Landscape Ecology **18:** 347–361, 2003.

Research article

Modeling the spatially dynamic distribution of humans in the Oregon (USA) Coast Range

Jeffrey D. Kline[1,*], David L. Azuma[2] and Alissa Moses[3]

[1]*Forestry Sciences Laboratory, Pacific Northwest Research Station, 3200 SW Jefferson Way, Corvallis, OR 97331, USA;* [2]*Forestry Sciences Laboratory, Pacific Northwest Research Station, Portland, OR 97208, USA;* [3]*Department of Forest Science, Oregon State University, Corvallis, OR 97331, USA; *Author for correspondence (e-mail: JKline@fs.fed.us)*

Received 23 November 2001; accepted in revised form 30 April 2002

Key words: Ecological economics, Forest/urban interface, Land use change, Landscape modeling, Western Oregon, USA

Abstract

A common approach to land use change analyses in multidisciplinary landscape-level studies is to delineate discrete forest and non-forest or urban and non-urban land use categories to serve as inputs into sets of integrated sub-models describing socioeconomic and ecological processes. Such discrete land use categories, however, may be inappropriate when the socioeconomic and ecological processes under study are sensitive to a range of human habitation. In this paper, we characterize the spatial dynamic distribution of humans throughout the forest landscape of western Oregon (USA). We develop an empirical model describing the spatial distribution and rate of change in historic building densities as a function of a gravity index of development pressure, existing building densities, slope, elevation, and existing land use zoning. We use the empirical model to project changes in building densities that are applied to a 1995 base map of building density to describe future spatial distributions of buildings over time. The projected building density maps serve as inputs into a multidisciplinary landscape-level analysis of socioeconomic and ecological processes in Oregon's Coast Range Mountains.

Introduction

A common approach to multidisciplinary landscape-level analysis of socioeconomic and ecological processes is to treat humans largely as separate from the forest landscape. Empirical models of land use change commonly have been used in landscape analyses to delineate discrete forest and nonforest, forest and urban, or other similar discrete land use categories, to serve as inputs into sets of integrated sub-models describing socioeconomic and ecological processes and conditions (see, for example, Bockstael (1996) and Irwin and Geoghegan (2001), Kline et al. (2001), Schoorl and Veldkamp (2001)). Such delineations often are intended to identify where humans are and are not present on the landscape. We are aware of only two studies that attempt to treat humans as part of the landscape, by describing a range of human habitation. Wear and Bolstad (1998) develop an empirical model of building densities to describe the ″spatial diffusion″ of human populations, but ultimately use their building density to describe discrete forest and nonforest land use categories. Fagan et al. (2001) suggest several modeling approaches for describing housing starts near cities, but lack empirical data with which to estimate and test empirical versions of their models. We build upon these works by examining the spatial distribution and rate of change in historic building densities in western Oregon, USA and use this information to characterize the future spatial distributions of humans throughout Oregon's Coast Range Mountains.

For many applications, a discrete treatment of land use may be appropriate, when the landscape-level socioeconomic and ecological processes under study are relatively insensitive to low levels of human habitation. For example, in many studies land use modeling is focused more on characterizing changes in land (or vegetative) cover than on characterizing the level of human habitation. Examples of such studies include models of agricultural cropping patterns (Serneels and Lambin 2001; Walsh et al. 2001), forest succession (Turner et al. 1996; Helmer 2000), or deforestation (Geoghegan et al. 2001; Schneider and Pontius 2001) to name a few. Also, the specific intent of many studies is to characterize the probability of a particular type of land use change occurring, to identify potential priority conservation areas for example (Swenson and Franklin 2000), or to identify causal factors of land use change (for example, Nelson and Hellerstein (1997)), rather than projecting potential future land use scenarios. However, in other applications, where a more explicit characterization of potential future land uses is desired or where land use projections will serve as inputs to other models of socioeconomic and ecological processes that may be sensitive to a range of human habitation, discrete land use categories may be inadequate to characterize the spatial and temporal interactions of humans as agents affecting landscape-level processes under study.

For example, multidisciplinary studies of forest landscapes commonly delineate discrete forest and non-forest land categories as key inputs in sub-models describing timber management and production, both as an economic activity and as an important factor affecting landscape-level ecological processes such as habitat viability. Research, however, suggests that the intensity of timber management and production activities conducted by private forestland owners can be negatively correlated with human population densities such that they vary across forest landscapes depending on human population levels (Barlow et al. 1998; Wear et al. 1999). Habitat viability for certain species itself may vary according to a range of human habitation, in addition to land cover characteristics associated with general land use categories. Fire also may be an important factor in landscape-level modeling. Some forestry analysts hypothesize that increasing numbers of residences located in forested landscapes increase the likelihood of wildfire and increase fire suppression costs when firefighting resources are re-directed to save homes instead of containing fires (Milloy 2000). In these and perhaps other examples, discrete land use categories may be less useful as inputs into landscape-level models of socioeconomic and ecological processes than would be more detailed information describing ranges of human habitation on the landscape.

The empirical methods used to model changes among discrete land use categories can involve other difficulties. Empirical land use models based on discrete land use data commonly are estimated using logit or probit techniques that result in projected probabilities of land use change rather than projections of discrete land use categories. These projected probabilities can be difficult to interpret or incorporate into other socioeconomic and ecological models. Discrete land use models also may be limited by the specific characteristics of available land use data. Discrete land use models often are estimated using data collected from land inventories, such as the National Resources Inventory (Nusser and Goebel 1997) and the USDA Forest Service's Forest Inventory and Analysis Program inventories (see, for example, Frayer and Furnival (1999)), which may be designed to meet specific informational objectives. These inventories may categorize land according to criteria or definitions that may be imperfect or inappropriate for examining socioeconomic and ecological processes of interest. Ideally, a modeling approach that allows for a range of human habitation, more definitive projections of change, and greater flexibility in its applicability to issues under study is desirable.

In this paper, we build upon the work of (Wear and Bolstad 1998; Fagan et al. 2001) by characterizing the spatial distribution of humans throughout the landscape comprising Oregon's (USA) Coast Range Mountains. We develop an empirical model describing the spatial distribution and rate of change in historic building densities in western Oregon as a function of a gravity index of development pressure, existing building densities, slope, elevation, and existing land use zoning. We use the empirical model to project pixel-level changes in building densities that are applied to a 1995 base year building density base map to describe the future spatial distributions of buildings through 2055. The building density maps are key inputs in other socioeconomic and ecological sub-models comprising the Coastal Landscape Analysis and Modeling Study in western Oregon (USA).

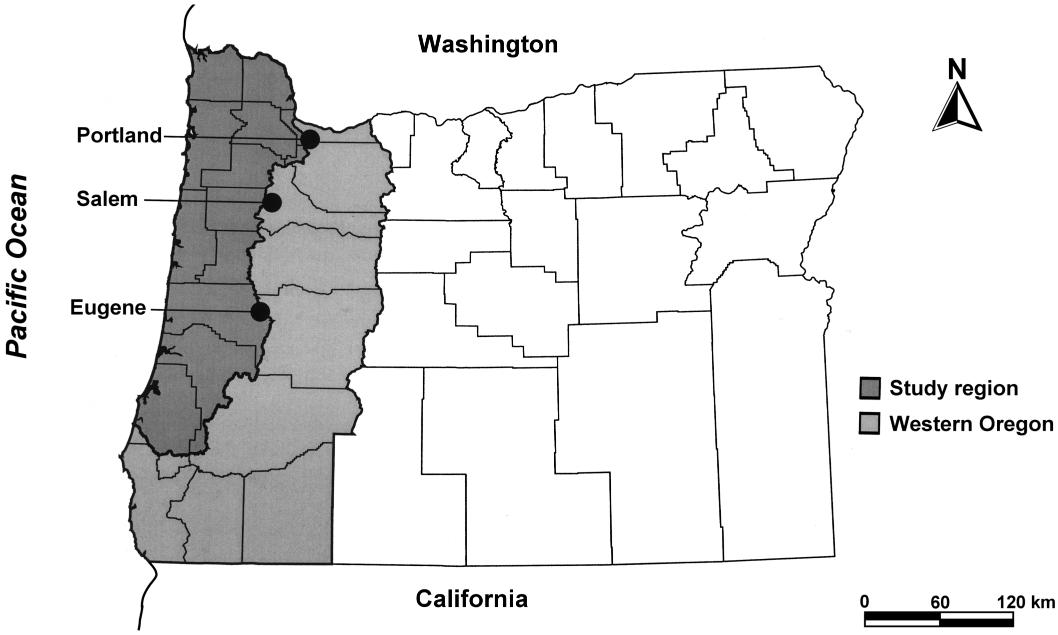

Figure 1. Coastal landscape analysis and modeling study region in Western Oregon.

Study Area

The Coastal Landscape Analysis and Modeling Study (Spies et al. 2002) is a multidisciplinary research effort to analyze the aggregate ecological, economic, and social consequences of forest policies in western Oregon's Coast Range Mountains. The study region borders the Pacific Ocean on the west and the Willamette Valley on the east (Figure 1). Current forest policies in the region attempt to achieve a particular mix of forest goods and services by spatially distributing different forest practices over watersheds or landscapes, and across multiple ownerships. A particular policy concern in recent years has been ensuring sufficient habitat for spotted owls (*Strix occidentalis caurina*) and coho salmon (*Oncorhynchus kisutch*).

The project is intended to provide quantitative analyses testing the assumptions of current forest policies to determine if projected future outcomes are consistent with policy goals. Specific objectives include: 1) characterizing current spatial patterns and historical dynamics of ecological, economic, and social components of the Coast Range ecosystem; 2) developing ecological, economic, and social models describing these components, and the linkages among each; and 3) projecting the aggregate impacts of current forest policies in the Coast Range on ecosystem conditions and economic outputs over time.

One socioeconomic factor that is expected to have a significant impact on projected forest policy outcomes in the Coast Range is land-use change resulting from the conversion of forestland to residential, commercial, and industrial uses. Currently, seventy percent of Oregon's 3.4 million people live in the Willamette Valley, with the valley population expected to grow by 1.3 million new residents in the next forty years (McGinnis et al. 1996; Franzen and Hunsberger 1998). Projected population growth has motivated increasing interest in examining where land-use changes are most likely to affect forests and the goods and services they provide throughout the region. Urbanization potentially can cause the forestland base to become more fragmented, adversely impacting ecosystem conditions and economic outputs. Ecological impacts could include direct loss of habitat or diminished habitat quality. Economic impacts could include less intensive forest management for commercial timber production resulting in reduced economic output. The goal of land-use modeling in the Coastal Landscape Analysis and Modeling Study is to place current and future forest policies in an appropriate socioeconomic context by accounting for the future distribution of humans throughout the study region.

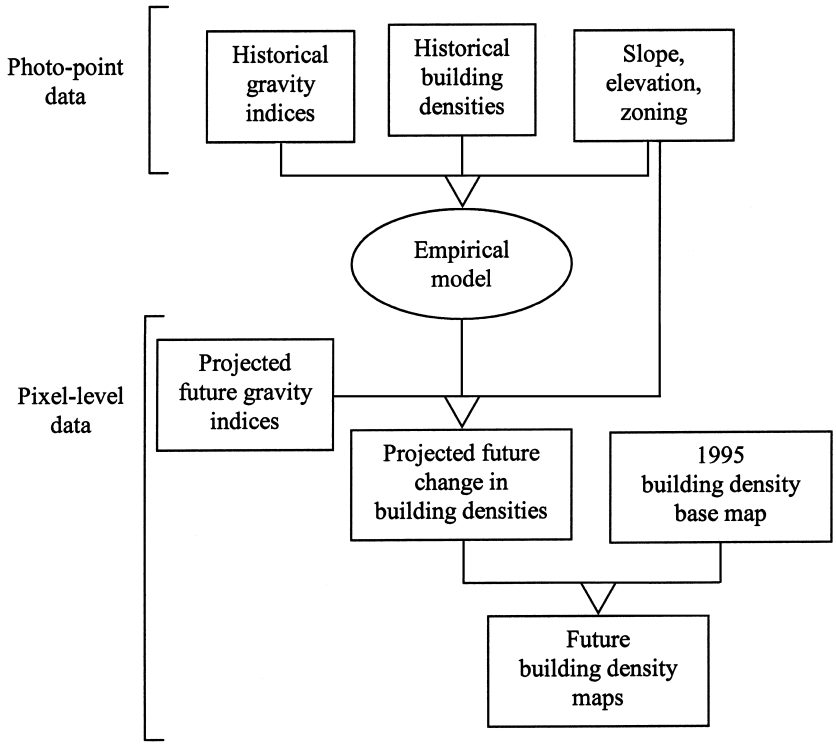

Figure 2. General modeling procedure.

Methods

Initial land use models developed for the study were based on readily available plot-level data describing historical changes among discrete forest, agriculture, and urban land-use categories provided by the USDA Forest Service's Forest Inventory and Analysis Program (Kline and Alig 1999; Kline et al. 2001). These data were used to estimate probit models describing the probability that forest and agriculture plots converted to urban uses in western Oregon and western Washington, as a function of several explanatory variables. Integrating the projected probabilities into other sub-models, however, presented difficulties. A specific need of the study is the delineation of future forestland areas at each modeling time interval. In western Oregon, the proportion of land in forest use historically has been quite high relative to the proportion in urban uses, based on Forest Inventory and Analysis land use definitions. As a result, projected probabilities describing the likelihood of future conversions of forestland to urban uses generally are quite low. However, analysis based on more recently available data describing building densities in western Oregon suggests that although the conversion of land from the discrete forest to urban use categories historically has been a relatively slow process, land use change has occurred in the form of relatively dispersed, low-density development (Azuma et al. 1999). Characterizing this particular form of development is the focus of the current land use modeling effort.

We use spatial photo-point data depicting historical building densities to estimate an empirical model describing historical changes in building densities in western Oregon as a function of several explanatory variables, including a gravity index of development pressure (Figure 2). We combine the empirical model with projected future gravity index values to project future changes in building densities that are applied to a 1995 map of building density to compute projected future building densities through 2055. We convert projected population densities into discrete land use classes using a decision rule that defines the conversion of forestland to low-density and urban development as a building density threshold.

Building density data

Data describing building densities in western Oregon were developed by the Pacific Northwest Research Station's Forest Inventory and Analysis Program. The data consist of photo-point observations of building density (number of buildings in an 80-acre (32-ha) vicinity) on non-federal land taken from aerial photographs in 1974, 1982, and 1994 (Azuma et al. 1999). With nearly 24,000 photo-points, the data provide almost 72,000 photo-point observations of building density varying in space at three points in time. By tracking building densities on individual photo-points at each of the three points in time, we were able to construct a data set comprised of two observations of change in building density for each photo-point.

A relatively large proportion of the photo-points show building densities of zero and do not change over the three time points described by the data. This results in a large number of zero's in the data set that complicates estimation of the empirical models. To alleviate these problems, we omitted observations showing building densities of zero. Computations of projected values for these omitted observations based on estimated coefficients of the estimated empirical models suggest that areas where building densities are less than 1 building per 80-acre (32 ha) are relatively unlikely to gain a sufficient number of buildings to ″convert″ to low density or urban development as defined by the landscape modeling study, due to poor physical access and steep slopes. We combined the building density data with other spatially-referenced socioeconomic and other data using a geographic information system to develop explanatory variables including slope, elevation, and land use zoning adopted under Oregon's land-use planning program. The resulting data set used to estimate the empirical models is comprised of 12,866 observations of changes in building densities from one time point to the next.

Characterizing development pressure

The value of land for residential, commercial, or industrial uses is perhaps the single most important factor affecting whether or not land is converted from a forest use to a developed use. Conceptually, the value of land in developed uses has been viewed as a function of the spatial proximity to city centers (Mills 1980; Miyao 1981; Fujita 1982; Wheaton 1982; Capozza and Helsley 1989). The traditional Von Thunen view of spatial proximity to cities had been viewed as affecting the profitability of non-developed land uses, such as agriculture and forestry, in terms of the costs associated with transporting forest and agricultural commodities to market (Barlow (1978), p. 37). However, modern society associates spatial proximity more with maximizing the difference between quality of life factors such as housing and neighborhood amenities, and the costs associated with commuting to employment locations. Additionally, other physical and institutional land characteristics may also affect the value of land in developed uses and the decisions of individual landowners regarding land use. Steeper slopes may increase building construction costs. Land use zoning may restrict certain types of development. We use a gravity index that integrates information about the sizes and locations of cities to describe the influence that cities have on land's development potential. We combine this with other variables describing physical land characteristics, such as slope and elevation, and institutional factors, such as land use zoning regulation, to develop an empirical model characterizing the value of land in developed uses.

Gravity models were initially developed by Reilly (1929) to describe the degree to which cities attract retail trade from surrounding locations (see for example Haynes and Fotheringham (1984)). A common gravity index specification for a single city is

$$Gravity \quad index = \frac{Population}{(Distance)^2} \tag{1}$$

and is directly proportional to the population of the city and inversely proportional to the square of the distance between the city and the location of interest. Gravity indices also have been used to account for the combined influence of population and proximity as economic forces effecting land-use change. For example, Shi et al. (1997) include a gravity index as an explanatory variable in a county-level hedonic model of farmland prices. Their 'urban influence potential variable' is constructed as the sum of the gravity indices computed for each of the three major cities nearest to each county and is a statistically significant variable in their empirical model of farmland value. Mathematical specifications other than Equation (1) are possible by including multiple cities in the gravity index computation and by varying the exponents on *population* and *distance*. In this way, gravity indices can be adapted to the specific conditions or 'social context' of the geographic region under study (Haynes and Fotheringham (1984), pp. 12–16).

One of the most important factors affecting land's development potential is its commuting proximity to employment opportunities available in existing cities. Land within a short commuting distance to a given city likely will have a greater development potential than land within a relatively longer commuting distance. Similarly, land within a reasonable commuting distance of a large city likely will have a greater development potential than land within the same commuting distance of a relatively small city. Cities beyond a reasonable commuting distance likely will have very little, if any, influence on development potential. We describe these influences using a single gravity index computed as

$$\text{GRAVITY} \quad \text{INDEX}_i = \sum_1^K POPULATION_k \left(\frac{60 - TIME_{ik}}{60} \right) \tag{2}$$

where K represents the number of cities within a 60-minute drive (or commute) of each photo-point i, *POPULATION* is the population (U.S. Bureau of Census 1992) of each city k, and *TIME* is the driving time in minutes between photo-point i and city k. As computed, the gravity is the sum of the populations of all cities within a 60-minute commute of each photo-point, weighted by the estimated driving time to each city's edge. The index sets a 60-minute threshold on the 'reasonable' commuting time, based on our assumption that most Oregonians are probably unwilling to commute more than one-hour to work. Varying this threshold to reflect somewhat shorter or longer maximum reasonable commuting times did not substantially affect the sign, magnitude, or statistical significance of the gravity index variable in model estimation.

The complete set of cities incorporated into the gravity index computation included 45 western Oregon cities comprising 5,000 or more persons in 1990 (U.S. Bureau of Census 1992). Adjacent cities were combined and treated as larger metropolitan areas, reducing the total number of cities and metropolitan areas included in the analysis to 30. Driving times used to calculate the gravity index were estimated using a geographic information system map of existing roads to create a friction surface based on average driving times assumed for different types of roads. We assumed that drivers could average speeds of 60 miles per hour (97 km h^{-1}) on primary roads, 25 mile per hour (40 km h^{-1}) on secondary roads, and 10 miles per hour (16 km h^{-1}) where there are no roads. The driving times are based on roads data from a single point in time, because we lack data describing new roads and improvements. As a consequence, we ignore potential endogeneity between land use change and road building noted by Irwin and Geoghegan (2001) among others.

Model estimation

The building density data consist of observations taken at three points in time (1974, 1982, and 1994), resulting in two observations of building density change per sample point. The dependent variable $\Delta DENSITY$ was constructed by computing changes in building densities observed at each sample point at ten-year intervals between 1974 and 1984, and between 1984 and 1994. Building density data for 1984 were estimated by interpolation between 1982 and 1994 values, and rounding to the nearest whole number. The dependent variable $\Delta DENSITY$ is measured as a count and so is not continuous. Assuming $\Delta DENSITY$ is distributed as a Poisson leads to the negative binomial model

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

where γ is a random variable and $\exp(\gamma)$ has a gamma distribution with mean 1 and variance α, x_i is a vector of independent variables, and β' is a vector of coefficients to be estimated (Greene 1997).

The panel nature of the data – generally two observations of building density change per photo-point – creates the potential for correlation among the pairs of time-series observations for individual photo-points to deflate standard errors and bias estimated coefficients. These potential correlations can be accounted for using a random effects negative binomial model (see Greene (1995), pp. 570–571 for a derivation). Since the group effects are conditioned out (not computed), projected values cannot be computed using the random effects model (Greene (1995), p. 567). However, the estimated model coefficients can be

Table 1. Descriptions of Explanatory Variables Tested in the Empirical Model

Variable	Description
GRAVITY INDEX	Equal to the average of the gravity index computed (using Equation (2)) at the beginning of each time period and the gravity index computed at the end of each time period (times 1/100,000). City populations for study years for non-Census years estimated by interpolating between populations reported for Census years (U.S. Bureau of Census 1992).
BUILDING DENSITY	Number of buildings within an 80-acre circle surrounding photo-point (Azuma et al. 1999) at the beginning of each time period (times 1/100).
SLOPE	Percent slope at the sample point (times 1/100).
ELEVATION	Elevation in meters.
URBAN GROWTH BOUNDARY	Variable equals 1 if plot is located in an urban growth boundary or rural residential land use zone; 0 otherwise.
FARM ZONE	Variable equals 1 if plot is located in a farm zone; 0 otherwise.
FOREST ZONE	Variable equals 1 if plot is located in a forest zone; 0 otherwise.

used for comparison with those of the negative binomial model estimated without random effects.

A final estimation issue arises from our use of spatial observations of land use. Spatial autocorrelation can result from omitted spatial variables that influence the land-use decisions of landowners, such as weather-related variables, and spatial behavioral relationships, such as common ownership of neighboring photo-points. The first leads to inefficient but asymptotically unbiased estimated coefficients, while the second can lead to inefficient and biased estimated coefficients (Nelson and Hellerstein 1997). Although no standard statistical protocols exist, methods to treat spatial autocorrelation in land-use analyses have been devised and tested, including the use of spatial lag (or neighborhood) variables based on the variable values of neighboring pixels (see, for e.g., Bockstael (1996) and Turner et al. (1996), Nelson and Hellerstein (1997), Wear and Bolstad (1998), Schneider and Pontius (2001)) and purposefully sampling to reduce the potential of autocorrelation arising from spatial behavioral relationships (see, for e.g., Fortin et al. (1989) and Haining (1990), Helmer (2000)). In our case, building density data are based on a systematic sampling of photo-points roughly spaced on a 2.4-kilometer grid. We are unable to construct a spatial lag variable because pixel-level information regarding the actual building density between sample photo-points is unavailable. Given the 2.4-kilometer sample spacing, we assume that the effects of any spatial behavioral relationships not accounted for by the gravity index and other variables are minimal.

Results

The general regression equation describes the change in building density occurring on individual photo-points from one time point to the next as

$$\begin{aligned}&\Delta DENSITY\\&= f(GRAVITY\quad INDEX, BUILDING\quad DENSITY,\\&SLOPE, ELEVATION, URBAN\quad GROWTH\\&BOUNDARY, FARM\quad ZONE, FOREST\quad ZONE)\end{aligned} \tag{4}$$

where the specific explanatory variables are described in Table 1. Results from Poisson regression, negative binomial regression, and negative binomial regression with random effects are shown in Table 2. All models are highly significant ($P < 0.01$). Random effects coefficients are reasonably consistent with negative binomial coefficients, though the statistical significance of the beta coefficient in the negative binomial regression with random effects suggests that statistically significant random effects may be present.

Estimated coefficients for the linear and quadratic GRAVITY INDEX variables are statistically significant ($P < 0.01$) and together suggest that, over time, building densities increase at an increasing rate with greater proximity to existing cities within commuting distance and greater population sizes of those cities (Table 2). Estimated coefficients for the linear and quadratic BUILDING DENSITY variables are statistically significant ($P < 0.01$) and together suggest that existing building densities have a positive but diminishing impact on future building density changes. The estimated coefficients for SLOPE are negative, sug-

Table 2. Estimated Coefficients of the Empirical Models Describing Changes in Building Densities in Western Oregon

Variable	Poisson regression coefficient	Negative binomial regression		Negative binomial regression with random effects
		Coefficient	Marginal effect	
GRAVITY INDEX	−0.167 (−17.54)	−0.293 (−15.59)	−0.432	−0.183 (−9.125)
GRAVITY INDEX^2	0.037 (26.44)	0.051 (16.90)	0.075	0.027 (9.913)
BUILDING DENSITY	10.813 (84.13)	13.453 (36.71)	19.800	9.333 (31.64)
BUILDING DENSITY^2	−10.070 (−48.86)	−13.436 (−30.30)	−19.774	−12.207 (−26.22)
SLOPE	−0.338 (−1.97)	−0.191 (−0.56)	−0.281	−0.078 (−0.21)
ELEVATION	0.517 (9.96)	0.470 (4.08)	0.691	0.400 (3.62)
URBAN GROWTH BOUNDARY	−0.531 (−22.06)	−0.359 (−7.04)	−0.528	−0.503 (−13.05)
FARM ZONE	−1.314 (−49.02)	−1.339 (−32.17)	−1.748	−1.027 (−25.08)
FOREST ZONE	−1.210 (−34.18)	−1.188 (−21.00)	−1.970	−0.861 (−16.50)
Alpha	–	2.207 (45.94)	–	2.753 (26.59)
Beta	–	–	–	2.075 (18.09)
Summary Statistics:	N = 12,866	N = 12,866		N = 12,866
	Log-L = −24,687	Log-L = −16,479		Log-L =−16,550
	χ^2 = 16,708	χ^2 = 16,415		
	df = 8	df = 1		
	P < 0.0001	P < 0.0001		

Note: The *t*-statistics for each estimated coefficient are in parentheses.

gesting that slope has a negative impact on changes in building densities, but the coefficients' statistical significance is notable only in the Poisson regression model (P < 0.05) and the random effects negative binomial regression (P < 0.10). As defined, it is likely that the slope variable only poorly represents the impact of slope on average building density within the 80-acre vicinity of each sample point.

The estimated coefficients for ELEVATION are positive and statistically significant (P < 0.01, P < 0.01, and P < 0.05) in each of the three models, suggesting that elevation has a positive impact on changes in building densities. This finding is consistent with that of Wear and Bolstad (1998) who attribute their positive elevation coefficient to the possibility that higher elevations command better views, making them more attractive as building sites. Explanatory variables included to account for the potential impacts of land-use zoning adopted under Oregon's land-use planning program are negative and statistically significant (P < 0.01), suggesting that the implementation of land use zoning may have reduced the rate at which building densities increase over time (Table 2).

Model validation procedures

We evaluated the forecasting performance of the estimated negative binomial model in three ways: 1) examining the percentage of correct projections within-sample; 2) estimating auxiliary models after reserving validation data sets; and 3) examining several information indices and statistics based on model projections. First, we used the estimated negative binomial model coefficients (Table 2) to compute projected changes in building densities, then added the projected changes to the initial building densities to compute within-sample projected ending building densities for each observation (N = 12,866). We compared projected ending building densities to actual ending building densities to compute the percentage of correct projections.

The percentage of correct projections diminishes as ending building density increases, from a high of 52.0% for observations having an ending building density of 2 buildings per 80-acre (32-ha) to a low of 15.3% for observations having an ending building density of 8 (Table 3). The percentage of model projections correct within one building is higher, ranging from 99.5% for observations having an ending building density of 1 building per 80-acre (32-ha) to a low of 63.6% for observations having an ending building

Table 3. Percentage of Correct Projections of Ending Building Density and Ending Broad Building Density Class

Class	Percent in class	Percent correctly projected	Percent correctly projected within one building
Ending building density[a]			
1	25.4	50.9	99.5
2	17.0	52.0	99.0
3	12.8	43.7	91.5
4	8.6	36.7	83.1
5	6.2	23.6	71.2
6	5.2	21.9	66.8
7	3.6	18.9	63.6
8	3.1	15.3	75.1
> 8	18.1	86.7	89.6
Ending broad building density class			
⩽ 8	81.9	97.0	–
> 8	18.1	86.7	–

[a] Number of buildings per 80-acre (32-ha), rounded to nearest whole building if less than or equal to 8.

density of 7. Greater accuracy of projections in the lower range of ending building densities likely is due in part to the relatively large proportion of sample observations comprising relatively low ending building densities.

The immediate use of the model within the Coastal Landscape Analysis and Modeling Study is to locate forestland in the study region comprising ending building densities of greater than 8 buildings per 80-acre (32-ha) (64 per square mile) – the point at which timber management and production is assumed to cease in study sub-models. This threshold is consistent with an average forest parcel size of 10 acre (0.04 km^2) building (house), which is the minimum forest parcel size eligible for preferential assessment as forestland for property tax purposes in the State of Oregon (Oregon Department of Revenue 1998). Based on an average household size of 2.45 persons (Azuma et al. 1999), the 64 buildings per 2.59 km^2 threshold is equivalent to 157 people per square mile, which also is relatively consistent with the population density found by Wear et al. (1999) to be the point at which commercial timber production ceases. The percentage of correct projections for the two classes is relatively high – 97.0% for the ⩽ 8 class and 86.7% for the > 8 class – suggesting that the model is probably adequate for the immediate purposes to which it is used in the Coastal Landscape Analysis and Modeling Study (Table 3).

As a second model evaluation, we estimated five auxiliary models after omitting roughly 20% of the observations from the full sample (N = 12,866) as validation data sets. A common approach to evaluating the forecasting performance of empirical models is to reserve a portion of sample data prior to model estimation for later use as a validation data set. We initially declined to do this so that we could take full advantage of the relatively limited number of observations of actual building density changes. The five auxiliary models, however, enable us to evaluate our model specification by examining the sensitivity of coefficient estimates to the omission of the validation data sets and by examining the percentage of correct projections resulting from the five auxiliary models when applied to the validation data sets.

The five auxiliary models are highly significant (P < 0.01) and all coefficient estimates are consistent in sign, magnitude, and statistical significance with those of the main model estimated with the full data sample (N = 12,866), with the exception of the SLOPE coefficient estimates that are statistically insignificant (P > 0.20) in all models (Table 4). We compared coefficient estimates of the five auxiliary models to 95% confidence bounds computed for the coefficient estimates of the main model. All auxiliary model coefficient estimates fall within the 95% confidence bounds, with the sole exception of the BUILDING DENSITY^2 coefficient estimate from auxiliary model 2, which falls outside the lower bound for that variable. Together, these factors suggest that the five auxiliary models do not differ significantly from the main model. The weighted average percentage of correct projections of ending building classes of ⩽ 8 and > 8 buildings per 80 acre

Table 4. Estimated Coefficients of Five Auxiliary Negative Binomial Models Compared to 95% Confidence Bounds Computed for Main Model Coefficients (Table 3)

	Auxiliary negative binomial model estimated coefficients[a]					95% confidence bounds of main model coefficients[b]	
Variable	1	2	3	4	5	Lower	Upper
GRAVITY INDEX	−0.265	−0.312	−0.280	−0.315	−0.297	−0.330	−0.257
GRAVITY INDEX^2	0.047	0.053	0.050	0.054	0.051	0.045	0.057
BUILDING DENSITY	13.114	14.092	13.553	13.260	13.349	12.735	14.172
BUILDING DENSITY^2	−12.956	−14.754	−13.421	−12.935	−13.394	−14.305	−12.567
SLOPE	−0.124	0.154	−0.465	−0.109	−0.412	−0.856	0.475
ELEVATION	0.468	0.388	0.541	0.509	0.435	0.244	0.695
URBAN GROWTH BOUNDARY	−0.374	−0.365	−0.368	−0.379	−0.317	−0.459	−0.259
FARM ZONE	−1.374	−1.355	−1.343	−1.322	−1.301	−1.420	−1.257
FOREST ZONE	−1.192	−1.180	−1.190	−1.170	−1.207	−1.298	−1.077
Alpha	2.256	2.218	2.210	2.135	2.206	2.113	2.301
Summary statistics:	N = 10,240	N = 10,261	N = 10,275	N = 10,322	N = 10,366	–	–
	LL = −13,192	LL = −13,132	LL = −13,309	LL = −13,130	LL = −13,146		
	χ^2 = 13,424	χ^2 = 13,558	χ^2 = 13,637	χ^2 = 11,749	χ^2 = 13,219		
	df = 1	df = 1	df = 1	df = 1	df = 1		
	P < 0.0001	P < 0.0001	P < 0.0001	P < 0.0001	P < 0.0001		

[a] Estimated by omitting roughly 20% of the sample observations as validation data sets based on a random selection process. [b] Computed for negative binomial model estimated coefficients reported in Table 2.

(32-ha) resulting from the five auxiliary models is 97.0% for the ⩽ 8 class and 86.6% for the > 8 class.

As a third model evaluation, we computed several information indices and statistics suggested by Wear and Bolstad (1998), based on the ending building density projections from the main and auxiliary models. The index H(A) describes the total uncertainty that potentially can be explained by the estimated models, and is defined as

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

where $p(a_j)$ is the proportion of observations in the validation data set actually observed in building density class a_j and J is the total number of building density classes projected. The index I(A;X) describes the additional information contained in the estimated models, and is defined as

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

where $\delta_{ij} = 1$ if class j is observed at observation i ($\delta_{ij} = 0$ otherwise), x_i is the vector of independent variables describing observation i, $p(a_j|x_i)$ is the model-estimated probability of building density class j occurring at observation i, and m is the number of observations in the validation data set. The index EI(A;X) describes the expected information provided by the estimated models, and is defined as

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

(Wear and Bolstad 1998).

Table 5. Information Indices and Statistics Computed for the Main Model Projections Applied to the Sample Data and Auxiliary Model Projections Applied to Omitted Validation Data Sets

Index or statistic[a]	Main Model m = 12,866	Auxiliary Models 1 m = 2,626	2 m = 2,605	3 m = 2,591	4 m = 2,544	5 m = 2,500
Projecting ending building density (1, 2, 3, 4, 5, 6, 7, 8, > 8)						
I(A;X)	0.093	0.098	0.084	0.083	0.093	0.102
EI(A;X)	0.106	0.099	0.106	0.107	0.115	0.111
V(A;X)	0.038	0.025	0.038	0.040	0.044	0.043
H(A)	1.288	1.270	1.286	1.237	1.302	1.338
U^2 = I(A;X)/H(A)	0.072	0.078	0.066	0.067	0.071	0.076
t-statistic	0.066	0.004	0.113	0.119	0.105	0.042
LLR = 2nI(A;X)	2,387.9	517.3	439.2	429.6	473.2	509.5
Projecting ending broad building density class (≤ 8, > 8)						
I(A;X)	0.356	0.334	0.355	0.353	0.388	0.348
EI(A;X)	0.106	0.321	0.348	0.332	0.374	0.350
V(A;X)	0.345	0.298	0.278	0.289	0.275	0.270
H(A)	0.472	0.443	0.480	0.459	0.494	0.484
U^2 = I(A;X)/H(A)	0.754	0.753	0.740	0.769	0.785	0.719
t-statistic	0.022	0.024	0.014	0.039	0.028	0.002
LLR = 2nI(A;X)	9,168.3	1,753.6	1,850.6	1,828.2	1,976.2	1,742.0

[a] Computed following Wear and Bolstad (1998). The statistics define a pseudo-r^2 (U^2) measure of usefulness, a t-test of accuracy (H_O: I(A;X) = EI(A;X) where V(A;X) =), and a χ^2 test (LLR) of statistical significance of the model projections.

The three indices enable computation of three test statistics with which to evaluate the models as predictors of ending building density classes. The proportion of uncertainty explained by the models is a pseudo-r^2 defined as U^2 = I(A;X)/H(A) and is a test of the usefulness of the models at projecting ending building density classes. The index I(A;X) is normally distributed with a mean of EI(A;X) and a variance of V(A;X), enabling a t-test of the null hypothesis H_O: I(A;X) = EI(A;X), that provides a test of the accuracy of the empirical models. The log-likelihood ratio defined as LLR = 2nI(A;X) is distributed as a chi square with degrees of freedom equal to the number of estimated coefficients in the estimated models, and is a test of the overall significance of the empirical models (Hauser 1978; Wear and Bolstad 1998).

Information indices and statistics computed based on projections of ending building density and ending broad building density class resulting from the main and auxiliary models are provided in Table 5. The log-likelihood ratios (LLR) and t-statistics computed based on the ending building density projections suggest that the empirical models are both statistically significant and accurate, but each of the pseudo-r^2 (U^2) values suggest that the proportion of uncertainty explained by the empirical models is relatively low. The log-likelihood ratios (LLR) and t-statistics computed based on the projected ending broad building density classes suggest that the empirical models also are both statistically significant and accurate. However, in these cases, the pseudo-r^2 (U^2) values suggest that the proportion of uncertainty explained by the empirical models of ending broad building density class is much higher, ranging between 74.0% and 78.5%. Consistent with our earlier examination of the percentage of correct projections, the U^2 values suggest that the model is better at projecting coarser (or less precise) ending building density classes. Greater accuracy in projecting less precise ending building density classes, however, is not the result of a spatial scale (or 'grain size') effect (see, for example, Jenerette and Wu (2001)). Rather, it is the result of reducing through aggregation the number of building density classes we are attempting to project with the model, from nine (1, 2, 3, 4, 5, 6, 7, 8, > 8) to two (≤ 8, > 8).

Integrating building densities with ecological models

The empirical model was used to create geographic information system maps depicting spatial projections of future building density distributions throughout the Coastal Landscape Analysis and Modeling Study region. A base year 1995 map of building densities was developed from the 1994 photo-point data by interpolating between photo-point building density values. The estimated negative binomial model coefficients (Table 2) were combined with projected gravity index values based on population projections for western Oregon cities to project changes in building densities at 10-year time intervals. Projected population figures are based on county-level projected population growth through 2010 (McGinnis et al. 1996) and on state-level projected population growth for 2010 to 2050 reported by the U.S. Bureau of Census. Population projections for the years 2050 to 2095 are estimated by extrapolation. Projected changes in building densities for each 10-year time interval were added to the beginning building density map for that interval to obtain the ending building density map. For example, the projected changes occurring between the 1995 base year and 2005 were added to the 1995 base year building density map, to obtain a 2005 building density map. The 2005 map was combined with 2005 to 2015 projected changes in building densities to obtain a 2015 map. The resulting maps enable projected future changes of human habitation of forestland, as described by building densities, to be incorporated into other Coastal Landscape Analysis and Modeling Study sub-models describing other socioeconomic and ecosystem processes and conditions.

For the specific purposes of the Coastal Landscape Analysis and Modeling Study, the building density maps are incorporated into sub-models describing timber production and habitat viability according to building density thresholds. Initial land use conditions distinguish forestlands from agricultural lands using a vegetation map depicting forest and non-forest cover in 1995. These delineations remain constant throughout the modeling time horizon. Forestlands are distinguished from lands characterized by residential, commercial, or industrial uses by applying a set of decision rules to the building density maps at each modeling time interval. For timber production modeling purposes, timber production is assumed to cease on forestlands once a building density of 64 buildings per 2.59 km^2 is attained. For habitat viability modeling purposes, habitat is assumed to cease functioning once a building density of 640 buildings per 2.59 km^2 is attained. Land areas comprised of building densities between 64 and 640 buildings per 2.59 km^2 are assumed to comprise relatively low-density residential and other development. Land areas comprised of building densities of greater than 640 buildings per square mile are assumed to comprise predominantly high-density urban development (Figure 3).

Once the forestland area contributing to timber production and habitat viability sub-models is delineated, $1.0{\cdot}10^3$ m^2 open vegetation patches (or building footprints) are created for each projected new building. The building footprints are intended to represent the indirect impact of buildings on timber production and habitat viability in terms of their direct impacts on vegetative cover. The 1.0 10^3 m^2 footprints are consistent with the average vegetation patch sizes found among a sampling of buildings in the study area. The footprints also are roughly equivalent in size to the basic simulation unit used in Coastal Landscape Analysis and Modeling Study sub-models. The specific locations of building footprints are selected randomly according to the estimated building density for each unit at each 10-year modeling time interval.

Maps of projected future building densities for western Oregon suggest significant expansion of low-density and urban development (Figure 3). The proportion of western Oregon land in low-density and urban developed uses is projected to increase from 4.8% and 2.0% in 1995 to 5.6% and 3.7% in 2025, and to 6.2% and 6.6% in 2055. Although the majority of new buildings are projected in locations surrounding existing cities, greater numbers of buildings also are indicated in forested areas that remain below the low-density development threshold of 64 buildings per square mile (8 per 80-acre (32-ha)). These projections suggest greater numbers of people living in closer proximity to forestlands in the Coastal Landscape Analysis and Modeling Study region in the future.

The projected building densities are based on population values that are outside the range of data used to estimate future building density distributions. To evaluate the reasonableness of the building density projections, we compared the amount of low-density and urban development per capita indicated by our spatial projections with per capita land use rates indicated by the 1997 National Resources Inventory data for Oregon (NRCS (Natural Resources Conservation Service) 1999). Our projections suggest that low-density and urban development will increase an average of $2.7{\cdot}10^3$ m^2 per new resident from 1995 to 2055. This rate is reasonably close to the average $2.1{\cdot}10^3$ m^2 increase in ″developed land″ per new resi-

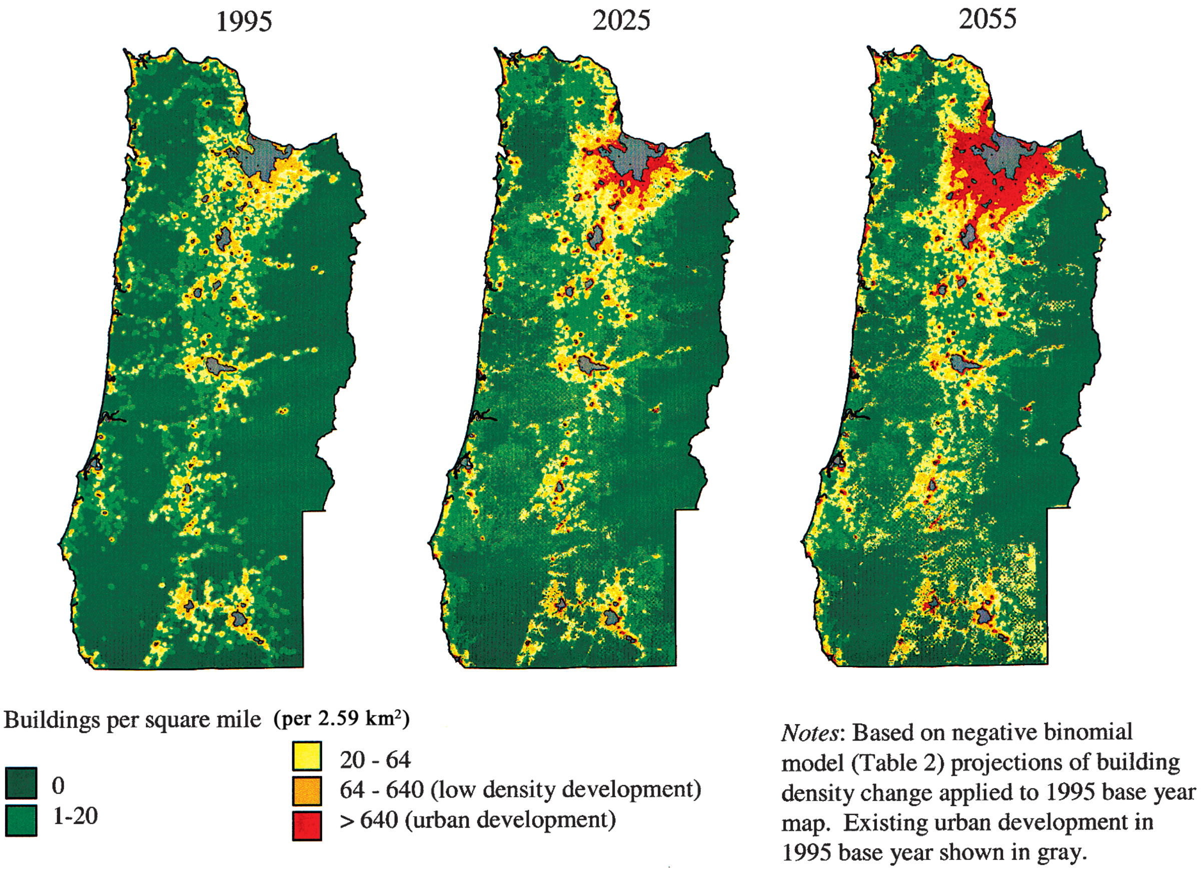

Figure 3. Projected building density classes in Western Oregon.

dent between 1982 to 1997 in Oregon and below the national average of 0.82 acres per new resident, based on National Resources Inventory data (Kline 2000).

Summary

Our empirical model of building density change is an improvement over the discrete land use modeling approach initially used by the Coastal Landscape Analysis and Modeling Study. The new model acknowledges that human habitation of forestland is not defined by discrete boundaries, but rather occurs along a continuum. The model describes a range of human habitation impacts that potentially can be incorporated into other sub-models describing socioeconomic and ecological conditions. Because the model is not limited to discrete delineations of forest and urban land, it potentially can be applied to a broader range of research issues. Also, the estimated negative binomial model provides projected values that are estimated changes in building densities, which are easier to interpret than projected probabilities provided by initial probit models based on changes among discrete land use categories.

In this particular application, the specific needs of the Coastal Landscape Analysis and Modeling Study called for the aggregation of projected building densities above 64 buildings per 2.59 km^2 into discrete land use categories of low-density (64 to 640 buildings per 2.59 km^2) and urban development (> 640 buildings per 2.59 km^2). Sub-models describing timber production activity and habitat viability were unable to use more detailed information regarding human habitation. However, modeling building densities, rather than discrete land use categories, enabled study researchers to select appropriate criteria with which to define these categories, rather than having to accept land use categories established by national land inventories or other data sources. The potential to incorporate the impacts of relatively low-density human habitation levels into landscape-level ecological models could motivate greater interest in multidisciplinary examinations of human interactions with landscape-level ecological and socioeconomic processes. The recent and increasing migration of humans to forested landscapes (noted by Egan and Luloff (2000) among others) likely will increase the need for such research in the future.

The results of model validation procedures suggest that the likelihood of correctly projecting ending building density classes using the model improves with the increasing coarseness of ending building density classes desired. To some extent, the reason for this result is intuitively obvious, and stems from the error inherent in estimation of a statistical model of building density change. The resulting model will be better at projecting *close to* the actual ending building density class than it will be at projecting the actual ending building density class *exactly*. However, the validation result also illustrates the tradeoff inherent in choosing between the precision and the accuracy with which building density classes are projected.

Our particular modeling approach was made possible by the ready availability of building density data, which is not available from existing national land inventories or other common sources. Although obtaining such data through photo-interpretation of aerial photography or satellite imagery is possible, it can be an expensive process and may not be feasible in certain applications. When limitations exist, researchers are advised to consider the tradeoffs associated with different types of data and models when evaluating the necessity for the additional detail contained in building density or similar data over more readily available discrete land use data. Among the many important factors to consider are the potential sensitivity of the socioeconomic or ecological processes under study to ranges of human habitation and the specific purposes of land use modeling in the overall landscape model context.

References

Azuma D.L., Birch K.R., DelZotto P., Herstrom A.A. and Lettman G.J. 1999. Land use change on non-federal land in western Oregon, 1973–1994. Oregon Department of Forestry, Salem, Oregon, USA, 55 pp.

Barlow R. 1978. Land Resource Economics: The Economics of Real Estate. Prentice-Hall, Inc, Englewood Cliffs, New Jersey, USA, 653 pp.

Barlow S.A., Munn I.A., Cleaves D.A. and Evans D.L. 1998. The effect of urban sprawl on timber harvesting. Journal of Forestry 96: 10–14.

Bockstael N.E. 1996. Modeling economics and ecology: the importance of a spatial perspective. American Journal of Agricultural Economics 78: 1168–1180.

Capozza D.R. and Helsley R.W. 1989. The fundamentals of land prices and urban growth. Journal of Urban Economics 26: 295–306.

Egan A.F. and Luloff A.E. 2000. The exurbanization of America's forests: research in rural social science. Journal of Forestry 98: 26–30.

Fagan W.F., Meir E., Carroll S.S. and Wu J. 2001. The ecology of urban landscapes: modeling housing starts as a density-dependent colonization process. Landscape Ecology 16: 33–39.

Fortin M.-J., Drapeau P. and Legendre P. 1989. Spatial autocorrelation and sampling design in plant ecology. Vegetatio 83: 209–222.

Franzen R. and Hunsberger B. 1998. Have we outgrown our approach to growth? The Sunday Oregonian, Portland, Oregon, USA, December 13.

Frayer W.E. and Furnival G.M. 1999. Forest survey sampling designs: a history. Journal of Forestry 97: 4–10.

Fujita M. 1982. Spatial patterns of residential development. Journal of Urban Economics 12: 22–52.

Geoghegan J., Villar S.C., Klepeis P., Mendoza P.M., Ogneva-Himmelberger Y., Chowdhury R.R. et al. 2001. Modeling tropical deforestation in the southern Yucatan peninsular region: comparing survey and satellite data. Agriculture. Ecosystems and Environment 85: 25–46.

Greene W.H. 1997. LIMDEP Version 7.0: User's Manual. Econometric Software, Inc., Bellport, New York, USA, 850 pp.

Greene W.H. 1995. Econometric Analysis. Prentice Hall, Upper Saddle River, New York, USA, 1075 pp.

Haining R. 1990. Spatial Data Analysis in the Social and Environmental Sciences. Cambridge University, Cambridge, UK, 409 pp.

Hauser J.R. 1978. Testing the accuracy, usefulness, and significance of probabilistic choice models: an information-theoretic approach. Operations Research 26: 406–421.

Haynes K.E. and Fotheringham A.S. 1984. Gravity and Spatial Interaction Models. Sage Publications, Beverly Hills, California, USA, 88 pp.

Helmer E.H. 2000. The landscape ecology of tropical secondary forest in montane Costa Rica. Ecosystems 3: 98–114.

Irwin E.G. and Geoghegan J. 2001. Theory, data, methods: developing spatially explicit economic models of land use change. Agriculture, Ecosystems and Environment 85: 7–23.

Jenerette G.D. and Wu J. 2001. Analysis and simulation of land-use change in the central Arizona – Phoenix region, USA. Landscape Ecology 16: 611–626.

Kline J.D. 2000. Comparing states with and without growth management: analysis based on indicators with policy implications, comment. Land Use Policy 17: 349–355.

Kline J.D. and Alig R.J. 1999. Does land use planning slow the conversion of forest and farmlands? Growth and Change 30: 3–22.

Kline J.D., Moses A. and Alig R.J. 2001. Integrating urbanization into landscape-level ecological assessments. Ecosystems 4: 3–18.

McGinnis W.J., Phillips R.H. and Connaughton K.P. 1996. County portraits of Oregon and northern California. General Technical Report PNW-GTR-377. U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station, Portland, Oregon, USA, 315 pp.

Milloy R.E. 2000. Population trends heighten West's fire woes. New York Times, New York, New York, USA, August 10.

Mills E.S. 1980. Urban Economics. Scott, Foresman and Co., Glenview, Illinois, USA, 241 pp.

Miyao T. 1981. Dynamic Analysis of the Urban Economy. Academic Press, New York, New York, USA, 188 pp.

NRCS (Natural Resources Conservation Service) 1999. Summary Report 1997 National Resources Inventory. U.S. Department of Agriculture, Washington, DC, USA, 84 pp.

Nusser S.M. and Goebel J.J. 1997. The National Resources Inventory: a long-term multi-resource monitoring programme. Environmental and Ecological Statistics 4: 181–204.

Nelson G.C. and Hellerstein D. 1997. Do roads cause deforestation? using satellite images in econometric analysis of land use. American Journal of Agricultural Economics 79: 80–88.

Oregon Department of Revenue 1998. Specially assessed forestland. Property Tax Division, Valuation Section, Oregon Department of Revenue., Salem, Oregon, USA.

Reilly W.J. 1929. Methods for the study of retail relationships. University of Texas Bulletin Number 2944. University of Texas, Austin, Texas, USA.

Schoorl J.M. and Veldkamp A. 2001. Linking land use and landscape process modeling: a case study for the Alora region (south Spain). Agriculture, Ecosystems and Environment 85: 281–292.

Schneider L.C. and Pontius R.G. Jr 2001. Modeling land-use change in the Ipswich watershed, Massachusetts, USA. Agriculture, Ecosystems and Environment 85: 83–94.

Serneels S. and Lambin E.F. 2001. Proximate causes of land-use change in Narok District, Kenya: a spatial statistical model. Agriculture, Ecosystems and Environment 85: 65–81.

Shi Y.J., Phipps T.T. and Colyer D. 1997. Agricultural land values under urbanizing influences. Land Economics 73: 90–100.

Spies T.A., Reeves G.H., Burnett K.M., McComb W.C., Johnson K.N., Grant G. et al. 2002. Assessing the ecological consequences of forest policies in a multi-ownership province in Oregon. In: Liu J. and Taylor W.W. (eds), Integrating Landscape Ecology into Natural Resource Management. Cambridge University Press, New York, New York, USA, pp. 179–207.

Swenson J.J. and Franklin J. 2000. The effects of future urban development on habitat fragmentation in the Santa Monica Mountains. Landscape Ecology 15: 713–730.

Turner M.G., Wear D.N. and Flamm R.O. 1996. Land ownership and land-cover change in the southern Appalachian Highlands and the Olympic Peninsula. Ecological Applications 6: 1150–1172.

U.S. Bureau of Census 1992. 1990 Census of Population and Housing. U.S. Department of Commerce, Washington, DC, USA.

Walsh S.J., Crawford T.W., Welsh W.F. and Crews-Meyer K.A. 2001. A multiscale analysis of LULC and NDVI variation in Nang Rong district, northeast Thailand. Agriculture Ecosystems and Environment 85: 47–64.

Wear D.N. and Bolstad P. 1998. Land-use changes in southern Appalachian landscapes: spatial analysis and forecast evaluation. Ecosystems 1: 575–594.

Wear D.N., Lui R., Foreman J.M. and Sheffield R. 1999. The effects of population growth on timber management and inventories in Virginia. Forest Ecology and Management 118: 107–115.

Wheaton W.C. 1982. Urban residential growth under perfect foresight. Journal of Urban Economics 12: 1–12.