the area-weighted median patch size increased for three of the four study areas, with Grayson County as the lone exception. Forests comprised the largest patches in all study areas, and the increase in patch size reflects an increase in forest cover. The mean patch size is approximately 2–4 orders of magnitude smaller than the area-weighted patch size indicating a large number of small patches. These small patches represent a small portion of the aggregate study area but, when averaged, significantly reduce mean patch size. When weighted by area, the influence of large forest patches is observed. Grayson County differs from the other study areas in that more than half of its area was in agriculture in 1950, with larger, contiguous blocks of agriculture. As these become more broken with forest regrowth, the median area-weighted patch size decreases. The percent core area also increased, representing the proportion of the landscape more than 100 m from a change in land use. Contagion also increased, reflecting greater patch connection across the landscape. All of these measures represent the reduction in agricultural land, both the complete “absorption” of small, isolated farm par-

Table 4. Marginal Effects Coefficients of Referenced Variable on Building Counts from Estimated Negative Binomial Regression Models and Log-likelihood Ratio for Each Model

	Study Area			
Variable	LTRB	Cane Creek	Grayson County	Madison County
BD_{50}	0.53570	0.44363	0.31559	0.60746
DtM_{50}	-0.03070	0.01920	NS	-0.01735
DtR_{50}	-0.00257	-0.00526	-0.00240	-0.00139
Slope	-0.04014	-0.05871	NS	-0.02971
Elevation	0.00126	-0.00351	NS	NS
LLR-Poisson	7378.62	17957.23	6987.59	6419.12
LLR-Neg. binomial	639.28	798.73	12.16	208.38

Reported values are significant at the 5% level based on the z statistic for coefficients, a chi-squared test with 5 df for the log-likelihood ratio testing overall significance of the Poisson model, and a chi-squared test with 1 df for the log-likelihood ratio testing significance of the negative binomial model against the null of the Poisson model. NS, that thr coefficient is not significant; the other abbreviations are defined in the text

cels into the surrounding forest matrix, and the growth of forest along agricultural edges. As this change occurs, the remaining forests become larger (larger patch size), have less edge (larger core areas), and are more likely to span the landscape (increased contagion).

Other-indices reflect the greater urbanization in the study areas. The mean distance from any point in the landscape to the nearest road decreased for all four study areas. Initial distances varied from 0.25 to 0.5 km, values typical of rural areas, and decreased from slightly (Grayson County) to substantially (Cane Creek), reflecting little to substantial new road construction. There was a concomitant increase in mean building density at all four study sites. Urban densities were typically from 3 to 18 structures/ha and, when averaged over the nonurban land uses, yield study-area averages of 0.04–0.16/ha in the 1950s. In three of four study areas, the number of buildings nearly doubled (Table 1). Grayson County was the exception, with only a slight increase in building density. These trends were observed when considering all land but were amplified when restricted to private land, because nearly all public land was and has remained in forest.

Estimation Results

Building-density equation. Estimates of the negative binomial regression for building density [Eq. (3)] indicate that equations for all study sites were significant. Chi-squared tests (Table 4) indicate that the Poisson regressions were all significant ($df = 5$ and $P \leq 0.05$) and that the negative binomial regression provided a significant improvement over the Poisson ($df = 1$ and $P \leq 0.05$). We therefore reject the hypothesis that the building-density models do not have explanatory power.

Tests regarding the significance of explanatory variables in the building-density equations varied across sites. For all sites, the marginal effects coefficient for lagged building density was significantly positive ($P \leq 0.05$) consistent with expectations. Slope had a significant and negative effect on building density for all study sites except Grayson County (where it was not a significant variable), and distance to road had a significant and negative effect on all study sites. Travel time to MC was significant and negative for the LTRB and the Madison County study sites, consistent with expectations, but it was significant and positive for the Cane Creek site. It was insignificant in the Grayson County site. Elevation was significantly positive in the LTRB but significant and negative in Cane Creek. It was insignificant for the other two study areas. Of the 20 estimated coefficients, 16 were significant and, of these, 14 were consistent with expectations based on theory. For the three study sites with substantial development (LTRB, Cane Creek, and Madison County), only one coefficient among 15 did not have a significant effect on building density.

Land-cover equations. Logit regression models of the probability of forest versus nonforest land cover were all significant (Table 5). We rejected no explanatory power based on the log-likelihood ratio test ($df = 4$ and $P \leq 0.05$). Marginal effects coefficients listed in Table 5 were measured with respect to the probability of observing nonforest cover. All coefficients for expected building density (EB) were significant and positive, consistent with theory. The slope coefficient was significant and negative for all four areas. The elevation coefficients were significant and negative in the LTRB, Cane Creek, and Madison County areas.

Table 5. Marginal Effects Coefficients of Referenced Variable on the Probability That Land Is Nonforest Derived from Estimated Logit Models of Land Cover and Log-likelihood Ratio for Each Mode1

Variable	Study Area: LTRB	Cane Creek	Grayson County	Madison County
NF_{50}	0.00155	0.00351	0.00845	0.00166
EB_{90}	0.01 148	0.02977	0.03784	0.00959
Slope	-0.00457	-0.03004	-0.01944	-0.006589
Elevation	-0.00004	-0.00056	NS	-0.00007
LLR	1568.1	2656.59	1643.23	1389.39

Reported values are significant at the 5% level based on the z statistic for coefficients, and on a chi-squared test with four degrees of freedom for the log-likelihood ratio testing the significance of the estimated logit model. NS, the coefficient is not significant; the other abbreviations are defined in the text.

The coefficients on the share of nonforest area in the 9-ha neighborhood of each cell (NF) were significant and positive for all study sites. This variable provides a measure of the spatial contagion of land cover, separate from the effects of the other variables, and therefore defines a test for spatial autocorrelation of land cover. The finding of significant and positive effects indicates that we reject the null hypothesis of no spatial autocorrelation. Of 16 coefficient estimates in the land-cover models, 15 were significant, and all signs were consistent with our expectations.

Plots of predicted probabilities (Figure 5) provide additional insight into the influence of location and topographic variables on land uses. Figure 5 charts the probability of land uses and the diversity of predicted land uses in 1990 for land that was forest with no buildings in 1950 in the Little Tennessee River Basin. Comparisons between rows show the effect of slope while comparisons between columns show the effect of distance to roads on these values. The probability of forested land converting to nonforest is small throughout the range of scenarios and is discernible only on flat land (slope = 0°). On this portion of the landscape, the probability of forest with buildings is also highest. The effect of distance to roads can be evaluated by comparing graphs in the top row of Figure **5.** With decreasing distance to road, forest with buildings dominates on land further from the MC. At $DtR_{50} = 0$, forest without buildings dominates from the MC [Pr(FNB) >0.5] up to about 50 min from the center; up to 35 min at $DtR_{50} = 250$ and up to 15 min at $DtR_{50} = 500$. At slope = 30, forest without buildings dominates over nearly the entire range of DtR_{50} and travel time. The exception is where $DtR_{50} = 0$ and travel time is less than 10 min. At slope = 60, forest without buildings unambiguously dominates the landscape. The diversity of land uses is also influenced by DtR, slope, and travel time. Land use is most diverse on flat land at a moderate distance from the MC. It declines with travel time to an asymptote of 0.2 for slope = 0 and to 0.0 for slope = 30 and 60.

Forecast Evaluation

Forecasts of land use for observations in the validation data sets indicated mixed success in predicting specific land-use categories. For the fine-resolution building classes, the share of the landscape correctly predicted by the forecasting mode1 ranged from 67.63% in the LTRB to 71.61% in Grayson County (Table 6). The mode1 had the most success in predicting land in forest without buildings (correct predictions of FNB ranged from 87.65% to 96.86%). The model also correctly predicted the location of a majority of nonforest cover (between 61.63% and 79.67%) for the LTRB, Cane Creek, and Grayson County areas. The mode1 had less success in predicting the specific building density for forested land with buildings, with success declining as building density increased. Predictions for the broad building classes (Table 6) shows that, in the Little Tennessee and Cane Creek, the models correctly predict roughly one-half of the forest area with buildings (55.31% and 49.59%, respectively). In contrast to these areas with the highest growth over the study period, the model correctly predicted much less of the forest with buildings in Grayson County (25.50%) and Madison County (38.54%). Results for the land-cover classes alone (Table 6c) indicate that the mode1 correctly predicted land-cover classes for between 79.50% (Grayson County) and 89.38% (Little Tennessee) of the validation data.

The information-theoretic statistics (Table 7), which are based on an assessment of improvements in the accuracy of the predicted probabilities, indicate that the forecast models are significant and that

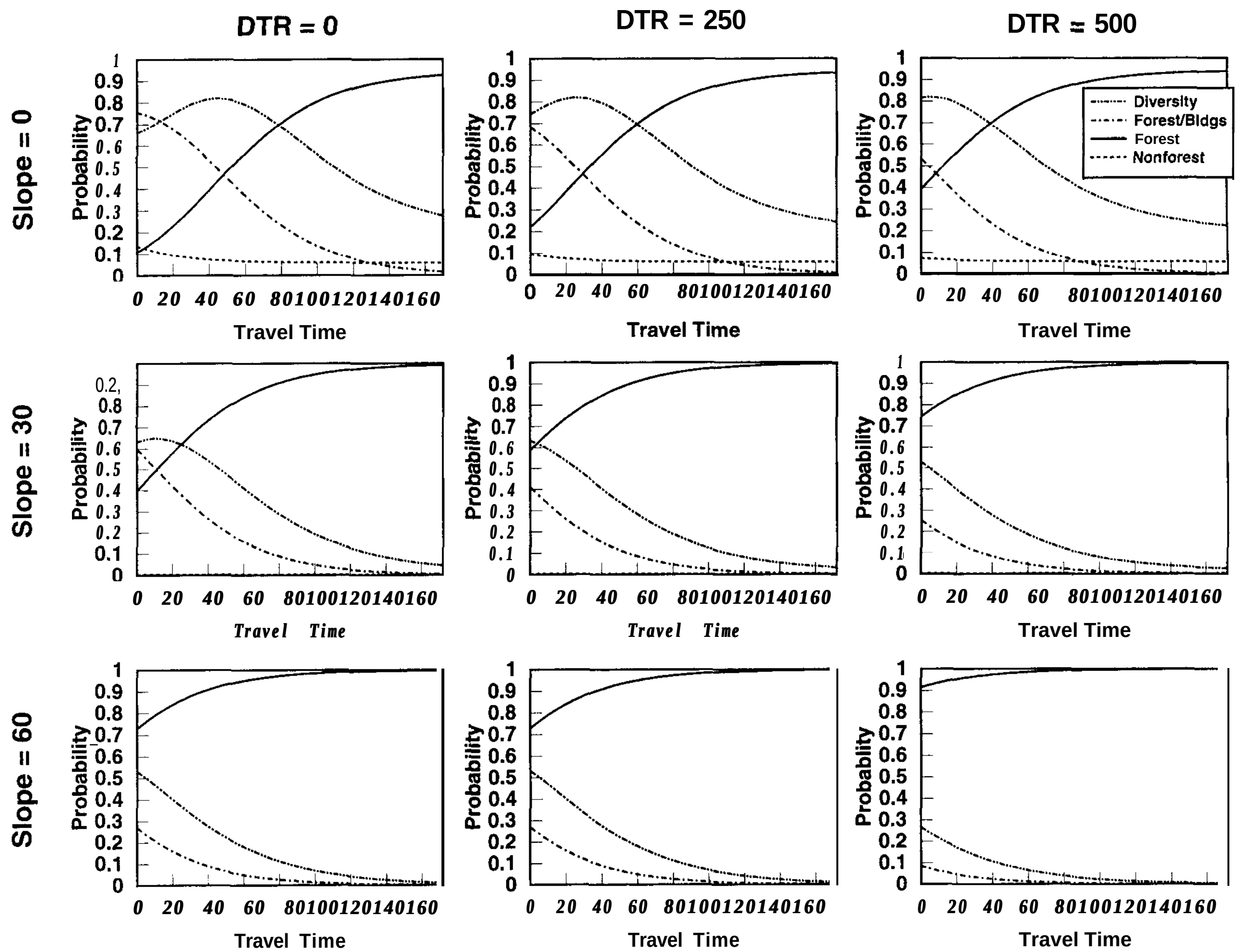

Figure 5. Predicted probability of broad land-use classes in 1990 for land that was forested without buildings in 1950 based on models for the Little Tennessee River Basin. Scenarios are defined with NF_{50} = 0 and elevation = 800. Other variables [slope (degrees) and DtR_{50}, and DtM_{50} (minutes of travel time)] are varied as shown in the panels. Diversity is measured relative to maximum possible diversity using the Shannon-Weaver index. NF_{50}, DtR_{50}, and DtM_{50} are defined in the text.

they provide significant improvements over null models of land use. Hauser's *significame* test based on the log-likelihood ratio test indicates rejection of the hypothesis that the model provides no improvement over the null mode1 for all four of the study areas. U^2 statistics (Hauser's *usefulness* measure; Table 7) indicate that the forecasting model reduces the residual uncertainty of the null model-defined by the frequencies of land uses in the estimation data sets-by between 47% and 66%. Results of the *accuracy* tests (Table 7), which compare I(A;X,Z) with the confidence interval around the expected information EI(A;X,Z), are, however, mixed. For the LTRB and Cane Creek study sites, we cannot reject the hypothesis that the estimated models explain the observations in the validation data set (z test with $P \leq 0.05$). For Grayson and Madison Counties, the hypothesis is rejected, indicating that, although the estimated models have explanatory power (based on usefulness and significance tests), additional work on their specification is warranted.

Landscape Simulations

To demonstrate application of these models further, we developed land-use forecasts for our four study sites. Equations (3) and (6) were applied to GIS data layers for all cells in the referenced landscapes. Conditions in 1990 were then used to forecast land use in 2030, in effect applying the processes behind land-use changes over the 1950–90 period to simulate the next 40 years.

Results (Figures 6 and 7) forecast increases in building density and some change in land cover. Overall, land cover appears relatively stable, while human populations continue to grow across these landscapes. One exception is in Cane Creek, where

Table 6. Percentage of Correct Predictions for Three Classification Schemes: Fine-Resolution Building Classes, Broad Building Classes, and Land Cover Alone

	Percent Correctly Predicted			
Land-Use Class	LTRB	Cane Creek	Grayson County	Madison County
Fine-resolution building classes				
Forest				
0 Buildings	90.32	88.13	87.65	96.86
l-2 Buildings	46.73	43.18	19.13	33.03
3-5 Buildings	17.78	20.00	9.09	6.90
6-8 Buildings	13.64	4.35	0.00	28.57
>8 Buildings	0.00	0.00		0.00
Nonforest	61.63	79.67	71.98	49.39
Total	67.63	70.87	71.61	68.91
Broad building classes				
Forest				
0 Buildings	90.32	88.13	87.65	96.86
With buildings	55.31	49.59	25.50	38.54
Nonforest	61.63	79.69	71.98	49.39
Total	74.15	75.67	72.93	72.14
Land cover alone				
Forest	95.16	88.44	83.41	94.58
Nonforest	61.63	79.67	71.98	49.39
Total	89.38	85.29	79.50	87.19

Predictions were developed for validation data sets by using the forecasting model defined by Eqs. (3) and (6). The dash indicates no observations in this category.

Table 7. Information Indices and Derived Statistics for the Forecasting Model Applied to Validation Data Sets by Using the Fine-Resolution Building Classes

	Study Area			
Index/ Statistic	LTRB	Cane Creek	Grayson County	Madison County
I(A;X,Z)	0.598	0.728	0.418[b]	0.532[b]
El(A;X,Z)	0.602	0.713	0.333	0.441
V(A;X,Z)	0.0011	0.0011	0.0005	0.0007
H(A)	0.988	1.102	0.890	0.982
U^2	0.605	0.660	0.469	0.541
LLR	1188.3”	1446.3”	830.06”	1057.0”

[a]*Significant at $P = 0.05$.*
[b]*Indicates that I(A;X,Z) is significantly different from EI(A;X,Z).*

projected growth in building density would lead to a substantial agglomeration of nonforest cover. This finding suggests that land-cover trends in other areas may be reversed as well. Shifts toward forest cover have been fueled by declining agricultural uses over the past 40 years. However, as indicated by the land-cover model, as building densities increase in response to population growth, land will eventually shift from forest with buildings to nonforest.

These simulations highlight portions of the landscapes where land uses would be relatively stable. For example, areas in public ownership (almost exclusively National Forests; Figure 8), where land use is restricted by law and administrative rules, are highly stable. Additionally, areas with high slopes and low access (for example, see simulations for Cane Creek) remain in forest without buildings in spite of strong development pressures. As suggested by the coefficients of the estimated models, change is most likely to occur in areas close to roads and with low slopes.

DISCUSSION

Estimates of both building-density and land-cover equations reinforce previous findings regarding the effects of location on land uses. We find that topographic features-that is, slope and elevation—hold significant influence on land uses/cover and therefore on landscape configurations, as did Wear and Flamm (1994), Turner and colleagues (1996), Nelson and Hellerstein (1997), and Chomitz and Gray (1995). Topography is, therefore, a significant constraining factor in these heterogeneous landscapes, limiting intensive land uses to only certain portions (Figure 5). Accordingly, some areas will likely persist in forest cover regardless of development pressures. However, these constraints will also concentrate agricultural and residential/urban uses onto specific portions of the landscape.

Similarly, we find, as have others (Chomitz and Gray 1995; Nelson and Hellerstein 1997; Turner and others 1996), that land-use changes are also strongly influenced by the location of sites measured in terms of distance to roads and travel time to the closest MC. These findings further corroborate the dominant Von Thiinen model of rural land-use specialization and central business district models of urban development [for example, see Capozza and Helmsley (1989)]. While distance and travel time variables influence land-use patterns in a manner similar to topography, it is important to remember that distance factors are mutable. Road construction and improvements have ramifications for the subsequent evolution of rural landscapes. Our findings highlight that constructing roads for one purpose—for example, access to a new subdivision-will likely

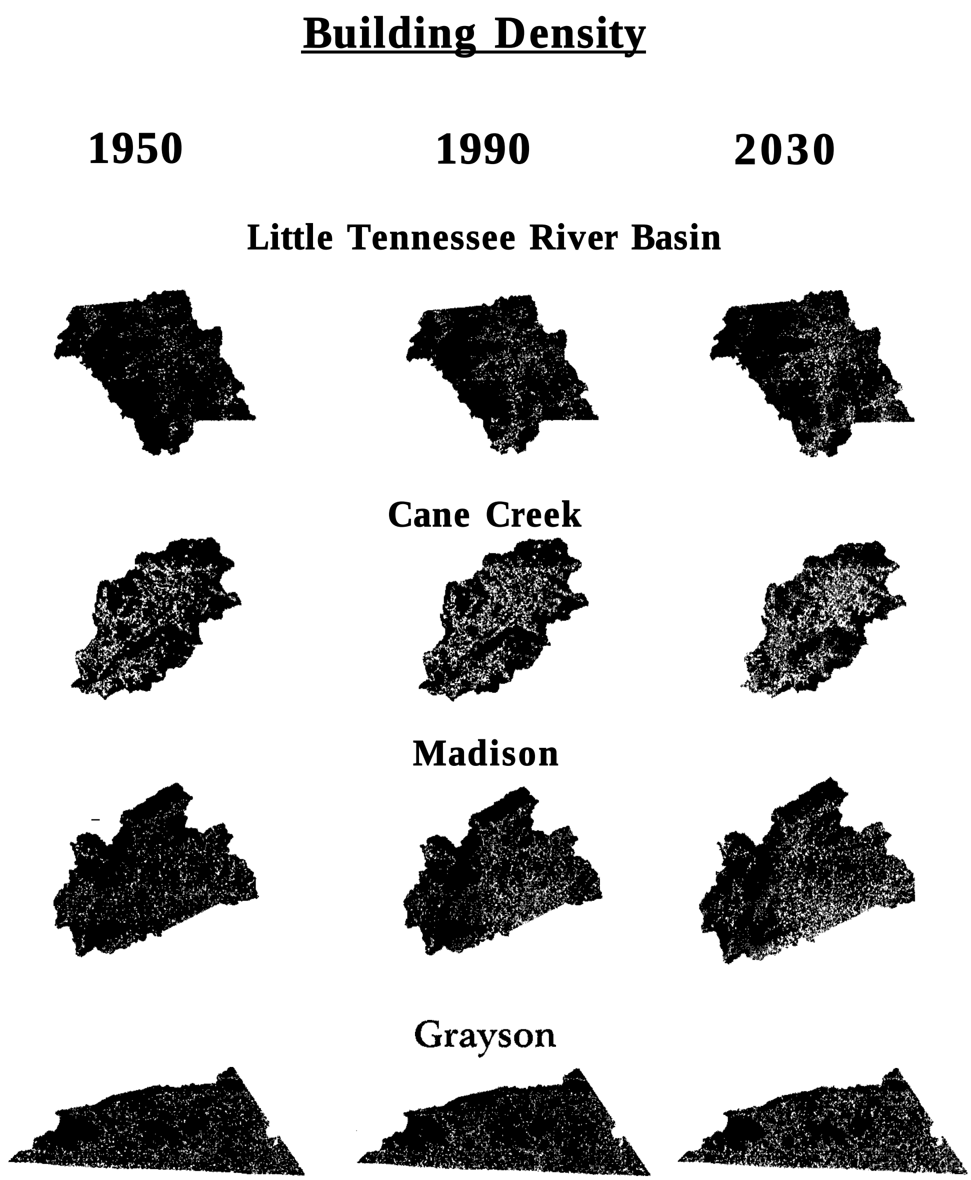

Figure 6. Observed (1950 and 1990) and predicted (2030) building-density classes for the four study sites.

have secondary and tertiary impacts on a larger portion of the referenced landscape.

There has been little expansion of the road network in all study sites between 1950 and 1990. Rather, it appears that the primary road network has been fairly stable since 1950 in all four study areas. Construction generally filled in the road network, thereby bringing land closer to the road network. As a result, the average distance to road declined substantially between 1950 and 1990 in the three study areas experiencing growth. In contrast, there was little change in Grayson County.

Taken together, influences of topographic and distance variables on land use may concentrate development in riparian areas, an especially important portion of the landscape. Our models indicate that the highest probability of intensive land use occurs where land is relatively level, close to roads, and close to MCs. Since MCs are located on the rivers of the region, these factors unambiguously concentrate development around water courses. Water quality and the structure and function of aquatic ecosystems in this region have been heavily influenced by massive deforestation at the turn of the century, followed by decades of agricultural uses. Our analysis of land-use changes indicates that these areas will further be effected by shifts toward residential uses. Conversely, a large portion of upland areas on moderate-to-steep slopes are likely to be retained in a forest cover without buildings. While occasional timber harvesting is likely, our models identify large portions of the study areas

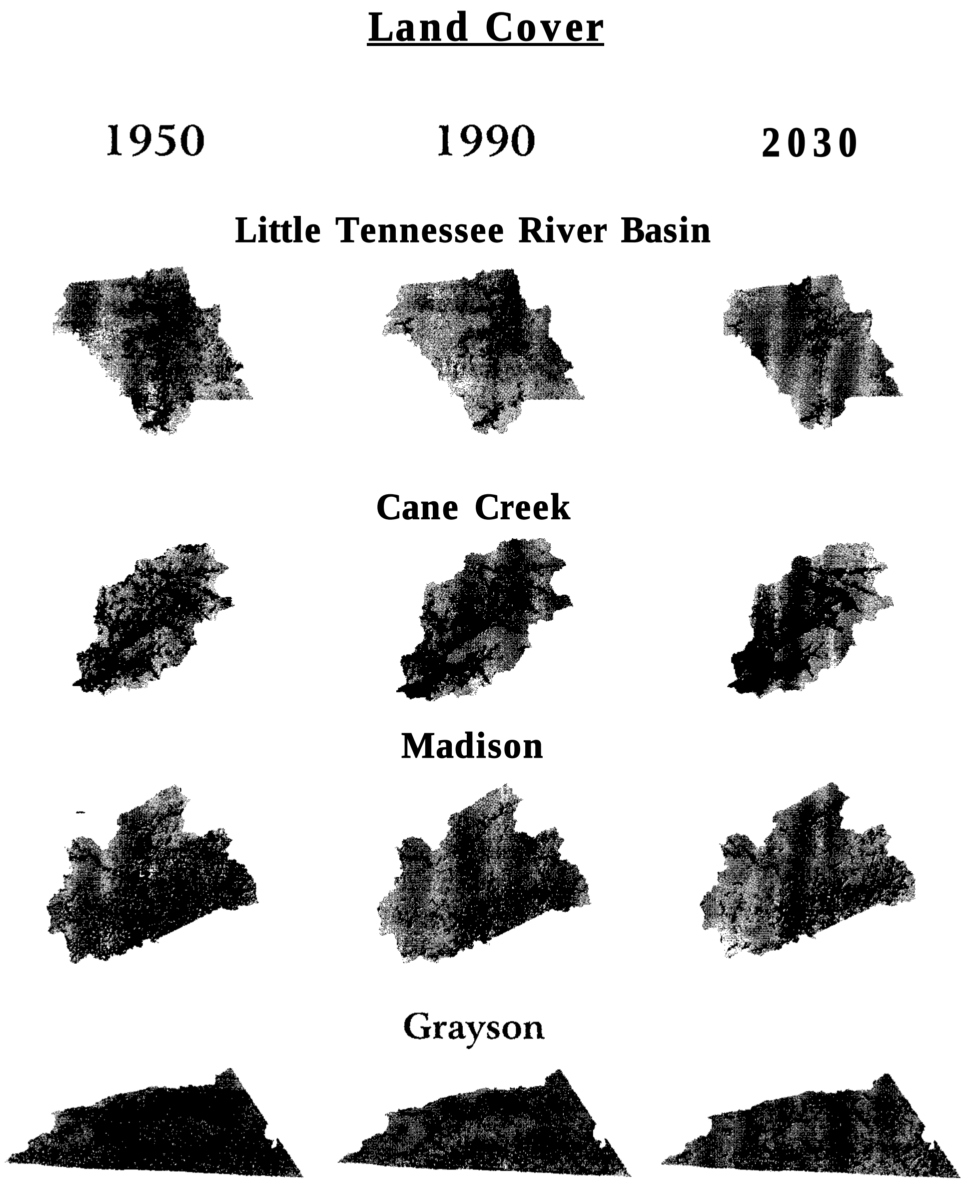

Figure 7. Observed (1950 and 1990) and predicted (2030) land-cover classes for the four study sites (*green* is forest and *orange* is nonforest cover).

where land cover and usage will be relatively stable in the future.

These findings, along with our land-use projections, suggest how spatial models such as these could be used to design development and conservation strategies that address specific ecological concerns in specific places. For example, in rapidly developing study areas, it appears that the ecological consequences of land-use changes will be most severe in riparian areas. The potential ecological benefits of buffering development in this relatively small portion of the landscape-that is, mitigating temperature gains, changes in stream chemistry, and consequent alterations in species diversity—could be high relative to the area impacted. In comparison, efforts to maintain connectivity in upland areas may have relatively low returns if these areas have fairly stable land-cover conditions. In this way, modeling human dynamics can identify where ecological benefits will be relatively scarce or plentiful. Costs need also to be considered. Because the value of land is generally much greater in accessible riparian areas than in upland areas, the per-acre cost of conservation easements or regulations may be relatively high in these important areas and therefore more difficult to implement. These types of cost-benefit trade-offs are a critical part of designing effective conservation plans. Of course, development of effective conservation strategies requires more than just projections of land uses. It requires meaningful insights into the effects of land use on ecosystem structure and function. Nonethe-

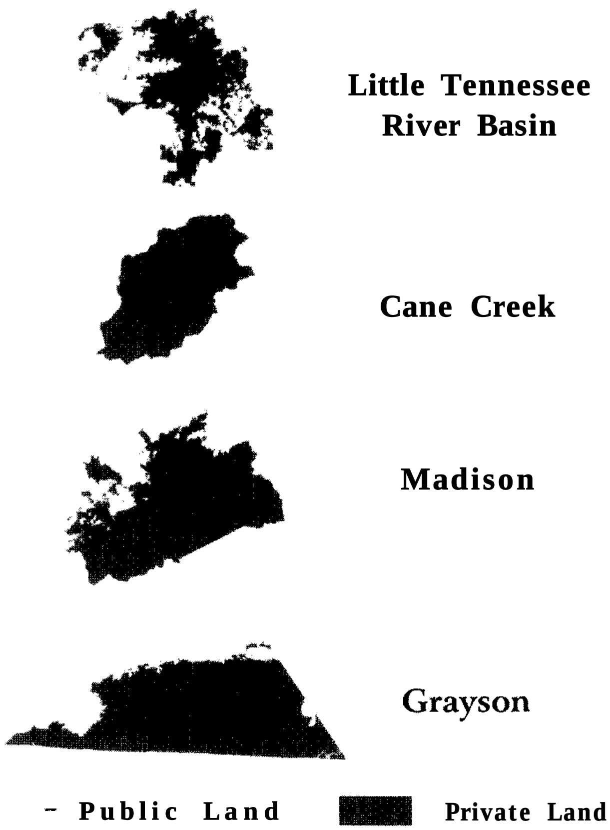

Figure 8. Broad land-ownership categories for the four study sites.

less, landscape models can help focus investigation on portions of the landscape that are most likely to change.

Estimates of the land-cover logit models all show a significant positive effect of neighborhood nonforest cover (NF,,). As a result, we reject the hypothesis of no spatial autocorrelation. Dubin (1995) has shown that unaddressed spatial autocorrelation can be a significant misspecification of the logit maximum-likelihood function, leading to inconsistent estimators. We have applied a plausible correction (using a simple, spatially weighted explanatory variable) without directly modeling the cause of the autocorrelation. Clearly, other formulations may be appropriate, and this defines an issue for future research.

One of our objectives was to refine the definition of land use by using building density to approximate human presence. Incorporation of this layer of data leads to fundamentally different insights into landscape change in the Southern Appalachians. Although all study sites experienced increases in forest cover between 1950 and 1990, ostensibly reflecting the reversion of agricultural land to forests, this apparent transition to a more natural cover masked significant human impacts revealed by changes in building densities. In the three study sites experiencing development pressures (LTRB, Cane Creek, and Madison County), the amount of forest without buildings declined substantially, while the area of forest with buildings and the average building density increased. Observations of land-cover change alone may prove misleading because they measure the net outcome of more than one process (in this case, agricultural decline combined with population growth.)

These results also show that changes measured at coarse scales may not reflect change at spatially fine scales. This effect can be seen especially in Madison County, where total population actually declined between 1950 and 1990 (-17.4%) with nearly all of this change occurring between 1950 and 1960. The total population of the county was essentially stable between 1960 and 1990. However, the spatial distribution of the population clearly changed. The average building density more than doubled over this period (from 0.45 to 1 .O), and shifts in building-density classes were comparable to changes observed in areas with substantial population growth (LTRB and Cane Creek). These changes in building densities suggest a shift toward fewer people per household in addition to a spreading out of residences within the county (Figure 6). Such a change is confirmed by data from the US Census Bureau. In Madison County, persons per household fell 22%, from 3.17 to 2.48, between 1970 and 1990. In effect, then, no change in total population masks an increase in the number and spatial distribution of households and, therefore, in the area effected by human presence.

Although composite land-use classes offer improvement over land cover alone for explaining land-use change, they are untested as ecological units. Additional research is needed to define the ecological implications of building density and other measures of human presence. For example, at what level does building density impede the dispersal of organisms or reduce forest-interior species? These kinds of insights are needed to define ecologically meaningful land-use categories and, more generally, to understand the effects of urban-rural gradients on ecosystem structure and function. In addition, it might be possible to refine our building categories either by recognizing different types of buildings (for example, dwellings vs barns vs commercial facilities) or by incorporating altogether different measures of human presence (for instance, population density).

While we sought to evaluate land-use change over a long period, the 40-year time step may have proved a limitation of our analysis. We expect that structural changes in factors such as access (for example, development of interstate highways) and agricultural markets (for example, reform of farm subsidies) may have shifted land-use regimes within this time frame. Future research should focus not only on shortening the time step but, more generally, on defining appropriate temporal and spatial scales for land-use change analysis.

In spite of these limitations, the forecasting model developed here was able to explain 80%–89% of land cover, 73%–76% of broad land-use classes, and 68%–75% of the fine land-use classes in 1990 based on conditions observed in 1950. These validation tests, along with the results of significance and accuracy tests, indicates that these models are powerful tools for explaining observed changes over the historical period. As is the case with any forecasting model, however, there is unquantifiable uncertainty regarding a model's ability to forecast change beyond this historical period. These models apply the processes that changed landscapes between 1950 and 1990 to current landscapes. But, just as changes in access, agricultural markets, and population growth altered demands for land over this 40-year period, we could reasonably expect additional structural changes in the future. Forecasts necessarily should be viewed with trepidation.

However, in the context of evaluating future ecological impacts or conservation strategies, we are not as interested in predicting a specific land use in a particular place as we are interested in understanding where the relative probability of change is high and where it is low. Land-use forecasting models can be used to drive hazard or risk assessments where land-use maps are linked to ecological impact models. In this kind of analysis, the focus shifts to understanding where human activities are most likely to generate significant ecological consequences. Thie type of hazard analysis could focus subsequent research and planning in areas that are the most critical in supporting ecological health. Linking the spatial dynamics of human populations to potential ecological impacts defines an important emphasis for future research.

ACKNOWLEDGMENTS

This research was supported by an National Science Foundation grant and by the Economics of Forest Protection and Management Research Work Unit, Southern Research Station, USDA Forest Service. The authors thank Dr. Frank Benford for his careful review of an earlier version of the manuscript and Diane Riggsbee for her help with the figures.

REFERENCES

Alig RJ. 1986. Econometric analysis of the factors influencing acreage trends in the Southeast For Sci 32:119–34.

Bernhardsen T. 1992. Geographic information systems. Myrene (Norway): VIAK. 318 p.

Cameron AC, Trivedi PK. 1986. Econometric models based on count data: comparisons and applications of some estimators and tests. J Appl Econom 1:29–53.

Capozza DR, Helsley RW. 1989. The fundamentals of land prices and urban growth. J Urban Econ 26:295–306.

Chomitz KM, Gray DA. 1995. Roads, land, markets and deforestation: a spatial model of land use in Belize. World Bank Econ Rev 10:487–5 12.

Colwell RK. 1974. Predictability, constancy, and contingency of periodic phenomena. Ecology 55:1148–53.

Costanza R, Sklar FH, White ML. 1990. Modeling coastal landscape dynamics. Bioscience 40:91–107.

Dale VH, O'Neill RV, Pedlowski M, Southworth F. 1993. Causes and effects of land-use change in Central Rondônia, Brazil. Photogrammetric Eng Remote Sens 59:997–1005.

Dubin R. 1995. Estimating logit models with spatial dependence. Anselin L, Horax JGM, editors. New directions in spatial econometrics. New York: Springer-Verlag. p 229-42.

Geoghegan J, Bockstael N. 1997. Human behavior and ecosystem valuation: an application to the Patuxent watershed of the Chesapeake Bay. In: Simpson RD, Christensen Jr NL, editors. Ecosystem function and human activities. New York: Chapman and Hall. p 147-74.

Gottfried RR, Wear DN, Lee RG. 1996. Institutional solutions to market failure on the landscape scale. Ecol Econ 18: 133–40.

Greene WH. 1995. LIMDEP: user's manual and reference guide. Version 7.0. Bellport (NY): Econometric Software.

Hardie IW, Parks PJ. 1997. Land use in a region with heterogenous land quality: an application of an area base model. Am J Agric Econ 79:299–3 10.

Hauser JR. 1978. Testing the accuracy, usefulness and significance of probabilistic choice models: an information-theoretic approach. Operations Res 26:406–21.

Judge GG, Griffiths WE, Hill RC, Lutkepolh WE, Lee TC. 1985. The theory and practice of econometrics. New York: John Wiley and Sons.

Katzman MT. 1974. The Von Thünen paradigm, the industrial–urban hypothesis, and the spatial structure of agriculture. Am J Agric Econ 56:683–96.

Kullback S. 1959. Information theory and statistics. New York: John Wiley and Sons.

Lee RG, Flamm RO, Turner MG, Bledsoe C, Chandler P, DeFerrari C, Gottfried R, Naiman RJ, Schumaker N, Wear D. 1992. Integrating sustainable development and environmental vitality. In: Naiman RJ, editor. Watershed management: balancing sustainability and environmental change. New York: Springer-Verlag. p 499-52 1.

Lillesand TM, Kiefer RW. 1994. Remote sensing and image interpretation. 3rd ed. New York: John Wiley and Sons. 750 p.

Lubchenco J, Olson AM, Brubaker LB, Carpenter SR, Holland MM, Hubbell SP, Levin SA, McMahon JA, Matson PA, Melillo JM, and others. 1991. The sustainable biosphere initiative: an ecological research agenda. Ecology 72:371–412.

Nelson GC, Hellerstein D. 1997. Do roads cause deforestation? Using satellite images in econometric analysis of land use. Am J Agric Econ 79:80–8.

Palmquist RB, Danielson LE. 1989. A hedonic study of the effects of erosion control and drainage on farmland values. Am J Agric Econ 7:55–62.

[SAMAB] Southern Appalachian Man and the Biosphere. 1996. The Southern Appalachian assessment social/cultural/economic technical report: report 4 of 5. Atlanta: US Department of Agriculture, Forest Service, Southern Region.

Samuelson PA. 1983. Thünen at two hundred. J Econ Lit 21:1468–8.

Sokal RR, Rohlf FJ. 1981. Biometrics. San Francisco: WH Freeman.

Spies TA, Ripple WJ, Bradshaw GA. 1994. Dynamics and pattern of a managed coniferous forest landscape in Oregon. Ecol Appl 4:555–68.

Turner MG, Costanza R, Sklar FH. 1989. Methods to evaluate the performance of spatial simulation models. Ecol Modell 48: 1-18.

Turner MG, Wear DN, Flamm RO. 1996. Land ownership and land-cover change in the Southern Appalachian Highlands and the Olympic Peninsula. Ecol Appl 6: 1150–72.

Wear DN, Flamm RO. 1993. Public and private disturbance regimes in the Southern Appalachians. Nat Rcsour Model 7:379–97.

Wear DN, Turner MG, Flamm RO. 1996. Ecosystem management with multiple owners: landscape dynamics in a Southern Appalachian watershed. Ecol Appl 6:1173–88.

Wolf PR. 1983. Element of photogrammetry. New York: McGraw-Hill. 628 p.